

UNIVERSITE PARIS I – PANTHEON-SORBONNE
U.F.R. D'ECONOMIE

THESE
POUR LE DOCTORAT DE L'UNIVERSITE PARIS I

Discipline : Sciences Economiques

ANALYSE ET PREVISION EN PRESENCE DE RUPTURE
DU DEGRE D'INTEGRATION :
APPLICATION AUX TRAFICS DE LA RATP

réalisée dans le cadre d'une CIFRE (RATP et laboratoire EUREQua)

sous la tutelle de

Philippe Jolivaldt ¹

présentée et soutenue publiquement par

Emilie Muzereau ²

le 6 Décembre 2004

¹Laboratoire EUREQua UMR 8594 du CNRS, Université de Paris 1 Panthéon - Sorbonne, Maison des Sciences Economiques, 106/112 boulevard de l'Hôpital 75647 Paris Cedex 13, France. Mail : philippe.jolivaldt.univ-paris1.fr

²RATP, Département Commercial, Université Tarification Ventes Résultats, LAC A73, 54 quai de la Râpée, 75599 Paris Cedex 12. Mail : emilie.muzereau@ratp.fr

Publication AJD - 82 Agora Jules Dupuit,
Université de Montréal
www.ajd.umontreal.ca

Directeur de thèse :
Monsieur Jean-Pierre LAFFARGUE,
Professeur à l'Université de Paris I

JURY

M. Alain AYONG LE KAMA	Professeur à l'Université de Lille I Conseiller scientifique au Commissariat général du Plan
M. Marc GAUDRY	Professeur à l'Université de Montréal (Quebec) Chercheur associé à l'INRETS/DEST
M. Philippe JOLIVALDT	Maître de conférences à l'Université de Paris I
M. Jean-Pierre LAFFARGUE	Professeur à l'Université de Paris I
M. Eric TARDIVEL	Responsable de l'Unité Etudes et Modélisation à la RATP

L'UNIVERSITE DE PARIS I PANTHEON-SORBONNE n'entend donner aucune approbation ni improbation aux opinions émises dans les thèses ; ces opinions doivent être considérées comme propres à leurs auteurs.

Analyse et prévision en présence de rupture du degré d'intégration :

Application aux trafics de la RATP

Résumé :

Les déplacements en transport en commun sont soumis en permanence à de nombreux chocs, positifs (amélioration de la qualité de service, de l'offre de titres de transport, modifications du réseau) ou négatifs (perturbations, menaces d'attentats). Ces chocs affectent de manière distincte les sous-réseaux de la RATP (métro, RER, bus de Paris et de banlieue) et amènent à prendre en compte des évolutions non stationnaires et des ruptures.

Nous construisons des tests de modification du degré d'intégration à une date inconnue (hypothèses, statistiques, valeurs critiques, puissance et qualité de l'estimation de la date de rupture) et les appliquons aux différents trafics. Suivant le résultat de ces tests, les trafics sont modélisés de façon auto-régressive, sur l'ensemble de la période ou sur la seconde sous-période uniquement.

Afin d'améliorer l'analyse et la prévision des trafics mensuels, des variables décrivant la situation économique en Ile-de-France sont utilisées dans des modélisations issues de l'analyse de la cointégration. Des tests de changement du degré de cointégration sont construits, prenant en compte une rupture de l'équilibre de long terme (apparition / disparition de l'équilibre).

Un travail de rétro-prévision sur l'année 2003 montre la qualité explicative et prévisionnelle des modèles comportant des variables économiques. Les trafics sont prévus pour les années 2004 et 2005 : les prévisions ponctuelles sont calculées à partir de chaque modèle et l'incertitude est prise en compte par la construction d'intervalles de prévision simulés et calibrés.

Classification JEL : C15, C22, C53, R41

Mots clés : demande de transport, racine unitaire, cointégration, changement de structure, évaluation de la prévision, intervalles de confiance simulés

Analysis and forecasting in presence of a break of the integration degree. Application to the traffics of the RATP

Abstract :

Travels in public transport are continuously affected by crisis (perturbations, threat of attack) or improvements (service quality, tickets, network changes). These shocks have a different impact on the networks (subway, express rail, bus serving Paris or its suburbs) and lead to take in consideration non-stationary evolutions that can be submitted to breaks.

We construct tests to detect change in the integration degree at unknown break date (hypothesis, statistics, critical values, power and quality of the estimated date) and apply them to the traffic series. According to the test conclusions, we estimate each series with an auto-regression on the whole period or on a second sub-period only. In order to improve analysis and forecasting of monthly traffics, we take into account economic variables that describe the situation in Paris and its suburbs.

We construct tests of modification of cointegration degree, which consider a break of the long term relation (appearance/ disappearance of the equilibrium).

Back forecasting (year 2003) shows the quality of models with economic variables to explicate and forecast traffics. We calculate trajectories from each model and construct simulated and calibrated confidence intervals in order to consider uncertainty.

Keywords : : transportation demand, auto-regression, unit root, cointegration, change of structure, forecasting evaluation, simulated confidence intervals.

Remerciements

En premier lieu, je tiens tout particulièrement à remercier les deux personnes à l'origine de cette thèse. La peur de les décevoir a été mon plus fort moteur pour accomplir ce travail. Philippe JOLIVALDT, mon co-directeur de thèse, a été le premier à m'accorder sa confiance, bien avant la signature de la convention. Sa disponibilité à mon égard, sa gentillesse et sa patience, ses conseils (sa fermeté parfois !) m'ont poussée à travailler régulièrement. Georges FOURCADE m'a ouvert les portes de la RATP en acceptant d'être mon tuteur scientifique au sein de l'entreprise. Son intérêt vif pour la recherche scientifique et l'importance de ses connaissances en économie et les mathématiques appliquées m'ont beaucoup impressionnée et motivée.

Je remercie aussi Jean-Pierre LAFFARGUE, co-directeur de cette thèse, pour sa réactivité au moment opportun et ses nombreuses remarques. Je suis très honorée que Marc GAUDRY et Alain AYONG LE KAMA aient accepté d'être rapporteurs au sein de mon jury et que Eric TARDIVEL ait examiné ce travail dans des délais très raccourcis.

Je tiens aussi à remercier mes collègues de la RATP, qui m'ont offert une ambiance de travail des plus agréables.

Mes supérieurs hiérarchiques à la RATP, Michel BECK, Jacques DELORME ainsi que Philippe DIDIER, ont réussi à la fois à m'intégrer dans leur équipe et à me laisser l'autonomie nécessaire à l'élaboration d'une thèse.

La confiance inconditionnelle de Gérard DEMEOTHIS dans mes capacités et son soutien dans les moments plus difficiles m'ont particulièrement aidée. Nos promenades pendant l'heure du déjeuner m'ont permis de structurer mes horaires de travail, de me détendre et de reprendre des forces pour l'après-midi.

Je remercie aussi Patricia DUBUS, devenue mon tuteur scientifique en cours de route, pour l'intérêt qu'elle a porté à mes travaux, ainsi que pour la qualité de ses relectures. Son enthousiasme a grandement facilité la rédaction de cette thèse.

Merci aussi aux autres doctorants CIFRE de la RATP : le partage de nos expériences m'a permis d'anticiper de nombreuses difficultés.

Enfin, mes pauses ont été égayées par l'ensemble des personnes que j'ai pu rencontrer à la RATP. Ils se reconnaîtront !

En dernier lieu, je tiens à remercier mes amis et ma famille. Nos (trop) rares rencontres ont été des moments de véritable détente, qui m'ont redonné coeur à l'ouvrage.

Table des matières

Introduction générale	13
I Marché du transport collectif urbain et bases de données	23
1 L'offre de transport collectif urbain par la RATP	27
1.1 Le marché des transports collectifs urbains, un monopole naturel?	28
1.1.1 L'importance des coûts fixes et la fixation des prix	29
1.1.2 L'existence d'économies d'échelle	31
1.2 L'intervention nécessaire des pouvoirs publics	34
1.2.1 La notion de "service public"	35
1.2.2 L'existence d'externalités positives au développement de l'offre de transport collectif urbain	37
1.2.3 Le lien entre transport et aménagement du territoire	38
1.3 L'introduction d'éléments de compétitivité dans le marché des transports collectifs urbains	43
1.3.1 Les imperfections de la régulation étatique dans une grande entreprise	43
1.3.2 Est-ce un monopole "contestable" ?	45
1.3.3 Les objectifs et incitations du contrat STIF - RATP	49
2 La demande de transport collectif urbain	55
2.1 La demande de transport collectif urbain, une demande dérivée	56
2.1.1 Les motifs de déplacement	56

2.1.2	Motifs contraints et non contraints : conséquences sur la fréquentation des transports collectifs urbains	60
2.1.3	Les déterminants de la demande de transport	74
2.2	La demande de transport collectif urbain et l'équilibre sur ce marché	79
2.2.1	L'offre de transport crée la demande de transport	80
2.2.2	La demande de transport est plus ou moins élastique au niveau de l'offre de transport collectif	81
2.2.3	L'offre intervient comme une caractéristique du bien de transport collectif	83
2.3	L'inadéquation des modélisations existantes de la demande de transport collectif urbain	86
2.3.1	Le trafic inter-urbain	87
2.3.2	Le trafic intra-urbain	88
2.3.3	Le faible nombre existant de modèles économétriques "agrégés"	93
2.3.4	Modèles utilisés à la RATP	100
3	Les données relatives à la RATP	107
3.1	Caractéristiques des réseaux RATP et unités de mesure du trafic	108
3.1.1	Réseaux et sous-réseaux à la RATP	108
3.1.2	Unités de mesure du trafic à la RATP	122
3.2	Collecte des données brutes	124
3.2.1	Réseau ferré (métro + RER)	124
3.2.2	Réseau de surface (autobus + sites propres)	126
3.3	Correction des données par le calcul du jour ouvrable moyen sur chaque sous-réseau	129
3.3.1	Le calcul du trafic du jour ouvrable moyen	129
3.3.2	Jour ouvrable moyen et désaisonnalisation	134
3.3.3	Utilité pour la prévision	134
3.3.4	Evolution des poids des samedis et dimanches/jours fériés	135
3.4	Trafics étudiés et évolutions	137
3.4.1	Le trafic du jour ouvrable sur le métro	139
3.4.2	Le trafic du jour ouvrable sur le RER	142

3.4.3	Le trafic du jour ouvrable sur le réseau des autobus de Paris	144
3.4.4	Le trafic du jour ouvrable sur le réseau des autobus de banlieue	146
3.4.5	Le trafic du samedi sur le métro	148
3.4.6	Le trafic du dimanche/jour férié sur le métro	150
3.4.7	Le trafic du samedi sur le RER	152
3.4.8	Le trafic du dimanche/jour férié sur le RER	154
4	Les données d'environnement	157
4.1	L'emploi en Ile-de-France	158
4.1.1	Taux de chômage fournis par l'INSEE	158
4.1.2	Les Demandes d'Emploi en Fin de Mois (DEFM)	166
4.1.3	Evolution du taux de chômage et des demandes d'emploi en fin de mois depuis 1990	167
4.1.4	Les DEFM à Paris	171
4.1.5	Les DEFM en Ile-de-France	173
4.2	Le commerce en Ile-de-France	175
4.2.1	Evolution de la consommation en Ile-de-France depuis 1990	175
4.2.2	L'indice du chiffre d'affaires des grands magasins en région parisienne . .	176
4.3	Le tourisme en Ile-de-France	180
4.3.1	La situation du tourisme en Ile-de-France depuis 1990	181
4.3.2	Le trafic grandes lignes de la SNCF à Paris	186
4.3.3	Le taux d'occupation des hôtels à Paris	189
4.3.4	Le taux d'occupation des hôtels en Ile-de-France	191
4.4	La concurrence automobile en Ile-de-France	193
4.4.1	La consommation de carburants en Ile-de-France	194
4.4.2	Le rapport du prix du super au prix du ticket T	196
4.4.3	Le volume et la vitesse de circulation dans Paris	200
4.4.4	L'indice d'évolution de la circulation en zone agglomérée de la région Ile- de-France sur les autoroutes et routes nationales	203
4.5	Le manque de données sur la situation immobilière en Ile-de-France	205

II	Tests de changement du degré d'intégration au cours de la période	211
5	Les tests de changement de structure existants	219
5.1	Tests de changement de structure à une date inconnue et estimation de la date de rupture	220
5.1.1	Test classique de changement de structure dans un modèle stationnaire	220
5.1.2	Test de changement de structure unique à date inconnue	222
5.1.3	Tests de changements de structure multiples	223
5.2	La distinction entre une marche aléatoire et l'existence d'un changement de structure	224
5.2.1	Tests avec date de changement de structure connue	225
5.2.2	Tests avec date de changement de structure inconnue	228
5.2.3	Tests de changement de la persistance de la série à date inconnue.	229
5.3	Inadaptation de ces tests à notre problème	232
5.3.1	Le manque de travaux sur les changements de structure dans les processus auto-régressifs	232
5.3.2	L'inadaptation des tests existants de changements du degré d'intégration	232
5.3.3	L'inadaptation des tests ayant pour hypothèse nulle l'existence de racine unitaire	233
6	Tests de changement du degré d'intégration	235
6.1	Changement du paramètre associé au premier retard	236
6.1.1	Modèle de base : AR(1)	236
6.1.2	Hypothèses de test	236
6.2	Changement du degré d'intégration et de la constante avec correction des auto-corrélations	240
6.2.1	Hypothèse alternative de série intégrée puis stationnaire	241
6.2.2	Hypothèse alternative de série stationnaire puis intégrée	244
6.2.3	Valeurs critiques des tests de changement de degré d'intégration avec éventuelle rupture de la constante	247
6.3	Puissance des tests	249

6.3.1	Puissance des tests suivant la valeur du paramètre non contraint	250
6.3.2	Puissance suivant la valeur de la constante sur la partie stationnaire	256
6.3.3	Puissance des tests suivant la véritable date de rupture	260
7 Application des tests aux données de la RATP et aux séries économiques franciliennes 267		
7.1	Application des tests de changements du degré d'intégration aux données de la RATP	268
7.1.1	Trafic du jour ouvrable sur le métro	268
7.1.2	Trafic du jour ouvrable sur le RER	272
7.1.3	Trafic du jour ouvrable sur le réseau des bus de Paris	274
7.1.4	Trafic du jour ouvrable sur le réseau des bus de banlieue	276
7.1.5	Trafic du samedi sur le métro	278
7.1.6	Trafic du dimanche/jour férié sur le métro	280
7.1.7	Trafic du samedi sur le RER	282
7.1.8	Trafic du dimanche/jour férié sur le RER	284
7.1.9	La modification du degré de stationnarité des séries de trafic sur les sous-réseaux RATP	286
7.2	Application des tests de changements du degré d'intégration aux séries relatives à l'environnement économique en Ile-de-France	287
7.3	La pertinence de nos tests de modification du degré d'intégration par rapport aux tests de Perron	288
III Test de modification du degré de cointégration au cours de la période 293		
8 Tests de changement de structure de la relation de cointégration : synthèse 297		
8.1	Les tests de cointégration standards pour une relation de cointégration	298
8.1.1	Les tests de cointégration standards ayant pour hypothèse nulle l'absence de relation de cointégration	298
8.1.2	Tests de cointégration ayant pour hypothèse nulle l'existence de cointégration	302

8.2	Les effets d'un changement structurel sur les tests de cointégration	305
8.2.1	Niveau réel et puissance des tests de cointégration en présence de changement de structure au sein de la relation de cointégration	305
8.2.2	Niveau réel et puissance des tests de cointégration dans le cas de séries soumises à un changement de structure	307
8.3	Les tests de changement structurel dans une relation de cointégration	313
8.3.1	Les tests de Gregory et Hansen	314
8.3.2	Les tests de spécification de la relation de cointégration	317
8.3.3	Tests dans le cas d'un système de relations de cointégration	328
8.4	Inadaptation des tests de cointégration à notre problème	329
8.4.1	La non prise en compte de séries ayant subi une modification du degré d'intégration au cours de la période considérée	329
8.4.2	L'absence de tests de changement du degré de cointégration	330
9	Tests de modification du degré de cointégration	333
9.1	Motivation et hypothèses de test	334
9.2	Estimation et statistiques de test	335
9.3	Tables de valeurs critiques	337
9.3.1	Processus générateur de données sous l'hypothèse nulle	337
9.3.2	Valeurs critiques du test supérieur de Wald	338
9.3.3	Valeurs critiques du test inférieur de Student	339
9.4	Calculs de puissance	339
9.4.1	Puissance des tests suivant la valeur du paramètre auto-régressif des résidus de la relation de long terme sur la période où il y a cointégration . . .	340
9.4.2	Puissance des tests suivant la valeur de la constante dans la modélisation auto-régressive des résidus pour la sous-période sur laquelle il y aurait cointégration	346
9.4.3	Puissance des tests suivant la véritable date de rupture	350

10 Recherche des équilibres de long terme et des dynamiques de court terme	359
10.1 Analyse préliminaire de la corrélation et de la causalité	360
10.1.1 Analyse des corrélations	360
10.1.2 Analyse de la causalité au sens de Granger	363
10.2 Application des tests de changement du degré de cointégration et modèles de trafic à long terme	365
10.2.1 Trafic du jour ouvrable sur le métro	366
10.2.2 Trafic du jour ouvrable sur le RER	373
10.2.3 Trafic du jour ouvrable sur le réseau des bus de Paris	377
10.2.4 Trafic du jour ouvrable sur le réseau des bus de banlieue	381
10.2.5 Trafic du samedi sur le métro	385
10.2.6 Trafic du dimanche/ jour férié sur le métro	389
10.2.7 Trafic du samedi sur le RER	395
10.2.8 Trafic du dimanche/jour férié sur le RER	399
10.3 Dynamiques de court terme	403
10.3.1 Trafic du jour ouvrable sur le métro	403
10.3.2 Trafic du jour ouvrable sur le RER	405
10.3.3 Trafic du jour ouvrable sur le réseau des bus de Paris	407
10.3.4 Trafic du jour ouvrable sur le réseau des bus de banlieue	409
10.3.5 Trafic du samedi sur le métro	411
10.3.6 Trafic du dimanche/jour férié sur le métro	413
10.3.7 Trafic du samedi sur le RER	415
10.3.8 Trafic du dimanche/jour férié sur le RER	417
 IV Prévisions	 423
 11 Des modèles cointégrés sur une sous-période plus adaptés aux prévisions	 427
11.1 Le bilan de l'année 2003	428
11.1.1 Evolution de la situation économique en Ile-de-France au cours de l'année 2003	428

11.1.2	Evolution des trafics de la RATP en 2003	430
11.2	Les rétro-prévisions de trafic sur l'année 2003	431
11.2.1	Construction des prévisions ponctuelles	431
11.2.2	Critères de comparaison de la qualité des rétro-prévisions	437
11.2.3	Trafic du jour ouvrable sur le métro	440
11.2.4	Trafic du jour ouvrable sur le RER	445
11.2.5	Trafic du jour ouvrable sur le réseau des autobus de Paris	448
11.2.6	Trafic du jour ouvrable sur le réseau des autobus de banlieue	451
11.2.7	Trafic du samedi sur le métro	454
11.2.8	Trafic du dimanche/jour férié sur le métro	456
11.2.9	Trafic du samedi sur le RER	459
11.2.10	Trafic du dimanche/jour férié sur le RER	461
11.3	Apports des rétro-prévisions	464
11.3.1	Recommandations pour le choix du modèle de prévision	464
11.3.2	Impact de l'estimation des paramètres de la modélisation auto-régressive sur une sous-période uniquement	466
11.3.3	L'amélioration des explications des variations de trafic passées	470
12	Prévisions à l'horizon 2004 - 2005	483
12.1	Mise à jour des estimations des modèles	484
12.1.1	Trafic du jour ouvrable sur le métro	484
12.1.2	Trafic du jour ouvrable sur le RER	485
12.1.3	Trafic du jour ouvrable sur le réseau des autobus de Paris	487
12.1.4	Trafic du jour ouvrable sur le réseau des autobus de banlieue	488
12.1.5	Trafic du samedi sur le métro	489
12.1.6	Trafic du dimanche/jour férié sur le métro	490
12.1.7	Trafic du samedi sur le RER	491
12.1.8	Trafic du dimanche/jour férié sur le RER	491
12.1.9	Données décrivant l'environnement économique en Ile-de-France	492
12.2	L'introduction de scénarios	493

12.2.1	Prévisions auto-régressives avec scénarios	493
12.2.2	Evolution des variables dans le passé prise en compte dans le scénario implicite	495
12.2.3	Eléments de perspective en Ile-de-France à fin 2003	496
12.2.4	Un scénario d'évolution des données plus plausible	498
12.3	Les intervalles de prévision	500
12.3.1	Modèle auto-régressif avec racine unitaire	500
12.3.2	Modèle de long terme	502
12.3.3	Modèle de court terme (à correction d'erreur)	503
12.4	Prévisions des trafics RATP par type de journée et de sous-réseau	504
12.4.1	Résultats des prévisions pour chaque type de trafic journalier	504
12.4.2	Evolution future des différents trafics journaliers	521
12.4.3	Les prévisions sur chaque sous-réseau pour les années 2004 et 2005	528
 Conclusion générale		 537
 Bibliographie		 545
 ANNEXES		 563
 A Tests de racine unitaire sur période entière ou sous-période éventuelle pour l'ensemble des 8 trafics		 563
 B Propriétés des tests de changement de structure dans un modèle auto-régressif stationnaire de taille finie		 569
B.1	Modèle et estimation	570
B.2	Les tests utilisés	570
B.3	Valeurs critiques	573
B.4	Puissance des différents tests et estimation de la date de rupture par rapport à la vraie date	576

C	Tables de puissance des tests de changement du degré d'intégration	581
C.1	Processus générateur de données et calculs de puissance	581
C.2	Puissance des tests suivant la valeur du paramètre auto-régressif sur la sous-période stationnaire	582
C.3	Puissance des tests suivant la valeur de la constante sur la sous-période stationnaire	589
C.4	Puissance des tests suivant la véritable date de rupture	591
 D	 Application des tests de Perron et de nos tests de modification du degré d'intégration	 593
D.1	Test pour une date de rupture k connue	593
D.1.1	Hypothèses de test	593
D.1.2	Modèles pris en considération	594
D.2	Test pour une date de rupture inconnue	596
D.2.1	Sélection du nombre p de variables en différences retardées	597
D.2.2	Valeurs critiques	598
D.3	Application des tests de Perron aux données de la RATP	598
D.3.1	Mise en garde préliminaire	599
D.3.2	Application aux séries de trafic	602
D.4	Application des tests de Perron et de nos tests de modification du degré d'intégration aux séries économiques	611
 E	 Tables de puissance des tests de changement du degré de cointégration	 643
E.1	Processus générateur de données et calculs de puissance	643
E.2	Puissance des tests suivant la valeur du paramètre auto-régressif des résidus de la relation cointégrée	645

E.3	Puissance des tests suivant la valeur de la constante dans la modélisation des résidus pour la sous-période cointégrée	657
E.4	Puissance des tests suivant la véritable date de rupture	661
F	Causalité au sens de Granger	665
F.1	Définitions relatives à la causalité	665
F.2	Application des tests de causalité à un modèle bivarié	667
F.3	Liens entre cointégration et causalité	668
G	Historique des faits marquants ayant une incidence sur le trafic de la RATP	671
H	L'amélioration des méthodes d'analyse et de prévision à la RATP	687

Introduction générale

Selon l'étude "Life Resolution" d'AIRPARIF [1], plus de 60% des Franciliens respirent un air de qualité "non satisfaisante". Les résultats de l'étude concernent la pollution moyenne de l'ensemble de l'Ile-de-France, i.e. ne sont pas influencés par des sources d'émission locale. Dans Paris intra-muros et le nord est de la petite couronne, les taux de benzène (cancérigène) et de dioxyde d'azote sont supérieurs aux normes. Certains secteurs du Sud du Val d'Oise et du Nord de l'Essonne sont aussi touchés de façon ponctuelle. L'évolution technique des moteurs à essence ainsi que l'amélioration de la qualité des carburants ont permis de réduire les rejets de benzène depuis janvier 2000, mais 740 000 personnes sont touchées régulièrement par des dépassements de la norme, principalement dans Paris intra-muros et la banlieue nord et nord-est. Au final, l'agglomération parisienne présente des résultats en termes de pollution nettement moins bons que ceux de Rome ou de Madrid¹. Les études épidémiologiques relèvent une hausse des décès et hospitalisations lorsque la pollution augmente, les personnes âgées et enfants asthmathiques étant les plus touchés. Or 54% du dioxyde d'azote et 34% du benzène proviennent du trafic routier. Cette forte pollution observée a conduit à l'émergence et au renforcement de la volonté de réduire la place de l'automobile en Ile-de-France.

La capitale est résolue à décourager la circulation dans le centre-ville et à promouvoir les transports en commun. Cette volonté de la mairie de Paris se traduit par la mise en place d'un système "Mobilien" : la création de couloirs protégés par des "banquettes" en béton doit permettre au réseau des autobus de devenir un véritable métro de surface, alliant rapidité et fiabilité. Dans les couloirs de bus, les émissions polluantes ont baissé de 5 à 15% lors de leur

¹L'étude AIRPARIF portait sur les quatre agglomérations suivantes : Dublin, Madrid, Rome et Paris.

mise en place, et ont encore été réduites par le remplacement progressif des autobus par des autobus écologiques. La politique de verbalisation des stationnements s'est en outre fortement renforcée en 2003 (selon la Préfecture de Paris, le nombre de procès verbaux a augmenté de +21.4% entre février 2002 et 2003), le nombre de places de stationnement gratuit a fortement diminué, de même que le nombre de places dans leur ensemble (lors de la mise en place des couloirs de bus notamment). Autre élément dissuasif, les horodateurs n'acceptent plus les pièces mais nécessitent l'achat d'une carte à puce. La circulation automobile a connu à Paris une baisse d'environ -5% entre avril 2001 et avril 2003, conformément aux prévisions du Plan de Déplacements Urbains en Ile-de-France (PDUIF)².

Le nombre moyen de déplacements effectués par chaque Francilien (âgé de plus de 6 ans, au cours d'un jour de semaine) est demeuré stable entre 1976 et 2001 : 3.5 déplacements par jour et par personne [25]. L'augmentation de la population en Ile-de-France a engendré, de façon mécanique, une augmentation globale du nombre de déplacements, mais un infléchissement de la progression de la voiture a été constaté. Entre 1991 et 2001, le nombre de déplacements en automobile a augmenté en moyenne de +0.7% par an (contre +2.6% entre 1976 et 1991). La diminution du nombre de déplacements réalisés à pied et à l'aide de deux-roues a été stoppée, et l'utilisation des transports en commun a poursuivi son augmentation. Les transports collectifs ont toutefois vu leur part de marché diminuer légèrement, de +20.1 en 1991 à +19.4% en 2001. Cette légère diminution s'est accomplie dans une période fortement défavorable aux transports collectifs : entre 1991 et 2001, la population a augmenté fortement en grande couronne (+9.8%), là où la voiture est le mode de transport prédominant (6 trajets sur 10 y sont effectués en voiture particulière).

La première question que nous nous poserons est la suivante : comment analyser cette évolution plutôt favorable aux transports en commun en Ile-de-France ? Quels facteurs influencent de façon positive les transports collectifs urbains, ou au contraire les pénalisent ? Jouent-ils un rôle aussi important sur le métro que sur les autobus ? Peut-on quantifier l'impact de ces différents

²Celui-ci prévoyait en outre une diminution de la circulation automobile de -3% sur l'ensemble de la région, de -3% en petite couronne et de -2% en grande couronne, tandis que l'usage des transports en commun devait augmenter de +2% et représenter un tiers des déplacements domicile-travail et domicile-école.

éléments et améliorer l'analyse des variations passées des trafics de la RATP ?

En Ile-de-France, les transports collectifs bénéficient d'un atout : la forte densité du réseau ferré à Paris, supérieure à des villes telles que Londres, Tokio ou New-York. Depuis le début des années 1990, le succès des lignes de tramway et de la ligne 14, ou encore la montée en puissance des abonnements (notamment la carte Imagine R créée en 1998) ont eu un impact favorable sur la fréquentation des sous-réseaux de la RATP. A l'inverse, les grandes grèves dans les transports en décembre 1995 ou les attentats sur le métro et le RER ont terni la réputation des transporteurs publics. Ces différents éléments ont pu constituer une réelle rupture, et modifier le comportement des utilisateurs de la RATP. Un deuxième ensemble de questions auxquelles devra répondre la thèse sera donc le suivant : quels événements ont pu modifier le comportement des utilisateurs des transports en commun ? Ces modifications ont-elles été ponctuelles, ou durables ?

L'amélioration constante de l'offre de transport de la RATP (forte diminution du nombre de grèves après 1995, augmentation de la fréquence sur les sous-réseaux, développement de nouveaux titres...) correspond à un élément de réponse aux questions précédentes. L'offre de transport influence la demande de transport en commun : par exemple, une augmentation de la fréquence permet de diminuer la congestion dans les rames de métro, les wagons de RER ou encore les autobus. Pour attirer de nouveaux usagers dans les transports en commun et accroître la mobilité des clients de la RATP, l'offre de service doit s'adapter aux besoins des utilisateurs et à l'évolution de leurs comportements³. Pour cela, il est nécessaire de prévoir les niveaux futurs de fréquentation des sous-réseaux, le plus fréquemment possible. Le troisième but de la thèse résidera dans la construction d'un outil opérationnel de prévision mensuelle des trafics sur les sous-réseaux de la RATP. Ces prévisions serviront en outre de base à l'élaboration du budget annuel de la RATP, à un horizon de 12 et 18 mois, ainsi qu'aux re-prévisions effectuées en milieu d'année (horizon : 6 mois).

³L'évolution des rythmes urbains depuis une dizaine d'années a entraîné une forte augmentation de la fréquentation sur le métro en début de soirée après 20 heures (+60% entre 1990 et 2000) et au cours des week-ends (+14%). Pour répondre aux attentes des Franciliens en termes de mobilité et de confort, la RATP a mis en application une série de mesures améliorant l'offre de service en réduisant le temps d'attente pour les clients : augmentation du nombre de rames en circulation, embauches de conducteurs supplémentaires...

L'objet de la **première partie** de la thèse est d'analyser les caractéristiques du marché des transports collectifs en Ile-de-France, et de déterminer les principaux facteurs agissant sur les niveaux de trafic, afin d'améliorer l'analyse des variations de trafic à la RATP. Cette partie constitue une introduction au monde des transports collectifs urbains en Ile-de-France : elle permet de définir le périmètre d'action de la RATP et de circonscrire les variables économiques qui peuvent influencer sur le nombre de voyageurs qu'elle transporte.

Le marché des transports collectifs franciliens est régulé par une autorité organisatrice, le Syndicat des Transports en Ile-de-France (STIF). Celui-ci fixe à la RATP un volume d'offre à mettre en oeuvre, et décide de la politique tarifaire. Le **premier chapitre** de cette partie détaille de façon précise le contexte institutionnel des transports collectifs en Ile-de-France et rappelle les théories économiques qui ont pu conduire à ce type de régulation.

Dans le **deuxième chapitre** de cette partie sera analysée la demande de transport collectif urbain. Face à cette offre fixée par le STIF (volume de service offert et tarification), la demande de transport collectif correspond à une demande dérivée : le transport n'est pas demandé pour lui-même, mais parce qu'il apporte un "bénéfice" à sa destination. Les motifs à l'origine des déplacements sont des déterminants de la fonction de demande de transport. Sur ce marché, l'offre et la demande ne s'ajustent pas par les prix, qui sont fixes, mais plus par la congestion du système de transport : l'offre intervient comme un ensemble de caractéristiques de ce bien de transport. De nombreuses modélisations existantes prennent en compte cette spécification du marché des déplacements, nous nous demanderons s'il est possible de les appliquer directement au trafic de la RATP que nous désirons analyser et prévoir.

L'existence de modifications de la contexture des réseaux de la RATP ainsi que la disponibilité des données nous conduisent à construire une base de données mensuelles à partir de janvier 1990. Ces deux éléments sont abordés dans les deux chapitres suivants.

Dans le **chapitre 3**, nous nous intéresserons plus précisément à l'activité de transporteur public qu'exerce la RATP, et définirons de façon précise les séries que nous désirons modéliser. La contexture⁴ des réseaux, leurs caractéristiques en terme de qualité de service ou leurs évolutions

⁴Le terme "**contexture**" désigne la façon dont sont assemblées les différentes parties d'un tout ; un synonyme pourrait être la "structure" ou le "maillage".

passées et à venir, les unités de mesure du trafic ou leur méthode de collecte, les différentes clientèles ainsi que les différents titres qui leur sont proposés, seront détaillés au cours de ce chapitre. La thèse sera centrée sur 4 sous-réseaux (le métro, le RER, les bus de Paris ainsi que les bus de banlieue (y compris tramway)) et trois types de journées moyennes : le jour ouvrable, le samedi et le dimanche/jour férié. La prise en compte de ces séries typiques d'un type de journée permet d'éliminer l'impact ponctuel des mouvements sociaux au sein de l'entreprise ou les effets du calendrier (nombre plus ou moins grands de week-ends, existence éventuelle de ponts ou de vacances scolaires...). Les effets de saisonnalité sont aussi écartés par une désaisonnalisation des variables de trafic.

Dans le **chapitre 4**, la RATP sera replacée au centre de sa zone d'influence historique, l'Ile-de-France. Le but de ce chapitre est de présenter les séries chronologiques associées aux motifs de déplacement dégagés dans le deuxième chapitre. Pour que le modèle d'analyse et de prévision soit opérationnel, il est indispensable que les données les plus récentes de ces séries soient disponibles rapidement et sous leur forme définitive. Néanmoins, aucune série pouvant correspondre à un déterminant de la fonction de demande de transport collectif ne doit être omise. Comment concilier l'exigence théorique et le caractère opérationnel d'un modèle d'analyse et de prévision des différents types de trafic ? Nous nous demanderons s'il est utile de prendre en compte des données économiques relatives à cette seule région, et difficiles à collecter, ou s'il serait possible d'utiliser les mêmes variables au niveau national. L'analyse des évolutions des variables économiques depuis 1990, au niveau national et régional, permettra de mettre en exergue la spécificité de la région Ile-de-France et de sa conjoncture économique et, par conséquent, la nécessité de prendre en compte des séries franciliennes dans nos modèles.

Dans la **partie II**, les trafics de la RATP seront abordés sous un angle économétrique, plus technique. La description des séries de trafic dans la partie précédente montre que de nombreux chocs ont affecté les données. En particulier, les fortes grèves de décembre 1995 ou encore la mise en place de la carte Imagine R en 1998 semblent avoir modifié le comportement des utilisateurs des sous-réseaux de la RATP. Alors que l'analyse des données effectuée à la partie précédente montre que les séries sont intégrées⁵ sur l'ensemble de la période, nous nous sommes

⁵Les notions de "non-stationnarité", "d'intégration" ou de présence de "racine unitaire" formalisent le fait de

demandés si ces séries n'étaient pas soumises à un changement de structure. Au sein d'une analyse préliminaire, le test qui nous semblait le plus intéressant à appliquer était le test de Perron (1989 [134] et 1990 [135]) : pour Perron, la plupart des séries macroéconomiques peuvent être modélisées sous la forme de fluctuations stationnaires autour d'une tendance déterministe ayant subi une rupture. Sur la période considérée, l'application de ce type de test permettait de prendre en compte l'éventualité d'un cycle économique. Pour la majorité des séries, les tests de Perron concluaient à la stationnarité autour d'une tendance dont le niveau ou la pente avaient subi une rupture. L'application de tests de racine unitaire sur les sous-périodes dégagées infirmaient en partie les conclusions de Perron : s'il n'existait pas de racine unitaire sur chaque sous-période, une racine unitaire était présente sur l'une des sous-périodes. Cette analyse préliminaire nous a conduit à nous interroger, tout au long de la partie II, sur l'éventualité d'une modification du degré d'intégration au cours de la période considérée. Notre idée était la suivante : il est possible que, au cours du temps, un événement de grande ampleur (grève, attentat, nouvel abonnement...) ait eu pour conséquence de modifier de façon durable le comportement des voyageurs, en les rendant plus sensibles à leurs expériences passées sur les réseaux de transport collectif.

Pour répondre à cette question, nous rechercherons dans la littérature existante un test qui nous permette de vérifier cette hypothèse. Cette littérature ayant trait aux changements de structure et à la présence de racine unitaire sera décrite dans le premier chapitre de la partie II (**chapitre 5**). L'état actuel de la recherche n'apporte pas de solution opérationnelle à ce problème : les seuls tests prenant en compte une modification de la persistance sont les tests de Kim [97] ou de Busetti et Taylor [32] mais ils possèdent une hypothèse nulle stationnaire.

Cette recherche nous pousse à construire nos propres tests de modification du degré d'intégration, dans le deuxième chapitre de cette partie II (**chapitre 6**). Nous nous demanderons quels tests peuvent être modifiés pour tester un changement du degré d'intégration, quelles pourraient être les hypothèses d'un tel test ainsi que la statistique de test, en supposant que la date de rupture nous est inconnue. Pour permettre une utilisation opérationnelle de ce test, il sera nécessaire de créer les tables de valeurs critiques mais aussi d'évaluer la qualité de ces tests en terme de puissance et d'estimation de la date de rupture. Bien plus, il semble indispensable de

garder en mémoire les chocs passés, de façon pérenne.

créer un mode d'emploi pour l'utilisation de ces tests, contenant des recommandations liées à leur application et les conséquences en terme d'estimation auto-régressive des séries.

Une fois ces outils construits, nous appliquerons nos tests de modification du degré d'intégration aux séries de trafic de la RATP mais aussi aux séries décrivant l'environnement économique en Ile-de-France (**chapitre 7**). L'application de nos tests à ces séries est justifiée : elles ont été testées intégrées à l'ordre 1 sur l'ensemble de la période lors de leur analyse à la partie I, mais nous soupçonnons qu'elles ont subi une rupture de leur degré d'intégration. Cette application devra être comparée aux résultats des tests de Perron, de façon à être persuadés de la nature de la modification des séries, mais aussi à mettre en valeur la qualité de nos tests. Enfin, les conclusions des tests nous permettront de conduire une estimation auto-régressive différente suivant les cas : sur la période entière ou sur la deuxième sous-période uniquement, sur les variables en niveau ou bien en différence première.

Dans la **partie III**, nous nous intéresserons moins à la nature de chacune des séries qu'aux liens entre les trafics de la RATP et les variables décrivant l'activité économique en Ile-de-France. L'application de nos tests de modification du degré d'intégration aux séries de trafic sur les sous-réseaux de la RATP montre l'existence d'une rupture du processus de mémoire des séries, rupture à mettre en lien avec les modifications de l'environnement économique en Ile-de-France. Une modélisation des trafics en fonction de variables économiques s'avèrera nécessaire, pour mieux expliquer la rupture des séries mais aussi en expliquer les retournements. Comment prendre en compte des séries ayant subi une rupture de leur degré d'intégration ?

Dans le premier chapitre de cette partie (**chapitre 8**), une synthèse de la littérature existante sur la cointégration et les changements de structure sera dressée. Le fait que les tests de racine unitaire concluent que les séries étudiées sont intégrées sur l'ensemble de la période nous pousse à envisager l'approche de la cointégration, mais la modification du degré d'intégration correspond à une rupture. Une première étape indispensable résidera dans la vérification que l'analyse de la cointégration peut s'appliquer à des séries ayant subi une modification de leur degré d'intégration, sans altération des performances des tests usuels de cointégration. En outre, nous montrerons qu'un type de test de cointégration avec rupture n'a jamais été envisagé jusqu'ici : il s'agit des tests de modification du degré de cointégration au cours d'une période, qui correspondent à une

application de nos tests de modification du degré d'intégration aux résidus d'une relation non cointégrée.

Dans le deuxième chapitre (**chapitre 9**), nous montrerons l'intérêt de tels tests. L'idée est la suivante : il semble possible de considérer qu'un équilibre de long terme a été valide pendant une certaine période, puis ne l'a plus été, ou inversement. Une variable a pu intervenir pendant un temps dans la relation et ne jouer plus aucun rôle ensuite, tandis que de nouveaux facteurs peuvent apparaître déterminants à partir d'une certaine période. Par exemple, on peut supposer que la vitesse de circulation dans Paris pénalise de moins en moins le transport en autobus à Paris du fait de la mise en place de couloirs protégés. A l'inverse, la qualité de l'information sur les lignes ou la propreté semblent des critères de plus en plus importants aux yeux des voyageurs. Comme dans le chapitre 6, nous présenterons les hypothèses de test, les statistiques, et tabulerons les valeurs critiques, en supposant que la date de rupture est inconnue. Nous nous intéresserons en outre aux propriétés (puissance, qualité de l'estimation de la date de rupture) de ces tests de modification du degré de cointégration et élaborerons une procédure de test, et mettrons en lumière les implications des conclusions des tests sur la modélisation de la relation de cointégration.

Dans le chapitre suivant (**chapitre 10**), nous appliquerons les tests de modification du degré de cointégration aux relations entre une série de trafic et des variables décrivant l'environnement économique en Ile-de-France. Cette application est justifiée : pour la majorité des trafics étudiés, quel que soit le nombre de variables économiques prises en compte, il n'existe pas de relation de cointégration sur la période considérée. Mais nous pensons qu'un tel équilibre existe, au moins sur une sous-période. Les conclusions de nos tests de modification du degré de cointégration auront des conséquences sur l'estimation des équilibres de long terme (valides sur l'ensemble de la période ou une sous-période uniquement), mais aussi sur celle des dynamiques de court terme. Les élasticités de chaque trafic par rapport aux variables économiques significatives, qui mesurent la variation de trafic induite par la variation d'une variable économique, devront être analysées d'un point de vue économique.

Les modèles de trafic en fonction de leur seul passé (partie II) ou à long et court terme en fonction de variables économiques (partie III) ainsi construits pourront permettre la prévision,

thème central de la **partie IV** de la thèse. Au cours de cette partie, la principale question sera la suivante : laquelle de ces modélisations doit-être choisie pour prévoir ? En l'absence de modification de l'environnement, la modélisation en fonction du passé de la variable ("modélisation auto-régressive") conduit à de bonnes prévisions. Elle sert à la fois à réaliser une première prévision des données de trafic, et à prévoir l'évolution des variables économiques qui apparaissent dans les modèles cointégrés de trafic. Elle est d'utilisation aisée et immédiatement opérationnelle. Toutefois, ce type de prévision a pour défaut majeur de ne pas permettre de prendre en compte un retournement conjoncturel envisagé par les experts. L'introduction de variables économiques dans les modélisations de la partie III devrait permettre d'analyser et de prévoir les retournements des séries de trafic.

Dans le premier chapitre de cette dernière partie (**chapitre 11**), nous nous intéresserons de façon précise à la qualité des prévisions issues des différentes modélisations. Ces prévisions seront effectuées pour l'année 2003, dont les données n'ont pas été prises en compte au cours des estimations alors qu'elles sont d'ores et déjà disponibles. Pour chaque type de trafic et chaque type de modélisation, nous évaluerons la qualité de ces rétro-prévisions. Cette étape permettra de mettre en évidence les modèles de prévision à privilégier, et constituera un moyen de valider les différentes modélisations proposées dans la thèse. En outre, au cours de ce chapitre, nous montrerons que l'analyse des variations passées de trafic est améliorée par la prise en compte de variables économiques au sein des modèles.

Dans le dernier chapitre (**chapitre 12**), enfin, nous réaliserons la prévision des différents trafics pour les années 2004 et 2005. Les données décrivant l'environnement économique en Ile-de-France ne seront pas connues mais devront être prévues. Nous considérerons deux types de scénario : un scénario "implicite" qui suppose que les évolutions futures des variables économiques seront similaires à celles observées en moyenne par le passé, mais aussi un scénario plus positif, décrit par une reprise économique de même ampleur que celle observée en 1997. Nous nous intéresserons en outre au risque associé à ces trajectoires prévues, par la construction d'intervalles de prévision. Toute la difficulté consistera à obtenir des intervalles suffisamment réduits pour qu'ils soient informatifs.

De façon à faciliter la lecture de la thèse, de nombreux résultats (tableaux, graphiques,

méthodes, ...), un historique des faits marquants ayant eu une influence possible sur les trafics de la RATP, ainsi que les premiers résultats de l'appropriation de cette thèse au sein de l'entreprise ont été reportés dans les annexes.

Première partie

**Marché du transport collectif urbain
et bases de données**

Introduction à la première partie

Le marché des transports en commun en Ile-de-France est un marché bien spécifique :

- l’offre de transport est déterminée par l’autorité organisatrice : le STIF (Syndicat des Transports en Ile-de-France) fixe un volume d’offre (nombre de véhicules*kilomètres) à mettre en oeuvre, les titres offerts par les entreprises de transport collectif (ticket T, abonnements...) ainsi que leurs prix. C’est ce que nous examinerons dans le premier chapitre.
- la demande de transport est une demande dérivée : le transport n’est pas demandé pour lui-même, mais parce qu’il apporte un ”bénéfice” à sa destination. Il s’agit en outre d’une demande effective, i.e. qui résume l’équilibre de marché. Dans la littérature des transports, cette demande de transport collectif urbain a été modélisée de différentes manières. Les caractéristiques de la demande de transport collectif urbain seront abordées dans le deuxième chapitre.

Les modèles présentés au cours du chapitre 2 ne peuvent être utilisés de façon directe pour expliquer et prévoir à court et moyen terme les trafics sur le métro, le RER, les autobus de Paris ou les autobus de banlieue. Les hypothèses sous-jacentes (absence de rupture des données par exemple, ou propriétés statistiques telles que la stationnarité) ne sont pas valides pour les données dont nous disposons. Ces modélisations permettent néanmoins de dresser l’inventaire des déterminants de la demande.

Les données utilisées lors de la modélisation seront détaillées dans les chapitres suivants : le chapitre 3 présentera les différents trafics de la RATP, le chapitre 4 les données d’environnement. L’origine des données, les corrections apportées ou encore les propriétés statistiques de chacune des bases de données seront détaillées. Le choix des données de trafic a été réalisé en fonction des

besoins du département Commercial de la RATP. La base de données décrivant l'environnement économique a été construite de façon à prendre en compte l'ensemble des déterminants de la demande, sous réserve de disponibilité des données. Nous avons choisi de faire débuter ces bases de données mensuelles en janvier 1990, après l'interconnexion de la ligne du RER A à la ligne SNCF de Cergy-Pontoise, de manière à limiter le nombre de ruptures dans les séries de trafic, et à avoir des séries environnementales complètes.

Chapitre 1

L'offre de transport collectif urbain par la RATP

Le marché des transports collectifs intra-urbains en Ile-de-France est constitué de trois producteurs principaux de service : la RATP (Régie Autonome des Transports Parisiens), la SNCF Ile-de-France (Société Nationale des Chemins de Fer) et OPTILE (Organisation Professionnelle des Transports d'ILE-de-France)¹. Ces trois organismes sont placés sous la tutelle d'une autorité organisatrice, le STIF (Syndicat des Transports d'Ile-de-France)², qui réunit l'Etat, la Région Ile-de-France et les 8 départements d'Ile-de-France³.

L'existence d'une autorité organisatrice sur le marché des transports collectifs urbains amène naturellement à s'interroger sur la nature de ce marché des transports intra-urbains, en particulier sur les justifications et objectifs de la régulation économique, c'est-à-dire de l'intervention publique en matière de prix et de service offert par une entreprise telle que la RATP.

¹Optile regroupe l'ensemble des entreprises privées exploitant des lignes régulières (bus et navettes) inscrites au Plan de Transport d'Ile-de-France, soit, à ce jour, 93 entreprises ou centres d'exploitation. Ces entreprises exercent, pour la plupart, trois types d'activités : du transport public sur lignes régulières, du transport scolaire, et du transport de personnel. Les recettes réalisées sur les lignes régulières exploitées par ses adhérents font l'objet d'une compensation du STIF (compensation carte orange).

²Le STIF a été officiellement créé le 13 décembre 2000, en remplacement de l'ancien STP (Syndicat des Transports Parisiens). Cette réforme institutionnelle a pour origine l'entrée de la région au conseil d'administration, ainsi que la création de nouveaux circuits de financement. Par souci de simplification, nous parlerons toujours par la suite du "STIF".

³Un processus de décentralisation est à l'oeuvre. En janvier 2005, l'Etat, majoritaire, devrait quitter le STIF alors que la Région devrait prendre un poids plus important.

1.1. LE MARCHÉ DES TRANSPORTS COLLECTIFS URBAINS, UN MONOPOLE NATUREL ?

La majorité des analyses économiques relatives au transport en commun de voyageurs ont été réalisées en prenant en considération le chemin de fer. Elles sont transposables au cas du réseau ferré (RER et métro), parfois au réseau routier (lignes d'autobus) de la RATP. Selon Nash et Preston (1993) [125], l'organisation des transports collectifs correspond naturellement à un monopole public, pour les raisons suivantes :

1. c'est un monopole naturel,
2. les pouvoirs publics doivent intervenir pour des raisons de "service public".

Nous examinerons successivement ces deux points, dans les deux premières sections.

Une dernière question, relative à cette intervention de l'Etat, concerne les objectifs poursuivis : équité sociale ou efficacité économique ? Est-il possible de concilier au mieux ces deux éléments ? Ceci nous amènera à nous intéresser, dans la troisième section, à la recherche de compétitivité dans ce monopole public.

Cette analyse nous permettra de mieux connaître l'offre sur le marché des transports collectifs intra-urbains d'une part, et l'évolution récente du mode de fonctionnement de la RATP d'autre part. Le but sous-jacent est de montrer que la RATP n'est pas décisionnaire en matière de tarification et de service offert, fixés par le STIF.

1.1 Le marché des transports collectifs urbains, un monopole naturel ?

Dans certaines activités (énergie, activité ferroviaire, télécommunications...), le coût fixe des infrastructures n'est pas couvert par un prix de vente maintenu au plus bas par la libre concurrence. Pour atteindre un profit non négatif, les entreprises doivent développer leur taille jusqu'à occuper une part essentielle du marché, voire la totalité du marché. Progressivement, les producteurs se concentrent, jusqu'à une situation de monopole, appelé dans ce cas précis "monopole naturel".

Dans le premier paragraphe, nous mettrons en avant l'importance des coûts fixes sur le marché des transports collectifs. L'existence d'économies d'échelle, qui poussent à la recherche de la taille maximale, sera évoquée dans le second paragraphe.

1.1.1 L'importance des coûts fixes et la fixation des prix

1.1.1.1 L'importance des coûts fixes

Les réseaux ferrés sont caractérisés par l'importance des coûts fixes, notamment du fait des infrastructures (voies, véhicules, stations et gares, points de vente des titres de transport, garages pour les véhicules à l'arrêt, signalisation...) et de leur entretien.

A titre d'exemple, la construction de la ligne 14 du métro, appelée "Météor" (METro Est - Ouest Rapide), entre les stations "Madeleine" et "Bibliothèque François-Mitterrand" a coûté l'équivalent de 941 millions d'euros pour les infrastructures hors matériel roulant. A cette somme s'ajoutent 124 millions d'euros pour le matériel roulant et 29 millions d'euros pour la salle d'échanges entre la ligne 14 ("Bibliothèque François - Mitterrand") et la nouvelle gare Masséna de la ligne C du RER⁴. Cette ligne, de 7.2 kilomètres et 7 stations, a été mise en service en octobre 1998 et est entièrement automatisée. Le prolongement de "Madeleine" à "Saint-Lazare", mis en service à la fin 2003, devait coûter plus de 130 millions d'euros (convention tripartite Etat - RATP - Région du 15 janvier 1999), tandis que le prolongement de 0.7 kilomètre au sud-est vers une nouvelle station de métro appelée "Olympiades" a un coût prévisionnel de 111 millions d'euros selon le STIF.

Alors que les coûts fixes sont exorbitants, le coût marginal est en général négligeable pour transporter une personne de plus⁵. Ce type de situation est typique d'un monopole naturel, et a des conséquences sur la tarification qui sera appliquée.

1.1.1.2 Les conséquences théoriques en termes de tarification

L'équilibre général optimal au sens de Pareto (équilibre dit de "premier rang" [159]), est obtenu quand le prix de chaque produit est égal à son coût marginal. L'importance des coûts

⁴La construction de ce premier tronçon a duré 5 ans et a permis la création d'un tunnel de 8.7 kilomètres de long et de 8.6 mètres de diamètre, à environ 25 mètres de profondeur. Un tunnelier à pression de boue a été utilisé entre la station "Madeleine" et la "Gare de Lyon", tandis que la traversée de la Seine a été rendue possible par l'utilisation de 4 caissons immergés dans une tranchée creusée dans le lit du fleuve [42].

⁵En effet, le coût marginal associé au transport d'un voyageur supplémentaire peut être considéré comme nul si les capacités de transport sont suffisantes. Si les capacités sont insuffisantes pour transporter ce voyageur en plus et qu'un bus ou un métro supplémentaire doit être mis en place, ou que les fréquences doivent être accrues, il est bien plus important.

1.1. LE MARCHÉ DES TRANSPORTS COLLECTIFS URBAINS, UN MONOPOLE NATUREL ?

fixes a pour conséquence que la tarification ne peut s'effectuer au coût marginal. Pour un tel prix, le profit de l'entreprise serait négatif, et l'entreprise n'aurait pas intérêt à produire. Il n'est pas non plus souhaitable de laisser le monopole naturel pratiquer un prix de monopole, bien trop élevé.

Pour concilier un prix faible et un profit non négatif, deux types d'approche en termes de tarification sont possibles :

1. Si l'entreprise ne reçoit pas de subventions, elle doit réaliser un profit non négatif, ce qui implique de proposer un prix supérieur ou égal au coût moyen. Si elle doit fournir un service à toute personne disposée à payer ce service, le prix d'équilibre est obtenu par égalité au coût moyen. Dans ce contexte, l'organisme gouvernemental de contrôle fixe le prix que l'entreprise de service est autorisée à pratiquer, ce qui implique de connaître les véritables coûts de l'entreprise. L'inconvénient de cette solution, qui respecte la contrainte d'équilibre budgétaire et est qualifiée d'équilibre de "second rang", provient du fait que l'offre réellement produite sera inférieure au niveau optimal.
2. Le gouvernement peut fournir lui-même les services. Il offre la quantité de services telle que le prix est égal au coût marginal (le niveau de service est alors optimal) et alloue une subvention globale qui permet de maintenir l'entreprise en activité. Cette subvention globale n'a pas pour but de corriger un fonctionnement inefficace, mais doit refléter l'importance des coûts fixes associés à cette mission de service public. Elle implique une fois de plus de surveiller les coûts.

La RATP est plutôt dans ce deuxième type de configuration : le STIF définit les conditions générales d'exploitation, détermine les itinéraires, la durée et la fréquence des services. Il assure en outre l'équilibre financier du fonctionnement des transports publics en versant chaque année, grâce au versement transport et aux contributions des collectivités publiques, 3.7 milliards d'euros aux transporteurs. Il crée les titres de transport et fixe les tarifs. Enfin, il approuve les contrats triennaux qui définissent les obligations réciproques du STIF (financement), de la RATP et de la SNCF (offre et qualité de service). Nous y reviendrons dans la troisième section de ce chapitre.

1.1.2 L'existence d'économies d'échelle

La taille importante d'une entreprise telle que la RATP a pour conséquence première de permettre l'achat en gros de ses équipements, presque systématiquement au terme d'appels d'offre, ce qui en diminue le prix d'acquisition. Produire sur une plus grande échelle permet de faire baisser le prix de revient unitaire, et conduire à une hausse de la recette unitaire (si le prix de vente ne diminue pas) ou à une augmentation des quantités vendues induite par une diminution du prix de vente. Cette baisse du coût unitaire d'un produit suite à un accroissement des quantités produites traduit une "économie d'échelle".

En matière de transport collectif urbain, l'existence d'économies d'échelle est confortée par deux facteurs majeurs : l'importance des coûts fixes évoqués précédemment, et la prise en considération de la qualité de service.

1.1.2.1 La répartition des coûts fixes et l'absence de taille limite

Dans le cas d'infrastructures de transports collectifs, les économies d'échelle résultent de la répartition des coûts fixes (siège social, bâtiments, réseau social, coût salarial de l'encadrement, équipements,...) sur une production qui augmente.

Au fur et à mesure que croît la taille des équipements, les contraintes de production se complexifient, et peuvent conduire à un accroissement de leur coût. Une taille maximale peut être atteinte, au delà de laquelle il n'est plus raisonnable d'augmenter la capacité de l'équipement. Pour augmenter la production, il deviendrait nécessaire de doubler les équipements, et d'accroître conjointement les dépenses de maintenance. L'existence d'une telle "taille limite" est remise en cause dès que la diversification des services offerts est permise, ce qui est le cas à la RATP. L'exemple le plus frappant est celui de la construction de la ligne 14. A l'origine, cette nouvelle ligne avait notamment pour but de remédier à la congestion sur la ligne A du RER, particulièrement dans la partie centrale comprise entre les gares d' "Auber" et de "Gare de Lyon"⁶. En effet, alors que la capacité à l'heure de pointe du matin était estimée à 44 000

⁶Le deuxième but à la construction de cette ligne 14 était la volonté de mieux desservir les 12^{ème} et 13^{ème} arrondissements, en particulier le Ministère de l'Economie et des Finances à Bercy, la future Bibliothèque Nationale et la ZAC Paris Rive Gauche en construction.

1.1. LE MARCHÉ DES TRANSPORTS COLLECTIFS URBAINS, UN MONOPOLE NATUREL ?

voyageurs, ce tronçon était emprunté par 50 000 à 60 000 voyageurs en 1989, et 60 000 à 70 000 voyageurs étaient prévus pour le milieu des années 90. Météor, dont le tracé est parallèle à la ligne A en suivant la Seine, devait délester la ligne A d'au moins 13 000 voyageurs par heure. Le prolongement jusqu'à Saint-Lazare n'ayant été mis en service que fin 2003, cet objectif n'a pas encore été atteint et la saturation sur le tronçon central est encore d'actualité, dans des proportions toutefois moindres que les prévisions réalisées au début des années 1990. Les fréquences entre les trains ont été réduites à un intervalle minimum durant l'heure de pointe, et le moindre incident dans un train (un malaise de voyageurs par exemple) entraîne des retards sur tous les trains suivants. A tel point que certains voyageurs préfèrent prendre la ligne 1 du métro, qui dessert l'ensemble des gares du tronçon central à l'exception de la gare "Auber", plutôt que le RER. En l'absence d'incidents, le trajet effectué par le métro est plus long que celui sur le RER, mais il est moins soumis aux aléas d'exploitation.

1.1.2.2 La meilleure gestion des ressources dans une optique de qualité de service

Dans le domaine des transports en commun, l'existence d'économies d'échelle se vérifie d'autant plus que la qualité de service est prise en compte.

Selon Keeler (1983) [96], les économies d'échelle dans ce contexte se traduisent plus par une meilleure utilisation des équipements et du personnel de conduite que par une meilleure utilisation des voies. Ces deux éléments sont pourtant parfaitement compatibles. Par exemple, dans un centre bus de la RATP, une partie des machinistes-receveurs est affectée à une ligne bien précise, tandis que les machinistes "hors-lignes" sont affectés à une ligne ou à une autre, selon les besoins de l'offre de service. Ils permettent de maintenir l'offre de service définie par les tableaux de marche en remplaçant, par exemple, les personnes malades ou en retard. Plus l'offre de service est importante, plus le nombre de machinistes-receveurs nécessaires est grand, et plus le risque d'avoir au moins un service non couvert augmente. L'utilité d'une personne hors-ligne s'accroît alors, car la probabilité d'avoir besoin de lui est plus grande. En outre, dans un centre bus, le départ des bus est échelonné au cours du temps, une même personne pourra donc remplacer plusieurs machinistes en retard et assurer le service normal. Enfin, certaines lignes de bus, notamment celles qui desservent Paris, ont des fréquences théoriques très élevées.

En pratique, ces fréquences sont soumises à des éléments extérieurs tels que les conditions de circulation, soumises à de nombreux aléas : manifestations, encombrements⁷, travaux, accidents... Sur les points d'arrêts des lignes à forte fréquence, les heures de passage des bus ne sont pas marquées : seul un intervalle de temps (minimal/maximal) entre deux bus est indiqué. Dans le cas de perturbations comme de défaillances de personnes ou d'équipement, les régulateurs vont pouvoir modifier à la marge les tableaux de marche, jusqu'à atteindre l'intervalle de temps maximal indiqué sur les points d'arrêt de la ligne. Ce type de régulation peut être transparent pour le client : d'une certaine façon la qualité de service est maintenue. Cette possibilité de réaffecter bus et machinistes-receveurs sur d'autres lignes, en jouant sur les fréquences entre deux autobus, permet de respecter les horaires indiqués pour les lignes à plus faible fréquence, comme certaines lignes de banlieue. Au total, la qualité de service peut être maintenue en utilisant au mieux les ressources en hommes et matériel, ce qui ne serait pas forcément le cas dans une petite entreprise, ayant un service moins étalé au cours du temps.

Pour Grimm et Harris (1983) [68], l'augmentation de la contexture du réseau permet d'offrir des trains ne s'arrêtant pas à un certain nombre de terminaux, mais ne justifie pas forcément l'utilisation de trains plus longs. A la RATP, les deux types d'économies d'échelle sont possibles. Sur le RER, par exemple, lors des heures creuses, sont mis en place des trains courts, ce qui permet de maintenir une certaine fréquence, tout en limitant les frais de nettoyage des rames restées au garage. En heures pleines sont mis en service des trains longs, ces services longs ne desservant pas forcément toutes les gares. Par exemple, sur la ligne ouest du RER A, en direction de Saint-Germain en Laye, les RER ayant pour terminus la gare du "Vésinet - Le Pecq" ne desservent pas tous la gare de "Nanterre - Ville" et/ou de "Nanterre-Préfecture" pendant les heures d'affluence, tandis que les RER ayant pour terminus Saint-Germain-en-Laye ne s'arrêtent pas aux gares du "Vésinet-Centre" et de "Chatou-Croissy". Ces trains, plus directs, améliorent la qualité de service : tout en maintenant une forte fréquence sur le tronçon central,

⁷La construction des couloirs de bus a permis de réduire l'impact de ces facteurs, du moins en partie (problèmes du non respect des couloirs pour livraison, par exemple). Dans Paris, la vitesse moyenne de circulation d'un bus est de 9 km/h en l'absence de couloirs protégés, contre 12 km/h lorsque le parcours est protégé. Le principal avantage des couloirs de parcours réside dans la diminution des temps de parcours (5 à 7 minutes pour la liaison Châtelet - Gare du Nord contre 15 minutes avant la création des couloirs de bus) mais surtout dans la plus grande régularité de ces temps de parcours, qui devrait être prévu de façon aussi fiable que sur le métro.

très fréquenté, ils permettent d'atteindre plus rapidement certaines destinations. Sur le réseau des bus, certains services ne desservent pas l'ensemble du parcours, mais effectuent un service "partiel", c'est-à-dire la partie la plus fréquentée à ce moment là. Ces services partiels, affectés aux secteurs où la demande est plus importante, maintiennent ainsi une qualité de service en terme de charge à bord des véhicules. Enfin, sur le réseau du métro, dans le cas de la ligne 14, la longueur des quais en stations est accrue par rapport aux lignes plus anciennes, même si pour l'instant circulent des rames de longueur standard. Quand cette ligne sera saturée, il sera possible d'accroître la capacité de chaque rame de métro.

Du fait de l'existence de coûts fixes extrêmement importants, le prix du transport collectif urbain ne peut s'établir au coût marginal : l'entreprise réaliserait un profit négatif. Cette inefficacité du marché du transport collectif urbain, associée à d'importantes économies d'échelle, incite à penser que l'organisation de ce marché correspond à un monopole naturel. Les effets vertueux de la concurrence, en termes de bas prix, de recherche de qualité et d'efficience, n'existent plus. Le monopole a la possibilité d'augmenter le prix pour une qualité de service moindre, d'introduire une discrimination entre les clients pour faire payer plus ceux qui ne peuvent se passer du service offert, ou d'exclure les clients ou les territoires trop peu rentables. Cette situation est particulièrement dommageable pour les services d'intérêt général, comme le transport collectif urbain, et amène à soulever la question de l'intervention des pouvoirs publics.

1.2 L'intervention nécessaire des pouvoirs publics

Selon Pigou (1932 [143]), les pouvoirs publics ont un rôle important à jouer dans la régulation d'un monopole naturel. L'équilibre budgétaire et la couverture de l'ensemble de la demande à un prix optimal ne peuvent être atteints qu'à la condition que les pouvoirs publics accordent des subventions à l'entreprise pourvoyeuse de services.

En plus de considérations sociales en termes de tarification et de service offert, les pouvoirs publics doivent penser en termes d'accessibilité régionale, de lutte contre la congestion excessive et de coûts environnementaux des autres modes de transports en contexte urbain. Par exemple, la volonté d'une desserte géographique étendue peut ne pas être considérée comme efficace au

sens économique, c'est néanmoins une volonté politique importante, à mettre en liaison avec la notion de "service public".

Pour mettre en avant la nécessaire intervention des pouvoirs publics en matière de transport collectif urbain, nous examinerons successivement la notion de service public (premier paragraphe)⁸, l'existence d'externalités positives au développement de l'offre de transport collectif (deuxième paragraphe), ainsi que la prise en compte nécessaire des autres types d'externalités (troisième paragraphe).

1.2.1 La notion de "service public"

Selon Duguit (1928) [49], *"toute activité dont l'accomplissement doit être assuré, réglé et contrôlé par les gouvernants, parce que l'accomplissement de cette activité est indispensable à la réalisation et au développement de l'interdépendance sociale, et qu'elle est de nature telle qu'elle ne peut être réalisée complètement que par l'intervention de la force gouvernante"* relève du service public. Pour qu'un service collectif puisse être considéré comme faisant partie du service public, celui-ci doit répondre à un besoin d'intérêt général. Cette notion d'intérêt général ne possède pas de définition juridique précise : il revient au législateur de la définir.

La notion de "service public" est une conception française. Dans les textes européens, les notions utilisées sont celles de "service d'intérêt économique général"⁹ et, plus récemment, de "service universel". La notion de "service universel" apparaît dans les réglementations du service des télécommunications (Livre Vert 1987) et concerne les services de base auxquels tout citoyen doit pouvoir accéder, à un prix abordable. Les principes auxquels doit se soumettre un service universel sont les suivants :

- l'universalité, c'est-à-dire l'accès de tous les citoyens à des conditions abordables. Ainsi, service public et service universel n'impliquent pas la gratuité, et englobent les services publics à caractère industriel et commercial ;

⁸ Ce paragraphe tire certains éléments d'un dossier réalisé par la Documentation Française sur la régulation des services publics, disponible sur www.ladocumentationfrancaise.fr/dossier_polpublic/regulation_sp/index.html

⁹ Ainsi, l'article 90§2 du Traité de Rome, qui légifère la concurrence dans les services marchands, pose que *"les entreprises chargées de la gestion de services d'intérêt économique général (...) sont soumises (...) aux règles de la concurrence, dans les limites où l'application de ces règles ne fait pas échec à l'accomplissement (...) de la mission particulière qui leur a été impartie"*.

- l'égalité qui correspond à l'idée de non-discrimination des clients, c'est-à-dire à un accès garanti au service quelle que soit la situation géographique¹⁰. Au delà du principe même, pour que les réseaux soient véritablement efficaces, ils doivent être accessibles au plus grand nombre. Ces investissements très lourds ne peuvent être réalisés sans l'intervention de l'Etat, entrepreneur par défaut au service de l'intérêt général ;
- la continuité, c'est-à-dire le fonctionnement permanent du service public. Au sein de la RATP, une forme de continuité du service public est réalisée grâce au réseau Noctambus, qui dessert un certain nombre de destinations la nuit. On peut y ajouter les premiers bus du matin et derniers bus du soir, qui transportent peu de voyageurs, mais répondent aux besoins de certaines personnes ayant par exemple des horaires de travail décalés.

Ce troisième principe est régulièrement au coeur des débats depuis une vingtaine d'années, car il peut justifier la mise en place d'un "service minimum" lors de conflits sociaux. L'instauration d'un tel service est néanmoins problématique, pour des raisons de sécurité et de respect du droit de grève. En effet, si le service minimum est défini en posant un nombre minimal de rames de métro tout au long de la journée, le réseau risque d'être rapidement saturé, ce qui comporte de nombreux risques pour les voyageurs (cohue, quais bondés, risques de malaise et d'accidents...). Si le service est assuré uniquement aux heures de pointe, le problème de saturation demeure. Surtout, pour assurer ce service, plus de 80% du personnel doit travailler, ce qui revient à restreindre le droit de grève.

A la RATP, en 1996 a été mis en place "l'alarme sociale", inspirée de ce que les Allemands ont défini comme le "principe de refroidissement". Avant d'en arriver à la grève, les syndicats et l'entreprise se sont engagés à se rencontrer pour trouver une solution ou, sinon, mettre noir sur blanc les raisons du conflit dans un document officiel. Le préavis de grève ne peut être déposé que cinq jours plus tard. Cet engagement n'est pas réglementé par une loi mais par un accord syndical, et a contribué à une forte diminution des grèves.

La notion de service public ajoute un principe directeur : l'adaptation au changement, i.e. l'ajustement du contenu du service public au progrès technique d'une part, et aux besoins des

¹⁰Du point de vue des utilisateurs des transports en commun, une discrimination existe entre les habitants de Paris, où le maillage des transports collectifs est dense, et la banlieue, moins bien desservie.

usagers d'autre part.

1.2.2 L'existence d'externalités positives au développement de l'offre de transport collectif urbain

L'une des raisons sociales de l'intervention publique en matière de transport collectif urbain provient du fait que ce type de transport produit moins de coûts externes que les autres modes de transport mécanisés, en particulier les véhicules particuliers. Amener les gens à prendre les transports en commun peut réduire la congestion sur les routes ou les accidents. Des externalités négatives concernant l'ensemble des personnes, y compris celles qui ne se déplacent pas par les transports mécanisés, peuvent être diminuées : le bruit et la pollution atmosphérique, ou encore la coupure du tissu urbain et du paysage, et la consommation d'espace.

1.2.2.1 La monétarisation des externalités en Ile-de-France

Le groupe de travail du Commissariat Général au Plan - Transports a calculé des valeurs tutélaires permettant de valoriser des éléments tels que les accidents, la pollution ou l'effet de serre. Par exemple, la monétarisation des accidents de personnes se fait à partir de la valeur tutélaire de la vie humaine (1.5 million d'euros pour un tué, 0.225 million pour un blessé grave, 0.033 million pour un blessé léger). Le montant des travaux nécessaires à la diminution de la pollution sonore permet de monétiser le bruit, et s'élève à 1 140 millions d'euros pour l'ensemble des voitures particulières, et 53 millions pour les transports collectifs. La congestion implique une perte de temps dans les transports collectifs par autobus, qui peut être évaluée et valorisée sur la base du salaire horaire moyen en Ile-de-France. Des valeurs tutélaires sont calculées pour la pollution et l'effet de serre.

Selon le STIF [156], les nuisances liées au transport de personnes ont coûté, tous modes confondus et pour la seule Ile-de-France, plus de 5.2 milliards d'euros en 2000 à la collectivité. 96% des coûts externes sont imputables aux seuls transports individuels. Dans ce coût total associé aux transports individuels, 41% proviennent des accidents, 23% de la pollution générée, 20% de la pollution sonore, 10% de la congestion et 6% de l'effet de serre.

1.2.2.2 Les externalités non monétarisées en Ile-de-France

D'autres externalités auraient pu être monétisées de la même façon, à savoir les effets de coupure ou les effets sur le paysage urbain [84]. Les effets de coupure regroupent l'ensemble des effets négatifs induits par la présence de voies routières à grande circulation, de voies ferrées ou de voies d'eau, infranchissables soit pour des raisons matérielles soit pour des raisons de sécurité. Ils sont subis essentiellement par les piétons, et se traduisent par un allongement des délais des parcours (contournement), des risques accrus, une diminution des relations de voisinage avec séparation des communautés et perte de lisibilité du territoire (le périphérique isole Paris et les banlieues proches par exemple). Ces effets de coupure pourraient être monétisés en calculant la disponibilité à payer des citoyens, ou en évaluant monétairement la perte de temps des piétons. Autre externalité négative, les infrastructures de transports ont un impact négatif sur le paysage urbain. Par exemple, les infrastructures de transport constituent une intrusion visuelle qui doit faire l'objet d'une insertion paysagère coûteuse. En outre, la rue se réduit à sa fonction de circulation, et constitue une gêne pour le piéton. La pollution atmosphérique, outre ses effets sur la santé, dégrade le patrimoine urbain et accroît le nombre de ravalement. Enfin, le paysage est dégradé à l'entrée des villes : l'effet de grands axes de circulation attire les commerces certes, mais aussi les panneaux publicitaires. Les infrastructures de transport collectif ont sans aucun doute des effets négatifs moins importants dans ces domaines.

1.2.3 Le lien entre transport et aménagement du territoire

L'extension des réseaux de transport, en commun ou individuel, est à mettre en relation, à long terme, avec l'aménagement du territoire. Il est sans doute judicieux qu'une seule personne morale gère l'ensemble de ces différents aspects.

1.2.3.1 Le phénomène de péri-urbanisation

Il existe des effets indirects à la modification du réseau (que ce soit son allongement ou l'accroissement de la qualité de service, en termes de fréquence pour les transports collectifs par exemple), le plus frappant étant le phénomène de péri-urbanisation ou étalement urbain des villes. Selon l'Institut d'Aménagement et d'Urbanisme de la Région Ile-de-France (IAURIF [87]),

les années 1970 sont caractérisées par la reconquête des centres anciens d'habitat et la création de nouveaux pôles (Bobigny, Créteil), de façon concomitante au prolongement des lignes de métro. L'étalement urbain se poursuit avec la réalisation du RER et le maillage autoroutier. Les améliorations en termes de temps de transport impliquent que, pour un temps de parcours stable (ce qui est le cas pour les relations domicile - travail depuis de nombreuses années), les distances parcourues peuvent être plus importantes. Or, en s'éloignant de Paris, les prix immobiliers ont tendance à diminuer et les surfaces foncières disponibles sont plus importantes, ce qui favorise l'accès à la propriété. Ces facteurs poussent à s'installer de plus en plus loin du centre de l'agglomération. Ainsi, la politique de transport a des conséquences sur l'aménagement du territoire.

Le nombre important des surfaces disponibles, ou le plus faible prix des loyers immobiliers, attire aussi les entreprises, qui peuvent se délocaliser de Paris. La proximité des transports collectifs et des dessertes autoroutières est un facteur important de délocalisation, comme le montre la montée en puissance de la ville de Saint-Denis. Or l'arrivée ou la création de nouvelles entreprises est un moyen important de créer des facteurs de croissance dans un endroit (recrutement de personnel, mais aussi création de commerces ou restaurants de proximité...).

1.2.3.2 Densité urbaine et utilisation de la voiture dans les zones péri-urbaines

Un des problèmes associés à la péri-urbanisation est qu'elle s'accompagne généralement d'un accroissement de la mobilité automobile, génératrice d'externalités négatives. Les zones péri-urbaines sont peu denses : emplois, commerces, administrations(...) sont des lieux éloignés les uns des autres. Pour ce type de micro-mobilité, la marche à pied n'est pas toujours envisageable, et les transports collectifs, peu présents, sont généralement mal adaptés. Ainsi, si les transports collectifs sont majoritaires pour les déplacements dans ou vers Paris, la voiture particulière est majoritaire pour les déplacements de banlieue à banlieue.

La répartition modale pour la mobilité interne au quartier de domicile est la suivante [95] :

	Urbain central	Sub-urbain	Ville nouvelle	Péri-urbain
Marche à pied et vélo	72%	50%	19%	25%
Automobile	18%	50%	81%	75%

Cette répartition s'explique par le fait que le choix du mode de transport est lié principalement à la structure urbaine (localisation et densité des populations et des emplois, type d'habitat), la structure de la population (composition des ménages), la contexture des réseaux (forte densité des transports collectifs dans Paris, réseau de lignes radiales en banlieue), et la disponibilité d'une place de stationnement à la destination pour la voiture particulière.

1.2.3.3 Des risques moins marqués lors de l'amélioration des transports collectifs

La faible densité dans les zones péri-urbaines est moins marquée lorsque l'allongement du réseau se réalise dans le domaine des transports collectifs. Selon Pierre Merlin [122], dans une même agglomération, la création d'une autoroute sub-urbaine (comme l'autoroute du sud) ouvre de vastes espaces à l'urbanisation, tandis qu'une ligne de métro régional comme le RER (RER B Sud, anciennement ligne de Sceaux, par exemple) entraîne une urbanisation plus concentrée, particulièrement dans le périmètre autour de la gare accessible par la marche à pied. L'urbanisation est plus vaste quand elle est associée à l'automobile : il est possible de prolonger son trajet à partir d'un échangeur sans changer de mode de transport, tandis qu'un autre mode de transport (ou une correspondance entre deux bus) est nécessaire pour se rabattre sur les transports en commun.

Dans les zones péri-urbaines, les prix immobiliers augmentent à proximité des infrastructures de transport. Il existe des courbes d'iso-prix des transactions foncières autour des pôles d'emplois et des infrastructures de transport. La différence d'accessibilité trie les catégories sociales péri-urbaines et pousse les moins fortunées en marge des zones urbaines, entraînant une diminution de la mixité sociale. Mais l'allongement des transports collectifs ne crée pas de réelle exclusion : à proximité des infrastructures de type RER, par exemple, le rabattement par bus est fréquent. Au contraire, pour les banlieues accessibles presque uniquement par l'automobile, les personnes

ne pouvant en acquérir une sont exclues.

Enfin, dans le cas d'entreprises installées en dehors de Paris, l'existence de moyens de transports collectifs performants peut diminuer le recours à l'automobile pour les trajets domicile - travail. C'est par exemple le cas des villes le long du RER A telles que la Défense, Rueil-Malmaison ou la nouvelle gare de Val d'Europe. Ainsi, 300 000 personnes se rendaient chaque jour à La Défense au milieu des années 90, et le seuil des 450 000 sera prochainement dépassé. Parallèlement, la place de la voiture n'a cessé de reculer depuis dix ans et plus de 80% des salariés empruntent aujourd'hui les transports en commun.

1.2.3.4 Une approche cohérente des politiques urbaines et de déplacement

Le transport public est nécessaire pour garantir le développement durable d'une ville et la qualité de vie de ses habitants. Ces objectifs sont atteints à travers une approche cohérente des politiques urbaines et de déplacement, et supposent la coopération de tous les protagonistes : élus politiques, autorités organisatrices, entreprises de transport...

La politique de développement des transports publics en Ile-de-France s'inscrit dans le cadre de la politique de développement durable et s'appuie sur les principes définis dans trois documents officiels :

- le Schéma Directeur de la Région Ile-de-France 1994 - 2015 ("SDRIF"),
- le Contrat de Plan Etat - Région 2000 - 2006 ("CPER"),
- le Plan de Déplacements Urbains ("PDU").

Le SDRIF fixe la stratégie d'aménagement de la région pour une vingtaine d'années, en indiquant les grandes orientations en matière d'usage des sols et d'architecture des réseaux routiers et des transports collectifs. Il met l'accent sur la priorité donnée aux transports en commun, avec la réalisation de liaisons de banlieue à banlieue, par le développement de transports en commun de type tramways ou bus en sites propres en complément des infrastructures ferrées.

Le CPER définit la mise en oeuvre du SDRIF en termes d'infrastructures de transport. Il précise les actions que l'Etat et les régions doivent mener de façon conjointe, et leur mode de financement. Le XII^{ème} plan prévoit l'allocation d'un budget de 4 milliards d'euros aux transports collectifs franciliens (pour un budget total de 9 milliards d'euros). D'ici la fin 2006 ont été prévus

la construction de 50 kilomètres de lignes de tramway, le prolongement de 4 lignes de métro vers la banlieue ainsi que le prolongement de Meteor, ainsi que la création de lignes de bus en site propre (voir chapitre 3).

La loi du 30 décembre 1996 sur l'air et l'utilisation rationnelle de l'énergie a rendu obligatoire l'élaboration d'un plan de déplacements urbains pour toutes les agglomérations de plus de 100 000 habitants en Ile-de-France. Le Plan de Déplacements Urbains d'Ile-de-France (PDUIF) a été élaboré par l'Etat, et approuvé par arrêté interpréfectoral le 15 décembre 2000. Il définit les principes d'organisation des déplacements de personnes et du transport de marchandises, de la circulation et du stationnement. Trois orientations principales sont soulignées :

- la diminution du trafic automobile de façon différenciée selon les zones de l'agglomération et leur desserte en transport collectif (-3% sur l'ensemble de la région, -5% dans Paris, -3% en Petite Couronne, -2% en Grande Couronne),
- l'augmentation de la part des transports collectifs,
- l'accroissement du recours aux modes non motorisés (marche à pied et vélos).

La mise en oeuvre du PDUIF s'effectue au sein de comités, à des échelles différentes (région, département, localité voire quartier). En particulier, le PDU définit un réseau de lignes d'auto-bus qui doit constituer le réseau régional de transports collectifs en complément du réseau de tramway et des lignes ferrées à grand gabarit. Les aménagements de voirie seront réalisés sous la maîtrise d'ouvrage des gestionnaires des voiries concernées, pour accroître vitesse, régularité, fréquence, accessibilité de ces lignes. L'Etat et la Région participeront au réaménagement des pôles d'échanges, afin de favoriser l'intermodalité et le maillage des réseaux.

En présence d'un monopole public, l'intervention des pouvoirs publics apparaît souhaitable, d'autant plus que l'on se situe dans des services ayant trait à l'intérêt général, et nécessaires à l'intégration des personnes et des territoires. A défaut de pouvoir rétablir la concurrence, la nationalisation ou la régulation par les pouvoirs publics semble préférable au maintien d'un monopole privé. Néanmoins, la mise en place d'un monopole public ne règle pas l'ensemble des problèmes, puisqu'il faut remplacer le moteur du profit et la pression de la concurrence par l'exigence de rentabilité sociale, en introduisant la notion de compétitivité.

1.3 L'introduction d'éléments de compétitivité dans le marché des transports collectifs urbains

Pour éviter les dérives de l'intervention publique (rappelées dans le premier paragraphe), il est nécessaire d'introduire des éléments de compétitivité. Ce n'est pas seulement "l'efficacité" (atteinte des objectifs fixés par la collectivité) qui est recherchée, mais "l'efficience", c'est-à-dire l'optimisation des coûts nécessaires en vue de cet objectif donné. Plusieurs manières d'atteindre un certain niveau d'efficience sont possibles :

1. la concurrence, réelle ou potentielle, qui relève de la notion de marché "contestable", est abordée dans le deuxième paragraphe ;
2. l'introduction d'objectifs d'efficacité par l'autorité régulatrice (troisième paragraphe).

1.3.1 Les imperfections de la régulation étatique dans une grande entreprise

1.3.1.1 Les risques d'inefficacité dans l'utilisation des facteurs de production

En matière de théorie des organisations, l'absence de compétition induite par le contrôle public dans un monopole naturel peut conduire, selon Liebenstein (1966) [115], à une "inefficacité X". Cette inefficacité concerne principalement le manque de motivation des employés. En effet, sachant que l'équilibre budgétaire sera atteint en fin d'exercice quelle que soit l'action des salariés, ceux-ci sont moins incités à produire efficacement et, en particulier, à minimiser les coûts. De même, sans compétition, ils préféreront sans doute maximiser leur utilité propre, c'est-à-dire minimiser la désutilité de l'effort dans une certaine mesure, d'autant plus que la nature des contrats de travail est généralement incomplète (le salaire y est écrit, pas la productivité attendue). Ainsi, l'accroissement des subventions diminue dans une certaine mesure l'efficacité totale, et dans une large mesure les productivités partielles, essentiellement celle du travail (selon Oum et Yu (1994) [130], l'élasticité de la productivité du travail par rapport au niveau de subvention serait égale à - 0.12).

Ainsi, l'asymétrie d'information entre le producteur et l'autorité régulatrice offre au premier de multiples possibilités pour tromper l'autre quant à la réalité de ses besoins financiers et de

1.3. L'INTRODUCTION D'ÉLÉMENTS DE COMPÉTITIVITÉ DANS LE MARCHÉ DES TRANSPORTS COLLECTIFS URBAINS

ses performances de gestion : il n'est pas sûr que l'autorité organisatrice connaisse la réelle productivité des facteurs.

1.3.1.2 Les risques de "déséconomies d'échelle", inhérents à toute grande entreprise

Les grandes entreprises sont généralement organisées selon une hiérarchie des qualifications pour que les fonctions de contrôle puissent s'exercer correctement. Par exemple, traditionnellement, les machinistes-receveurs, directement productifs, sont sous l'autorité d'un responsable d'équipe de ligne, dépendant lui-même du chef du centre bus. Les directeurs de centre bus sont dirigés par le directeur du département "Bus" et ainsi de suite. D'autres personnes, à priori indépendantes, sont sous le contrôle hiérarchique direct du directeur du centre bus : le responsable commercial, celui des ressources humaines, le contrôleur de gestion, le responsable de la voirie et des points d'arrêts, celui de la prévention - sécurité... Dans ce même centre bus, une équipe de maintenance peut être présente, et appartient à un département différent donc une hiérarchie différente. Selon Williamson (1967) [163], la transmission de l'information entre les différents niveaux hiérarchiques est en général imparfaite, et peut conduire à des distortions qui sont source de déséconomies d'échelle. La multiplication et la relative inefficacité des fonctions de contrôle peut augmenter les coûts salariaux de façon plus que proportionnelle à l'accroissement de la production.

Au sein de la RATP, l'organisation de l'entreprise a été pensée de manière à réduire ce type d'inefficacité : en favorisant la décentralisation. Ainsi, chaque centre bus a des objectifs individuels, ce qui revient implicitement à introduire une certaine concurrence entre les différents centres. En outre, une certaine liberté d'organisation est permise au sein de chaque centre, permettant ainsi d'adapter la gouvernance aux problèmes rencontrés sur place, et d'innover. Par exemple, les machinistes-receveurs ont droit à 4 semaines de vacances consécutives entre les mois de juin et de septembre. Pour des raisons de continuité de service, tout le monde ne peut partir en même temps. Ainsi, une personne peut avoir à prendre la totalité de ses vacances au mois de juin, alors que ses enfants sont encore à l'école. Le centre bus de Malakoff a introduit un nouveau système de vacances d'été : au lieu de 4 semaines garanties, les machinistes ont droit à 3

semaines seulement, mais au cours des mois de juillet-août, les plus prisés. Cette innovation, qui satisfait à la fois les machinistes et leurs encadrants (diminution des tensions lors de l'attribution des congés et meilleur dialogue social) a été permise par la relative liberté de gestion au sein d'un centre bus. Elle sera peut être adoptée ultérieurement par d'autres centres.

Enfin, l'obtention de résultats dans chaque centre bus est facilitée par les orientations actuelles de gestion, à savoir : la diminution de la taille des équipes (entre 30 et 40 machinistes-receveurs), l'allègement de la charge de travail des chefs de ligne pour leur permettre d'aller plus souvent sur le terrain, et l'octroi aux chefs de ligne de plus grands leviers d'action (congés ou avancement des machinistes).

La lourdeur bureaucratique, dans une grande entreprise, peut entraîner une plus grande difficulté à connaître l'ensemble des paramètres qui conditionnent les prises de décision. Néanmoins, dans le centre bus de Malakoff, la hiérarchie n'est pas aussi verticale que celle présentée en début de paragraphe. La volonté de transversalité est forte. Ainsi, l'équipe des encadrants du département "bus" mais aussi du département de la "maintenance" se réunit périodiquement pour examiner l'ensemble des implications d'une décision. En outre, lorsqu'un machiniste a un problème, il peut s'adresser à son chef de ligne, qui est bien informé et peut directement résoudre son problème, ou l'envoyer vers la personne compétente.

1.3.2 Est-ce un monopole "contestable" ?

La théorie de "marché contestable" spécifie un certain nombre de conditions selon lesquelles les entreprises agiront de façon efficace quel que soit le nombre d'entreprises sur le marché. La compétition potentielle peut être préférable à la compétition réelle, et conduire à un monopole naturel qui sert le marché de façon efficace étant donné la contrainte d'équilibre financier.

1.3.2.1 La notion de "marché contestable"

Selon la définition introduite par Baumol, Panzar et Willig (1982) [23], un marché est dit "contestable" si :

1. L'entrée y est absolument libre. La liberté d'entrer sur le marché est caractérisée d'une part par l'absence d'obstacle institutionnel. D'autre part, les entreprises entrant sur ce

1.3. L'INTRODUCTION D'ÉLÉMENTS DE COMPÉTITIVITÉ DANS LE MARCHÉ DES TRANSPORTS COLLECTIFS URBAINS

marché ne doivent avoir aucun désavantage par rapport à la firme déjà installée, en termes de techniques de production, de qualité des produits ou d'image de marque.

2. La sortie du marché s'effectue sans aucun coût. Cette condition implique que tous les facteurs physiques acquis peuvent être revendus sans autre perte que leur dépréciation physique normale. En d'autres termes, il n'y a pas de coût irréversible.

En outre, un marché contestable doit être caractérisé par l'absence de profit. Sinon, une autre entreprise pourrait entrer sur le marché, réaliser la même production pour des coûts identiques, tout en pratiquant un prix légèrement inférieur pour détourner tout ou partie de la clientèle de la firme installée. Cette nouvelle entreprise toucherait alors un profit positif. La firme installée, pour récupérer sa clientèle, diminuerait son prix à son tour. Deux issues sont alors possibles : soit les deux entreprises diminuent tour à tour le prix de vente de leur produit jusqu'à ce que le profit soit nul, et deux entreprises coexistent, soit la deuxième entreprise quitte le marché après avoir réalisé un profit positif. Dans ce deuxième cas, si la première entreprise maintient son prix de vente initial, qui lui permet de réaliser des profits positifs, elle s'expose à l'arrivée de nouveaux entrants. En diminuant son prix de façon à avoir un équilibre budgétaire, elle reste seule sur le marché, sans perte de bien-être pour la société.

Dans le cas d'un monopole, cette idée peut être exprimée de façon différente. Sur un marché contestable, aucune inefficacité en termes d'allocation des ressources et de détermination des coûts ne devrait avoir lieu : des coûts anormalement élevés incitent d'autres entreprises à entrer sur le marché.

1.3.2.2 "Contestabilité" du quasi-monopole de la RATP

Les barrières à l'entrée du marché des transports collectifs urbains Jusqu'en décembre 2000, la RATP était en situation de monopole public dans les transports parisiens, de façon institutionnelle. Elle était présente à l'étranger via la société Systra, filiale commune avec la SNCF, qui propose des services d'ingénierie (par exemple sur les métros de Hongkong, de Buenos Aires ou du Caire). En France, elle pouvait fournir une assistance technique ou du conseil aux exploitants et élus affiliés au réseau Agir pour le transport public, qui regroupe des régies municipales ou des sociétés d'économie mixte contrôlées par les communes (Marseille, Poitiers,

Annecy, Niort...). Mais son activité de conception et d'exploitation des réseaux demeurait limitée à l'Ile-de-France.

La loi Solidarité et Renouveau Urbain (SRU) du 13 décembre 2000 permet à la RATP, par le biais de ses filiales, de répondre aux appels d'offre dans l'hexagone et à l'étranger¹¹. Ces filiales ne peuvent pas bénéficier de subventions de l'Etat, des collectivités locales. Son activité de conception et d'exploitation de réseaux de transport public n'est plus cantonnée à l'Ile-de-France comme c'était le cas auparavant¹². En contrepartie de la conquête possible de nouveaux marchés, existe une réciprocité éventuelle : les transporteurs urbains et inter-urbains de Province et de l'étranger devraient pouvoir jouer un rôle sur le marché francilien. L'échéance n'est pas fixée, mais, selon le GART (Groupement des Autorités Responsables des Transports), ce bastion devrait tomber au cours des années 2008 - 2010. Jusque là, il semble justifié de considérer qu'il existe une barrière à l'entrée sur le marché des transports collectifs franciliens.

Une des raisons au fait que la RATP est l'entreprise dominante en termes de transports collectifs en Ile-de-France, provient de raisons historiques. Entrée en premier sur ce marché, la RATP possède sans doute des avantages en termes de coût suffisants pour décourager l'entrée de nouvelles entreprises sur le marché, notamment en termes de coûts d'équipements. Le "savoir-faire" acquis au cours des années en terme de recrutement des agents, de gestion de conflits internes ou externes, la diversité des points de vente jusqu'à leur implantation chez certains burocrates, ainsi que le développement de cartes d'abonnements, notamment annuelles, constituent

¹¹Le texte précise que : "En dehors de la Région Ile-de-France et à l'étranger, la RATP peut également par l'intermédiaire de filiales, construire, aménager, exploiter des réseaux et des lignes de transport public de voyageurs dans le respect réciproque des règles de concurrence. Ces filiales ont le statut de sociétés anonymes. Leur gestion est autonome sur le plan financier dans le cadre des objectifs du groupe. Elles ne peuvent notamment pas bénéficier de subvention attribuée par l'Etat, le Syndicat des transports d'Ile de France (STIF), et les autres collectivités publiques au titre du fonctionnement et de l'investissement des transports dans la région Ile-de-France".

¹²L'ouverture des marchés extérieurs a conduit la RATP, fin avril 2002, à finaliser son alliance avec Transdev, premier exploitant mondial de tramways sur le marché concurrentiel (Nantes, Montpellier, Strasbourg et Orléans) et dont les développements à l'étranger l'ont notamment mené en Australie. La RATP est entrée à hauteur de 25% dans Transdev, via une augmentation de capital de 50 millions d'euros. De son côté, Transdev a pris 25% de RATP-Développement, qui regroupe les filiales France et international. De façon concrète, les deux entreprises répondent conjointement aux appels d'offre internationaux. Pour les appels d'offre au niveau national, la RATP répond à ceux concernant le métro, et Transdev à ceux concernant les tramways et les bus.

Cette alliance devrait favoriser la réussite des réponses aux appels d'offre. Pour l'instant, la RATP est déjà implantée dans plusieurs agglomérations française (Annemasse, Aix en Provence, Mulhouse, Clermont-Ferrand...) mais aussi à l'étranger : Italie (ex : tramway de Florence) ou Grèce (ex : réseau RER d'Athènes), mais aussi en Angleterre, en Allemagne et en Europe de l'Est par le biais de sa filiale ferroviaire avec Transdev nommée Eurailco.

1.3. L'INTRODUCTION D'ÉLÉMENTS DE COMPÉTITIVITÉ DANS LE MARCHÉ DES TRANSPORTS COLLECTIFS URBAINS

un frein à l'entrée de nouvelles entreprises.

Les barrières à la sortie du marché des transports collectifs urbains Une autre barrière à l'entrée, mais surtout à la sortie du marché, est à mettre en relation avec l'importance de l'infrastructure à mettre en place, particulièrement sur le réseau ferré. L'aménagement des voies, des tunnels, les terminaux de contrôle, la signalisation ou même le matériel roulant représentent un investissement très coûteux, et ont une vie très longue. Ce type de matériel est difficile à redéployer ailleurs, et peut être considéré comme un investissement irréversible.

Selon Tye (1990) [158], cette objection est peu recevable : les coûts irréversibles existent, mais ils dépendent de l'offre d'infrastructure et non des opérations. Il est envisageable, comme c'est le cas pour la SNCF, que les activités de fourniture d'infrastructures lourdes et de service commercial soient séparées. Autrement dit, la RATP pourrait gérer la mise en place et l'entretien des voies et autres équipements, une autre entreprise, régulée, s'occupant de transporter les voyageurs, ou inversement. Sur l'ensemble du RER A, par exemple, les trains qui circulent appartiennent tous à la RATP. Néanmoins, il existe une interconnexion avec la SNCF : sur les branches de Cergy et de Poissy, les trains sont conduits par des agents SNCF jusqu'à la gare de Nanterre-Préfecture, où a lieu un changement de conducteur, le nouveau conducteur étant un agent RATP. La SNCF dispose en outre de lignes de RER qui lui sont entièrement propres, à savoir les lignes C et D. Ainsi, malgré des signalisations différentes et d'autres spécificités qui devraient être apprises, il serait possible d'envisager que le transport de voyageurs par le RER soit réalisé par la SNCF sur des lignes RATP, et inversement.

En outre, si les infrastructures fixes ont une durée de vie très longue, les infrastructures mobiles, telles que les bus, durent moins longtemps. Leur remplacement s'explique par leur obsolescence physique plus que par leur obsolescence technique. Leur redéploiement est aussi plus aisé : un bus peut être réaffecté à une autre ligne de bus, même non-RATP, ou à un autre type de service que celui qui lui était assigné (transport d'enfants...).

L'existence d'une forme actuelle de concurrence Malgré ces barrières à l'entrée ou à la sortie du marché, le marché des transports collectifs urbains est d'une certaine manière soumis à une concurrence.

- Au sein même de la RATP, existe une concurrence entre les différents modes de transports : métro, RER et bus. Les reports entre les modes sont fréquents, surtout à l'intérieur de Paris où le maillage des transports collectifs est très fin.
- En outre, d'autres modes de transports sont possibles : la concurrence de l'automobile est indéniable. Elle dépend entre autres de la localisation et densité des populations et des emplois, de la structure de la population, de la contexture des réseaux, et la disponibilité d'une place de stationnement à la destination pour la voiture particulière.
- Enfin, le savoir-faire de la RATP est mis en concurrence avec celui d'autres entreprises dans des domaines ne touchant pas directement au domaine de l'offre de service. Certains types de matériel roulant, ferroviaire ou pour les bus, peuvent durer jusqu'à 45 années s'ils font l'objet d'une rénovation pendant la 25^{ème} année. A cette occasion, le secteur de la maintenance propre à la RATP peut être mis en concurrence avec d'autres entreprises par le biais d'appels d'offre.

Ainsi, il existe certains éléments qui poussent à penser que le marché des transports collectifs offerts par la RATP est un monopole contestable, puisque une concurrence potentielle ne peut être écartée théoriquement. Institutionnellement, néanmoins, il existe une barrière légale à l'arrivée de nouveaux entrants¹³. Les risques d'inefficacité demeurent, d'où l'intérêt pour l'autorité régulatrice de mettre en place des mécanismes incitatifs.

1.3.3 Les objectifs et incitations du contrat STIF - RATP

Une solution aux problèmes d'inefficacité réside dans la mise en place de mécanismes de rémunération qui incitent les salariés à contribuer le plus efficacement possible au bon fonction-

¹³Notons qu'il existe en outre une barrière légale à l'arrivée de modes alternatifs sur le marché monopolistique de la RATP. Par exemple, les taxis "collectifs privés" qui se promèneraient sur des itinéraires quasi-fixes ou libres sont interdits. Ces taxis jouent un rôle très important dans de grandes capitales (par exemple Santiago du Chili) et, plus généralement, dans beaucoup de pays en développement. Pour plus de renseignements, le lecteur est invité à lire l'article de T. Zhiri Oualalou intitulé "*les Requins blancs*" de Casablanca et la sécurité" et présenté à la conférence ATEC-ITIS en janvier 2004 à Issy les Moulineaux.

Une plus grande contestabilité existe dans d'autres pays tels que le Nouvelle Zélande ou le Chili, où existe un autre mode de régulation : l'autorité organisatrice définit la structure des lignes, leurs fréquences et horaires et ouvre ensuite chaque marché (ligne ou groupe de lignes) à la concurrence.

1.3. L'INTRODUCTION D'ÉLÉMENTS DE COMPÉTITIVITÉ DANS LE MARCHÉ DES TRANSPORTS COLLECTIFS URBAINS

nement de l'entreprise. Cette efficacité se traduit par la non-rétention de leur information et, surtout, par la fourniture du niveau d'effort minimal requis par leur poste de travail.

Le 12 juillet 2000, le régime des relations entre le STIF et les entreprises publiques de transport a été modifié par la contractualisation ([156]). Jusqu'alors, la RATP et la SNCF n'avaient pas de description précise de l'offre de service qu'elles étaient censées mettre en place. En outre, les comptes étaient automatiquement équilibrés en fin d'année. Cette indemnité compensatrice ne constituait en rien à une incitation à produire plus efficacement : un accroissement des coûts ou une diminution des recettes conduisaient à une augmentation de la subvention, telle que l'équilibre budgétaire soit réalisé en fin d'année.

Les contrats RATP - STIF et SNCF - STIF sont triennaux et comportent deux volets : la description de l'offre de service à fournir, et la rémunération prévue pour le service rendu. Examinons chacun de ces points pour les contrats RATP - STIF 2000 - 2003 et 2004 - 2007.

1.3.3.1 La description du service demandé par l'autorité régulatrice de transport.

Ce premier volet du contrat qui lie la RATP au STIF stipule d'une part le service de référence¹⁴, ainsi que les composantes qui y sont liées en matière d'information des voyageurs, d'accessibilité aux personnes à mobilité réduite, de sécurité ou de traitement des réclamations des voyageurs. Ces objectifs annuels de production sont assortis d'une franchise, de façon à tenir compte des écarts normaux qui peuvent apparaître entre le service théorique et la mise en oeuvre quotidienne. Cette franchise s'élève à 3% pour le métro, 2% pour le RER, 3% pour les TCSP, 3.5% pour les lignes PDU, 3% pour le réseau de bus de banlieue et 4% pour le bus de Paris.

D'autre part, les modalités de modification de ce service de référence sont énoncées :

- Les modifications importantes ayant une incidence financière doivent faire l'objet d'un avenant au contrat. Par exemple le 5^{ème} avenant (2002) définit les conditions dans les-

¹⁴Le contrat fixe un volume annuel de production de service sur les réseaux :

	Volume de l'offre de référence en 2004
Métro	44.37 millions de trains*kilomètres
RER A	8.19 millions de trains*kilomètres
RER B	4.16 millions de trains*kilomètres
Réseau Régional de Surface	111.85 millions de kilomètres

quelles la RATP doit mettre en place une offre de transport renforcée lors de nuits festives exceptionnelles (Nuit de la Saint-Sylvestre, Fête de la Musique ou autre comme la Nuit Blanche)¹⁵, prend en compte la création du Ticket et du titre "Congrès", ou précise les contreparties financières à la gratuité des transports lors de la nuit de Saint-Sylvestre.

- Les modifications de moyenne importance qui ont trait à des problèmes de coordination entre les transporteurs ont des procédures inchangées par la contractualisation.
- Enfin, les modifications mineures de service doivent simplement faire l'objet d'une information au STIF.

Dans le nouveau contrat 2004 - 2007, l'offre de service n'est plus appréciée globalement, mais ligne par ligne, chacune faisant l'objet d'un suivi mensuel.

1.3.3.2 La rémunération prévue pour le service rendu à la RATP.

Le second volet du contrat précise les composantes de la rémunération du transporteur, à savoir :

1. Les recettes totales du trafic, calculées sur la base de barèmes de référence des titres et non sur la base des tarifs publics des titres. Le prix de référence contractuel du voyage est calculé sur la base du prix des services, de façon à couvrir le coût global de production de la RATP. Il est différent du prix public des titres, qui dépendent d'un choix public, et est réactualisé chaque année. Les recettes totales comprennent à la fois des recettes directes venant des usagers et des compensations tarifaires venues du STIF. Elles ont augmenté de +1.9% entre 2000 et 2001 (soit + 46.9 millions d'euros après mise en oeuvre du mécanisme de partage avec le STIF) et de +3.3% entre 2002 et 2001.

Le nouveau contrat 2004 - 2007 a prévu une augmentation des prix de référence de 0.3% par an en sus de l'inflation contre les 0.8% requis en moyenne dans la profession. Pour atteindre le résultat prévisionnel calibré, la RATP devra diminuer ses frais d'exploitation

¹⁵ Ainsi, pendant les nuits festives exceptionnelles, les services sont maintenus entre 1h15 et 5h30 sur les lignes 1, 2, 4, 6, 9 et 14 avec ouverture des stations importantes et/ou de correspondance (+ 10 900 trains-km), les RER A et B doivent fonctionner dans leur totalité (+ 2 219 trains-km), tandis que le service Noctambus est aménagé et renforcé (+ 5 300 véhicules-km). Ce complément d'offre donne lieu à une rémunération additionnelle.

1.3. L'INTRODUCTION D'ÉLÉMENTS DE COMPÉTITIVITÉ DANS LE MARCHÉ DES TRANSPORTS COLLECTIFS URBAINS

ou accroître les recettes totales de trafic grâce à une augmentation de la fréquentation payante supérieure à celle prévue dans le contrat. Ceci explique que la lutte contre la fraude soit devenue un enjeu crucial pour la RATP, et que l'entreprise ait accepté de financer seule le projet "Bus Attitude"¹⁶.

2. Une rémunération de la vente correspondant à 6% du chiffre d'affaires des ventes, qui lui a permis en 2001 d'obtenir 2.4 millions d'euros, soit une progression de + 3.3% par rapport à l'année précédente.
3. Et une compensation forfaitaire qui permet de tenir compte des contraintes spécifiques des entreprises publiques (financement des investissements, ou contraintes sociales).

En outre est prévu un triple mécanisme d'intéressement :

1. L'intéressement lié au volume des ventes, au taux de 40% à l'intérieur d'une "bande passante" de 2%, avec un intéressement supplémentaire de 10% au dessus et en dessous de cette bande de 2%. Ce mécanisme incite à accroître le volume des ventes, et réduit les pertes en cas de mauvaise conjoncture. Supposons par exemple que le contrat prévoit un volume de vente égal à 100. La bande passante est l'intervalle compris entre 98 et 102. Si le volume des ventes est égal à 100, l'intéressement est nul. Si le résultat obtenu est égal à 102, l'intérêt est égal à $(102 - 100) * 40\% = 0.8$; si le résultat obtenu est égal à 105, et dépasse la borne supérieure de la bande passante de 102, l'intérêt est égal à $(102 - 100) * 40\% + (105 - 102) * 10\% = 1.1$. Dans le cas où l'entreprise n'atteint pas ses objectifs et obtient 95, au lieu de perdre 5, elle perd $(100 - 98) * 40\% + (98 - 95) * 10\% = 1.1$.
2. Une réduction de la rémunération de l'offre en cas de non-réalisation de l'offre (dans le cas

¹⁶Expérimentée par la RATP sur une dizaine de lignes de bus dans le nord-ouest de Paris depuis septembre 2003, la "Bus Attitude" est une démarche consistant à redonner du sens aux règles de base d'utilisation du bus et à certains principes régissant la vie commune à bord.

Les règles de conduite sont les suivantes :

- en attendant le bus : *"Au point d'arrêt, préparez votre titre de transport et faites signe au conducteur"*
- en montant dans le bus : *"Montez par la porte avant, présentez votre titre ou composez votre ticket et avancez vers le fond du bus"*.
- en sortant du bus : *"Demandez l'arrêt et sortez par la porte arrière"*.

Parallèlement à cette campagne d'amélioration de la relation de service entre la RATP et les voyageurs, le développement des valideurs NAVIGO dans les bus permettra de généraliser la "validation systématique" du titre de transport.

de mouvements sociaux, par exemple) avec une franchise de 3%. Ainsi, le renforcement progressif prévu au dernier trimestre 2001 pour améliorer l'offre du métro le week-end et durant l'heure de pointe du soir n'a pu que partiellement être réalisé, ce qui a donné lieu à un remboursement par la RATP de 0.3 millions d'euros en 2001.

3. Et un système de bonus-malus lié à la qualité de l'offre. Des critères de qualité de service ont été fixés au début du contrat, en prenant en compte la demande des clients, mais aussi le maintien de coûts raisonnables. Des mesures régulières sont effectuées. Si la situation constatée est meilleure que l'objectif à atteindre, la RATP voit sa rémunération affectée d'un bonus, dans le cas contraire elle est affectée d'un malus¹⁷.

En outre, la RATP dispose d'autres recettes commerciales, les recettes hors-transport, telles que les recettes publicitaires, les recettes issues des commerces dans le domaine RATP, les commissions sur les produits annexes vendus (cartes de téléphone, passeports Disneyland Resort Paris) ou encore les amendes payées dans les temps à la RATP.

Les marges dégagées au cours d'un exercice permettent à la RATP de financer son développement en dehors de l'Ile-de-France. Elles favorisent aussi la motivation du personnel, car les agents de la RATP sont intéressés aux résultats de l'entreprise. Tous ces éléments concourent à motiver les agents, et à favoriser la recherche de compétitivité.

* * *

¹⁷Les principaux critères de qualité de service pour le réseau ferré (métro + RER) portent sur les 5 points suivants : la régularité, l'accueil au guichet, la netteté des stations, la disponibilité des équipements de vente, et des escaliers mécaniques. Pour le réseau bus sont pris en compte les facteurs suivants : le service rendu par le machiniste-receveur et l'information aux points d'arrêt. A titre d'exemple, dans au moins 96% des cas, le client du métro doit attendre moins de 3 minutes en heure de pointe. En outre, lorsqu'il se présente devant un escalier mécanique ou un ascenseur, celui-ci doit fonctionner dans au moins 9 cas sur 10.

En 2000 et 2001, ces critères de qualité ont permis à la RATP de bénéficier de bonus, au niveau élevé de 85% du maximum prévu par le contrat, soit 9.1 millions d'euros en 2000, et 8.6 millions d'euros en 2001. En 2002, le bonus sur la qualité de service a diminué de - 14% et s'élève à 7.4 millions d'euros. Cette diminution est imputable à deux facteurs dépendant peu des agents RATP : la mauvaise évaluation de la netteté des stations de métro (le nettoyage est sous-traité) et des problèmes de régularité sur le RER B (60% des points perdus s'expliquent par des problèmes d'exploitation sur la portion SNCF de la ligne).

Le nouveau contrat STIF -RATP pour les années 2004 - 2007 a accru le nombre de ces indicateurs de qualité de service. Il a en outre mis en place un indicateur hebdomadaire de situation inacceptable par ligne de métro. A performances inchangées, le bonus dont bénéficie la RATP ne devrait plus représenter que 30% du maximum possible (de 3 à 3.5 millions d'euros) au lieu des 80% connus antérieurement (de 7.5 à 8 millions d'euros).

1.3. L'INTRODUCTION D'ÉLÉMENTS DE COMPÉTITIVITÉ DANS LE MARCHÉ DES TRANSPORTS COLLECTIFS URBAINS

La récente contractualisation a permis de clarifier les rôles et de définir les responsabilités. L'Etat est actionnaire de la RATP et fixe les objectifs (marge d'autofinancement, endettement...) comme n'importe quel actionnaire. Le STIF est l'autorité organisatrice, il définit la politique de transport, veille à sa mise en oeuvre et décide du prix de vente public du transport. La RATP est responsable de la production industrielle du service et de la performance. Ainsi, nous pouvons considérer que l'offre de la RATP est fixée de manière exogène, de même que la tarification. Ces éléments amènent à s'intéresser à la demande de transport collectif urbain.

Chapitre 2

La demande de transport collectif urbain

La principale caractéristique du transport réside dans le fait qu'il n'est pas demandé pour lui-même, mais parce qu'un "bénéfice" est attendu à la destination finale. Le voyage en lui-même représente un coût et se doit d'être le plus court possible car il n'apporte pas de satisfaction. Ainsi, la demande de transport collectif urbain est une "demande dérivée" et dépend des facteurs qui ont motivé le déplacement.

Nous examinerons tout d'abord les caractéristiques de cette demande dérivée (première section), en détaillant les différents motifs qui génèrent un déplacement, et leurs poids relatifs. La demande sur le marché des transports collectifs urbains résume l'équilibre de marché (deuxième section) : il s'agit d'une demande effective. Cette demande a été modélisée économétriquement à l'étranger, en France, et au sein même de la RATP (troisième section), mais les modélisations existantes présentent deux inconvénients : elles ont généralement pour but une prévision uniquement à long terme ou reposent sur des hypothèses qui ne sont pas valides pour nos données.

2.1 La demande de transport collectif urbain, une demande dérivée

Pour satisfaire un certain nombre de besoins (activités professionnelles ou scolaires, consommation de biens et services...), les personnes sont amenées à se déplacer. Le transport, collectif ou individuel, est une "demande dérivée" des facteurs qui ont motivé le déplacement. Il correspond à une "consommation intermédiaire", i.e. se définit comme un bien ou service consommé dans le cadre d'activités auxiliaires.

Nous examinerons tout d'abord les motifs de déplacement pris en compte dans les principales enquêtes de transport (première sous-section). Ces motifs peuvent être regroupés en deux classes, suivant l'existence de contraintes (heure d'arrivée et/ou destination imposée) : les motifs contraints, et les motifs non contraints. Cette distinction a des conséquences sur la fréquentation des transports collectifs urbains (deuxième sous-section). Cette analyse a pour but de mettre en évidence les principaux déterminants de la demande de transport (troisième sous-section).

2.1.1 Les motifs de déplacement

Deux types d'enquête permettent d'appréhender les motifs de déplacement : les enquêtes "TJRF" (Trafic Journalier sur le Réseau Ferré) et "EGT" (Enquête Globale de Transport).

2.1.1.1 Les enquêtes "TJRF" (Trafic Journalier sur le Réseau Ferré)

Les enquêtes TJRF sont des enquêtes "minute" réalisées pour la RATP dans chaque station de métro et gare de RER. Elles fournissent des renseignements sur les premières directions empruntées, les modes de rabattement, les motifs ou encore les communes à l'origine et à la destination, le titre utilisé, ainsi que des indications relatives à l'âge et au sexe de la personne interrogée.

Au cours de l'année 2001 et lors du premier trimestre 2002, l'enquête de 1997 a été actualisée par des enquêtes ponctuelles réalisées dans les stations ou gares ouvertes depuis 1997¹, ainsi que

¹Ligne Météor, prolongement des lignes 13 et A.

dans les stations à partir desquelles les déplacements ont été modifiés par l'ouverture puis le prolongement de Météor² ou la mise en service d'Eole³.

2.1.1.2 Les enquêtes "EGT" (Enquête Globale de Transport)

Les EGT sur les déplacements des Franciliens sont réalisées pour le compte de la DREIF (Direction Régionale de l'Équipement de la région Ile-de-France), du Ministère des transports, de la Région Ile-de-France, de la Ville de Paris, de la RATP, de la SNCF, de la Direction Régionale de l'INSEE, du STIF et de Cofiroute. Ces enquêtes lourdes sont effectuées depuis 1969⁴ et permettent le recueil d'informations sur quatre types de données :

- les caractéristiques des ménages (adresse, motorisation, revenu...) et des personnes composant le ménage (activité, âge, lieu de travail...),
- les caractéristiques des déplacements⁵ pour un jour de la semaine : origine, destination, mode de transport, durée... Ce questionnaire "déplacement jour de semaine" est destiné uniquement aux personnes de plus de 6 ans et recense l'ensemble des déplacements ayant au moins une extrémité en Ile-de-France, et effectués au cours de la journée précédente (jour ouvrable compris entre le lundi et le vendredi).
- Le troisième volet correspond à une enquête sur les "déplacements fin de semaine" et concerne les déplacements du samedi et du dimanche des personnes de 6 ans ou plus (pour seulement un ménage sur trois).
- Le dernier volet est un questionnaire "situation de handicap" pour toute personne qui s'estime gênée dans ses déplacements.

Le dernier sondage totalement exploité a été réalisé sur un échantillon approprié de ménages tirés au sort de novembre 1997 à mai 1998, en excluant les jours de vacances scolaires. Les résultats de l'enquête 2001, réalisée entre octobre 2001 et avril 2002, commencent à être publiés.

²Ligne 1, ligne A et ligne 6 autour de Bercy.

³Gare de l'Est, Saint-Lazare, pôle "Opéra", à l'exception de la Gare du Nord en travaux.

⁴Des enquêtes ont été réalisées en 1976, 1983, 1991, 1997 et 2001.

⁵Un "déplacement" est défini comme le "mouvement motivé d'une personne entre une origine et une destination, mouvement réalisé selon un itinéraire, et pendant une certaine durée".

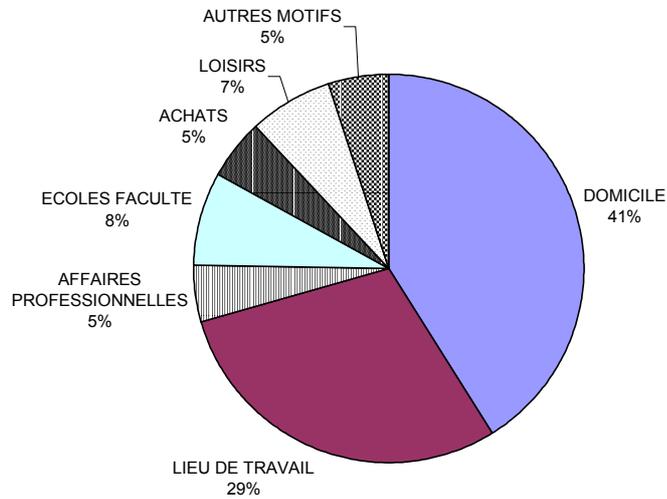
2.1.1.3 Motifs de déplacement

L'enquête TJRF distingue 7 facteurs de déplacements contre 19 pour l'enquête EGT, plus détaillée. En synthétisant, nous obtenons les motivations suivantes à un déplacement :

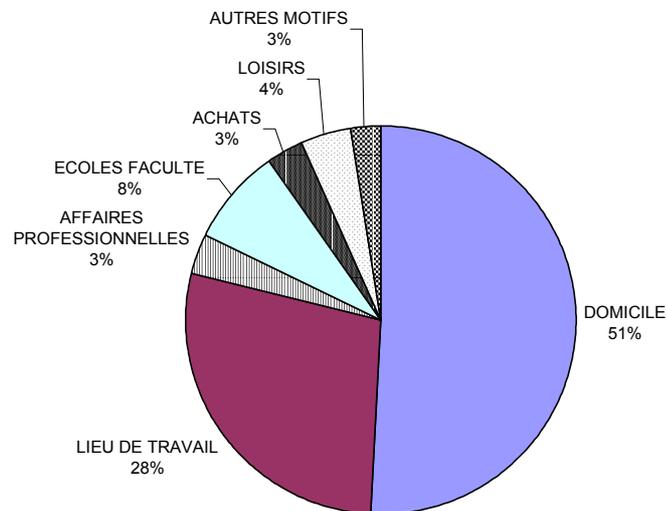
1. Domicile
2. Travail fixe ou habituel
3. Affaires professionnelles : travail non fixe et inhabituel
4. Ecoles et facultés : maternelle ou école primaire, enseignement secondaire ou technique, enseignement supérieur
5. Achats : achats quotidiens (pain, journal), achats hebdomadaires ou bi-hebdomadaires (achats en grandes surfaces), ou achats exceptionnels
6. Loisirs : participation à une activité de loisirs (sports,...), promenade, visite familiale ou amicale, spectacles (exposition, cinéma...) ou autres loisirs (restaurant, café...)
7. Autres : santé, affaires personnelles autres que loisirs (administration...), accompagnement ou fait d'aller chercher quelqu'un.

Les graphiques suivants indiquent les parts respectives de ces différents motifs à l'origine d'un déplacement, pour les stations de métro et les gares de RER RATP.

Motifs à l'origine des entrants dans le métro



Motifs à l'origine des entrants dans le RER



Source : enquête TJRF 1997 actualisée en 2002

2.1.2 Motifs contraints et non contraints : conséquences sur la fréquentation des transports collectifs urbains

Les motifs énumérés précédemment (hors domicile) peuvent être classés en deux catégories :

- Les motifs "obligés" ou "contraints" : travail, affaires professionnelles, écoles et facultés.

Ce type de motif est un motif "contraint" dans le sens où la destination et l'horaire d'arrivée sont imposés.

- Les motifs "non obligés" ou "non contraints" : achats et loisirs mais aussi "autres motifs" (excepté l'accompagnement de personnes).

Suivant le type de catégorie, la demande de transport qui en dérive possède des caractéristiques différentes en termes de fréquentation et de typologie de clientèle. Examinons ces caractéristiques pour la demande de transport collectif.

2.1.2.1 Caractère obligatoire des motifs et conséquences sur la fréquentation des transports

Deux types d'impact sur la fréquentation peuvent être mis en avant :

- l'existence de fluctuations régulières liées aux motifs obligés (travail, formation, affaires professionnelles),
- l'accroissement de fréquentation en lien avec le développement des déplacements non obligés (loisirs, achats, autres motifs).

Des fluctuations régulières liées aux motifs obligés Chaque jour ouvrable, les liaisons pendulaires "domicile - travail" ou "domicile - écoles/facultés" conduisent à l'existence d'une pointe du matin et d'une pointe du soir. Ainsi, les fluctuations dans la demande de transport collectif urbain reflètent les impératifs de la demande pour le bien final accessible par les transports publics.

Néanmoins, le rôle central du travail dans la synchronisation des temps sociaux a tendance à se modifier de façon lente depuis plus de 15 ans. La répartition par demi-heure du trafic des

entrants-correspondants sur le métro ou le RER, établie d'après les relevés des péages un jour ouvrable d'hiver de 1990 et de 2002, fait apparaître les éléments d'évolution suivants :

- un décalage de la pointe du matin avec augmentation de l'hyperpointe,
- un écimage de la pointe du soir,
- une croissance significative du trafic entre le flanc de pointe du soir et la fin de service.

Le tableau suivant ainsi que les graphiques précisent ces évolutions :

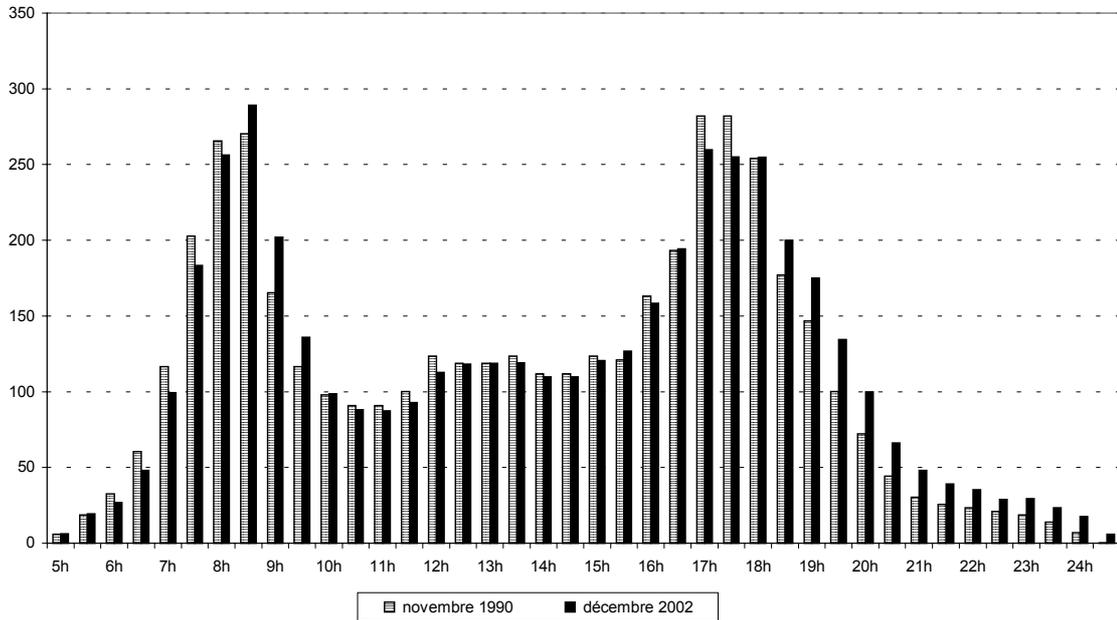
Part des entrants par tranche horaire

	METRO		RER	
	Nov. 1990	Déc. 2002	Nov. 1990	Déc. 2002
Avant 8h30	15.82%	13.93%	19.22%	17.34%
8h30 - 9h30	9.81%	10.68%	7.55%	9.44%
9h30 - 17h00	40.65%	38.98%	32.97%	34.57%
17h00 - 18h00	12.69%	11.20%	16.20%	12.76%
18h00 - 20h00	15.26%	16.63%	18.54%	18.14%
Après 20h00	5.77%	8.57%	5.52%	7.75%

Source : RATP, données corrigées des incidents techniques aux péages

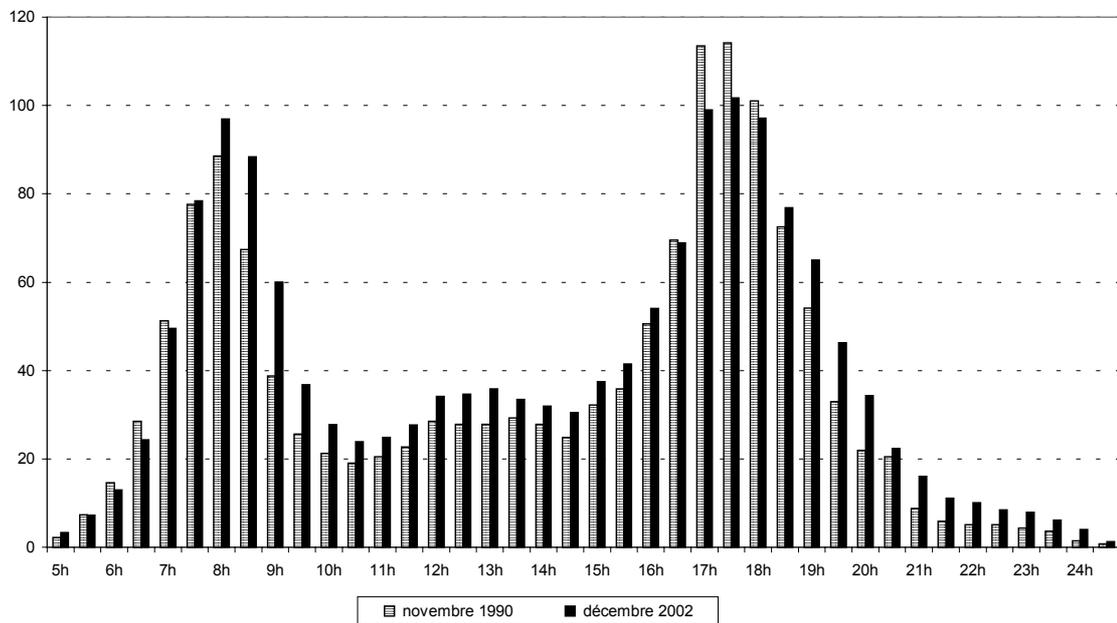
2.1. LA DEMANDE DE TRANSPORT COLLECTIF URBAIN, UNE DEMANDE DÉRIVÉE

Entrants directs et correspondants sur le métro (en milliers)



Source : RATP, données corrigées des incidents techniques des relevés de péage

Entrants directs et correspondants sur le RER (en milliers)



Source : RATP, données corrigées des incidents techniques des relevés de péage

Ce glissement des "migrations alternantes" résulte vraisemblablement de la diminution de l'activité industrielle au profit du secteur des services, dans lequel les horaires de travail sont plus tardifs voire plus flexibles.

Entre 1991 et 2001, le volume de déplacements "domicile-travail" a moins augmenté que la moyenne (+2.8% contre +6%), et les déplacements pour affaires professionnelles ont diminué de -3.5%. En 2001, travail et affaires professionnelles représentent 33% des déplacements, contre 35% en 1991. Les transports en commun assurent 40% de ces déplacements pendulaires.

De façon équivalente, le volume de déplacement pour études est demeuré stable entre 1991 et 2001, mais la part de ces déplacements a diminué : ce motif contribuait à 16.1% des déplacements en 1991 contre 15.2% en 2001.

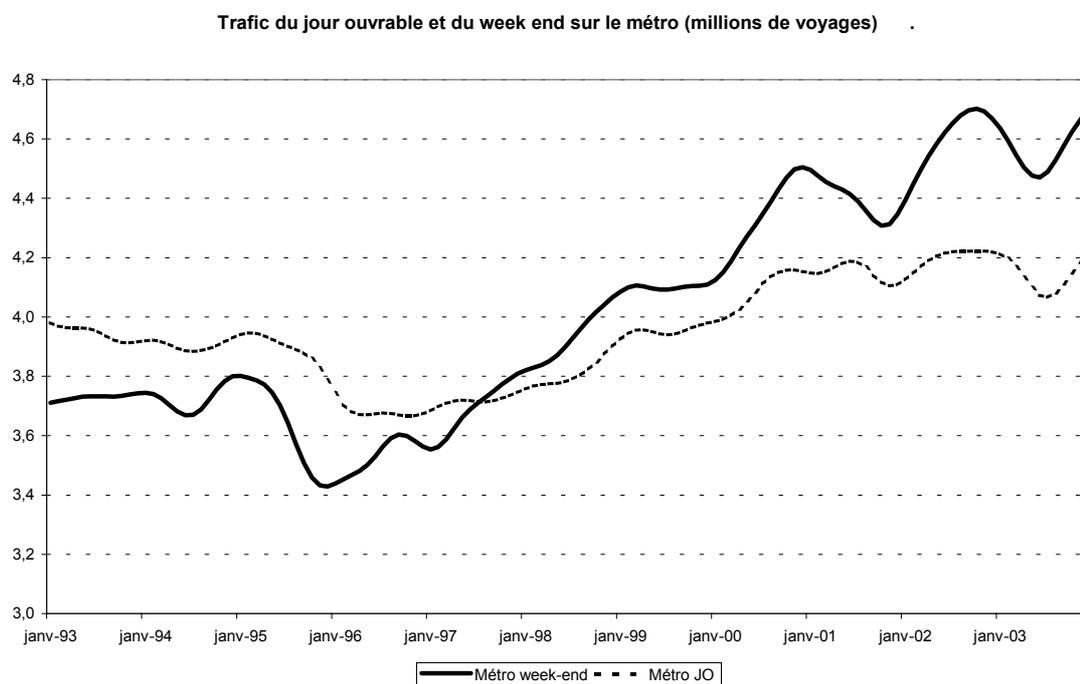
L'ensemble des motifs obligés expliquait 1.77 déplacements par jour et par personne en 1991, contre 1.67 en 2001.

L'accroissement de fréquentation lié à l'essor des motifs non-contraints [25] En 2001, les Franciliens se déplacent davantage pour des motifs autres que le travail ou l'école qu'en 1991. Cette évolution reflète les modifications démographiques, sociales et professionnelles, telles que la stagnation voire la baisse de l'emploi à taux plein (développement du travail à temps partiel), et l'augmentation du nombre de retraités (diminution du taux d'activité), mais aussi la forte augmentation de la mobilité liée aux loisirs depuis 1991. Les déplacements pour affaires personnelles (loisirs, achats, santé, visites familiales ou amicales) ont progressé de +12% entre 1991 et 2001, contre +6% pour l'ensemble des déplacements. Ce type de déplacements, pour motifs privés, représente plus de la moitié des mouvements effectués par les Franciliens et 47% des déplacements motorisés (i.e. excluant la marche à pied).

Les principaux faits marquants relatifs à la diversification de la demande de transport résident dans la hausse du trafic en soirée, mais aussi dans l'accroissement du trafic du week-end.

2.1. LA DEMANDE DE TRANSPORT COLLECTIF URBAIN, UNE DEMANDE DÉRIVÉE

- Sur le métro, le trafic du week-end est supérieur à celui d'un jour ouvrable depuis mi-1997.



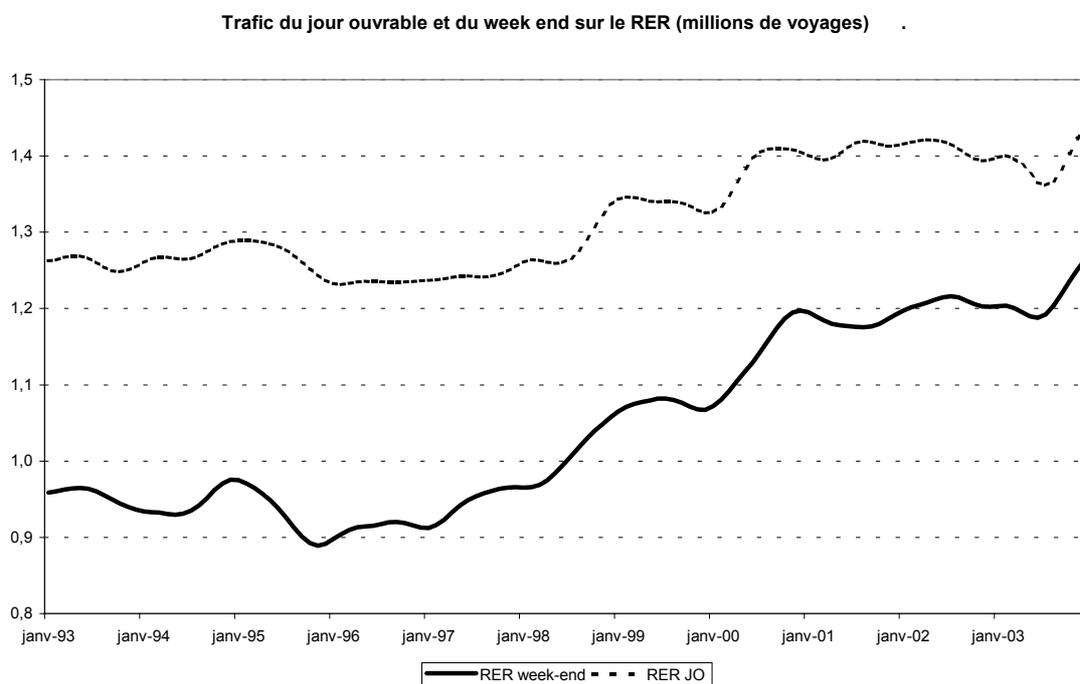
Source : RATP (données corrigées des perturbations, des effets calendaires et saisonniers)⁶

Trafic sur le métro par type de journée (millions de voyages)				
	1993	2003	Variation 2003/1993	Taux annuel moyen
Jour ouvrable	3.941	4.150	+5.28%	+0.47%
Samedi	2.321	2.748	+18.39%	+1.55%
Dimanche/jour férié	1.430	1.834	+28.27%	+2.29%

Source : RATP (données corrigées des incidents techniques aux péages)

⁶Les différentes corrections apportées au trafic seront détaillées dans le chapitre suivant.

- Sur le RER, l'écart entre le trafic du jour ouvrable et le trafic du week-end diminue.



Source : RATP (données corrigées des perturbations, des effets calendaires et saisonniers)

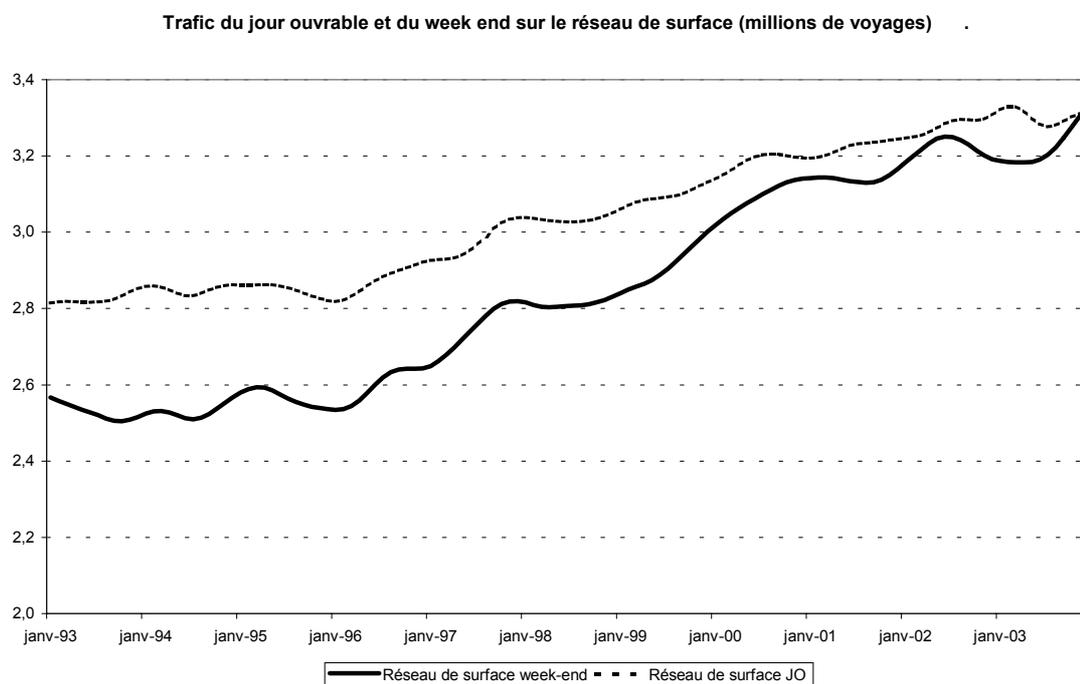
Trafic sur le RER par type de journée (millions de voyages)				
	1993	2003	Variation 2003/1993	Taux annuel moyen
Jour ouvrable	1.256	1.396	+11.17%	+0.97%
Samedi	0.602	0.742	+23.22%	+1.92%
Dimanche /jour férié	0.354	0.471	+32.93%	+2.62%

Source : RATP (données corrigées des incidents techniques aux péages)

- Pour le réseau de surface (bus de Paris et bus de banlieue), l'écart entre le trafic du jour ouvrable et le trafic du week-end s'est résorbé au fil des mois, malgré une inversion

2.1. LA DEMANDE DE TRANSPORT COLLECTIF URBAIN, UNE DEMANDE DÉRIVÉE

passagère en 2003.



Source : RATP (données corrigées des perturbations, des effets calendaires et saisonniers)

Trafic sur le bus par type de journée (millions de voyages)				
	1993	2003	Variation 2003/1993	Taux annuel moyen
Bus Paris				
Jour ouvrable	1.159	1.200	+3.49%	+0.31%
Samedi	0.729	0.798	+9.48%	+0.83%
Dimanche/jour férié	0.286	0.318	+11.07%	+0.96%
Bus banlieue				
Jour ouvrable	1.664	2.106	+26.55%	+2.16%
Samedi	0.979	1.355	+38.44%	+3.00%
Dimanche /jour férié	0.527	0.766	+45.41%	+3.46%

2.1.2.2 Motifs de déplacement et typologie de clientèle

Les titres de transport ciblent des motifs de déplacement et des catégories de clientèles bien spécifiques. De façon générale on distingue :

- une clientèle régulière, utilisatrice d’abonnements, dont les déplacements relèvent principalement de motifs contraints,
- une clientèle occasionnelle, qui achète des billets ou des forfaits commerciaux pour des déplacements ayant souvent des motifs non contraints.

Présentons les principaux titres de transport destinés aux utilisateurs réguliers ou occasionnels des transports collectifs urbains.

Les abonnements

La Carte Hebdomadaire de Travail 12 voyages permet d’effectuer, sur certaines relations, 12 voyages à raison de deux voyages par jour pendant une période maximale de 7 jours. Elle peut être utilisée à partir de n’importe quel jour de la semaine et donne droit à des parcours déterminés :

- soit entre une gare RER ou SNCF Ile-de-France et Paris (métro, RER et prolongements de métro en Ile-de-France)
- soit entre deux gares RER ou SNCF Ile-de-France.

Depuis 1990, les ventes de ce titre ont diminué fortement jusqu’à un pallier d’environ 500 000 cartes par an, qui correspond à une clientèle fidèle qui voyage peu (un aller - retour sur le même trajet) et par conséquent n’a pas intérêt à acheter des Cartes Orange.

La Carte Orange correspond à un coupon unique valable un mois ou une semaine accompagné d’une carte nominative avec photo. Elle s’adresse aux Franciliens qui utilisent régulièrement les transports urbains sur les réseaux de la RATP (sauf Orlyval), la SNCF Ile-de-France et le réseau des autobus OPTILE. Cet abonnement permet au client d’effectuer autant de voyages qu’il le désire (du 1er au dernier jour du mois pour le coupon mensuel, du lundi au dimanche pour le coupon hebdomadaire). Pour les salariés, le coupon est remboursé à 50 % par l’employeur.

2.1. LA DEMANDE DE TRANSPORT COLLECTIF URBAIN, UNE DEMANDE DÉRIVÉE

Entre le 1er décembre 1990 et la fin de l'année 2002, le prix de la carte orange mensuelle pour les zones 1 – 2 s'est accru de +75%, soit un taux de variation annuel moyen de l'ordre de +4,8%. Cette hausse des tarifs est néanmoins moins importante que pour les cartes orange hebdomadaires, dont le prix pour les zones 1 – 2 a augmenté de +84%, soit +5.2% en moyenne annuelle.

La Carte Intégrale correspond à un coupon unique valable 12 mois, accompagné d'une carte nominative avec photo. Ce titre s'adresse principalement aux Franciliens qui utilisent les transports en commun de façon régulière, au moins 10 mois sur 12, et est valable sur l'ensemble du réseau RATP (à l'exception d'Orlyval), la SNCF Ile-de-France, et le réseau OPTILE.

Ce titre possède certains avantages par rapport à la carte orange mensuelle : le client ne réalise qu'un seul achat par an, et évite les files d'attente chaque mois ; en cas de perte ou de vol de la carte ou du coupon, le remplacement est immédiat et n'engendre pas de frais pour lui. Le prix de l'abonnement est égal à celui de 10.2 coupons de Carte Orange pour les zones 1 – 2, et est remboursé à 50% par l'employeur. Enfin, ce titre bénéficie d'une souplesse importante dans la mesure où il est possible de changer de zones de validité en cours d'année, d'interrompre l'abonnement pour au moins deux mois à tout moment, ou de résilier sans frais l'abonnement en cours (les mois non utilisés sont remboursés).

La Carte Intégrale donne enfin droit à certains avantages commerciaux : la réception bimestrielle du magazine Intégralement vôtre qui offre des informations sur les transports et les loisirs en Ile-de-France, des réductions à la GMF, AVIS , Radio France et l'Institut Océanographique.

La carte Imagine R, scolaire ou étudiant, a été mise en place en septembre 1998 par le Syndicat des Transports d'Ile-de-France, la région Ile-de-France, le ministère des Transports et le ministère de l'Education nationale avec l'aide des départements. Le coupon unique est valable une année scolaire (de septembre à septembre pour les collégiens et lycéens, pendant 12 mois avec un début de validité compris entre septembre et janvier pour les lycéens de moins de 26 ans), et accompagné d'une carte nominative avec photo. Il permet d'effectuer la semaine un nombre illimité de voyages sur l'ensemble des transports en commun en Ile-de-France (sauf Orlyval) en fonction des zones choisies, pour un prix presque deux fois inférieur à celui de la

Carte Orange (sans compter certaines possibilités de subventions par les Conseils Généraux des huit départements ou de tarifs spécifiques pour les boursiers).

Le week-end, les jours fériés, lors des petites vacances scolaires (depuis juillet 2000) et des grandes vacances (depuis juillet 2004), le "dézonage" permet d'aller partout en Ile-de-France (sauf Orlyval et certaines lignes à tarification spéciale d'OPTILE), sans limitation de zone. En cas de déménagement ou de changement d'établissement scolaire, il est possible de changer de zones de validité sans frais supplémentaires.

La carte Imagine "R" offre aussi des réductions et des avantages exclusifs pour les concerts, le cinéma, des soirées. Le magazine des abonnés Imagine "R", "a/r" donne des informations sur l'actualité, la mode, les "Bons Plans" et les villes qui "bougent". La gestion de cette carte, sa commercialisation et la recherche des bons plans sont assurées par OPTILE, la RATP et la SNCF.

Les billets Les billets permettent de se déplacer dans le métro, le bus, le tramway et le RER. Ils sont vendus soit à l'unité en plein tarif, soit en carnet de 10 tickets plein tarif ou tarif réduit (sur présentation d'une carte famille nombreuse par exemple). Ils sont principalement destinés aux clients "occasionnels" : les Franciliens qui se déplacent peu, ainsi qu'aux provinciaux et aux étrangers de passage à Paris qui ne comptent pas se déplacer beaucoup au cours de la journée. On distingue :

1. les tickets T⁷ qui sont vendus en carnet ou à l'unité, à tarif normal ou à tarif réduit⁸ ;

⁷Le 1er janvier 2003, les titres de transports "section urbaine" classiques (ticket de métro - autobus, billets des transporteurs privés affiliés à OPTILE) ont disparus au profit du "ticket T". Ce nouveau titre est valable sur l'ensemble des réseaux d'Ile-de-France : sur les bus OPTILE et RATP, le métro, le tramway et le RER dans Paris (RATP et SNCF). Cette modification facilite l'utilisation des transports collectifs (surtout le bus) des voyageurs occasionnels, qui auparavant devaient utiliser plusieurs types de ticket pour se déplacer en Ile-de-France.

⁸Les personnes ayant droit au tarif réduit à 50 % sont :

- les enfants, de 4 ans à moins de 10 ans ;
- les titulaires d'une carte Améthyste "demi-tarif" ;
- les titulaires d'une carte "Famille nombreuse" de couleur bleue délivrée par la SNCF ;
- les groupes d'au moins 10 jeunes de moins de 16 ans et leur accompagnateur, ou les élèves des collèges et lycées, classe de terminale incluse, qui se déplacent en groupes de 10 et qui sont accompagnés d'un enseignant (cette réduction n'est pas valable dans les bus et trams).

2.1. LA DEMANDE DE TRANSPORT COLLECTIF URBAIN, UNE DEMANDE DÉRIVÉE

2. les billets section urbaine + banlieue qui sont vendus à tarif normal ou à tarif réduit en carnets, uniquement à tarif normal à l'unité.

Sur les lignes RER et SNCF Ile-de-France au-delà des limites de la ville de Paris, la tarification varie en fonction de la distance. Il est possible d'acheter :

- soit des billets valables entre deux gares déterminées du RER ou du réseau SNCF Ile-de-France (hors de Paris) ;
- soit des billets permettant des trajets entre n'importe quelle gare de RER ou de la SNCF dans Paris ou n'importe quelle station du métro d'une part et des gares déterminées du RER ou du réseau SNCF Ile-de-France d'autre part ;
- soit des billets pour des trajets entre deux gares du RER ou du réseau SNCF Ile-de-France (hors de Paris), qui comportent un parcours en métro ou RER dans Paris.

Il existe par ailleurs des billets "desserte aéroports", qui permettent de voyager entre Paris et ses aéroports. Les dessertes des aéroports exploitées par la RATP sont au nombre de 3 :

1. "Orlyval" relie Paris à Orlyval via Antony, le billet associé permet de voyager sur la totalité du réseau ferré. L'utilisation d'Orlyval est comprise dans le titre Paris Visite zones 1 – 5 ; les personnes titulaires d'une Carte Orange ne paient que la partie Antony - Orly.
2. "Orlybus" permet de se rendre directement de Paris à Orly par le bus, au départ de Denfert-Rochereau et d'Orly. Le billet "Orlybus" s'adresse aux personnes qui ne disposent pas d'un Paris Visite zones 1 – 5, d'une Carte Orange ou Carte Intégrale ou Imagine "R" couvrant les zones 1 à 4.
3. "Roissybus" relie en bus Paris (Opéra) et Roissy - Charles de Gaulle. Il est accessible aux porteurs de Paris Visite, Carte Orange, Carte Intégrale, Imagine "R" couvrant les zones 1 – 5.

En outre, la SNCF dessert les aéroports parisiens par le biais de :

- "Orlyrail" (ligne C du RER), accessible aux porteurs de Paris Visite zones 1 à 5, de Carte Orange, Carte Intégrale ou Imagine "R" zones 1 à 4. Les billets sont en vente à la SNCF.

- "Roissyrail" (ligne B du RER), accessible aux porteurs de Paris Visite zones 1 à 5, de Carte Orange, Carte Intégrale ou Imagine "R", mais couvrant cette fois ci les zones 1 à 5. Les billets sont vendus dans les gares SNCF mais aussi dans les gares de RER et les stations de métro.

Enfin, il existe un billet "Noctambus". Sur les lignes Noctambus, un tarif spécial unique est appliqué, quel que soit le parcours emprunté, avec ou sans correspondance entre les lignes Noctambus. Ce billet est vendu exclusivement dans les bus.

Le "funiculaire de Montmartre" est quant à lui accessible avec un billet simple "Métro, RER, Bus, Tram", qui permet d'effectuer un trajet (montée ou descente) sans correspondance avec le métro ou le bus.

Entre 1990 et 2002, le prix du carnet de billets plein tarif a plus que doublé, du fait d'un accroissement annuel moyen égal à +6%. La plus forte hausse a eu lieu en 1998 (+8.3%).

Les forfaits commerciaux

Le Billet l'OpenTour est un produit valable 1 ou 2 jours, destiné aux personnes désirant découvrir Paris et ses principaux monuments du haut d'un bus aménagé, avec des commentaires en plusieurs langues. Il permet d'utiliser les bus de cette ligne à volonté, de 10 heures à 18 heures. Son tarif est réduit sur présentation de la Carte Orange, de la Carte Intégrale, d'un Paris Visite ou d'un Pass Batobus⁹.

Formule 1 et Mobilis Le titre "Formule 1" se composait d'une carte nominative et d'un coupon valable dans les zones 1 – 2, 1 – 3, 1 – 4 (avec accès à l'aérogare d'Orly sans pouvoir utiliser Orlyval) ou 1 – 5 qui permet la desserte des aéroports d'Orly et de Roissy - Charles de Gaulle sur les dessertes exploitées par la RATP (lignes régulières à tarification générale ou Orlybus et Roissybus à tarification spéciale, ainsi que Orlyval) ou par la SNCF (notamment

⁹Les Batobus constituent un service régulier de navettes fluviales, mis en place en direction des îles de la Cité et Saint-Louis afin d'inciter les touristes à découvrir la capitale par ce moyen plutôt que par des autocars. Mis en service le 5 avril 2003, ce projet a été réalisé par la Préfecture de Police, la Ville de Paris et le Port Autonome de Paris. 8 escales sont desservies (Tour Eiffel, Champs-Élysées, Musée d'Orsay, Louvre, Saint-Germain-des-Prés, Hôtel de Ville, Notre Dame et Jardin des Plantes).

2.1. LA DEMANDE DE TRANSPORT COLLECTIF URBAIN, UNE DEMANDE DÉRIVÉE

Orlyrail et Roissyrail incluant les navettes routières aéroportuaires). Ce titre permettait d'aller à Marne - la Vallée - Chessy (parc "Eurodisneyland").

Ce titre a été totalement remplacé en 1998 par le titre de transport appelé "Mobilis" (créé en avril 1997) destiné aux Franciliens qui utilisent occasionnellement plusieurs modes de transport sur une même journée. Le titre "Mobilis" est un coupon valable une journée entière, qui permet d'effectuer un nombre illimité de déplacements sur tous les modes de transport collectif, à l'exception des dessertes des aéroports telles que Orlyval, Orlybus, Roissybus, JetBus, Roissy Rail et Orly Rail. Ce titre est valable sur toutes les zones choisies, depuis les zones 1 – 2 jusqu'aux zones 1 – 8. Il est destiné aux voyageurs qui doivent effectuer plusieurs déplacements dans la même journée sur les réseaux RATP, SNCF Ile-de-France et OPTILE. Son tarif est particulièrement intéressant en zone 1 – 4 et 1 – 5, il est inférieur à celui d'un aller - retour Banlieue - Paris au delà des zones 1 – 6.

Le "Paris Visite" a été créé par décision du Syndicat des Transports Parisiens le 28 mai 1989. Il est utilisable sur tous les réseaux de la RATP et de la SNCF et donnait accès à la 1^{ère} classe jusqu'à sa suppression. Ce titre comporte un coupon valable soit dans les zones 1 – 3 de la carte orange, soit dans les zones 1 – 5 ou 1 – 8 avec, dans ces deux derniers cas, accès aux aéroports d'Orly et de Roissy - Charles de Gaulle sur les dessertes exploitées par la RATP (lignes régulières à tarification générale ou Orlybus, Roissybus et Orlyval, à tarification spéciale) ou par la SNCF (notamment Orly Rail et Roissy Rail incluant les navettes routières aéroportuaires). Il est possible de choisir un coupon valable 1, 2, 3 ou 5 jours.

2.1.2.3 Motifs de déplacement et typologie des stations et gares

La distinction entre les déplacements ayant un "motif obligé" et les déplacements "non contraints" est corroborée par la construction d'une classification des stations de métro et gares de RER.

Pour présenter le bilan annuel du trafic observé, l'unité Tarification Vente Résultats du département Commercial de la RATP a décidé de mettre en place une nouvelle classification des stations de métro et des gares de RER. Une analyse factorielle des composantes ainsi qu'une

classification hiérarchique ascendante ont été menées pour l'ensemble des stations (ou gares) en prenant en compte à la fois le motif à l'origine¹⁰ et le mode de rabattement¹¹ à la station (respectivement gare) issus de l'enquête TJRF actualisée.

Les résultats sur le métro indiquent une classification en 6 classes :

- Les stations de rabattement du domicile par la SNCF (grandes lignes ou trains de banlieue) sont une porte d'entrée au coeur de Paris pour des Franciliens aux habitations éloignées. Cette classe prend en compte l'ensemble des gares SNCF (Gare d'Austerlitz, Gare de l'Est, Gare de Lyon...) mais aussi les stations ayant une correspondance par la marche à pied telles que la station de métro "La Chapelle" située à proximité de la Gare du Nord.
- Les stations de rabattement du domicile par le tramway sont au nombre de 3 : Basilique Saint-Denis, Bobigny - Pablo Picasso et La Courneuve - 8 mai 1945.
- Les stations de rabattement du domicile par le bus correspondent à des zones d'habitation périphériques. Cette classe regroupe des stations telles que Porte d'Orléans, Malakoff - Rue Etienne Dolet, Boulogne - Pont de Saint Cloud,... situées aux abords de Paris, et telles qu'il existe un fort rabattement à ces stations par les autobus de banlieue.
- Les stations "domicile" sont des stations d'habitation dans Paris.
- Les stations à motifs non contraints (achats et loisirs) sont des stations où le motif domicile n'est pas le motif principal à l'origine. En moyenne, seuls 22.56% des entrants proviennent de leur domicile, contre 41.26% au niveau de l'ensemble des stations. Les motifs "achats" + "loisirs" représentent 26.09% des motifs à l'origine, soit plus de deux fois plus que la moyenne des stations de métro (10.58%). Les stations les plus représentatives de cette classe sont "Louvre - Rivoli", "Hôtel de Ville", "Châtelet", "Havre - Caumartin", "Pont Neuf"...
- Les stations à motifs contraints (travail, affaires professionnelles et écoles/facultés) sont caractérisées par l'importance de ces 3 motifs à l'origine (59.26% des entrants sur cette

¹⁰Ce sont les motifs évoqués précédemment : domicile, travail, affaires professionnelles, écoles/facultés, achats, loisirs et autres.

¹¹Les modes de rabattement pris en compte sont les suivants : marche à pied, autobus, tramway, SNCF, voiture ou deux roues.

2.1. LA DEMANDE DE TRANSPORT COLLECTIF URBAIN, UNE DEMANDE DÉRIVÉE

classe de stations de métro contre 42.56% au niveau de l'ensemble des stations de métro). En particulier, 47.60% des entrants ont pour motif à l'origine le travail et les affaires professionnelles, contre 34.22% sur toutes les stations. Cette classe regroupe des stations telles que "Quatre Septembre" (où 72.92% des entrants ont pour motif à l'origine le travail), "Esplanade de la Défense" (69.65%), "Kleber" (63.01%)... mais aussi des stations universitaires comme "Jussieu", "Censier - Daubenton", "Porte Dauphine", ..., où le motif à l'origine est l'école ou la faculté pour plus de la moitié des entrants.

Sur le RER, les critères de rabattement à la gare de RER n'apportent rien de plus à la classification que la seule prise en compte des motifs de déplacement à l'origine. 4 classes peuvent être distinguées :

- Les gares de rabattement à partir du domicile, situées en banlieue, telles que Le Vésinet Centre,
- Les gares de RER à proximité des universités, telles que Luxembourg ou Nanterre-Université,
- Les gares de RER dans les zones touristiques (motifs "achats" et "loisirs") telles que Châtelet-les-Halles ou Marne-la-Vallée,
- Et les gares de RER situées dans des quartiers de bureaux (La Défense, Gare de Lyon, Rueil Malmaison...).

Les analyses distinguent différents types de stations d'habitation, les stations à motifs contraints et les stations à motifs non contraints. Ces motifs de déplacements constituent des déterminants de la demande de transport.

2.1.3 Les déterminants de la demande de transport

Une analyse des transports doit prendre en compte chacun des motifs de déplacement à un niveau macro-économique. Les principaux déterminants de la demande de transport associés aux motifs sont à mettre en relation avec :

- les motifs obligés (situation du marché du travail, situation de l'immobilier de bureau, rythmes urbains, ...)
- les motifs non obligés (situation du tourisme, de la consommation et de l'inflation, ...)

En outre, les changements sociaux-économiques et culturels affectent la localisation, la taille et le mode de vie des familles. Par exemple, les changements familiaux tels que l'augmentation du nombre de divorces ou de célibataires ont entraîné une multiplication des domiciles, ce qui accroît la demande de transport.

2.1.3.1 La situation sur le marché du travail

La situation sur le marché du travail a un impact important sur le niveau de la demande de transport, par différents biais.

Le premier élément qui amène à prendre en compte le travail réside dans l'importance des liaisons "domicile - travail", déjà évoquées. Ils constituent une part importante, quoique décroissante, du trafic dans les transports collectifs urbains. Rappelons que ces liaisons pendulaires ont lieu à certains moments de la journée, et que le trafic de ces "heures de pointe" est utilisé pour dimensionner l'offre de trafic.

En outre, la situation sur le marché du travail a un impact sur le revenu, donc des conséquences sur les activités d'achats et de loisirs. Dans une certaine mesure, les déplacements pour des motifs contraints induisent les déplacements pour les motifs non contraints.

2.1.3.2 L'implantation géographique des lieux d'origine et de destination des déplacements

Sur l'ensemble de la région Ile-de-France, la voiture particulière représente le moyen de transport le plus utilisé. Elle assure 45% des déplacements quotidiens contre 19% pour les transports en commun. Une analyse plus détaillée montre que le choix modal dépend fortement de l'origine et de la destination du déplacement [25] :

- Lors des déplacements internes à la banlieue (petite couronne et grande couronne), où habitent 80% des Franciliens, 6 trajets sur 10 sont réalisés avec une voiture particulière. Entre 1991 et 2001, les liaisons internes à la banlieue ont continué à augmenter à un rythme plus soutenu que la croissance de la population résidentielle et que celle du nombre de ses déplacements. La capitale a perdu des habitants et des emplois au profit des couronnes.

2.1. LA DEMANDE DE TRANSPORT COLLECTIF URBAIN, UNE DEMANDE DÉRIVÉE

- Pour les déplacements à l’intérieur de Paris, le mode de déplacement le plus utilisé correspond à la marche à pied (plus de la moitié des déplacements), suivi des transports en commun (3 déplacements sur 10).
- Les transports en commun constituent le mode privilégié des déplacements entre Paris et la banlieue (6 déplacements sur 10), surtout à moyenne et longue distance. Face aux difficultés de circulation, les transports en commun lourds tels que le RER l’emportent facilement.

En conséquence, un phénomène de délocalisation des entreprises en dehors de Paris a un impact sur les trafics de la RATP. Le besoin d’agrandir les capacités de l’unité existante, la minimisation des coûts (foncier, construction, réglementaires, accès aux réseaux...), le temps d’accès au centre de leur aire de marché, la proximité des fournisseurs... sont les facteurs explicatifs des délocalisations des entreprises retenus dans de nombreuses études [110]. A ces facteurs peuvent être ajoutés, de façon non exhaustive, le niveau des coûts des services (fiscalité, exonérations, primes...) ou les ressources du territoire (qualité des infrastructures de recherche et d’éducation, intensité de la mise en réseau des acteurs aux travers ”d’incubateurs” ou de ”pépinières”...).

2.1.3.3 L’activité touristique et les loisirs

Paris est la première destination touristique au monde. 36 millions de personnes visitent la ville chaque année, et plus de la moitié des déplacements touristiques s’effectue sur le réseau de la RATP. De nombreux éléments favorisent l’attractivité de Paris, dans un contexte de diminution de la durée des séjours concomitant à l’essor du tourisme urbain.

La région parisienne est desservie par de nombreuses infrastructures de transport, et occupe une place centrale en Europe :

- Deux aéroports internationaux (Orly et Roissy - Charles de Gaulle) permettent d’accéder depuis Londres, Amsterdam, Bruxelles ou Bonn (une heure) et Copenhague, Dublin, Madrid, Berlin, Rome ou Vienne en 2 heures. Ces aéroports internationaux, et les 12 aéroports d’affaires franciliens voient passer 73 millions de passagers par an, et 183 villes européennes sont desservies au départ de Paris par 7400 vols hebdomadaires.

- Le système ferroviaire français a une structure assez semblable au réseau routier : toutes les lignes partent de Paris et sont connectées aux villes européennes d'importance. Ainsi, le réseau TGV (Trains à Grande Vitesse) relie l'Ile-de-France à 145 destinations françaises et 43 destinations européennes.
- Le développement des autoroutes dans les années 1980 - 1990 a consacré la prééminence de Paris pour la desserte du territoire national, et permet des accès rapides de villes comme Bruxelles (300 km), Londres (430 km), Amsterdam (480 km), Francfort (550 km), Milan (800 km)...

La bonne desserte de Paris au niveau national et international explique l'attraction touristique de Paris. Lorsque le touriste arrive à Paris par un mode de transport collectif, il peut faire partie d'un groupe organisé, auquel cas la probabilité est forte qu'il visite Paris en autocar de tourisme. S'il s'agit d'un touriste individuel, il peut décider de se déplacer en taxi, de louer une voiture ou d'emprunter les transports en commun. Par exemple, 18% des passagers empruntent les transports en commun pour se rendre de l'aéroport de Roissy à Paris (d'après la revue Transport Public de février 2002). Ces passagers ont une forte probabilité de continuer à emprunter les transports en commun lors de leur séjour en région parisienne.

2.1.3.4 L'influence des éléments sociaux-démographiques sur la mobilité

La mobilité Entre 1991 et 2001, la mobilité individuelle mesurée par le nombre de déplacements par personne et par jour a peu varié. Cette faible variation confirme la stabilité observée depuis 1976. Un Francilien âgé de plus de 6 ans effectue en moyenne 3.50 déplacements par jour ouvrable (tous moyens de transport et tous motifs confondus).

Le sexe et l'âge ont un impact sur la fréquence et le motif de déplacement, le mode de transport choisi ou encore la distance parcourue. Par exemple, la fréquence des déplacements diminue avec l'âge, et le motif de déplacement qui a la plus forte progression est le motif "achat".

L'impact des nouvelles techniques de communication et d'information (NTIC) Le développement des NTIC, comme Internet, le téléphone mobile ou plus anciennement le fax, auraient pu diminuer la demande de transport de voyageurs. Ces nouvelles méthodes de communi-

2.1. LA DEMANDE DE TRANSPORT COLLECTIF URBAIN, UNE DEMANDE DÉRIVÉE

cation permettent de réaliser certaines actions sans se déplacer de chez soi. C'est en particulier le cas du télétravail ou du téléachat, ou encore de la recherche d'information, administrative notamment. Il semble au contraire que les nouvelles technologies sont à mettre en rapport avec le développement de la mobilité : plus la communication est facile, plus le besoin de mettre un visage sur un nom ou une voix semble se faire ressentir, et plus la rencontre est facile à organiser.

La modification des rythmes urbains En France, la Réduction du Temps de Travail (RTT) s'est majoritairement traduite par l'attribution d'une journée ou demi-journée de manière régulière (37% des salariés), ou par des jours de congés supplémentaires (31 % des salariés, principalement cadres). Pour 24% des salariés, généralement employés et ouvriers non qualifiés, la RTT a pris la forme d'une réduction de la durée du travail quotidien.

Chez 60% des salariés, la demi-journée ou journée accordée correspond à un jour fixe de la semaine : le vendredi (33%), le mercredi (15%), ou le lundi (11%). La structure de la semaine semble modifiée : 2 jours de travail (mardi et jeudi), 3 jours plus partagés entre travail et temps libre (lundi, mercredi, vendredi), et 2 jours de temps libre (samedi et dimanche). La RTT a globalement peu modifié les horaires de travail, même atypiques, et l'impact de la réduction du temps de travail sur les trajets domicile - travail est quasi - inexistant.

Ce temps libéré est surtout utilisé pour " libérer le week-end " : il est consacré à 50% aux tâches quotidiennes (courses, cuisine, ménage), au repos (35%), à la famille et aux enfants (19%), à la pratique d'une activité culturelle ou sportive (18%). Le temps supplémentaire accordé à la vie associative (4%) ou au loisir (1%) est très faible. Une conséquence visible est un lissage des flux de consommation et de déplacement : en permettant aux salariés d'échapper aux heures et jours de pointe, les vendredi, mercredi et lundi deviennent des jours de pleine charge pour les commerces.

Lorsque la RTT se traduit par des jours de congés supplémentaires, les salariés disposent en moyenne de 16 jours par an. Néanmoins, la RTT ne semble pas avoir d'incidence importante sur les pratiques de tourisme. Selon les professionnels du tourisme, les 35 heures ont amplifié une tendance préexistante : le morcellement des vacances. A la SNCF, il existe une hausse de trafic le vendredi matin (voire le jeudi) et le lundi soir, et une baisse sensible de la durée moyenne du séjour, autour de 12 jours. En outre, si le tourisme reste marqué par le rythme des vacances

scolaires, le nombre de séjours l’hiver et au début du printemps augmente, et la fréquentation des infrastructures de loisir s’accroît hors vacances scolaires.

Ces différents éléments sont détaillés dans de nombreux ouvrages, dont les rapports de Viard [160] et de l’IRIS [31]. Ils peuvent avoir un impact sur les trafics observés à la RATP.

Dans cette première section, nous avons mis en avant les principaux déterminants de la demande de transport. Pour cela, nous avons rappelé que la demande de transport urbain est une ”demande dérivée”, i.e. que le transport n’est pas demandé pour lui même, mais pour la satisfaction d’un besoin ou d’un désir à sa destination. L’examen des motifs de déplacements utilisés dans les principales enquêtes franciliennes de transport (1^{ère} sous-section), ainsi que la distinction entre le caractère contraint/non contraint des déplacements (2^{ème} sous-section) nous ont permis de mieux appréhender les causes de déplacement en transport en commun ainsi que leurs conséquences sur les trafics à un niveau journalier ou hebdomadaire. Ils nous ont en outre permis de justifier l’abondance des titres disponibles pour emprunter les sous-réseaux RATP. Cette analyse nous a conduit à mettre en avant un certain nombre de déterminants de la demande de transport urbain (3^{ème} sous-section), déterminants à mettre en lien avec ces motifs de déplacements.

Ces différents déterminants ne sont pas exhaustifs : les caractéristiques de l’offre doivent être prises en compte. Ceci amène naturellement à se demander comment se réalise l’équilibre sur le marché des transports collectifs.

2.2 La demande de transport collectif urbain et l’équilibre sur ce marché

Dans le premier chapitre, nous avons mis en avant le fait suivant : sur le marché des transports collectifs, l’offre (volume de service et prix) est fixée par une autorité régulatrice appelée Syndicat des Transports d’Ile-de-France. La section précédente présente la demande comme une fonction de différents déterminants, à mettre en lien avec les motifs qui génèrent les déplacements. Il s’agit de mieux définir la nature de l’équilibre sur le marché des transports collectifs urbains.

2.2. LA DEMANDE DE TRANSPORT COLLECTIF URBAIN ET L'ÉQUILIBRE SUR CE MARCHÉ

Nous montrerons que la demande de transport résume l'équilibre sur le marché des transports collectifs. Une augmentation de l'offre de transport conduit à un accroissement de la demande de transport (1^{ère} sous-section), tandis qu'une diminution de l'offre de transport n'entraîne pas forcément une baisse de la fréquentation, mais de la congestion (2^{ème} sous-section). L'offre de transport collectif apparaît plus comme un ensemble de caractéristiques du bien de transport collectif urbain (3^{ème} sous-section). La qualité de service, la régularité, la fréquence, ..., correspondent à de nouveaux déterminants de la demande de transport en commun.

2.2.1 L'offre de transport crée la demande de transport

Dans le chapitre précédent, nous avons souligné les interactions entre la construction d'infrastructures de transport et l'aménagement du territoire. La modification des infrastructures de transport, par ses impacts sur la nature ou l'importance de la population résidentielle et de l'activité future, influence la demande future.

2.2.1.1 La péri-urbanisation et les interactions entre offre et demande de transport collectif urbain

Une première conséquence de cette interaction réside dans le phénomène de péri-urbanisation. L'augmentation de la vitesse des moyens de transports a eu pour conséquence une dispersion plus grande de l'habitat, dans un contexte de mobilité toujours plus grande et de recours accru à la voiture particulière. En effet, l'accroissement des vitesses de parcours a permis de parcourir des distances de plus en plus longues pour un temps de trajet relativement stable depuis de nombreuses années. L'amélioration passée de l'offre de transport (par exemple le développement du RER dans les années 70) a eu un impact sur la localisation actuelle de la demande de transport.

Cette modification de la demande de transport, autour de la métropole, est un enjeu pour la RATP, dans la mesure où il peut permettre à l'entreprise une certaine diversification de ses activités et de son offre de transport. Dans les zones peu denses, la RATP doit adapter son offre de transport collectif à la demande, en développant par exemple les bus à la demande.

2.2.1.2 L'existente d'une demande latente dans les zones peu desservies

La modification des infrastructures de transport collectif a aussi des implications immédiates en terme de demande : l'ouverture d'une nouvelle ligne de métro (ou un prolongement) attire des voyageurs qui auparavant utilisaient un mode de transport différent. En outre, la modification de la contexture du réseau peut encourager le déplacement de personnes qui ne voyageaient pas (diminution du temps de parcours, réduction du nombre de correspondances et simplification du trajet) et constituent une "demande latente". Cette demande latente est présente en banlieue, mais aussi parfois au centre de l'agglomération, où la desserte peut être difficile (rues étroites...).

Ce phénomène de "création" de la demande suite à une amélioration de l'offre est un phénomène économique connu : il se retrouve par exemple sur le marché du travail (l'embellie sur le marché de l'emploi incite des femmes au foyer à re-travailler) ou en médecine (l'installation d'un médecin à proximité incite à aller le consulter). Il constitue en outre la base du marketing : un nouvel objet doit créer/révéler un nouveau besoin.

2.2.2 La demande de transport est plus ou moins élastique au niveau de l'offre de transport collectif

Une partie de la clientèle des transports en commun peut être considérée comme une clientèle "captive", i.e. n'ayant pas le choix de prendre leur automobile pour effectuer leurs déplacements. La première cause de l'existence d'une telle demande captive réside dans la faible motorisation des Parisiens. Le tableau suivant indique la proportion de ménages "motorisés", c'est-à-dire disposant d'au moins une voiture de tourisme :

Pourcentage des ménages motorisés			
	1990	1997	2002
Paris	47%	47%	45%
Ile-de-France hors Paris	77%	76%	77%
France entière	76.8%	78.8%	80.2%

Sources : INSEE, SOFRES, EGT

2.2. LA DEMANDE DE TRANSPORT COLLECTIF URBAIN ET L'ÉQUILIBRE SUR CE MARCHÉ

Les ménages sont d'autant plus motorisés qu'ils résident dans des zones d'habitation peu denses. Plus de la moitié des ménages parisiens ne possède pas de voiture, et peut être considérée comme une clientèle captive des transports en commun (ou de la marche à pied) lors de leurs déplacements.

La faible motorisation de l'Ile-de-France et de Paris par rapport au niveau national, de même que sa stabilité dans un contexte de motorisation croissante des ménages, s'explique en partie par les difficultés de circulation voire de stationnement, surtout à Paris, qui conduisent à choisir de ne pas avoir de voiture. Pour les déplacements ayant une extrémité à Paris, nous avons déjà souligné l'importance de la part des transports collectifs, qui reflète un avantage en termes de temps de transport des transports collectifs, en particulier les transports en commun en site propre.

En outre, certaines populations ne sont pas motorisées, pas forcément par choix : les jeunes qui ne possèdent pas de permis de conduire, les personnes très âgées qui ne conduisent plus ou n'ont jamais appris à conduire, ainsi que toutes les personnes qui ont un revenu insuffisant pour faire l'acquisition d'un véhicule. De même, quand un ménage composé de plus de deux personnes ne possède qu'une voiture et que les deux personnes n'empruntent pas le même itinéraire pour aller travailler, une des deux personnes peut être obligée d'utiliser les transports en commun.

L'existence d'une telle demande captive correspond à une demande incompressible. Lorsque le service est perturbé, du fait de grèves ou d'incidents techniques imprévus, l'adaptation de la demande n'est pas immédiate, et ne concerne qu'une partie de la population, non captive. Cette captivité de la demande se traduit alors par des effets de cohue, des métros/RER/bus bondés et amène à penser que, sur le marché des transports collectifs urbains, la demande est en partie inélastique au niveau de l'offre.

Le trafic observé ne résulte pas d'une confrontation offre/demande, mais se trouve sur la courbe de demande : sauf perturbations extrêmement importantes, les personnes qui ont décidé d'emprunter les transports collectifs ont la possibilité de le faire, soit en se serrant, soit en acceptant d'attendre le métro, le RER ou le bus suivant. Le trafic observé correspond à la demande "effective". L'offre de service est une caractéristique du transport collectif.

2.2.3 L'offre intervient comme une caractéristique du bien de transport collectif

Selon Lancaster [103], tout bien est décrit comme une combinaison de caractéristiques : performances, qualité, esthétique, localisation, disponibilité... Les produits sont désirés non pas pour eux-mêmes, mais pour les caractéristiques dont ils sont pourvus. Chaque consommateur a en outre une hiérarchie personnelle de préférence de ces différentes variables.

Le transport collectif offert par la RATP possède de telles caractéristiques. Les principales caractéristiques du transport par la RATP sont évaluées périodiquement par des indicateurs de qualité et concernent :

- la régularité (tous les sous-réseaux)
- la propreté des stations, la disponibilité des équipements de vente et des escaliers mécaniques (réseau ferré)
- l'accueil au guichet (sur le réseau ferré) ou par le machiniste receveur (sur le réseau routier).

Sur le réseau des autobus de Paris et de banlieue, un indicateur particulier est évalué : l'information au point d'arrêt ainsi que la conformité des points d'arrêts. La qualité de l'information, pour les autobus comme pour le réseau ferré, est une caractéristique importante du bien de transport collectif.

2.2.3.1 L'information relative au déplacement

L'offre d'informations sur les déplacements est un élément primordial dans l'évaluation du bien qu'est le transport. Lors d'un déplacement, le voyageur demande des informations tout au long de son parcours [88]. Examinons la réponse de la RATP à chacune de ces demandes.

Lors de voyages occasionnels ou de trajets réguliers ayant une grande importance pour le voyageur (examen, importante réunion...), celui-ci va rechercher des informations précises sur son déplacement (itinéraire optimal et itinéraires alternatifs, prix, horaires, nombre de correspondances, perturbations...). Cette recherche peut s'effectuer par différents intermédiaires : plans des réseaux distribués dans les gares de RER et stations de métro, centres d'information téléphoniques (08 92 68 77 14, 0.34 euros la minute), ou internet (www.citefutee.com). De façon

2.2. LA DEMANDE DE TRANSPORT COLLECTIF URBAIN ET L'ÉQUILIBRE SUR CE MARCHÉ

générale, lorsqu'il existe des trajets alternatifs, le voyageur préfère des itinéraires plus longs mais directs aux itinéraires plus courts mais avec plusieurs correspondances. Il évite ainsi les temps d'attente et des incertitudes.

Tout au long du déplacement apparaissent des besoins d'information. Lorsqu'il commence son déplacement, le voyageur doit localiser le point d'entrée sur le réseau (gare de RER, station de métro ou arrêt de bus), puis s'orienter dans la bonne direction. Le besoin d'information concerne alors l'horaire du véhicule (arrivée, départ, retard éventuel) et le temps de voyage (perturbations sur les autres réseaux en correspondance). Une fois à l'intérieur du véhicule apparaissent des besoins d'informations géographiques (nom de l'arrêt suivant) et d'informations en temps réel (arrêt soudain du véhicule, modification de l'itinéraire, retards). A la RATP, ces différents éléments d'information sont disponibles grâce au SIEL (Système d'Information En Ligne), qui fournit les informations en temps direct sur les temps d'attente et les perturbations¹².



SIEL RER : prochains passages et perturbations ; SIEL métro : temps d'attente



SIEL BUS : temps d'attente, et annonce visuelle des prochains arrêts

¹²Les panneaux et bornes SIEL équipent les infrastructures suivantes : lignes 1, 4, 7, 13 et 14 du métro, lignes A et B du RER RATP, lignes T1 et T2 du tramway, ainsi que le Trans Val-de-Marne, et une cinquantaine de lignes de bus (soit la moitié du parc d'autobus de la RATP).

Le déploiement est prévu sur l'ensemble des lignes du métro avant la fin 2006. En complément, un système pour les non-voyants est à l'étude : dans un premier temps, un système sonore indiquera la direction sur les lignes à fourche.

Source : www.ratp.fr

Lors de perturbations sur le trajet en réseau ferré RATP, le voyageur doit être informé par haut-parleur.

Enfin, une fois le déplacement sur le réseau de transport en commun effectué, il existe une demande d'indications géographiques aux alentours de la station, de la gare ou de l'arrêt de bus, ainsi que des indications sur les horaires de retour. Un plan des alentours est affiché au sein de chaque station de métro et gare de RER RATP, de même que sur de nombreux arrêts de bus.

Lors de la préparation et de la réalisation de son déplacement, l'utilisateur des transports collectifs urbains acquiert des connaissances utiles pour ses déplacements ultérieurs. Il développe des préjugés (positifs ou négatifs) à l'usage du mode de transport utilisé, qui le conduiront à des processus de préférence ("le métro plutôt que le bus") ou d'évitement ("pas de trajet durant les heures de pointe").

Régularité, propreté, information sont des éléments qui caractérisent le bien de transport collectif urbain. Le prix est une autre de ces caractéristiques.

2.2.3.2 Le coût du transport collectif

Selon Janson (1998) [89], le marché du transport est un marché bien spécifique. Les utilisateurs demandent certes un service, mais ils offrent leur temps, en se rendant à la station, en attendant leur métro, leur RER ou leur bus, puis en voyageant sur les modes choisis (voire en prenant des correspondances). Le temps "perdu" dans les transports peut être monétisé, et inclus dans un "coût généralisé" du transport supérieur au simple prix du voyage. Ce coût généralisé intègre par conséquent la notion de qualité de service et les prix des transports.

Au niveau microéconomique, une première difficulté provient du fait que l'évaluation du temps est subjective, donc propre à chaque individu...

Elle dépend en outre des conditions dans lesquelles s'effectue le voyage, en particulier, la présence d'autres voyageurs dans les rames de métro ou de RER, ou dans les bus. Ainsi, la congestion, qui implique parfois de voyager debout, serré contre d'autres personnes, accroît la pénibilité du trajet et augmente le coût du transport perçu par le voyageur. A l'inverse, lors des derniers voyages, le faible nombre de voyageurs peut entraîner un sentiment d'insécurité.

2.3. L'INADÉQUATION DES MODÉLISATIONS EXISTANTES DE LA DEMANDE DE TRANSPORT COLLECTIF URBAIN

La valeur du temps varie en outre en fonction du motif de déplacement et de sa fréquence. La valeur du temps est plus faible pour des motifs de déplacements réguliers.

Sur le marché des transports collectifs urbains en Ile-de-France, l'offre est fixée en niveau et en prix par l'autorité organisatrice, le STIF. La présence d'une demande captive entraîne une certaine inélasticité par rapport à l'offre de la RATP. D'un autre côté, l'offre crée sa propre demande, et un fort accroissement de demande peut pousser le STIF à accroître l'offre de service. L'équilibre sur le marché des transports collectifs se résume donc à la courbe de demande, l'ajustement entre offre et demande est réalisé par la congestion.

Cette demande effective est fonction de déterminants liés aux motifs qui ont généré le déplacement ; l'offre intervient dans cette fonction de demande comme un ensemble de caractéristiques du bien de transport collectif urbain. Ce phénomène explique pourquoi la majorité des études relatives au transport collectif urbain traite de la demande de transport.

2.3 L'inadéquation des modélisations existantes de la demande de transport collectif urbain

Selon Bonnafous (1972 [27]), un modèle économétrique correspond à une représentation simplifiée de la réalité, délimitée dans une première approche. Le choix du type de modèle dépend des performances attendues.

Dans notre cas,

- l'objet du modèle est le nombre de voyageurs sur différents sous-réseaux de la RATP, et son évolution en fonction de variables économiques
- l'objectif est de prévoir à court et moyen terme ces différents trafics.

Le choix d'un modèle dépend de l'objectif de la modélisation. Les prévisions des trafics à la RATP sont réajustées plusieurs fois par an, voire chaque mois, et concernent un horizon à court et moyen terme.

La prise en compte de ces horizons de prévision implique de réaliser un modèle sur des données mensuelles. La mise à jour mensuelle d'un grand nombre de séries, en particulier explicatives, a

un coût pour l'entreprise, du moins en temps. Pour obtenir une grande opérationnalité, le nombre de variables utilisés dans les différents modèles de trafic doit être assez réduit. Les conditions d'obtention des données ont aussi leur importance : fiabilité des données (absence si possible de chiffres provisoires), rapidité de leur obtention (durant le mois en cours ou le mois d'après), continuité des données (pas de modification des définitions ou des périmètres), ... Ainsi, il existe un perpétuel arbitrage entre l'opérationnalité, contrainte par les données disponibles, et le niveau d'exigence théorique.

Cet arbitrage a été réalisé lors de la construction postérieure des modèles de trafics (marchandises et voyageurs). Les modélisations existantes ne peuvent être utilisées directement pour réaliser notre objectif : en grande majorité, ces modèles ont été construits en prenant en compte un horizon à long terme, de façon à dimensionner au mieux des infrastructures coûteuses et construites pour durer. En outre, de nombreux modèles considèrent des liaisons inter-urbaines, et non intra-urbaines. Enfin, le faible nombre d'études économétriques repose sur des hypothèses qui ne sont plus valides aujourd'hui.

2.3.1 Le trafic inter-urbain

A l'échelle macroscopique inter-urbaine, un flux entre deux villes peut être estimé de façon relativement satisfaisante par un modèle dit "gravitaire", construit par analogie avec la loi de Newton. Le flux est fonction des populations des villes, ainsi que de la distance qui les sépare. La concurrence entre les différents modes est ensuite introduite. Cette analyse peut être améliorée en substituant aux seuls chiffres de population, des "variables d'émissivité et d'attractivité", c'est-à-dire des variables décrivant plus précisément la quantité de trafic originaire des différentes zones et y aboutissant, en affinant la notion de distance entre les villes en prenant par exemple des effets de frontière (entre Lyon et Turin par exemple), ou en jouant sur les formes fonctionnelles entre les distances et le flux. Dans certains contextes, la position relative des villes peut aussi avoir un impact, dans la mesure où des localités intermédiaires peuvent capter une partie des flux.

2.3.2 Le trafic intra-urbain

Le trafic quotidien se concentre pour l'essentiel dans les zones urbaines, mais la modélisation est plus ardue car il existe une multitude d'itinéraires possibles entre deux destinations. En outre, les temps de déplacements ne sont pas fixes, mais dépendent de la congestion du système, et il existe plusieurs modes de transport concurrents, organisés eux-mêmes en réseaux. Le type de modélisation le plus utilisé pour modéliser le trafic intra-urbain est la modélisation en quatre étapes. Ce type de modélisation est "désagrégée" : elle considère les choix des individus suivant les zones géographiques ou les caractéristiques des voyageurs.

2.3.2.1 Modèles à quatre étapes

Dans les années 1950, aux Etats-Unis, sont mis au point des modèles de représentation et de prévision du trafic urbain. Ce sont les modèles à quatre étapes, "calibrés" sur une ville donnée grâce à des données de comptage. Dans la pratique, des formes particulières des fonctions sont utilisées à chacune de ces étapes¹³ :

1. **La génération du trafic** définit les pôles de génération et d'absorption des déplacements à partir de données sur les structures spatiales (habitat, activité...) et la mobilité. Cette étape répond à la question "se déplacer ou non?".

En pratique, les modèles de générations (i.e. de fréquence des déplacements) sont le plus souvent des modèles linéaires de type :

$$Y_z = a + b_1 X_{1z} + b_2 X_{2z} + \dots + b_n X_{nz}$$

¹³Selon Gaudry [58], il est possible de diminuer le nombre de formes fonctionnelles associées à chaque étape, en utilisant la transformation de Box et Tukey, dont les formulations directes et indirectes sont, respectivement :

$$y^{(l,m)} = \frac{(y-m)^l - 1}{l} \text{ si } l \neq 0 \\ = \ln(y-m) \text{ si } l \rightarrow 0$$

et

$$y^{(l,m)} = (ly+1)^{-l} + m \text{ si } l \neq 0 \\ = \exp(y) + m \text{ si } l \rightarrow 0$$

où Y_z représente le nombre de déplacements générés ou attirés par la zone z , et les variables X_{iz} sont les caractéristiques de cette zone z (i : emploi, population...)

2. **La distribution** définit la répartition des déplacements entre ces pôles, et répond à la question "vers quelle destination se déplacer?". Le but de cette étape est de modéliser les volumes de déplacements entre une origine et une destination précise, à l'aide d'un modèle gravitaire de la forme

$$T_{ij} = f(O_i, O_j, F_{ij})$$

avec

- T_{ij} le nombre de déplacements entre la zone i et la zone j
- O_i le nombre de déplacements ayant pour origine la zone i
- O_j le nombre de déplacements ayant pour origine la zone j
- F_{ij} une mesure du coût généralisé du déplacement entre la zone i et la zone j (distance, temps, prix...)

Ces modèles de distribution sont généralement multiplicatifs.

3. **Le choix du mode** ("par quel moyen se déplacer?") décrit la répartition des déplacements, pour chaque origine - destination, sur chacun des modes.

Les modèles de choix du mode sont souvent des modèles probabilistes de type LOGIT associés à des fonctions d'utilité linéaire¹⁴. Ce type de modèle a la formulation suivante :

$$p_m = \frac{\exp(U_m)}{\sum_{m=1}^M U_m}$$

dans laquelle p_m représente la probabilité de choisir le mode m .

¹⁴Le modèle LOGIT, est le modèle probabiliste le plus populaire. Il s'agit d'un cas particulier au sein d'une famille plus large. Par exemple, les familles BOX-COX LOGIT [60], DOGIT [43] ou INVERSE POWER TRANSFORMATION [57] permettent d'introduire des phénomènes de captivité à un ou plusieurs modes, ainsi que des effets de seuil, qui introduisent une asymétrie dans la courbe de réaction de la probabilité à des modification des variables explicatives ou des fonctions d'utilité.

2.3. L'INADÉQUATION DES MODÉLISATIONS EXISTANTES DE LA DEMANDE DE TRANSPORT COLLECTIF URBAIN

Les fonctions "d'utilité représentative" U_m associées au mode m peuvent s'écrire

$$U_m = b_{m0} + b_{m1} (f_{m1} (X_1)) + \dots + b_{mK} (f_{mK} (X_K))$$

où les variables explicatives X_k , ($k = 1, \dots, K$), représentent les caractéristiques des modes et les caractéristiques socio-économiques des consommateurs. Les fonctions f_{mk} sont souvent des fonctions linéaires ou encore le logarithme népérien, mais peuvent être une transformation de Box-Cox, plus générale.

4. **L'affectation** décrit, à l'intérieur d'un mode, la répartition des déplacements entre les différents itinéraires possible sur le réseau considéré. Cette étape répond à la question suivante : "quel chemin emprunter?".

Ces modèles ont été appliqués en France à partir des années 1960, après avoir été adaptés. En particulier, la RATP et l'IAURIF prendront en compte le fait que les modèles américains sont fortement orientés "automobile" et donc difficilement applicables en région parisienne.

Ce type de modèle s'appuie sur le comportement moyen d'un individu. Or, la variabilité des comportements au sein d'une agglomération aussi importante que l'Ile-de-France rend cet individu moyen peu représentatif de l'ensemble de la population. Des recherches [166] se proposent de construire des typologies d'individus de façon à réduire la variabilité au sein des classes dégagées. Une autre solution consiste à prendre en compte des modèles désagrégées, qui considèrent les décisions au niveau individuel.

2.3.2.2 Modèles désagrégés, à choix discrets

Les modèles désagrégés considèrent les décisions du mode de déplacement au niveau individuel. Des données sont recueillies sur les caractéristiques des individus (composition familiale, revenu, statut socio-professionnel...), le type de déplacements effectués (travail, achats, loisirs...), les caractéristiques de ces déplacements (origine, destination, temps disponible...), ou les caractéristiques des modes alternatifs (coût, temps, fiabilité, confort...). Les choix réellement effectués sont observés afin de calibrer les paramètres de la fonction d'utilité qui maximisent une fonction de vraisemblance sous contrainte des différents facteurs (revenu, temps...) [24]. La probabilité de

choisir un mode donné est ainsi déterminée pour un individu, et, par agrégation, il est possible de prédire les choix collectifs.

Il existe une variante méthodologique à ces modèles de "préférences dévoilées" : en laboratoire, des individus sont confrontés à des choix hypothétiques, concernant par exemple l'émergence d'un nouveau mode de transport, ou la modification radicale de l'offre de transport. On parle de modèles à "préférences exprimées".

Le problème de ce type de modèle réside dans la difficulté à recueillir les données adéquates : les individus doivent être sélectionnés selon une base géographique, temporelle, modale... L'estimation des modèles de choix discrets pose aussi de nombreux problèmes dans la mesure où l'ensemble des individus n'est pas soumis au même ensemble de choix, et où les décisions des différents individus ne sont pas indépendantes. Enfin, ces modèles sont délicats à mettre en oeuvre en présence d'alternatives nouvelles (par exemple, mise en place d'un nouveau mode de transport) dont l'utilité ne peut être calculée à partir de données existantes.

Les travaux de recherche actuels [51] insistent de plus en plus sur la nécessité de prendre en compte de véritables boucles de déplacement, i.e. des chaînes d'activités effectuées par l'individu entre sa sortie du domicile et son retour au domicile (par exemple : "domicile - travail - loisir - domicile"), et non des déplacements isolés ("domicile - travail", "travail - loisir", "loisir - domicile"). La prise en compte de ces boucles de déplacement a un impact sur le choix modal : le mode utilisé pour le premier déplacement a un impact fort sur les modes choisis pour les déplacements suivants.

L'accumulation des décisions individuelles engendre une situation de congestion face à laquelle l'individu va rétroagir. Les conséquences collectives des comportements individuels doivent être pris en compte, ce qui est réalisé dans le cadre "de modèles de trafic".

2.3.2.3 Modèles de trafic

La demande de transport est connue et donnée sous la forme d'une matrice origines - destinations, éventuellement assortie de contraintes temporelles. Le problème est de déterminer les itinéraires empruntés par les utilisateurs du réseau. Ces itinéraires dépendent du contexte dans lequel se prennent les décisions.

2.3. L'INADÉQUATION DES MODÉLISATIONS EXISTANTES DE LA DEMANDE DE TRANSPORT COLLECTIF URBAIN

- Dans le cas de décisions centralisées, une autorité centrale a décidé de façon autoritaire quel itinéraire devait être emprunté pour chaque déplacement. Le réseau est alors représenté par un graphe valué, qui indique les capacités de flux pouvant circuler à l'intérieur de chaque arête. La formulation des programmes mathématiques avec structure particulière, même dans le cas de problèmes de grande taille, ne pose pas de problèmes.
- Dans le cas de décisions décentralisées, l'individu choisit lui-même son itinéraire. Les modèles prennent en compte des graphes valués dans lesquels la vitesse de parcours sur un arc est fonction de la congestion sur cet arc et cherchent la situation telle que chaque agent économique choisit l'itinéraire le plus court pour le déplacement souhaité, compte-tenu de l'ensemble des choix effectués par les autres agents économiques. Cette situation d'équilibre est caractérisée par le fait qu'aucun individu ne peut modifier son choix d'itinéraire sans détériorer la situation. A un niveau macroscopique, on aboutit à une représentation de la circulation dans le réseau urbain et du degré d'engorgement de ces artères. La résolution numérique de ces problèmes fait appel à des techniques "d'inégalités variationnelles" extrêmement complexes.

Les premiers modèles de ce type avaient pour défaut majeur leur côté statique, alors que le temps est une composante essentielle d'un déplacement. Dans les modèles dynamiques, les agents économiques choisissent à la fois leur itinéraire et le moment de leur départ, ce qui accroît fortement la complexité des modèles et de leur résolution numérique. Pourtant, ces modèles restent partiels dans la mesure où les motifs à l'origine du déplacement, mais aussi le déplacement lui-même, sont donnés. Or, pour un motif de déplacement donné, comme les achats, un individu peut choisir entre plusieurs destinations.

2.3.2.4 Vers de nouveaux modèles de mobilité

Pour mieux exprimer les interactions qui existent entre les décisions des agents économiques que la simple agrégation, tendent à se développer depuis une quinzaine d'années des modèles intégrés de trafic, tels le modèle QUATUOR. Ces modèles, canadiens ou britanniques à l'origine, possèdent en outre l'avantage d'une plus grande convivialité du fait de leur interface graphique notamment, une mise à jour constante des modules, les avantages liés à leur utilisation par des

milliers d'utilisateurs dans le monde. On peut distinguer de véritables modèles de trafic tels EMME /2 ou TRANSCAD, et des systèmes d'information géographique ayant une vocation généraliste, mais capables d'intégrer des calculs de flux de transport sur des réseaux.

Ces nombreux modèles désagrégés ne peuvent être utilisés dans le cadre de notre étude : leur objectif principal est la prévision à long terme. Ils reposent sur des données souvent anciennes, car coûteuses à obtenir. Pour prévoir à court et moyen terme, les données utilisées doivent être mensuelles, et le modèle ne doit pas être lourd opérationnellement. Ceci nous conduit à privilégier une approche macro-économique.

2.3.3 Le faible nombre existant de modèles économétriques "agrégés"

Les recherches bibliographiques ont montré la prépondérance des modèles désagrégés, à l'exception de pays tels que les Etats-Unis et le Canada, où existaient des modèles macro-économiques assez similaires à celui que nous comptons construire.

Ainsi, en France, les principaux modèles de demande de transport collectif urbain peuvent être synthétisés dans le tableau suivants :

	Modèles macro-économiques	Modèles désagrégés
Transport en commun	Thèse de LENORMAND [109]	TERESE (SEMALY)
Transport en commun et voiture	QUINQUIN (SNCF)	ANTONIN (STIF) EMME II GLOBAL V4 (RATP) IMPACT3 (RATP) METROPOLIS MINUTP (SNCF) MODUS (DREIF) QUATUOR (SNCF) TRIPS

Informations issues de l'Inventaire des Modèles de Transport, PREDIT 1996 -2000[144]

Classification semblable à celle de BONNAFOUS et alii (1995) [28]

2.3. L'INADÉQUATION DES MODÉLISATIONS EXISTANTES DE LA DEMANDE DE TRANSPORT COLLECTIF URBAIN

2.3.3.1 La modélisation du trafic à Calgary (Canada) [33]

En 1995, deux types de modélisations sur données mensuelles étaient utilisées de façon conjointe à Calgary :

- un modèle univarié de Box et Jenkins pour les prévisions à court terme (moins de 6 mois)

De façon plus précise, le modèle retenu était un modèle *SARIMA* $(1, 1, 1)(1, 0, 1)_{12}$ sur la série en logarithme :

$$(1 - \phi_1 L)(1 - \phi_2 L)^{12} \ln(\text{trafic})_t = \alpha + (1 - \varphi_1 L)(1 - \varphi_2 L)^{12} \varepsilon_t$$

où L désigne l'opérateur retard, tel que $Lx_t = x_{t-1}$.

Ainsi, la modélisation prend en compte la valeur du trafic réalisée le mois précédent, et la valeur au même mois de l'année précédente (de façon à éliminer la saisonnalité). En outre sont ajoutées les erreurs de prévision ayant eu lieu dans le passé : si le modèle, le mois précédent, a prévu un trafic trop important par rapport à la valeur réalisée, le modèle corrige ce biais en réduisant la prévision du mois en cours. Le même type de correction a lieu pour les erreurs de prévision réalisées l'année précédente, de façon à corriger une modification de la saisonnalité.

- un modèle de régression prenant en compte les variables exogènes retardées suivantes : le taux de chômage à Calgary, la moyenne des prix du ticket et de la carte forfaitaire pondérée par leur poids d'utilisation, le nombre de salariés dans les principaux secteurs de l'économie, un indice du coût moyen de l'utilisation d'une automobile à Calgary. En outre étaient introduites des variables indicatrices de façon à tenir compte des périodes de grèves importantes, des Jeux Olympiques de 1988, des périodes de vacances... Les conditions climatiques exceptionnelles (neige, températures très faibles), le nombre de véhicules immatriculés pendant l'année ou la dispersion géographique n'apparaissaient pas comme des facteurs significatifs, de même que l'offre de service.

Ce type de modèle était plus performant que la modélisation Box et Jenkins pour la réalisation de prévisions à l'horizon d'un an.

2.3.3.2 La modélisation du trafic à Edmonton (Canada) [86]

La demande de transport dans la ville, pour un individu moyen, s'écrit :

$$q = \frac{Q}{P} = f\left(\frac{F}{G}, \frac{Y}{G}, \frac{K}{P}, \frac{Z}{P}\right)$$

avec

- F le prix du voyage, qui a un impact négatif sur le transport collectif
- Y le revenu moyen des habitants de la ville

Le revenu a un impact sur le trafic des transports en commun qui peut être positif (accroissement du nombre de déplacements liés au loisir par exemple) ou négatif (achat d'une voiture...).

- G le niveau global des prix.

Le niveau général des prix est un déflateur du prix du voyage ou du revenu.

- K le nombre de kilomètres réalisés sur le réseau, i.e. la capacité du réseau.

Si la demande de transport collectif est supérieure à cette capacité disponible, les voyageurs subissent un surpeuplement des rames ou à l'intérieur des bus (d'où une pénibilité accrue) voire des délais supplémentaires de transport (attente du métro, RER ou bus suivant). Cet élément apparaît aussi dans la modélisation du trafic à Montréal, et constitue une fonction approximative du coût généralisé de transport.

- Z un ensemble de facteurs qui influencent *a priori* le trafic (places de stationnement, parc automobile, évolution de la population...).
- P représente le nombre d'habitants : on considère dans le modèle la capacité de transport et les autres facteurs d'influence par personne.

Le modèle a été estimé par les moindres carrés ordinaires sur la période 1965-1992 (données annuelles soit 28 points), en appliquant la technique de Cochrane-Orcutt pour corriger l'existence

2.3. L'INADÉQUATION DES MODÉLISATIONS EXISTANTES DE LA DEMANDE DE TRANSPORT COLLECTIF URBAIN

d'auto-corrélations des résidus, et en appliquant la transformation de Box-Cox

$$\begin{aligned} y^{(\lambda)} &= \frac{y^\lambda - 1}{\lambda} \text{ si } \lambda \neq 0 \\ &= \ln(y) \text{ si } \lambda = 0 \end{aligned}$$

L'équation du nombre de voyage par personne s'écrit de la façon suivante :

$$\begin{aligned} \left(\frac{Q_t}{P_t}\right)^{(-0.1)} &= -25.374 - 0.205 \left(\frac{F_t}{P_t}\right) + 0.001 \left(\frac{Y_T}{P_t}\right) + 0.976 \left(\frac{K_t}{P_t}\right) \\ &+ 0.718 \left(\frac{P'_t}{P_t}\right) + \sum_{i=1}^3 \alpha_i D_{it} - 0.367 \varepsilon_{t-1} \end{aligned}$$

où P'_t représente le nombre d'habitants du centre ville de Edmonton, et les D_{it} sont des variables indicatrices visant à corriger l'impact des grèves, un changement tarifaire, et la modification de la méthode de mesure du trafic.

2.3.3.3 La demande de transport en commun à Montréal : le modèle DEMTEC

A partir de 1975, la Commission de Transport de la Communauté Urbaine de Montréal (CTCUM) a utilisé un modèle appelé DEMTEC (DEMande de Transport En Commun) pour l'analyse des facteurs déterminants de sa clientèle et la prévision mensuelle. Ce modèle est un modèle agrégé en séries chronologiques mensuelles. Deux types de marchés sont pris en compte : les adultes et les scolaires. Sur chaque marché, à la date t , la demande de transport en commun s'écrit sous la forme d'une régression linéaire (voir [54], [55], [56] et [44]) :

$$q_t^D = f_D(P_t, T_t, W_t, Y_t, A_t, \dots)$$

avec

- P_t les niveaux des prix des transports en commun, mais aussi le coût d'un véhicule particulier, le niveau général des prix...
- T_t les niveaux de service dans les différents modes (vitesse, confort, temps d'attente et de correspondance...)

- W_t les conditions atmosphériques
- Y_t le revenu
- A_t les variables d'activités (emplois, achats, ...)
- ainsi que d'autres variables pour prendre en compte des évènements particuliers (grèves, salons...) ou des changements ayant un impact sur la mesure de la variable expliquée (éléments calendaires par exemple).

L'offre de service est une fonction du trafic observé 12 mois auparavant et des facteurs définissant la fonction de demande de transport collectif :

$$q_t^O = f_O(q_{t-12}, P_t, T_t, W_t, Y_t, A_t, \dots)$$

Le nombre de personnes transportées résulte de la confrontation des courbes d'offre et de demande. Si le nombre de voyageurs potentiels excède la capacité réelle du réseau, l'excès de demande se traduit par des délais d'attente (i.e. une hausse du coût généralisé du transport collectif) voire par un report vers un autre mode de transport tel que la marche à pied (le nombre de voyages réellement effectués est alors inférieur au nombre de voyages désirés). En moyenne néanmoins, la capacité du réseau est suffisante pour répondre à la demande. Et l'offre est introduite dans la demande effective par le biais des temps d'attente

$$TW_t = f_T(q_t^O) = f_{T2}(q_{t-12})$$

L'estimation est corrigée des auto-corrélations par la méthode de Cochrane Orcutt, et effectuée sur la période décembre 1956 - décembre 1971. Les principales élasticités de la modélisation sur le marché des adultes sont :

2.3. L'INADÉQUATION DES MODÉLISATIONS EXISTANTES DE LA DEMANDE DE TRANSPORT COLLECTIF URBAIN

Elasticité de la demande	
• au prix du voyage	-0.15
• au nombre de voitures	-0.10
• au prix des autres biens	-0.16
• au revenu	0.08
• au temps d'attente sur la CTCUM	-0.54
• au temps de voyage sur la CTCUM	-0.27
• au temps de voyage en voiture	0.42

Source : Marc Gaudry [55] p.254

Ce modèle a été amélioré, en introduisant des variables retardées, en modifiant la forme fonctionnelle ou changeant de méthode d'estimation pour corriger l'hétéroscédasticité. Ce type de modélisation a été implanté à Toronto en 1981, et a été envisagé à la RATP. Les performances du modèle ont été excellentes jusque dans les années 80 (moins de 1% d'erreur), où a été créé un pass mensuel (assez similaire à la carte Orange) donnant droit à un nombre illimité de voyages sur les autobus et le métro.

Le problème des laissez-passer a été traité en 1996 par Arbic et Gaudry [9]. Le consommateur représentatif décide à la fois du nombre de déplacements effectués dans le mois, et le type de titre choisi pour chaque déplacement (ticket ou pass mensuel). Si le modélisateur s'intéresse uniquement au choix du titre de transport, une approche probabiliste prenant en compte les caractéristiques des alternatives peut être effectuée. Cette approche est semblable à celle utilisée pour le choix du mode de transport dans l'étape 3 du modèle désagrégé à 4 étapes.

Pour connaître la demande associée à chaque titre, une modélisation de Format Quasi Direct (QDF) peut être utilisée. Cette méthode est détaillée en 1998 par Gaudry et Liem [59], pour le choix du mode de transport. La demande par titre de déplacements (T_m) est estimée par le produit de deux estimations successives :

- une fonction des parts de marchés des différents titres (S_m), obtenue par exemple par un

modèle de probabilité

$$S_m = \frac{U_m}{\sum_m U_m}$$

- une fonction de demande totale indépendante des titres (T). Chez Gaudry et Liem, la fonction de demande totale est une régression en niveau, toutes les variables étant soumises à une transformation de Box-Cox distincte :

$$T^{(\lambda_y)} = \beta_0 + \sum_k \beta_k X_k^{(\lambda_k)} + \beta_U U^{(\lambda_U)} + \varepsilon$$

avec U un indice non-linéaire des utilités associées aux différents titres et ε les résidus, éventuellement hétéroscédastiques ou auto-corrélés.

Ces différents modèles macro-économiques sont proches du modèle que nous désirons construire, car leur mise à jour est assez aisée, de même que la réalisation de prévisions à court et moyen terme. Néanmoins, ces différentes modélisations sont réalisées sur des séries en niveau, ce qui ne peut être le cas au sein de la RATP. En effet, nous montrerons par la suite que les séries ont subi des ruptures qui doivent être prises en compte, mais aussi que les séries utilisées ne sont pas stationnaires, et que ce type de modélisation risque de conduire à des régressions fallacieuses.

En outre, nous montrerons aussi que les séries qui sont utilisées pour les analyses de la RATP correspondent généralement à des séries corrigées des effets calendaires et saisonniers. Ces différents effets n'ont plus lieu d'apparaître dans la modélisation par le biais de variables indicatrices (nombre de week-ends dans la modélisation de Gaudry, par exemple) ou de variables retardées d'une année.

Les différents modèles utilisés à la RATP pour prévoir la demande de trafic ne constituent pas une réponse à notre problème.

2.3.4 Modèles utilisés à la RATP

2.3.4.1 Les prévisions conjoncturelles de la RATP réalisées au sein du Département Commercial

Jusqu'à l'année 1996, les prévisions pour les quatre sous-réseaux traditionnels (Métro, RER, bus de Paris et bus de banlieue) étaient effectuées en appliquant les méthodes de Box et Jenkins univariées aux séries corrigées des perturbations exceptionnelles et des incidences calendaires¹⁵. Pour les réseaux plus récents, tels que le tramway, Trans Val-de-Marne ou Orlyval, des méthodes de lissage exponentiel étaient utilisées.

Ces méthodes extrapolent le passé sans utiliser de variables explicatives, aussi les prévisions subissaient quelques corrections de façon à prendre en compte l'impact des actions commerciales, des modifications du réseau, des politiques anti-fraudes ou encore des modifications de l'environnement économique.

Ce type de modélisation a été abandonné suite aux grèves de 1995 car il conduisait à des prévisions toujours en dessous des trafics observés sur les sous-réseaux. Actuellement, la RATP réalise ses prévisions en prolongeant la tendance observée au cours des dernières périodes, et en corrigeant les modifications précédemment évoquées.

2.3.4.2 Le modèle global de la RATP V4 (1997)

Développé par Bernard LABBE pour la première version (1970), puis par Claudine MAZEL, le Modèle Global de la RATP est un modèle de prévisions de trafic classique à quatre étapes, qui a pour but d'évaluer les conséquences de la mise en service d'une nouvelle ligne de transport en commun, du prolongement ou de la modification de l'offre de service. En particulier, sont évaluées les modifications en terme d'origine, de destination, de volume, de gains de temps des utilisateurs, de reports entre les modes (voiture particulière, transport en commun routier ou ferroviaire, marche à pied), de charge maximum de la ligne...

Le Modèle Global étudie les déplacements en Région Ile-de-France à l'heure de pointe du matin, ce qui permet de dimensionner l'offre de transport. Les données de base sont les popu-

¹⁵Ces corrections sont présentées de façon détaillée dans le chapitre 3.

lations et emplois par communes, ainsi que les migrations alternantes, car le motif travail est le motif prépondérant des déplacements en transport en commun à l'heure de pointe du matin. Les niveaux de services des transports en commun avec temps de correspondance, les vitesses sur les tronçons des voies, la tarification carte orange, le coût d'utilisation de la voiture particulière (dont coût de stationnement) ... sont aussi prises en compte. Le zonage de référence est découpé à la commune pour les étapes de génération - distribution, du fait de la facilité d'obtention des données sur un zonage administratif. Il est redécoupé en zones plus petites pour les étapes "choix modal" et "affectation" de façon à mettre en adéquation le zonage et l'infrastructure étudiée. La population est segmentée en trois catégories qui ont un comportement différent pour se rabattre sur les transports en commun (rabattement au plus près, rabattement de façon à minimiser le nombre de correspondances, rabattement au moyen de la voiture particulière), par conséquent, il existe trois itinéraires différents pour une même liaison. Le pourcentage des catégories est différent suivant la couronne d'habitation.

La recherche du plus court chemin sur un graphe est réalisée par un algorithme de Johnson "DIKT", le choix modal entre la marche à pied et un déplacement motorisé est modélisé par une courbe de type exponentielle, le choix modal entre voiture particulière et transports en commun est issu d'une tangente hyperbolique (ces deux courbes étant calées sur l'Enquête Globale de Transport). Les recherches en cours portent sur la modélisation du motif de déplacement école - université, et la meilleure prise en compte des déplacements intra-communaux.

En pratique, ce modèle a été utilisé pour prévoir les trafics de prolongements des lignes de métro dans les années 1970 ou lors de la création du RER, et son amélioration a servi aux études de restructuration du réseau d'autobus et de création de site propre en banlieue.

La RATP a développé deux logiciels de cartographie (EDGAR et GRAFFITI) pour cartographier les données et les résultats (charge des réseaux, montants/ descendants/ correspondants aux stations...) du modèle Global.

2.3.4.3 Le modèle IMPACT3 (1996)

Le modèle IMPACT3 a été développé pour la RATP par The MVA Consultancy dans sa version 3, les versions 1 et 2 ayant été construites au début des années 1980, avec l'assistance

2.3. L'INADÉQUATION DES MODÉLISATIONS EXISTANTES DE LA DEMANDE DE TRANSPORT COLLECTIF URBAIN

de Cambridge Systematics Europe devenu Hague Consulting Group. Il s'agit d'un modèle de tests de politique de transport mettant en oeuvre des modèles désagrégés de choix discrets, qui supposent que l'individu choisit l'alternative qui maximise son utilité¹⁶. Sa spécificité réside dans le fait que l'estimation des résultats d'un test de politique de transport est réalisée à partir d'un échantillon de déplacements et non à partir de matrices de déplacement.

Le but de ce modèle est la simulation à court et moyen terme de l'effet d'une politique de transport en région Ile-de-France sur le choix des modes de transports pour les motifs obligés, ou sur le choix du mode et de la destination pour les motifs non obligés (achats, loisirs, affaires personnelles...). L'origine et la destination observées des déplacements des individus sont connues dans un carroyage (division, habituellement géodésique en carrés ou secteurs sur une carte ou un plan) de la région de 300 mètres de côté. Quand il y a changement de destination, un zonage en 700 zones, dont 200 à Paris, est utilisé. Les phénomènes de congestion utilisent un zonage en 30 zones environ, et les déplacements sont connus à l'heure où ils ont été effectués. IMPACT3 applique aux nouvelles conditions de transport le comportement connu et modélisé des individus. Les comportements des individus ont été observés lors de l'Enquête Globale de Transport.

Ce type de modèle probabiliste estimé par la méthode du maximum de vraisemblance a été développé car, au début des années 1980, les modèles classiques à quatre étapes ne permettaient pas d'étudier correctement des problèmes tels que l'augmentation du coût de l'essence. L'interaction entre l'offre et la demande, jusqu'à l'équilibre, est obtenue par des itérations successives en faisant varier certaines variables de niveau de service prises en compte dans les modèles.

De façon plus précise, le type de modèle utilisé est un LOGIT hiérarchique comportant 5 modes :

1. Marche à pied (y compris les vélos)
2. Voiture Particulière (y compris les deux - roues hors vélos)
3. Bus uniquement
4. Réseau ferré uniquement

¹⁶Un modèle très proche de celui-ci est le modèle ANTONIN, utilisé par le Syndicat des Transports d'Ile-de-France. Face à plusieurs alternatives, l'individu choisit celle qui maximise son utilité. L'utilité du consommateur se compose d'une partie déterministe fonction de variables économiques et d'une partie aléatoire.

5. Réseau ferré et bus.

Deux types principaux de modèles désagrégés ont été construits :

1. Les modèles "domicile - motif obligé", pour lesquels les motifs obligés sont : le travail, l'école, et les affaires professionnelles. Le choix que l'individu doit réaliser est le choix du mode, la destination est imposée.
2. Les modèles "domicile - motif non imposé", tel que les achats, les loisirs (sport, visite à des amis, cinéma, promenade...) et les affaires personnelles (santé, démarches administratives...). Le choix porte alors sur le choix du mode mais aussi de la destination. Ces deux choix sont intimement liés : pour effectuer un achat, un individu peut choisir d'aller à pied dans une zone proche de son domicile, prendre un bus pour une zone peu éloignée ou sa voiture pour aller à un supermarché en grande banlieue, ou choisir le RER pour rejoindre les grands magasins parisiens.

Pour résumer l'univers de choix d'un individu dans ce contexte, il a été nécessaire de décrire de façon exhaustive l'offre de transport en Ile-de-France, et de recueillir des indices d'attractivité des zones auprès de l'APUR (Atelier Parisien d'URbanisme). En outre, les motifs de déplacement dépendent des caractéristiques sociaux-économiques de l'individu : ces éléments (sexe, activité, motorisation du ménage) ont été introduits dans le modèle.

Des phénomènes de reports de mode et/ou de destination peuvent être engendrés par la modification de la politique de transport, par le biais de phénomènes de congestion/décongestion. Par exemple, une amélioration de l'infrastructure des transports collectifs se traduit par un transfert modal de la voiture vers les transports en commun. Ceci conduit, toutes choses égales par ailleurs, à une diminution de la circulation automobile, i.e. à une baisse du coût généralisé pour les transports routiers, qui deviennent plus attractifs eux aussi. Le modèle IMPACT3 prend en compte de tels impacts en comparant, par tranche horaire, l'offre et la demande par zone. Les éléments de l'offre sont réajustés à l'aide de 4 sous modèles :

1. Ajustement des vitesses des véhicules particuliers
2. Ajustement des vitesses des bus
3. Ajustement des temps de recherche de place de stationnement

2.3. L'INADÉQUATION DES MODÉLISATIONS EXISTANTES DE LA DEMANDE DE TRANSPORT COLLECTIF URBAIN

4. Congestion dans les transports en commun (pénibilité plus ou moins grande des temps passés dans les transports en commun, possibilité de devoir attendre la voiture suivante).

L'équilibre offre - demande est obtenu par itérations successives. L'impact d'une politique est évalué en comparant cet équilibre et la situation de base, autant en termes d'évolution des déplacements (modes, durées, type de liaisons...) qu'en termes d'incidence globale (pollution, accessibilité, recettes).

Au sein de la RATP, les modèles développés de façon à prendre en compte l'environnement économique sont utilisés pour prévoir à moyen et à long terme. Leur but premier est d'aider à la décision lors des modifications de l'offre de transport, ou d'évaluer l'impact de mesures gouvernementales visant à réduire la part de l'automobile. Ils ne sont par exemple pas utilisés pour construire les objectifs budgétés sur chacun des sous-réseaux.

Lors de la construction du budget, le département Commercial réalise ses prévisions en prolongeant la tendance observée au cours des années précédentes, et en prenant en compte des éléments tels que la lutte contre la fraude, l'impact d'actions commerciales, les gains de trafic provenant des modifications futures de l'offre... Ces différents éléments ne sont pas pris en compte au sein d'un véritable modèle.

* * *

Sur le marché des transports collectifs urbains, l'offre est gérée par une autorité organisatrice et ne dépend pas de la RATP. L'équilibre ne se réalise pas par les prix, mais se confond avec la fonction de demande de transport en commun : il est réalisé par la congestion. L'offre de service constitue une caractéristique du transport proposé par la RATP : une offre insuffisante de voitures-kilomètres génère de la congestion (voyageurs serrés et/ou temps d'attente), et augmente le coût généralisé du transport en commun. L'offre apparaît comme un déterminant du choix modal en faveur des transports collectifs, même s'il existe une demande captive de ces transports en commun.

En amont de ce choix modal, les différents motifs de déplacement sont des déterminants de la demande de transport. Les liaisons "domicile-travail" expliquent l'existence d'heures de pointe,

tandis que le récent développement des affaires personnelles a conduit à l'accroissement du trafic pendant les heures creuses, en particulier le week-end et en soirée.

Les modélisations passées ont pris en compte ces différents déterminants, mais généralement dans une optique de prévision de long terme. Les modèles macro-économiques examinés ont une problématique similaire à la nôtre, à savoir la prévision à court et moyen terme. Néanmoins, les hypothèses sous-jacentes à ces modélisations (en particulier, l'absence de mémoire des chocs, et l'absence de rupture) ne sont pas compatibles avec les séries de trafic de la RATP. En outre, les séries de la RATP qui servent aux analyses et prévisions ne sont pas des données brutes. Le chapitre suivant précise les définitions des séries à modéliser, ainsi que leurs différentes propriétés statistiques.

2.3. L'INADÉQUATION DES MODÉLISATIONS EXISTANTES DE LA DEMANDE DE TRANSPORT COLLECTIF URBAIN

Chapitre 3

Les données relatives à la RATP

Le but de ce chapitre est de présenter de façon détaillée les séries qui seront utilisées et expliquées au cours des différentes modélisations.

En premier lieu, il s'agit de préciser la nature des sous-réseaux étudiés. Ces précisions ne sont pas inutiles : si le métro est un sous-réseau identifiable (l'ensemble des lignes est du ressort de la RATP), il n'en est pas de même pour le RER (Réseau Express Régional). Par exemple, sur la ligne A du RER, une portion de la ligne est gérée par la SNCF. Les sous-réseaux ont en outre connu de nombreuses modifications de leur contexture, le prolongement de la ligne 14 à Saint-Lazare en est l'exemple le plus récent. Les différentes modifications de l'offre de service seront abordées pour les différents sous-réseaux au cours de ce chapitre. Nous montrerons en outre que les unités de mesure sur les différents réseaux ne sont pas comparables : on parle "d'utilisation" du réseau ferré (RER et métro) mais de "parcours" sur le réseau des autobus (1^{ère} section).

La méthode de collecte des données brutes sera détaillée dans la deuxième section, de façon à mieux appréhender la fiabilité des données. Le fait que le métro et le RER soient des sous-réseaux fermés, contrairement aux sous-réseaux des bus de Paris ou de banlieue, ou encore du tramway, explique la plus grande fiabilité des comptages sur le réseau ferré. Elle explique aussi la richesse plus grande des données de trafic sur ce réseau : des données sont disponibles à la demi-heure, pour les différents types de titres, par station ou gare...

3.1. CARACTÉRISTIQUES DES RÉSEAUX RATP ET UNITÉS DE MESURE DU TRAFIC

Les différentes corrections apportées afin de mieux suivre l'évolution des variables mensuelles de trafic seront abordées dans la troisième section. En effet, les données utilisées pour les analyses et la prévision sont des données corrigées des perturbations, mais aussi des effets calendaires et des effets saisonniers. Un exemple détaillé nous permettra de voir ce qu'apportent ces différentes corrections.

Enfin, la quatrième section présentera de façon détaillée chacune des séries mensuelles étudiées, pour la période janvier 1990 - décembre 2002, ainsi que leur évolution au cours de la période considérée. Les fortes variations auxquelles sont soumises ces différentes variables, visibles sur les graphiques présentés, nous ont amené à tester leur degré de stationnarité, i.e. la mémoire des chocs passés. Pour chaque série, les résultats des tests de racine unitaire sur la série en niveau et la série en différence première seront reportés dans un tableau.

3.1 Caractéristiques des réseaux RATP et unités de mesure du trafic

3.1.1 Réseaux et sous-réseaux à la RATP

Deux grands réseaux sont distingués à la RATP :

- le réseau ferré (métro et RER),
- le réseau de surface (autobus et tramway).

Présentons chacun de ces réseaux ainsi que les principales évolutions de sa contexture depuis 1990. Un historique mensuel depuis 1990 sur les principaux faits marquants ayant un lien avec les trafics de la RATP, est disponible en annexe.

3.1.1.1 Le réseau ferré

Le réseau ferré est subdivisé en deux sous-réseaux :

- le métro, y compris le funiculaire rattaché à la ligne 2,
- le RER RATP.

En 1990, la longueur des lignes exploitées par la RATP est de 198.9 kilomètres sur le métro et 102.7 kilomètres sur le RER, soit une longueur de 301.6 kilomètres pour l'ensemble du réseau ferré.

Le métro Le métro ou "métropolitain" est un chemin de fer électrique, généralement souterrain, qui dessert une grande ville. Les agglomérations les plus importantes, où la densité de la population est élevée, en sont pourvues : en dehors de la région parisienne, la plupart des grandes villes françaises possèdent un réseau de métro (Lille, Lyon, Marseille, Rennes et Toulouse).

Le sous-réseau du métro que nous prenons en compte est celui de la région parisienne uniquement, géré par la RATP.

Aux lignes de métro est ajouté le **funiculaire** de Montmartre, rattaché à la ligne 2. Un funiculaire se compose généralement de deux trains reliés par un câble, et circulant sur des rails. Ce type de matériel, économe en énergie, est utilisé pour des lignes comportant de fortes pentes : le poids du train descendant fournit la quasi-totalité de l'énergie nécessaire pour tracter le train montant. Le moteur du câble ne sert qu'à fournir l'impulsion nécessaire au démarrage et, dans le cas d'un train montant plus chargé (donc plus lourd) que le train descendant, à apporter un complément d'énergie. Le funiculaire de Montmartre permet de monter au sommet de la butte Montmartre et d'accéder à la basilique du Sacré Coeur ; il existe depuis 1900 et a été modernisé en 1991.

Les caractéristiques du sous-réseau métro sont précisées dans le tableau suivant :

Caractéristiques du métro	
Vitesse commerciale moyenne à l'affluence du soir	24.9 km/h
Heure de départ du premier train	5h30
Heure d'arrivée du dernier train	1h15
Nombre de trains en service à l'affluence du matin	548
Pourcentage de la partie aérienne	≈ 8%

Source : RATP, Mémento de statistiques annuelles 2003

3.1. CARACTÉRISTIQUES DES RÉSEAUX RATP ET UNITÉS DE MESURE DU TRAFIC

A un niveau plus détaillé, les lignes de métro sont détaillées dans le tableau suivant :

Lignes de métro, situation au 31 décembre 2003			
n° de ligne	Ligne	Longueur	Nb de stations
1	Grande Arche de la Défense - Château de Vincennes	16.6 km	25
2	Porte Dauphine - Nation	12.3 km	25
3	Pont de Levallois Bécon - Gallieni	11.7 km	25
3 bis	Gambetta - Porte des Lilas	1.3 km	4
4	Porte de Clignancourt - Porte d'Orléans	10.6 km	26
5	Bobigny Pablo Picasso - Place d'Italie	14.6 km	22
6	Charles de Gaulle Etoile - Nation	13.6 km	28
7	La Courneuve 8 mai 1945 - Villejuif Louis Aragon	22.4 km	38
	La Courneuve 8 mai 1945 - Mairie d'Ivry		
7 bis	Louis Blanc - Pré Saint-Gervais	3.1 km	8
8	Balard - Créteil-Préfecture	22.1 km	37
9	Pont de Sèvres - Mairie de Montreuil	19.6 km	37
10	Gare d'Austerlitz - Boulogne Pont de Saint Cloud	11.7 km	23
	Gare d'Austerlitz - Porte d'Auteuil		
11	Châtelet - Mairie des Lilas	6.3 km	13
12	Porte de la Chapelle - Mairie d'Issy	13.9 km	28
13	Châtillon Montrouge - Saint-Denis-Université	22.5 km	30
	Châtillon Montrouge - Asnières-Gennevilliers		
14	Saint-Lazare - Bibliothèque François Mitterand	km	8

Source : RATP, Mémento de statistiques annuelles 2003

Modification de la contexture du métro depuis 1990 La contexture du métro s'est agrandie le 1er avril 1992 par le prolongement de la ligne 1, du "Pont de Neuilly" à la "Grande Arche de la Défense", sur une longueur de 1.8 kilomètres, avec l'ouverture de deux nouvelles stations : "Esplanade de la Défense" et "Grande Arche de la Défense". Grâce à ce prolongement, existe une nouvelle correspondance avec la ligne A du RER ainsi qu'avec les lignes SNCF qui

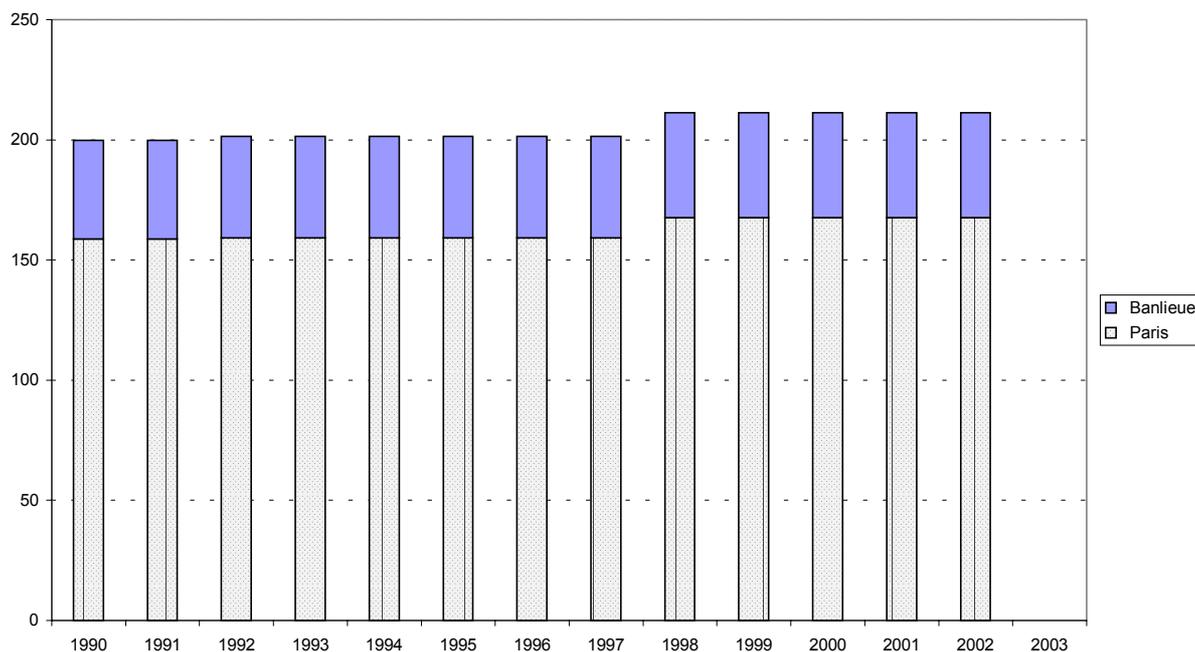
relient Saint-Lazare à Versailles ou Saint-Nom la Bretèche. La longueur du réseau métro est de 200.7 kilomètres à la fin de l'année 1992.

En 1993, l'outil de calcul du kilométrage a été réajusté, amenant la longueur exploitée sur le réseau métro à 201.5 kilomètres alors qu'il n'y a eu aucune modification de ce réseau.

Enfin, le 15 octobre 1998 est ouverte le premier tronçon de la nouvelle ligne 14, qui relie la place de la Madeleine (station "Madeleine") à la Bibliothèque Nationale de France (station "Bibliothèque François Mitterrand") en 9.8 kilomètres, portant la longueur du réseau métro à 211.3 kilomètres. Météor¹ est la première ligne de métro parisienne à pilotage automatique intégral : il n'y a plus de "pilote" humain à bord du train même, l'ensemble est supervisé par un poste de commandes et de contrôles centralisés.

Depuis le 16 décembre 2003, la ligne 14 a été prolongée jusqu'à gare Saint-Lazare, et est en correspondance avec les lignes 3, 9, 12, 13, le RER SNCF E, les lignes SNCF et plusieurs lignes de bus.

Longueur des lignes en km (voies doubles) sur le métro



Source : RATP, Mémento de statistiques annuelles

¹METEOR, pour "METro Est Ouest Rapide", est l'autre nom de la ligne 14 du métro.

3.1. CARACTÉRISTIQUES DES RÉSEAUX RATP ET UNITÉS DE MESURE DU TRAFIC

Les projets de développement du réseau de métro, engagés dans le cadre du Contrat de Plan Etat-Région 2000 - 2006, sont présentés dans le tableau suivant :

Prolongements prévus sur le métro			
Ligne	De :	A :	Mise en service prévue
4	Porte d'Orléans	Bagneux	fin 2007 jusqu'à Mairie de Montrouge
8	Créteil Préfecture	Créteil Parc des Sports	2007
12	Porte de la Chapelle	Mairie d'Aubervilliers	fin 2007 jusqu'à Proudhon Gardinoux
13	Gabriel Péri	Port de Genevilliers	fin 2007 jusqu'à Asnières Genevilliers le Luth
14	Bibliothèque F. Mitterrand	Olympiades	fin 2007

Source : RATP sur www.extension-reseau.ratp.fr

Le RER Depuis sa création en 1969, le RER (Réseau Express Régional) s'est étendu sur toute l'Ile-de-France. Il se compose de 5 lignes de chemin de fer de banlieue reliées entre elles par des tunnels passant par le centre ville. Le gabarit des lignes de RER est souvent compatible avec celui des grandes lignes de chemin de fer, mais il est supérieur à celui des lignes de métro. Il se distingue en outre du métro par son confort et sa plus grande rapidité.

En France, il existe un seul RER, celui de l'Ile-de-France. Il présente la particularité d'avoir des tronçons exploités par la RATP, et d'autres par la SNCF. Les lignes A et B sont exploitées par la RATP, à l'exception de la partie nord-ouest de la ligne A (à partir de Nanterre-Préfecture jusqu'à Poissy ou Cergy), et de la partie nord de la ligne B (à partir de la Gare du Nord jusqu'à Aéroport Charles de Gaulle ou Mitry-Claye). Ces tronçons, ainsi que les lignes C, D et E sont gérés par la SNCF.

La ligne A (RATP) dessert 6 départements et 35 gares sur 75 kilomètres de trajet. Cette ligne ferroviaire transversale est - ouest représente un quart du trafic banlieue de la région Ile-de-France grâce à des gares telles que La Grande Arche de la Défense, les pôles universitaires de Nanterre et de Marne-La-Vallée, ou les parcs de loisir Disneyland.

La ligne B (RATP) couvre 5 départements et dessert 32 gares sur un trajet de 40 kilomètres. Cette transversale nord - sud dessert les aéroports d'Orly (via Orlyval) et de Roissy-Charles de Gaulle, le Stade de France, et conduit vers la Vallée de la Chevreuse.

Les principales caractéristiques du RER RATP par rapport au métro résident dans la plus grande vitesse commerciale moyenne, et une partie aérienne plus importante :

Caractéristiques du RER RATP		
	Ligne A	Ligne B
Vitesse commerciale moyenne	49.6 km/h	38.6 km/h
Heure de départ du premier train	5h00	5h00
Heure d'arrivée du dernier train	1h20	1h15
Nombre de trains en service à l'affluence	66	52
Pourcentage de la partie aérienne	≈ 67%	≈ 85%

Source : RATP, Mémento de statistiques annuelles 2003

Les lignes de RER de la RATP sont présentées dans le tableau suivant :

Lignes de RER RATP, situation au 31 décembre 2003			
Indicatif	Tronçon commun et branches	Longueur	Nb de gares
A		75.4 km	35
	Nanterre Préfecture - Vincennes		
A1	Saint Germain en Laye - Nanterre Préfecture		
A2	Vincennes - Boissy Saint Léger		
A3	Vincennes - Marne la Vallée Chessy		
B		39.7 km	32
	Bourg la Reine - Gare du Nord		
B2	Robinson - Bourg la Reine		
B4	Saint Rémy les Chevreuses - Bourg la Reine		

3.1. CARACTÉRISTIQUES DES RÉSEAUX RATP ET UNITÉS DE MESURE DU TRAFIC

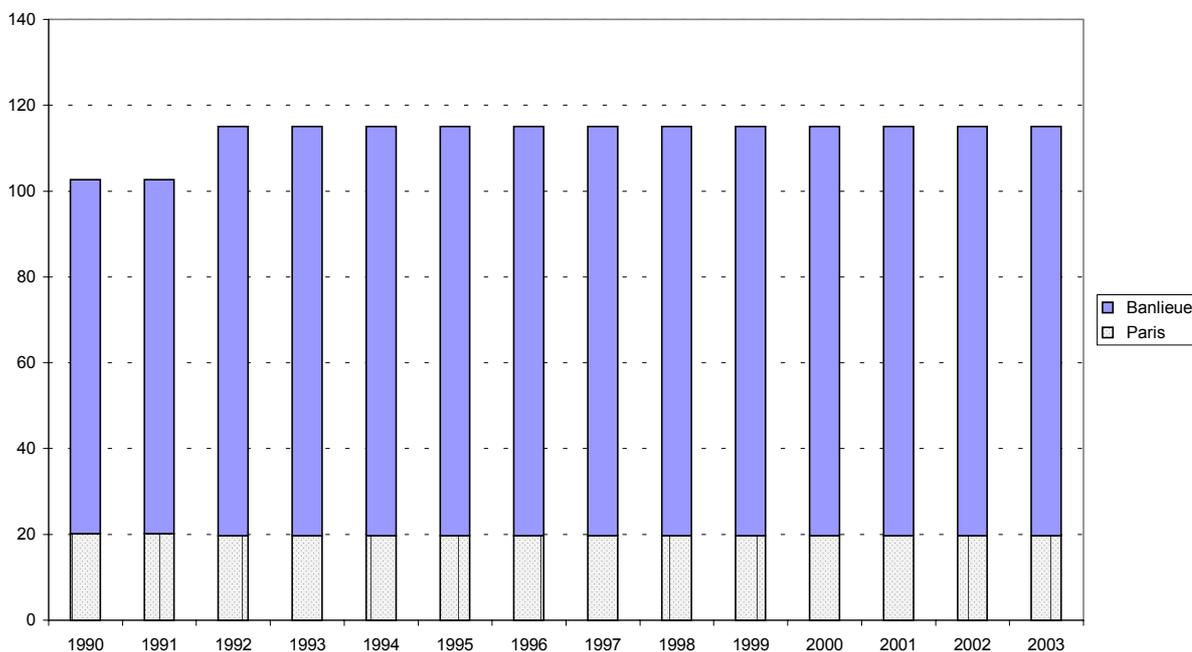
Modification de la contexture du sous-réseau RER RATP depuis 1990

Le 1^{er} avril 1992, la contexture du réseau RER s'est agrandie de 13 kilomètres du fait du prolongement de la ligne A vers l'est, de "Torcy" à "Chessy-Marne la Vallée", permettant ainsi la desserte du parc d'attractions Eurodisneyland. Sur ce prolongement, le 21 décembre, était ouverte au public la gare de "Bussy-Saint Georges", portant à 115.7 kilomètres la longueur des voies exploitées, pour 68 points d'arrêts et 65 stations nominales.

En avril 2001 a été ouverte une nouvelle gare sur le RER A, "Val d'Europe".

Il semble toutefois utile de préciser qu'il existe des modifications de la contexture du RER SNCF au cours de cette période. Sans parler de la création et des prolongements d'EOLE (ligne E), du prolongement des lignes C et D, les portions SNCF des lignes A et B ont été prolongées. En mai 1994, la ligne A du RER SNCF a été prolongée vers l'ouest, avec ouverture de 2 nouvelles gares : "Cergy le Haut" et "Neuville-Université". La ligne B du RER est prolongée au nord à "Aéroport Charles de Gaulle 2-TGV" en novembre 1994, tandis que la gare "La Plaine Stade de France" est ouverte en janvier 1998.

Longueur des lignes en km (voies doubles) sur le RER



Source : RATP, Mémento de statistiques annuelles

L'extension du réseau de RER, dans les domaines RATP comme SNCF, semble marquer une pause. Le Contrat de Plan Etat-Région 2000 - 2006 ne prévoit aucune extension.

3.1.1.2 Le réseau de surface

Jusqu'en 2003, le réseau de surface était subdivisé en trois sous-réseaux ;

- les lignes d'autobus de Paris,
- les lignes d'autobus de banlieue,
- et les lignes en sites propres.

Les principales caractéristiques sur le réseau de surface de la RATP sont résumées dans le tableau suivant :

Caractéristiques du réseau de surface de la RATP		
	Paris	Banlieue + services communaux
Vitesse commerciale moyenne	12.9 km/h	17.9 km/h
Premiers départs	Avant 6h00 : 40%	Avant 5h30 : 61%
	6h00 - 6h30 : 55%	5h30 - 6h00 : 15%
	Après 6h30 : 5%	Après 6h00 : 24%
Derniers départs	Avant 21h00 : 7%	Avant 20h30 : 22%
	21h00-21h30 : 50%	20h30-21h30 : 24%
	21h30-00h00 : 12%	21h30-00h00 : 22%
	Après 00h00 : 31%	Après 00h00 : 32%

Source : RATP, Mémento de statistiques annuelles 2003

Les lignes d'autobus de Paris comprennent les lignes remisées² de Paris, et le Noctambus.

- Il existe 57 lignes de bus RATP dans Paris, pour une longueur déployée de 554 kilomètres.
- La ligne PC ("Petite Ceinture" longeant le périphérique) est aujourd'hui scindée en 3 tronçons.

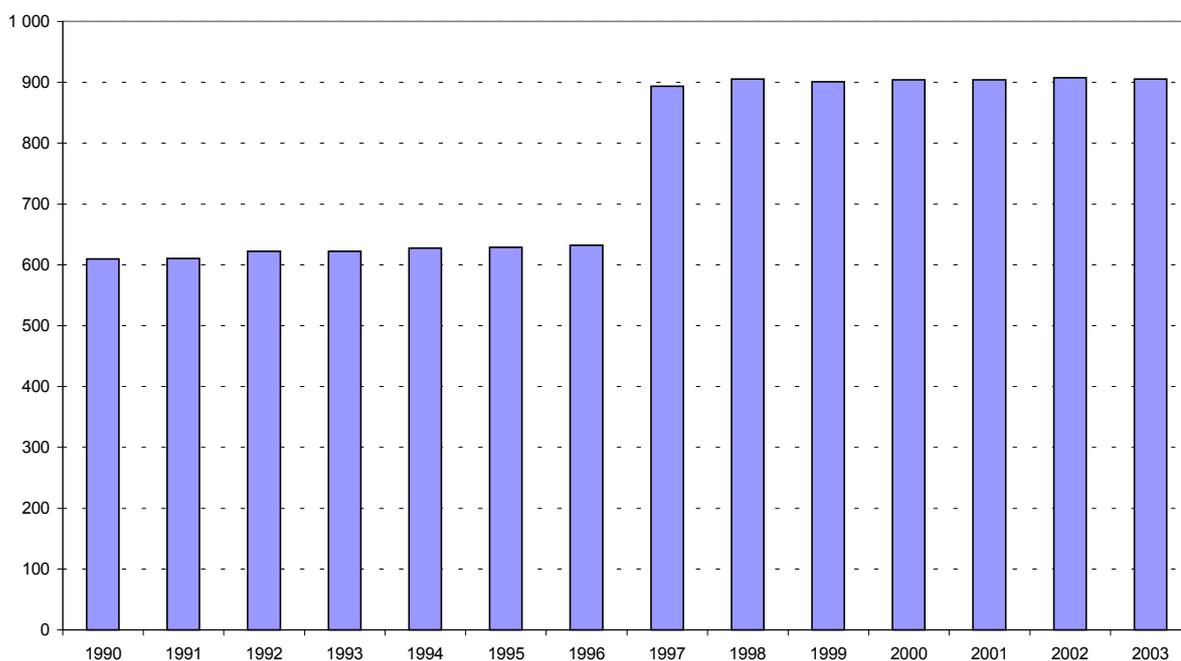
²On parle de **ligne remisée** lorsque la ligne de bus utilise du matériel et des agents RATP. 23 centres bus, anciennement appelés "dépôts", assurent le remisage et la maintenance des 4000 bus de la Régie.

3.1. CARACTÉRISTIQUES DES RÉSEAUX RATP ET UNITÉS DE MESURE DU TRAFIC

- La ligne **Balabus** est exploitée uniquement les dimanches et jours fériés d’avril à septembre, elle permet de traverser la capitale d’Ouest (La Défense) en Est (Gare de Lyon) en passant par les sites et les monuments les plus intéressants.
- Le réseau **Noctambus** a totalement été refondu en septembre 1997. Il est composé de 18 lignes de bus, circulant 7 jours sur 7 entre 1h et 5h35 du matin (1 bus par heure) dans Paris et dans un rayon de 15 kilomètres autour. La longueur de la ligne déployée est de 322 kilomètres. Le Noctambus est accessible par un billet spécial, le billet Noctambus (2.70 euros), valable sur l’ensemble du trajet (y compris en cas de correspondance avec un autre Noctambus), ou un titre valable sur les zones traversées.

Le détail et les modifications de la contexture du réseau de bus de Paris seraient trop longues à détailler. Le graphique suivant indique la modification du nombre de kilomètres parcourues par les lignes déployées. La forte augmentation du nombre de kilomètres en 1997 s’explique principalement par la refonte du réseau Noctambus (nombre de kilomètres multiplié par 4).

Longueur déployées des lignes de Paris (Lignes remisées, Balabus et Noctambus) en km



Source : RATP, Mémento de statistiques annuelles

Les lignes d'autobus de banlieue comprennent des lignes remisées³, les services communaux, les lignes affrétées⁴ de la CEA-CGEA, d'Athis Cars et du service à la demande ainsi que le STL (Services de transports et locations).

De façon plus précise, il existe

- 193 lignes remisées de bus en banlieue, pour une longueur déployée de 2092 kilomètres.
- 39 services communaux de banlieue, pour une longueur déployée de 297 kilomètres.
- des lignes affrétées :
 - 1 ligne CEAT (ligne 421 de 18 kilomètres),
 - 7 lignes CGEA - Transport qui constituent le réseau Traverciel (sauf ligne 467) d'une longueur déployée de 63 kilomètres,
 - 5 lignes ATHIS-CAR pour un total de 41 kilomètres,
 - 1 service départemental pour handicapés dans le Val de Marne. Le **SERVAL** est de longueur indéterminée : il correspond à un service à la demande de minibus équipés de 5 fauteuils ancrés et conduits par du personnel spécialisé.
- le STL regroupe 26 services communaux, 8 dessertes en contrat de villes (Villiers, Pantin, Nanterre, Sarcelles...), 2 dessertes de zones d'activités (Port de Genevilliers, Plaine-Saint-Denis, Plessis-Robinson), et 2 dessertes scolaires (Saint-Cloud et Drancy).

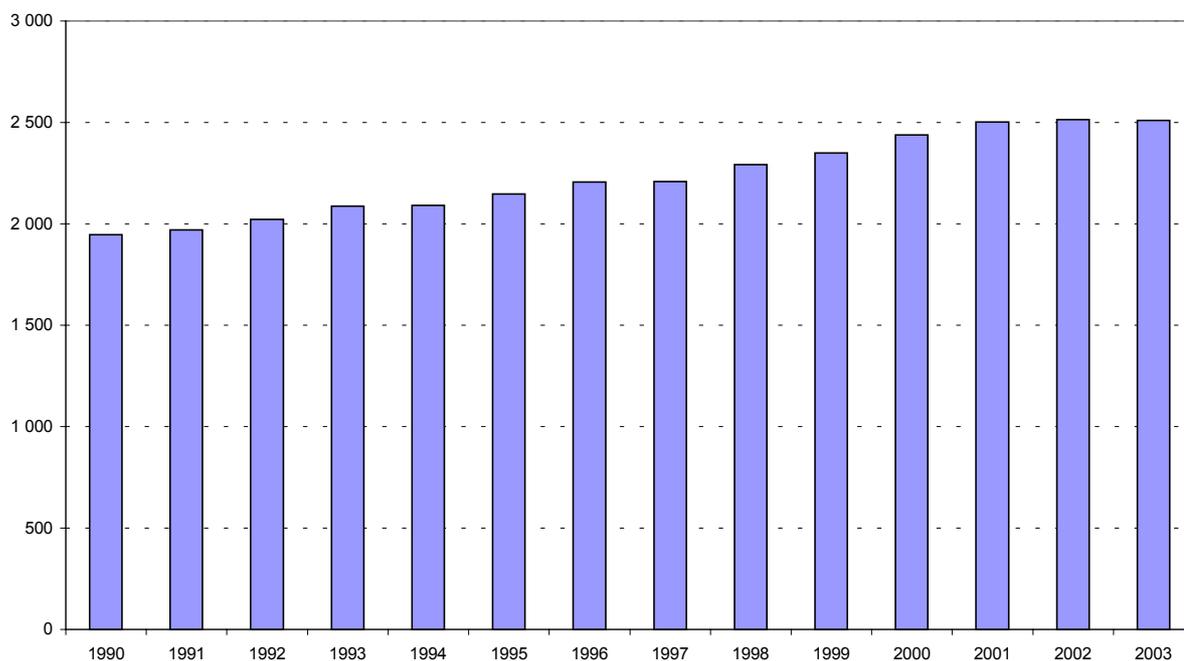
Les modifications de la longueur déployée du réseau des autobus de banlieue sont synthétisées dans l'histogramme suivant. Le détail de chacune des lignes et des changements de leurs itinéraires aurait été trop fastidieux, et aurait masqué le plus important, à savoir l'augmentation continue de la desserte en banlieue par la RATP.

³On parle de **ligne remisée** lorsque la ligne de bus utilise du matériel et des agents RATP. 23 centres bus, anciennement appelés "dépôts", assurent le remisage et la maintenance des 4000 bus de la Régie.

⁴Une **ligne affrétée** est telle que ni le matériel ni les agents n'appartiennent à la RATP. La billetterie est RATP, de même que l'encadrement, la mise en place des tableaux de marche, la régulation...

3.1. CARACTÉRISTIQUES DES RÉSEAUX RATP ET UNITÉS DE MESURE DU TRAFIC

Longueur déployée des lignes de banlieue (Lignes remisées et affrétées, services communaux et SERVAL) en km



Source : RATP, Mémento de statistiques annuelles

En particulier, des restructurations sectorielles du réseau bus dans les zones de banlieue ont été engagées en janvier 1989 sous le nom "Autrement Bus", dans le but de simplifier et de moderniser le réseau. Pendant des dizaines d'années, le réseau des autobus de banlieue s'était fortement complexifié, du fait notamment de nombreuses extensions de lignes. Par ailleurs, le réseau était peu lisible du fait de numérotation mixte (combinaison d'un numéro et d'une lettre, par exemple : lignes 144A, 144B). Pour mieux répondre à la demande, faciliter la compréhension et l'utilisation du réseau, les trajets ont été modifiés, prolongés ou raccourcis et la numérotation a été simplifiée par la suppression de la numérotation mixte.

Les lignes en site propre circulent en site protégé par rapport à la circulation automobile. A la RATP, les lignes en site propre regroupent le tramway, appelé aussi "mode T" ainsi que Orlyval. Le tramway est une forme de transport en commun souvent à traction électrique, circulant sur des rails plats (ceux des trains sont légèrement inclinés vers l'intérieur de la voie).

- La mise en service du **tramway** en site propre⁵, reliant sur 9 kilomètres la gare de Saint Denis (RER) à la station de métro "Bobigny Pablo Picasso" (ligne 5), s'est réalisée en deux étapes. Le 6 juillet 1992 a été mise en service la liaison "Bobigny Pablo Picasso" à "La Courneuve-8 mai 1945", tandis que le 21 décembre 1992 était ouverte la liaison "La Courneuve-8 mai 1945" à "Saint-Denis Gare RER". Cette ligne dessert 21 stations, dont quatre possèdent des correspondances directes ou rapprochées avec des infrastructures radiales : la SNCF à la gare de Saint-Denis, et le métro (ligne 13 pour "Saint-Denis Basilique", ligne 7 pour "La Courneuve-8 mai 1945" et ligne 5 pour "Bobigny Préfecture", cette dernière station s'intégrant dans un complexe d'échanges plus vaste comprenant un terminal d'autobus et un parc de stationnement). Le 15 décembre 2003, la ligne **T1** de Bobigny a été prolongée jusqu'à la gare de Noisy-le-Sec, sur 2.9 kilomètres, avec desserte de 5 nouvelles stations. 6 lignes de bus du secteur ont adapté leur offre à ce prolongement.
- En octobre 1993, la première ligne d'autobus en site propre continu a été mise en place, sous le nom de "Trans Val de Marne" (**TVM**). Cette ligne relie cinq communes du Val de Marne entre le Marché International de Rungis et la gare RER de Saint-Maur Créteil, à savoir Chevilly-Larue, Thiais, Choisy-le-Roi, Créteil et Saint-Maur. Elle assure en outre la correspondance avec plusieurs lignes de métro, le RER et la SNCF ainsi qu'une vingtaine de lignes de bus.
- La ligne **T2** du tramway est inaugurée en 1997 sous le nom de "Tram Val de Seine" (TVS) entre La Défense et Issy Val de Seine.
- Depuis le 4 février 1993, la RATP est chargée de la responsabilité du service **Orlyval** et a confié à sa filiale RATP Val Service (RVS) l'exécution de l'exploitation et de la maintenance de la liaison entre Antony et les aéroports d'Orly. OrlyVal est un métro entièrement automatique et sans conducteur, et non un tramway. Il permet de relier l'aéroport à Denfert-Rochereau en 20 minutes seulement grâce à une vitesse commerciale moyenne de 35 km/h et offre un service régulier (intervalle de 4 minutes entre 2 rames en heure de pointe, et 7

⁵ Il serait plus juste de parler d'un renouveau du tramway à Paris dans les années 1990, plutôt que d'une mise en service. Jusqu'en 1926, la ville de Paris était sillonnée par un réseau de tramway dense dans la ville et ses faubourgs. Sous la pression des intérêts automobiles et pétroliers, le démantèlement du réseau parisien de tramway a fini a été achevé en 1938 au profit d'une desserte par les autobus.

3.1. CARACTÉRISTIQUES DES RÉSEAUX RATP ET UNITÉS DE MESURE DU TRAFIC

minutes en heure creuse).

Les lignes en site propre peuvent être synthétisées dans le tableau suivant :

Lignes de transport collectif en site propre			
Indicatif	Ligne	Longueur	Nombre d'arrêts
T1	Gare de Saint-Denis - Noisy-le-Sec	12 km	26
T2	La Défense - Issy Val de Seine	11.4 km	13
TVM	Rungis - Saint Maur-Créteil	12.7 km	22
Orlyval	Antony - Orly Sud	7.3 km	3

Source : RATP Département Bus, Statistiques annuelles 2003

Contrairement aux autobus, les tramways ne sont pas pénalisés par les contraintes liées à l'accroissement de la circulation routière, qui limitent la vitesse commerciale. Le gain de temps par rapport aux autobus qui desservaient auparavant les trajets de T1 est par exemple de l'ordre de +50%, et s'accompagne d'une grande régularité. La vitesse commerciale moyenne sur le tramway est égale à 24 km/h.

En plus de ces qualités, le tramway possède une capacité à remodeler la ville, en requalifiant les espaces dévolus aux transports collectifs, aux piétons ou à la voiture. Il bénéficie d'aménagements facilitant son accessibilité, de climatisation dans ses voitures, ... et son coût de réalisation au kilomètre est près de 3 fois plus faible que celui du métro (il s'agit d'aménager la voirie et non de creuser un tunnel). Ces différents éléments expliquent la présence croissante de tramways dans les villes françaises⁶ et le nombre de prolongements envisagés en Ile-de-France, présentés dans les

⁶Nantes, Grenoble, Rouen, Strasbourg, Montpellier, Lyon, Orléans, Nancy, Bordeaux... sont desservies par un tramway, et ce sera bientôt le cas de villes telles Le Mans, Mulhouse, Nice et Valenciennes.

CHAPITRE 3. LES DONNÉES RELATIVES À LA RATP

Prolongements envisagés sur les lignes de transport en commun en site propre

Ligne	De :	A :	Mise en service prévue
T1	Gare Saint Denis-Asnières	Genevilliers AGIII	fin 2008
T1	Noisy-le-Sec	Montreuil	?
T2	Issy-Val de Marne	Porte de Versailles	2007
T2	La Défense	Pont de Bezons	2008-2009
TVM	Rungis	Croix de Berny	2006
TVM	Saint-Maur-Créteil	Champigny-les-Boullereaux	?

Source : RATP sur www.extension-reseau.ratp.fr

Créations envisagées de lignes de tramway

Ligne	De :	A :	Mise en service prévue
Maréchaux Sud	Pont du Garigliano	Porte d'Ivry	mi -2006
Maréchaux Est	Porte d'Ivry	Porte de Charenton	2007
	Châtillon	Viroflay	2009
	Villejuif	Athis-Mons	2009
	Place du 8 mai 1945 (St Denis)	Sarcelles (RER D)	2007-2008
	St-Denis Porte de Paris	Villetaneuse	2010
	Meudon-sur-Seine	Parc de Saint-Cloud	?

Source : RATP sur www.extension-reseau.ratp.fr

Créations envisagées de lignes de bus en site propre

De	A :	Mise en service prévue
Pompadour	Sucy-Bonneuil	2008
Vitry-sur-Seine	Choisy-le-Roi	?

Source : RATP sur www.extension-reseau.ratp.fr

3.1.2 Unités de mesure du trafic à la RATP

Un "déplacement" est un mouvement entre une origine et une destination, caractérisé par un motif, pouvant être réalisé à l'aide d'un ou plusieurs modes de transports (y compris la marche à pied). Pour mesurer le nombre de déplacements sur les réseaux de la RATP, la notion de "voyage effectué" est utilisée.

3.1.2.1 Voyage effectué

Depuis 1975, date de la création de la Carte Orange qui donne droit à un nombre illimité de voyages, l'unité de mesure est le "voyage effectué" alors que l'on utilisait auparavant les "voyages payés".

Les voyages effectués sont mesurés par le nombre de personnes physiques qui franchissent les lignes de contrôle du réseau ferré (comptages enregistrés aux péages magnétiques installés en entrée de station de métro ou de gare RER), ou qui pénètrent dans les voitures sur le réseau routier (comptages enregistrés par les "marches sensibles" ou "tapis - contacts" installés à la montée dans les autobus).

Pour le réseau RATP, un voyage effectué correspond à un "parcours" sur le réseau de surface et à une "utilisation" sur le réseau ferré. Précisons ces deux notions.

3.1.2.2 Parcours et utilisation

- Il y a "utilisation" d'un réseau quand un voyageur utilise au moins une ligne de ce réseau au cours de son déplacement. Sur le réseau ferré, l'utilisation est l'unité de mesure qui correspond au voyage effectué.

Sur le métro, les utilisations sont calculées en sommant le nombre d'entrants directs⁷ sur le métro et le nombre de correspondants⁸ en provenance du RER. Les correspondants en

⁷Un "entrant direct" est un voyageur pénétrant pour la première fois sur le réseau ferré de la RATP au cours de son déplacement. Il peut provenir aussi bien du domaine public que des points d'arrêt des réseaux de surface de la RATP (lignes d'autobus et sites propres) ou de la SNCF (grandes lignes, lignes de banlieue ou RER SNCF).

⁸Un "correspondant" est un voyageur qui passe d'un sous-réseau du réseau ferré RATP (métro et RER RATP) à l'autre. Un voyageur SNCF qui accède au réseau RATP en empruntant une correspondance, en passant par le domaine public, en changeant de quai dans une gare commune, ou en restant dans le même véhicule, est considéré comme un entrant direct.

provenance du RER SNCF sont comptabilisés comme des entrants directs (lignes B et D du RER SNCF à Gare du Nord, ligne D du RER à Gare de Lyon).

Sur le RER, les utilisations sont calculées en ajoutant le nombre d'entrants directs sur le RER RATP (y compris correspondants en provenance du RER SNCF) et le nombre de correspondants provenant du métro.

Les utilisations du réseau ferré ne correspondent pas obligatoirement à la somme des utilisations sur le métro et sur le RER dans la mesure où les voyageurs peuvent utiliser soit le métro exclusivement, soit le RER exclusivement, soit combiner métro et RER avec une ou plusieurs correspondances. Toutefois, il est possible d'additionner les utilisations du métro et du RER (et compter deux fois les personnes effectuant un trajet combinant métro et RER RATP) à partir du moment où l'on décide de les dissocier et de les comparer aux parcours autobus.

- Le "parcours" est une unité de mesure qui correspond au "voyage effectué" : un voyageur utilisant plusieurs lignes d'autobus est compté pour autant de parcours. Les parcours sur chaque ligne de Paris et de banlieue sont additionnés pour obtenir le nombre de parcours sur le réseau de surface.

3.1.2.3 Voyageurs - kilomètres

On considère aussi les "voyageurs-kilomètres" sur un réseau, c'est-à-dire le kilométrage parcouru par l'ensemble des clients sur ce sous-réseau.

Le nombre de "voyageurs-kilomètres" est calculé comme le produit du nombre de voyages effectués sur un réseau multiplié par le parcours moyen correspondant, i.e. la distance moyenne en kilomètres parcourue par un voyageur sur un réseau ou sous-réseau (estimée à l'aide de données recueillies par enquête pour un type de journée donnée).

Il existe 6 parcours moyens : "métro", "RER", "réseau ferré", "autobus de Paris", "autobus de banlieue", "réseau d'autobus". Pour les raisons évoquées plus haut concernant les "utilisations" ou les "parcours", le parcours moyen du réseau ferré (respectivement réseau d'autobus), n'est pas égal à la somme des parcours moyens métro et RER (respectivement lignes de Paris et lignes de banlieue).

3.2 Collecte des données brutes

La cellule Statistiques de trafic de la Mission Sécurité Ferroviaire du département RER transmet au département Commercial les données des sous-réseaux ferrés, corrigées des éventuelles défaillances du système informatique.

L'entité Enquêtes et Statistiques du Pôle commercial du département BUS communique au département Commercial les données brutes relatives aux différents sous-réseaux de surface, à l'exception du trafic mensuel d'Orlyval, communiqué par CGF (Contrôle de Gestion et Finances). Ces données sont corrigées des problèmes de comptage.

Examinons la manière dont sont collectées ces données et leur fiabilité.

3.2.1 Réseau ferré (métro + RER)

De façon générale, le trafic sur le réseau ferré est mesuré par le nombre de voyageurs qui franchissent les lignes de contrôle.

3.2.1.1 Les lignes de contrôle

- Il existe des lignes de contrôle d'accès direct aux quais dans toutes les stations de métro et pratiquement toutes les gares du RER.
- Lorsqu'il existe une correspondance entre une station de métro et une gare de RER, les lignes de correspondance sont situées dans des salles différentes.

Quelques cas particuliers d'absence de lignes de contrôle pour les gares de RER en correspondance peuvent être mis en avant :

- Sur les lignes A, B et D, le passage est libre entre les domaines RATP et SNCF : les voyageurs qui transitent entre les domaines SNCF et RATP ne passent pas par des lignes de contrôle et ne changent pas de véhicules :
 - Interconnexion à Nanterre Préfecture avec la ligne A ouest jusqu'à Poissy ou Cergy le Haut (portions SNCF).
 - Interconnexion à Gare du Nord pour la ligne B Nord jusqu'à Aéroport Charles-de-Gaulle 2 ou Mitry-Claye (portions SNCF).

- Il n'existe en outre pas de lignes de péage entre ces différentes lignes, à Gare du Nord (lignes B et D), à Gare de Lyon (lignes A et D), à Saint-Michel-Notre Dame (lignes A et B, mais aussi B et C), mais aussi entre les lignes de banlieue de la SNCF et la ligne A du RER à la gare Val de Fontenay.
- Les gares du sud de la ligne B au delà de la gare des Baconnets ne sont équipées que de valideurs, et non de lecteurs magnétiques.

La collecte des données de trafic sur le réseau ferré (et chaque sous-réseau) s'effectue par comptage des entrants et correspondants lorsqu'il existe des lignes de contrôle, et par estimation en son absence.

3.2.1.2 Les données brutes de trafic sont collectées aux lignes de contrôle.

Les appareils de contrôle lisent la piste magnétique des titres de transport et archivent les données par type de titre et par station dans les mémoires du CTK (Centre de Traitement du Contrôle) ou du SAK (Système d'Aquisition des Contrôles) pour Météor. La cellule Statistiques de trafic de la Mission Sécurité Ferroviaire du département RER corrige ces données lors d'éventuelles défaillances du système informatique (pannes des péages, du CTK ou des interfaces entre les péages et le CTK) ou de la mise hors tension des appareils de contrôle (lors de grèves par exemple, ou d'un flux massif exceptionnel de voyageurs après une manifestation sportive ou autre...). Une alarme est émise lorsque se produisent des incidents dans le système de traitement et de recueil des données de trafic (données manquantes ou anormales).

3.2.1.3 En l'absence de lignes de contrôle, les données brutes de trafic sont estimées.

Le trafic des gares du RER assurant l'interconnexion avec la SNCF, le trafic des gares du sud de la ligne B au delà des Baconnets, ou le trafic sur la portion RATP de la ligne D sont estimés à l'aide de coefficients correspondants à des pourcentages du trafic des autres gares de la ligne. Les pourcentages utilisés sont remis périodiquement à jour par enquête. Pour ces stations, le trafic par type de titres est obtenu en appliquant les pourcentage de ventilation observés aux lignes de contrôle pour les entrants directs.

3.2.1.4 Les données du réseau ferré sont des données fiables

Selon l'étude effectuée en 1990 par le CGS de l'Ecole des Mines, les trafics annuels du métro et du RER sont estimés avec une incertitude très faible : environ $\pm 0.03\%$ pour le métro et $\pm 0.08\%$ pour le RER. La légère sous-estimation des trafics sur le réseau ferré semble avoir un caractère systématique : s'il existe un biais, celui-ci n'enlève aucune pertinence à l'analyse des évolutions entre des périodes différentes.

La fiabilité des données sur le réseau ferré provient de la fermeture de ce réseau par les lignes de contrôle. La fraude n'est pas comprise dans ces données ; son taux a été évalué à 4.9% sur le métro et 3.6% sur le RER, en 1999. La seule incertitude résulte des trafics provenant de la SNCF, qui doivent être estimés. Il n'en est pas de même pour le réseau de surface, qui est un réseau ouvert.

3.2.2 Réseau de surface (autobus + sites propres)

Le trafic sur le réseau de surface est mesuré par un comptage grâce à des marches sensibles, à quelques exceptions près. Le nombre de parcours sur le réseau de surface est divisé en deux catégories :

- Le "trafic pointé" correspond au trafic billet. Il est mesuré par les validations dans les appareils oblitérateurs BC30 dans l'autobus.
- Le "trafic à vue" correspond à la différence entre le trafic total sur le réseau de surface et le trafic pointé⁹. Des enquêtes permettent de ventiler le trafic à vue par type de titres hors billets.

⁹Aujourd'hui, sur le réseau de surface, un ticket permet de réaliser un parcours. Avant septembre 1999, le nombre de billets à valider (appelé "terme") était fonction du nombre de sections parcourues. Un voyageur sur le réseau de surface compostait entre un (lignes urbaines sauf PC, et lignes de banlieue pour les trajets d'au plus trois sections) et six billets pour un parcours. Le trafic billet était calculé en divisant le nombre de titres validés par le nombre moyen (estimé par enquête) de termes par trajet.

3.2.2.1 Le trafic sur le réseau de surface est généralement mesuré à l'aide de "marches sensibles"

Le trafic des lignes d'autobus à plancher haut et du Trans-Val-de-Marne ("TVM") est calculé à partir d'un système automatique de comptage des voyageurs par des "marches sensibles" ou "tapis contact". Le relevé des compteurs Almex est corrigé d'éventuelles erreurs de saisie puis de différentes perturbations (piétinement, descente par la porte avant...) à l'aide de coefficients correcteurs propres à chaque ligne et remis à jour périodiquement grâce à des enquêtes.

3.2.2.2 Il existe quelques cas particuliers

- Sur les lignes des autobus à plancher bas (principalement des autobus de grand gabarit de type AGORA ou GX 317), les marches sensibles sont remplacées par des cellules photo-électriques.
- Sur les lignes de tramway T1 et T2, le trafic est estimé sur la base de recensements.
- Sur Orlyval, le trafic est issu du relevé des appareils de contrôle.

3.2.2.3 Des données relativement fiables

L'étude effectuée en 1990 par le CGS de l'Ecole des Mines a conclu que les trafics annuels sur les sous-réseaux du réseau de surface étaient connus avec une faible incertitude : $\pm 0.08\%$ pour l'ensemble des lignes d'autobus de Paris et $\pm 0.05\%$ pour l'ensemble des lignes d'autobus de banlieue. Sur données mensuelles, cette incertitude augmente, mais reste limitée à $\pm 0.3\%$ pour le sous-réseau de Paris, et $\pm 0.2\%$ pour le sous-réseau de banlieue.

L'imprécision est plus importante pour le trafic d'une ligne donnée ($\pm 0.6\%$ pour les données annuelles et $\pm 2\%$ en données mensuelles).

La moindre fiabilité des données brutes de trafic sur le réseau de surface par rapport au réseau ferré provient de la non-fermeture de ce réseau et de la nécessité de recourir à des coefficients corrigeant piétinements, descente par la porte avant... En outre, la mesure du réseau de surface incorpore la fraude, ce qui n'est pas le cas du trafic sur le réseau ferré.

3.2.2.4 Les résultats de la lutte contre la fraude sur le réseau routier

En 1999, le taux de fraude observé est égal à 16.1% sur le réseau de bus. Maîtriser cette fraude est devenu un enjeu majeur depuis la signature du contrat STIF - RATP, car la rémunération se base désormais sur le volume des ventes de titres, et non plus sur le trafic global. Une des conséquences est la mise en place progressive de la validation systématique à l'avant des bus, dans le cadre du projet "**Bus Attitude**".

En 1999, la ligne T1 (Bobigny - Saint Denis) transporte 85 000 voyageurs par jour, avec un taux de fraude réel de près de 29%. En octobre 2002, ce taux de fraude était tombé à 12.3% grâce aux effets du "**Nouveau Service Tram**" instauré en septembre 2001. Des équipes de quatre agents patrouillent toute la journée sur cette ligne, et vérifient à chaque porte si les usagers possèdent un titre de transport et, le cas échéant, les invitent à s'en procurer un. Lorsque la personne refuse de s'acquitter d'un billet, le contrôleur dresse un procès verbal. Cette opération a été étendue à la ligne T2 (La Défense - Issy-les-Moulineaux), et a permis de réduire le taux de fraude de près de 20% en juin 2000 à 9.8% deux années après. Sur la ligne TVM (Trans Val-de-Marne), qui devrait bientôt disposer de personnel à bord, le taux de fraude a pu être ramené d'environ 30% en juin 2000 à 20% grâce à la simple augmentation des contrôles de 1 à 3%.

Aujourd'hui, avec la généralisation du passe Navigo¹⁰ et la "**Tram'attitude**", la Validation Systématique à l'Entrée devrait permettre la poursuite de la régression de l'incivilité et de la fraude. Les voyageurs seront invités à présenter leur passe au contact de valideurs sans contacts, lors de l'accès au quai. La validation entraînera l'émission d'un signal sonore et lumineux. Les contrôles seront facilités dans la mesure où seuls les fraudeurs seront abordés.

Les données collectées par les réseaux sont transmises au département Commercial de la RATP. Ces données sont des données brutes, la seule correction effectuée par les réseaux ayant pour but de corriger des dysfonctionnements liés aux modes de comptage.

¹⁰Le passe Navigo est une carte à puce de format classique mais issue de la dernière génération. La technologie retenue est celle du standard CALYPSO développé en Europe et déjà choisi par 24 villes. Le passage des lecteurs est plus rapide qu'avec les cartes magnétiques puisqu'il suffit de présenter le passe en face d'un lecteur qui décode la carte à quelques centimètres : plus besoin d'extraire son coupon de l'étui!

Avant d'être utilisées et analysées, les séries brutes sont corrigées des perturbations, du calendrier et des effets de saisonnalité.

3.3 Correction des données par le calcul du jour ouvrable moyen sur chaque sous-réseau.

Les séries collectées sont des séries brutes, pour lesquelles les seules corrections apportées concernent les défaillances des systèmes de comptage.

Les données mensuelles disponibles concernent :

- le nombre d'entrants directs et de correspondants pour le réseau ferré
 - par sous-réseau : métro et RER,
 - à un niveau encore plus fin : par lignes, par stations de métro ou gares de RER,
 - par type de titres,
- le nombre de parcours pour le réseau de surface
 - par sous-réseau : autobus de Paris, autobus de banlieue et sites propres,
 - en distinguant les voyages pointés et les voyages à vue.

Les séries mensuelles décrivant les trafics par sous-réseaux sont corrigées des grèves et des effets calendaires grâce au calcul du trafic du jour ouvrable moyen. Ces calculs nécessitent l'utilisation de données journalières brutes de trafic sur les sous-réseaux.

3.3.1 Le calcul du trafic du jour ouvrable moyen

L'évolution du trafic entre deux mois s'explique en partie par l'existence de facteurs calendaires : certains mois sont plus courts que d'autres, comportent des jours fériés, un plus grand nombre de week-end... Cet impact du calendrier est un élément qui perturbe l'analyse mensuelle des variations de trafic. Pour éliminer ce facteur perturbateur, le trafic mensuel d'un jour ouvrable moyen est calculé en divisant le trafic mensuel brut par le coefficient calendaire associé.

Cette méthode permet en outre de calculer un trafic corrigé des événements ponctuels tels que les grèves de faible ampleur.

3.3. CORRECTION DES DONNÉES PAR LE CALCUL DU JOUR OUVRABLE MOYEN SUR CHAQUE SOUS-RÉSEAU.

Nous examinerons tout d'abord la méthode de calcul des coefficients calendaires et du trafic du jour ouvrable (premier paragraphe), et détaillerons un exemple précis (deuxième paragraphe).

3.3.1.1 Calcul des coefficients calendaires et trafic du jour ouvrable : méthodologie

- Chaque mois, à partir des données de trafic journalières corrigées des incidents techniques, sont calculées les moyennes des jours ouvrables, des samedis, et des dimanches ou jours fériés, en distinguant les périodes de vacances scolaires et hors vacances scolaires et en écartant les journées spéciales (journées de grèves ou journées dont le trafic a été modifié de façon importante par un évènement ponctuel exceptionnel).
- Le poids 1 est affecté à la moyenne des jours ouvrables hors vacances scolaires, des poids relatifs sont attribués aux samedis, dimanches, et jours ouvrables de vacances scolaires. Ces poids relatifs sont calculés comme le quotient de la moyenne de ce type de journée sur la moyenne du jour ouvrable hors vacances scolaires. La somme des poids de chaque journée donne le coefficient calendaire pour le réseau considéré.

Ainsi, le coefficient calendaire correspond à un équivalent en jours ouvrables hors vacances scolaires, qui reflète à la fois la composition du calendrier et le poids du trafic du week-end. Il existe des coefficients calendaires différents suivant les réseaux : le poids du week-end est par exemple plus important sur le métro que sur les bus de banlieue. Ces coefficients sont calculés " *a posteriori* ", à partir du trafic réalisé.

- Le trafic du jour ouvrable pour le mois considéré est alors obtenu par la formule :

$$\text{Trafic du jour ouvrable} = \frac{\text{Trafic mensuel corrigé des événements ponctuels}}{\text{Coefficient calendaire}}$$

Ce trafic correspond à la moyenne des trafics des jours ouvrables du mois considéré.

3.3.1.2 Exemple détaillé du calcul du coefficient calendaire et du trafic du jour ouvrable

Prenons l'exemple du mois de décembre 2001 sur le métro.

- Distinguons tout d'abord les jours de vacances scolaires des jours hors vacances scolaires.

La date légale des vacances en 2001 est comprise entre le lundi 24 décembre et le 2 janvier 2002 inclus. En pratique, il est fréquent que les gens partent en vacances dès le samedi. Pour cette raison, l'unité TVR (Tarification Vente Résultats) considère que les vacances scolaires commencent dès le samedi 22 et se terminent après le lundi 31.

- Pendant les vacances scolaires, plusieurs journées spéciales sont à mettre à part : le lundi 24, qui est une journée de Pont, le jour de Noël, ainsi que le lundi 31 décembre qui est à la fois une journée de pont et une journée spéciale dans la mesure où les transports ont été gratuits à partir de 17 heures (aucune donnée de trafic n'a été enregistrée sur le réseau ferré). Aucune journée de grève (RATP ou SNCF) n'est à écarter ce mois là pour la période de vacances scolaires.

3.3. CORRECTION DES DONNÉES PAR LE CALCUL DU JOUR OUVRABLE MOYEN SUR CHAQUE SOUS-RÉSEAU.

	date	voyages	
	samedi 1 décembre 2001	3 005 530	
	dimanche 2 décembre 2001	1 921 233	
	lundi 3 décembre 2001	4 171 057	
	mardi 4 décembre 2001	4 489 168	
	mercredi 5 décembre 2001	4 483 741	
	jeudi 6 décembre 2001	4 602 099	
	vendredi 7 décembre 2001	4 673 041	
	samedi 8 décembre 2001	3 344 243	
	dimanche 9 décembre 2001	2 144 281	
	lundi 10 décembre 2001	4 153 834	
	mardi 11 décembre 2001	4 461 808	
	mercredi 12 décembre 2001	4 493 132	
	jeudi 13 décembre 2001	4 526 885	
	vendredi 14 décembre 2001	4 556 607	
	samedi 15 décembre 2001	3 177 628	
	dimanche 16 décembre 2001	2 099 768	
	lundi 17 décembre 2001	4 086 929	
	mardi 18 décembre 2001	4 518 972	
	mercredi 19 décembre 2001	4 577 156	
	jeudi 20 décembre 2001	4 620 123	
	vendredi 21 décembre 2001	4 602 178	
	samedi 22 décembre 2001	2 832 840	} Week - end de vacances scolaires
	dimanche 23 décembre 2001	1 860 805	
	lundi 24 décembre 2001	2 435 069	} Jour de pont
	mardi 25 décembre 2001	1 069 521	
	mercredi 26 décembre 2001	3 080 403	} Vacances scolaires légales
	jeudi 27 décembre 2001	3 408 660	
	vendredi 28 décembre 2001	3 448 073	} Jour de pont
	samedi 29 décembre 2001	2 459 294	
	dimanche 30 décembre 2001	1 986 810	
	lundi 31 décembre 2001	1 832 448	

Traffics par jour observés sur le métro en décembre 2001 (RATP, relevés de comptage)

- La moyenne du trafic le lundi est calculée en divisant la somme des trafics effectués les lundis normaux de vacances scolaires par le nombre de lundis de ce type. Ce calcul est effectué pour chaque type de jour de vacances scolaires. En décembre 2001, le nombre de lundis de vacances scolaires est nul, de même que le nombre de mardis (Noël est écarté), il existe un mercredi (le 26), un jeudi (le 27), un vendredi (le 28), deux samedis (les 22 et 29) et deux dimanches (les 23 et 30) de vacances scolaires.
- Le même travail est effectué pour les journées hors vacances scolaires, en écartant la journée de grève du lundi 19.

CHAPITRE 3. LES DONNÉES RELATIVES À LA RATP

- Une estimation du "trafic du jour ouvrable hors vacances scolaires" est obtenue en calculant la moyenne du trafic des lundis aux vendredis moyens hors vacances scolaires.

	Trafic	Nombre de jours	Moyenne du jour (moy)	- X JO	Poids des jours (moy/X)	Total
<u>Vacances Scolaires</u>						
lundi	0	0				
mardi	0	0				
mercredi	3 080 403	1	3 080 403		0,6887	0,6887
jeudi	3 408 660	1	3 408 660		0,7621	0,7621
vendredi	3 448 073	1	3 448 073		0,7709	0,7709
samedi	5 292 134	2	2 646 067		0,5916	1,1832
dimanche	3 847 615	2	1 923 808		0,4301	0,8602
grève		0				
Noel : 25 dec 2001	1 069 521	1	1 069 521		0,2391	0,2391
jour spécial : 31 dec 2001	2 582 448	1	2 582 448		0,5774	0,5774
pont : 24 dec 2001	2 435 069	1	2 435 069		0,5444	0,5444
<u>Hors Vacances Scolaires</u>						
lundi	8 324 891	2	4 162 446	Trafic du jour ouvrable hors vacances scolaires	0,9306	2,7918
mardi	13 469 948	3	4 489 983		1,0038	3,0115
mercredi	13 554 029	3	4 518 010		1,0101	3,0303
jeudi	13 749 107	3	4 583 036		1,0246	3,0739
vendredi	13 831 826	3	4 610 609		1,0308	3,0924
				4 472 816		
samedi	9 527 401	3	3 175 800		0,7100	2,1301
dimanche	6 165 282	3	2 055 094		0,4595	1,3784
grève (lundi 19 dec 2001)	4 086 929	1				
						Coefficient Calendaire
Total	107 873 336	31				24,1344

Coefficient calendaire "a posteriori" et du trafic du JO sur le métro pour décembre 2001

- Les poids associés à chaque jour de la semaine sont alors calculés, en divisant le trafic moyen associé à ce jour par la moyenne du trafic du jour ouvrable hors vacances scolaires.
- Les poids relatifs de chaque type de journée sont sommés, en réintégrant les jours de grève avec un poids adéquat. Dans notre exemple, il y a eu une journée de grève, le lundi 19 décembre, qui a été exclue du calcul du jour ouvrable moyen. En l'absence d'incident ponctuel, le trafic ce jour là aurait été identique aux autres lundis hors vacances scolaires du mois, avec un poids relatif de 0.9306. La somme des poids relatifs de chaque journée

3.3. CORRECTION DES DONNÉES PAR LE CALCUL DU JOUR OUVRABLE MOYEN SUR CHAQUE SOUS-RÉSEAU.

donne le coefficient calendaire du mois considéré.

- Le trafic du jour ouvrable pris en compte est égal au trafic mensuel corrigé des événements ponctuels (correction apportée par le réseau ferré) divisé par le coefficient calendaire.

3.3.2 Jour ouvrable moyen et désaisonnalisation

La série des jours ouvrables est désaisonnalisée à l'aide du logiciel STATISTICA, pour permettre des comparaisons mois par mois et non d'une année sur l'autre pour le même mois : par exemple, le trafic brut moyen d'un jour ouvrable d'été est par exemple toujours très fortement inférieur à celui d'hiver.

L'utilisation de moyennes mobiles permet de calculer un ensemble de 12 coefficients de somme égale à 12, compris entre 0.6 environ pour le mois d'août à 1.1 environ pour les mois d'hiver. Ces coefficients servent à obtenir les courbes du trafic du jour ouvrable moyen.

Les évolutions constatées entre 2 mois successifs ne dépendent ni de la composition calendaire, ni de l'évolution du poids des week-ends, ni des effets saisonniers récurrents.

3.3.3 Utilité pour la prévision

Le trafic du jour ouvrable est utilisé lors du calcul des prévisions, par exemple pour l'établissement du budget :

- En prolongeant la tendance observée et en prenant en compte les effets probables d'éventuelles modifications de réseau, l'évolution du trafic du jour ouvrable est estimée.
- Les coefficients calendaires futurs, dit coefficients calendaires "*a priori*", sont calculés à partir des seules modifications de calendrier, en supposant que le poids des week-ends est identique à l'année précédente. L'évolution des poids des samedis et dimanches est connue seulement "*a posteriori*", mais il aurait été possible de prendre une autre hypothèse sur leur évolution *a priori*, par exemple en tenant compte du fait que pendant la seconde moitié des années 1990, le trafic des week-ends a crû plus rapidement que celui du jour ouvrable.

– Enfin, les trafics mensuels prévus sont calculés de la façon suivante :

$$\begin{aligned} \text{trafic prévu le mois } m &= \text{trafic du jour ouvrable prévu} \\ &\quad * \text{coefficient calendaire "a priori" pour le mois } m \\ &\quad * \text{coefficient saisonnier associé au mois } m \end{aligned}$$

3.3.4 Evolution des poids des samedis et dimanches/jours fériés

La différence entre le coefficient calendaire "*a priori*" pour le mois courant et le coefficient calendaire "*a posteriori*" du même mois de l'année précédente provient uniquement de la modification du poids relatifs des samedis, dimanches et jours fériés, jours de vacances scolaires. En comparant ces deux coefficients, il est possible de connaître quel est l'impact du seul changement des mobilités des jours non ouvrables ou hors période scolaire.

Les séries de samedis et dimanches/jours fériés sont obtenues en multipliant les séries historiques de trafic du jour ouvrable sur le réseau de surface par les séries de coefficients calendaires des samedis et dimanches/ jours fériés.

Pour les données relatives au réseau ferré (métro et RER), d'autres séries historiques sont disponibles : les séries de trafic du jour ouvrable, des samedis et dimanches/jours fériés issues des documents "MF 14" édités par le département RER, calculées sans prendre en compte l'évolution du poids des week-ends.

3.3.4.1 Exemple détaillé du calcul du coefficient calendaire et du trafic du jour ouvrable (suite)

Les calculs précédemment effectués donnent le coefficient calendaire du mois de décembre 2001 sur le métro "*a posteriori*" : il dépend des données de trafic observées en décembre 2001, du poids relatifs des différentes journées ce mois là. De la même façon, le coefficient calendaire "*a posteriori*" sur le métro de décembre 2000 a été calculé l'année d'avant.

Le coefficient calendaire "*a posteriori*" est plus élevé que le coefficient "*a priori*" (24.1344 contre 24.0735). Cette différence signifie que le poids relatifs des week-ends ou des journées de vacances scolaires a été plus important que l'année précédente. C'est effectivement le cas du poids du dimanche pendant les vacances scolaires (0.4301 en 2001 contre 0.3503 en 2000).

Cet effet peut être expliqué en partie par un report des achats en fin d'années, suite notamment à "l'effet du 11 septembre" : en 2000 ainsi que les années précédentes, les achats de Noël avaient lieu en moyenne entre le 8 et le 10 décembre, en 2001, ces achats ont été retardés jusqu'au 19 - 20 décembre.

Entre décembre 2000 et décembre 2001, le trafic mensuel a diminué de -5.29% en données brutes. La composition du calendrier était plus défavorable en 2001, et -4.02% de cette diminution sont expliqués par la seule modification du calendrier entre les deux années (évolution entre le coefficient 2001 "*a priori*" et le coefficient 2000 "*a posteriori*").

L'accroissement du poids relatifs des journées non ouvrables ou des vacances scolaires a compensé en partie cet effet défavorable, de $+0.25\%$ (évolution entre le coefficient 2001 "*a priori*" et le coefficient 2001 "*a posteriori*"). Cet effet est sans doute en partie imputable à l'effet des 35 heures, qui ont diminué l'importance du trafic des journées hors vacances scolaires.

Pour les données corrigées des variations calendaires et des événements ponctuels, c'est-à-dire la variation du trafic moyen du jour ouvrable, la diminution du trafic n'est plus que de -1.51% .

La décalendarisation est ainsi un moyen d'expliquer les variations de trafic induites par les modifications calendaires, les grèves et les poids relatifs des différents types de journées.

3.4 Trafics étudiés et évolutions

Le travail de décalendarisation et de désaisonnalisation a été appliqué aux séries brutes de trafic sur les sous-réseaux ferrés (données fournies par la cellule Statistiques de trafic de la Mission Sécurité Ferroviaire du département RER) et aux séries de trafic sur certains sous-réseaux de surface (données obtenues auprès de l'entité Enquêtes et Statistiques du Pôle commercial du département BUS).

Les séries suivantes sont obtenues :

- le trafic du jour ouvrable sur le métro ("*jomsa*"),
- le trafic du jour ouvrable sur le RER ("*jorersa*"),
- le trafic du jour ouvrable sur le bus de Paris ("*joparsa*"),
- le trafic du jour ouvrable sur le bus de banlieue ("*jobansa*").

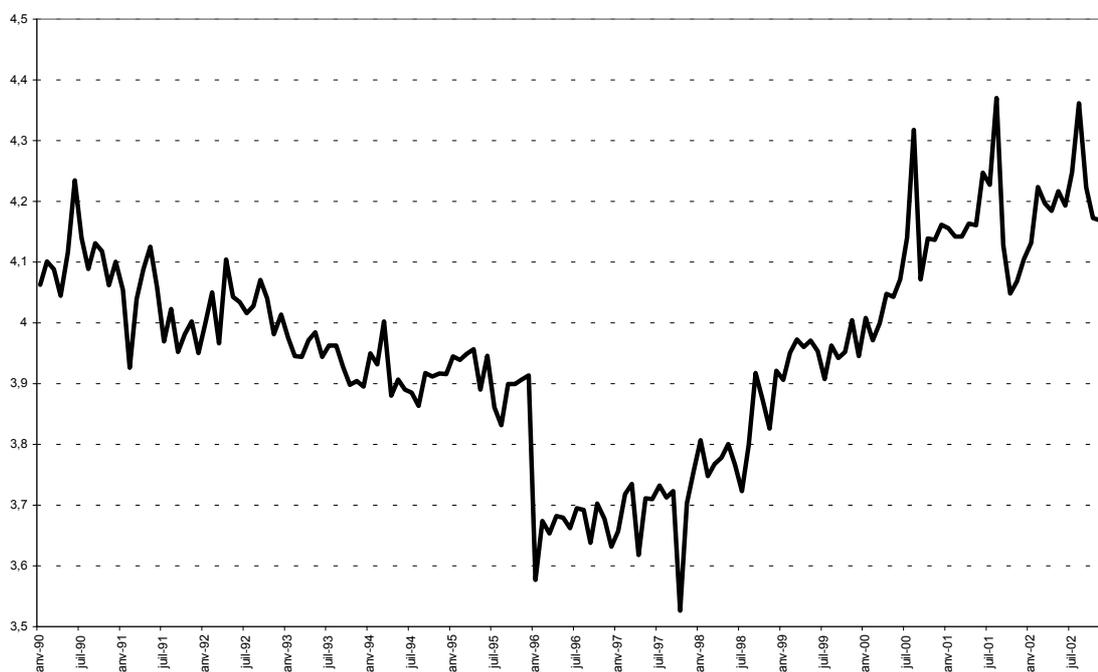
Ces séries sont corrigées des variations calendaires et des perturbations ("*CVCP*") et des variations saisonnières ("*CVS*")¹¹.

Les trafics des samedis et dimanches/jours fériés sur les sous-réseaux ferrés sont issus des documents "*MF14*" :

- trafic du samedi sur le métro ("*smsa*"),
- trafic du dimanche/jour férié sur le métro ("*dfmsa*"),
- trafic du samedi sur le RER ("*srersa*"),
- trafic du dimanche/jour férié sur le RER ("*dfrersa*").

¹¹En langue anglaise, une série corrigée des variations saisonnières (ou CVS) est dite "seasonally adjusted", ou SA. Le nom de série de trafic RATP "*jomsa*" signifie "trafic du Jour Ouvrable sur le Métro, Seasonally Adjusted".

3.4.1 Le trafic du jour ouvrable sur le métro



jomsa : Trafic du jour ouvrable sur le métro (données CVCP - CVS) en millions de voyages

Entre 1990 et 2002, le trafic du jour ouvrable a augmenté de +2.5% (taux d'accroissement annuel moyen : +0.2%), pour atteindre une moyenne de 4.21 millions de voyages (ou "utilisations") en 2002. Sur l'ensemble de la période, plus de 15.5 milliards de voyages ont été effectués sur le métro, soit une moyenne de 99.6 millions de voyages par mois, pour un trafic relatif au jour ouvrable moyen de 3.96 millions de voyages.

En 1996, le trafic annuel sur le métro connaît sa plus forte baisse (−6.1% par rapport à l'année précédente, pour une baisse du trafic du jour ouvrable de −6.5% en valeur annuelle moyenne). Le trafic du jour ouvrable atteint ses valeurs minimales : sur l'année, il s'établit en moyenne à 3.66 millions de voyages. Les fortes grèves de l'année 1995, et leur rémanence les mois suivants, expliquent cette forte chute. Elles ont eu lieu à la fin d'une sous-période de faible baisse tendancielle du trafic : entre 1990 et 1995, le trafic du jour ouvrable diminuait en moyenne de −0.8% par an. Cette baisse de la fréquentation du métro, due à la diminution des emplois dans Paris (délocalisation des entreprises de Paris vers la banlieue), a été à peine compensée par la

progression du trafic dans les gares de banlieue du RER ou sur les lignes d'autobus de banlieue. Un autre élément explicatif de la baisse du trafic du jour ouvrable sur le métro résulte de la concurrence des modes en site propre comme le tramway au nord - est ou le TVM au sud.

L'accroissement du trafic du jour ouvrable entre 1990 et 2002 est imputable à la forte reprise du trafic sur le métro en 1997. Entre 1997 et 2002, le trafic du jour ouvrable s'est accru de +14.0%, soit un accroissement annuel moyen d'environ +2.2%. Cette augmentation s'explique notamment par la baisse des loyers de bureaux, qui a entraîné le retour d'un certain nombre d'entreprises vers Paris. La plus forte croissance a été observée durant l'année 1999 (+3.7% par rapport à l'année précédente). Elle s'explique par la baisse du chômage, la poursuite de la croissance économique, mais aussi par l'impact favorable sur l'image des transports en commun de l'ouverture de Météor. En 2001, le trafic du jour ouvrable a retrouvé son niveau de début de période, et a continué à croître jusqu'en 2002.

Les pointes observées en fin de période correspondent aux mois d'été : ces dernières années, le poids du trafic observé en juillet et en août s'est accru du fait du morcellement croissant des vacances. Ce phénomène est encore trop récent pour que la méthode de désaisonnalisation en écarte les effets.

La rémanence des grèves de l'hiver 1995, ainsi que le fort rattrapage observé depuis, amènent à penser que la série est intégrée, i.e. conserve en mémoire chacun des chocs passés.

Les **tests de racine unitaire** ont été effectués sur la série désaisonnalisée log-linéarisée. La transformation des séries par le logarithme népérien est une transformation courante dans les analyses économétriques, pour au moins deux raisons :

- elle réduit l'écart-type des données et l'impact des données aberrantes,
- elle permet en outre d'interpréter les coefficients des modèles de trafic en fonction de variables économiques comme des élasticités (i.e. des pourcentages de variation de trafic induits par une variation de 1% d'une variable économique).

Deux types de tests sont envisagés :

1. Le test Augmenté de Dickey - Fuller [45], noté test "ADF". Ce test a pour hypothèse nulle l'existence d'une racine unitaire.

Si cette statistique est inférieure à la valeur critique associée au niveau du test choisi, l'hypothèse nulle de racine unitaire est rejetée.

2. Le test de Kwiatkowski, Phillips, Schmidt et Shin de 1992 [102], noté test "KPSS".

L'hypothèse nulle de stationnarité, pour une largeur de fenêtre, est rejetée si la statistique de test est supérieure à la valeur critique au seuil choisi.

Les tests de racine unitaire ont été effectués :

- sur la variable en niveau pour vérifier la présence d'une racine unitaire,
- sur la variable en différence pour préciser que la série est intégrée à l'ordre 1, i.e. qu'il n'existe qu'une seule racine unitaire.

$\ln(jomsa)$	Variable en niveau	Variable en différence
TEST ADF		
valeur de ρ	0.9749	-1.0740
t_{ADF}	-0.7940	-8.6430
VC_{ADF} à 5%	-2.9124	-2.9124
TEST KPSS		
Largeur de fenêtre	Statistique	
3	0.9150	0.2270
6	0.5415	0.2979
9	0.3910	0.3306
VC_{KPSS} à 5%	0.4572	0.4572
Conclusion	$I(1)$	$I(0)$

La série de trafic du jour ouvrable sur le métro semble intégrée à l'ordre 1.

3.4.2 Le trafic du jour ouvrable sur le RER



jorersa : Trafic du jour ouvrable sur le RER (données CVCP - CVS) en millions de voyages

Entre 1990 et 2002, près de 4.9 milliards de voyages ont été effectués sur le RER pour une moyenne mensuelle de 31.4 millions de voyages ou "utilisations". Le trafic du jour ouvrable moyen s'élève à 1.30 million de voyages sur la période considérée.

Le trafic du jour ouvrable observé en 2002 est supérieur de +11.6% à celui de 1990, ce qui correspond à un taux d'accroissement annuel moyen de +0.8%, quatre fois supérieur à celui du métro. Cette forte augmentation n'a toutefois pas été linéaire : le trafic était relativement stable entre 1990 et 1995 (taux d'accroissement annuel moyen de +0.1%), alors que le trafic du jour ouvrable sur le métro diminuait. Cette stagnation s'explique par la poursuite de l'accroissement du trafic RER dans les gares de prolongement de Chessy et dans les zones résidentielles (branches de Marne-la-Vallée et de Boissy-Saint-Léger sur le RER A, ainsi que la partie du RER B située entre Cité Universitaire et Bagneux). Cet accroissement a compensé les pertes de trafic significatives dans les quartiers d'affaire (comme Auber) et les gares en correspondance avec la SNCF à l'exception de Saint-Michel-Notre-Dame.

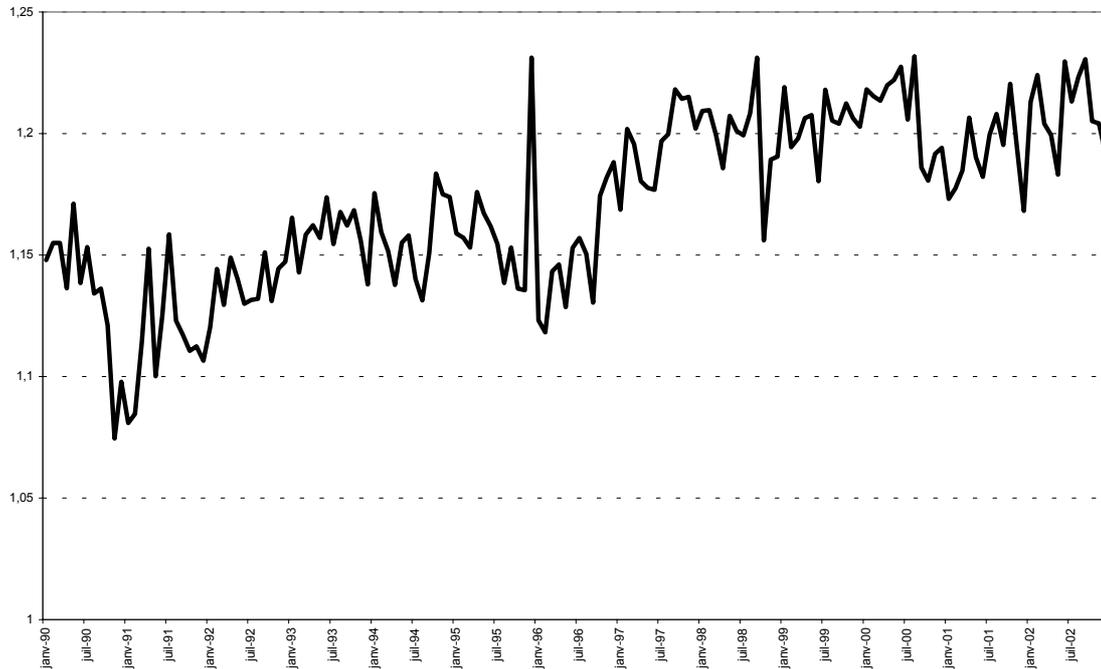
Le trafic du jour ouvrable sur le RER a subi une forte chute à la suite des grèves de novembre - décembre 1995 (baisse de -3.0% du trafic du jour ouvrable sur le RER entre 1995 et 1996, en moyenne annuelle). Entre 1997 et 2002, le trafic du jour ouvrable sur le RER a augmenté de $+14.2\%$, soit $+2.2\%$ en moyenne annuelle, comme sur le métro. Le retour des entreprises à Paris suite à la baisse des loyers de l'immobilier de bureaux explique en partie la reprise du trafic. Le retour des touristes et l'essor de la consommation et des loisirs est un autre facteur explicatif. L'année 1999 correspond à l'année de plus forte croissance du jour ouvrable sur le RER ($+4.7\%$) mais aussi à une année de renforcement de la croissance économique et de forte baisse du taux de chômage. Cette croissance semble se tasser : en 2002, le trafic du jour ouvrable moyen était inférieur de -2.0% à celui de l'année précédente. Les pointes observées en fin de période correspondent aux mois d'été dont le poids s'est accru, mais de façon trop récente pour que les méthodes de désaisonnalisation en tiennent compte.

La forte croissance depuis 1997 amène à tester le degré d'intégration de cette série du trafic du jour ouvrable sur le RER.

$\ln(jorersa)$	Variable en niveau	Variable en différence
TEST ADF		
valeur de ρ	0.9776	-0.7299
t_{ADF}	-0.9228	-9.8596
VC_{ADF} à 5%	-2.9124	-2.9124
TEST KPSS		
Largeur de fenêtre	Statistique	
3	2.4342	0.0739
6	1.4302	0.0848
9	1.0280	0.0967
VC_{KPSS} à 5%	0.4572	0.4572
Conclusion	$I(1)$	$I(0)$

La série du trafic du jour ouvrable sur le RER est intégrée à l'ordre 1.

3.4.3 Le trafic du jour ouvrable sur le réseau des autobus de Paris



joparsa : Trafic du jour ouvrable sur les bus de Paris (données CVCP - CVS) en millions de voyages

Entre 1990 et 2002, plus de 4,4 milliards de parcours ont été réalisés sur le sous-réseau des autobus de Paris, soit une moyenne de 28.6 millions de parcours par mois, et 1.17 million par jour ouvrable sur l'ensemble de la période.

En 2002, un jour ouvrable, 1.21 million de voyage ont été effectués sur les autobus de Paris, contre 1.13 en 1990. Cette variation de trafic de +6.6% correspond à un accroissement annuel moyen de l'ordre de +0.5%, mais reflète une évolution différenciée par sous-périodes. Entre 1990 et 1995, le trafic sur le réseau des bus de banlieue connaît une faible croissance de l'ordre de +0.4% par an, et augmente de +2.5% sur l'ensemble de la sous-période. Cette évolution contraste avec la baisse constatée sur le métro et la stagnation sur le RER, et est imputable à l'accroissement de trafic sur le PC ("Petite Ceinture" qui fait le tour de Paris).

Les grèves de décembre 1995 ont eu un effet rémanent plus faible sur les bus de Paris que sur le réseau ferré : en 1996, le trafic du jour ouvrable sur le bus de Paris diminue seulement de -1.1% par rapport à l'année précédente, ce qui limite l'effet de rattrapage les années d'après.

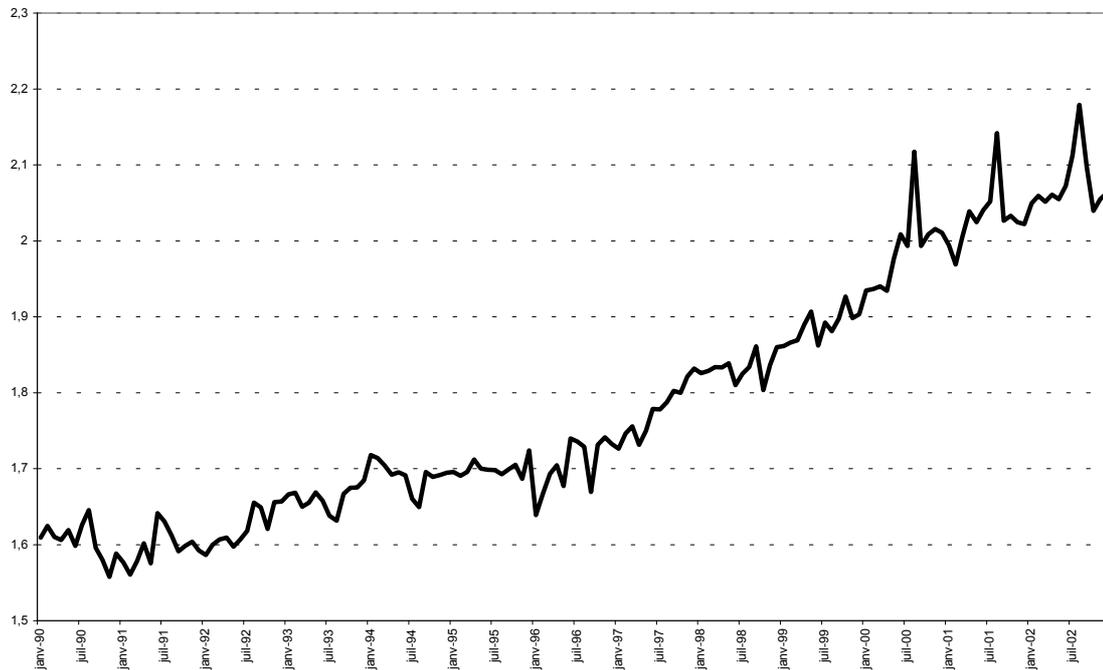
La croissance à partir de 1997 est plus faible que sur le métro ou le RER : entre 1997 et 2002, le trafic du jour ouvrable sur le bus de Paris a augmenté de +1.1% contre des taux supérieurs à +14% sur le réseau ferré, soit un accroissement annuel moyen de +0.2%. L'année de plus forte croissance correspond à l'année 1997 (+4.0%) et est plus précoce que sur le réseau ferré. Cette reprise plus rapide du trafic du jour ouvrable sur le réseau des bus de Paris s'explique par le report d'une partie de la clientèle du réseau ferré, dont le trafic était perturbé par les alertes au colis piégé. Les autobus constituaient, pour une partie de la clientèle, un mode de transport plus sécurisé que le réseau ferré, sur lequel avait eu lieu les attentats de Saint-Michel ou de Port Royal. En outre, une réforme tarifaire a été mise en place à partir du 1^{er} août 1995 : pour les lignes situées à l'intérieur du périphérique, voire desservant la proche banlieue, un seul billet est nécessaire pour parcourir toute la ligne, contre deux billets auparavant pour les parcours d'au moins 3 sections.

Les résultats des tests d'intégration sont résumés dans le tableau suivant.

$\ln(joparsa)$	Variable en niveau	Variable en différence
TEST ADF		
valeur de ρ	0.9381	-1.4906
t_{ADF}	-1.2961	7.3300
VC _{ADF} à 5%	-2.9124	-2.9124
TEST KPSS		
Largeur de fenêtre	Statistique	
3	3.1705	0.0355
6	1.9006	0.0513
9	1.3802	0.0611
VC _{KPSS} à 5%	0.4572	0.4572
Conclusion	$I(1)$	$I(0)$

La série du trafic du jour ouvrable sur le réseau des autobus de Paris est intégrée à l'ordre 1.

3.4.4 Le trafic du jour ouvrable sur le réseau des autobus de banlieue



jobansa : Trafic du jour ouvrable sur les bus de banlieue (données CVCP - CVS) en millions de voyages

Plus de 6.8 milliards de voyages ont été effectués entre 1990 et 2002 sur le réseau des autobus de banlieue, soit une moyenne de 43.7 millions de voyages (ou "parcours") par mois, et 1.8 million de voyages un jour ouvrable moyen.

Le trafic du jour ouvrable sur les bus de banlieue est passé de 1.60 million de voyages en 1990 à 2.07 millions de voyages en 2002, soit une variation de +29.1%. Le taux d'accroissement annuel moyen s'élève à +2,0%, sans commune mesure avec ceux du métro, du RER et du réseau des autobus de Paris. Un élément explicatif réside dans la mise en service du tramway, comptabilisé avec le réseau des autobus de banlieue, mais aussi de l'amélioration constante du service offert sur ce sous-réseau.

Entre 1990 et 1995, le trafic du jour ouvrable sur le réseau des bus de banlieue a connu un accroissement annuel moyen égal à +1.0%.

La forte augmentation du trafic du jour ouvrable sur les bus de banlieue est essentiellement postérieure à 1997, avec un taux d'accroissement annuel moyen de +2.6% entre 1997 et 2002. La réforme tarifaire du 1^{er} août 1995 constitue un facteur de croissance du trafic sur le réseau des

autobus de banlieue : un seul billet est nécessaire pour les trois premières sections parcourues (contre deux sections auparavant), deux billets suffisent pour parcourir quatre sections (contre au moins trois sections auparavant).

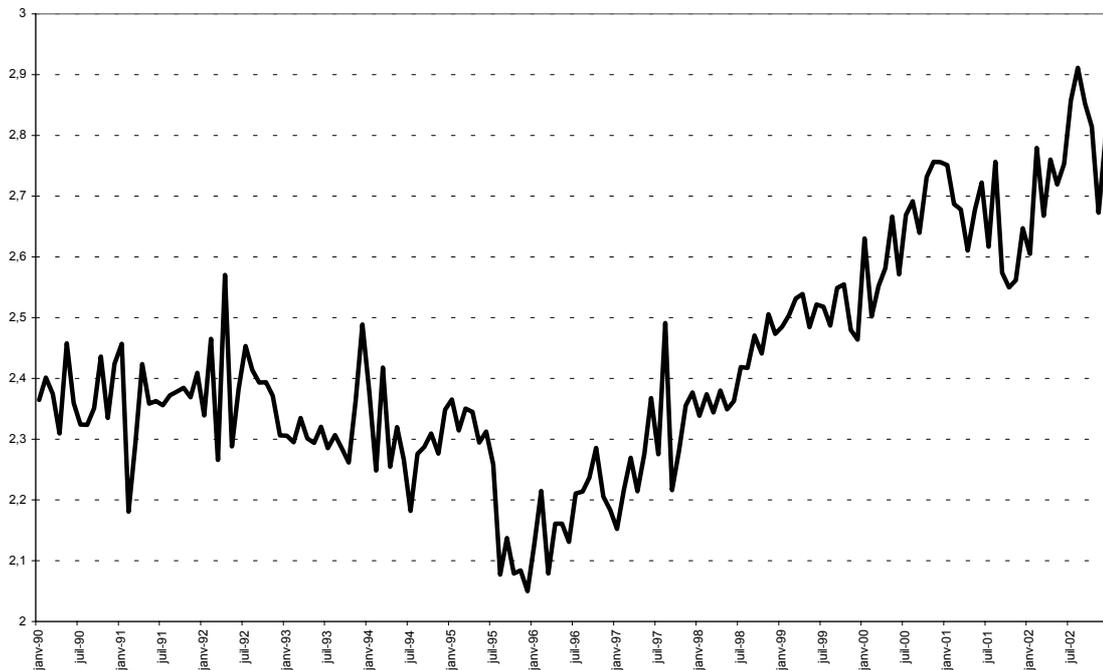
En 2003, le trafic d'un jour ouvrable sur le réseau des autobus de banlieue a augmenté de +1.3% par rapport à 2002, malgré un léger fléchissement durant l'été à mettre en lien avec la canicule.

Le phénomène d'accroissement du trafic au cours des mois d'été depuis quelques années se retrouve sur le réseau des autobus de banlieue. Cette similitude avec le réseau ferré s'explique par le fait que de nombreux autobus de banlieue ont pour principal objectif le rabattement sur les stations de métro et gares de RER.

L'accroissement important du trafic sur le réseau des autobus de banlieue amène à soupçonner l'existence d'une racine unitaire. Le caractère intégré à l'ordre 1 est confirmé par les tests de racine unitaire : les tests sur la variable en niveau concluent à la présence d'une racine unitaire, racine unitaire qui disparaît lorsque la série en différence première est testée.

$\ln(jobansa)$	Variable en niveau	Variable en différence
TEST ADF		
valeur de ρ	1.0048	-0.9807
t_{ADF}	0.3541	-7.6987
VC_{ADF} à 5%	-2.9124	-2.9124
TEST KPSS		
Largeur de fenêtre	Statistique	
3	3.7859	0.1377
6	2.2072	0.1716
9	1.5762	0.2279
VC_{KPSS} à 5%	0.4572	0.4572
Conclusion	$I(1)$	$I(0)$

3.4.5 Le trafic du samedi sur le métro



msa : Trafic du samedi sur le métro (données CVCP - CVS) en millions de voyages

Entre 1990 et 2002, en moyenne, 2.42 millions de voyages sont effectués un samedi sur le métro, avec un taux d'accroissement annuel moyen égal à +1.2%. Sur l'ensemble de la période, l'augmentation de +16.5% du trafic du samedi moyen ne s'est pas réalisée de façon linéaire. Entre 1990 et 1995, ce trafic a diminué en moyenne de -1.1% par an, avec une baisse marquée en 1995 (-3.3%). Cette diminution importante s'explique par la diminution des déplacements non obligés (achats et loisirs) dans un contexte économique morose, où le risque d'attentat et les perturbations de service suite à des alertes au colis piégé sont importants.

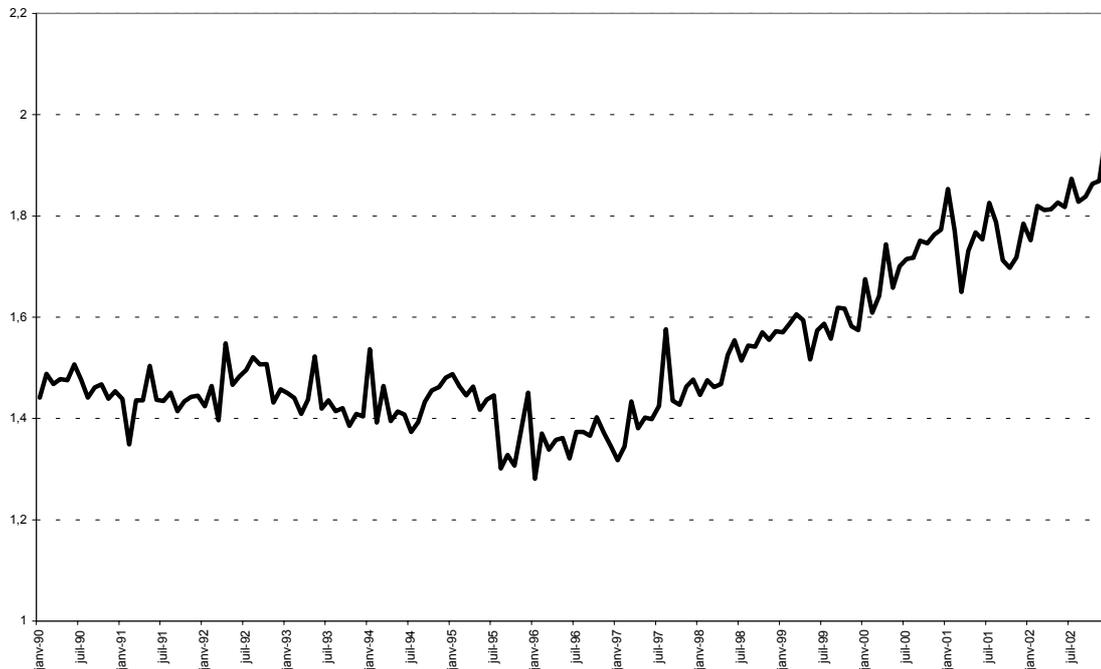
Le trafic du samedi sur le métro connaît une très forte augmentation après 1997 : entre 1997 et 2002, il a augmenté de +20.5%, ce qui correspond à un taux d'accroissement annuel moyen de +3.2%. Ce taux est supérieur d'un point au taux d'accroissement relevé sur la même période pour le trafic du jour ouvrable sur le métro, et s'explique par la reprise de la consommation et des dépenses de loisirs d'une part, et par la création du titre Imagine R d'autre part. Ce titre permet en effet aux scolaires et étudiants abonnés d'effectuer un nombre illimité de voyages le week-end sur l'ensemble du réseau RATP, sans limitation de zone.

La forte hausse du trafic du samedi sur le métro peut s'expliquer par la présence d'une racine unitaire dans la série.

L'intégration à l'ordre 1 est confirmée par les tests de racine unitaire synthétisés dans le tableau suivant.

$\ln(smsa)$	Variable en niveau	Variable en différence
TEST ADF		
valeur de ρ	0.9795	-1.0154
t_{ADF}	-0.5738	-7.3291
VC_{ADF} à 5%	-2.9124	-2.9124
TEST KPSS		
Largeur de fenêtre	Statistique	
3	2.1423	0.1239
6	1.2765	0.1714
9	0.9286	0.2032
VC_{KPSS} à 5%	0.4572	0.4572
Conclusion	$I(1)$	$I(0)$

3.4.6 Le trafic du dimanche/jour férié sur le métro



dfmsa : Trafic du dimanche / jour férié sur le métro (données CVCP - CVS) en millions de voyages

En moyenne, sur la période 1990 - 2002, 1,53 million de voyages sont réalisés au cours d'un dimanche ou jour férié moyen. Ce type de trafic a augmenté de +25.3% entre 1990 et 2002, soit un taux d'accroissement annuel moyen égal à +1.8%, alors que le trafic du jour ouvrable sur le même sous-réseau n'a bénéficié que d'une faible croissance (+0.2% par an).

Entre 1990 et 1995, le trafic sur le métro a diminué de -0.6% par an le dimanche/jour férié, de façon moins prononcée que le jour ouvrable (-0.8%) ou le samedi (-1.1%) du fait d'un impact moindre de la diminution de la consommation ou des délocalisations des entreprises. En 1996, le trafic du dimanche (-4.0%) subit sa plus forte baisse sur la période considérée : les déplacements pour les motifs non obligés ont régressé fortement dans un climat d'insécurité.

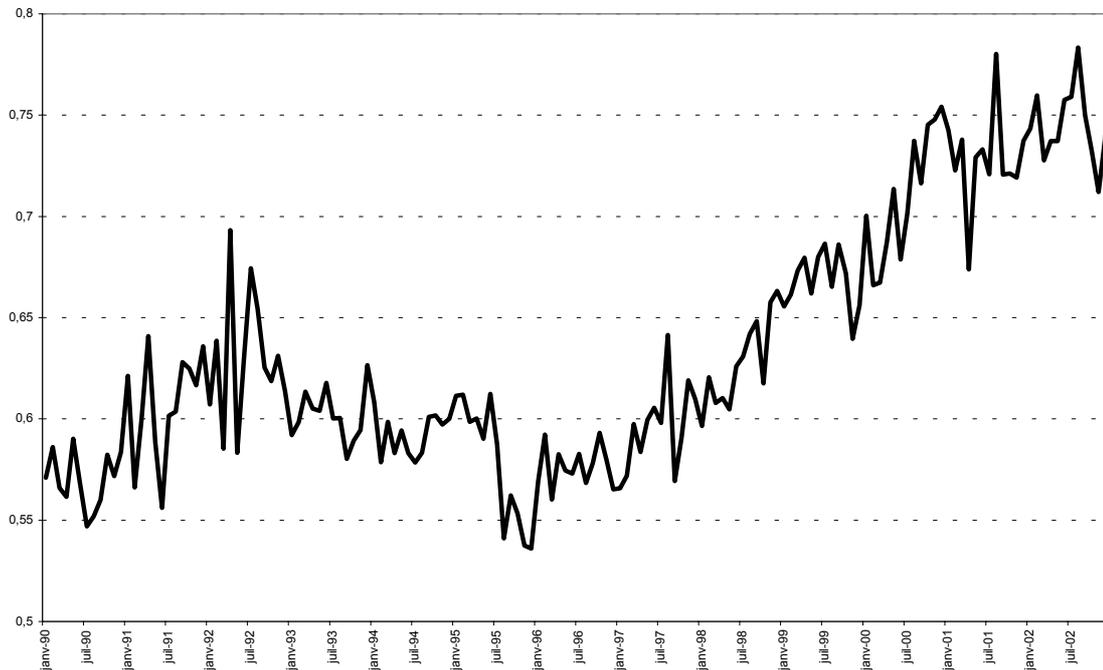
A partir de 1997, le trafic repart à la hausse, et augmente en moyenne de +4.4% par an jusqu'en 2002, soit une augmentation de +29.4% sur la sous-période considérée. La reprise de la consommation et des dépenses de loisirs, le retour des touristes, ainsi que la création du titre Imagine R à l'automne 1998 expliquent ce fort accroissement. En 2002, le trafic du dimanche ou

jour férié s'élève en moyenne à 1.84 million de voyages.

Le fort taux de croissance observé depuis 1997 pour le trafic du dimanche ou jour férié sur le métro résulte de la présence d'une racine unitaire dans la série, comme le montre le tableau synthétique suivant.

$\ln(dfmsa)$	Variable en niveau	Variable en différence
TEST ADF		
valeur de ρ	1.0279	-1.3868
t_{ADF}	0.9934	-8.993
VC_{ADF} à 5%	-2.9124	-2.9124
TEST KPSS		
Largeur de fenêtre	Statistique	
3	2.6016	0.2492
6	1.5359	0.3387
9	1.1062	0.4055
VC_{KPSS} à 5%	0.4572	0.4572
Conclusion	$I(1)$	$I(0)$

3.4.7 Le trafic du samedi sur le RER



srersa : Trafic du samedi sur le RER (données CVCP - CVS) en millions de voyages

Sur la période 1990 - 2002, le trafic du samedi moyen sur le RER s'élève à 0.6 million de voyages. Ce trafic a augmenté de +30.3% entre 1990 et 2002, soit un accroissement annuel moyen égal à +2.1%, supérieur à celui observé sur le métro pour le même type de journée.

Alors que le trafic du samedi a subi une baisse de -6.4% entre 1990 et 1995 sur le métro, celui-ci a augmenté de +1.5% sur le sous-réseau du RER (taux d'accroissement annuel moyen : +0.2%). Cette relative stagnation correspond à une évolution similaire à celle observée un jour ouvrable sur le RER. La baisse des déplacements pour les loisirs et achats est en partie compensée par les augmentations de trafic induites par le prolongement vers l'Est du RER.

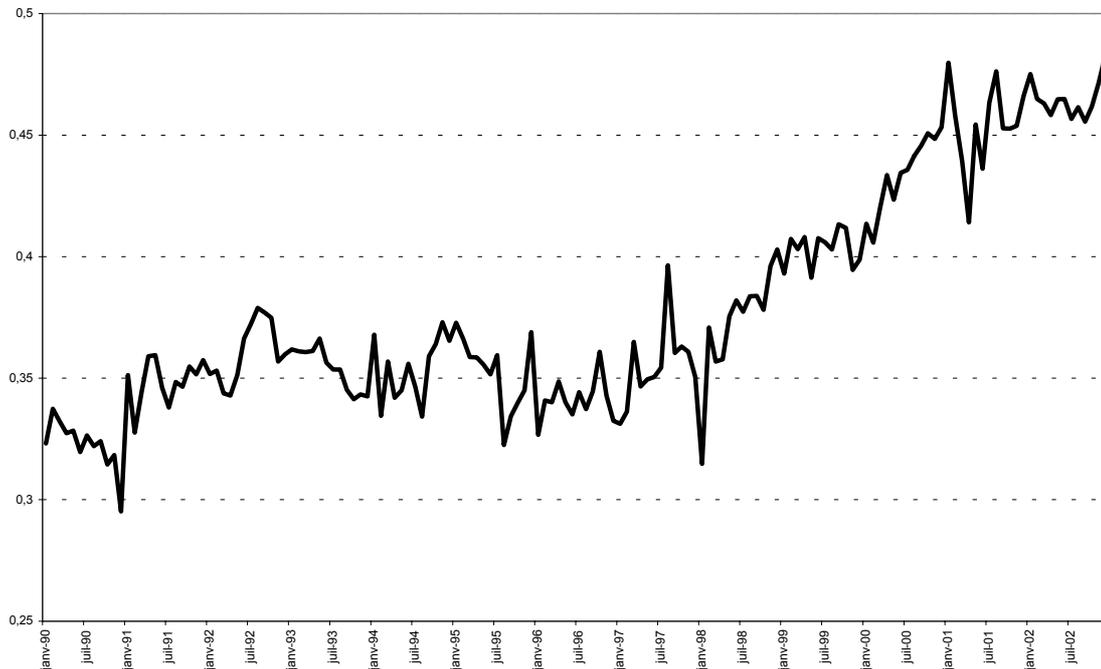
Le retour de la croissance, l'essor de la consommation et du tourisme, puis la création de la carte Imagine R entraînent une forte augmentation de trafic à partir de 1997. Entre 1997 et 2002, le trafic du samedi sur le RER a augmenté de +24.9%, au rythme annuel moyen de +3.8%. Cette augmentation est légèrement plus marquée que sur le métro, pour le même type de journée : le tronçon central du RER A dessert presque exclusivement des zones de commerce et de tourisme, avec de nombreuses salles de cinéma (le centre commercial de La Défense, les

Champs-Élysées, les grands magasins, le Forum des Halles...), et il faut emprunter le RER pour se rendre à Eurodisney.

Les tests de racine unitaire confirment que cette série en forte croissance est intégrée à l'ordre 1.

$\ln(srersa)$	Variable en niveau	Variable en différence
TEST ADF		
valeur de ρ	0.9774	-1.2192
t_{ADF}	-0.7017	-7.8708
VC_{ADF} à 5%	-2.9124	-2.9124
TEST KPSS		
Largeur de fenêtre	Statistique	
3	2.6435	0.0654
6	1.5641	0.0932
9	1.1314	0.1083
VC_{KPSS} à 5%	0.4572	0.4572
Conclusion	$I(1)$	$I(0)$

3.4.8 Le trafic du dimanche/jour férié sur le RER



dfrensa : Trafic du dimanche / jour férié sur le RER (données CVCP - CVS) en millions de voyages

Le trafic du dimanche/jour férié sur le RER a augmenté de +44.3% entre 1990 et 2002, soit un taux d'accroissement annuel moyen égal à +2.9%. Le trafic du dimanche/jour férié sur le RER correspond à la série qui a le plus augmenté sur la période considérée, mais aussi à celle dont le trafic est le plus faible : en moyenne, 0.38 million de voyages sont effectués un dimanche/jour férié sur le RER.

Entre 1990 et 1995, il s'est accru de +1.5% par an, en moyenne. Cet accroissement, dans un contexte économique morose, s'explique par l'augmentation du trafic en 1991 et 1992 (respectivement +8.2% et +3.4% par rapport à l'année précédente). L'année 1996, qui fait suite aux fortes grèves dans le secteur public et aux attentats dans les transports en commun, correspond à une année de forte baisse (-3.4% par rapport à 1995) : une partie de la clientèle a déserté les transports en commun d'une part, d'autre part les déplacements liés aux loisirs (dont les déplacements touristiques) ont fortement chuté. La reprise du trafic le dimanche sur le RER date de 1997 : entre 1997 et 2002, ce type de trafic a cru en moyenne de +4.6% par an, soit une

augmentation de +31.0% sur la sous-période considérée. Cette évolution est comparable à celle observée sur le métro, et est expliquée par des effets identiques : la reprise des déplacements liés aux achats et aux loisirs (notamment le cinéma ou le tourisme) ou le succès de la carte Imagine R. Les années 1999 et 2000 correspondent aux années de plus forte progression de ce trafic, avec des accroissements de trafic par rapport à l'année précédente de, respectivement, +8.0% et +7.6%. La poursuite de l'accroissement de trafic observé le dimanche s'effectue désormais à un rythme inférieur.

La forte croissance du trafic observé sur le métro au cours des dimanches et jours fériés incite à penser que la série est intégrée. Les tests de racine unitaire, synthétisés dans le tableau suivant, confirment que la série de trafic du dimanche/jour férié est intégrée à l'ordre 1 sur l'ensemble de la période.

$\ln(df\text{rersa})$	Variable en niveau	Variable en différence
TEST ADF		
valeur de ρ	0.9900	-0.8564
t_{ADF}	-0.3730	-9.4984
VC_{ADF} à 5%	-2.9124	-2.9124
TEST KPSS		
Largeur de fenêtre	Statistique	
3	3.0593	0.0744
6	1.8086	0.1140
9	1.3044	0.1415
VC_{KPSS} à 5%	0.4572	0.4572
Conclusion	$I(1)$	$I(0)$

* * *

3.4. TRAFICS ÉTUDIÉS ET ÉVOLUTIONS

Les séries de trafic RATP, corrigées des variations calendaires et des perturbations ("CVCP") mais aussi corrigées des variations saisonnières ("CVS" ou "SA" en langue anglaise) étudiées au sein de ce chapitre sont rappelées dans le tableau suivant. Chacune de ces séries mensuelles, transformée par le logarithme népérien, est intégrée à l'ordre 1.

Identifiant	Nom de la série corrigée des variations calendaires, des perturbations, et des variations saisonnières	Conclusions des tests de RU sur la série log-linéarisée
<i>jomsa</i>	Trafic du jour ouvrable sur le métro	I(1)
<i>jorersa</i>	Trafic du jour ouvrable sur le RER	I(1)
<i>joparsa</i>	Trafic du jour ouvrable sur le bus de Paris	I(1)
<i>jobansa</i>	Trafic du jour ouvrable sur le bus de banlieue	I(1)
<i>smsa</i>	Trafic du samedi sur le métro	I(1)
<i>dfmsa</i>	Trafic du dimanche/jour férié sur le métro	I(1)
<i>srersa</i>	Trafic du samedi sur le RER	I(1)
<i>dfwersa</i>	Trafic du dimanche/jour férié sur le RER	I(1)

Chaque série de trafic, pour un type de jour et un sous-réseau défini, est caractérisée au début des années 1990 par une relative stabilité, voire une légère décroissance. Les grèves de décembre 1995 et leurs effets rémanents ont causé une brusque rupture des séries étudiées, avec le départ brutal d'une partie de la clientèle. Dès 1997, cependant, les trafics sur les sous-réseaux connaissent une forte croissance, à mettre en lien avec l'essor de la consommation mais aussi le retour de la croissance économique et des touristes à Paris. L'amélioration constante de l'offre de service est une autre explication de l'essor des transports collectifs.

Ces différents facteurs économiques, dont nous avons vu au chapitre 2 qu'ils constituent des déterminants de la demande de transport collectif, sont étudiés dans le chapitre suivant.

Chapitre 4

Les données d'environnement

Au sein du département Commercial de la RATP, une base de données d'environnement relatives à la région Ile-de-France a été constituée depuis quelques années. Cette base a pour principal intérêt de donner des explications qualitatives aux variations de trafic observées ; elle est périodiquement remise à jour, dans la limite des données disponibles. En effet, le suivi de certaines variables a été interrompu du fait de l'arrêt de leur publication (c'est par exemple le cas pour le nombre d'entrées à Disneyland Resort Paris), d'autres ne sont pas homogènes dans le temps car le mode de calcul ou le périmètre de référence ont changé. En outre, il est plus difficile d'obtenir des données périodiques pour une région que pour l'ensemble de la France, il n'existe pas d'organisme régional centralisant l'ensemble des données économiques nécessaires. Enfin, toutes ces variables économiques ne possèdent pas forcément la même périodicité : certaines ne sont qu'annuelles ou trimestrielles, d'autres sont mensuelles.

Le but de ce chapitre est quadruple :

1. présenter les variables d'environnement relatives à l'Ile-de-France (définition, source, unité...), qui sont disponibles mensuellement sur la période 1990 - 2002,
2. exposer leur évolution sur cette période,
3. expliquer pourquoi il est important de ne considérer que des variables parisiennes ou franciliennes, dans la mesure où elles ont une dynamique distincte des variables considérées au niveau national,

4. présenter les modifications appliquées aux séries disponibles (désaisonnalisation, log-linéarisation,...) et mettre en avant certaines propriétés statistiques comme le degré d'intégration des séries.

Pour cela, nous nous intéresserons tout d'abord aux données relatives à l'emploi en Ile-de-France, puis nous aborderons, dans l'ordre, le commerce, le tourisme et la concurrence automobile, toujours en Ile-de-France. Nous évoquerons enfin le manque de données concernant l'évolution de la situation immobilière.

4.1 L'emploi en Ile-de-France

L'activité est le moteur principal des déplacements, puisque après le domicile, le travail est le premier motif à l'origine (voir enquête TJRF). Le travail est en outre lié à d'autres motifs de déplacements : par le biais des revenus, il influe sur le niveau des déplacements associés aux loisirs ou aux achats.

Pour représenter l'évolution mensuelle du marché du travail en Ile-de-France, deux types de données disponibles périodiquement peuvent être envisagés : le taux de chômage au sens du Bureau International du Travail (BIT) fourni par l'INSEE¹ d'une part, et, d'autre part, les demandes d'emploi en fin de mois (DEFM) issues des publications de la DRTEFP² d'Ile-de-France.

Nous nous proposons de présenter la différence entre les séries de chômage au sens du BIT (1^{er} paragraphe) et celles relatives aux demandes d'emploi en fin de mois (2^{ème} paragraphe). Cette analyse a pour but d'expliquer pourquoi, dans le reste de l'étude, nous ne retiendrons que cette deuxième approche.

Les différences tiennent autant dans la définition même du concept de chômage que dans la constitution et le traitement des séries.

4.1.1 Taux de chômage fournis par l'INSEE

Après avoir rappelé la définition du chômage au sens du BIT, et le calcul du taux de chômage, nous examinerons la méthode d'estimation de la série trimestrielle au niveau de la région Ile-

¹INSEE : Institut National de la Statistique et des Etudes Economiques

²DRTEFP : Direction Régionale du Travail, de l'Emploi et de la Formation Professionnelle

de-France. La périodicité de cette série ne convient pas : pour être en cohérence avec les séries expliquées (i.e. les séries de trafic), les variables explicatives doivent être des données mensuelles.

Toutefois, si les évolutions du chômage sont identiques au niveau national et au niveau régional, la série de chômage pour l'ensemble de la France pourra servir de base à la mensualisation de la série de chômage en Ile-de-France. La comparaison entre les évolutions sur l'ensemble de la France et sur la Région met en lumière les spécificités de l'emploi en région Ile-de-France, et explique pourquoi cette méthode de mensualisation n'a pas été retenue.

4.1.1.1 Chômage au sens du BIT, et taux de chômage

La définition du Bureau International du Travail (BIT), correspond à la définition du chômage la plus restrictive.

Elle considère qu'une personne en âge de travailler (plus de 15 ans) est au chômage si :

- Elle est sans emploi, c'est-à-dire qu'elle n'a pas travaillé (même pas une heure) durant la semaine de référence,
- Elle est disponible pour prendre un emploi dans les 15 jours,
- Elle cherche activement un emploi, ou en a trouvé un qui commence ultérieurement.

Le taux de chômage est obtenu en rapportant le nombre de chômeurs au sens du BIT à la population active totale, somme des effectifs salariés et non salariés, des militaires du contingent et des chômeurs.

L'emploi qui figure au dénominateur du taux de chômage est établi par sexe à la date du 31 décembre de chaque année dans chaque région, à partir de sources administratives diverses. Un certain nombre de fichiers administratifs fournissent des évolutions annuelles d'emploi, comme l'UNEDIC³, l'URSSAF⁴, le MSA⁵, les fichiers de paye des agents de l'Etat... Ces résultats sont complétés par des enquêtes, telles les enquêtes annuelles sur les collectivités locales, ou les enquêtes Emploi de l'INSEE, réalisées en mars chaque année jusqu'en 2002. Depuis 2003, l'enquête Emploi est réalisée en continu [26]. A partir de ces données, il est possible de calculer en

³UNEDIC : Union Nationale pour l'Emploi Dans l'Industrie et le Commerce

⁴URSSAF : Union pour le Recouvrement des cotisations de la Sécurité Sociale et des Allocations Familiales

⁵MSA : Mutualité Sociale Agricole

fin d'année l'évolution de l'emploi par secteur d'activité ou par région. Les estimations annuelles servent alors de base aux estimations trimestrielles d'emploi, en complément du fichier UNEDIC et de l'enquête ACEMO⁶ réalisée par le Ministère de l'emploi et de la solidarité auprès des entreprises. Ces deux sources ont néanmoins pour inconvénient de ne couvrir qu'un champ partiel et ne sont pas régionalisées.

4.1.1.2 Estimation du chômage régional

Les taux de chômage régionaux et départementaux calculés par l'INSEE, d'après la définition préconisée par le BIT, découlent de la synthèse de diverses sources statistiques et administratives. L'enquête Emploi permet une analyse qualitative des disparités régionales de chômage par âge et sexe et le calcul de coefficients de passage. Le chômage par région au sens du BIT est estimé en appliquant ces coefficients de passage aux demandes d'emplois en fin de mois ("DEFM") fournies par l'ANPE.

"L'emploi au lieu de travail", sur lequel s'appuie l'Enquête Régionale sur l'Emploi, est ensuite transformé en "emploi au lieu de résidence" à partir des résultats des recensements, puis trimestrialisé par interpolation linéaire. Les taux sont corrigés des variations saisonnières. Ils font l'objet d'une révision deux fois par an : une première fois à l'été de l'année n , lorsque les résultats de l'enquête Emploi de mars sont disponibles, puis, en début d'année $n + 1$, le calcul des variations saisonnières est actualisé et les estimations d'emploi mises à jour.

Ainsi, un premier inconvénient des séries de l'INSEE, outre leur périodicité qui ne convient pas à notre étude, est le fait que ces données sont déjà désaisonnalisées, par une méthode qui n'est pas homogène à celle que nous appliquerons aux données brutes de certaines autres séries, dont les séries de trafic. En outre, ces données ne sont validées que très tardivement.

4.1.1.3 Une estimation mensuelle du chômage au niveau national

Depuis octobre 1996, l'INSEE a mis en place une nouvelle méthode d'estimation mensuelle du chômage BIT au niveau national. Cette méthode a été appliquée rétroactivement depuis le mois d'avril 1994, les estimations du mois de mars restant inchangées puisqu'elles sont directement

⁶ACEMO : Activités et Conditions d'Emploi de la Main d'Oeuvre

issues des enquêtes annuelles sur l'emploi. Cette méthode vise à rendre compte le mieux possible de l'évolution conjoncturelle du chômage BIT, en minimisant les révisions annuelles dues au calage sur l'enquête emploi, et en atténuant l'impact des perturbations accidentelles des données de base (les demandes d'emplois en fin de mois, notées DEFM). Un modèle économétrique est utilisé, et relie les glissements mensuels du chômage BIT à ceux de deux catégories de variables :

- les DEFM de la catégorie 1, et les DEFM des catégories 1 + 6 (définies page 166),
- deux variables de conjoncture décalées de 3 mois : l'emploi intérimaire et les créations d'entreprises employeuses exprimées en écart par rapport à leur moyenne.

En novembre 2001, les taux de chômage ont été révisés à partir du 2^{ème} trimestre 2001, du fait de l'application d'une nouvelle méthode permettant d'affiner les estimations. Ce sont l'ensemble des DEFM de catégories 1 (demandes d'emploi à durée indéterminée à temps plein), 2 (à durée indéterminée à temps partiel) et 3 (à durée déterminée) hors activité réduite, et non plus uniquement celles de catégorie 1, qui sont mobilisées pour estimer l'évolution du chômage.

La relation est estimée en glissements annuels bruts, puis mensualisée et appliquée aux données mensuelles désaisonnalisées. Elle fournit alors une estimation du chômage BIT corrigée des variations saisonnières. Quand les résultats de l'enquête annuelle sur l'emploi sont connus, la série provisoire ainsi constituée est révisée, et les modèles sont ré-estimés pour servir de base aux estimations des 12 mois suivants.

Les données mensuelles de chômage, même au niveau national, ne sont pas mesurées, mais estimées. Les méthodes d'estimation ont changé au cours du temps, à tel point que l'on ne peut être sûr de l'homogénéité de la série sur une longue période. On se demande néanmoins si elles ne pourraient pas servir de base pour mensualiser les séries trimestrielles de taux de chômage en Ile-de-France.

4.1.1.4 Vers une mensualisation des taux de chômage trimestriels franciliens à partir de données nationales ?

Dans la mesure où le modèle que nous désirons construire est un modèle à données mensuelles, les taux de chômage trimestriels franciliens devraient être mensualisés. Une solution consisterait à mensualiser les données franciliennes trimestrielles à l'aide de la série mensuelle nationale.

Cette solution est néanmoins inapplicable dans la mesure où les évolutions entre le chômage francilien et le chômage au niveau national sont contrastées.

Les dynamiques contrastées du chômage en Ile-de-France et au niveau national sur la période 1990-2002 Depuis 1990, le taux de chômage national est en moyenne supérieur au taux francilien de 1.2 points. Néanmoins, cet écart s'est fortement réduit depuis le 2^{ème} trimestre 2001, surtout au cours de l'année 2002, où l'écart moyen sur l'année n'est plus que de 0.5 point. La hausse du chômage a été plus forte en Ile-de-France qu'au niveau national sur cette dernière période. En décembre 2003, le taux de chômage francilien s'établit à 9.7% (données CVS) contre 8.7% en décembre 2002 : il a rejoint le niveau sur France entière. A Paris, le taux de chômage atteint 11.5%.

Il existe certains décalages dans les retournements des courbes de chômage français et franciliens, et des évolutions plus ou moins fortes suivant le périmètre considéré. Ainsi, l'accroissement du taux de chômage a été plus rapide en Ile-de-France que dans l'ensemble de la France au début des années 90. Sa décroissance durant l'année 1994 est apparue plus précocement quoique de façon moins forte, de même que la reprise de la montée du chômage dans les années qui ont suivi les grandes grèves de la fin de l'année 1995.

Les évolutions du chômage au niveau de l'Ile-de-France et au niveau national ne sont pas identiques, il semble exister une spécificité régionale.

La spécificité de l'Ile-de-France en terme de chômage et de marché de l'emploi

L'Ile-de-France constitue le pôle de développement de la France, et l'une des plus importantes métropoles européennes et mondiales par l'ampleur de sa population et le nombre d'emplois. D'après le recensement général de la population en 1999, la région comporte près de 11 millions d'habitants, soit 18.7% de la population française, sur 2.2% du territoire français (12 012 km²) et le nombre d'emplois est supérieur à 5 millions (25% des emplois salariés nationaux), pour une population active totale égale à 5 980 000 personnes. 623 509 entreprises y sont présentes, soit 18.1% de l'effectif national. Sa valeur ajoutée représente 28.7% de la richesse nationale, et le Produit Intérieur Brut annuel par habitant est l'un des plus élevés de l'Union Européenne et

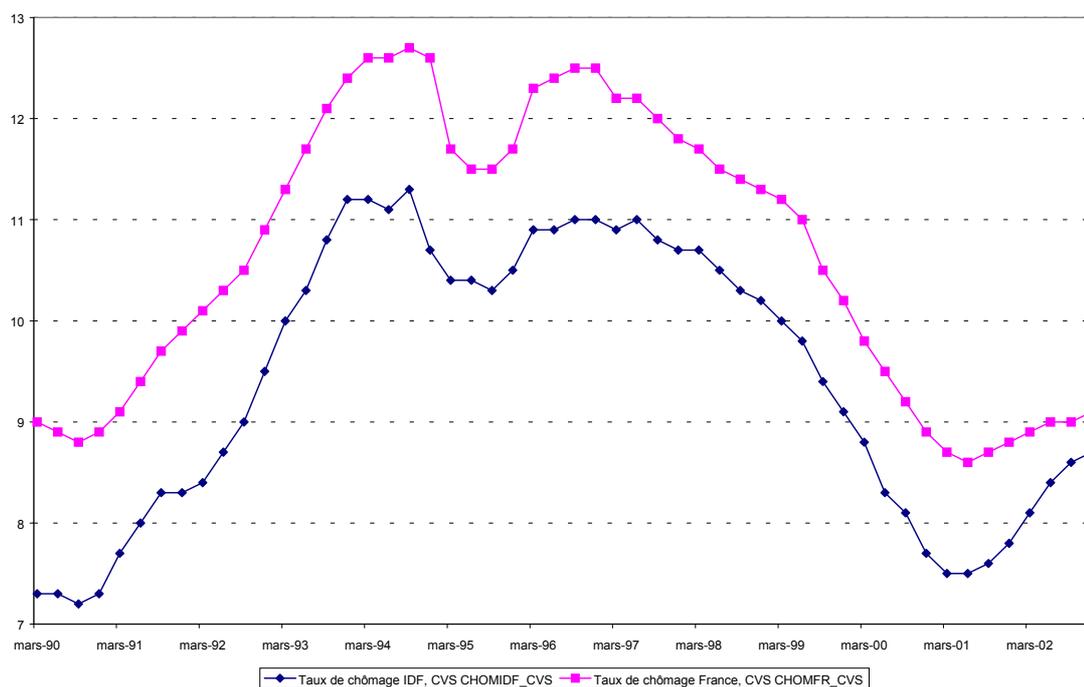


FIG. 4-1: Taux de chômage en France et en Ile-de-France

s'élève en 2002 à 38 739 euros courants. Après Tokyo, c'est le deuxième pôle d'implantation des 500 plus grandes entreprises du monde.

La prédominance du secteur tertiaire L'économie francilienne se caractérise par une prédominance du secteur tertiaire, comprenant administrations, services marchands, établissements financiers à vocation internationale, services aux entreprises. Ainsi, 81% des effectifs franciliens travaillent dans le secteur tertiaire, 14% dans l'industrie, près de 5% dans la construction et 0.2% dans l'agriculture. Cette répartition se retrouve dans la nature des établissements : 82.1% appartiennent au secteur tertiaire, 8.4% sont des établissements industriels, 8.1% appartiennent au secteur de la construction et 1.4% à l'agriculture.

Composée de 8 départements, la région attire 38% des sièges sociaux implantés sur le territoire national, notamment les trois-quarts des sièges des banques et compagnies d'assurance et plus du quart des sièges sociaux de l'industrie. Des zones fortement spécialisées en activités tertiaires, comme la Défense ou Marne-La-Vallée traduisent cette prédominance du secteur

tertiaire. Ce type d'activités tertiaires, notamment celles relatives aux services rendus aux entreprises, sont très sensibles à la conjoncture. Cette sensibilité explique la plus forte diminution du chômage dans les années 1997-2000 en Ile-de-France qu'au niveau national, et la dégradation plus marquée de la situation de l'emploi en Ile-de-France depuis le retournement conjoncturel de 2001.

L'Ile-de-France demeure néanmoins la première région industrielle française en employant 24% des effectifs nationaux dans ce secteur (642 066 emplois). Deux secteurs importants de l'industrie francilienne sont l'édition-imprimerie (40% des effectifs nationaux), qui a su faire face au numérique et progresse depuis 1997, et l'industrie automobile, dont la restructuration a débuté dans les années 80. Deux secteurs stratégiques enregistrent une diminution de leurs effectifs : le secteur de la Recherche et Développement, atteint par les mesures de rééquilibrage national, et l'industrie pharmaceutique fortement restructurée.

Les firmes Renault, Citroën, Thomson CSF ou IBM sont implantées en Ile-de-France même si l'industrie francilienne est essentiellement constituée de PME/PMI. Un tiers des effectifs industriels travaille pour le compte d'une entreprise étrangère.

D'après une étude intitulée "Géographie de l'emploi en 2000", qui s'appuie sur l'Enquête Régionale sur l'Emploi menée conjointement par l'IAURIF⁷, l'INSEE, la DREIF⁸, l'APUR⁹, la Ville de Paris et le CRIF¹⁰, les emplois industriels demeurent très concentrés en région Ile-de-France, avec plus de 60% dans les Hauts-de-Seine, à Paris, et dans le département des Yvelines, qui a bénéficié du desserrement de la petite couronne.

En fait, la situation de l'industrie en Ile-de-France est paradoxale : première région industrielle française malgré le recul continu (moins sur la période 1997/2000) de son poids relatif, c'est aussi une des régions françaises les moins industrialisées. Ce phénomène résulte de la transformation du système productif, en particulier du processus d'externalisation des fonctions tertiaires des entreprises industrielles qui ne sont donc plus comptabilisées dans l'industrie et font croître largement les effectifs des services aux entreprises.

⁷IAURIF : Institut d'Aménagement et d'Urbanisme de la Région Ile de France

⁸DREIF : Direction Régionale de l'Équipement en Ile de France

⁹APUR : Atelier Parisien d'URbanisme

¹⁰CRIF : Conseil Régional d'Ile de France

L'agriculture occupe 50% du territoire régional et se concentre principalement en Seine-et-Marne.

Les industries agroalimentaires franciliennes réalisent le plus fort chiffre d'affaire et la plus forte valeur ajoutée au niveau national.

Une main d'oeuvre plus qualifiée La région offre une main d'oeuvre hautement qualifiée, avec 47.8% des cadres français (34% des cadres supérieurs) et 42% du personnel de recherche, ainsi que 30% des établissements d'enseignement supérieur. La sur-représentation des cadres et professions intellectuelles supérieures est à mettre en lien avec l'importance du secteur tertiaire, la forte implantation des sièges sociaux (6500 sièges sociaux en Ile-de-France) et la concentration des filières de formation supérieure dans la région. On y dénombre notamment 17 universités, de grandes écoles et de grands instituts de recherche scientifique : l'Ecole Centrale de Paris, l'Ecole Nationale des Ponts et Chaussées, l'Ecole Nationale Supérieure des Télécommunications, l'Ecole Polytechnique, l'Ecole Supérieure de Physique et Chimie Industrielle, l'Institut d'Etudes Politiques de Paris, l'école des Hautes Etudes Commerciales, l'Ecole Normale Supérieure des Mines de Paris, l'ENS, l'Ecole Supérieure Commerce de Paris...

Un marché plus fluide et plus actif Le marché du travail est plus fluide et plus actif en Ile-de-France qu'au niveau national : les établissements du secteur privé renouvellent plus fréquemment leur personnel. Pour les établissements de plus de 50 salariés, les embauches en contrat à durée indéterminée, rapportées au personnel en place, sont plus nombreuses. Le recours au licenciement économique y est aussi plus fréquent, proportionnellement à la population active.

Pour prendre en compte la spécificité du chômage et du marché de l'emploi, et avoir des données mensuelles, il ne paraît pas judicieux d'extrapoler la série trimestrielle du chômage francilien à l'aide des données nationales mensuelles. Une solution pour représenter l'évolution de l'emploi est de considérer les séries départementales de Demandes d'Emplois en Fin de Mois, séries qui par ailleurs servaient de base aux calculs des taux de chômage franciliens.

Ces séries, qui présentent l'avantage d'être mensuelles et prélevées directement au niveau du département (sans correction saisonnière, calculs, rectifications ou interpolation linéaire), sont

présentées dans le paragraphe suivant.

4.1.2 Les Demandes d'Emploi en Fin de Mois (DEFM)

4.1.2.1 Définition et homogénéité de la série des DEFM

La mesure du chômage fournie par l'ANPE¹¹ correspond aux demandes d'emploi en fin de mois (DEFM). Cette mesure présente l'avantage d'être une mesure mensuelle exhaustive et localisée.

Avant juin 1995, il existait 3 catégories principales de DEFM :

- la catégorie 1, qui comprend les personnes sans emploi à la recherche d'un emploi à durée déterminée à temps plein,
- la catégorie 2, constituée des personnes sans emploi à la recherche d'un emploi à durée déterminée à temps partiel,
- et la catégorie 3 qui regroupe les personnes sans emploi à la recherche d'un emploi à durée déterminée.

Un arrêté du Conseil d'Etat daté du 13 mai 1994 a conduit à modifier ces catégories de demandeurs d'emploi à partir du mois de juin 1995. Les nouvelles catégories 1, 2 et 3 ne comprennent plus les personnes ayant travaillé plus de 78 heures, qui sont considérées comme indisponibles immédiatement, et sont classées dans les nouvelles catégories 6, 7 et 8 selon qu'elles recherchent, respectivement, un emploi à durée indéterminée à temps plein ou à temps partiel, ou un emploi à durée déterminée.

La somme des nouvelles catégories 1 et 6 correspond exactement à l'ancienne catégorie 1, et sert de référence. La série est homogène dans le temps, et les données sont rapidement disponibles, dans les publications mensuelles de la DRTEFP¹² d'Ile-de-France intitulées "Marché du travail" jusqu'en 2002, sur demande ensuite.

Le nombre de DEFM et celui des chômeurs au sens du BIT ne sont pas identiques dans la mesure où des chômeurs au sens du BIT peuvent ne pas être inscrits à l'ANPE. C'est le cas en

¹¹ANPE : Agence Nationale Pour l'Emploi

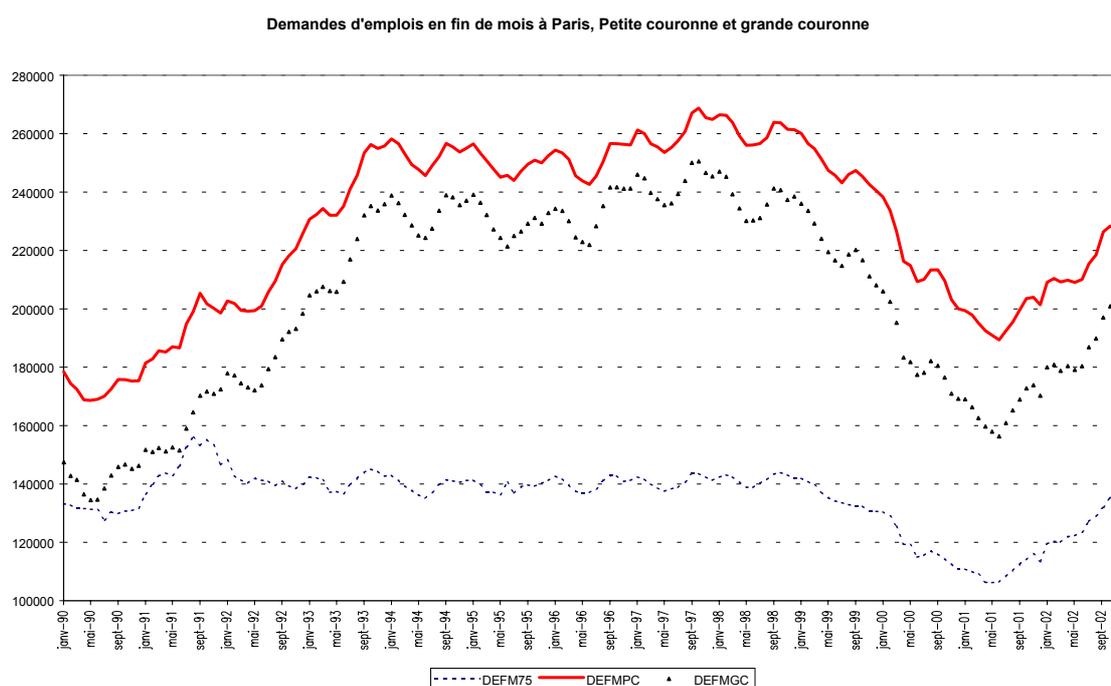
¹²DRTEFP : Direction Régionale du Travail, de l'Emploi et de la Formation Professionnelle

particulier pour des femmes qui reviennent sur le marché du travail, ou de jeunes qui y rentrent pour la première fois. A l'inverse, des demandeurs d'emploi peuvent avoir un emploi (en général de durée très réduite) ou être inactifs au sens du BIT dans le sens où ils ne recherchent pas "activement" un emploi.

Compte-tenu de la nécessité de prendre des données mensuelles, homogènes dans le temps, et relatives à la seule région Ile-de-France, nous avons choisi de considérer les séries relatives aux DEFM dans l'ensemble des travaux économétriques. Les taux de chômage régionaux, calculés par l'INSEE, ont pour but d'éclairer les analyses d'évolution du chômage sur la période 1990-2002.

4.1.3 Evolution du taux de chômage et des demandes d'emploi en fin de mois depuis 1990

On dispose du graphique suivant :



4.1.3.1 La montée du chômage en Ile-de-France entre 1990 et 1997

En 1990, en Ile-de-France, le nombre des effectifs salariés a atteint son point culminant, et la tendance s'inverse les années suivantes. Entre 1990 et 1997, l'Ile-de-France va perdre -2.5%

de ses emplois salariés et connaître une profonde mutation dans la nature de ses activités : poursuite du gonflement du secteur tertiaire, de la désindustrialisation et de la récession du secteur du bâtiment. En outre, le chômage commence à croître. Au premier semestre 1994, une tendance à la baisse de la croissance du chômage puis à une stagnation de celui-ci se dessine en Ile-de-France, grâce à la baisse des licenciements économiques et à la diminution du chômage chez les jeunes. Ces derniers poursuivent plus longtemps leurs études ou effectuent davantage de "petits boulots". Par contre, le nombre de chômeurs de longue durée continue à croître, ainsi que le chômage chez les personnes de plus de 50 ans. Pour ces personnes, les possibilités d'insertion sur le marché du travail sont de plus en plus incertaines.

L'aggravation du chômage en 1996 touche tous les départements, l'évolution est identique au niveau régional et au niveau départemental. Au premier semestre 1997, le chômage continue d'augmenter en Ile-de-France, et de façon plus rapide que dans le reste de la France, du fait d'une moindre reprise francilienne de l'emploi dans le secteur des services aux entreprises et aux personnes. En outre, contrairement aux trois années passées, la baisse du chômage chez les jeunes ne s'observe plus, ce qui conduit à un accroissement du nombre de personnes de moins de 25 ans se présentant pour la première fois à l'ANPE. La progression du chômage chez les personnes de plus de 50 ans reste vraie. Ainsi, en juin 1997, un chômeur sur 10 a moins de 25 ans et 2 chômeurs sur 10 sont âgés de plus de 49 ans. Cet accroissement du chômage s'explique enfin par des pertes d'emploi salarié dans l'industrie et la construction, pertes non compensées par le secteur tertiaire, qui n'a pas offert d'emplois nouveaux sur cette période.

A partir de juillet 97, les jeunes sont plus nombreux à choisir de s'insérer dans le monde professionnel au moyen de contrat par alternance, notamment contrat d'apprentissage et de qualification. D'autres éléments expliquent la baisse du nombre de demandes d'emplois en fin de mois : l'offre de travail temporaire progresse vivement et le chômage partiel s'avère moins important que par le passé. Malgré tout, le chômage de longue durée continue à croître fortement : le pourcentage de chômeurs de longue durée dans la population de demandeurs d'emploi atteint 37.1% en décembre 1997, leur nombre a augmenté de +10.3% par rapport à l'année précédente. Au total, entre 1990 et 1997, le taux de chômage en Ile-de-France est passé de 7.3% à 10.8%, ce qui correspond à un accroissement du nombre de chômeurs au sens du BIT de +60%. La période

de chômage moyenne s'est allongée sur cette même période, passant de 11.4 mois en 1990 contre 14.7 mois en 1997.

4.1.3.2 L'embellie du chômage entre 1997 et 2001

L'année 1997 marque un tournant, celui de la reprise économique. Par un effet d'hystérésis, l'impact sur le chômage n'est pas ressenti immédiatement. La courbe de chômage, qui s'infléchit depuis le mois de septembre se retourne véritablement en décembre 1997, premier mois de l'année à enregistrer une variation du chômage négative. L'amélioration est, au départ, moins importante en Ile-de-France qu'au niveau national et ne porte pas sur tous les départements : le nombre de demandeurs d'emploi continue à augmenter en Seine-Saint-Denis jusqu'en septembre 1998 alors qu'il diminue fortement dans l'Essonne et les Yvelines. Le chômage des jeunes poursuit sa décroissance, de même que le chômage chez les personnes de 25 à 49 ans, mais la baisse ne concerne ni le chômage de longue durée (dont la part dans le chômage dépasse les 40% à Paris et en Seine-Saint-Denis), ni le chômage chez les personnes âgées de 50 ans et plus

Pendant l'année 1998, la baisse du chômage est plus importante en Ile-de-France qu'au niveau national. Le chômage de longue durée commence à diminuer, et baisse de -17.4% en 1999, c'est-à-dire plus vite que le chômage dans son ensemble. En mars 1999, le taux de chômage passe en dessous de la barre des 10% de chômeurs qui avait été franchie en 1993. Le recours à l'alternance commence à se stabiliser et les embauches sont soutenues par les mesures d'exonération pour la première embauche et les embauches à temps partiel. En novembre 2000, le nombre francilien de demandeurs d'emploi inscrits en catégorie 1 et 6 (ancienne catégorie 1) passe en dessous de la barre des 500 000, et retrouve son niveau de juin 1991, du fait des nombreuses créations d'emploi au cours de l'année 2000, notamment dans le secteur tertiaire, secteur à l'origine de la moitié des emplois créés en Ile-de-France. La baisse du chômage en 2000 est plus importante dans les départements de grande couronne ($-18,0\%$) qu'à Paris (-16.4%) et en petite couronne (-16.7%). Les dispositifs destinés à favoriser l'insertion des chômeurs de longue durée ou en voie d'exclusion, tels le programme "nouveaux départs" inscrit dans le plan national d'action pour l'emploi ou les programmes départementaux globalisés, portent leurs fruits : le chômage de longue durée continue à diminuer. La mise en oeuvre des 35 heures a peut-être contribué aussi à ces créations

d'emploi. Pourtant, le nombre de bénéficiaires des mesures en faveur de l'emploi diminue de -3.4% en 2000, à l'exception des formations en alternance, telles les contrats de qualification et les contrats d'orientation. Certains employeurs ont anticipé leurs embauches compte-tenu de la suppression de la prime à l'embauche pour les contrats de qualification postérieurs au 1^{er} janvier 2001.

4.1.3.3 Le retournement du cycle en 2001

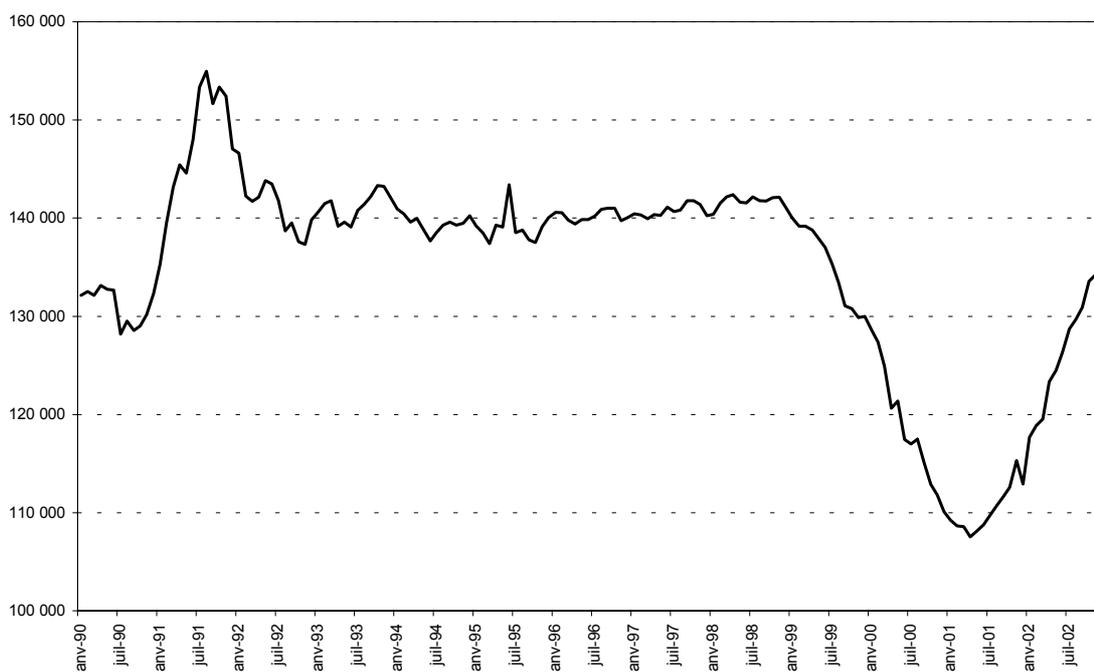
Après 4 années de croissance économique, plus riche en emplois que les cycles précédents, l'économie française ainsi que francilienne a été marquée par un net ralentissement au cours de l'année 2001. La dégradation de la conjoncture internationale s'est propagée à partir des Etats-Unis à l'ensemble des pays européens. Le maintien de la consommation des ménages, soutenue par la progression du pouvoir d'achat du revenu disponible suite aux mesures d'allègement fiscal, aux primes et aides et à la faiblesse de l'inflation, a empêché un décrochage brutal de l'activité. Après une baisse régulière du nombre de chômeurs franciliens depuis plusieurs années, celui-ci est reparti à la hausse au second semestre 2001, de façon bien plus marquée en Ile-de-France que sur l'ensemble du territoire national. La région subit les conséquences de l'effondrement des technologies de l'information et de la consommation, un secteur prédominant en Ile-de-France. Par ailleurs, différentes opérations de restructuration ont eu lieu dans des secteurs tels que la maintenance industrielle, entraînant des annonces de plans sociaux à l'été et l'automne 2001, et une forte augmentation du nombre de licenciements pour motif économique. Le nombre de chômeurs de longue durée a continué à diminuer, tandis que le chômage des moins de 25 ans a augmenté. L'arrêt anticipé du service national ainsi que l'inscription plus rapide des jeunes diplômés à l'ANPE, expliquent en partie cette augmentation, plus marquée dans cette classe d'âge que chez les autres.

Au deuxième trimestre 2002, les demandes d'emplois ont continué de croître, mais à un rythme légèrement moins soutenu. La relance de l'économie est insuffisante pour infléchir la courbe du chômage et, sur l'année, le chômage francilien s'est accru de $+17\%$ soit près du triple de l'évolution nationale. L'Ile-de-France est la troisième région de France enregistrant la hausse la plus marquée, après l'Alsace ($+31.3\%$) et la Franche-Comté ($+28.1\%$). Cette spécificité est due

à la forte présence en Ile-de-France d'activités liées aux nouvelles technologies de l'information et de la communication, alors que la période est caractérisée par de faibles recrutements et un grand nombre de cessations d'activité des start-ups.

Au cours de l'année 2003, la dégradation économique est telle que l'on parle de récession ; le taux de chômage régional a continué à augmenter.

4.1.4 Les DEFM à Paris



defm75 : Demandes d'emplois en fin de mois à Paris (données DRTEFP CVS)

Sur la période 1990 - 2002, à Paris, il existe en moyenne près de 135 000 demandes d'emplois en fin de mois. Entre 1990 et 2002, le nombre de demandes d'emplois a globalement décru (-3.21%), à un rythme moyen très faible de l'ordre de -0.25% par an.

Plusieurs périodes peuvent être distinguées :

- le fort accroissement au début des années 1990 ($+12.48\%$ entre 1990 et 1991),
- la stagnation des demandes d'emplois en fin de mois à un niveau élevé entre 1992 et 1997,
- une décroissance importante à partir de 1998, au taux de croissance annuel moyen de $-4,95\%$, l'année de plus forte décroissance correspondant à l'année 2000 ($-12,15\%$ par

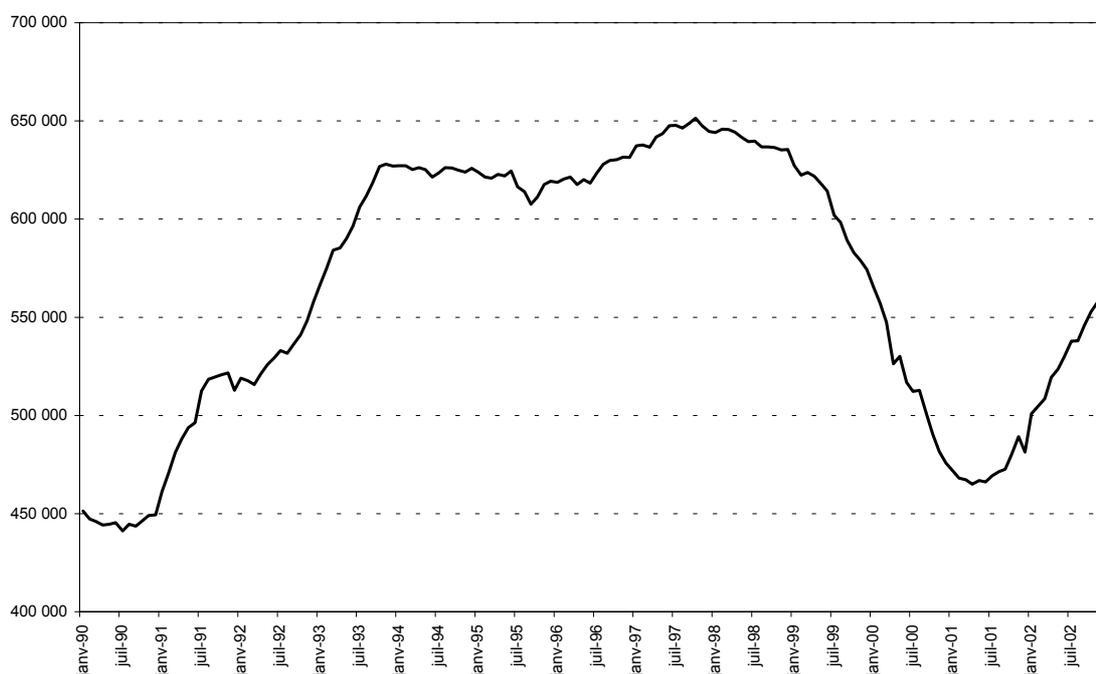
rapport à 1999),

- la recrudescence du nombre de demandes d'emplois au cours de l'année 2002 (+14,99% par rapport à 2001), à un niveau bien inférieur à celui de début de période.

La forte décroissance du nombre de demandes d'emplois en fin de mois lors de la reprise économique peut s'expliquer par la non-stationnarité de la série. Le degré de stationnarité est testé à l'aide des tests de racine unitaire déjà utilisés dans le chapitre précédent, sur la série en niveau et en différence première, pour la série log-linéarisée. Les tests de racine unitaire, présentés dans le tableau suivant concluent à l'existence d'une unique racine unitaire sur la période janvier 1990 - décembre 2002 : la série logarithmique des demandes d'emplois en fin de mois à Paris est intégrée à l'ordre 1.

$\ln(defm75)$	Variable en niveau	Variable en différence
TEST ADF		
valeur de ρ	0.9770	0.5513
t_{ADF}	-1.9453	-3.9364
VC_{ADF} à 5%	-2.9124	-2.9124
TEST KPSS		
Largeur de fenêtre	Statistique	
3	1.7852	0.1705
6	1.0474	0.1257
9	0.7589	0.1039
VC_{KPSS} à 5%	0.4572	0.4572
Conclusion	I(1)	I(0)

4.1.5 Les DEFM en Ile-de-France



defmidf : Demandes d'emplois en fin de mois en Ile-de-France (données DRTEFP CVS)

Sur l'ensemble de la région Ile-de-France, le nombre de demandes d'emplois en fin de mois est plus de quatre fois supérieur à celui sur Paris. Entre 1990 et 2002, il a augmenté de +19.20%, soit +1.36% par an en moyenne, alors qu'il a globalement diminué sur le département - capitale. Les évolutions en Ile-de-France sont distinctes de celles de Paris :

- l'augmentation du nombre de demandes d'emplois est quasi continue jusqu'en 1997, avec un taux d'accroissement annuel moyen égal à +4.62% (contre +0.98% pour Paris),
- la diminution du nombre de demandes d'emplois en fin de mois est plus marquée au niveau de la région que du département de Paris : entre 1998 et 2002, les demandes d'emplois ont diminué en moyenne annuelle de -5.97%, contre -4.95% au niveau de Paris.

Cette évolution différenciée entre Paris et l'ensemble de la région s'explique en partie par l'inégalité des densités d'emplois : il existe plus de 200 emplois salariés à l'hectare construit à Paris contre 12 en grande couronne. Paris regroupe 32.4% des emplois, la petite couronne 35.7% et la grande couronne 31.9%. Il existe en outre de fortes disparités entre les départements

franciliens, liées au tissu économique et industriel de chaque département et aux caractéristiques socio-démographiques des populations. La Seine Saint-Denis concentre une proportion élevée de personnes n'ayant peu ou pas de diplôme, ainsi qu'une forte population étrangère, et affiche le taux de chômage le plus élevé de la région Ile-de-France. En Grande Couronne, à l'exception du Val d'Oise, les départements enregistrent au contraire des taux de chômage assez bas. La situation de l'emploi y a été favorisée par la redistribution des emplois du centre de la région vers la périphérie.

La forte augmentation du nombre de demandes d'emplois en Ile-de-France, puis sa décline à un niveau élevé conduisent à s'interroger sur l'existence d'une racine unitaire au sein de la série.

$\ln(defmidf)$	Variable en niveau	Variable en différence
TEST ADF		
valeur de ρ	0.9845	0.7965
t_{ADF}	-2.6814	-2.4363
VC_{ADF} à 5%	-2.9124	-2.9124
TEST KPSS		
Largeur de fenêtre	Statistique	
3	0.8921	0.6987
6	0.5223	0.4509
9	0.3768	0.3449
VC_{KPSS} à 5%	0.4572	0.4572
Conclusion	I()?	I()?

Les résultats des tests de racine unitaire demeurent difficiles à interpréter : sur la série en niveau comme sur la série en différence, le test ADF ne rejette pas l'hypothèse de racine unitaire, i.e. laisse croire que la série est intégrée à un ordre supérieur à 1. Le test KPSS n'apporte pas de conclusion identique quelle que soit la largeur de fenêtre choisie. Selon le test de Phillips et Perron, toutefois, la série est intégrée à l'ordre 1.

4.2 Le commerce en Ile-de-France

4.2.1 Evolution de la consommation en Ile-de-France depuis 1990

4.2.1.1 Le contexte de la consommation des ménages entre 1990 et 2002

Au début des années 1990, la consommation des ménages est atone (+1.2% entre 1991 et 1992), dans un contexte conjoncturel de désinflation et de déflation. Les produits "premiers prix" font leur apparition. En 1993, le fort taux de chômage chez les cadres, et la pause des salaires poussent les ménages à accroître leur épargne de précaution au détriment de la consommation. Au niveau national, la consommation des ménages baisse de -0.6%. Entre 1994 et 1996, le taux de croissance de la consommation des ménages est de l'ordre de +1.2%, ce qui révèle un attentisme des ménages d'une part et une stagnation de leur pouvoir d'achat d'autre part. L'année de véritable reprise est 1998 avec une progression de +3.4% (ce qui ne s'était pas vu depuis 12 ans) puis +3.2% en 1999. Cette reprise est caractérisée par une recherche de sécurité, de sens et de valeur dans les achats. En réaction à la standardisation des années 50-60 de la consommation dite de masse, des comportements contradictoires se développent : le consommateur peut désirer un produit puis son opposé (par exemple, des produits du terroir et des produits ethniques). En 2001, la consommation demeure dynamique avec un taux de croissance annuel de +2.6% (+2.5% en 2000), du fait de l'accroissement du pouvoir d'achat (+3.5%). Mais, avec une menace de reprise de l'inflation, un accroissement du taux d'épargne et un redémarrage de la montée du chômage, la conjoncture de 2002 est moins favorable.

4.2.1.2 Une consommation francilienne plus sensible à la conjoncture internationale

L'indice des ventes en volume dans les grands magasins en région parisienne est moins stable que celui sur la France entière : il diminue de 1990 à 1995, remonte jusqu'en l'an 2000, pour redescendre ensuite. Les variations les plus marquées ont lieu en 1995, où l'indice brut chute de près de -6% par rapport à l'année précédente (contre -0.26% au niveau national) et en 2000 (+7.89% contre +0.66% en France entière). Les évolutions en région francilienne et au niveau national ne sont en outre pas toujours de même sens : au début des années 1990, la consommation

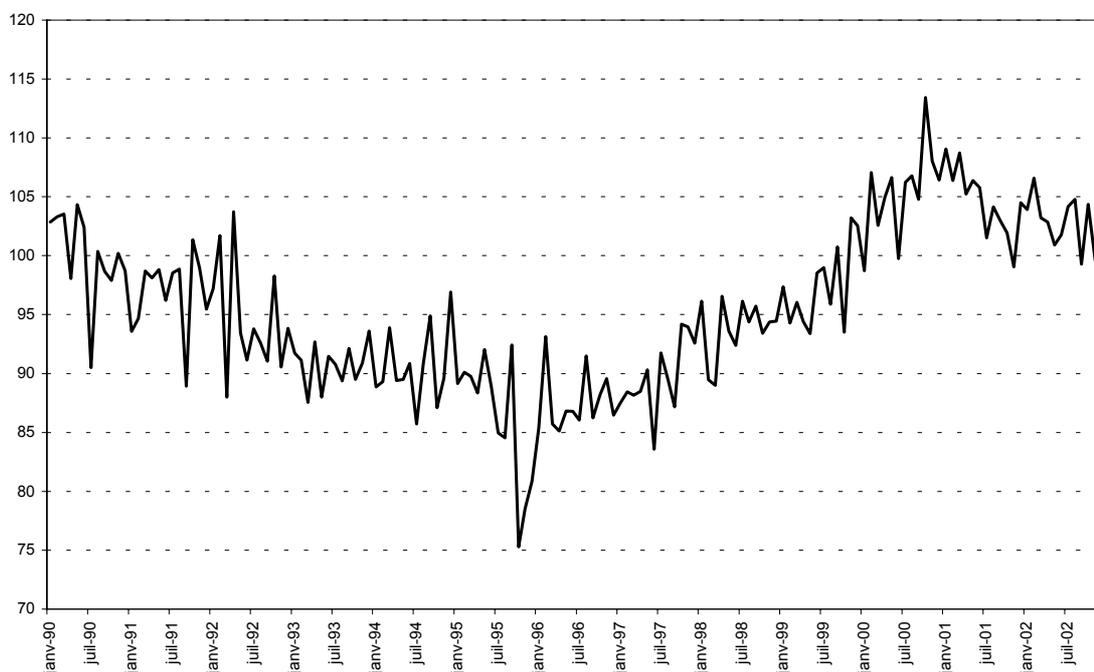
diminue en Ile-de-France mais augmente faiblement au niveau national. La reprise a lieu dès 1996 en Ile-de-France, une année plus tard au niveau national. En 1997, la reprise économique est caractérisée par une augmentation de la consommation des ménages, qui se poursuivra jusqu'au début de l'année 2001. La diminution au cours de l'année 2001 est en partie causée par le relâchement de la consommation des ménages face à une conjoncture nationale et internationale incertaine, mais la plus grande volatilité de la série relative à la région Ile-de-France provient surtout de la dynamique liée à l'affluence touristique en région parisienne, elle-même reliée au cours de l'euro. Les attentats du 11 septembre 2001 et les répercussions sur le tourisme ont entraîné une diminution bien plus accrue du commerce dans les magasins parisiens que dans le reste de la France.

Au début de l'année 2002, l'annonce de résultats économiques favorables aux Etats-Unis accroît la confiance des ménages, qui augmentent leur consommation malgré une détérioration de la situation sur le marché de l'emploi. En outre, la clientèle étrangère revient en région Ile-de-France, ce qui explique le différentiel de performance du commerce en faveur de l'Ile-de-France par rapport à la province. A cette époque, les experts pensent que la période de repli des ventes s'éloigne, et que se profile un arrêt de la dégradation de la situation économique. Néanmoins, en septembre 2002, une nouvelle diminution des ventes en volume des grands magasins franciliens a lieu, et fait écho au retournement conjoncturel associé à la détérioration du marché du travail d'une part, et aux conditions de change défavorables à l'euro d'autre part. Les incertitudes économiques mais aussi politiques favorisent en outre l'attentisme des chefs d'entreprise à l'embauche, et le manque de dynamisme de la consommation des ménages.

4.2.2 L'indice du chiffre d'affaires des grands magasins en région parisienne

Pour représenter les activités d'achat en Ile-de-France, la base de données franciliennes comprend une série d'indices du chiffre d'affaires des grands magasins en région parisienne, en valeur ou en volume, ayant une base 100 en 1990. Ces séries excluent les hypermarchés, les supermarchés, les magasins populaires, les jardineries et maisons de bricolage. Ces indicateurs sont fournis de façon mensuelle par le Centre d'Observation de Paris - Chambre de Commerce et d'Industrie de Paris (COE - CCIP), sous forme d'indices en volume bruts et d'indices en volume corrigés

des variations saisonnières et de l'effet des jours ouvrés.



gdmagr : Ventes dans les grands magasins de la région parisienne (COE CVS, base 100 en 90)

La consommation en volume dans les grands magasins de la région parisienne a faiblement augmenté entre 1990 et 2002 (+2.26% sur la période, soit +0.17% par an). Cette faible augmentation n'a pas été linéaire : trois sous-périodes peuvent être dégagées.

- Entre 1990 et 1995, l'indice du chiffre d'affaire en volume dans les grands magasins de la région parisienne a diminué de -14.44% , soit une diminution annuelle moyenne de -2.57% .
- Entre 1996 et 2000, il a augmenté en moyenne de $+3,88\%$, pour une augmentation globale de $+20,95\%$.
- Depuis la fin de l'année 2000, cet indice a tendance à diminuer.

La diminution importante de la série de ventes entre 1990 et 1995, puis sa forte remontée jusqu'en 2000 amènent à penser que la série n'est pas stationnaire. Les tests de racine unitaire sont appliqués sur la série désaisonnalisée (non log-linéarisée car il s'agit d'un indice) :

<i>gdmagrp</i>	Variable en niveau	Variable en différence
TEST ADF		
valeur de ρ	0.9392	-1.5848
t_{ADF}	-1.3636	-8.6127
VC_{ADF} à 5%	-2.9124	-2.9124
TEST KPSS		
Largeur de fenêtre	Statistique	
3	1.4186	0.0850
6	0.8420	0.1234
9	0.6063	0.1519
VC_{KPSS} à 5%	0.4572	0.4572
Conclusion	I(1)	I(0)

La série des indices des volumes de ventes dans les grands magasins de la région parisienne est intégrée à l'ordre 1.

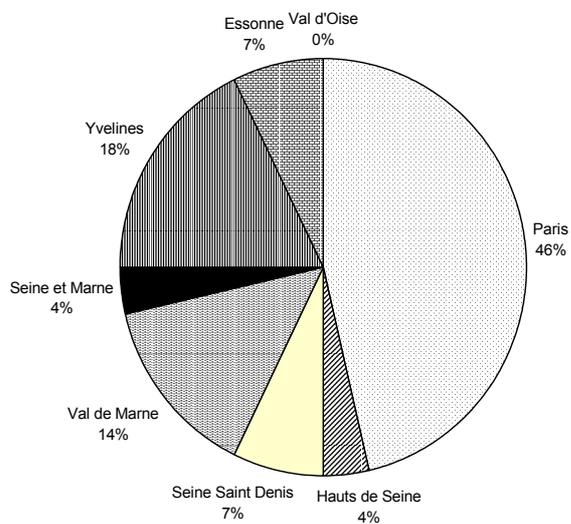
4.2.2.1 Une série représentative des ventes des grands magasins à Paris... plus "chahutée" que la même série au niveau national

La série est essentiellement représentative de l'activité des grands magasins parisiens, dans la mesure où 46% de ces magasins se trouvent à Paris et représentent 70% de la surface francilienne en milliers de mètres carrés.

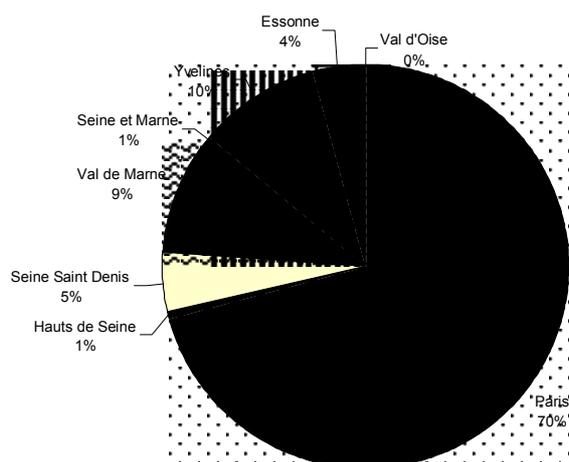
Les variations du chiffre d'affaires en volume des grands magasins sont amplifiées en Ile-de-France car 25% de l'activité y résulte du tourisme.

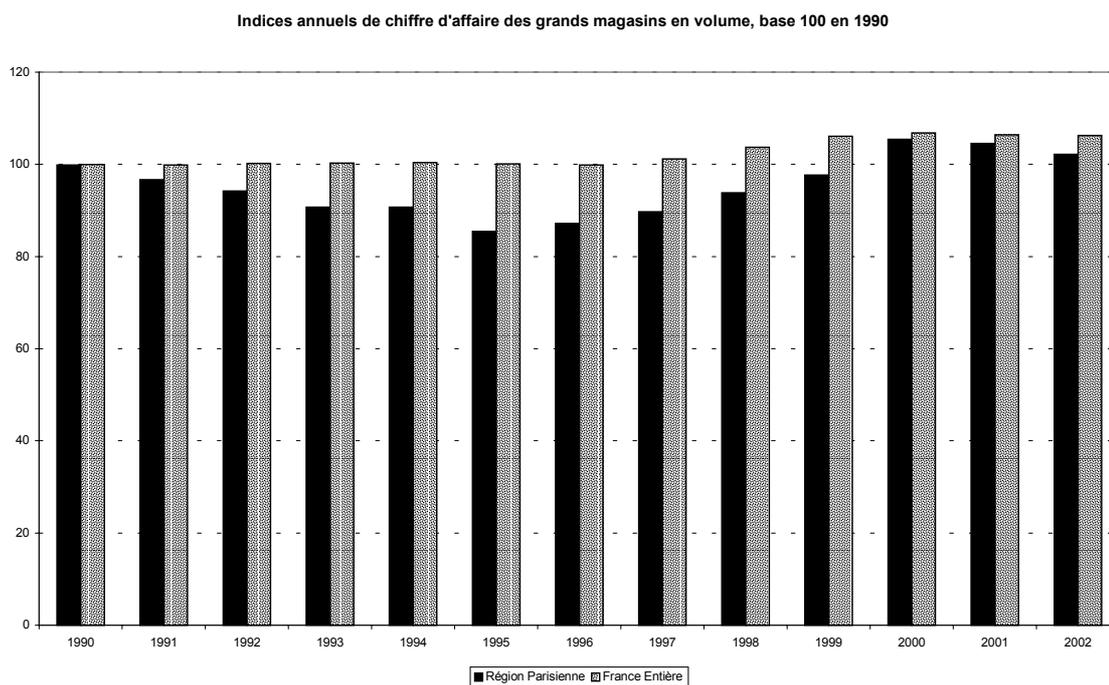
CHAPITRE 4. LES DONNÉES D'ENVIRONNEMENT

Répartition des grands magasins en nombre dans la région Ile de France au 1er septembre 2001



Répartition des grands magasins en surface dans la région Ile de France au 1er septembre 2001





4.3 Le tourisme en Ile-de-France

Le tourisme occupe une place importante dans le développement régional, en générant en particulier un grand nombre d'emplois.

La situation du tourisme en Ile-de-France, et son évolution depuis 1990, seront examinées dans une première sous-section. Les atouts de l'Ile-de-France, en matière d'offre d'hébergement et de sites culturels, mais aussi du fait de la modification des habitudes touristiques seront examinés à cette occasion.

Deux types de données mensuelles permettent d'appréhender le tourisme en Ile-de-France :

- les données relatives au trafic des grandes lignes de la SNCF dans les gares de Paris,
- le taux d'occupation des hôtels.

Ces séries seront examinées dans les deux dernières sous-sections.

4.3.1 La situation du tourisme en Ile-de-France depuis 1990

Entre 1990 et 2002, le tourisme en Ile-de-France a fortement évolué du fait de la croissance du tourisme international, de l'augmentation de l'offre touristique, mais aussi de l'évolution des comportements et des modes de vie. Le tourisme dans la région francilienne peut être qualifié de "tourisme urbain".

4.3.1.1 Le tourisme urbain

Deux types de clientèle pratiquent le tourisme urbain en France [152] :

- les résidents en France eux-mêmes,
- et les non-résidents.

Les caractéristiques du tourisme urbain en France des résidents En France, 30% des séjours des Français sont réalisés en milieu urbain, pour 22% des nuitées¹³. Le milieu urbain correspond au deuxième espace fréquenté après la campagne. Ce type de tourisme attire les personnes jeunes (moins de 25 ans), et dans des proportions moindres, la classe 25 - 34 ans et les plus de 50 ans. La voiture est un mode d'accès moins répandu que pour les autres destinations : 73% des séjours en milieu urbain contre 83% en moyenne. Les difficultés de circulation et de stationnement, mais aussi la qualité de la desserte aérienne et ferroviaire, poussent de nombreux Français à utiliser les transports en commun pour se rendre en Ile-de-France.

68% des séjours en milieu urbain ont pour motivation principale la visite à des familles ou des amis. L'importance de l'hébergement non marchand explique le fait que les dépenses dans les magasins représentent le principal poste de dépenses (46% du budget). Ces séjours ne sont pas réellement soumis à la saisonnalité, même s'ils se concentrent principalement au printemps et en été, du fait du nombre de ponts éventuels important. 65% des séjours ont une durée inférieure à 4 nuits, et la durée moyenne du séjour s'élève à 4.1 nuitées pour le tourisme urbain, contre 5.5 nuitées en moyenne pour l'ensemble des séjours.

¹³Les **nuitées** sont définies comme le nombre total de nuits passées par les clients dans un hôtel.

Les caractéristiques du tourisme urbain des non-résidents Selon l'enquête aux frontières de 1996, le milieu urbain correspond à l'espace privilégié des séjours des étrangers en France (48% des séjours pour 38% des nuitées). Chez les touristes résidant loin de la France, la quasi totalité des séjours se déroulent en ville et en particulier à Paris (82% des Américains et 90% des Japonais).

La majorité des touristes provient de l'Europe (79% des séjours) : les habitants d'Europe centrale, de l'Est et du Sud privilégient les séjours en ville.

Les séjours sont généralement de courte durée : en moyenne, ils ont une durée de 5.1 nuits mais 2/3 des séjours ont une durée inférieure ou égale à 3 nuits. Un séjour sur deux est un voyage d'agrément, plus d'un voyage sur 6 est un voyage à motif professionnel.

L'accès à la ville ne s'effectue par l'automobile que dans 40% des cas : l'avion transporte 19% des touristes en provenance de l'étranger, l'autocar est prisé par 22% des non-résidents (population jeune ou âgée, ou en provenance de l'Europe de l'Est).

Un voyage sur 4 correspond à un voyage en groupe, et 2/3 des nuitées sont réalisées dans un hébergement marchand (49% des nuitées sont réalisées en hôtels).

4.3.1.2 L'accroissement du tourisme international

Selon l'Organisation Mondiale du Tourisme, le nombre d'arrivées¹⁴ de touristes internationaux dans le monde est passé de 461 millions en 1990 à 715 millions en 2001. La France correspond à la première destination mondiale, avec 101.3 millions d'arrivées en 2001, dont 28.8 millions pour la seule région Ile-de-France.

D'ici à 2020, selon l'OMT, le nombre de touristes mondiaux devrait doubler, du fait de l'élargissement de l'Europe et de l'effet "euro", de la hausse du pouvoir d'achat des pays de l'Est et de l'Asie. La France resterait le principal pays de destination, et l'Ile-de-France devrait accueillir une grande part de ces touristes.

La fréquentation étrangère a plus que doublé entre 1989 et 2000 : 10 millions d'arrivées d'étrangers étaient constatées en 1989, contre 24 millions en 2000. Près de 14.4 millions d'étran-

¹⁴Le nombre d'arrivées correspond au nombre de clients différents qui séjournent une ou plusieurs nuits consécutives dans un même hôtel.

gers ont été accueillis en hôtellerie homologuée, la clientèle étrangère représente plus de la moitié du total des nuitées sur la région.

Pays de provenance	Part dans la clientèle étrangère des hôtels homologués
Royaume - Uni	21.3%
Etats - Unis	16.1%
Allemagne	8.4%
Italie	6.9%
Japon	6.5%
Espagne	6.3%

Source : ORTIF, memento de l'activité touristique d'Ile-de-France en 2001

Au niveau national, l'augmentation du nombre de vols aériens à faible prix (permise en partie par le développement d'internet), ou le développement du TGV, favorisent le tourisme en Ile-de-France. L'augmentation du temps libre grâce à la réduction du temps de travail et l'essor du tourisme urbain de courte durée correspondent à d'autres facteurs de croissance (11% des courts séjours ont pour destination l'Ile-de-France).

4.3.1.3 La présence de sites culturels

L'accroissement de la demande de tourisme en Ile-de-France s'explique par l'attractivité culturelle de la Région Ile-de-France. Plus de 42 millions de touristes (Français et étrangers) visitent chaque année l'Ile-de-France, pour les nombreux sites culturels qui y sont implantés. Le parc Disneyland Paris Resort, créé en 1992, est le site le plus visité (13 millions de visiteurs en 2002). Les sites culturels les plus prisés apparaissent dans le tableau page suivante.

A ces pôles culturels payant peuvent être ajoutés les lieux gratuits suivants : la cathédrale de Notre-Dame de Paris ou le Sacré Coeur. Le SDRIF¹⁵ souligne la forte évolution des comportements depuis 1990 : en plus de ces sites culturels, ce sont désormais les grandes manifestations, les foires, ou les fêtes qui attirent les touristes, de même que les nouveaux équipements liant la culture et les loisirs au sport et au commerce (par exemple, le Val d'Europe ou le Carré Sénart).

¹⁵SDRIF : Schéma Directeur Régional pour l'Ile-de-France.

Lieu culturel	Millions de visiteurs
Tour Eiffel	6.1
Centre Georges - Pompidou	5.3
Musée du Louvre	5.2
Domaine de Versailles	3.0
Cité des Sciences de la Villette	2.6
Musée d'Orsay	1.7

Source : ORTIF, mémento de l'activité touristique d'Ile-de-France en 2001

4.3.1.4 L'importance du tourisme d'affaires

La Région est aussi la première place mondiale en matière de tourisme d'affaires (11.3 millions de visiteurs par an en Ile-de-France).

Le "tourisme d'affaire" regroupe les déplacements individuels ou organisés, effectués pour des motifs professionnels, et dont la durée est au moins égale à 24 heures. Le marché du tourisme d'affaires comporte trois secteurs :

- les congrès, conventions d'entreprise, séminaires,
- les foires et salons,
- les voyages d'affaires individuels (prospection de clientèle, chantier, négociation, formation...).

18% des salons européens sont tenus en Ile-de-France, 5% des congrès mondiaux et 90% des manifestations nationales (Mondial de l'Automobile, Foire de Paris, Salon international de l'Agriculture...) ont lieu à Paris.

En plus des infrastructures traditionnelles, de nombreux hôtels ont développé une prestation de service "affaires" dans leur établissements, avec des salles de rencontre de grande capacité ou des salles de réunions avec un matériel spécifique. Selon l'International Congress and Convention Association, 80% des réunions d'entreprises se déroulent dans un hôtel urbain : l'assiduité des participants est garantie par la proximité entre le lieu d'hébergement et le lieu de travail. La région Ile-de-France concentre 43% de ces hôtels avec salle de réunion.

D'autres lieux de réunion se sont créés au cours de la décennie : des complexes spécialisés proposant espace de rencontre (salles de réunion et d'expositions), espace affaires (salles de travail disposant du matériel nécessaire), à proximité d'hôtels et restaurants. Par exemple, des groupes de loisirs (Center Parc, Club Med, Disneyland Paris, VVF Vacances...) misent sur l'accueil de réunions professionnelles pour diversifier leurs activités.

Type de site	Nombre	Surface à louer en m ²
Palais des congrès et parcs d'exposition	22	90 319
Centres de conférence ou salles de réunion	134	121 074
Hôtels avec salle de réunion	574	162 162
Autres lieux	337	202 833

Source : ORTIF, memento de l'activité touristique d'Ile-de-France en 2001

L'hôtellerie d'affaires représente 44.2% de l'activité touristique de l'Ile-de-France en 2002. Paris perd néanmoins des parts de marché dans ce type de tourisme du fait de l'émergence de nouveaux pôles dynamiques : Londres, Bruxelles, Vienne ou encore Singapour. Le développement des infrastructures de tourisme d'affaires en Ile-de-France a accusé un retard certain par rapport aux prévisions du SDRIF : seul le Stade de France a vu le jour en Plaine Saint-Denis. Le Stade de France est un grand équipement sportif, mais il dispose aussi de 10 lieux à géométrie variable (7500 m²) destinés aux congrès, séminaires ou expositions.

4.3.1.5 L'accroissement de l'offre de tourisme

La forte croissance du nombre de touristes a eu pour conséquence un accroissement de l'offre en hébergement, en particulier de l'offre marchande. Entre 1990 et 2001, l'hôtellerie francilienne a enregistré un gain d'environ 39 500 chambres, soit un accroissement de +39%. L'accroissement de l'offre d'hébergement s'est traduite par une augmentation de la part des hôtels non homologués en dehors de Paris ainsi que des hôtels 4 étoiles et 4 étoiles luxe, aux dépens des catégories intermédiaires (2 et 3 étoiles).

Catégorie	Nombre d'établissements	Nombre de chambres
Non homologuée	188	13230
1 étoile	228	8311
2 étoiles	1015	51698
3 étoiles	774	47357
4 étoiles	156	19924
4 étoiles luxe	9	1248

Source : ORTIF, memento de l'activité touristique d'Ile-de-France en 2001

L'accroissement de l'offre de tourisme a eu des répercussions en termes d'aménagement du territoire : l'essor de l'hôtellerie économique s'est réalisé principalement en dehors de Paris, autour du périphérique et à proximité des aéroports d'Orly et de Roissy, tandis que le développement de l'offre de luxe (quatre étoiles et plus) s'est ancré au coeur de Paris. La part des hôtels dans Paris a fortement diminué : selon le SDRIF, si 70% des chambres se situaient dans la capitale en 1994, cette proportion est tombée à 55% en 2002. L'essor des pôles hôteliers à la Défense ou à Marne la Vallée illustrent ce phénomène.

Les emplois liés au tourisme ont fortement augmenté depuis 1990 : en 2000, 4.6% de l'emploi salarié francilien est en lien avec l'activité touristique (environ 238 000 emplois salariés soit 13.6% de plus qu'en 1997). Ces différents emplois concernent l'hébergement, la restauration, les agences de voyages, les activités de loisirs et de culture, les parcs d'attraction... Le tourisme a en outre des impacts en terme d'emplois dans des domaines tels que le transport ou le commerce, non pris en compte.

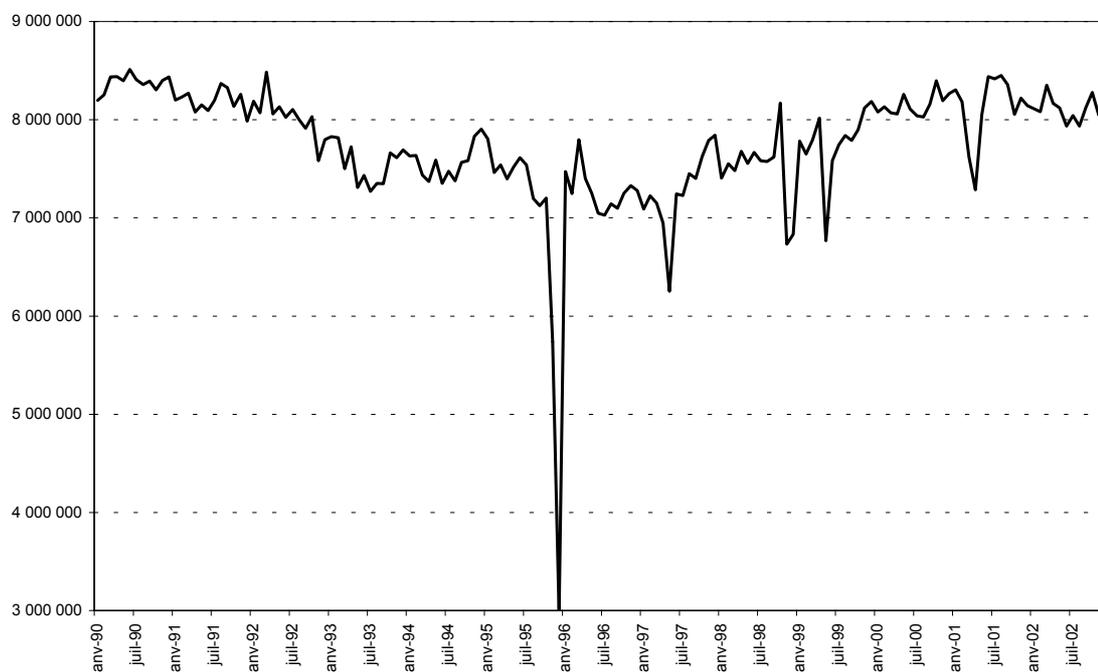
Les départements de l'Ile-de-France qui n'en possédaient pas se sont dotés de comités départementaux de tourisme, les offices de tourisme et syndicats d'initiative se sont multipliés. Ces outils de développement touristique permettent d'accueillir les touristes, et de diffuser l'information.

4.3.2 Le trafic grandes lignes de la SNCF à Paris

Les données mensuelles concernent le trafic des grandes lignes (Trains à Grande Vitesse et Trains Rapides Nationaux) et sont obtenues auprès de la SNCF à la fin de chaque année. Elles

sont exprimées en nombre de voyageurs.

Après désaisonnalisation nous obtenons le graphique suivant :



totgl : Nombre de voyageurs grandes lignes dans l'ensemble des gares de Paris (données SNCF CVS)

Sur l'ensemble de la période, le trafic grandes lignes de l'ensemble des gares de Paris est en moyenne de 7 750 000 voyageurs par mois. Au début de la décennie et depuis l'année 2000, il est supérieur à 8 millions de voyageurs mensuels. En 2002, plus de 97 millions de voyageurs grandes lignes sont passés par une gare parisienne.

Quatre périodes peuvent être envisagées :

- **Les années 1990 - 1995 de diminution du trafic des grandes lignes dans les gares de Paris** (-2.2% par an en moyenne, même en corrigeant la série des grèves de la fin d'année 1995).

La baisse de trafic grandes lignes dans l'ensemble des gares de Paris peut être expliquée par les vagues d'attentats, la très mauvaise perception de Socrate¹⁶ pendant une longue

¹⁶Créé en 1972, l'ancien système de vente et de réservation de la SNCF ne permettait plus de répondre aux besoins commerciaux générés par le développement du TGV, et son extension au niveau international. Pour

période de mise en place en 1993-1994, la grève de 1995 et, liant tout cela, une image du service assombrie et qui ne peut être restaurée que progressivement.

- **La stagnation des années 1995-1997**, à un niveau faible de 7,3 millions de voyageurs annuels.
- **La reprise pendant les années 1998-2000**, à un rythme élevé (en moyenne +3.9% par an).

Entre 1997 et 2001, le trafic grandes lignes a connu une forte croissance (+ 17% au niveau national d'après le Rapport du Sénat de 2002 [148]), tirée essentiellement par le TGV (+35% au niveau national). La mise en service réussie du TGV Méditerranée en juin 2001 a permis de conserver un fort taux de croissance du trafic TGV (+7.6% en 2001), qui représente plus de la moitié de l'ensemble du trafic voyageurs de la SNCF. Les trafics relatifs aux autres lignes de TGV ont augmenté plus modérément ou se sont stabilisés (TGV Nord-Europe, Eurostar).

Une partie de cette croissance peut être expliquée par la mise en place d'une nouvelle gamme tarifaire en 1997, et la création des "prix découverte 12 - 25 ans" (réduction allant jusqu'à -60% du tarif normal), les "prix découverte à deux" (réduction de -25% sur le prix de base, dans la majorité des trains, dans la limite des places disponibles), les "prix découverte séjour SNCF" sur un aller - retour en TGV commencé en période bleue (-25%)... En 1999 a été lancé le programme de fidélisation "Grand Voyageur", qui permet de cumuler des points à chaque voyage, points qui pourront être convertis en billets de trains gratuits, permet un échange facile des réservations...

- **La stagnation depuis 2000, à un niveau élevé**, comparable celui du début des années 1990.

Le trafic des autres trains "grandes lignes" a fortement décliné depuis 2001, du fait de la suppression de relations lors de la mise en service du TGV Méditerranée et de la requalification de certaines relations vers les services régionaux.

répondre à d'importants volumes de places de trains, la SNCF s'est dotée en janvier 1993 d'un nouvel outil de distribution, appelé SOCRATE (Système Offrant à la Clientèle des Réservations d'Affaires et de Tourisme en Europe). Les multiples difficultés techniques ont causé de nombreux désagréments pour la clientèle.

Nous avons appliqué les tests de racine unitaire sur la série désaisonnalisée log-linéarisée, en niveau et en différence.

$\ln(totgl)$	Variable en niveau	Variable en différence
TEST ADF		
valeur de ρ	0.6018	-1.8736
t_{ADF}	-4.4547	-8.9834
VC_{ADF} à 5%	-2.9124	-2.9124
TEST KPSS		
Largeur de fenêtre	Statistique	
3	0.6016	0.0242
6	0.4251	0.0435
9	0.3316	0.0614
VC_{KPSS} à 5%	0.4572	0.4572
Conclusion	I(0)	I(0)

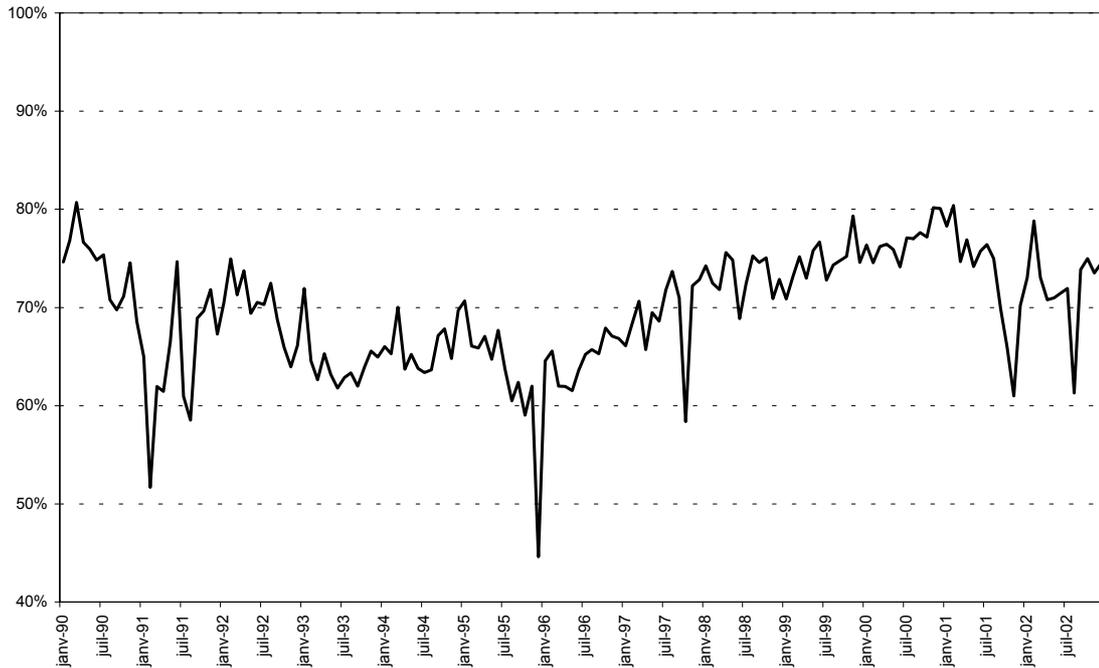
Malgré ses fortes variations, la série est stationnaire sur la période 1990 - 2002.

4.3.3 Le taux d'occupation des hôtels à Paris

Le taux d'occupation des hôtels correspond au rapport du nombre de chambres occupées par le nombre de chambres effectivement disponibles. Il est issu des enquêtes de fréquentation dans l'hôtellerie réalisées au sein d'une convention entre la Direction du Tourisme et l'INSEE. L'objectif de ces enquêtes est la collecte de données sur la fréquentation des hébergements touristiques : capacité d'accueil, taux d'occupation, nombre de nuitées selon le pays de résidence du tourisme, durées moyennes des séjours... Elles sont réalisées chaque mois au sein d'un échantillon représentatif des hôtels homologués (de 0 à 4 étoiles luxe).

Les données relatives au taux d'occupation des hôtels de Paris sont disponibles chaque mois sur le site de l'INSEE.

Après désaisonnalisation, nous obtenons le graphique suivant :



txhot75 : Taux d'occupation en % des hôtels à Paris (données INSEE CVS)

En moyenne sur la période 1990 - 2002, 70% des chambres des hôtels de Paris étaient occupées. Entre 1990 et 2002, le taux d'occupation des hôtels de Paris a diminué d'environ 2 points, i.e. est resté relativement stable.

Cette stabilité ne se vérifie pas à un niveau plus fin :

- Le taux d'occupation a diminué de 11 points entre 1990 et 1995 du fait de la guerre du Golfe (1991), de la récession économique (1993), des attentats à Paris et des grèves en 1995.
- L'augmentation du taux d'occupation des hôtels entre 1996 et 2000 s'explique par la reprise économique au niveau international. La diminution brutale de la série en décembre 1996 s'explique par l'attentat sur le métro, en gare de Port-Royal. Cette période est en outre marquée par les Journées Mondiales de la Jeunesse en 1997, et la Coupe du Monde de football en France en 1998.
- Depuis 2001, le remplissage des hôtels diminue à Paris. Les attentats du 11 septembre 2001 aux Etats-Unis ont un impact négatif sur l'activité touristique. La chute du dollar

par rapport à l'euro, la peur des attentats, la crise économique ou encore le développement du SRAS (Syndrome Respiratoire Aigu Sévère) ont limité les voyages des Japonais en Ile-de-France. Au cours de cette période, la clientèle européenne (Anglais, Allemands, Italiens et Espagnols) revient en Ile-de-France.

La forte diminution de la série au début des années 1990 pousse à vérifier la non-stationnarité éventuelle de la série. Les tests de racine unitaire ont été appliqués à la série des taux d'occupation des hôtels à Paris (données corrigées des variations saisonnières) sur la variable en niveau, mais aussi sur la variable en différence. La série correspond à un rapport, elle n'est pas log-linéarisée. La série du taux d'occupation des hôtels à Paris est intégrée à l'ordre 1.

<i>txhot75</i>	Variable en niveau	Variable en différence
TEST ADF		
valeur de ρ	0.8281	-1.4190
t_{ADF}	-2.6461	7.4991
VC_{ADF} à 5%	-2.9124	-2.9124
TEST KPSS		
Largeur de fenêtre	Statistique	
3	1.2999	0.0496
6	0.8254	0.0702
9	0.6246	0.0831
VC_{KPSS} à 5%	0.4572	0.4572
Conclusion	I(1)	I(0)

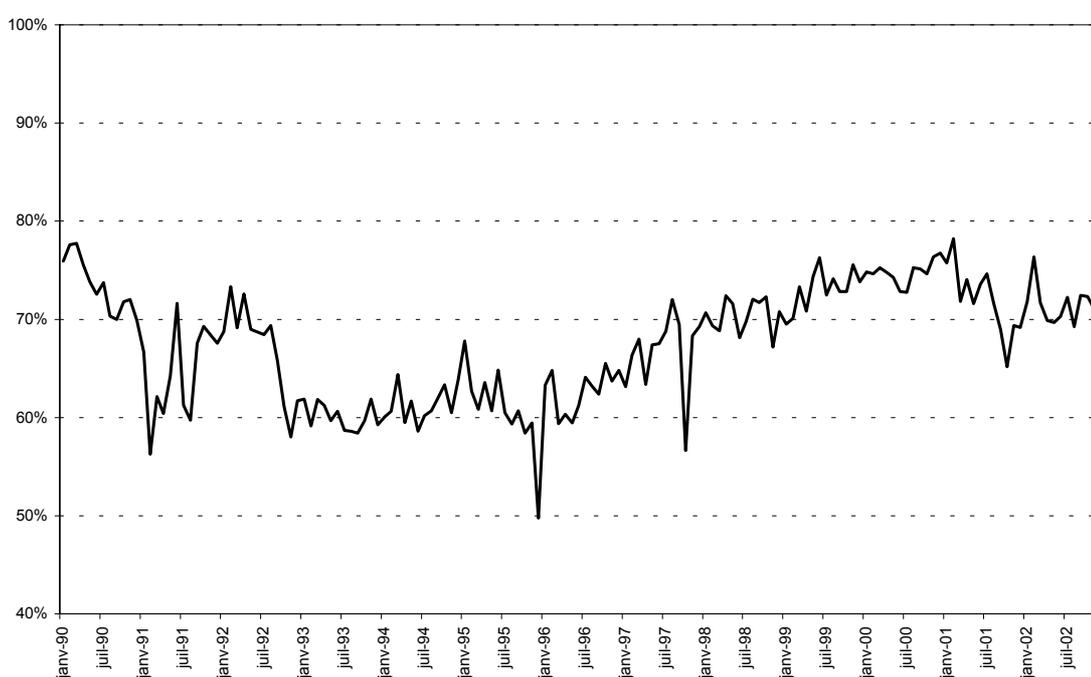
4.3.4 Le taux d'occupation des hôtels en Ile-de-France

En moyenne sur la période 1990 - 2002, 68% des chambres des hôtels d'Ile-de-France étaient occupées. Entre 1990 et 2002, le taux d'occupation des hôtels d'Ile-de-France a diminué d'environ 2 points, i.e. est resté relativement stable.

L'équivalent en termes d'arrivées et de nuitées est présenté dans le tableau suivant

Fréquentation dans l'hôtellerie homologuée d'Ile-de-France en 2002		
	Niveau	Evolution 2002/2001
Arrivées	28.8 millions	+1.9%
Nuitées	60.9 millions	+2.1%
Durée moyenne du séjour	2.10 jours	-0.01 jour
Taux d'occupation moyen	71.5%	-0.3 points

Source : Enquêtes de fréquentation touristique (INSEE DT ORTIF)



txhotidf : Taux d'occupation en % des hôtels en Ile-de-France (données INSEE CVS)

L'évolution du taux d'occupation des hôtels en Ile-de-France présente des caractéristiques assez similaires à celle de Paris, quoique moins marquées :

- Diminution jusqu'en 1995, forte durant les premières années.
- Reprise entre 1996 et 2000.
- Légère décroissance depuis 2001.

Les variations moins marquées de la série des taux d'occupation en Ile-de-France par rapport à ce même indicateur au niveau de Paris uniquement s'expliquent par la part moins importante des touristes non-résidents en dehors de Paris. Les variations de l'euro ou les menaces internationales ont un impact plus faible en Ile-de-France.

La série du taux d'occupation des hôtels en Ile-de-France correspond à une série intégrée à l'ordre 1. Le tableau suivant présente les résultats des tests de racine unitaire appliqués à cette série.

<i>txhotidf</i>	Variable en niveau	Variable en différence
TEST ADF		
valeur de ρ	0.9154	-1.3212
t_{ADF}	-1.7755	-9.6670
VC_{ADF} à 5%	-2.9124	-2.9124
TEST KPSS		
Largeur de fenêtre	Statistique	
3	1.4719	0.1176
6	0.8973	0.1547
9	0.6621	0.1779
VC_{KPSS} à 5%	0.4572	0.4572
Conclusion	I(0)	I(0)

4.4 La concurrence automobile en Ile-de-France

La concurrence (ou complémentarité) de l'automobile en Ile-de-France est appréhendée par trois types de statistiques :

- les consommations de carburants,
- les prix de vente de ces carburants ramenés au prix du ticket RATP,
- et les indicateurs de circulation dans Paris, et dans l'Ile-de-France hors Paris.

4.4.1 La consommation de carburants en Ile-de-France

La consommation de carburant est mesurée par la vente de carburant en mètre cube, des sociétés pétrolières à leur clientèle (consommateurs directs ou revendeurs).

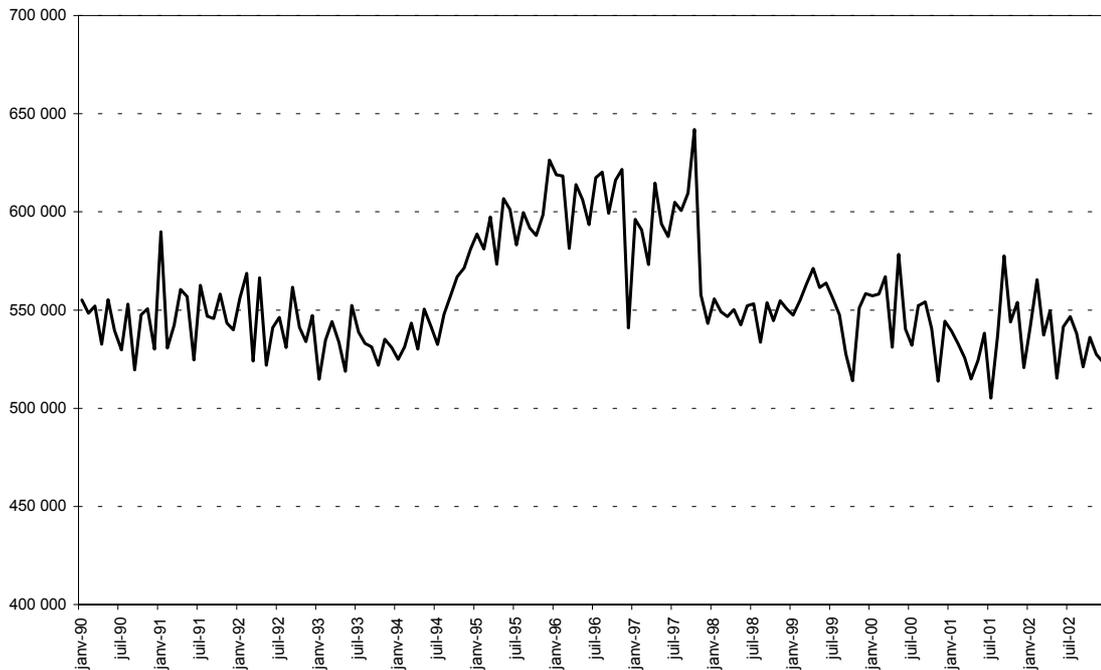
Les quantités déclarées comme vendues dans un département peuvent être consommées dans un autre département.

Plusieurs séries sont suivies chaque mois et concernent :

- la vente de super et de sans plomb, pour tous les départements franciliens,
- la vente de gazole, pour tous les départements franciliens.

Ces séries sont fournies par le Comité Professionnel du Pétrole, qui reconstitue tous les mois les ventes dans tous les départements français.

Après sommation de ces séries et désaisonnalisation, nous obtenons la série *carbifdf*, qui représente la vente de carburant en m³ des différents types de carburants au niveau de l'Ile-de-France.



carbifdf : Ventes de carburants en Ile-de-France (données CPDP CVS) en m³

L'ensemble de la consommation des carburants routiers a augmenté de +11% entre 1990

et 1995, puis a diminué de -7% jusqu'en 2000. Cette baisse de la consommation d'essence s'explique, selon la DRIRE [48], par trois facteurs :

- la diminution du taux de motorisation des ménages,
- la baisse du kilométrage moyen parcouru
- et la diminution de la consommation unitaire des véhicules. La consommation moyenne de l'ensemble des carburants, exprimée en litres aux 100 kilomètres, a diminué de -9.5% entre 1990 et 2001. Elle s'explique en partie par l'augmentation de la part des voitures diesel dans le parc.

La forte augmentation de la vente des carburants au milieu des années 1990 est imputable à l'augmentation de l'utilisation de l'automobile dans un contexte incertain de fortes grèves dans les transports collectifs et d'alertes aux attentats.

Les tests de racine unitaire ont été appliqués à la série désaisonnalisée log-linéarisée $\ln(carbidf)$, en niveau et en différence première.

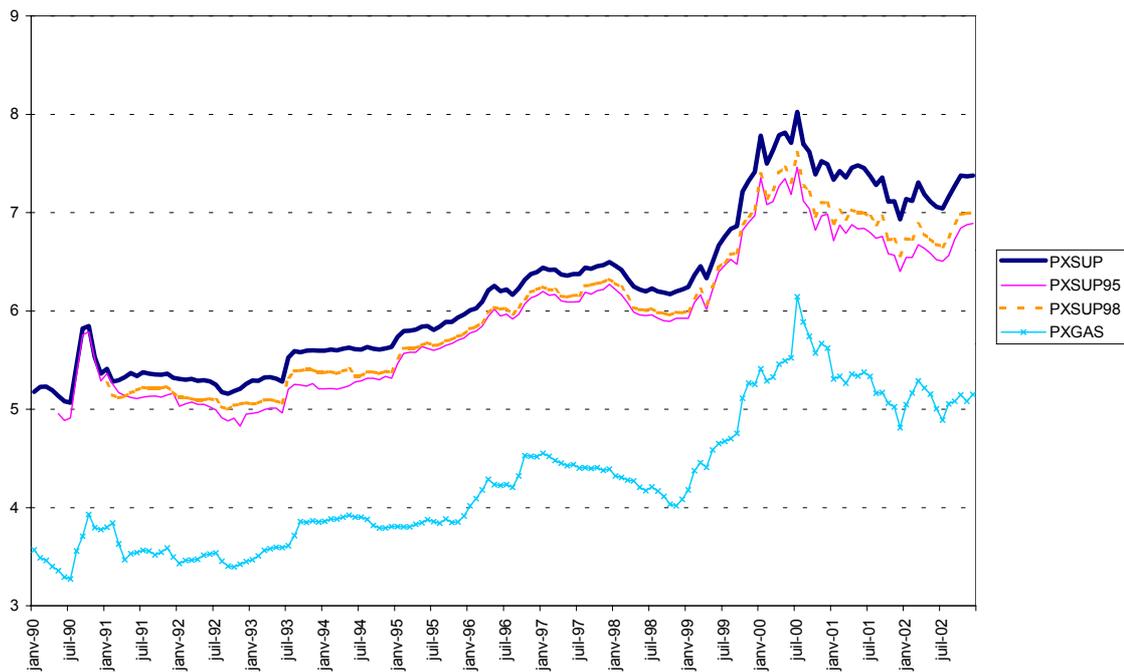
$\ln(carbidf)$	Variable en niveau	Variable en différence
TEST ADF		
valeur de ρ	0.9900	-1.4040
t_{ADF}	-0.3730	-8.7263
VC_{ADF} à 5%	-2.9124	-2.9124
TEST KPSS		
Largeur de fenêtre	Statistique	
3	3.0593	0.0449
6	1.8086	0.0672
9	1.3044	0.0820
VC_{KPSS} à 5%	0.4572	0.4572
Conclusion	I(1)	I(0)

La série de ventes des carburants est intégrée à l'ordre 1.

4.4.2 Le rapport du prix du super au prix du ticket T

Le prix des carburants est un prix moyen relatif à l'ensemble de l'Ile-de-France, l'indice "DHYCA" étant mesuré sur un panel de stations services franciliennes. Ce prix est exprimé en francs courants par hectolitre, toutes charges comprises, jusqu'au 1^{er} janvier 2002. A partir de janvier 2002, les prix sont exprimés en euros courants par hectolitre, toutes charges comprises. Dans un premier temps, pour avoir des séries homogènes dans le temps, les prix en euros ont été divisé par le cours officiel de l'euro, à savoir 1 euro = 6.55957 francs français. Le Comité Professionnel Du Pétrole distingue plusieurs séries, relatives :

- au prix du super plombé,
- au prix du super sans plomb 95,
- au prix du super sans plomb 98,
- au prix du gazole.



Prix des différents carburants en francs (données CPDP CVS)

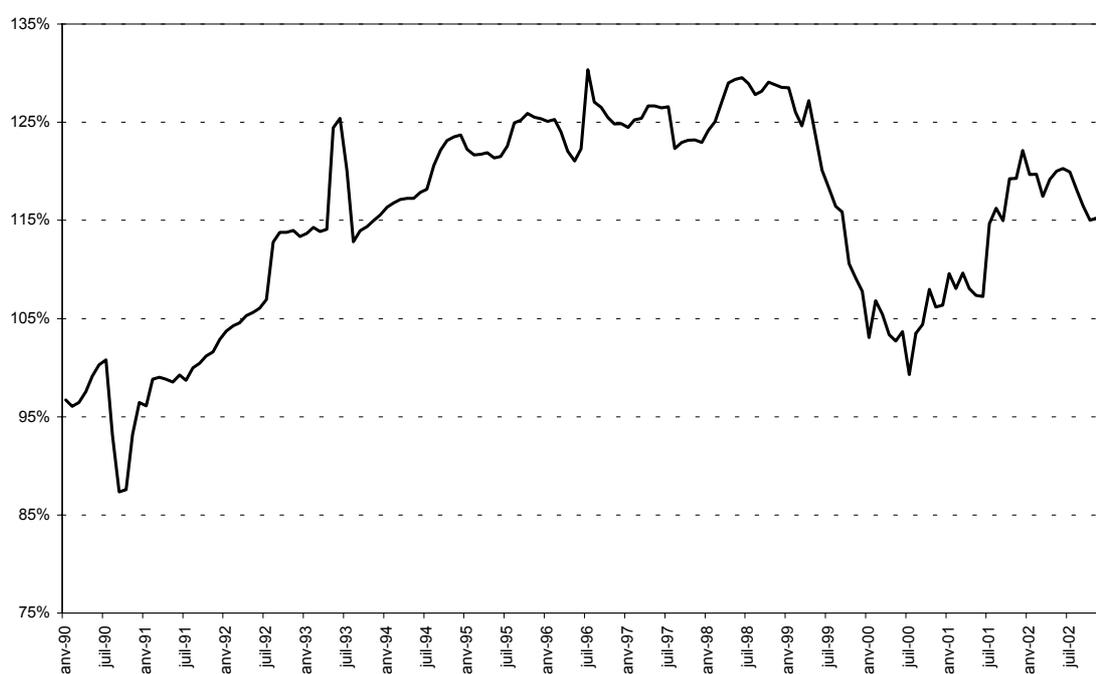
Le choix du carburant a peu d'importance : les variations des prix des différents types de carburants ont suivi une évolution assez comparable, comme le montre le graphique précédent.

Le choix du super provient du fait que c'est la seule série complète depuis janvier 1990.

Le prix RATP choisi correspond au prix du ticket T : il s'agit d'un prix de référence des transports collectifs, au même titre que le prix de la baguette comme prix de référence du pain.

L'indice de compétitivité des transports en commun correspond au rapport du prix du ticket T par rapport au prix du super.

Après désaisonnalisation, nous obtenons la série *compet* représentée dans le graphique suivant.



compet : Compétitivité des TC (rapport des prix du ticket T et du super) (données CPDP RATP CVS)

Sur l'ensemble de la période, l'indice de compétitivité des transports en commun par rapport à l'automobile s'est globalement dégradé : il a augmenté de +23.71%, soit un accroissement annuel de +1.65%. Cette forte augmentation provient de la hausse continue du prix du ticket T durant l'été, tandis que le prix du super est soumis à des variations plus erratiques. Le prix des carburants dépend du prix du brut, mais aussi du taux de change du dollar, des marges de raffinage et de distribution, de la fiscalité..

Les variations du prix du pétrole brut Plusieurs éléments relatifs à la situation internationale concourent aux variations importantes des cours du pétrole brut.

- L’augmentation progressive des cours du pétrole brut jusqu’en 1997 est imputable aux éléments suivants :
 - La guerre du Golfe en 1990 - 1991 a conduit à un embargo sur le pétrole irakien dès le mois d’août 1990.
 - En août 1991, la crise politique en Russie entraîne une hausse momentanée des prix.
 - En 1993, la crainte du retour de l’Irak sur le marché cause un rebond éphémère.
 - En 1995, les américains imposent un embargo sur le pétrole irakien.
 - En 1996, une hausse des cours du pétrole est imputable à une forte vague de froid, à la maladie du roi Fahd d’Arabie, ainsi qu’aux réductions des stocks d’essence aux Etats-Unis .
- Le deuxième contre-choc pétrolier a lieu en 1997. Entre janvier et mai, le cours du brut passe de 24 \$ le baril à 17.7 \$. Ce phénomène est amplifié par la crise asiatique durant l’été, qui conduit à une diminution de la demande des pays d’Asie, ainsi que par la reprise des exportations irakiennes en août 1998, qui accroissent l’offre. De nombreux pays de l’OPEP décident de réduire leurs quotas de production pour enrayer la chute des prix.
- A partir de 1999, le prix du pétrole brut recommence à croître.

L’inégale répercussion des variations du cours du baril sur les prix de vente des carburants Selon l’INSEE, le prix des carburants est plus sensible à une hausse qu’à une baisse du prix du pétrole brut. La répercussion des hausses des prix du baril sur le prix à la pompe est plus rapide que la répercussion des baisses.

La première hypothèse explicative serait que si le prix du brut baisse, les coûts induits par les destockages se répercutent sur les prix à la pompe, et compensent la baisse du prix du baril.

La deuxième hypothèse revient à prendre en compte la faiblesse des marges des distributeurs. Du fait de l’importance de la distribution par les grandes surfaces, qui font de l’essence un produit

d'appel, et de la concurrence entre les distributeurs, les marges sur l'essence sont très faibles. En conséquence, une hausse des prix doit immédiatement être répercutée.

Selon l'INSEE, "face à une baisse des coûts, la décision des distributeurs de réduire leurs prix est tributaire de considérations stratégiques", ce que conteste l'Union Française des Industries Pétrolières.

L'accroissement de cet indice au rythme de +4.38% entre 1990 et 1995, puis sa diminution de -1.34% à partir de 1996 amènent à se demander si la série possède une racine unitaire.

Les tests de racine unitaire ont été appliqués à la série en niveau et à la série en différence première.

Comme il s'agit d'un indice, la log-linéarisation préalable ne sera pas nécessaire pour les modélisations ultérieures.

Le tableau suivant indique la présence d'une racine unitaire dans la série en niveau, et son absence pour la variable en différence. La série est intégrée à l'ordre 1.

<i>compet</i>	Variable en niveau	Variable en différence
TEST ADF		
valeur de ρ	0.9633	0.0467
t_{ADF}	-2.0297	-8.8403
VC_{ADF} à 5%	-2.9124	-2.9124
TEST KPSS		
Largeur de fenêtre	Statistique	
3	1.2609	0.2030
6	0.7486	0.2183
9	0.5436	0.2052
VC_{KPSS} à 5%	0.4572	0.4572
Conclusion	I(1)	I(0)

4.4.3 Le volume et la vitesse de circulation dans Paris

Pour observer la circulation automobile, des capteurs (boucles électromagnétiques) sont placés dans la chaussée, et mesurent en permanence les caractéristiques du trafic sur chaque voie. Les informations recueillies concernent les débits et taux d'occupation, elles sont transmises aux contrôleurs de carrefours et, toutes les 5 secondes, au poste central de régulation.

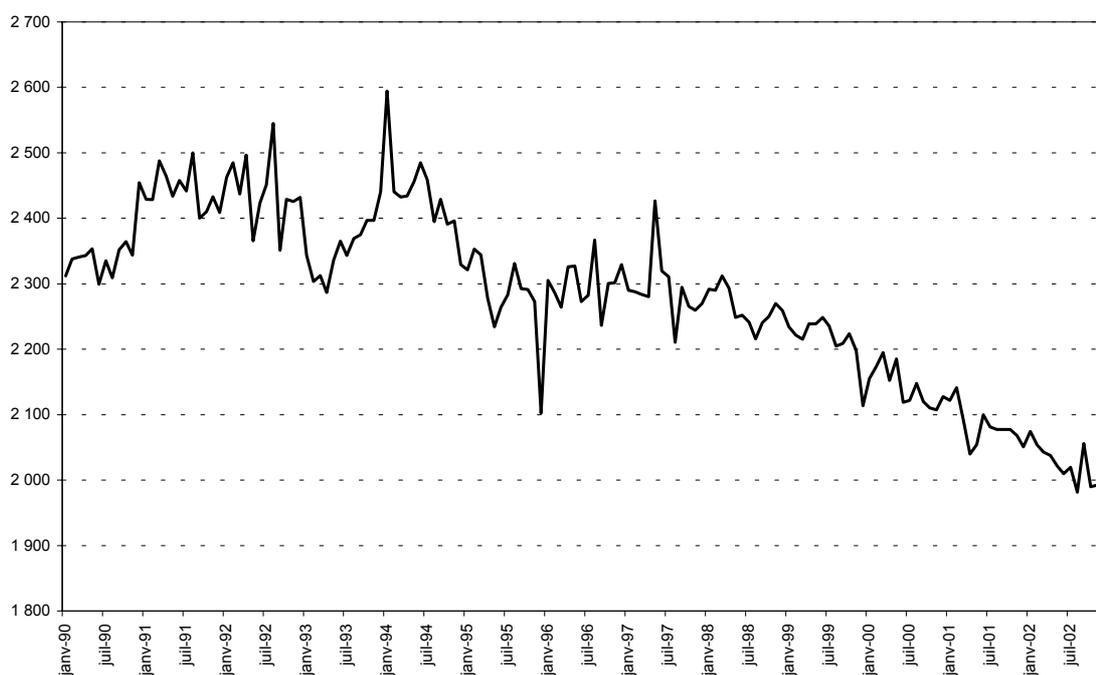
4.4.3.1 Le volume de circulation dans la zone centrale régulée de Paris

Le volume de circulation est observé les jours ouvrables, entre 7 heures du matin et 21 heures, dans la zone centrale régulée de Paris. Il est exprimé en milliers de véhicules par kilomètres et par heure.

La zone centrale régulée de Paris ou "réseau instrumenté" correspond à la zone où la circulation est fluidifiée grâce à une synchronisation informatisée des feux, à l'aide de SURF (Système Urbain de Régulation des Feux) et de SAGE (Système d'Aide à la Gestion des Embouteillages). Elle comprend environ 200 carrefours. De petite taille, elle est influencée par des perturbations telles que les travaux, la forte modification des flux de véhicules sur certaines voies, ...

La série est calculée par l'Observatoire des Déplacements de la Mairie de Paris. Les données mensuelles brutes utilisées sont confidentielles, mais les données en moyenne glissante sur 12 mois sont disponibles trimestriellement dans le bulletin trimestriel de l'Observatoire des Déplacements à Paris.

Entre 1990 et 2002, le volume de circulation dans la zone centrale régulée de Paris a diminué de -13.88% , soit une diminution annuelle moyenne de -1.14% . Cette forte diminution a été réalisée essentiellement après 1995 : entre 1990 et 1995, le taux d'accroissement annuel moyen du volume de circulation est égal à -0.48% , contre -1.82% entre 1996 et 2002. Elle amène à se poser la question de la stationnarité de la série.



circ75 : Circulation sur le réseau instrumenté de Paris (données de l'Observatoire des Déplacements CVS) par km et par heure.

$\ln(circ75)$	Variable en niveau	Variable en différence
TEST ADF		
valeur de ρ	1.0071	-1.0138
t_{ADF}	0.2462	-10.4139
VC_{ADF} à 5%	-2.9124	-2.9124
TEST KPSS		
Largeur de fenêtre	Statistique	
3	3.3135	0.1860
6	1.9619	0.2461
9	1.4181	0.2677
VC_{KPSS} à 5%	0.4572	0.4572
Conclusion	I(1)	I(0)

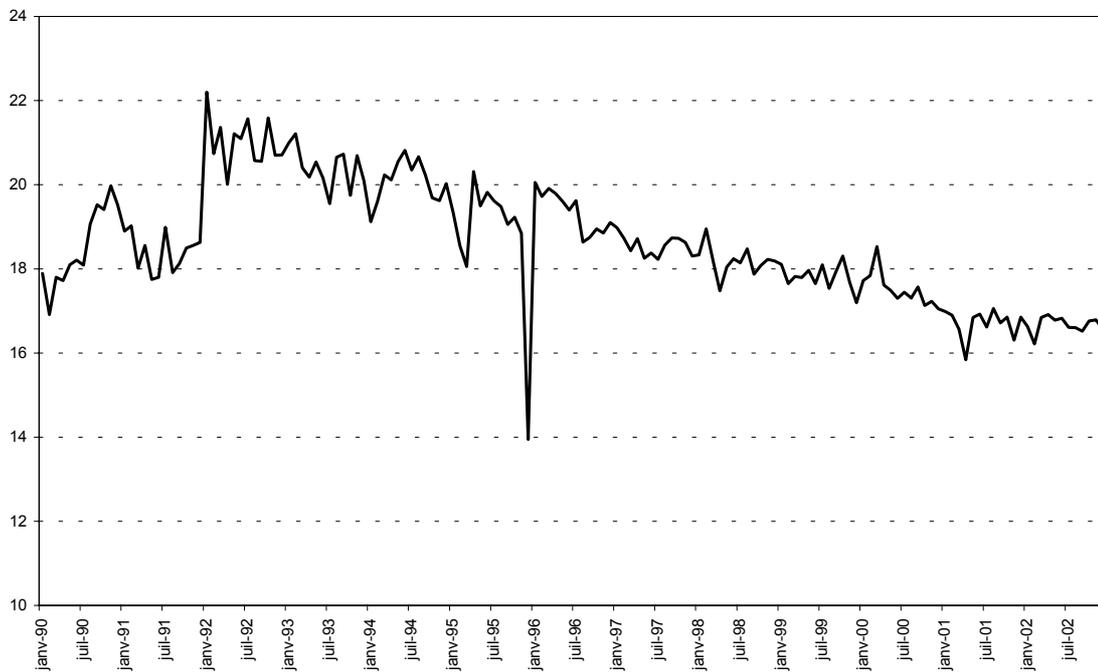
La série du volume de circulation dans Paris est intégrée à l'ordre 1.

4.4.3.2 La vitesse de circulation dans la zone régulée de Paris

La vitesse de circulation est observée les jours ouvrables, entre 7 heures du matin et 21 heures, dans la zone centrale régulée de Paris. Elle est exprimée en kilomètres-heure. La série est calculée par l'Observatoire des Déplacements de la Mairie de Paris.

Les données mensuelles brutes utilisées sont confidentielles, mais les données en moyenne glissante sur 12 mois sont disponibles trimestriellement dans le bulletin trimestriel de l'Observatoire des Déplacements à Paris.

Après désaisonnalisation, la série *speed75* est représentée dans le graphique suivant.



speed75 : Vitesse de circulation sur le réseau instrumenté de Paris (données de l'Observatoire des Déplacements CVS) en km/h

Entre 1990 et 2002, la vitesse dans Paris a globalement diminué de -9.91% , ce qui correspond à une diminution moyenne de -0.80% par an.

Cette diminution n'a pas été linéaire : entre 1990 et 1995, la vitesse de circulation a globalement augmenté de $+1.91\%$, au taux d'accroissement annuel moyen de $+0.32\%$, du fait de la mise en place des axes rouges. Sur ces axes, il est interdit de stationner, parfois même de s'arrêter, ce qui a eu pour conséquence un accroissement de la vitesse moyenne.

Entre 1996 et 2002, le taux d'accroissement moyen annuel de la vitesse de circulation dans Paris est égal à -2.13% . Malgré la diminution du volume de circulation, la création de couloirs de bus a diminué la voirie réservée aux véhicules particuliers, entraînant un accroissement des embouteillages, et une diminution de la vitesse dans Paris.

Les tests de racine unitaire ont été appliqués à la série log-linéarisée. Présentés dans le tableau suivant, ils concluent à une intégration à l'ordre 1.

$\ln(\text{speed75})$	Variable en niveau	Variable en différence
TEST ADF		
valeur de ρ	0.9263	-1.1608
t_{ADF}	-1.4340	-8.3030
VC_{ADF} à 5%	-2.9124	-2.9124
TEST KPSS		
Largeur de fenêtre	Statistique	
3	2.4738	0.0769
6	1.5096	0.1104
9	1.1091	0.1551
VC_{KPSS} à 5%	0.4572	0.4572
Conclusion	I(1)	I(0)

4.4.4 L'indice d'évolution de la circulation en zone agglomérée de la région Ile-de-France sur les autoroutes et routes nationales

Les coefficients mensuels d'évolution sont calculés par le Service Interdépartemental d'Exploitation Routière (SIER), appartenant à la Direction Régionale de l'Équipement en Ile-de-France. Le SIER est chargé depuis 1988 de l'exploitation de l'ensemble du réseau des voies rapides d'Ile-de-France. 300 kilomètres de ces voies rapides sont totalement équipés sur la partie Est du réseau, un niveau de service comparable sera atteint sur la partie Ouest avant la fin 2003. L'indice de trafic est calculé toutes les 6 minutes à partir d'environ 3 000 points de comptage.

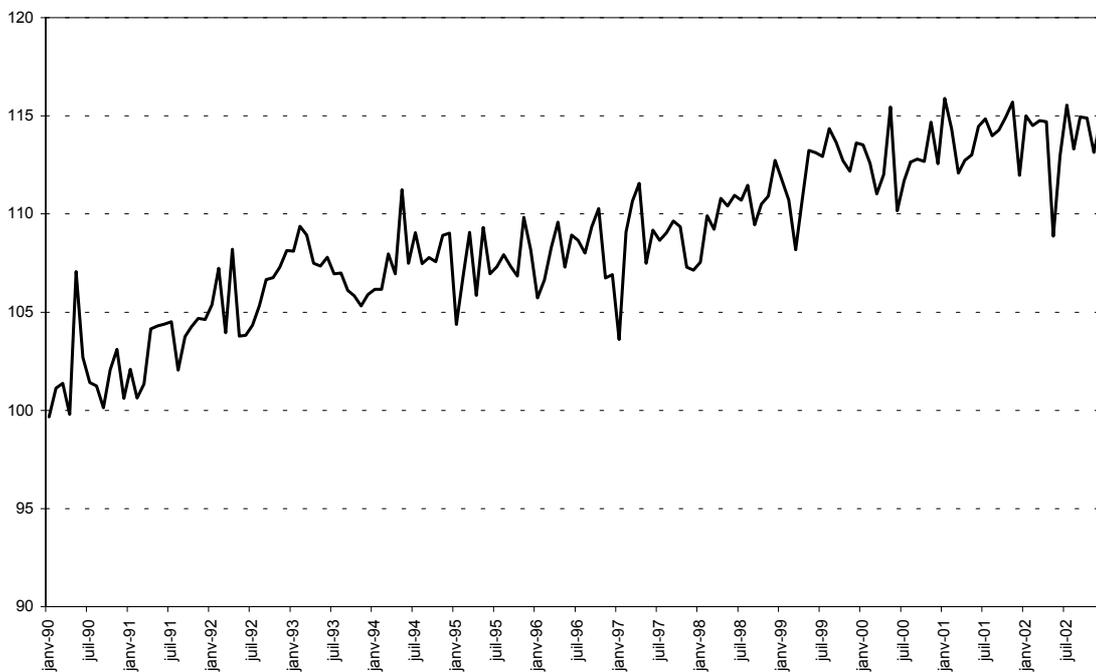
Deux zones sont distinguées :

- la zone agglomérée, délimitée par la Francilienne, Paris exclus,
- la rase campagne.

pour 2 types de routes :

- les routes nationales,
- et les autoroutes.

Nous avons choisi de conserver l'indice d'évolution de la circulation en zone agglomérée de la région Ile-de-France sur les autoroutes et routes nationales. Ce choix est motivé par le fait que les transports en commun ne desservent pas la rase campagne, et ne peuvent constituer une alternative à l'automobile.



circagar : Evolution de la circulation en zone agglomérée francilienne sur les autoroutes et routes nationales (données DREIF CVS)

Entre 1990 et 2002, la circulation en zone francilienne agglomérée sur les routes nationales et autoroutes a augmenté de plus de 12%, au rythme de +0.88% par an. L'augmentation, légèrement plus rapide au début de la décennie, s'explique en partie par la légère augmentation des distances parcourues, par l'amélioration du réseau routier en banlieue, ainsi que l'accroissement des déplacements en banlieue, dont la population s'est accrue.

Les tests de racine unitaire ont été effectués sur la série désaisonnalisée log-linéarisée.

<i>circagar</i>	Variable en niveau	Variable en différence
TEST ADF		
valeur de ρ	0.9550	-1.9071
t_{ADF}	-1.2250	-7.0338
VC_{ADF} à 5%	-2.9124	-2.9124
TEST KPSS		
Largeur de fenêtre	Statistique	
3	3.5828	0.0320
6	2.1450	0.0508
9	1.5512	0.0753
VC_{KPSS} à 5%	0.4572	0.4572
Conclusion	I(1)	I(0)

L'intégration à l'ordre 1, suggérée par l'augmentation constante du volume de circulation, est confirmée par les tests de racine unitaire présentés dans le tableau ci-dessus.

4.5 Le manque de données sur la situation immobilière en Ile-de-France

Il serait intéressant de prendre en compte dans les modélisations des variables décrivant l'immobilier en Ile-de-France (immobilier résidentiel et immobilier de bureau).

Comme nous l'avons souligné dans les chapitres précédents, l'implantation des habitats et des bureaux influence fortement le mode de transport choisi. Entre 1991 et 2001, Paris a perdu des habitants et des emplois au profit de la banlieue [25] dans laquelle de nombreuses zones se sont équipées de commerces, services... En conséquence, les flux de déplacements ont connu une modification de leur structure géographique : les déplacements ayant au moins une extrémité à Paris ont stagné voire diminué, tandis que les déplacements à l'intérieur de la banlieue (petite et

4.5. LE MANQUE DE DONNÉES SUR LA SITUATION IMMOBILIÈRE EN ILE-DE-FRANCE

grande couronnes) ont fortement augmenté. Cette modification géographique a un impact sur les trafics de la RATP : le mode de transport privilégié dépend fortement de l'itinéraire parcouru. 50% des déplacements internes à la capitale sont effectués par la marche à pied et 30% par les transports en commun ; 6 trajets sur 10 reliant Paris au reste de l'Ile-de-France sont réalisés en transports collectifs ; à l'intérieur de la petite couronne, 45% des déplacements sont effectués par la marche à pied et 40% à l'aide d'une voiture ; dans la grande couronne, six trajets sur 10 sont effectués en voiture particulière.

Il n'existe malheureusement pas de séries mensuelles décrivant le phénomène de délocalisation des entreprises ou la fuite des ménages de Paris.

Quelques pistes ont été envisagées : l'INSEE propose quelques indicateurs trimestriels (créations et défaillances d'entreprises, par département), le site www.webimm propose des études, ou informe trimestriellement des loyers moyens constatés, de la demande placée ou des investissements pour les différents quartiers de Paris et les zones de périphérie.

Concernant l'immobilier, la Chambre des notaires de Paris Seine-Saint-Denis - Val-de-Marne (www.paris.notaires.fr) offre des informations sur les volumes et les chiffres d'affaires, les prix... Ces données concernent uniquement la petite couronne et sont trimestrielles. Pour la grande couronne, les données sont moins détaillées et mises à jour tardivement (voir par exemple la chambre des notaires de Versailles sur www.chambre-versailles.notaires.fr/stat_evolution.htm). L'INSEE propose un suivi annuel de la population estimée dans chaque région mais ne descend pas au niveau du département (sauf pour les recensements)

* * *

Les séries présentées dans ce chapitre correspondent à des séries d'environnement économique, pour la seule région Ile-de-France. Ces séries ont été désaisonnalisées, voire log-linéarisées pour les séries en niveau. Elles sont intégrées à l'ordre 1, à l'exception de $\ln(totgl)$, qui représente le nombre de voyages par les grandes lignes de la SNCF ayant une extrémité dans l'une des gares de Paris.

Présentons un tableau synthétique de ces variables économiques.

Identifiant	Nom de la série CVS	Source	Tests de RU
$\ln(defm75)$	Demandes d'emplois en fin de mois à Paris	DRTEFP	I(1)
$\ln(defmidf)$	Demandes d'emplois en fin de mois en Ile-de-France	DRTEFP	I(1)
$gdmagr$	Indice du chiffre d'affaires des grands magasins de la Région Parisienne	COE CROCIS	I(1)
$\ln(totgl)$	Nombre de voyages par les grandes lignes de la SNCF ayant une extrémité dans l'une des gares de Paris	SNCF	I(0)
$txhot75$	Taux d'occupation des hôtels à Paris	INSEE	I(1)
$txhotidf$	Taux d'occupation des hôtels en Ile-de-France	INSEE	I(1)
$\ln(carbidf)$	Ventes de carburants en Ile-de-France	CPDP	I(1)
$compet$	Rapport du prix du ticket T au prix du super	RATP, CPDP	I(1)
$\ln(circ75)$	Volume de circulation dans la zone centrale régulée de Paris	Mairie de Paris	I(1)
$\ln(speed75)$	Vitesse de circulation dans la zone centrale régulée de Paris	Mairie de Paris	I(1)
$circagar$	Indice d'évolution de la circulation dans la zone agglomérée de l'Ile-de-France, sur les routes nationales et autoroutes	SIER DREIF	I(1)

Les évolutions du nombre de demandeurs d'emplois, de la consommation ou du tourisme dépendent moins de la véritable situation économique que du moral des entreprises et des consommateurs, i.e. d'un ressenti qui s'explique en partie par les chocs passés, au niveau national comme au niveau international. La diminution progressive des volumes et vitesses de circulation à Paris provient quant à elle d'une politique visant à restaurer la sécurité sur les routes (radars...) et à limiter la place de l'automobile (politique du stationnement à Paris par exemple).

Conclusion à la première partie

Dans cette première partie, l'environnement institutionnel ainsi que le fonctionnement du marché des transports collectifs urbains ont été détaillés pour la Région Ile-de-France. Le transport est un bien particulier qui s'apparente à une consommation intermédiaire, ou "demande dérivée" : il n'est pas demandé pour lui-même mais pour la satisfaction qu'il apporte à la destination. Sur ce marché monopolistique, l'offre de la RATP (volume de référence et tarification) est fixée par une autorité organisatrice, le Syndicat des Transports en Ile-de-France.

La demande de transport collectif urbain correspond à une demande "effective" : elle résume l'équilibre sur le marché, tandis que l'offre de service de la RATP constitue une caractéristique du transport et conditionne le choix de ce mode de transport. En amont de ce choix modal, les différents motifs de déplacement (travail, école et facultés, achats, loisirs, ...) sont des déterminants de la demande de transport.

Les modélisations existantes ont pris en compte ces déterminants, généralement dans une optique de prévision de long terme. Les modèles macro-économiques possèdent une problématique similaire à la nôtre (prévision à court et moyen terme) mais les hypothèses sous-jacentes sont incompatibles avec les séries de trafic de la RATP.

Les séries mensuelles que nous désirons expliquer et prévoir sont couramment utilisées par l'unité Tarification Ventes Résultats du département Commercial de la RATP : elles sont corrigées des journées de grèves, de l'impact du calendrier ainsi que des variations saisonnières. Chaque série de trafic, pour un type de jour et un sous-réseau défini, est intégrée à l'ordre 1. Les grèves de décembre 1995 et leurs effets rémanents ont causé une brusque rupture, avec le départ brutal d'une partie de la clientèle. Dès 1997, les trafics sur les sous-réseaux augmentent fortement, grâce au retour de la croissance économique et des touristes à Paris, et à l'amélioration

constante de l'offre de service (volume et qualité).

Pour expliquer ces variations à partir des déterminants de la demande de transport, une base mensuelle de données environnementales a été construite. Ces données décrivent l'environnement économique en Ile-de-France, fort différent de l'environnement national. La majorité des séries sont intégrées à l'ordre 1, i.e. dépendent des chocs passés qui influent sur le moral des ménages et des entreprises.

Deuxième partie

Tests de changement du degré d'intégration au cours de la période

Introduction à la deuxième partie

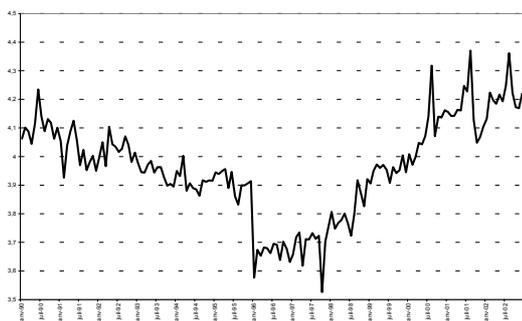
Face à une grève mais aussi un attentat, les utilisateurs de la RATP ont tendance à chercher un autre mode de transport. Ceci est particulièrement vrai pour les utilisateurs réguliers qui empruntent la RATP pour des motifs obligés, à savoir le travail, les affaires professionnelles ou le motif scolaire (école/université), pour lesquels la destination est imposée. Ce report peut prendre différentes formes, suivant la nature du choc :

- Dans le cas de menaces d’attentats ou d’un fort climat d’insécurité, certains utilisateurs du réseau ferré (métro et RER) vont se reporter vers le réseau de surface (les bus) qu’ils jugeront plus sûr.
- Dans le cas de grèves ponctuelles, ces reports pourront s’effectuer vers les réseaux moins perturbés de la RATP, ou se diriger vers des modes de transports concurrents : voiture, vélo, marche à pied...

Les voyageurs qui avaient choisi de se déplacer pour des motifs non obligés (achats, loisirs...) peuvent reporter leurs déplacements à une date ultérieure, ou changer de destination pour s’y rendre plus facilement.

En sens inverse, la fidélisation de la clientèle par le biais de campagnes sur les abonnements longs comme l’abonnement ”Intégrale” ou la carte ”Imagine R”, qui permettent d’effectuer un nombre illimité de voyages, peut avoir modifié les comportements et favorisé une mobilité accrue par les transports en commun. Les chocs ayant un impact sur les trafics de la RATP peuvent être plus exogènes : forte augmentation du chômage ou du tourisme en Ile-de-France...

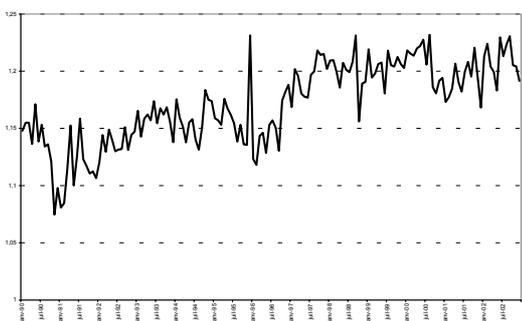
Les grèves de décembre 1995 semblent avoir profondément modifié les habitudes des usagers de la RATP. Les graphiques suivants, déjà présentés dans le chapitre 3 de la première partie semblent montrer un changement de comportement des séries, quel que soit le type de jour ou le sous-réseau considéré. Cette éventuelle rupture expliquerait les déboires des modélisations SARMA utilisées par le département Commercial de la RATP, dans les années qui ont suivi les grèves de décembre 1995.



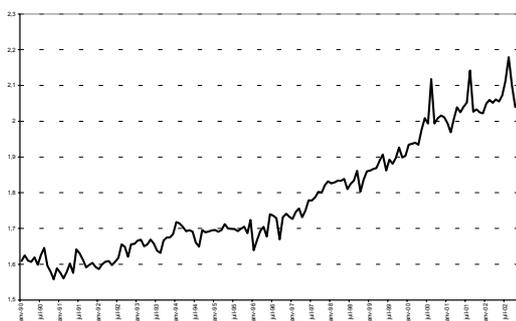
Trafic du jour ouvrable sur le métro



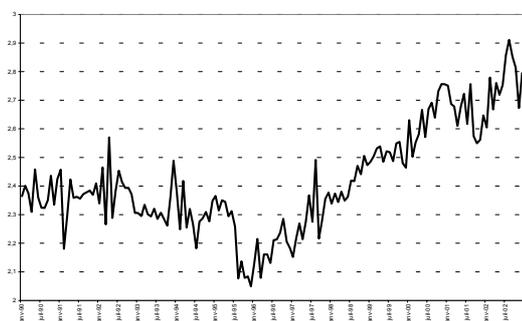
Trafic du jour ouvrable sur le RER



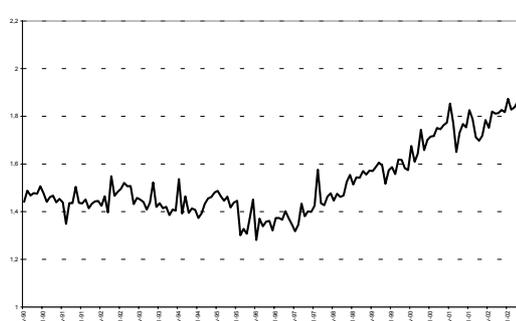
Trafic du jour ouvrable sur les bus de Paris



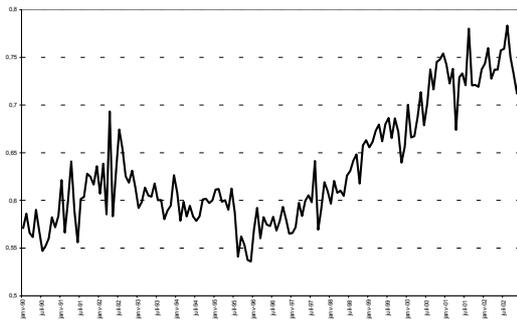
Trafic du jour ouvrable sur les bus de banlieue



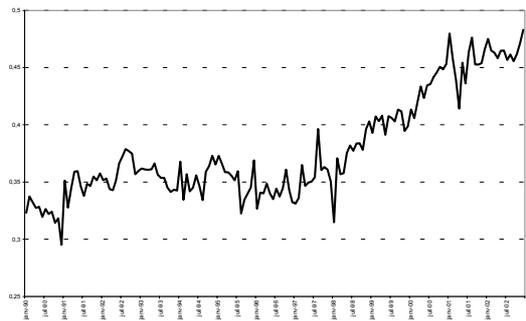
Trafic du samedi sur le métro



Trafic du dimanche/jour férié sur le métro



Trafic du samedi sur le RER



Trafic du dimanche/jour férié sur le RER

Les graphiques semblent montrer pour chaque série l'existence d'une rupture, parfois sur le niveau (jour ouvrable sur les bus de Paris) parfois sur la pente (trafic des week-ends et du jour ouvrable sur les bus de banlieue). Sur le métro, le niveau mais aussi la pente de la série du jour ouvrable semble s'être modifiés : le nombre de voyageurs a fortement diminué après les grèves, et s'est accru rapidement ensuite.

Les séries sont intégrées sur l'ensemble de la période, d'après le chapitre 3. Selon Perron (1989 [134]), le caractère intégré des variables peut provenir d'une rupture de la tendance et/ou de la constante, qui entraîne une modification des corrélations telle que les propriétés deviennent similaires à celles d'une marche aléatoire. Notre connaissance du trafic suggère pourtant que la forte croissance après 1997 est la résultante de différents chocs. Par exemple, l'essor des loisirs après la Coupe du Monde de Football ou la mise en place de Imagine R expliquent que les déplacements sont toujours plus nombreux au cours des week-ends.

Plus qu'une simple rupture de niveau et/ou de la pente de la tendance, une analyse préliminaire montre une modification du processus de mémoire à cette date. Le tableau suivant synthétise les conclusions des tests de racine unitaire sur l'ensemble de la période, et les résultats de ces tests sur deux sous-périodes : avant les grèves (janvier 1990 - décembre 1995) et après les grèves (janvier 1996 - décembre 2002). Les résultats détaillés sont disponibles dans l'annexe

	Période entière	Sous-périodes	
	janv 90 - déc 02	janv 90 - déc 95	janv 96 - déc 02
Jour ouvrable métro	I(1)	I(1)	I(1)
Jour ouvrable RER	I(1)	I(0)	I(1)
Jour ouvrable bus Paris	I(1)	I(1)	I(0)
Jour ouvrable bus banlieue	I(1)	I(1)	I(1)
Samedi métro	I(1)	I(1)	I(1)
Dimanche/Férié métro	I(1)	I(0)	I(1)
Samedi RER	I(1)	I(0)	I(1)
Dimanche/Férié RER	I(1)	I(0)	I(1)

Selon les tests ADF et KPSS, toutes les séries sont intégrées à l'ordre 1 sur l'ensemble de la période. L'analyse sur les sous-périodes *ante* et *post* grève de 1995 confirme l'intégration d'ordre 1 par morceaux pour les séries relatives au trafic du jour ouvrable sur le métro et le réseau du bus de banlieue, ainsi que pour la série de trafic du samedi sur le métro.

Pour les autres séries, un changement du degré d'intégration a eu lieu entre les deux sous-périodes :

- les séries décrivant le trafic sur le RER (jour ouvrable, samedi et dimanche/jour férié) et le trafic du dimanche/jour férié sur le métro sont stationnaires avant la grève et intégrées à l'ordre 1 ensuite,
- le trafic du jour ouvrable sur le réseau des autobus de Paris est intégré puis stationnaire.

La question est de savoir si les modifications des habitudes de déplacements suite à de tels évènements sont durables, et si cela a toujours été le cas sur la période d'étude, qui s'étend du mois de janvier 1990 au mois de décembre 2002. Observait-on auparavant une telle mémoire des chocs ? Autrement dit, ce choc a-t'il eu un impact sur le degré d'intégration ?

Mais la question la plus importante de ce chapitre est sans doute "L'éventuelle modification du phénomène de mémoire prend-elle source dans d'autres évènements que ces grèves ?"

Pour répondre à ces questions, nous établirons tout d'abord une revue de la littérature économétrique consacrée aux tests de changement de structure (chapitre 5). La modification du degré d'intégration n'a pas été envisagée dans cette littérature existante : il n'existe pas de test ayant pour hypothèse nulle l'existence d'une racine unitaire (i.e. de mémoire rémanente) et pour hypothèse alternative l'existence d'une racine unitaire sur une sous-période, l'autre sous-période étant stationnaire. Autrement dit, l'état actuel des recherches n'apporte pas de solution opérationnelle à ce problème.

La construction de tels tests est présentée dans le chapitre 6 de cette deuxième partie. Les hypothèses du test, les tables de valeurs critiques, mais aussi les propriétés de ces tests de modification du degré d'intégration sont détaillées. La présence d'auto-corrélations éventuelles a été prise en compte.

Enfin, les tests que nous avons construits sont appliqués aux données de trafic de la RATP ainsi qu'aux données économiques franciliennes (chapitre 7), de façon à préciser la nature des ruptures éventuelles.



Chapitre 5

Les tests de changement de structure existants

L'analyse des séries temporelles (modélisation, prévision...) s'appuie sur l'hypothèse de "stationnarité" au cours du temps, i.e. de la stabilité de la moyenne, de la variance ou de paramètres tels que la tendance. En outre, les auto-corrélations sont caractérisées par une décroissance exponentielle.

Considérons par exemple le plus simple des modèles dynamiques, un AR(1) avec constante :

$$\begin{aligned}y_t &= \alpha + \rho y_{t-1} + e_t \\e_t &\sim BB(0, \sigma^2)\end{aligned}$$

Les paramètres du modèle sont α , ρ , et σ . Le bruit blanc $BB(0, \sigma^2)$ désigne un processus formé de variables aléatoires centrées, de variance finie σ^2 et sans auto-corrélation. La "stationnarité" implique que ces paramètres sont constants au cours du temps.

Un "changement structurel" s'est produit si au moins un de ces paramètres s'est modifié au cours de la période à une date appelée "date de rupture". Si tous les paramètres sont modifiés, on parle de "pur changement de structure" contre un "changement de structure partiel" si une partie seulement des paramètres n'est pas constante au cours du temps.

5.1. TESTS DE CHANGEMENT DE STRUCTURE À UNE DATE INCONNUE ET ESTIMATION DE LA DATE DE RUPTURE

Les implications sont différentes suivant le type de paramètre soumis à rupture :

- Une modification du paramètre auto-régressif ρ entraîne une modification de la moyenne $E(y_t) = \frac{\alpha}{1-\rho}$, de la variance $V(y_t) = \frac{\sigma^2}{1-\rho^2}$ ainsi que des auto-corrélations temporelles

$$\frac{E((y_t - E(y_t))(y_{t-j} - E(y_t)))}{V(y_t)} = \rho^j$$

- La constante α intervient dans la moyenne du processus : $E(y_t) = \frac{\alpha}{1-\rho}$
- Un changement de la variance σ^2 correspond à une modification de la volatilité de la série.

Selon une synthèse bibliographique dressée en 2001 par Hansen [78], les plus importantes contributions à la littérature du changement de structure sont les suivantes :

1. Les tests de changement de structure à une date inconnue avec estimation de la date de rupture, dans un cadre stationnaire (premier paragraphe),
2. Les tests permettant de faire la distinction entre marche aléatoire et rupture de la tendance temporelle, et opposant stationnarité avec rupture contre intégration tout au long de la période (deuxième paragraphe).

Le troisième paragraphe montre que ces deux types de tests ne sont pas adaptés pour mettre en évidence un changement du degré d'intégration entre deux sous-périodes.

5.1 Tests de changement de structure à une date inconnue et estimation de la date de rupture

5.1.1 Test classique de changement de structure dans un modèle stationnaire

Point de départ de tous les développements ultérieurs, le test classique de changement de structure est le test de Chow (1960 [37]), qui divise l'échantillon en deux sous-périodes, estime les paramètres sur chacune de ces sous-périodes et teste l'égalité des deux ensembles de paramètres à l'aide d'une statistique de Fisher.

Supposons le modèle linéaire écrit sous la forme

$$y_t = \begin{cases} X_t\beta_1 + \varepsilon_{1t}, & t = 1, \dots, k \\ X_t\beta_2 + \varepsilon_{2t}, & t = k + 1, \dots, T \end{cases}$$

pour lequel on désire tester l'hypothèse nulle suivante : $\{H_0 : \beta_1 = \beta_2\}$. Le nombre total d'observations est égal à $T = k + (T - k)$, le nombre de paramètres est égal à p . Il existe deux manières équivalentes d'écrire cette hypothèse : en utilisant la somme des carrés résiduels ou bien en faisant appel à des variables indicatrices.

Si l'hypothèse nulle n'est pas vérifiée, il faut estimer séparément les deux régressions. La première régression permet d'estimer les paramètres ($\hat{\beta}_1$) et les résidus ($\hat{\varepsilon}_1$) de la première sous-période, et de calculer la somme des carrés résiduels sur cette sous-période ($SCR_1 = \hat{\varepsilon}_1'\hat{\varepsilon}_1$). La seconde régression donne $\hat{\beta}_2$, $\hat{\varepsilon}_2$ et $SCR_2 = \hat{\varepsilon}_2'\hat{\varepsilon}_2$. La somme des résidus non contrainte sur l'ensemble de la période est égale à $SCR_{nc} = SCR_1 + SCR_2$, son nombre de degrés de liberté est égal à $(k - p) + ((T - k) - p) = T - 2p$. Si l'hypothèse nulle est vérifiée, alors la procédure adéquate réside dans l'estimation de l'équation sur l'ensemble de la période, du paramètre $\hat{\beta}$, des résidus $\hat{\varepsilon}$ et de la somme des carrés des résidus $SCR_c = \hat{\varepsilon}'\hat{\varepsilon}$, qui a $T - p$ degrés de liberté. Sous l'hypothèse nulle, $SCR_{nc} \approx SCR_c$. La statistique de test est une F -statistique :

$$F = \frac{(SCR_c - SCR_{nc})/p}{SCR_{nc}/(T - 2p)} \stackrel{H_0}{\sim} F_{p, T-2p}$$

La procédure alternative faisant appel à des variables indicatrices est de réécrire le modèle sous la forme

$$y_t = X_t\beta_1 + Z_t d\beta + \varepsilon_t$$

où $d\beta = \beta_2 - \beta_1$ et Z_t est une matrice de format $(T \times p)$ telle que $Z_t = \begin{pmatrix} 0 & \text{si } t \leq k \\ X_t & \text{si } t > k \end{pmatrix}$. L'hypothèse $\{H_0 : \beta_1 = \beta_2\}$ est équivalente à $\{H_0 : d\beta = 0\}$ qui peut être testée avec un F -test usuel. La statistique de test est numériquement identique à la statistique précédente. Cette dernière spécification du test a pour avantage de permettre de tester l'égalité de tous les paramètres

5.1. TESTS DE CHANGEMENT DE STRUCTURE À UNE DATE INCONNUE ET ESTIMATION DE LA DATE DE RUPTURE

ou de seulement certains paramètres entre les deux périodes.

Les valeurs critiques sont issues d'une distribution de χ^2 . Ce test suppose l'égalité des variances sur les sous-périodes, ce qui est discutable : si les paramètres peuvent être différents, les variances peuvent l'être aussi. Une façon de corriger l'hétéroscédasticité est d'utiliser l'approche faisant appel à des variables indicatrices en association avec l'estimation de la matrice de variance-covariance robuste de White (1980 [165]), c'est-à-dire d'appliquer un test de Wald, généralisation du test de Chow à différentes estimations de la matrice de variance-covariance.

L'inconvénient principal de ce test réside dans l'hypothèse sous-jacente selon laquelle la véritable date de rupture est connue.

5.1.2 Test de changement de structure unique à date inconnue

Le test de Quandt (1960 [145]) suppose que la date de rupture est inconnue : la statistique utilisée correspond à la plus grande statistique de Chow sur toutes les dates de rupture possibles (les dates de rupture trop proches du début ou de la fin de l'échantillon sont traditionnellement omises, de façon à ce que les sous-périodes aient un nombre suffisant d'observations). L'estimation de la date de rupture et le test de changement de structure sont réalisés de façon simultanée. Les valeurs critiques de ce test n'ont été tabulées que dans les années 1990 par Andrews (1993 [6]), ou par Andrews et Ploberger (1994 [7]) qui montraient que l'on pouvait atteindre une puissance plus importante en prenant la moyenne exponentielle de la séquence des tests de Chow.

L'estimation de la date de rupture est de bonne qualité dans le cas d'un modèle linéaire quand le test de Chow est construit à partir de la matrice de variance - covariance homoscédastique. En outre, lorsque la statistique d'un test de Wald est construite en prenant en compte la matrice de variance - covariance homoscédastique, la date qui maximise la séquence de tests de Wald correspond algébriquement à la date qui minimise la somme des carrés résiduels.

De façon pratique, l'estimation peut être réalisée par les moindres carrés : pour chaque date de rupture possible, l'échantillon est divisé en deux sous-périodes, les autres paramètres sont estimés par les moindres carrés, et l'on conserve en mémoire la somme des carrés résiduels associée à cette estimation de la partition. La date de rupture estimée correspond à celle qui minimise la somme des carrés résiduels. Bai (1994 [12], 1997 [13]) a écrit la distribution asymptotique de

l'estimateur de la date de rupture, ainsi que la manière de construire des intervalles de confiance.

5.1.3 Tests de changements de structure multiples

Bai, Lumsdaine et Stock (1998 [17]) ont étendu cette analyse à des séries temporelles multivariées. Lorsque chacune des séries possède un changement de structure ayant lieu à la même date, il existe un gain à estimer l'ensemble du système au lieu de chacune des séries. Dans leur article, le test de changement de structure appliqué à chacune des séries (production réelle, consommation et investissement trimestriels aux Etats-Unis, pour la période 1959 - 1995) ne permettait pas de conclure à une rupture. L'hypothèse de rupture a pu être acceptée en prenant le vecteur de ces variables.

5.1.3.1 Estimation séquentielle des dates de rupture

Chong (1995 [36]) et Bai (1997 [13]) estiment de façon séquentielle les dates de rupture dans le cas de modèles possédant plusieurs ruptures. L'idée sous-jacente est la suivante : quand le modèle comporte plusieurs ruptures, la somme des carrés résiduels associée à une date de rupture atteint un minimum local au voisinage de la date de cette rupture. Par conséquent, le minimum global de la somme des carrés résiduels correspond à la date de la première rupture, les autres minima locaux sont interprétés comme des dates de rupture supplémentaires possibles. L'échantillon est divisé en deux sous-périodes à la date de rupture correspondant à la somme des carrés résiduels minimale, et un test de changement de structure est appliqué sur chaque sous-période dégagée. Bai montre que des améliorations de l'estimation des dates de rupture peuvent être obtenues à l'aide d'estimations successives des dates de rupture. Supposons par exemple qu'un modèle comporte deux ruptures. La première minimisation de la somme des carrés résiduels permet d'estimer la date de rupture a . Un test de changement de structure est effectué sur la première sous-période $[0, a]$, et accepte l'hypothèse de stabilité des paramètres ; un test de changement de structure sur la deuxième sous-période $[a + 1, T]$ met en évidence un changement de structure à la date $b, a < b < T$. Selon Bai, les dates de rupture estimées peuvent être affinées. L'application d'un test de changement de structure sur la sous-période $[0, b]$ permet l'obtention d'une date de rupture "affinée" \hat{a} . Un test de changement de structure est alors réalisé sur $[\hat{a}, T]$ et conduit

5.2. LA DISTINCTION ENTRE UNE MARCHE ALÉATOIRE ET L'EXISTENCE D'UN CHANGEMENT DE STRUCTURE

à l'obtention de la date de rupture estimée \hat{b} . L'itération continue jusqu'à ce que les dates de rupture estimées ne soient plus modifiées.

5.1.3.2 Estimation simultanée des dates de rupture

Bai et Perron (1998 [18] et [19] ou 2001 [20]) ont développé des tests de changement de structure multiple. La méthode utilisée correspond à une utilisation séquentielle de tests de changement de structure unique. Si le test rejette l'hypothèse nulle de stabilité des paramètres, l'échantillon est divisé en deux sous-périodes et un test de changement de structure unique est réalisé sur chacune de ces sous-périodes. Et ainsi de suite jusqu'à ce que le test ne puisse rejeter l'hypothèse d'absence de rupture sur une des sous-périodes dégagées.

Le grand intérêt de leurs travaux provient de l'utilisation d'un algorithme de programmation dynamique peu coûteux en termes de nombre d'estimations par les moindres carrés : pour un échantillon de taille T , quel que soit le véritable nombre de ruptures m , le nombre d'estimations est au maximum de l'ordre $O(T^2)$ contre $O(T^m)$ pour une procédure de balayage standard.

5.2 La distinction entre une marche aléatoire et l'existence d'un changement de structure

Les séries temporelles ont souvent été modélisées comme une somme d'une tendance et d'un cycle.

Jusqu'aux travaux de Nelson et Plosser (1982 [126]), il était courant de supposer une tendance linéaire, mais ces auteurs ont remis en cause cette hypothèse en montrant que la majorité des séries macro-économiques agrégées est caractérisée par une tendance correspondant à une marche aléatoire. En d'autres termes, au lieu d'avoir une tendance fixe autour de laquelle oscillent les cycles, les chocs aléatoires entraînent un changement du niveau de la tendance.

5.2.1 Tests avec date de changement de structure connue

5.2.1.1 Tests de racine unitaire contre stationnarité avec rupture de la tendance déterministe, à date de rupture connue.

Selon Perron (1989 [134]), le mouvement de cette tendance peut provenir d'un changement de structure unique parcimonieux dans la tendance linéaire, une rupture dans la tendance entraînant une modification des corrélations, dont les propriétés deviennent assez similaires à celles d'une marche aléatoire. Ignorer l'existence d'une rupture de la tendance peut biaiser les tests de racine unitaire usuels et conduire à accepter à tort l'hypothèse nulle de racine unitaire.

Les hypothèses testées par Perron sont les suivantes :

- Hypothèse nulle : marche aléatoire
- Hypothèse alternative : stationnarité autour d'un terme déterministe ayant subi une rupture à date connue.

Pour effectuer ce test, le processus auto-régressif est estimé en introduisant des variables indicatrices correspondant aux différentes spécifications de la rupture des termes déterministes¹. Si la tendance est une marche aléatoire, la somme des coefficients auto-régressifs est égale à l'unité, ce qui peut être détecté par l'application d'une statistique augmentée à la Dickey - Fuller. Perron a écrit la distribution asymptotique de ce test et en a calculé les valeurs critiques.

Shin (2001 [155]) propose un test de racine unitaire plus puissant que le test de Perron (1989 [134] et 1990 [135]) en utilisant la démonstration de Hansen (1995) selon laquelle il est possible d'obtenir des gains de puissance substantiels en incluant des variables corrélées avec la variable pour laquelle on teste l'existence d'une racine unitaire. Dans son modèle, un changement de structure unique peut survenir dans la composante déterministe, à une date connue.

¹Ces différentes spécifications sont rappelées dans l'annexe D page 593. Cette annexe est consacrée aux tests de Perron et à leur application aux séries de trafic RATP et aux séries relatives à l'environnement économique francilien.

5.2.1.2 Tests de racine unitaire avec rupture dans la partie déterministe contre stationnarité avec rupture, à date de rupture connue.

Leybourne, Newbold et Vougas (1998 [114]) estiment quant à eux la partie déterministe avec possible changement de structure et appliquent ensuite des tests de racine unitaire sur les résidus.

De façon similaire, Saikkonen & Lütkepohl (1999 [150]) considèrent un test de racine unitaire avec rupture en niveau à une date connue. La rupture apparaît sur le terme constant dans la composante déterministe, et non dans la composante stochastique. La forme de la fonction de rupture est assez générale : selon ces auteurs, connaître la date de rupture est en effet très courant en pratique, tandis que la forme précise de la rupture est généralement inconnue. Pour tester la racine unitaire dans ce contexte, les paramètres de nuisance sont estimés par les moindres carrés non linéaires, puis un test de type Dickey - Fuller est appliqué. Lütkepohl, Müller & Saikkonen (1999 [117]) ont légèrement modifié le test précédent, en supposant un modèle un peu plus général, dans lequel on considère une rupture graduelle, i.e. une transition douce entre deux états.

Modèles

1. De façon plus précise, Saikkonen et Lütkepohl considèrent le modèle suivant :

$$y_t = \mu_0 + \mu_1 t + f_t(\theta)' \gamma + x_t, \quad t = 1, 2, \dots, T$$

où les scalaires μ_0 , μ_1 , et les vecteurs θ de taille $(m \times 1)$ et γ de taille $(l \times 1)$ sont des paramètres inconnus. $f_t(\theta)$ est un vecteur de dimension $(l \times 1)$ composé de séquences déterministes de θ , et représente la fonction de rupture. Il est possible de considérer une rupture unique dans la moyenne du processus, avec par exemple $f_t(\theta) = d_t = \begin{pmatrix} 0 & \text{si } t \leq k \\ 1 & \text{si } t > k \end{pmatrix}$. x_t désigne le terme d'erreur stochastique non observable, supposé suivre une représentation

$AR(p)$:

$$\begin{aligned} a(L)(1 - \rho L)x_t &= \varepsilon_t \\ \text{où } a(L) &= 1 - a_1L - \dots - a_pL^p \\ \text{avec } a(L) &\neq 0 \text{ pour } |L| \leq 1 + \xi, \xi > 0 \\ \text{et } \varepsilon_t &\sim iid(0, \sigma^2) \end{aligned}$$

Si x_t possède une racine unitaire, celle-ci apparaît aussi dans le processus y_t . On réalise donc un test de racine unitaire sur le polynôme $a(L)$.

2. Dans le modèle de Lütkepohl, Müller & Saikkonen (1999), la flexibilité est accrue :

$$b(L)y_t = \mu_0 + \mu_1 t + f_t(\theta)' \gamma + v_t, \quad t = 1, 2, \dots, T$$

avec

$$\begin{aligned} v_t &= \rho v_{t-1} + \varepsilon_t \\ \varepsilon_t &\sim iid(0, \sigma^2) \\ -1 &< \rho \leq 1 \\ b(L) &= 1 - b_1(L) - \dots - b_p L^p \end{aligned}$$

Même en utilisant une fonction de rupture simple comme la variable indicatrice, la transition sera progressive par le biais de $\gamma [b(L)]^{-1} d_t$: si l'on pose $[b(L)]^{-1} = 1 + \sum_{l=0}^{\infty} \alpha_l L^l$ et que $t > k$, $[b(L)]^{-1} d_t = 1 + \sum_{l=1}^{t-k} \alpha_l$.

Test de racine unitaire. Les étapes des tests de Saikkonen et Lütkepohl, ou de Lütkepohl, Müller et Saikkonen sont les suivantes :

1. Estimation des paramètres de nuisance $\mu_0, \mu_1, \theta, \gamma, b_1, b_2, \dots, b_p$ ou a_1, \dots, a_p (ce qui correspond à une estimation avec rupture de la partie déterministe) par les moindres carrés généralisés, linéaires si la fonction de rupture est indépendante de θ , non-linéaires sinon.

5.2. LA DISTINCTION ENTRE UNE MARCHE ALÉATOIRE ET L'EXISTENCE D'UN CHANGEMENT DE STRUCTURE

2. Test de racine unitaire à la Dickey-Fuller sur les résidus. L'hypothèse nulle correspond à l'existence d'une racine unitaire ainsi que d'une rupture dans la partie déterministe ; sous l'hypothèse alternative, le processus est stationnaire avec rupture.

Lanne, Lütkepohl et Saikkonen (1999 [105]) comparent les deux types de test précédents. Des modifications sont apportées pour prendre en compte les erreurs d'estimation relatives aux paramètres de nuisance, ou encore pour ne pas faire d'hypothèse sur les valeurs initiales dans le cas d'échantillons de taille finie. Les distributions asymptotiques sont calculées dans le cas de présence ou d'absence de constante dans la régression auxiliaire, ainsi que la puissance locale et les propriétés en échantillon de petite taille. Lorsque les paramètres de nuisance sont estimés sous une alternative locale, des distorsions substantielles de niveau peuvent apparaître dans le cas de petits échantillons. Ce résultat amène Lanne, Lütkepohl et Saikkonen à considérer qu'il est préférable d'estimer le modèle sous l'hypothèse nulle de racine unitaire, ce qui contraste avec les résultats dans le cas de modèles sans changement de niveau. En outre, introduire une constante dans la régression auxiliaire permet d'éviter les effets indésirables de valeurs initiales réelles importantes.

5.2.2 Tests avec date de changement de structure inconnue

Les types de tests précédents ont été étendus aux cas où la date du changement de structure n'est pas supposée connue *a priori*, mais estimée de façon simultanée à la réalisation du test. La structure de ce paragraphe est parfaitement identique à celle du paragraphe précédent mais les tests supposent désormais que la date de rupture est inconnue.

5.2.2.1 Tests de racine unitaire contre stationnarité avec rupture de la tendance déterministe, à date de rupture inconnue.

De nombreux travaux ont mis en cause le test de Perron (1989 [134]), du fait de sa faible puissance et de l'hypothèse selon laquelle la date de rupture de la tendance est connue (donc dépendante des données) : citons par exemple Christiano (1992[38]), Zivot et Andrews (1992 [168]), Banerjee, Lumsdaine et Stock (1992 [22]), Perron et Vogelsang (1992 [137]), ou encore Perron (1997 [136]).

L'ensemble de ces auteurs considère qu'une procédure appropriée revient à sélectionner la date de rupture qui maximise en valeur absolue le t -ratio (plus le t -ratio est grand en valeur absolue et plus l'hypothèse de racine unitaire risque d'être rejetée).

Lumsdaine et Papell (1997 [116]) ont étendu ce type d'analyse en considérant deux ruptures de la tendance.

5.2.2.2 Tests de racine unitaire avec rupture dans la partie déterministe contre stationnarité avec rupture, à date de rupture inconnue.

A la suite de leurs travaux de 1999, Lanne, Lütkepohl et Saikkonen (2001 [107]) étendent les deux types de tests au cas où la date de rupture est inconnue. Des simulations de Monte Carlo sur les différents estimateurs des dates de rupture les amènent à faire des recommandations dans le cas de petits échantillons. Le choix du nombre de retards doit permettre de s'assurer de la blancheur des résidus et peut être choisi par un critère d'information. Pour un nombre de retards fixé, il est possible d'estimer la date de rupture soit en minimisant la somme des carrés résiduels, soit en minimisant la statistique de test de façon à choisir la date de rupture qui soit la moins favorable à l'hypothèse nulle. La première méthode est préférable pour obtenir une bonne approximation d'une part, et pour accroître d'autre part la performance des tests de racine unitaire. Les deux types de tests doivent être réalisés : l'hypothèse de racine unitaire doit être rejetée dès que l'un et/ou l'autre test rejette cette hypothèse.

5.2.3 Tests de changement de la persistance de la série à date inconnue.

Kim (2000 [97]), Kim, Belaire-Franch et Badillo Amador (2002 [98]), ou encore Busetti et Taylor (2002 [32]) ont construit des tests de changement de la persistance d'une série temporelle. En particulier, ces tests permettent de détecter si une série est intégrée puis stationnaire ou stationnaire puis intégrée, contre une hypothèse nulle de stationnarité tout au long de la période. Ils sont dérivés des tests de stationnarité sur les résidus établis entre autres par Kwiatkowski, Phillips, Schmidt et Shin (1992 [102]) (test de "KPSS"). La date de changement de structure est supposée connue, ou inconnue. Contrairement aux tests précédents, la date de rupture est estimée après avoir réalisé le test de changement de structure, et non de façon simultanée.

Présentons par exemple le test de Kim.

5.2.3.1 Hypothèses de test

L'hypothèse nulle correspond à une persistance constante au cours de la période considérée. Elle peut s'écrire sous la forme

$$\left\{ H0 : y_t = r_0 + z_t, \quad t = 1, \dots, T \right\}$$

où r_0 est une constante, et z_t est une variable stationnaire qui suit certaines conditions de régularité (mélange uniforme)².

L'hypothèse alternative considère un changement de persistance au cours de la période considéré, ayant lieu à la date $t = [\tau T]$, avec $\tau \in (0, 1)$. Si la persistance est plus forte en seconde période, l'hypothèse alternative s'écrit

$$\left\{ H1 : \begin{array}{ll} y_t = r_0 + z_{t,0}, & t = 1, \dots, [\tau T] \\ y_t = r_1 + z_{t,1}, & t = [\tau T] + 1, \dots, T \end{array} \right\}$$

où $z_{t,0}$ et $z_{t,1}$ sont des processus stationnaires possédant les conditions de régularité de mélange uniforme. $z_{t,1}$ possède une plus forte persistance que $z_{t,0}$, r_0 et r_1 sont des constantes. Si la persistance est plus forte en première période, alors l'hypothèse alternative s'écrit

$$\left\{ H1' : \begin{array}{ll} y_t = r_1 + z_{t,1}, & t = 1, \dots, [\tau T] \\ y_t = r_0 + z_{t,0}, & t = [\tau T] + 1, \dots, T \end{array} \right\}$$

²Le processus $\{z_t\}_0^\infty$ est tel que :

- $E(z) = 0$
- $E|z|^{\gamma+\varepsilon} < \infty$ pour $\gamma > 2$
- $\{z_t\}_0^\infty$ est un processus φ -mélangeant, avec des coefficients de mélange φ_m tels que $\sum_{m=1}^\infty \varphi_m^{1-\frac{2}{\gamma}} < \infty$
- La variance de long-terme $\sigma_z^2 = \sum_{j=0}^\infty E[z_{j+1}z'_1]$ existe
- $\lim_{T \rightarrow \infty} \text{var} \left(\sqrt{T}^{-1} \sum_{t=1}^{sT} z_t \right) = s\sigma_z^2$ et $\lim_{T \rightarrow \infty} \text{var} \left(\sqrt{T}^{-1} \sum_{t=sT+1}^T z_t \right) = (1-s)\sigma_z^2$ pour tout $s \in (0, 1)$.

5.2.3.2 Statistique de test quand la date de rupture est connue.

Soient $(\tilde{z}_{0,t} \ t = 1, \dots, [\tau T])$ les résidus de la régression de $(y_t, \ t = 1, \dots, [\tau T])$ sur la constante. Soient $(\tilde{z}_{1,t} \ t = [\tau T] + 1, \dots, T)$ les résidus de la régression de $(y_t, \ t = [\tau T] + 1, \dots, T)$ sur la constante. Définissons les sommes partielles suivantes :

$$S_{0,t}(\tau) = \sum_{i=1}^t \tilde{z}_{0,i}(\tau) \quad t = 1, \dots, [\tau T]$$

et $S_{1,t}(\tau) = \sum_{i=[\tau T]+1}^t \tilde{z}_{1,i}(\tau) \quad t = [\tau T] + 1, \dots, T$

La statistique de test, pour une date de rupture $[\tau T]$ s'écrit :

$$\Xi_T(\tau) = \frac{[(1-\tau)T]^{-2} \sum_{t=[\tau T]+1}^T S_{1,t}(\tau)^2}{[\tau T]^{-2} \sum_{t=1}^{[\tau T]} S_{0,t}(\tau)^2}.$$

5.2.3.3 Statistique de test quand la date de rupture est connue et estimation de la date de rupture.

Si la date de rupture $[\tau T]$ est inconnue, plusieurs statistiques peuvent être prises en compte :

- $\Xi_T(\hat{\tau}) = \max_{\tau \in \mathfrak{S}} \Xi_T(\tau)$, où \mathfrak{S} est un sous-ensemble compact de $(0, 1)$
- $\mathbf{E}(\Xi_T(\hat{\tau})) \equiv \int_{\tau \in \mathfrak{S}} \Xi_T(\tau) d\tau$
- $\log \mathbf{E}(\exp \Xi_T(\hat{\tau})) \equiv \log \left\{ \int_{\tau \in \mathfrak{S}} \exp(\Xi_T(\tau)) d\tau \right\}$

L'hypothèse nulle est rejetée en faveur de l'hypothèse alternative H1 (respectivement H1') si la valeur de la statistique $\Xi_T(\hat{\tau})$, $\mathbf{E}(\Xi_T(\hat{\tau}))$ ou $\log \mathbf{E}(\exp \Xi_T(\hat{\tau}))$ est supérieure (respectivement inférieure) à la valeur critique adéquate.

La date de rupture est estimée par

$$\Lambda_T(\tau) = \frac{[(1-\tau)T]^{-2} \sum_{t=[\tau T]+1}^T \tilde{z}_{1,t}(\tau)^2}{[\tau T]^{-1} \sum_{t=1}^{[\tau T]} \tilde{z}_{0,t}(\tau)^2}.$$

5.3 Inadaptation de ces tests à notre problème

5.3.1 Le manque de travaux sur les changements de structure dans les processus auto-régressifs

La majeure partie des tests évoqués précédemment se place sous une hypothèse nulle de stationnarité tout au long de la période considérée. L'existence d'une rupture sur les coefficients, à date de rupture connue, est envisagée au travers de tests classiques tels les tests de Chow, de Wald, de Fisher, du maximum de vraisemblance, de Student... Ces tests peuvent être modifiés, généralement par le biais d'une procédure de balayage, quand la date de rupture est inconnue.

La première lacune que nous avons constatée réside dans le manque de travaux sur l'application de ces tests dans le cas de processus auto-régressifs, surtout en échantillon de taille finie. Les séries que nous désirons étudier étant de taille finie (environ 150 observations), prendre les valeurs critiques asymptotiques n'était pas forcément adéquat.

Nous avons donc calculé, pour cette taille d'échantillon, les tables de valeurs critiques ainsi que les puissances relatives de ces tests en cas de rupture sur un processus auto-régressif. Nous avons en outre évalué la qualité de l'estimation des dates de rupture lorsque les tests refusent l'hypothèse de stabilité du paramètre auto-régressif au cours du temps (annexe B page 569).

5.3.2 L'inadaptation des tests existants de changements du degré d'intégration

Une autre hypothèse alternative à la stationnarité tout au long de la période correspond à l'hypothèse d'une racine unitaire sur l'ensemble de la période, elle conduit à l'application de test à la KPSS. Une extension intéressante de ce genre de test est l'examen des éventuels changements de persistance au cours du temps, par les tests de Kim [97] ou de Buseti et Taylor [32]. Une application de ces tests réside dans la détection du changement de l'ordre d'intégration des séries, qui peuvent être intégrées puis stationnaires, ou stationnaires puis intégrées. C'est précisément ce genre de phénomène qui semble se produire à l'intérieur de nos séries de trafic. Néanmoins, trois facteurs nous amènent à ne pas utiliser ce test :

- Ces tests ne s'appuient pas sur un changement éventuel du paramètre auto-régressif, cadre

de notre étude.

- Ils ne correspondent pas à la réalité de nos séries, qui sont intégrées à l'ordre 1 sur l'ensemble de la période.
- Enfin, en réalisant les mêmes simulations que Kim [97], nous n'avons pas retrouvé les mêmes tables de probabilités que lui. Nos simulations se rapprochent plus des valeurs critiques corrigées dans le corrigendum [98], elles demeurent néanmoins très sensibles à la valeur de la constante. Enfin, nos résultats indiquent que l'estimation de la date de rupture n'est pas très solide, faiblesse que les trois auteurs avait soulignée.

5.3.3 L'inadaptation des tests ayant pour hypothèse nulle l'existence de racine unitaire

Certains tests ont pour hypothèse nulle l'existence d'une racine unitaire tout au long de la période, à commencer par le test de Dickey-Fuller et ses extensions. L'hypothèse alternative correspond à une série stationnaire sur l'ensemble de la période. Le test de Perron oppose à l'existence d'une racine unitaire tout au long de la période une stationnarité sur chaque sous-période avec rupture de la pente et/ou du niveau de la tendance déterministe. Enfin, les travaux de Lütkepohl et Saikkonen (et de leurs différents co-auteurs) ont considéré que le modèle possède une rupture dans sa partie déterministe, qui doit être purgée avant d'appliquer un test ayant pour hypothèse nulle l'existence d'une racine unitaire, et pour hypothèse alternative la stationnarité.

La modification du degré d'intégration n'a pas été envisagée : il n'existe pas de pendant au test de Kim ou de Buseti et Taylor. Un tel test aurait pour hypothèse nulle la présence d'une racine unitaire sur toute la période et pour hypothèse alternative l'existence d'une racine unitaire sur une sous-période, tandis que l'autre sous-période serait stationnaire. Dans le cas d'un processus auto-régressif, le paramètre auto-régressif est égal à 1 sur une sous-période, et inférieur à l'unité sur l'autre. Ce type de rupture sur le paramètre auto-régressif peut être envisagé en étendant les tests de Wald ou Student à un paramètre auto-régressif unitaire.

Chapitre 6

Tests de changement du degré d'intégration

Les tests de changement de structure évoqués au chapitre précédent ont presque tous été développés dans le cadre stationnaire, selon une philosophie à la Perron. Selon lui, les séries pour lesquelles les tests d'intégration concluent à une intégration à l'ordre 1 correspondent souvent à des séries stationnaires autour d'une tendance déterministe soumise à rupture.

L'examen préliminaire de nos données indiquent effectivement que nos séries ne sont pas intégrées à l'ordre 1 sur toutes les sous-périodes, mais qu'elles conservent un degré d'intégration égal à 1 sur une sous-période.

Il n'existait pas d'outil opérationnel pour tester cette hypothèse; la construction d'un tel test est réalisée au sein de ce chapitre technique.

Les tests de changement d'ordre d'intégration seront tout d'abord (section 1) présentés dans le cas d'un processus AR(1). Le paramètre auto-régressif est contraint à 1 sur une des sous-périodes, de façon à être certain que c'est bien un changement du degré d'intégration qui est testé. La date de l'éventuelle rupture est supposée inconnue.

Dans la deuxième section, les "tests augmentés de rupture de l'ordre d'intégration" seront construits en utilisant une modélisation auto-régressive d'ordre p avec constante, soumise à un changement partiel des paramètres. A cette occasion, les tables de valeurs critiques seront

construites par simulation, de manière à permettre une utilisation opérationnelle des tests que nous avons construits.

Dans la troisième et dernière section, nous nous livrerons à des calculs de puissance de ces outils. Il s'agit d'évaluer la capacité de nos tests à détecter une modification de leur degré d'intégration quand celle-ci est avérée et de mesurer la qualité d'estimation de la date de rupture. Les enseignements de cette section seront utiles lors de l'application empirique de nos tests.

6.1 Changement du paramètre associé au premier retard

6.1.1 Modèle de base : AR(1)

Le modèle considère l'éventualité de deux sous-périodes a et b :

$$\begin{aligned} y_t &= \rho_a y_{t-1} + u_t, & t = 1, \dots, k \\ y_t &= \rho_b y_{t-1} + u_t, & t = k + 1, \dots, T \end{aligned}$$

soit

$$\begin{aligned} y_t &= \rho_a y_{t-1} + u_t, & t = 1, \dots, k \\ y_t &= \rho_a y_{t-1} + (\rho_b - \rho_a) y_{t-1} + u_t & t = k + 1, \dots, T \end{aligned}$$

ou, sous forme matricielle

$$Y = \rho_a Y_{-1} + \Delta\rho Y_{-1,k} + U$$

où le vecteur $Y_{0,k}$ est composé de 0 jusqu'à la ligne k , puis des y_t de la ligne $t = k + 1$ jusqu'à la dernière ligne.

6.1.2 Hypothèses de test

6.1.2.1 Un premier test sans contrainte

De façon générale, les hypothèses de test suivantes pourraient être prises en compte :

- Hypothèse nulle : absence de rupture sur le paramètre auto-régressif i.e. $H_0 : \{\Delta\rho = 0\}$.

- Hypothèses alternatives : rupture sur le paramètre auto-régressif à date inconnue c'est-à-dire $H_a : \{\Delta\rho \neq 0\}$. Ce type d'hypothèse alternative se teste à l'aide d'un "test supérieur de Wald" : pour chaque date de rupture, la statistique de Wald associée à $\Delta\rho \neq 0$ est calculée. La valeur finale du test est obtenue en prenant la valeur supérieure de ces statistiques de Wald, et la date de rupture correspond à la date à laquelle la valeur de la statistique de Wald est maximale.

Il est possible de considérer une hypothèse plus précise :

- soit $\{H_a : \Delta\rho > 0\}$, la date de rupture étant inconnue, qui sera testée à l'aide d'un "test supérieur de Student". Pour chaque date de rupture possible, la statistique de Student associée au paramètre de rupture $\Delta\rho$ est calculée. La statistique finale correspond à la plus grande valeur de la statistique de Student, la date de rupture est la date associée à cette valeur maximale.
- soit $\{H_a : \Delta\rho < 0\}$, la date de changement de structure étant inconnue, testée au moyen d'un "test inférieur de Student". La statistique finale correspond à la plus petite valeur de la statistique de Student, la date de rupture est la date associée à cette valeur inférieure.

Les valeurs critiques de ce type de test ont été calculées en prenant pour processus générateur de données une marche aléatoire sans constante, donc un processus intégré sur l'ensemble de la période. Nos simulations montrent que les valeurs critiques obtenues pour un paramètre auto-régressif unitaire sont plus élevées que les valeurs critiques associées à un processus stationnaire. En outre, la puissance des tests à détecter un processus stationnaire puis intégré ou intégré puis stationnaire (pour ces valeurs critiques dans le cas unitaire) est plus forte que la puissance des tests dans le cas d'un changement de paramètre auto-régressif dans un cas stationnaire (pour des valeurs critiques stationnaires). L'ensemble des résultats liés à l'application de ces tests dans le cas de processus auto-régressifs stationnaires pour un échantillon de taille finie est reporté à l'annexe B page 569.

Une modification des hypothèses de test permet la construction d'un test ayant pour hypothèse nulle l'existence d'une racine unitaire sur l'ensemble de la période contre l'hypothèse alternative d'un changement du degré de la stationnarité au cours de la période.

Hypothèses des tests de changements de degré d'intégration Pour être assuré de tester réellement un changement de l'ordre d'intégration, c'est-à-dire que la série est intégrée puis stationnaire ou stationnaire puis intégrée, le paramètre auto-régressif est posé égal à 1 sur une des deux sous-périodes. La contrainte, et les tests, sont différents suivant l'hypothèse alternative considérée.

Sous l'hypothèse alternative de série intégrée puis stationnaire :

Pour réaliser ce test, à chaque date de rupture possible, on suppose que $\rho_a = 1$. L'hypothèse nulle correspond à

$$\rho_a = \rho_b = 1$$

i.e. la série est intégrée sur l'ensemble de la période. Sous l'hypothèse alternative, la série est intégrée puis stationnaire, soit

$$\Delta\rho_{is} = \rho_b - 1 \neq 0$$

à une date de rupture inconnue. Ce test peut être réalisé en utilisant un "test de Wald supérieur", ou en réalisant un "test inférieur de Student", ayant l'hypothèse alternative plus précise (l'hypothèse d'une série explosive étant écartée)

$$\Delta\rho_{is} = \rho_b - 1 < 0$$

Soit le modèle de régression suivant :

$$\begin{aligned} Y &= Y_{-1} + \Delta\rho_{is}Y_{-1,k} + U \\ \Delta Y &= \Delta\rho_{is}Y_{-1,k} + U \end{aligned}$$

où le vecteur $Y_{-1,k}$ est composé de 0 jusqu'à la ligne k , puis des y_{t-1} de la ligne $t = k + 1$ jusqu'à la dernière ligne.

Un balayage des statistiques est effectué sur toutes les dates de rupture k possibles, en

excluant $\pi * 100\%$, $\pi \in (0, \frac{1}{2})$ des dates en début et en fin d'échantillon.

– Le test supérieur de Wald augmenté a pour statistique de test :

$$\sup_{k \in [\pi T, (1-\pi)T]} Wald_T(k), \pi \in \left(0, \frac{1}{2}\right)$$

$$\text{où } Wald_T(k) = \frac{\Delta \hat{\rho}'_{is} \left(Y'_{-1,k} Y_{-1,k} \right) \Delta \hat{\rho}_{is}}{\hat{\sigma}^2(k)}$$

$$\text{avec } \hat{\sigma}^2(k) = \frac{SCR_T(k)}{T-1}$$

– La statistique de test inférieur de Student est

$$\inf_{k \in [\pi T, (1-\pi)T]} Student_T(k) \text{ avec } Student_T(k) = \frac{\Delta \hat{\rho}_{is,k}}{\hat{\sigma}(\Delta \hat{\rho}_{is,k})}$$

Sous l'hypothèse alternative de série stationnaire puis intégrée :

L'équation peut se réécrire sous la forme suivante :

$$y_t = \rho_b y_{t-1} + (\rho_a - \rho_b) y_{t-1} + u_t, \quad t = 1, \dots, k$$

$$y_t = \rho_b y_{t-1} + u_t, \quad t = k+1, \dots, T$$

Pour chaque date de rupture possible, le paramètre auto-régressif est contraint à 1 sur la seconde sous-période, soit $\rho_b = 1$ dans cette équation. L'hypothèse nulle est inchangée :

$$\rho_a = \rho_b = 1$$

i.e. la série est intégrée sur l'ensemble de la période. L'hypothèse alternative devient

$$\Delta \rho_{si} = \rho_a - 1 \neq 0$$

à une date de rupture inconnue. Ce test peut être réalisé à l'aide d'un "test de Wald supérieur".

Pour tester l'hypothèse alternative

$$\Delta \rho_{si} = \rho_a - 1 > 0$$

6.2. CHANGEMENT DU DEGRÉ D'INTÉGRATION ET DE LA CONSTANTE AVEC CORRECTION DES AUTO-CORRÉLATIONS

il est possible de réaliser un "test inférieur de Student".

Le modèle de régression devient :

$$\Delta Y = \Delta \rho_{si} Y_{k,-1} + U$$

où le vecteur $Y_{k,0}$ est composé des y_t jusqu'à la ligne $t = k$, puis de 0 de la ligne $t = k + 1$ jusqu'à la dernière ligne.

– Dans ce contexte, le test supérieur de Wald augmenté a pour statistique de test :

$$\begin{aligned} \sup_{k \in [\pi T, (1-\pi)T]} Wald_T(k), \pi &\in \left(0, \frac{1}{2}\right) \\ \text{où } Wald_T(k) &= \frac{\Delta \hat{\rho}'_{si} \left(Y'_{k,-1} Y_{k,-1} \right) \Delta \hat{\rho}_{si}}{\hat{\sigma}^2(k)} \\ \text{avec } \hat{\sigma}^2(k) &= \frac{SCR_T(k)}{T-1} \end{aligned}$$

– La statistique de test inférieur de Student est

$$\inf_{k \in [\pi T, (1-\pi)T]} Student_T(k) \text{ avec } Student_T(k) = \frac{\Delta \hat{\rho}_{si,k}}{\hat{\sigma}(\Delta \hat{\rho}_{si,k})}$$

Un balayage des statistiques de test est réalisé sur toutes les dates de rupture k possibles, en excluant $\pi * 100\%$ des dates en début et en fin d'échantillon.

6.2 Changement du degré d'intégration et de la constante avec correction des auto-corrélations

Il est assez illusoire de penser que la constante pourrait demeurer constante entre deux sous-périodes en cas de modification du paramètre auto-régressif. Pour s'en convaincre, imaginons une série intégrée sur une première sous-période avec une constante positive, puis stationnaire (modélisation AR(1)). Même si la constante c est assez faible, l'existence d'une racine unitaire et la présence de cette constante conduiront à un niveau assez élevé de la série en fin de première

sous-période, surtout si la série a subi des chocs positifs. Il paraît difficile que, dans un état stationnaire, l'espérance du processus puisse être égale à $\frac{c}{1-\rho}$. Il semble donc intéressant de prendre en compte une rupture possible de la constante en même temps que l'on teste le changement du paramètre auto-régressif associé à la variable retardée.

Deux tests sont construits, selon que l'hypothèse alternative est :

- la série est stationnaire puis intégrée, avec rupture sur la constante, la date de changement étant inconnue,
- la série est intégrée puis stationnaire, avec rupture sur la constante, la date de changement étant inconnue.

La statistique de test a en outre été corrigée des auto-corrélations résiduelles en ajoutant des variables retardées, de façon à blanchir le bruit. Les variables retardées ne sont pas au centre du test : elles ont pour but de corriger la statistique de test, mais ne sont pas soumises elles-mêmes à un test de modification des paramètres.

6.2.1 Hypothèse alternative de série intégrée puis stationnaire

Le modèle général avec variables en différence retardées et constante s'écrit de la façon suivante :

$$\begin{aligned}
 y_t &= c_a + \rho_a y_{t-1} + \gamma_1 \Delta y_{t-1} + \dots + \gamma_{p-1} \Delta y_{t-p+1} + u_t, & t = 1, \dots, k \\
 y_t &= c_b + \rho_b y_{t-1} + \gamma'_1 \Delta y_{t-2} + \dots + \gamma'_{p-1} \Delta y_{t-p+1} + u_t, & t = k+1, \dots, T
 \end{aligned}$$

Un changement des paramètres auto-régressifs ainsi qu'un changement de la constante au cours du temps peuvent se produire. Dans le cas d'une hypothèse alternative de série "intégrée puis stationnaire", le paramètre auto-régressif de première période est contraint à 1. Le modèle de

6.2. CHANGEMENT DU DEGRÉ D'INTÉGRATION ET DE LA CONSTANTE AVEC
CORRECTION DES AUTO-CORRÉLATIONS

régression pour une date de rupture possible k est :

$$\begin{aligned}
 y_t &= c_a + y_{t-1} + \gamma_1 \Delta y_{t-1} + \dots + \gamma_{p-1} \Delta y_{t-p+1} + u_t & t \leq k \\
 y_t &= c_a + y_{t-1} + \gamma_1 \Delta y_{t-1} + \dots + \gamma_{p-1} \Delta y_{t-p+1} + u_t & t > k \\
 &+ (\gamma'_1 - \gamma_1) \Delta y_1 + \dots + (\gamma'_{p-1} - \gamma_{p-1}) \Delta y_{t-p+1} \\
 &+ (c_b - c_a) + (\rho_b - 1) y_{t-1}
 \end{aligned}$$

soit, matriciellement,

$$\begin{aligned}
 \Delta Y &= c + \sum_{j=1}^{p-1} \gamma_j \Delta Y_{-j} + U \\
 &+ \sum_{j=1}^{p-1} \Delta \gamma_{j,is} [\Delta Y_{-j}]_{t>k} + \Delta c_{is} [1]_{t>k} + \Delta \rho_{is} [Y_{-1}]_{t>k}
 \end{aligned}$$

Soit $Y_{k,-1}$ le vecteur composé de 0 jusqu'à la ligne k puis des y_{t-1} de la ligne $t = k + 1$ à la dernière ligne. De la même façon, fabriquons les vecteurs $\Delta Y_{k,-j}$ composés de 0 jusqu'à la ligne k , puis des y_{t-j} de la ligne $t = k + 1$ jusqu'à la dernière ligne.

L'équation de régression peut s'écrire sous la forme réduite

$$\Delta Y = W\theta + X_{is}\beta_{is} + U$$

où

$$\begin{aligned}
 W &= \begin{bmatrix} \mathbf{1} & \Delta Y_{-1} & \dots & \Delta Y_{-p+1} \end{bmatrix} \\
 \theta' &= \begin{bmatrix} c & \gamma_1 & \dots & \gamma_{p-1} \end{bmatrix} \\
 X_{is} &= \begin{bmatrix} \Delta Y_{k,-1} & \dots & \Delta Y_{k,-p+1} & [1]_{t>k} & Y_{k,-1} \end{bmatrix} \\
 \beta'_{is} &= \begin{bmatrix} \Delta \gamma_{1,is} & \dots & \Delta \gamma_{p-1,is} & \Delta c_{is} & \Delta \rho_{is} \end{bmatrix}
 \end{aligned}$$

6.2.1.1 Test supérieur de Wald sur le changement de constante et du paramètre auto-régressif

Le test supérieur de Wald porte à la fois sur le changement de la constante et du paramètre auto-régressif associé à la variable retardée.

- Sous l'hypothèse nulle,

$$\beta_{is} = 0$$

il n'y a ni changement de la constante, ni changement du paramètre auto-régressif (série intégrée sur l'ensemble de la période)

- Sous l'hypothèse alternative,

$$\beta_{is} \neq 0$$

il existe un changement de la constante et/ou du paramètre auto-régressif (égal à 1 avant, inférieur à 1 après).

Le test supérieur de Wald augmenté a pour statistique de test :

$$\sup_{k \in [\pi T, (1-\pi)T]} Wald_T(k), \pi \in \left(0, \frac{1}{2}\right)$$

$$\text{où } Wald_T(k) = \frac{\Delta \hat{\beta}'_{is} (X'_{is} M X_{is}) \Delta \hat{\beta}_{is}}{\hat{\sigma}^2(k)}$$

$$\text{avec } M = I - W(W'W)^{-1}W' \text{ et } \hat{\sigma}^2(k) = \frac{S_T(k)}{T-1}$$

Un balayage des statistiques de test est réalisé sur toutes les dates de rupture k possibles, en excluant $\pi * 100\%$ des dates en début et en fin d'échantillon.

6.2.1.2 Test inférieur de Student

Il est aussi possible d'utiliser un test inférieur de Student, pour tester plus précisément la rupture du paramètre auto-régressif, en prenant pour statistique de test

$$\inf_{k \in [\pi T, (1-\pi)T]} Student_T(k) \text{ avec } Student_T(k) = \frac{\Delta \hat{\rho}_{is,k}}{\hat{\sigma}(\Delta \hat{\rho}_{is,k})}$$

Par rapport au test supérieur de Wald précédent, les hypothèses de tests sont modifiées :

- H_0 : "Existence d'une racine unitaire avec éventuelle rupture sur la constante"
- H_a : "Processus intégré puis stationnaire avec rupture éventuelle sur la constante".

L'hypothèse nulle est similaire à un test de Perron avec rupture du niveau de la série, dans le cas d'une date de rupture inconnue. De façon similaire à son article de 1997, il serait possible de considérer un test légèrement modifié dans lequel la date de rupture estimée serait la date correspondant à la plus grande statistique de Student associée à Δc_{is} .

6.2.2 Hypothèse alternative de série stationnaire puis intégrée

Dans le modèle général, pour tester si la série est "intégrée puis stationnaire", le paramètre auto-régressif de deuxième période est contraint à 1. Le modèle de régression pour une date de rupture possible k est :

$$\begin{aligned} y_t &= c_b + y_{t-1} + \gamma'_1 \Delta y_{t-1} + \dots + \gamma'_{p-1} \Delta y_{t-p+1} + (c_a - c_b) + (\rho_a - 1)y_{t-1} + u_t, & t \leq k \\ y_t &= c_b + y_{t-1} + \gamma_1 \Delta y_{t-1} + \dots + \gamma_{p-1} \Delta y_{t-p+1} + u_t, & t > k \end{aligned}$$

soit

$$\begin{aligned} y_t &= c_b + y_{t-1} + \gamma_1 \Delta y_{t-1} + \dots + \gamma_{p-1} \Delta y_{t-p+1} + u_t & t \leq k \\ &+ (\gamma'_1 - \gamma_1) \Delta y_1 + \dots + (\gamma'_{p-1} - \gamma_{p-1}) \Delta y_{t-p+1} \\ &+ (c_a - c_b) + (\rho_a - 1)y_{t-1} \\ y_t &= c_b + y_{t-1} + \gamma_1 \Delta y_{t-1} + \dots + \gamma_{p-1} \Delta y_{t-p+1} + u_t & t > k \end{aligned}$$

d'où le modèle matriciel contraint

$$\begin{aligned} \Delta Y &= c + \sum_{j=1}^{p-1} \gamma_j \Delta Y_{-j} + U \\ &+ \sum_{j=1}^{p-1} \Delta \gamma_{j,si} [\Delta Y_{-j}]_{t \leq k} + \Delta c_{si} [1]_{t \leq k} + \Delta \rho_{si} [Y_{-1}]_{t \leq k} \end{aligned}$$

Soit $Y_{-1,k}$ le vecteur composé des y_{t-1} de la ligne 1 à la ligne k , puis de 0 de la ligne $k+1$ à la dernière ligne. Fabriquons les vecteurs $\Delta Y_{-j,k}$ composés des Δy_{t-j} jusqu'à la ligne $t = k$, puis de 0 de la ligne $t = k+1$ jusqu'à la dernière ligne.

L'équation de régression peut s'écrire sous la forme réduite :

$$\Delta Y = W\theta + X_{si}\beta_{si} + U$$

$$\begin{aligned} W &= \begin{bmatrix} \mathbf{1} & \Delta Y_{-1} & \dots & \Delta Y_{-p+1} \end{bmatrix} \\ \theta' &= \begin{bmatrix} c & \gamma_1 & \dots & \gamma_{p-1} \end{bmatrix} \\ X_{is} &= \begin{bmatrix} \Delta Y_{-1,k} & \dots & \Delta Y_{-p+1,k} & [1]_{t \leq k} & Y_{-1,k} \end{bmatrix} \\ \beta'_{is} &= \begin{bmatrix} \Delta \gamma_{1,si} & \dots & \Delta \gamma_{p-1,si} & \Delta c_{si} & \Delta \rho_{si} \end{bmatrix} \end{aligned}$$

6.2.2.1 Test supérieur de Wald sur le changement de constante et du paramètre auto-régressif

Le test supérieur de Wald porte à la fois sur le changement de la constante et du paramètre auto-régressif associé à la variable retardée.

– Sous l'hypothèse nulle,

$$\beta_{is} = 0$$

i.e. il n'y a ni changement de la constante, ni changement du paramètre auto-régressif. La série est intégrée tout au long de la période.

6.2. CHANGEMENT DU DEGRÉ D'INTÉGRATION ET DE LA CONSTANTE AVEC CORRECTION DES AUTO-CORRÉLATIONS

– Sous l'hypothèse alternative

$$\beta_{is} \neq 0$$

il y a changement de la constante et/ou du paramètre auto-régressif (inférieur à 1 sur la première sous période, égal à 1 sur la seconde).

Le test supérieur de Wald augmenté a pour statistique de test :

$$\begin{aligned} \sup_{k \in [\pi T, (1-\pi)T]} Wald_T(k), \pi &\in \left(0, \frac{1}{2}\right) \\ \text{où } Wald_T(k) &= \frac{\Delta \hat{\beta}'_{si} (X'_{si} M X_{si}) \Delta \hat{\beta}_{isi}}{\hat{\sigma}^2(k)} \\ \text{avec } M &= I - W (W'W)^{-1} W' \text{ et } \hat{\sigma}^2(k) = \frac{S_T(k)}{T-1} \end{aligned}$$

La statistique finale est obtenue par balayage sur toutes les dates de rupture k possibles, en excluant $\pi * 100\%$ des dates en début et en fin d'échantillon.

6.2.2.2 Test inférieur de Student

Il est aussi possible d'utiliser un test inférieur de Student, pour tester plus précisément la rupture du paramètre auto-régressif ayant pour hypothèses

- H_0 : "Existence d'une racine unitaire avec éventuelle rupture sur la constante",
- H_a : "Processus stationnaire puis intégré avec rupture éventuelle sur la constante".

La statistique de ce test correspond à

$$\inf_{k \in [\pi T, (1-\pi)T]} Student_T(k) \text{ avec } Student_T(k) = \frac{\Delta \hat{\rho}_{si,k}}{\hat{\sigma}(\Delta \hat{\rho}_{si,k})}$$

où les paramètres ont été calculés dans le modèle avec paramètre auto-régressif unitaire sur la seconde sous-période.

Il serait possible de considérer un test légèrement modifié dont la date de rupture estimée serait la date correspondant à la plus grande statistique de Student associée à Δc_{is} .

6.2.3 Valeurs critiques des tests de changement de degré d'intégration avec éventuelle rupture de la constante

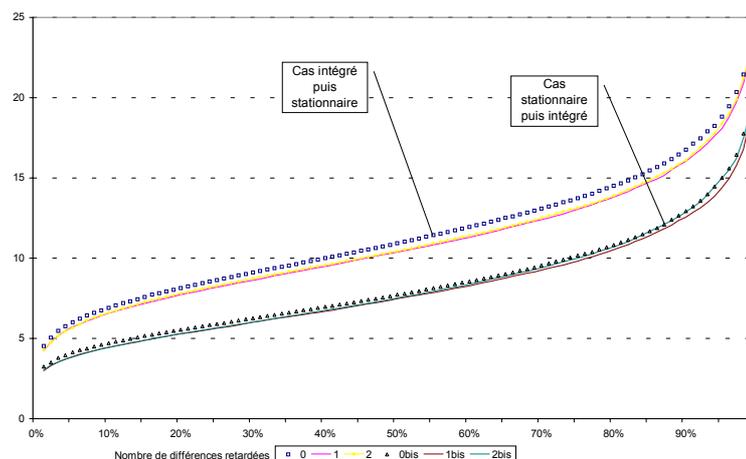
Pour chaque table, 10 000 simulations de Monte-Carlo pour un échantillon de taille 150 ont été réalisées. Le balayage des statistiques associées à chaque date de rupture a été effectué en enlevant 10% des observations en début d'échantillon et 10% des observations en fin d'échantillon.

Les valeurs critiques des tests sont présentées pour le cas intégré puis stationnaire comme pour le cas stationnaire puis intégré. Elles sont écrites pour un nombre de variables différenciées retardées compris entre 0 et 2. Nous disposons en outre de simulations pour un nombre de variables retardées égal à 6, montrant que l'on ne s'écarte jamais de plus de 3 points des valeurs critiques d'un processus à racine unitaire sans variables retardées. Ce travail a pour but de montrer que l'existence de variables retardées ne modifie pas les valeurs critiques obtenues en l'absence de corrélations temporelles.

6.2.3.1 Valeurs critiques du test supérieur de Wald

Nombre de variables différenciées	I(1) puis I(0)			I(0) puis I(1)		
	0	1	2	0	1	2
80%	14.505	13.804	13.893	10.825	10.518	10.698
90%	16.761	16.027	16.103	12.926	12.560	12.879
91%	17.127	16.394	16.505	13.239	12.854	13.191
92%	17.462	16.761	16.906	13.580	13.137	13.516
93%	17.898	17.154	17.344	13.986	13.467	13.955
94%	18.236	17.620	17.777	14.452	13.856	14.480
95%	18.816	18.076	18.386	15.014	14.393	14.998
96%	19.455	18.808	19.008	15.590	14.969	15.513
97%	20.342	19.718	19.871	16.439	15.760	16.202
98%	21.441	20.881	21.221	17.755	16.783	17.480
99%	23.234	22.929	22.844	19.595	18.906	19.214

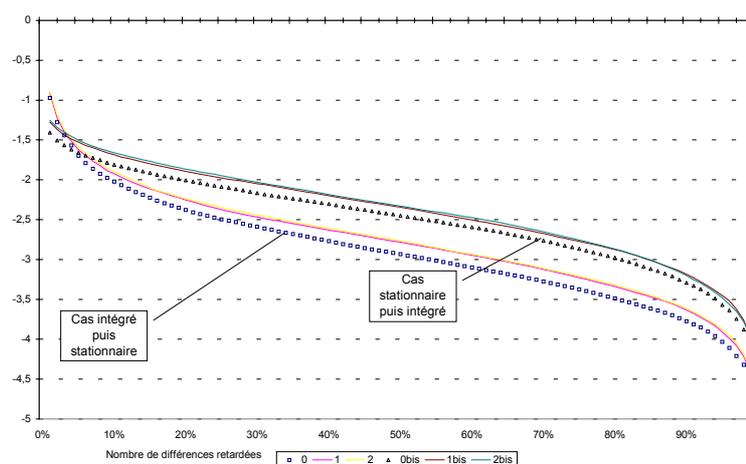
6.2. CHANGEMENT DU DEGRÉ D'INTÉGRATION ET DE LA CONSTANTE AVEC CORRECTION DES AUTO-CORRÉLATIONS



Valeurs critiques du test supérieur de Wald suivant l'hypothèse alternative testée, pour différents nombres de variables retardées. Le paramètre auto-régressif et la constante sont soumis à rupture, à une date inconnue.

6.2.3.2 Valeurs critiques du test inférieur de Student

Nombre de variables différenciées	I(1) puis I(0)			I(0) puis I(1)		
	0	1	2	0	1	2
80%	-3.488	-3.341	-3.322	-2.979	-2.876	-2.869
90%	-3.773	-3.634	-3.616	-3.294	-3.185	-3.202
91%	-3.811	-3.676	-3.660	-3.331	-3.229	-3.254
92%	-3.850	-3.726	-3.704	-3.372	-3.278	-3.297
93%	-3.904	-3.778	-3.760	-3.428	-3.333	-3.351
94%	-3.962	-3.831	-3.813	-3.490	-3.391	-3.411
95%	-4.033	-3.911	-3.878	-3.570	-3.449	-3.480
96%	-4.110	-3.990	-3.953	-3.636	-3.516	-3.560
97%	-4.211	-4.079	-4.051	-3.746	-3.621	-3.649
98%	-4.322	-4.216	-4.189	-3.879	-3.760	-3.783
99%	-4.511	-4.414	-4.411	-4.126	-4.001	-3.954



Valeurs critiques du test inférieur de Student suivant l'hypothèse alternative testée, pour différents nombres de variables retardées. Le paramètre auto-régressif et la constante sont soumis à rupture, à une date inconnue.

6.3 Puissance des tests

La puissance des tests est évaluée à l'aide de simulations de Monte-Carlo en examinant le rôle de trois facteurs :

1. la valeur du paramètre non contraint de la sous-période stationnaire,
2. la valeur de la constante associée à la sous-période stationnaire,
3. la place de la véritable date de rupture.

Pour chacun de ses facteurs, la puissance des tests est calculée en utilisant le processus générateur de données avec rupture à la date k

$$\begin{aligned}
 y_t &= c_1 + \rho_1 y_{t-1} + u_t \text{ si } t \leq k \\
 &= c_2 + \rho_2 y_{t-1} + u_t \text{ si } t > k \\
 u_t &\text{ est un bruit blanc}
 \end{aligned}$$

- dans le cas d'une série intégrée puis stationnaire, telle que $\rho_1 = 1$ et $|\rho_2| < 1$,
- dans le cas d'une série stationnaire puis intégrée, où $|\rho_1| < 1$ et $\rho_2 = 1$.

Le niveau du test est posé à 10%, la date de rupture est estimée en enlevant 10% des observations au début de l'échantillon et 10% des observations en fin d'échantillon. 2 500 simulations de Monte-Carlo sont réalisées pour chaque calcul de puissance. Les tables de puissance sont reportées à l'annexe C page 581.

6.3.1 Puissance des tests suivant la valeur du paramètre non contraint

La puissance des tests à 10% est calculée en posant une constante nulle sur chacune des sous-périodes, avec une véritable date rupture au point 75, i.e. à la moitié de l'échantillon. 2 500 simulations ont été réalisées. Dans les simulations suivantes, le paramètre sur la sous-période intégrée est égal à 1, tandis que le paramètre sur la sous-période stationnaire varie sur $]-1, 1[$.

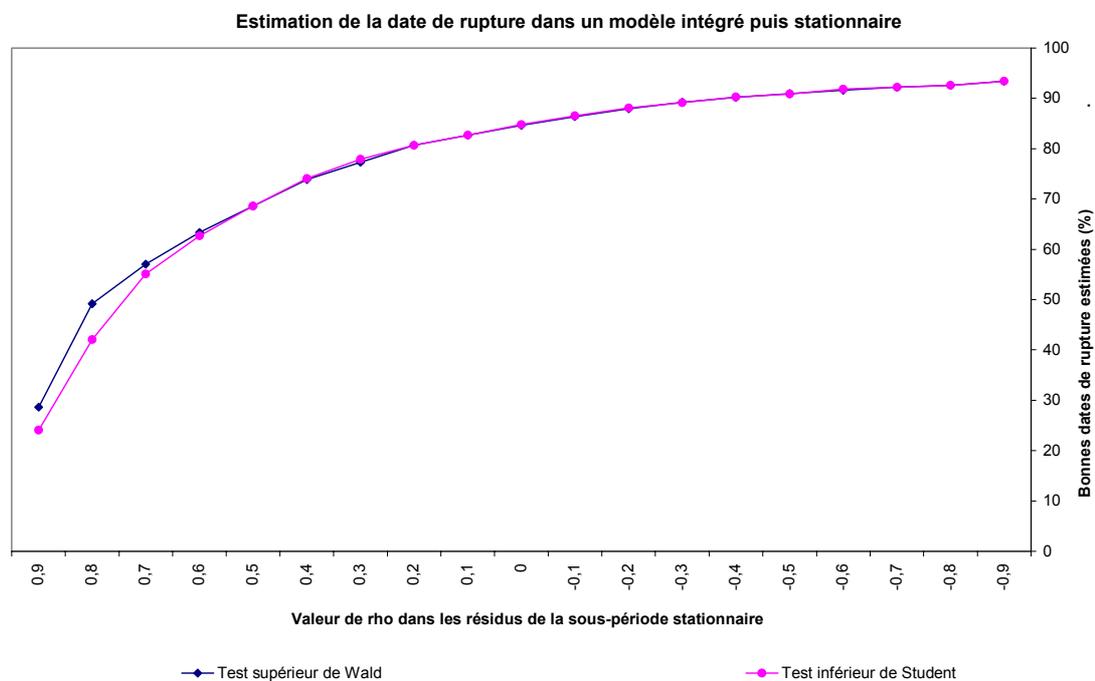
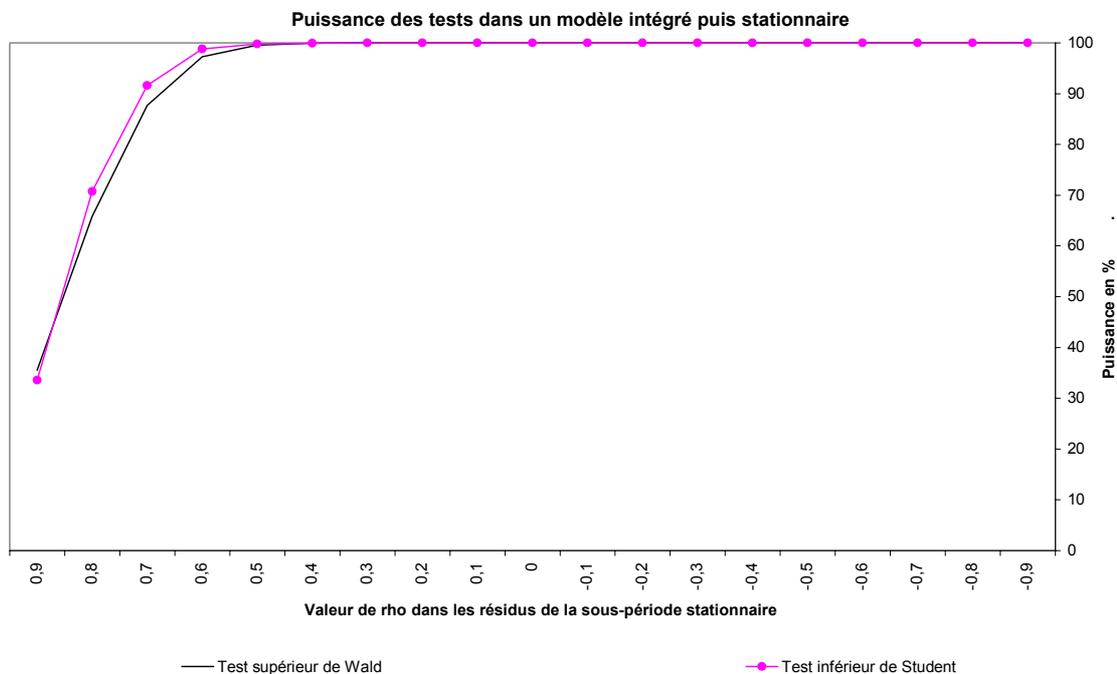
6.3.1.1 Cas d'une série intégrée puis stationnaire

Posons $\rho_1 = 1$, $c_1 = c_2 = 0$ et faisons varier ρ_2 pour une rupture à la date $k = 75$. Les résultats du test supérieur de Wald et du test inférieur de Student sont présentés sous forme de graphiques, les tableaux associés sont reportés à l'annexe C.2 page 582.

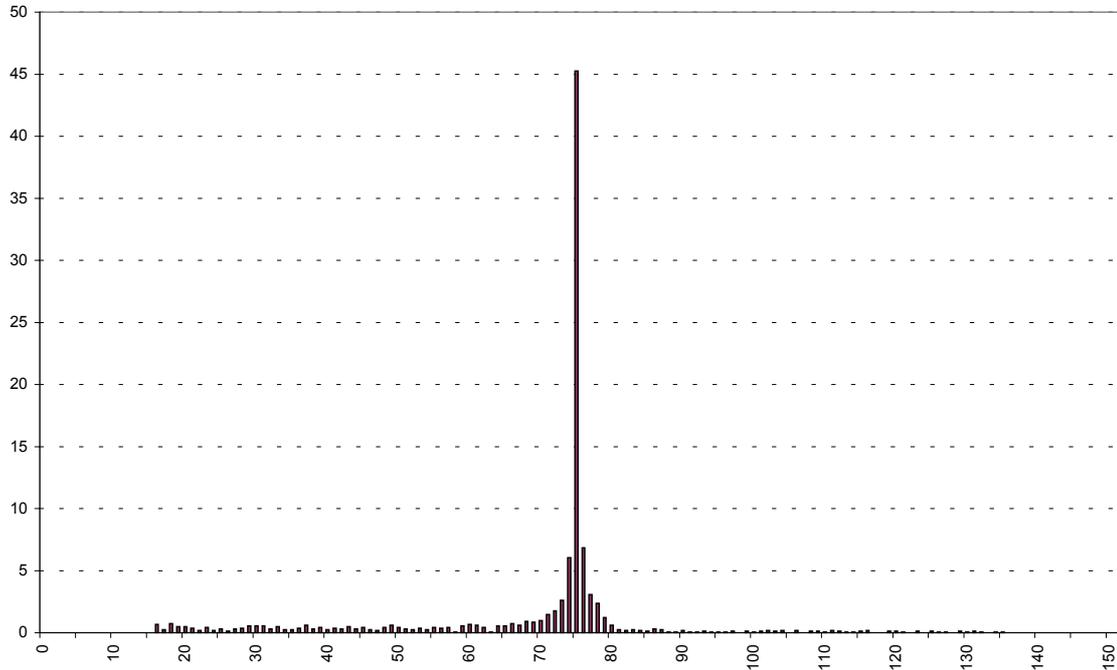
Les principaux faits marquants, en l'absence de constante sur chacune des sous-périodes et pour une date de rupture égale à 75, sont les suivants :

- Dans le cas d'un processus intégré puis stationnaire, la puissance du test supérieur de Wald est inférieure à la puissance du test de Student : si $\rho_1 = 1$ et $\rho_2 = 0.8$, la puissance du test supérieur de Wald est égale à 65.80% contre 70.72% pour le test inférieur de Student.
- Dans le cas d'un processus intégré puis stationnaire, l'estimation de la date de rupture est meilleure lorsqu'elle dérive du test supérieur de Wald : celui-ci conduit à 49.18% de bonnes dates estimées contre 42.03% pour le test inférieur de Student, pour un processus tel que $\rho_1 = 1$ et $\rho_2 = 0.8$. En outre, l'estimation de la date de rupture est de bonne qualité : si $\rho_2 = 0.8$, plus de 70% des dates estimées par ces tests sont situées à une distance inférieure à plus ou moins 6 mois de la véritable date de rupture.
- La puissance des tests supérieur de Wald et inférieur de Student diminue au fur et à mesure que le paramètre associé à la période stationnaire se rapproche de 1. Elle se rapproche du niveau nominal du test.

- De même la qualité de l'estimation de la date de rupture diminue au fur et à mesure que le paramètre de deuxième sous-période est proche de 1.



L'histogramme des dates de rupture estimées par le test supérieur de Wald dans le cas d'un processus généré par $\rho_1 = 1$, $\rho_2 = 0.8$, et $c_1 = c_2 = 0$, $k = 75$, est représenté ci-dessous.



Estimation de la date de rupture par le test supérieur de Wald ayant pour hypothèse alternative "série intégrée puis stationnaire" dans le cas d'un processus généré par $\rho_1 = 1$, $\rho_2 = 0.8$, $c_1 = 0$, $c_2 = 1$, $k = 75$.

6.3.1.2 Cas d'une série stationnaire puis intégrée

Soient $\rho_2 = 1$, $c_1 = c_2 = 0$. Faisons varier ρ_1 à l'intérieur du cercle unitaire pour une rupture à la date 75 (2 500 simulations).

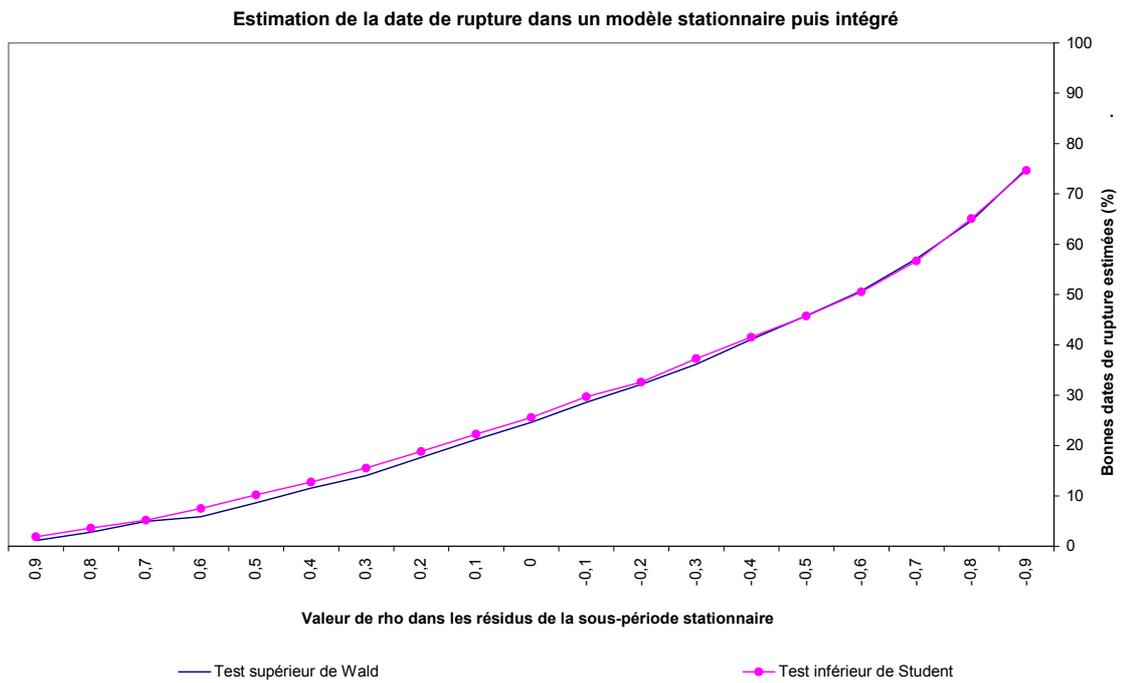
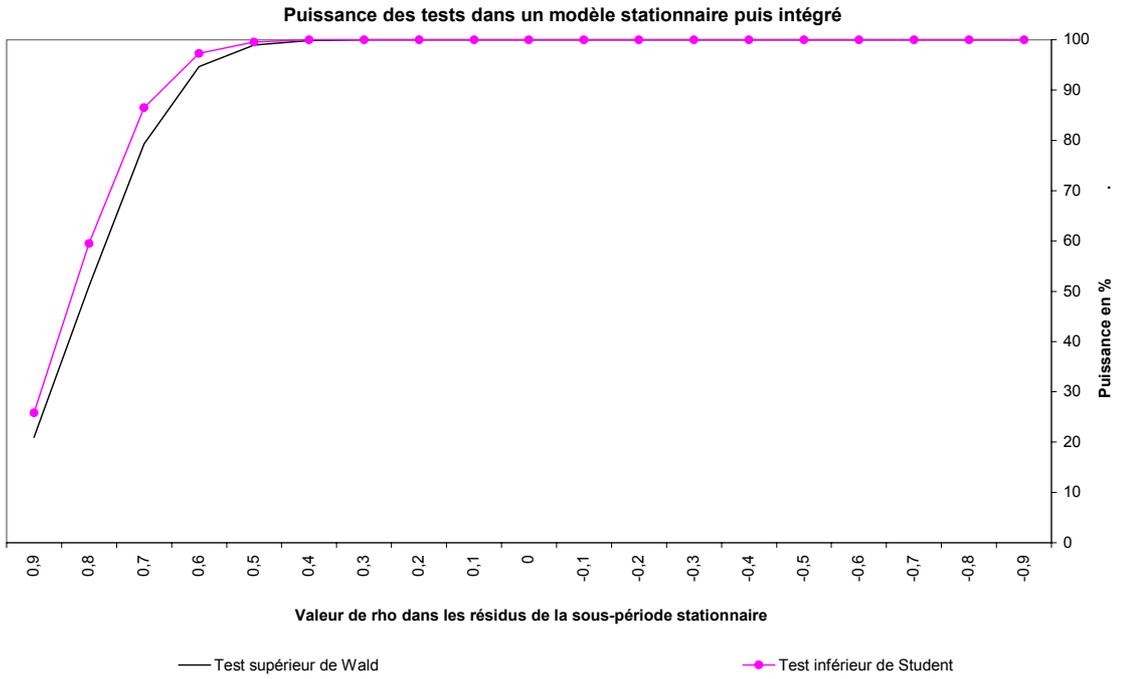
Les principaux faits marquants, en l'absence de constante sur chacune des sous périodes et pour une date de rupture égale à 75, sont les suivants :

- La puissance du test supérieur de Wald est plus faible dans le cas d'un processus "stationnaire puis intégré" que dans le cas d'un processus "intégré puis stationnaire", aux abords de l'hypothèse alternative. Pour un processus égal à 0.8 sur la sous-période "stationnaire" et 1 sur la sous-période "intégrée", la puissance du test supérieur de Wald est égale à 51.04% dans le cas I(0) puis I(1), contre 65.80% dans le cas I(1) puis I(0).

- De même, l'estimation de la date de rupture par le test supérieur de Wald est moins bonne pour un processus "stationnaire puis intégré" que pour un processus "intégré puis stationnaire" : seules 2.74% des bonnes dates de rupture sont estimées pour $\rho_1 = 0.8$ et $\rho_2 = 1$, contre 40.18% dans le cas d'un processus généré par $\rho_1 = 1$ et $\rho_2 = 0.8$. De même, la probabilité que la date estimée soit éloignée au maximum de plus ou moins 6 mois autour de la véritable date de rupture tombe à 27.82% dans le cas I(0) puis I(1) contre 70.21% dans le cas intégré puis stationnaire. La mauvaise qualité de l'estimation de la date de rupture lorsque ρ_2 est proche de 1 provient du fait que, le plus souvent, le test prend en compte la dernière date de rupture possible : il se rend compte très tardivement du passage d'un régime stationnaire à un régime intégré.
- Le test inférieur de Student dans ce cadre stationnaire puis intégré a une puissance supérieure au test supérieur de Wald, ce qui n'était généralement pas le cas pour un processus intégré puis stationnaire. Pour un processus généré par $\rho_1 = 0.8$ et $\rho_2 = 1$, ayant pour date de rupture la date 75, la puissance de ce test est égale à 59.48% contre 48.68% pour le test supérieur de Wald. La puissance demeure inférieure à celle obtenue pour une série I(1) puis I(0), où l'on obtenait 70.72% pour le test inférieur de Student et 65.80% pour le test supérieur de Wald.
- La qualité d'estimation de la date de rupture est identique pour chacun des deux tests.
- La puissance des tests supérieur de Wald et inférieur de Student diminue au fur et à mesure que le paramètre associé à la période stationnaire se rapproche de 1. Elle se rapproche du niveau nominal du test. La qualité de l'estimation de la date de rupture diminue au fur et à mesure que le paramètre de deuxième sous-période est proche de 1.

Les calculs de puissance dans le cas stationnaire puis intégré, ainsi que les mesures de qualité de l'estimation de la date de rupture sont reportés au tableau de l'annexe C.2 page 586.

6.3. PUISSANCE DES TESTS

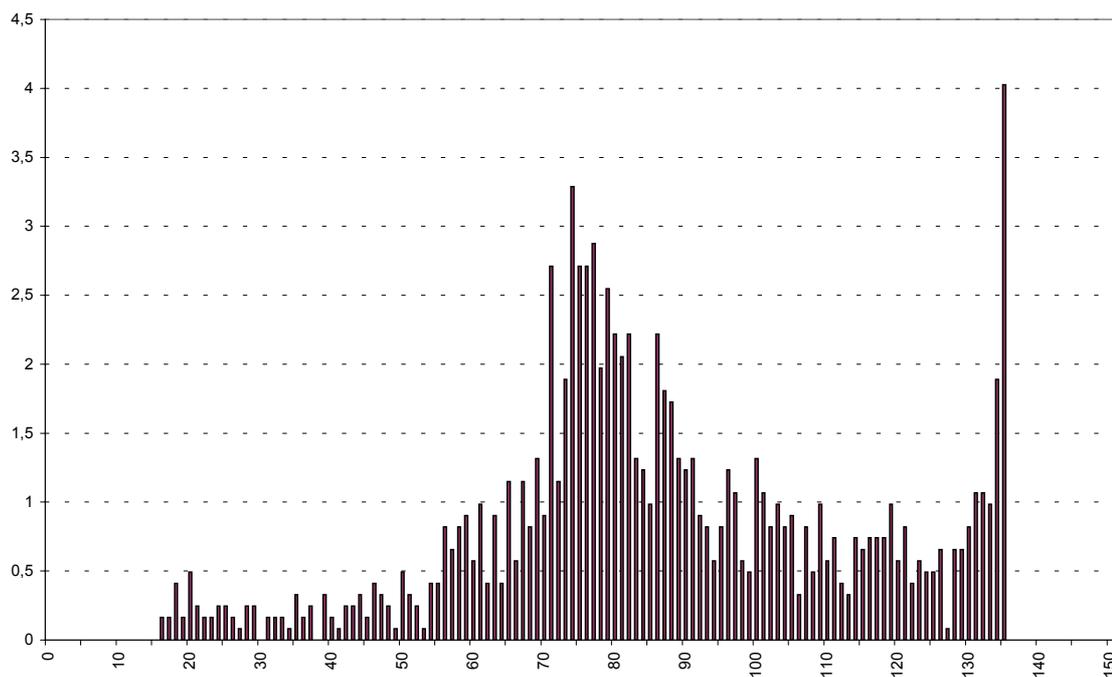


A l'approche de l'hypothèse alternative de racine unitaire sur l'ensemble de la période, les tests supérieur de Wald et inférieur de Student ont pour défaut majeur de s'apercevoir avec

retard que le régime stationnaire est devenu intégré. Les dates de rupture estimées sont plus souvent postérieures à la véritable date de rupture. Le test a tendance à prendre comme date de rupture estimée la dernière date de rupture possible.

En conséquence, lorsqu'un de ces tests ayant pour hypothèse alternative "la série est stationnaire puis intégrée avec éventuelle rupture de la constante" indique une modification du degré d'intégration à la fin de l'échantillon, il est judicieux d'examiner de façon détaillée l'ensemble des statistiques de chacun des tests (i.e. pour toutes les dates de rupture possibles), pour mettre en évidence un extremum local de la statistique.

L'histogramme des dates de rupture pour $\rho_1 = 0.8$ et $\rho_2 = 1$, $c_1 = 1$ et $c_2 = 0$, $k = 75$, synthétise cette lacune dans le cas du test supérieur de Wald :



Estimation de la date de rupture par le test supérieur de Wald ayant pour hypothèse alternative "série stationnaire puis intégrée" dans le cas

d'un processus généré par $\rho_1 = 0.8$, $\rho_2 = 1$, $c_1 = 1$, $c_2 = 0$, la date de rupture ayant lieu en 75.

6.3.2 Puissance suivant la valeur de la constante sur la partie stationnaire

6.3.2.1 Cas d'une série intégrée puis stationnaire

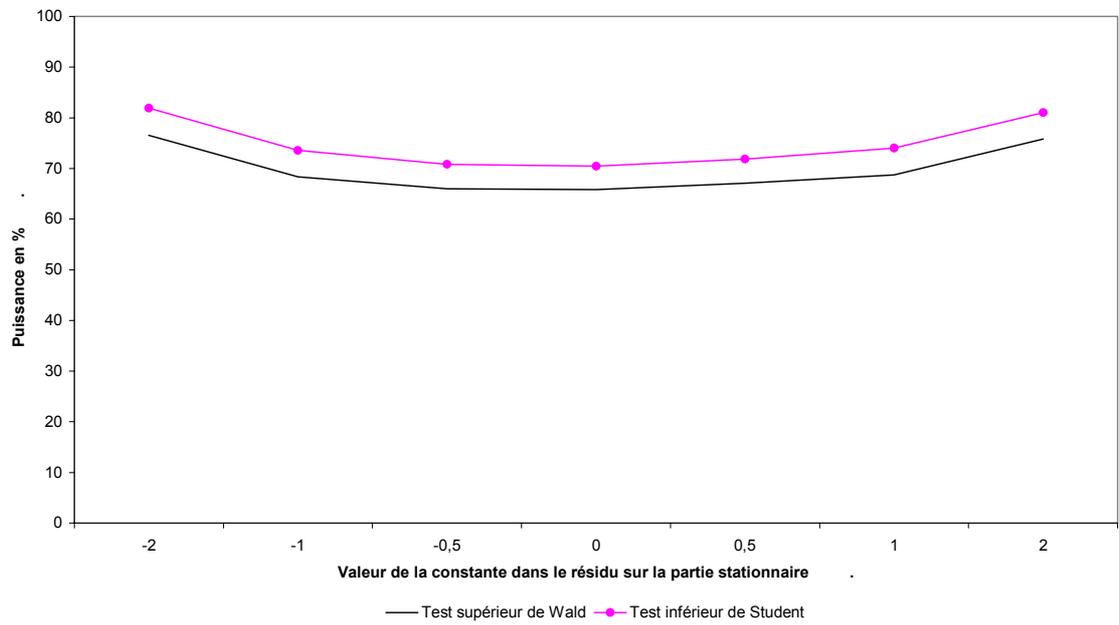
Soient $\rho_1 = 1$ et $\rho_2 = 0.8$ pour une véritable date de rupture égale à 75. La constante est nulle sur la partie intégrée ($c_1 = 0$), nous faisons varier la constante sur la partie "stationnaire" du processus .

Les principales remarques sont les suivantes :

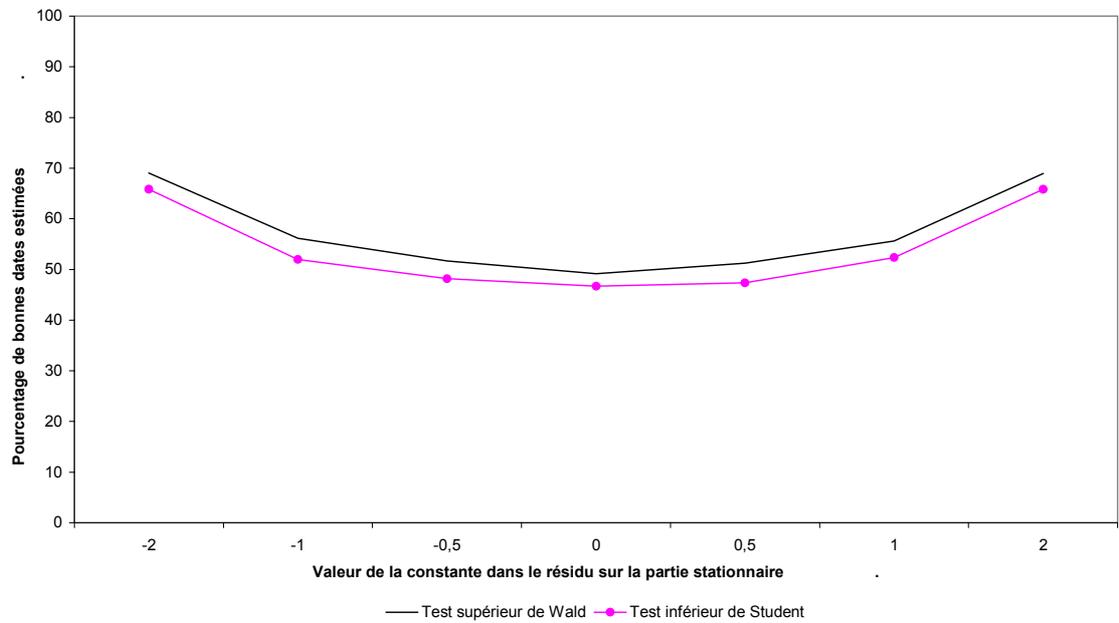
- La puissance des tests supérieur de Wald et inférieur de Student dans le cas d'une série intégrée puis stationnaire s'accroît quand la valeur absolue de la constante sur la partie stationnaire augmente. Le signe de la variation de la constante ne joue aucun rôle : courbes de puissance et mesures de la qualité de l'estimation de la date de rupture sont symétriques par rapport à $c_2 = 0$, où la puissance des tests est minimale. Par exemple, le test supérieur de Wald a une puissance égale à 66.12% dans le cas d'une constante nulle, et est de l'ordre de 79% quand la constante a une valeur absolue égale à 2. De même, le test inférieur de Wald a une puissance égale à 70.72% pour $c_2 = 0$, contre environ 84% pour $|c_2| = 2$.
- Dans le cas d'une série intégrée puis stationnaire, la qualité de l'estimation s'accroît pour chacun des tests quand la constante augmente en valeur absolue. Ainsi, si la constante sur la partie stationnaire est nulle le test supérieur de Wald estime la date de rupture correcte dans 45.25% des cas (42.03% pour le test inférieur de Student) contre 63.88% dans le cas d'une constante égale à 2 (respectivement 61.51%).
- Le test inférieur de Student a une puissance plus grande que le test supérieur de Wald quel que soit le niveau de la constante, mais conduit à une estimation de la date de rupture de moins bonne qualité (plus de 4 points d'écart pour les calculs de puissance, plus de 2 points d'écart pour l'estimation de la bonne date de rupture).

Les résultats détaillés de ces calculs de puissance sont reportés dans l'annexe C.3 page 589.

Puissance des tests dans un modèle intégré puis stationnaire



Estimation de la date de rupture dans un modèle intégré puis stationnaire



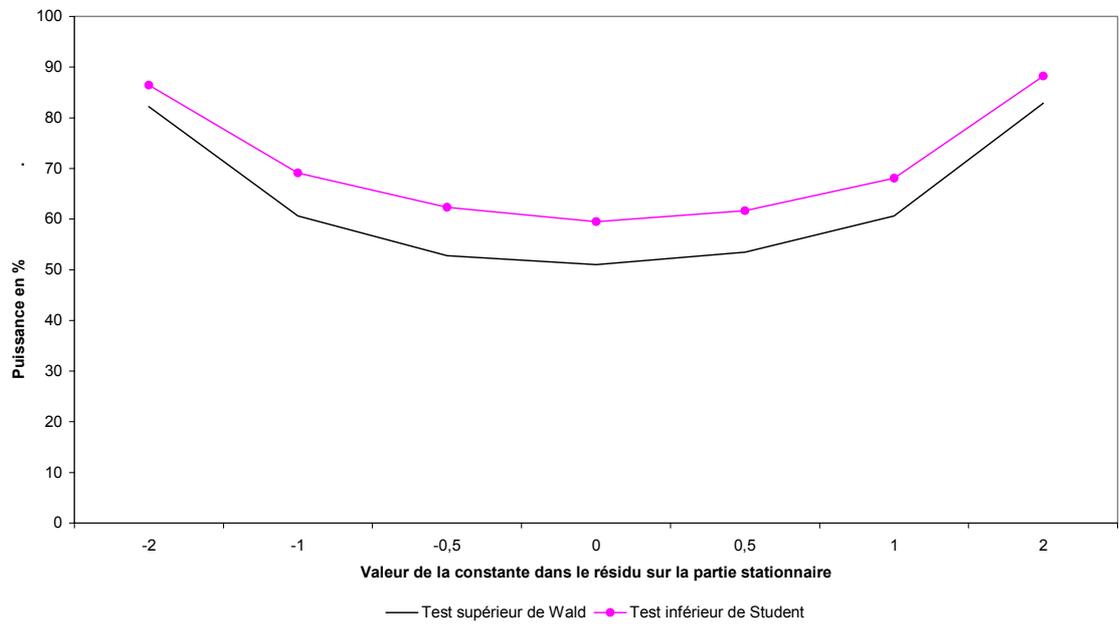
6.3.2.2 Cas d'une série stationnaire puis intégrée

Soient $\rho_1 = 0.8$ et $\rho_2 = 1$ pour une véritable date de rupture égale à 75. La constante est nulle sur la partie intégrée ($c_2 = 0$), elle varie sur la période stationnaire. L'annexe C.3 page 590 détaille les résultats de puissance.

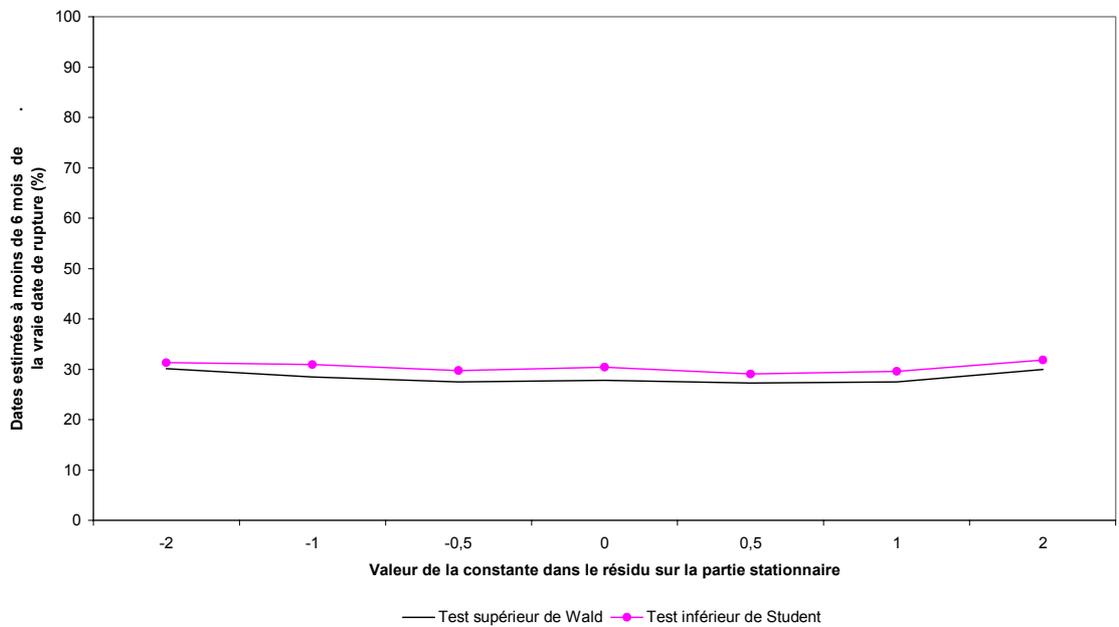
Les faits marquants associés aux tests ayant pour hypothèse alternative "la série est stationnaire puis intégrée avec rupture éventuelle de la constante" se résument de la façon suivante :

- De façon identique au cas intégré puis stationnaire, la puissance des tests supérieur de Wald et inférieur de Student s'accroît quand la valeur absolue de la constante sur la partie stationnaire augmente, de façon symétrique autour de $c_1 = 0$, où la puissance est minimale.
- La qualité de l'estimation de la date de rupture dans le cas stationnaire puis intégré ne varie pas la constante sur la partie stationnaire est modifiée. Ainsi, l'intervalle de confiance à plus ou moins 6 mois autour de la date de rupture estimée par le test supérieur de Wald contient dans 25-30% des cas la véritable date de rupture (autour de 30% pour le test inférieur de Student), quelle que soit la valeur de c_1 .
- Le test inférieur de Student a une puissance plus forte que le test supérieur de Wald, même si l'on se situe à des niveaux inférieurs à ceux obtenus dans le cas d'un régime intégré puis stationnaire. L'estimation des dates de rupture est médiocre par rapport aux tests dans le cadre I(1)-I(0).

Puissance des tests dans un modèle stationnaire puis intégré



Estimation de la date de rupture dans un modèle stationnaire puis intégré



6.3.3 Puissance des tests suivant la véritable date de rupture

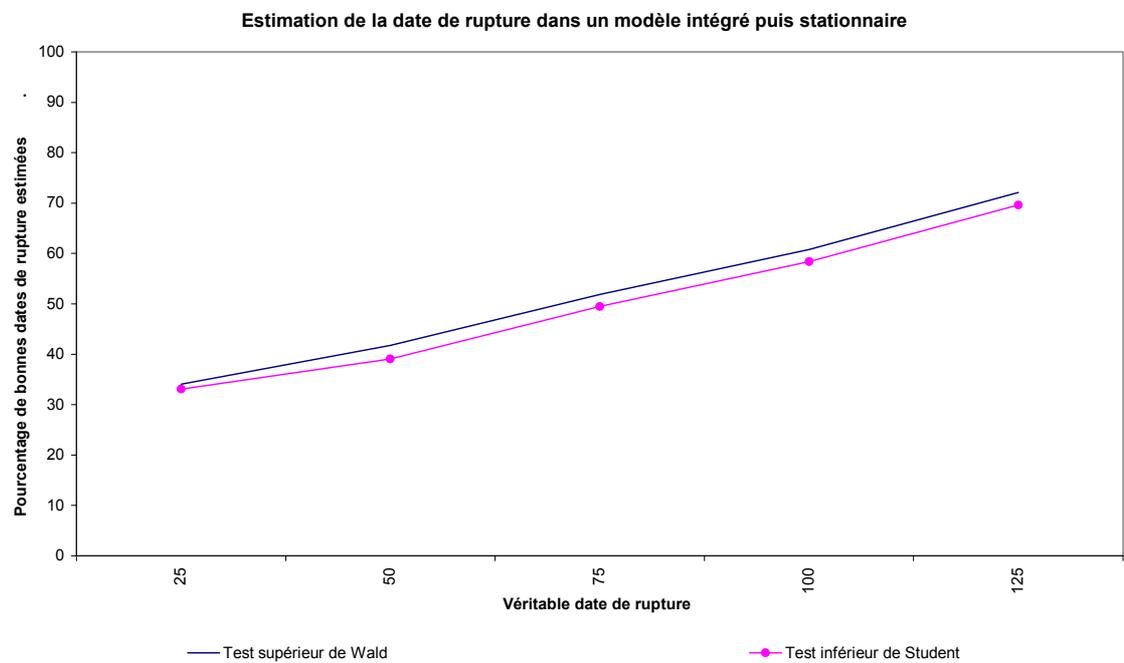
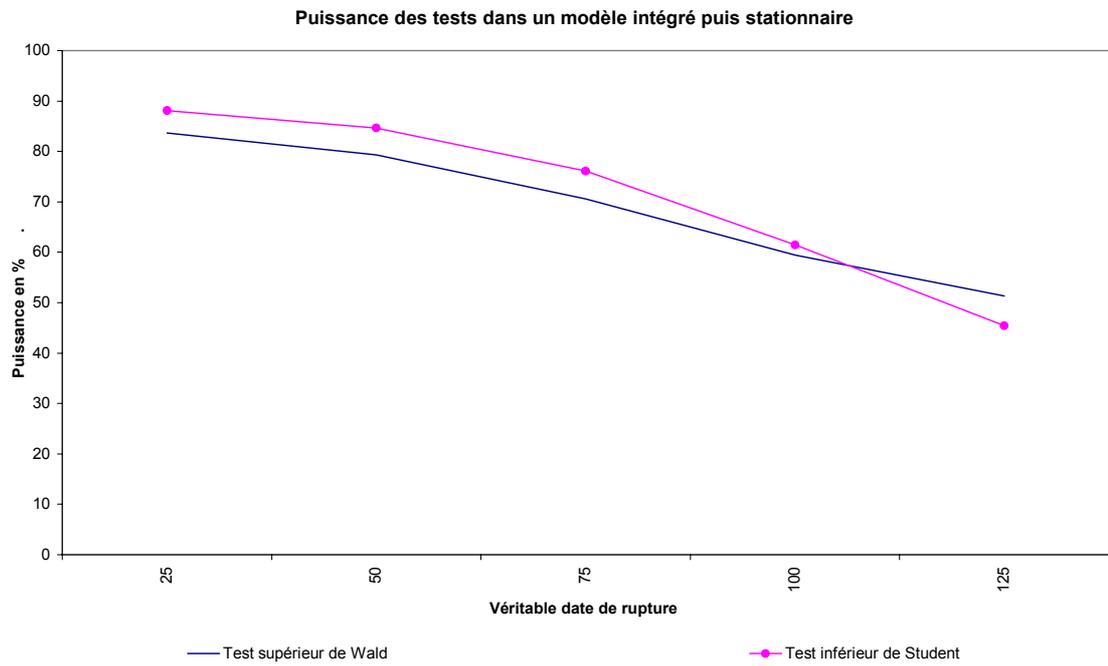
6.3.3.1 Cas d'une série intégrée puis stationnaire

Posons $\rho_1 = 1$, $\rho_2 = 0.8$, $c_1 = 0$, $c_2 = 1$ et calculons la puissance en faisant varier la date de la rupture.

Les principaux résultats sont les suivants :

- Dans le cas d'une série intégrée puis stationnaire, les tests supérieur de Wald et inférieur de Student ont une puissance plus importante quand la rupture a lieu au début de l'échantillon plutôt qu'à la fin. Ainsi, pour un processus généré par $\rho_1 = 1$, $\rho_2 = 0.8$, $c_1 = 0$, $c_2 = 1$, le test supérieur de Wald a une puissance de 83.72% quand la rupture se produit à la date 25, contre 51.32% pour une rupture à la date 125 (respectivement 88.08% et 45.40% pour le test inférieur de Student).
- Au contraire, la qualité de l'estimation de la date de rupture est meilleure quand la rupture a lieu en fin d'échantillon. Le test supérieur de Wald estime 34.07% des bonnes dates de rupture dans le cas d'une rupture à la date 25 contre 72.10% pour la date 125 (respectivement 33.11% contre 69.60% dans le cas d'un test inférieur de Student).
- La puissance du test inférieur de Student est plus grande que celle du test supérieur de Wald (excepté pour une rupture tardive), pour une estimation de la date de rupture de qualité presque équivalente.

Ces résultats de puissance et de qualité de l'estimation de la date de rupture pour différentes dates de rupture peuvent être synthétisés sous forme graphique. Les résultats détaillés sont reportés dans l'annexe C.4 page 591.



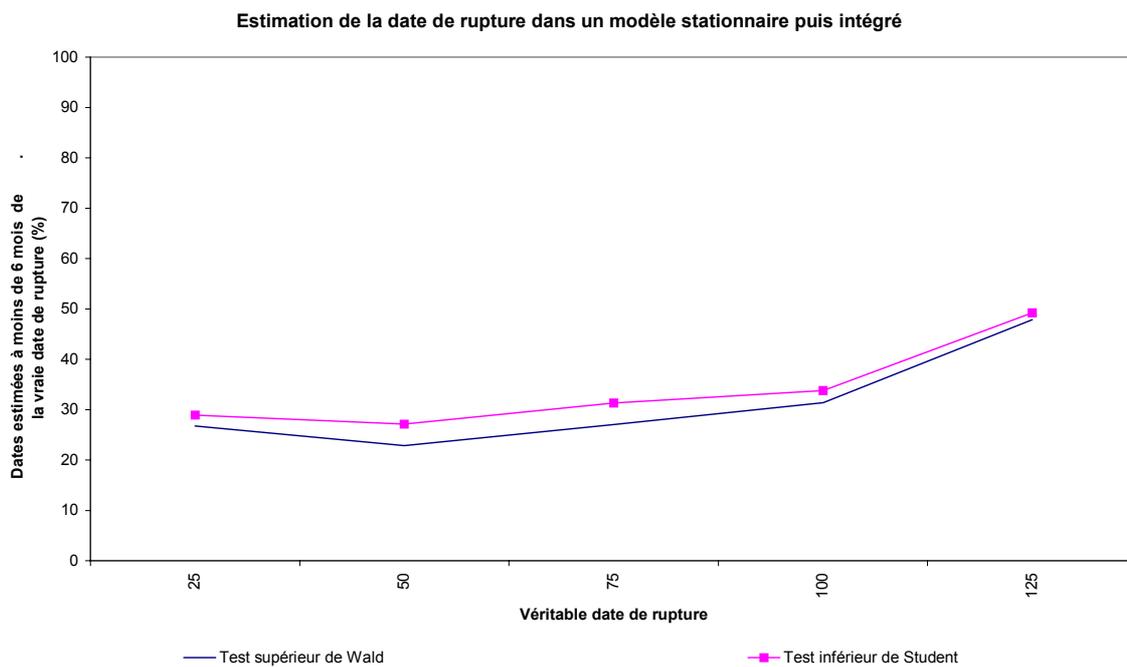
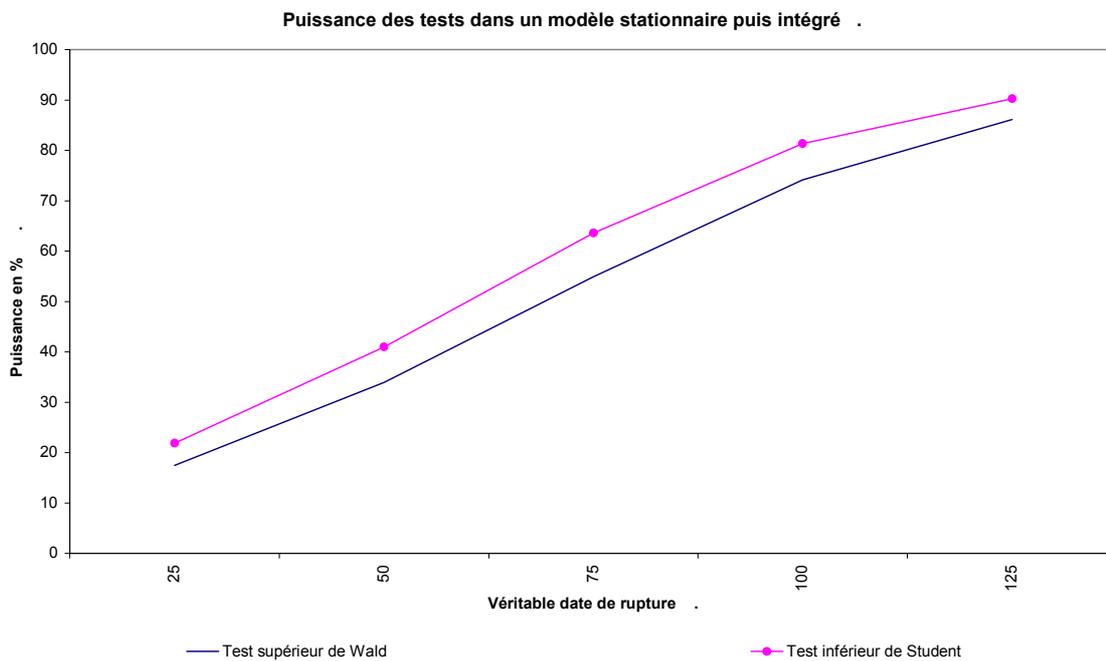
6.3.3.2 Cas d'une série stationnaire puis intégrée

Soient $\rho_1 = 0.8, \rho_2 = 1, c_1 = 1, c_2 = 0$. La puissance est calculée pour différentes dates de rupture véritables.

Dans le cas d'une série stationnaire puis intégrée, les faits marquants suivants sont mis en valeur :

- Contrairement au cas d'une série intégrée puis stationnaire, la puissance des tests supérieur de Wald et inférieur de Student augmente au fur et à mesure que la rupture est tardive. Ainsi, pour un processus généré par $\rho_1 = 0.8, \rho_2 = 1, c_1 = 1, c_2 = 0$, le test supérieur de Wald a une puissance de 17.48% quand la rupture se produit à la date 25, contre 86.12% pour une rupture à la date 125 (respectivement 21.84% et 90.28% pour le test inférieur de Student).
- L'estimation de la véritable date de rupture est mauvaise quel que soit le test considéré, elle augmente au fur et à mesure que la date de rupture augmente dans le temps. Cet accroissement de la qualité d'estimation quand la rupture est tardive s'explique par le fait que les tests ont tendance, dans le cas d'une série stationnaire puis intégrée, à estimer la date de rupture par la dernière date de rupture envisageable.
- Les tests ayant pour hypothèse alternative "la série est stationnaire puis intégrée avec éventuelle rupture de la tendance" ont une puissance assez similaire au cas d'un test sur une série "intégrée puis stationnaire". Par contre, la qualité d'estimation de la date de rupture est détériorée.

Les simulations de puissance effectuées pour le test supérieur de Wald ayant pour hypothèse alternative "la série est stationnaire puis intégrée avec rupture éventuelle de la constante" sont reportées dans le tableau de l'annexe C.4 page 592.



* * *

Ce chapitre avait pour objet de construire des tests de modification du degré d'intégration de la série, sous l'hypothèse nulle de racine unitaire sur l'ensemble de la période considérée.

Deux types de tests ont été construits :

- un test ayant pour hypothèse alternative "la série est intégrée puis stationnaire"
- un test ayant pour hypothèse alternative "la série est stationnaire puis intégrée".

Pour chacune de ces alternatives, deux tests peuvent être envisagés :

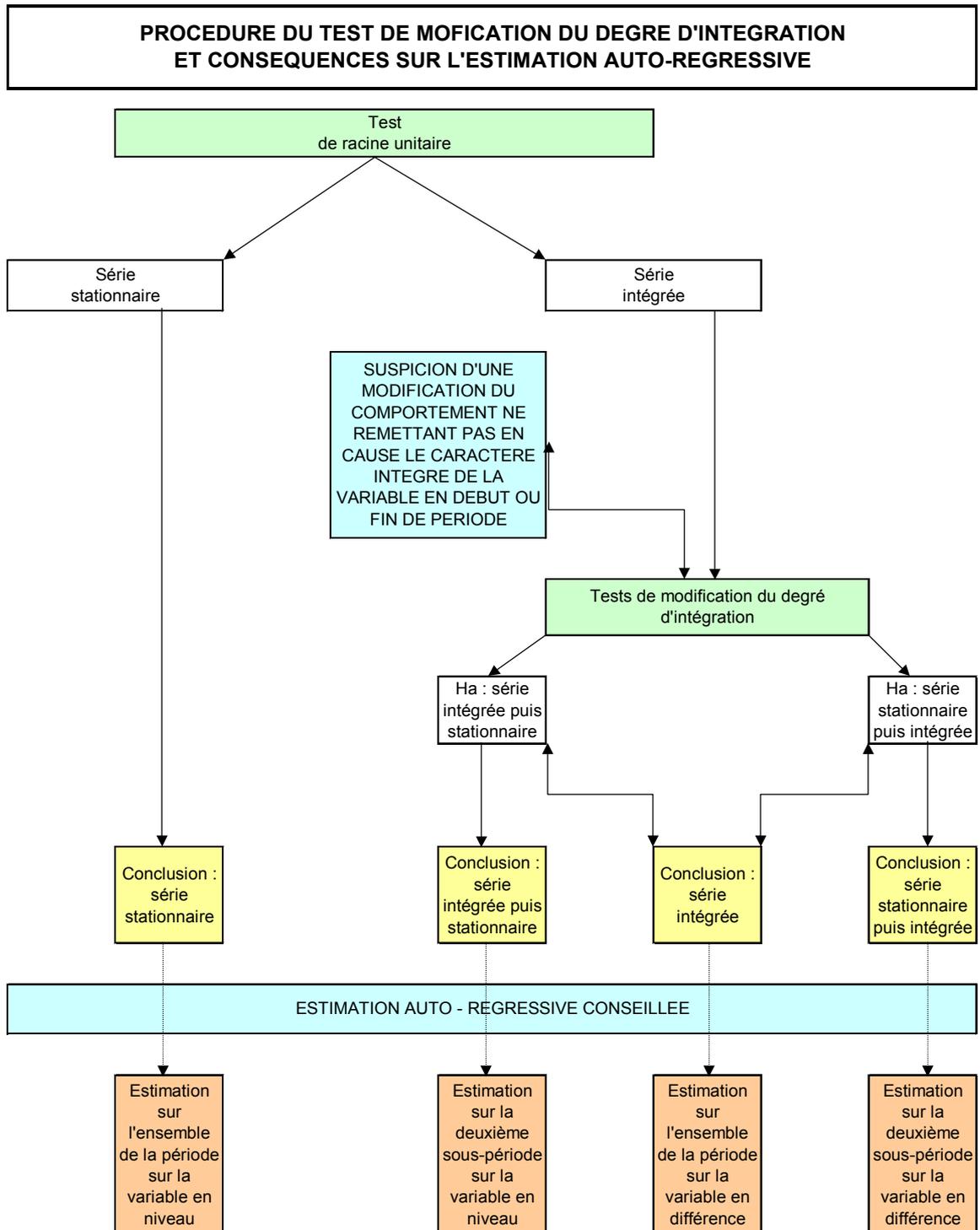
- le test supérieur de Wald, qui teste à la fois la modification de la constante et la modification du degré d'intégration
- et un test inférieur de Student, qui permet une modification de la constante, mais ne s'intéresse qu'à la modification du degré d'intégration.

Ainsi, si le test supérieur de Wald conclut à une modification du degré d'intégration mais que cette conclusion n'est pas validée par le test inférieur de Student, il est probable que le niveau de la série a changé, éventuellement le paramètre auto-régressif d'ordre 1, mais que cette variation ne constitue pas une modification du degré d'intégration au sens propre.

Pour tester uniquement la modification du degré d'intégration, nous conseillons d'utiliser un test inférieur de Student.

Ces tests ont pris en compte la nécessité de corriger les auto-corrélations résiduelles. Les tables de valeurs critiques ont été construites de façon à permettre une utilisation opérationnelle.

Nous avons élaboré une procédure de test de modification du degré d'intégration, présentée à la page suivante. Cette procédure est valide quel que soit le type de test choisi (test supérieur de Wald ou test inférieur de Student). Suivant les conclusions du test, les conséquences sur l'estimation auto-régressive de la série seront différentes.



Les calculs de puissance ont montré les résultats suivants, utiles lors de leurs applications :

- Le test inférieur de Student est un outil plus puissant que le test supérieur de Wald pour détecter la modification du degré d'intégration. Il procure en outre des dates de rupture estimées semblables au test supérieur de Wald.
- Le cas "intégré puis stationnaire" est mieux détecté que l'alternative "stationnaire puis intégré" ; l'estimation de la date de rupture est aussi bien meilleure dans ce premier cas. Dans le deuxième cas, les tests ont tendance à choisir la dernière date de rupture possible. Un examen judicieux des statistiques pour chaque date de rupture devra être effectué de façon à choisir pour date de rupture un extremum local antérieur.
- De façon générale, ces tests ont tendance à réduire la période intégrée : à s'apercevoir un peu tôt du passage d'un régime intégré à un régime stationnaire, ou à se rendre compte très tardivement que le régime est devenu intégré.

Cette procédure de test est appliquée à nos séries dans le chapitre suivant.

Chapitre 7

Application des tests aux données de la RATP et aux séries économiques franciliennes

Les tests construits dans le chapitre précédent sont appliqués aux séries de données relatives au trafic de la RATP ainsi qu'aux séries ayant trait à l'environnement économique en Ile-de-France.

L'application de nos tests est justifiée : les tests standards de racine unitaire concluent à une intégration à l'ordre 1 de chacune des séries, sur l'ensemble de la période janvier 1990 - décembre 2002.

Elle est motivée, pour les séries de trafic, par l'analyse de la stationnarité avant et après les grèves de 1995 effectuée en introduction de cette partie. Les tests de modification du degré d'intégration sont en outre appliqués aux séries décrivant l'environnement économique francilien de façon à permettre une meilleure connaissance des séries qui seront utilisées ultérieurement pour expliquer les trafics.

Pour chacune des séries, les résultats de nos tests sont présentés sous la forme d'un tableau synthétique indiquant la statistique finale, la valeur critique, la date de rupture éventuelle et la conclusion de chacun de ces quatre tests. Sont ensuite représentées graphiquement les valeurs

7.1. APPLICATION DES TESTS DE CHANGEMENTS DU DEGRÉ D'INTÉGRATION AUX DONNÉES DE LA RATP

des statistiques de Wald et de Student pour toutes les dates de rupture possibles, et pour les deux types d'hypothèses alternatives. Ces graphiques ont pour but d'affiner l'estimation de la date de rupture lorsque les tests de modification du degré d'intégration concluent que la série est stationnaire puis intégrée, cas pour lequel l'estimation de la date de rupture est parfois médiocre. Après une brève comparaison des résultats de nos tests avec les conclusions des tests de Perron et des tests standards de racine unitaire (renvoyées en annexe), nous indiquons la modélisation auto-régressive sur l'ensemble de la période ainsi que sur les sous-périodes éventuelles.

La procédure de test, les tableaux et graphiques sont expliqués en détail pour la première série étudiée, à savoir la série de trafic du jour ouvrable sur le métro. Les commentaires sont plus succincts pour les autres séries de la RATP et pour les séries décrivant l'environnement économique francilien.

7.1 Application des tests de changements du degré d'intégration aux données de la RATP

Dans un premier temps, nous avons appliqué nos tests de modification du degré d'intégration avec éventuelle rupture de la constante aux séries de trafic sur les différents réseaux de la RATP.

7.1.1 Trafic du jour ouvrable sur le métro

Les tests de modification du degré d'intégration sont appliqués en considérant successivement les deux hypothèses alternatives suivantes :

- la série est "intégrée puis stationnaire",
- la série est "stationnaire puis intégrée".

Pour chacun de ces cas, le tableau indique la statistique de test calculée pour le test supérieur de Wald et pour le test inférieur de Student, les valeurs critiques associées au seuil de 10%, et la date de la rupture éventuelle pour laquelle les tests atteignent un extrémum. La dernière colonne indique la conclusion de chaque test. Rappelons que si la statistique finale du test supérieur de Wald (respectivement, du test inférieur de Student) est supérieure (respectivement inférieure) à

CHAPITRE 7. APPLICATION DES TESTS AUX DONNÉES DE LA RATP ET AUX
SÉRIES ÉCONOMIQUES FRANCILIENNES

la valeur critique associée, le test rejette l'hypothèse nulle de présence d'une racine unitaire sur l'ensemble de la période.

$\ln(jomsa)$	Statistique	VC 10%	Date	Conclusion
Cas intégré puis stationnaire				
Wald supérieur	13.9428	16.7610	octobre 1997	I(1)-I(1)
Student inférieur	-2.5106	-3.7730	juin 2000	I(1)-I(1)
Cas stationnaire puis intégré				
Wald supérieur	8.3162	12.9260	octobre 1997	I(1)-I(1)
Student inférieur	-2.2890	-3.2940	mai 1991	I(1)-I(1)

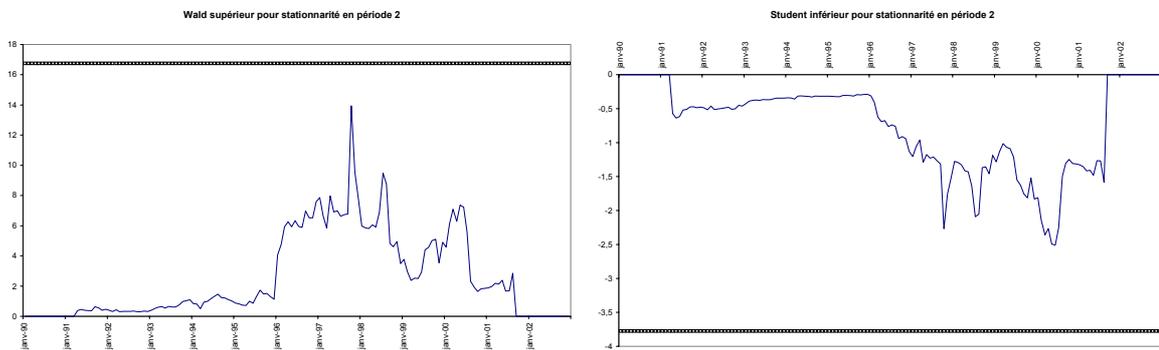
Pour la série du trafic du jour ouvrable sur le métro, les tests de modification du degré d'intégration ne permettent pas de rejeter l'hypothèse nulle au seuil de 10%. Les dates pour lesquelles la statistique finale atteint un extrêimum ne correspondent pas à une date de rupture, mais mettent en relief un évènement particulier, qui n'a pas eu d'impact sur le degré d'intégration de la série.

Les graphiques suivant présentent les statistiques (test de Wald, test de Student) sous l'hypothèse alternative retenue (série "intégrée puis stationnaire" ou "stationnaire puis intégrée"), pour chaque date de rupture possible. Ces graphiques ont double vocation :

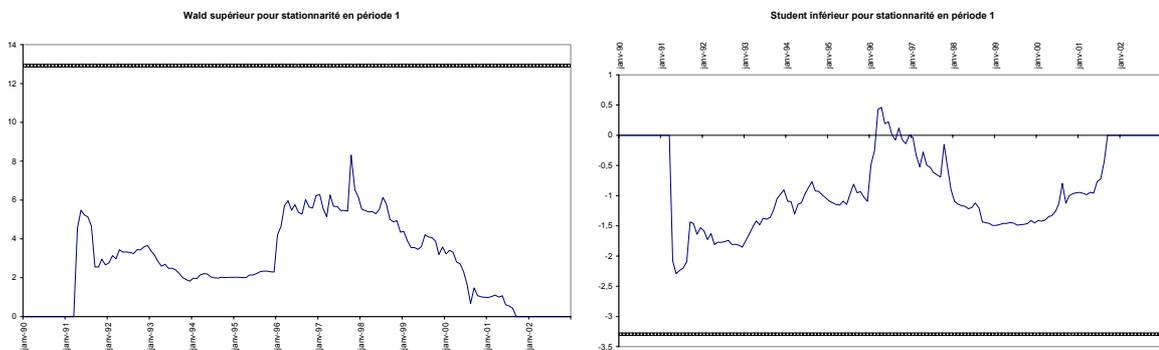
- permettre de repérer des chocs ponctuels de grande importance (les extrema locaux)
- affiner la date de rupture si le test conclut que la série est "stationnaire puis intégrée". Les calculs de puissance effectués au chapitre précédent ont mis en avant une lacune des tests sous cette hypothèse alternative : s'ils repèrent de façon satisfaisante la présence d'une rupture du degré d'intégration, la date de rupture estimée en prenant la date associée à la valeur extrême du test est souvent postérieure à la véritable date de rupture. L'examen des statistiques à chaque date de rupture possible permet de déterminer la position la plus probable de la rupture.

7.1. APPLICATION DES TESTS DE CHANGEMENTS DU DEGRÉ D'INTÉGRATION AUX DONNÉES DE LA RATP

Cas intégré puis stationnaire



Cas stationnaire puis intégré



Pour le trafic du jour ouvrable, le seul graphique des statistiques de Wald sur toutes les dates de rupture possibles indique les 3 points d'intérêt : octobre 1997, l'été 1998 et mai 2000. Ces deux dernières dates précèdent la mise en place de la carte imagine R, et de son dézouage au cours des petites vacances scolaires. Les courbes ne franchissent pas la droite en gras, qui représente la valeur critique associée à chaque test extrême.

Comparaison avec les tests de Perron Les tests de Perron ont pour hypothèse nulle l'existence d'une marche aléatoire et pour hypothèse alternative l'existence de stationnarité par morceaux, autour d'une tendance et /ou d'un niveau soumis à une rupture à une date connue ou inconnue. Ces tests sont expliqués en détail dans l'annexe D page 593, et appliqués à chacune des séries de trafic et d'environnement. Ils mettent en évidence une date de rupture éventuelle de la tendance et/ou du niveau et analysent la stationnarité autour de ce terme déterministe soumis

à rupture. Nous avons utilisé cette partition pour vérifier le degré d'intégration sur chacune des sous-périodes dégagées.

Pour la série de trafic du jour ouvrable sur le métro, le détail des tests de Perron et l'analyse de la stationnarité par sous-période sont reportés page 602. Ce paragraphe a pour but de comparer les résultats de nos tests d'intégration aux conclusions des tests de Perron, et d'en montrer la pertinence.

Le test de Perron avec simple rupture sur la constante ne permet pas de remettre en cause l'hypothèse de racine unitaire sur une partition en juillet 1998. Le test de Perron conclut à la stationnarité autour d'une constante et d'une pente ayant subi une rupture en novembre 1995. L'analyse de la stationnarité sur les deux sous-périodes (avant - après novembre 1995) ne permet pas de conclure à la stationnarité par sous-périodes. Nos tests de changement du degré d'intégration autour d'une constante avec rupture confirme le caractère intégré de la série. Ils suggèrent que si une rupture devait être prise en compte, la date de rupture serait postérieure aux grèves de l'hiver 1995 : elle serait à mettre en lien avec la reprise de la croissance économique et du tourisme en Ile-de-France, ainsi qu'à la fin de la rémanence des grèves.

Modélisation auto-régressive du trafic du jour ouvrable sur le métro Les tests de modification du degré d'intégration ont conclu à la présence d'une racine unitaire sur l'ensemble de la période, et à l'absence de rupture. La modélisation auto-régressive du trafic du jour ouvrable sur le métro est réalisée sur la période entière, en prenant la variable différenciée :

Equation retenue :

$$\begin{aligned} \Delta \ln(jomsa)_t = & 0.0005 - 0.4511 \Delta \ln(jomsa)_{t-1} \\ & (0.4209) \quad (-5.6139) \\ & - 0.2757 \Delta \ln(jomsa)_{t-2} - 0.1503 \Delta \ln(jomsa)_{t-3} \\ & (-3.2289) \quad \quad \quad (-1.8703) \end{aligned}$$

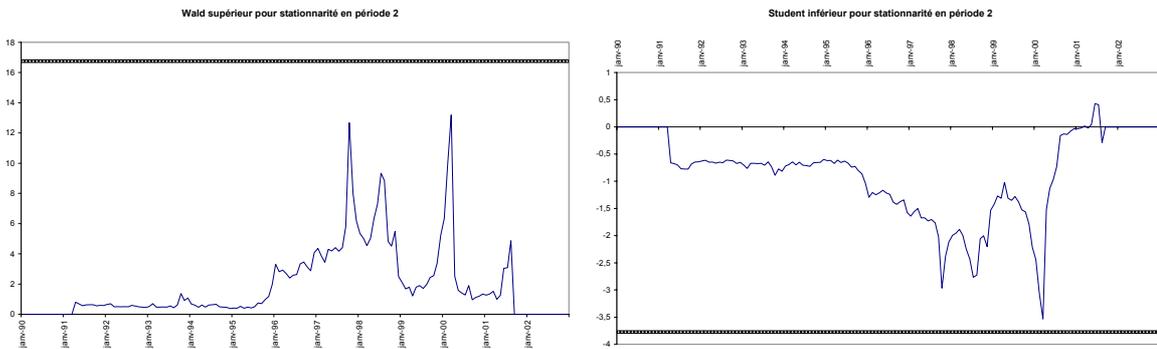
7.1. APPLICATION DES TESTS DE CHANGEMENTS DU DEGRÉ D'INTÉGRATION AUX DONNÉES DE LA RATP

7.1.2 Trafic du jour ouvrable sur le RER

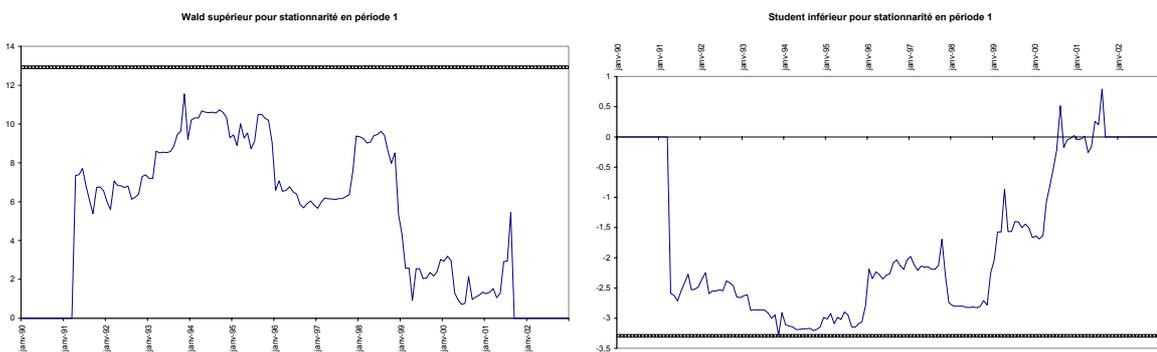
Application des tests de modification du degré d'intégration

$\ln(jorersa)$	Statistique	VC 10%	Date	Conclusion
Cas intégré puis stationnaire				
Wald supérieur	13.1992	16.7610	mars 2000	I(1)-I(1)
Student inférieur	-3.5323	-3.7730	mars 2000	I(1)-I(1)
Cas stationnaire puis intégré				
Wald supérieur	11.5515	12.9260	novembre 1993	I(1)-I(1)
Student inférieur	-3.2956	-3.2940	novembre 1993	I(0)-I(1)

Cas intégré puis stationnaire



Cas stationnaire puis intégré



Sur le dernier graphique, représentant les statistiques du test de Student pour l'alternative "la série est stationnaire puis intégrée", la modification du degré d'intégration se traduit gra-

phiquement par le passage de la courbe des statistiques à chaque date de rupture éventuelle en dessous de la valeur critique à 10%, représentée par la droite en gros pointillés.

Comparaison avec les tests de Perron Pour la série de trafic du jour ouvrable sur le réseau du RER, les tests de Perron (résultats détaillés page 604) conduisent aux conclusions suivantes :

- l'hypothèse de racine unitaire autour d'une constante avec rupture en octobre 1998 ne peut être écartée.
- La série semble stationnaire autour d'une tendance avec rupture de la pente en juillet 1998.

Pour ces deux partitions, les résultats des tests de racine unitaire ne permettent ni d'infirmer ni de confirmer ces résultats, mais amènent à se demander si la série n'est pas stationnaire puis intégrée. Au seuil de 10%, la série est stationnaire puis intégrée, pour une date de rupture en novembre 1993¹. Comme l'indique l'historique, la fin d'année 1993 est marquée par de nombreuses grèves, même si leur ampleur est faible. Notons que l'année 1998 correspond à un extrênum local de la statistique, signe de modifications importantes (mise en place de l'abonnement Imagine R par exemple, mais aussi coupe du monde de football).

Modélisation auto-régressive du trafic du jour ouvrable sur le RER Sur la période entière, la modélisation auto-régressive du trafic du jour ouvrable sur le RER s'écrit :

$$\begin{aligned} \Delta \ln(jorersa)_t = & 0.0013 - 0.3597 \Delta \ln(jorersa)_{t-1} \\ & (1.0198) \quad (-4.4758) \\ & - 0.2278 \Delta \ln(jorersa)_{t-2} - 0.1423 \Delta \ln(jorersa)_{t-3} \\ & (-2.7447) \quad (-1.7809) \end{aligned}$$

¹Au seuil de 20%, la série est en outre testée "intégrée puis stationnaire", pour une date de rupture estimée en mars 2000. Les tests de racine unitaire confirment l'existence de trois sous-périodes :

- janvier 1990 - novembre 1993 : la série est stationnaire,
- décembre 1993 - mars 2000 : la série est intégrée à l'ordre 1,
- avril 2000 - décembre 2002 : la série est de nouveau stationnaire.

7.1. APPLICATION DES TESTS DE CHANGEMENTS DU DEGRÉ D'INTÉGRATION AUX DONNÉES DE LA RATP

Les tests de modification du degré d'intégration au seuil de 10% indiquent une rupture en novembre 1993. Il apparaît plus judicieux de prendre en compte la modélisation sur la seconde sous-période uniquement :

Equation retenue :

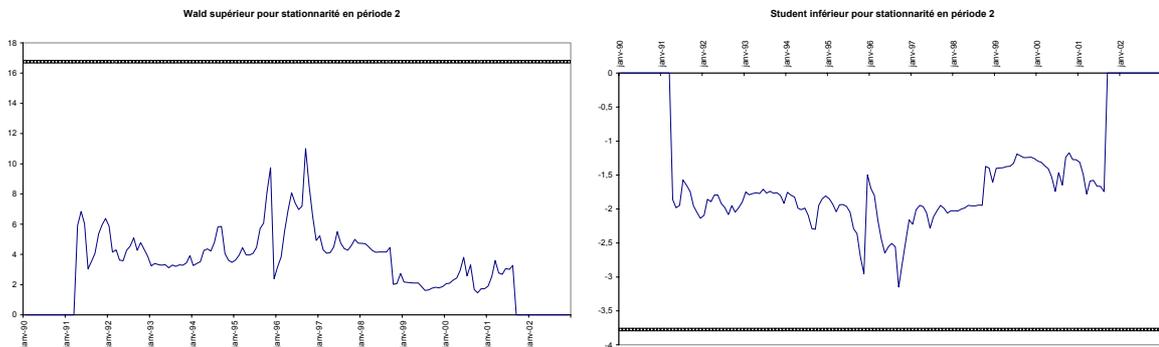
$$\begin{aligned} \Delta \ln(jorersa)_t = & 0.0018 - 0.3623 \Delta \ln(jorersa)_{t-1} \\ & (1.1832) \quad (-3.8573) \\ & - 0.1631 \Delta \ln(jorersa)_{t-2} - 0.1082 \Delta \ln(jorersa)_{t-3} \\ & (-1.6945) \quad (-1.1704) \end{aligned}$$

7.1.3 Trafic du jour ouvrable sur le réseau des bus de Paris

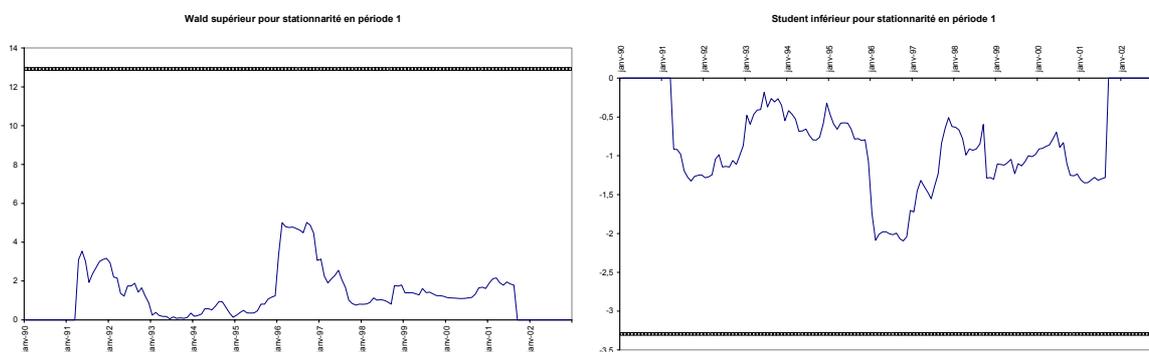
Application des tests de modification du degré d'intégration

ln(<i>joparsa</i>)	Statistique	VC 10%	Date	Conclusion
Cas intégré puis stationnaire				
Wald supérieur	11.0088	16.7610	septembre 1996	I(1)-I(1)
Student inférieur	-3.1494	-3.7730	septembre 1996	I(1)-I(1)
Cas stationnaire puis intégré				
Wald supérieur	5.0226	12.9260	septembre 1996	I(1)-I(1)
Student inférieur	-2.0961	-3.2940	octobre 1996	I(1)-I(1)

Cas intégré puis stationnaire



Cas stationnaire puis intégré



Comparaison avec les tests de Perron Pour la série de trafic du jour ouvrable sur le réseau des autobus de Paris, le test de Perron avec rupture de la constante conclut que la série est stationnaire autour d'une constante ayant subi une rupture en juillet 2000, et stationnaire autour d'une tendance dont la pente se serait modifiée en août 1996. Les résultats détaillés de ces tests sont reportés dans l'annexe D.3.2 page 605. Les tests de racine unitaire confirment que la série n'est pas intégrée sur l'ensemble de la période, mais plutôt intégrée puis stationnaire.

Nos tests n'ont pas permis de confirmer un tel changement de degré d'intégration.

Modélisation auto-régressive du trafic du jour ouvrable sur le bus de Paris Sur la période entière, la modélisation auto-régressive du trafic du jour ouvrable sur le réseau des autobus de Paris s'écrit :

Equation retenue :

$$\begin{aligned} \Delta \ln(joparsa)_t = & 0.0009 - 0.6714 \Delta \ln(joparsa)_{t-1} \\ & (0.6941) \quad (-8.4310) \\ & - 0.5021 \Delta \ln(joparsa)_{t-2} - 0.2181 \Delta \ln(joparsa)_{t-3} \\ & (-5.7478) \quad (-2.7317) \end{aligned}$$

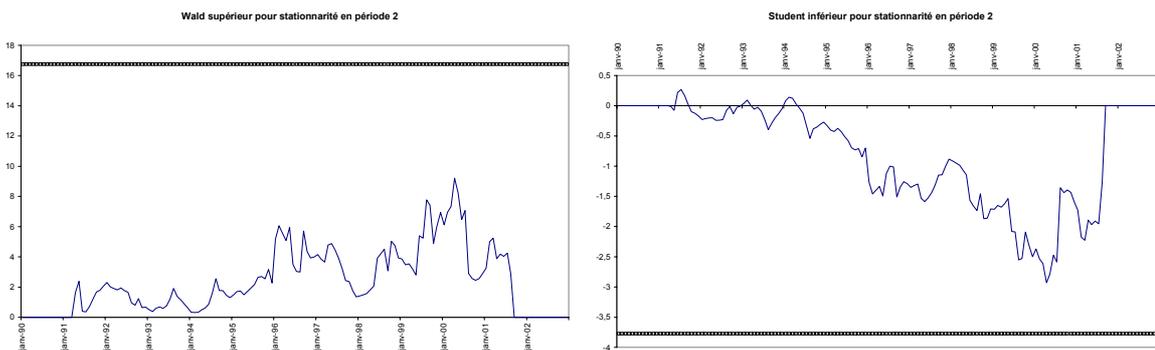
7.1. APPLICATION DES TESTS DE CHANGEMENTS DU DEGRÉ D'INTÉGRATION AUX DONNÉES DE LA RATP

7.1.4 Trafic du jour ouvrable sur le réseau des bus de banlieue

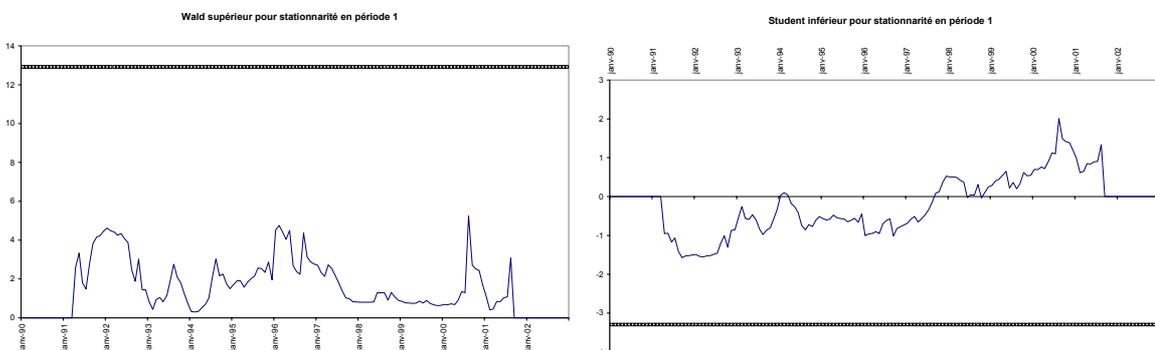
Application des tests de modification du degré d'intégration

$\ln(jobansa)$	Statistique	VC 10%	Date	Conclusion
Cas intégré puis stationnaire				
Wald supérieur	9.2133	16.7610	avril 2000	I(1)-I(1)
Student inférieur	-2.9312	-3.7730	avril 2000	I(1)-I(1)
Cas stationnaire puis intégré				
Wald supérieur	5.2442	12.9260	août 2000	I(1)-I(1)
Student inférieur	-1.5712	-3.2940	septembre 1991	I(1)-I(1)

Cas intégré puis stationnaire



Cas stationnaire puis intégré



Comparaison avec les tests de Perron Le test de Perron (détails page 606) avec rupture de la constante ne permet pas de remettre en cause l'hypothèse de racine unitaire avant et après avril 1994, date de rupture estimée. Une éventuelle rupture de la tendance amène le test de Perron à la conclusion d'une série stationnaire pour la partition au mois de novembre 1995.

Pourtant, les tests standards de racine unitaire concluent à la présence de racine unitaire sur chaque sous-période.

Nos tests confirment l'absence de changement du degré d'intégration des séries : la série est intégrée sur l'ensemble de la période. La présence d'un extrêmmum local à la fin de l'année 1995 confirme que cette période a connu de fortes perturbations. Les dates de rupture éventuelles estimées par nos tests sont en majorité plus tardives que celles issues des tests de Perron et correspondent à la fin d'une grande opération de renouveau du réseau des autobus de banlieue. En particulier, en avril 2000, le réseau des autobus du Sud des Hauts de Seine et du Nord de l'Essonne est en pleine restructuration.

Modélisation auto-régressive du trafic du jour ouvrable sur le réseau des autobus de banlieue Sur la période entière, la modélisation auto-régressive du trafic du jour ouvrable sur le réseau des autobus de banlieue s'écrit :

Equation retenue :

$$\begin{aligned} \Delta \ln(\text{jobansa})_t = & 0.0034 - 0.4956 \Delta \ln(\text{jobansa})_{t-1} \\ & (2.9430) \quad (-6.1720) \\ & - 0.3525 \Delta \ln(\text{jobansa})_{t-2} - 0.2052 \Delta \ln(\text{jobansa})_{t-3} \\ & (-4.1402) \quad (-2.5337) \end{aligned}$$

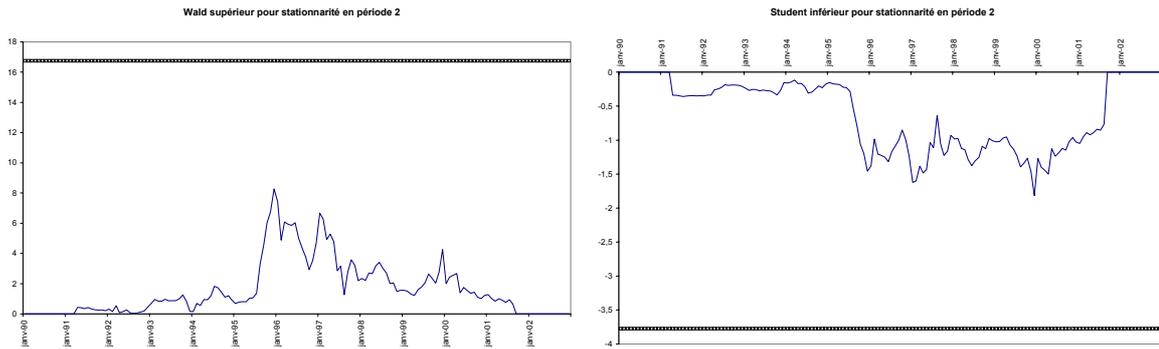
7.1. APPLICATION DES TESTS DE CHANGEMENTS DU DEGRÉ D'INTÉGRATION AUX DONNÉES DE LA RATP

7.1.5 Trafic du samedi sur le métro

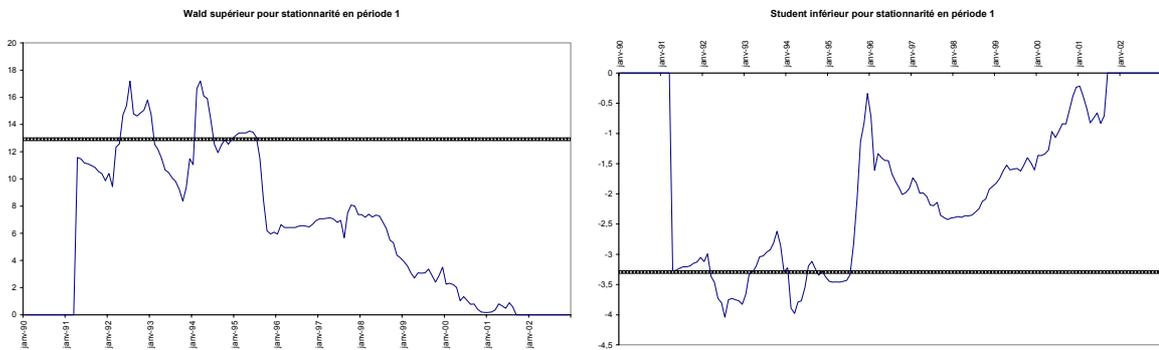
Application des tests de modification du degré d'intégration

$\ln(sm_{sa})$	Statistique	VC 10%	Date	Conclusion
Cas intégré puis stationnaire				
Wald supérieur	8.2731	16.7610	décembre 1995	I(1)-I(1)
Student inférieur	-1.8176	-3.7730	décembre 1999	I(1)-I(1)
Cas stationnaire puis intégré				
Wald supérieur	17.2016	12.9260	mars 1994	I(0)-I(1)
Student inférieur	-4.0421	-3.2940	juillet 1992	I(0)-I(1)

Cas intégré puis stationnaire



Cas stationnaire puis intégré



Comparaison avec les tests de Perron Le test de Perron avec rupture sur le niveau de la constante (date estimée : décembre 1993) ne permet pas de remettre en cause le caractère intégré de la série du trafic du samedi sur le métro. La présence d'une tendance avec rupture de la pente et du niveau (date estimée : juin 1995) amène à conclure que la série est stationnaire autour de cette tendance "flexible". Néanmoins, l'application des tests standards de racine unitaire laisse penser que, pour ces partitions, la série est stationnaire puis intégrée. L'ensemble de ces résultats est détaillé page 607.

Ce changement du degré d'intégration de la série est confirmé par l'application de nos tests dans le cas stationnaire puis intégré. La date de rupture estimée par le test inférieur de Student correspond au mois de juillet 1992, peu de temps après le prolongement de la ligne 1 du métro jusqu'à la Défense. Le test supérieur de Wald estime la date de rupture au mois de mars 1994, extrênum local pour le test inférieur de Student et mois de grèves à la RATP.

Modélisation auto-régressive du trafic du samedi sur le métro Sur la période entière, la modélisation auto-régressive du trafic du samedi sur le métro s'écrit :

$$\begin{aligned} \Delta \ln (smsa)_t &= 0.0018 - 0.68843 \Delta \ln (smsa)_{t-1} - 0.2987 \Delta \ln (smsa)_{t-2} \\ &\quad (0.7449) \quad (-8.8255) \quad (-3.8053) \\ \hat{\sigma} &= 0.030007 \end{aligned}$$

Du fait de la rupture, il est plus judicieux de considérer la modélisation sur la seconde sous-période uniquement :

Equation retenue :

$$\begin{aligned} \Delta \ln (smsa)_t &= 0.0030 - 0.6184 \Delta \ln (smsa)_{t-1} - 0.1973 \Delta \ln (smsa)_{t-2} \\ &\quad (1.1152) \quad (-6.6850) \quad (-2.1311) \\ \hat{\sigma} &= 0.027614 \end{aligned}$$

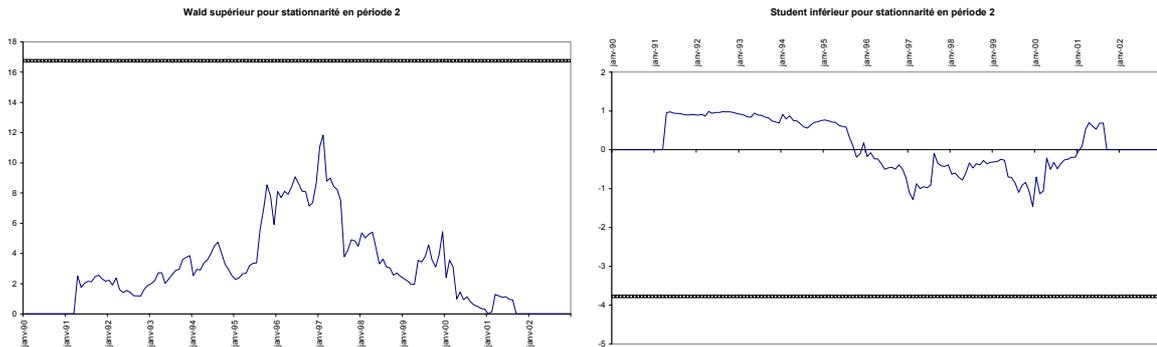
7.1. APPLICATION DES TESTS DE CHANGEMENTS DU DEGRÉ D'INTÉGRATION AUX DONNÉES DE LA RATP

7.1.6 Trafic du dimanche/jour férié sur le métro

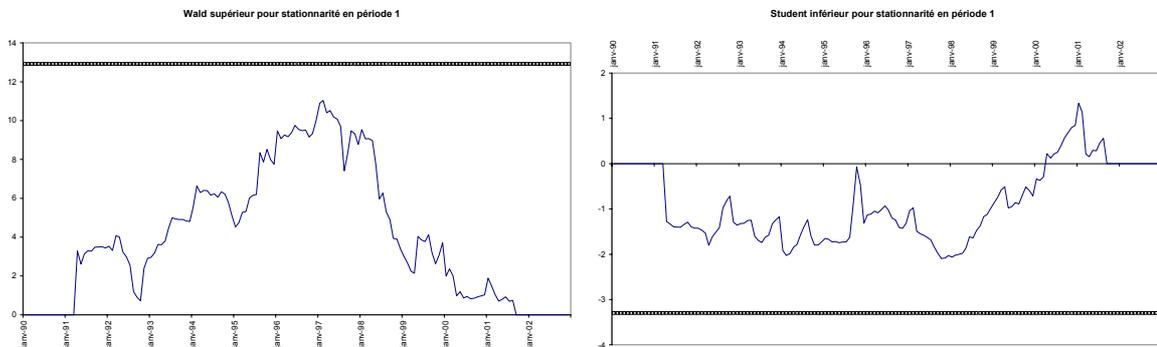
Application des tests de modification du degré d'intégration

$\ln(dfmsa)$	Statistique	VC 10%	Date	Conclusion
Cas intégré puis stationnaire				
Wald supérieur	11.8270	16.7610	février 1997	I(1)-I(1)
Student inférieur	-1.4589	-3.7730	décembre 1999	I(1)-I(1)
Cas stationnaire puis intégré				
Wald supérieur	11.0418	12.9260	février 1997	I(0)-I(1)
Student inférieur	-2.0953	-3.2940	octobre 1997	I(1)-I(1)

Cas intégré puis stationnaire



Cas stationnaire puis intégré



Comparaison avec les tests de Perron Les tests de Perron et la vérification de la stationnarité sur les sous-périodes dégagées pour le trafic du dimanche/jour férié sur le métro sont reportés en annexe à la page 608. Le test de Perron avec simple rupture de niveau conclut à la stationnarité de la série du trafic les dimanches ou jours fériés sur le métro autour de cette constante ayant subi une rupture en janvier 1994. Par contre, la série semble stationnaire autour d'une tendance dont le niveau et la pente auraient été soumis à une rupture en juin 1995 (date estimée). Les tests standards de racine unitaire concluent quant à eux à la stationnarité puis à la présence d'une racine unitaire.

Le test supérieur de Wald ayant pour hypothèse alternative "la série est stationnaire puis intégrée" détecte une rupture du paramètre auto-régressif et/ou de la constante au début de l'année 1997, i.e. après le retour des touristes à Paris et du retour de la croissance économique. Le test inférieur de Student conclut pourtant qu'il n'y a pas eu de modification du degré d'intégration : il y a eu une modification du niveau des déplacements effectués le dimanche sur le métro, sans doute à mettre en lien avec Imagine R quelques mois plus tard, mais pas de réelle modification du phénomène de mémoire des consommateurs.

Modélisation auto-régressive du trafic du dimanche/jour férié sur le métro La représentation auto-régressive du trafic du dimanche/jour férié sur le métro s'écrit, pour la période entière :

Equation retenue :

$$\begin{aligned}
 \Delta \ln(dfmsa)_t = & 0.0040 - 0.6434 \Delta \ln(dfmsa)_{t-1} \\
 & (1.6109) \quad (-7.8413) \\
 & - 0.3835 \Delta \ln(dfmsa)_{t-2} - 0.2084 \Delta \ln(dfmsa)_{t-3} \\
 & (-3.9801) \quad (-2.1611) \\
 & - 0.1515 \Delta \ln(dfmsa)_{t-4} \\
 & (-1.8465)
 \end{aligned}$$

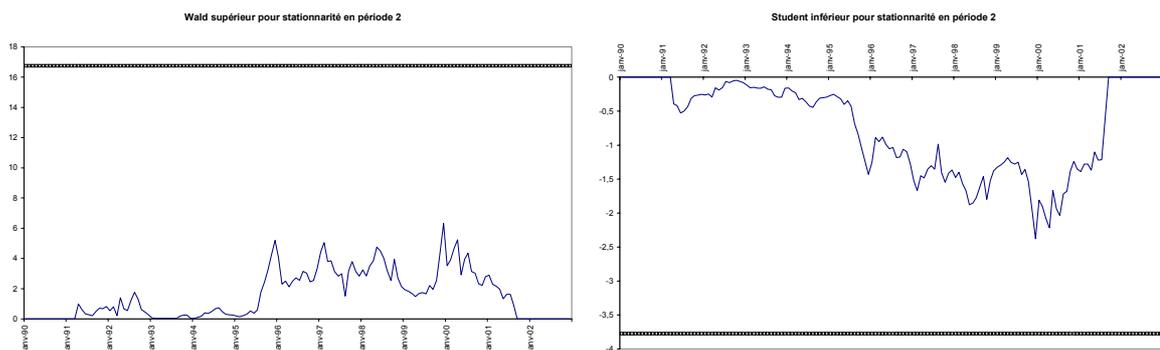
7.1. APPLICATION DES TESTS DE CHANGEMENTS DU DEGRÉ D'INTÉGRATION AUX DONNÉES DE LA RATP

7.1.7 Trafic du samedi sur le RER

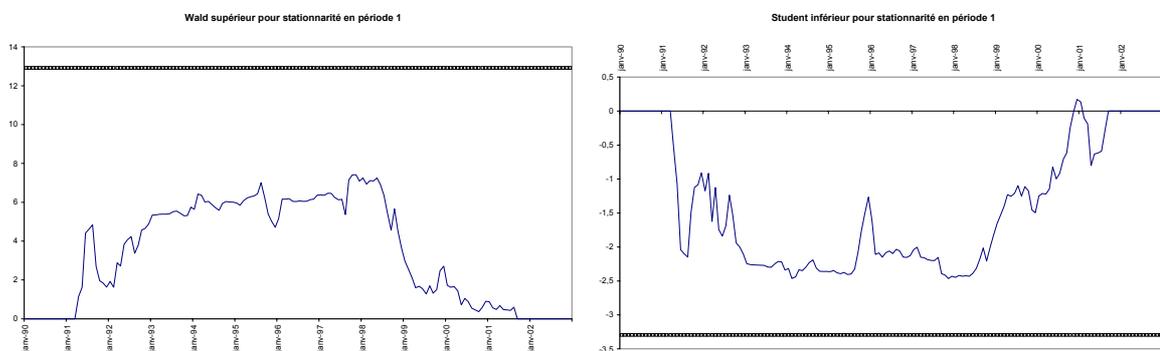
Application des tests de modification du degré d'intégration

$\ln(srersa)$	Statistique	VC 10%	Date	Conclusion
Cas intégré puis stationnaire				
Wald supérieur	6.3191	16.7610	décembre 1999	I(1)-I(1)
Student inférieur	-2.3800	-3.7730	décembre 1999	I(1)-I(1)
Cas stationnaire puis intégré				
Wald supérieur	7.4194	12.9260	novembre 1997	I(1)-I(1)
Student inférieur	-2.4644	-3.2940	novembre 1997	I(1)-I(1)

Cas intégré puis stationnaire



Cas stationnaire puis intégré



Comparaison avec les tests de Perron Les tests de Perron concluent que la série de trafic du samedi sur le RER est intégrée sur toute la période, même si l'on permet l'existence d'une rupture du niveau ou de la pente de la tendance (voir page 609). Ce résultat est confirmé par l'application des tests augmentés de Dickey-Fuller sur les deux partitions éventuelles en décembre 1993 ou mai 1995.

Nos tests ont une conclusion identique : la série est intégrée sur l'ensemble de la période. Les dates mises en relief sont novembre 1997 ou décembre 1999, mais l'analyse graphique des statistiques de Wald et de Student par date indiquent d'autres extrema locaux qui pourraient être pris en considération.

Pour le cas "stationnaire puis intégré", il existe un plateau élevé entre janvier 1993 et juillet 1995 (ce qui englobe les dates trouvées par l'application des tests de Perron et correspond à une période économique morose), puis entre septembre 1997 et mai 1998 (période de la reprise économique et du retour des touristes à Paris).

Modélisation auto-régressive du trafic du samedi sur le RER Sur la période entière, la série décrivant le trafic du samedi sur le RER s'écrit sous la forme du processus auto-régressif en différence suivant :

Equation retenue :

$$\begin{aligned} \Delta \ln (s\text{rersa})_t = & 0.0036 - 0.6740 \Delta \ln (s\text{rersa})_{t-1} - 0.4467 \Delta \ln (s\text{rersa})_{t-2} \\ & (1.2439) \quad (-8.1594) \quad \quad \quad (-4.4885) \\ & - 0.1141 \Delta \ln (s\text{rersa})_{t-3} + 0.0156 \Delta \ln (s\text{rersa})_{t-4} \\ & (-1.1479) \quad \quad \quad (0.1883) \end{aligned}$$

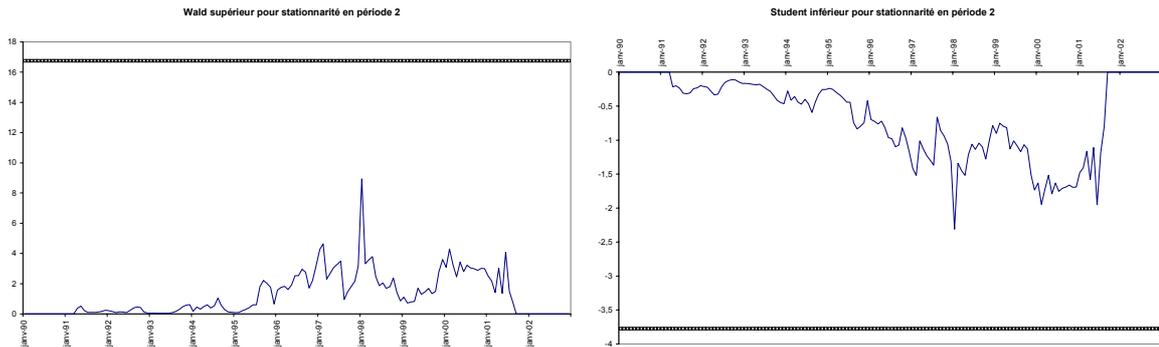
7.1. APPLICATION DES TESTS DE CHANGEMENTS DU DEGRÉ D'INTÉGRATION
AUX DONNÉES DE LA RATP

7.1.8 Trafic du dimanche/jour férié sur le RER

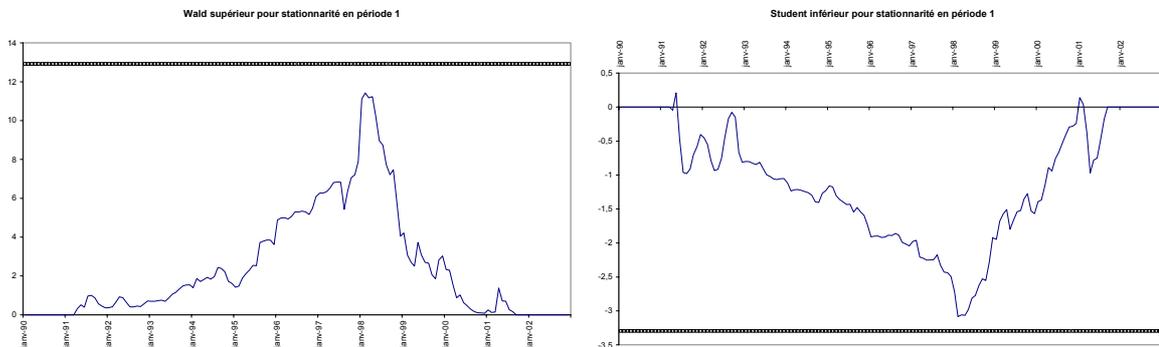
Application des tests de modification du degré d'intégration

$\ln(df\ rersa)$	Statistique	VC 10%	Date	Conclusion
Cas intégré puis stationnaire				
Wald supérieur	8.9359	16.7610	janvier 1998	I(1)-I(1)
Student inférieur	-2.3102	-3.7730	janvier 1998	I(1)-I(1)
Cas stationnaire puis intégré				
Wald supérieur	11.4235	12.9260	février 1998	I(1)-I(1) à 20%
Student inférieur	-3.0825	-3.2940	février 1998	I(0)-I(1) à 20%

Cas intégré puis stationnaire



Cas stationnaire puis intégré



Comparaison avec les tests de Perron Le test de Perron avec rupture sur la constante ne permet pas de rejeter l'hypothèse de racine unitaire sur l'ensemble de la période, la rupture de la constante ayant lieu en janvier 1995. La présence d'une tendance avec rupture sur le niveau et la pente (date estimée en juin 1995) dans le test de Perron permet de conclure à la stationnarité de la série de trafic du dimanche/jour férié sur le RER. L'application des tests de racine unitaire standard sur les partitions dégagées remet en cause cette stationnarité, au moins sur la seconde sous-période. Les tableaux détaillés associés à cette analyse sont disponibles à la page 610.

Nos tests confirment le caractère intégré en seconde période puisqu'ils concluent que la série est stationnaire, puis intégrée. La date de rupture estimée est plus tardive : février 1998. Cette date ressort très clairement dans les graphiques. Elle correspond à la période à laquelle le prolongement de la ligne B du RER à la "Plaine Stade de France" a été ouverte au public.

Modélisation auto-régressive du trafic du dimanche/jour férié sur le RER Sur période entière, la modélisation auto-régressive en différence conduit à l'équation suivante :

$$\begin{aligned} \Delta \ln (dfrersa)_t &= 0.0044 - 0.5451 \Delta \ln (dfrersa)_{t-1} \\ &\quad (1.4664) \quad (-6.6643) \\ &\quad - 0.2149 \Delta \ln (dfrersa)_{t-2} - 0.0964 \Delta \ln (dfrersa)_{t-3} \\ &\quad (-2.3466) \quad (-1.1801) \\ \hat{\sigma} &= 0.036576 \end{aligned}$$

Les tests de modification du degré d'intégration indiquent une rupture en février 1998. Sur la seconde sous-période, le processus auto-régressif en différence s'écrit :

Equation retenue :

$$\begin{aligned} \Delta \ln(df\text{rersa})_t &= 0.0120 - 0.5997 \Delta \ln(df\text{rersa})_{t-1} \\ &\quad (2.9176) \quad (-5.2268) \\ &\quad - 0.3596 \Delta \ln(df\text{rersa})_{t-2} - 0.1239 \Delta \ln(df\text{rersa})_{t-3} \\ &\quad \quad (-2.8688) \quad \quad (-1.0838) \\ \hat{\sigma} &= 0.030176 \end{aligned}$$

7.1.9 La modification du degré de stationnarité des séries de trafic sur les sous-réseaux RATP

Les principaux résultats de cette section sont les suivants :

- Les séries relatives au trafic du jour ouvrable sur le métro, le réseau des autobus de Paris et des autobus de banlieue sont intégrées sur l'ensemble de la période (janvier 1990 - décembre 2002). Sur le réseau routier, les modifications de l'offre de service sont généralement graduelles, et entraînent des modifications lentes des comportements des utilisateurs de ce réseau.
- La série relative au trafic du jour ouvrable sur le RER est stationnaire puis intégrée, pour un test au seuil de 10%, et une date de rupture en novembre 1993.
- Les séries relatives au trafic du samedi sur le métro, et du dimanche sur le RER sont stationnaires puis intégrées au seuil de 10%. La date de rupture du degré d'intégration de la série de trafic du samedi sur le métro est estimée par le mois de juillet 1992, ou, de façon plus probable, en mars 1994. La série relative au trafic du dimanche (ou jour férié) sur le RER a subi une rupture en février 1998, et est stationnaire avant puis intégrée ensuite.
- Enfin, les séries de trafic du dimanche/jour férié sur le métro et du samedi sur le RER sont intégrées sur l'ensemble de la période.

7.2 Application des tests de changements du degré d'intégration aux séries relatives à l'environnement économique en Ile-de- France

Nos tests de changement de l'ordre d'intégration avec éventuelle rupture de la constante sont ensuite appliqués aux séries relatives à l'environnement économique dans la région Ile-de-France.

L'ensemble des résultats a été renvoyé dans l'annexe D.4 page 611.

Les résultats principaux de cette analyse sont les suivants :

- Il existe une série intégrée puis stationnaire : le volume de ventes des carburants en Ile-de-France, pour une rupture en octobre 1997.
- Les séries relatives au volume de circulation dans Paris, ou sur l'ensemble du réseau aggloméré francilien ou le rapport entre le prix du ticket T et du super sont intégrées sur l'ensemble de la période considérée (janvier 1990 - décembre 2002).
- Les séries relatives à la vitesse de circulation dans Paris (rupture en février 1996), aux demandes d'emplois en fin de mois à Paris (rupture en avril 1999) ou en Ile-de-France (rupture en décembre 1991) ou encore au volume des ventes dans les grands magasins en région parisienne (rupture en avril 1992) sont stationnaires puis intégrées. Il en est de même pour le taux d'occupation des hôtels en Ile-de-France, pour une rupture en décembre 1997.
- Enfin, nous avons mis en valeur des séries à la fois "stationnaires puis intégrées" et "intégrées puis stationnaires". Parmi ces séries, le nombre de voyageurs grandes lignes passant par une gare de Paris ou le taux d'occupation des hôtels à Paris peuvent être considérés comme des séries stationnaires par morceaux sur l'ensemble de la période étudiée. Le taux d'occupation des hôtels en Ile-de-France est considéré comme "intégré puis stationnaire avec éventuelle rupture de la constante" par le test supérieur de Wald seulement, pour une date de rupture estimée par le mois de décembre 1995. Ce résultat n'est pas confirmé par le test de Student ayant pour hypothèse alternative "la série est intégrée puis stationnaire",

7.3. LA PERTINENCE DE NOS TESTS DE MODIFICATION DU DEGRÉ D'INTÉGRATION PAR RAPPORT AUX TESTS DE PERRON

laissant à penser que seule la modification de la constante a amené cette conclusion. Le test supérieur de Wald "stationnaire puis intégré" estime une rupture en avril 1991, tandis que le test inférieur de Student juge que la rupture du seul paramètre auto-régressif a été plus tardive, en novembre 1997.

7.3 La pertinence de nos tests de modification du degré d'intégration par rapport aux tests de Perron

Pour les séries étudiées, les résultats des tests de Perron sont en contradiction avec les tests de racine unitaire sur les sous-périodes dégagées. Cette contradiction est corrigée par l'application de nos tests de modification du degré d'intégration.

Résumons cette constatation dans des tableaux.

- La première colonne indique le nom de la série étudiée².
- Les trois colonnes suivantes indiquent les résultats du test A de Perron, qui teste la stationnarité en permettant une rupture du niveau de la série.
 - Dans la colonne "Test A" est indiquée la conclusion de ce premier test.
 - La colonne suivante indique la date de la rupture du niveau de la série.
 - La colonne d'après indique la conclusion des tests de racine unitaire sur les sous-périodes avant et après cette date de rupture.
- Le deuxième bloc indique les mêmes éléments, mais dans le cadre du test C de Perron, qui permet une rupture du niveau mais aussi de la tendance de la série.
- Enfin le troisième bloc indique la conclusion de nos tests de modification du degré d'intégration, ainsi que la date de rupture en cas de modification du degré d'intégration.

²Les sigles sont synthétisés à la fin du chapitre 3 pour les données de trafic et à la fin du chapitre 4 pour les données décrivant l'environnement économique en Ile-de-France.

CHAPITRE 7. APPLICATION DES TESTS AUX DONNÉES DE LA RATP ET AUX
SÉRIES ÉCONOMIQUES FRANCILIENNES

Conclusion des tests sur les séries de trafic								
	Test A	Date	Tests RU	Test C	Date	Tests RU	Nos tests	Date
jomsa	I(1)	juil. 98	I(1)-I(1)	I(0)	nov. 95	I(1)-I(1)	I(1)-I(1)	
jorersa	I(1)	oct. 98	I(0)-I(1)	I(0)	juil. 98	I(0)-I(1)	I(0)-I(1)	nov. 93
joparsa	I(0)	juil. 00	I(1)-I(0)	I(0)	août 96	I(1)-I(0)	I(1)-I(1)	
jobansa	I(1)	avr. 94	I(1)-I(1)	I(0)	nov. 95	I(1)-I(1)	I(1)-I(1)	
smsa	I(1)	déc. 93	I(0)-I(1)	I(0)	juin 95	I(0)-I(1)	I(0)-I(1)	mars 94
dfmsa	I(1)	janv. 94	I(0)-I(1)	I(0)	juin 95	I(0)-I(1)	I(1)-I(1)	
srersa	I(1)	déc. 93	I(0)-I(1)	I(1)	mai 95	I(0)-I(1)	I(1)-I(1)	
dfersa	I(1)	janv. 95	I(0)-I(1)	I(0)	juin 95	I(0)-I(1)	I(0)-I(1)	fév. 98

Conclusion des tests sur les séries économiques franciliennes								
	Test A	Date	Tests RU	Test C	Date	Tests RU	Nos tests	Date
carbidf	I(1)	juil. 97	I(1)-I(0)	I(0)	sept. 97	I(1)-I(0)	I(1)-I(0)	oct. 97
circ75	I(1)	oct. 99	I(1)-I(1)	I(0)	juil. 98	I(1)-I(1)	I(1)-I(1)	
speed75	I(1)	nov. 91	I(0)-I(0)	I(1)	fév. 92	I(1)-I(0)	I(0)-I(1)	fév. 96
circagar	I(0)	sept. 91	I(0)-I(0)	I(0)	juin 93	I(1)-I(1)	I(1)-I(1)	
defm75	I(0)	dec. 00	I(0)-I(1)	I(1)	nov. 99	I(0)-I(1)	I(0)-I(1)	avr. 99
defmidf	I(0)	mars 99	I(1)-I(0)	I(1)	mars 96	I(1)-I(1)	I(0)-I(1)	déc. 91
gdmagrp	I(1)	août 97	I(1)-I(1)	I(1)	nov. 94	I(1)-I(1)	I(0)-I(1)	avr. 92
totgl	I(0)	juin 99	I(0)-I(0)	I(0)	nov. 95	I(1)-I(0)	I(0)-I(0)	
txhot75	I(0)	sept. 97	I(0)-I(0)	I(0)	août 96	I(0)-I(0)	I(0)-I(0)	
txhotidf	I(1)	juil. 92	I(0)-I(1)	I(1)	juil. 92	I(0)-I(1)	I(0)-I(1)	déc. 97
compet	I(1)	mars 99	I(1)-I(0)	I(1)	mars 99	I(1)-I(0)	I(1)-I(1)	

L'application des tests de Perron aux séries de trafic de la RATP indique que l'hypothèse de racine unitaire autour d'une tendance avec rupture doit être rejetée, à l'exception du trafic

7.3. LA PERTINENCE DE NOS TESTS DE MODIFICATION DU DEGRÉ D'INTÉGRATION PAR RAPPORT AUX TESTS DE PERRON

du samedi sur le RER. Néanmoins, l'application des tests de racine unitaire à chacune de ses séries montre que dans la plupart des cas, si la présence de racine unitaire sur l'ensemble de la période doit être rejetée comme le font les tests de Perron, la série ne semble pas stationnaire sur l'ensemble de la période, mais stationnaire puis intégrée pour la partition considérée.

Il en est de même pour les séries économiques décrivant l'environnement économique en Ile-de-France, même si les conclusions des tests standards de racine unitaire sont parfois plus sujets à caution.

Ces contradictions avec les test de Perron proviennent en grande partie du fait que ceux-ci ont pour hypothèse nulle l'existence d'une racine unitaire sur l'ensemble de la période et pour hypothèse alternative une stationnarité pour la période entière. Ils ne prennent pas en compte la possibilité d'une autre alternative, à savoir le changement du degré d'intégration au cours de la période.

Conclusion à la deuxième partie

Dans cette deuxième partie, un inventaire des tests de changement de structure a été établi lors du premier chapitre (chapitre 5). Cette synthèse des tests existants a montré l'absence de tests de modification du degré d'intégration ayant pour hypothèse nulle la présence de racine unitaire sur l'ensemble de la période.

Dans le deuxième chapitre (chapitre 6), de tels tests ont été construits, et leurs propriétés ont été établies (tables de valeurs critiques et calculs de puissance). En particulier, le test ayant pour hypothèse alternative "la série est intégrée puis stationnaire" détecte particulièrement bien cette hypothèse, et estime la date de rupture avec une grande précision. Une procédure de test a été élaborée : suivant les conclusions du test d'intégration standard puis, éventuellement, des tests de modification du degré d'intégration, l'estimation auto-régressive sera effectuée sur tout ou partie de la période, sur la variable en niveau ou bien en différence.

Ces tests ont été appliqués dans le troisième chapitre (chapitre 7) aux séries de trafic et d'environnement économique dans la région Ile-de-France. Pour chacune de ces séries, détaillées précédemment dans la première partie, les conclusions des tests de modification du degré d'intégration ont été comparées aux conclusions des tests de Perron. Cette comparaison a montré la pertinence de nos tests lorsque l'on s'intéresse à la modification du degré de stationnarité.

Les tests de Perron détectent les modifications de la pente et/ou du niveau de la série, et testent le degré de stationnarité de la série purgée de cette rupture. La modification du degré d'intégration n'est pas envisagée dans ces tests. Les ruptures mises en lumière par ces tests correspondent à des événements majeurs souvent parfaitement identifiés. Par exemple, pour les séries de trafic sur les sous-réseaux de la RATP, les principales ruptures correspondent

7.3. LA PERTINENCE DE NOS TESTS DE MODIFICATION DU DEGRÉ D'INTÉGRATION PAR RAPPORT AUX TESTS DE PERRON

aux grandes grèves de l'hiver 1995, aux attentats sur le réseau ferré, à la mise en place de l'abonnement Imagine R.

Les tests de modification du degré d'intégration mettent en avant des dates de rupture plus difficilement interprétables. La modification du degré d'intégration provient d'éléments plus diffus et parfois moins en lien avec la seule activité de transporteur public. Par exemple, la modification du degré d'intégration de la série de trafic sur le RER un jour ouvrable est en partie imputable à la récession économique de 1993, à l'augmentation du chômage, ... En bref, des éléments parfois exogènes expliquent les modifications des comportements des utilisateurs.

Cette constatation explique pourquoi une modélisation en fonction de facteurs exogènes, telles que les séries relatives à l'environnement économique francilien, est indispensable. En outre, les modélisations auto-régressives, i.e. en fonction des valeurs passées de la variable, ont pour défaut majeur de ne pas permettre d'anticiper les retournements.

Ces retournements peuvent en partie être appréhendés par les modifications de l'environnement économique et conduisent à construire des modèles de trafic en fonction de variables économiques. Ce travail sera réalisé dans la partie suivante.

Troisième partie

Test de modification du degré de cointégration au cours de la période

Introduction à la troisième partie

Dans la partie précédente, nous avons étudié les tests de modification du degré d'intégration (la série est intégrée puis stationnaire ou inversement). L'application de ces tests aux séries de trafic sur les sous-réseaux de la RATP a montré l'existence d'une rupture du processus de mémoire des séries, rupture à mettre en lien avec les modifications de l'environnement économique en Ile-de-France. Une modélisation des trafics en fonction de variables économiques s'avère nécessaire, pour mieux expliquer la rupture des séries, mais aussi, dans un cadre plus général, pour en expliquer les retournements.

Le fait que la majorité des séries soient intégrées sur l'ensemble de la période amène à considérer l'analyse de la cointégration, i.e. l'existence d'un équilibre de long terme et d'une dynamique de court terme.

La présence de rupture ne peut être écartée de cette modélisation et nous conduit, dans un premier chapitre (chapitre 8), à dresser une synthèse des travaux existants sur les tests de cointégration avec rupture.

Dans le deuxième chapitre (chapitre 9), nous montrerons qu'une autre forme de rupture peut être prise en compte par des tests : la modification du degré de cointégration au cours de la période considérée. Autrement dit, en appliquant les tests de modification du degré d'intégration aux résidus d'une relation non cointégrée sur l'ensemble de la période, il est possible de montrer qu'un équilibre de long terme est valide sur une sous-période. Nous présenterons ces tests de modification du degré de cointégration dans le deuxième chapitre de cette partie : leur construction, leurs valeurs critiques, ainsi que leurs propriétés en termes de puissance.

Enfin, nos tests de modification du degré de cointégration seront appliqués aux équilibres de long terme décrivant les trafics des sous-réseaux de la RATP. Equilibres de long terme et dynamiques de court terme seront présentés dans le chapitre 10.

Chapitre 8

Tests de changement de structure de la relation de cointégration : synthèse

Formulé en 1987 par Engle et Granger [50], le concept de cointégration traduit le fait qu'à long terme une combinaison linéaire particulière de variables non stationnaires peut être stationnaire. Cette combinaison linéaire, relation de long terme de l'économie, est supposée invariante dans le temps. Les tests de cointégration usuels dans le cas d'une relation de cointégration sont rappelés dans une première section.

La présence d'une rupture de la relation de long terme peut conduire à des conclusions erronées lors de l'application des tests de rupture standards : l'hypothèse de cointégration est fréquemment rejetée en présence de cointégration avec rupture. Les propriétés des tests standards de cointégration lorsque les variables sont soumises à rupture seront abordées dans la deuxième section.

Plus récemment, la possibilité de changement de structure au sein de la relation de cointégration a été étudiée. De façon générale, la non prise en compte d'un changement de structure au sein d'une relation affecte l'inférence et engendre des prévisions de qualité médiocre. Ce risque de mauvaise spécification est d'autant plus important dans le cas d'une relation de cointégration :

8.1. LES TESTS DE COINTÉGRATION STANDARDS POUR UNE RELATION DE COINTÉGRATION

l'analyse de la cointégration s'intéresse aux équilibres de long terme, et la probabilité d'occurrence d'une rupture augmente avec la taille d'échantillon considérée. Les tests de cointégration avec rupture de la relation de cointégration seront abordés dans la troisième section.

Enfin, dans la quatrième et dernière section de ce chapitre, nous montrerons en quoi l'ensemble de ces tests existants ne répond pas à notre problématique. Leur vocation est de tester qu'un équilibre de long terme existe sur une période entière, alors que nous pensons que cet équilibre peut se réaliser sur une partie seulement de l'échantillon.

8.1 Les tests de cointégration standards pour une relation de cointégration

8.1.1 Les tests de cointégration standards ayant pour hypothèse nulle l'absence de relation de cointégration

Dans le cas d'une unique relation de cointégration, deux grands types de test de cointégration peuvent être distingués :

1. les tests de cointégration réalisés sur les résidus de la relation de cointégration,
2. les tests de cointégration s'appuyant sur un modèle à correction d'erreur.

8.1.1.1 Les tests de cointégration sur les résidus

Les premiers tests de cointégration s'appuient sur les résidus estimés d'une relation de cointégration et ont été formulés par Engle et Granger [50]. Si y_t et x_t sont des variables $I(1)$ et qu'il existe un vecteur β tel que $y_t - \beta x_t$ est stationnaire, alors $\begin{pmatrix} 1 \\ -\beta \end{pmatrix}$ est un "vecteur de cointégration", et $y_t - \beta x_t$ est appelée "relation de cointégration"¹.

Les tests de cointégration sur les résidus considèrent l'équation

$$y_t = \beta x_t + e_t$$

¹Cette relation de cointégration peut s'écrire de différentes formes, suivant la normalisation utilisée.

Si e_t possède une racine unitaire, $y_t - \beta x_t$ n'est pas une relation de cointégration. Tester l'existence d'une racine unitaire dans le résidu correspond à tester l'absence de cointégration. La significativité du paramètre ρ est examinée dans l'équation

$$e_t = \rho e_{t-1} + \varepsilon_t$$

En pratique, le vecteur cointégrant n'est pas connu, mais estimé, et les tests de cointégration sont appliqués sur les résidus estimés \hat{e}_t (et non e_t), conduisant à l'utilisation de valeurs critiques différentes par rapport aux tests d'intégration. Ces valeurs critiques, tabulées par Phillips et Ouliaris [141] pour les tests ADF, Z_t et Z_a , dépendent :

- du nombre de régresseurs, i.e. du nombre de variables composant x_t
- de l'inclusion éventuelle d'une tendance et/ou constante dans la relation de cointégration.

Les hypothèses de test peuvent être résumées sous la forme :

Test de Phillips et Ouliaris

$$H_0 : \{\hat{e}_t \sim I(1)\}, \text{ i.e. absence de cointégration}$$

$$H_a : \{\hat{e}_t \sim I(0)\}, \text{ i.e. cointégration}$$

Selon Kremers, Ericsson et Dolado [100], Zivot [167] ou Banerjee [21], les tests de cointégration sur les résidus de la relation de cointégration manquent de puissance, alors que ce sont les tests de cointégration les plus utilisés. Ce manque de puissance serait imputable à l'ignorance de la dynamique d'ajustement de la relation de long terme au profit de la dynamique des erreurs, ce qui revient à poser l'existence d'un facteur commun². Ce problème est évité lorsque l'on uti-

²Dans sa forme la plus élémentaire, selon Maddala et Kim [119], ce problème se trouve lorsque l'on considère les deux modèles suivants :

$$\text{Modèle 1 : } y_t = \beta x_t + u_t, u_t = \rho u_{t-1} + e_t$$

$$\text{Modèle 2 : } y_t = \alpha_1 y_{t-1} + \alpha_2 x_t + \alpha_3 x_{t-1} + e_t$$

Le premier modèle n'a pas de dynamique d'équation mais une dynamique d'erreur, à l'inverse du deuxième modèle. Or, le modèle 1 peut se réécrire sous la forme

$$y_t = \beta x_t + \rho(y_{t-1} - \beta x_{t-1}) + e_t = \rho y_{t-1} + \beta x_t - \rho \beta x_{t-1} + e_t$$

8.1. LES TESTS DE COINTÉGRATION STANDARDS POUR UNE RELATION DE COINTÉGRATION

lise les tests de cointégration sur les modèles à correction d'erreur, abordés dans le paragraphe suivant.

8.1.1.2 Les tests de cointégration sur les modèles à correction d'erreur (modèles "ECM")

Les tests de cointégration utilisant la modélisation à correction d'erreur s'appuient sur le théorème de représentation de Granger, selon lequel toute relation de cointégration peut être écrite sous une représentation à correction d'erreur. Kremers, Ericsson et Dolado [100] considèrent pour processus générateur de données un modèle à correction d'erreur conditionnel

$$\Delta y_t = a\Delta x_t + b(y_{t-1} - \beta x_{t-1}) + e_t$$

et un processus marginal à racine unitaire

$$x_t = x_{t-1} + u_t$$

tel que e_t et u_t sont deux bruits blancs orthogonaux.

Test de cointégration de Kremers, Ericsson et Dolado

H_0 : $\{b = 0\}$, i.e. absence de cointégration

H_a : $\{b \neq 0\}$, i.e. cointégration, β connu

Zivot [167] considère un cadre plus général, dans lequel le vecteur de cointégration β est inconnu, à l'aide d'un test en deux étapes :

- le vecteur β est estimé par les moindres carrés ordinaires dans la régression statique

$$y_t = \beta x_t + e_t$$

Ces deux modèles sont identiques si $\alpha_1 = \rho$, $\alpha_2 = \beta$, et $\alpha_3 = -\rho\beta$. Cette dernière condition correspond à l'existence d'un facteur commun et se réécrit sous la forme $\alpha_1\alpha_2 + \alpha_3 = 0$. Si cette restriction n'est pas valide, le premier modèle est invalide lui aussi, et la dynamique observée des erreurs est fallacieuse car elle ne résulte que de l'oubli de la dynamique de l'équation.

– La significativité de b est ensuite testée à l'aide d'une t -statistique pour la régression

$$\Delta y_t = a\Delta x_t + b(y_{t-1} - \hat{\beta}x_{t-1}) + e_t$$

Les hypothèses de ce test peuvent être résumées de la façon suivante :

Test de Zivot n°1 (test en deux étapes)

H_0 : $\{b = 0\}$, i.e. absence de cointégration

H_a : $\{b \neq 0\}$, i.e. cointégration, β inconnu

Il construit aussi un test en une unique étape, comparable au test construit par Boswijk [29, 30] en modifiant l'écriture du modèle à correction d'erreur :

$$\begin{aligned} \Delta y_t &= a\Delta x_t + b(y_{t-1} - \beta x_{t-1}) + e_t \\ &= a\Delta x_t + b \begin{pmatrix} 1 \\ -\beta \end{pmatrix} \begin{pmatrix} y_{t-1} & x_{t-1} \end{pmatrix} + e_t \\ &= a\Delta x_t + \Pi z_{t-1} + e_t \end{aligned}$$

L'hypothèse précédente d'absence de cointégration $\{b = 0\}$ implique $\Pi = 0$, hypothèse que l'on peut tester à l'aide d'un test de Wald.

Test de Zivot n°2 (test en une étape)

H_0 : $\{\Pi = 0\}$, i.e. absence de cointégration

H_a : $\{\Pi \neq 0\}$, i.e. cointégration, β inconnu

Les tests s'appuyant sur une modélisation à correction d'erreur ont une puissance plus importante que les tests de type ADF appliqués aux résidus de l'estimation à la Engle et Granger. Analytiquement, cette puissance plus importante s'explique de la façon suivante lorsque le vec-

8.1. LES TESTS DE COINTÉGRATION STANDARDS POUR UNE RELATION DE COINTÉGRATION

teur cointégrant est supposé connu : si le modèle ECM est réécrit sous la forme

$$\begin{aligned}\Delta y_t &= a\Delta x_t + b(y_{t-1} - \beta x_{t-1}) + e_t \\ \Delta y_t - \beta\Delta x_t &= a\Delta x_t - \lambda\Delta x_t + b(y_{t-1} - \beta x_{t-1}) + e_t \\ \Delta(y_t - \beta x_t) &= b(y_{t-1} - \beta x_{t-1}) + (a - \lambda)\Delta x_t + e_t\end{aligned}$$

et le modèle à la Engle et Granger sous la forme

$$\Delta(y_t - \beta x_t) = b(y_{t-1} - \beta x_{t-1}) + \varepsilon_t$$

on s'aperçoit que l'estimation à la Engle et Granger ignore l'information potentielle contenue dans Δx_t et impose l'existence d'un facteur commun entre y_t et x_t . Si $a = \lambda$, les deux types de tests ont une puissance identique. Les valeurs critiques correctes pour un test ECM dépendent du rapport q de l'écart-type de $(a - \beta)\Delta x_t$ par rapport à celui de e_t ainsi que de la taille de l'échantillon. Mais, en pratique, q n'est généralement pas connu. La présence de ce paramètre de nuisance dans la distribution de la statistique explique le fait que ce test ait été peu tabulé et, par conséquent, peu employé empiriquement malgré les gains de puissance.

8.1.2 Tests de cointégration ayant pour hypothèse nulle l'existence de cointégration

Un autre moyen de contourner le manque de puissance des tests de cointégration appliqués aux résidus de la relation de cointégration est de considérer des tests de cointégration ayant pour hypothèse nulle l'existence de cointégration. Trois grands types de tests ayant pour hypothèse nulle l'existence de cointégration peuvent être évoqués :

1. Le test de Leybourne et McCabe [111] s'appuie sur une modélisation à composante inobservable. La relation de cointégration demeure définie par

$$y_t = \beta x_t + e_t$$

et l'erreur e_t est telle que

$$e_t = \gamma_t + u_t, \quad u_t \sim i.i.N(0, \sigma_u^2)$$

$$\gamma_t = \gamma_{t-1} + \eta_t, \quad \eta_t \sim i.i.N(0, \sigma_\eta^2)$$

Les hypothèses de test s'écrivent sous la forme

Test de Leybourne et McCabe

$$H_0 : \{\sigma_\eta^2 = 0\}, \text{ i.e. cointégration}$$

$$H_a : \{\sigma_\eta^2 > 0\}, \text{ i.e. absence de cointégration}$$

2. Le test de Park et Choi [132] s'appuie sur la significativité de régresseurs superflus. La relation de cointégration avec régresseurs superflus peut s'écrire sous la forme

$$y_t = \beta x_t + \alpha_1 s_{1t} + \alpha_2 s_{2t} + e_t$$

où s_{1t} est un vecteur de q fonctions déterministes non-stationnaires possédant un ordre supérieur aux variables de la relation de cointégration³, et s_{2t} est un vecteur constitué de p variables intégrées à l'ordre 1. Soit SCR_1 la somme des carrés résiduels issue de l'estimation de la relation de cointégration⁴ sans variable superflue, et SCR_2 la somme des carrés résiduels sur la régression comprenant les variables s_{1t} et s_{2t} . Les hypothèses de test sont les suivantes :

Test de Park et Choi

$$H_0 : \{SCR_1 = SCR_2\}, \text{ i.e. cointégration}$$

$$H_a : \{SCR_1 - SCR_2 > 0\}, \text{ i.e. absence de cointégration}$$

³Par exemple, si l'on suppose la présence d'une tendance t dans la relation de cointégration, on pourra prendre les fonctions t^2, t^3, \dots, t^{q+1} dans s_{1t} .

⁴Dans le test de Park [131], la régression est effectuée après une correction non-paramétrique de la corrélation entre les paramètres de nuisance.

8.1. LES TESTS DE COINTÉGRATION STANDARDS POUR UNE RELATION DE COINTÉGRATION

3. Le test de Shin [154] correspond à l'application d'un test de type KPSS sur les résidus de l'éventuelle relation de cointégration. Pour éliminer les problèmes d'endogénéité des régresseurs (corrélacion instantanée), la relation de cointégration est corrigée de façon paramétrique :

$$y_t = \beta x_t + \sum_{j=-k_1}^{k_2} \gamma_j \Delta x_{t-j} + e_t$$

La statistique de test est égale à

$$C = T^{-2} \sum \frac{S_t^2}{\hat{\sigma}_e^2}$$

avec $S_t^2 = \sum_{i=1}^t \hat{e}_i^2$ et $\hat{\sigma}_e^2$ l'estimation semi-paramétrique convergente de la matrice de long-terme de e_t . Les hypothèses de tests se résument de la façon suivante :

Test de Shin

H_0 : $\{C = 0\}$, i.e. cointégration

H_a : $\{C > 0\}$, i.e. absence de cointégration

Le test de Harris et Inder [82] corrige l'estimation de la relation de cointégration de façon non-paramétrique (méthode des "moindres carrés totalement modifiés" voir page 317), mais utilise la même statistique du maximum de vraisemblance que le test de Shin.

Les tests de cointégration évoqués dans ce paragraphe ont été construits pour une relation de cointégration invariante au cours du temps. Le test ADF appliqué au résidu de l'éventuelle relation de cointégration (noté test t_{CADF}) et le test sur le modèle à correction d'erreur (test t_{ECM}) sont les plus couramment utilisés dans les analyses empiriques. Examinons leurs propriétés en cas de présence d'un changement structurel.

8.2 Les effets d'un changement structurel sur les tests de cointégration

La présence d'une rupture peut conduire à des conclusions erronées lors de l'application des tests de cointégration standards. Dans cette section, nous évaluerons l'impact de deux types de rupture :

- la présence d'une rupture au sein de la relation de cointégration (premier paragraphe).
- la présence d'une rupture des séries univariées intervenant dans la relation de cointégration (deuxième paragraphe).

8.2.1 Niveau réel et puissance des tests de cointégration en présence de changement de structure au sein de la relation de cointégration

Gregory, Nason et Watt [67] analysent la sensibilité du test augmenté de Dickey-Fuller de cointégration à la présence d'une rupture structurelle du vecteur de cointégration. Ils considèrent un modèle linéaire quadratique dont la cible de long terme s'écrit

$$y_t^* = \beta x_t + e_t$$

La valeur réelle choisie, dans le cas de deux variables univariées y_t et x_t , s'écrit sous la forme d'une relation de cointégration

$$\begin{aligned} y_t &= \beta x_t + \eta_t \\ \Delta x_t &= \varepsilon_t \end{aligned}$$

où

$$\eta_t = -\beta\lambda \cdot (1 - \lambda L)^{-1} \varepsilon_t + \delta\lambda \cdot (1 - \lambda L)^{-1} e_t$$

avec $\lambda < 1$ la racine stable de l'équation d'Euler issue des conditions de premier ordre.

8.2. LES EFFETS D'UN CHANGEMENT STRUCTUREL SUR LES TESTS DE COINTÉGRATION

Dans le cas d'un changement de structure, la relation de cointégration est re-spécifiée :

$$y_t = \beta_t \cdot x_t + \eta_t \text{ avec } \beta_t = \begin{cases} \beta_1, & t \leq [\pi T] \\ \beta_2, & t > [\pi T] \end{cases}$$

2 500 simulations de Monte Carlo ont été réalisées sur des échantillons de taille $T = 100$, 200 ou 500 au seuil de 5%, pour les valeurs critiques asymptotiques.

Le nombre de retards dans le test ADF est choisi par une procédure $t - sig$ (diminution du nombre de retards jusqu'à ce que le dernier retard soit significatif d'après un test de Student).

Les résultats sont reproduits pour

$$\begin{aligned} T &= 100 \\ \beta_1 &= 1, \beta_2 = 0.5 \\ \varepsilon_t &\sim i.i.N(0, 1) \end{aligned}$$

Puissance des tests ADF	$\Delta x_t = \varepsilon_t$			$\Delta x_t = 0.8\Delta x_t + \varepsilon_t$
λ	0.1	0.7	0.9	0.7
$\pi = 1$	96 %	80 %	18 %	44%
$\pi = 0.25$	68 %	73%	24 %	30 %
$\pi = 0.50$	39 %	51 %	21 %	7 %
$\pi = 0.75$	31 %	44 %	17 %	4 %

Source : Gregory, Nason et Watt [67]

La fréquence de rejet du test t_{ADF} (ayant pour hypothèse nulle l'absence de relation de cointégration) décroît en cas de présence d'une rupture du vecteur cointégrant. Cette fréquence est en outre fortement affectée par la présence de retards dans le résidu, ainsi que par les coûts d'ajustements λ à l'objectif. La faible fréquence de rejection lorsqu'il y a rupture du vecteur cointégrant ne correspond pas à un véritable manque de puissance du test dans la mesure où il signale bien qu'il n'existe pas de relation de cointégration avec des paramètres stables au cours

du temps.

8.2.2 Niveau réel et puissance des tests de cointégration dans le cas de séries soumises à un changement de structure

Au lieu de considérer des séries intégrées liées par une relation soumise à un changement structurel, nous nous intéressons désormais à une relation de cointégration sans rupture, mais faisant intervenir des variables ayant subi une rupture et pouvant passer à tort pour des variables intégrées. Deux cas sont envisagés :

- les séries apparaissant dans la relation de cointégration sont stationnaires avec rupture,
- les séries sont stationnaires puis intégrées ou intégrées puis stationnaires.

8.2.2.1 Prise en considération de séries stationnaires avec rupture

Selon Campos, Ericsson et Hendry [34], la présence d'un changement de structure sur les séries individuelles a peu d'impact sur le niveau des tests de cointégration, mais l'existence d'une telle rupture affecte la puissance de ces tests lorsque le processus générateur des séries ne suppose pas de facteur commun. Campos, Ericsson et Hendry posent comme processus générateur de données un modèle à correction d'erreur

$$\Delta y_t = a\Delta x_t + b(y_{t-1} - \lambda x_{t-1}) + e_t$$

et un processus marginal défini par

$$x_t = \rho x_{t-1} + \delta D_t + u_t$$

tel que e_t et u_t sont deux bruits blancs orthogonaux et

$$\begin{aligned} D_t &= 1 \text{ si } T_0 < t \leq T_1 \\ &= 0 \text{ sinon} \end{aligned}$$

8.2. LES EFFETS D'UN CHANGEMENT STRUCTUREL SUR LES TESTS DE COINTÉGRATION

Lorsque x_t est un processus stationnaire avec rupture (i.e. $|\rho| < 1, \delta \neq 0$), les tests standards de racine unitaire ont tendance à conclure à l'existence d'une racine unitaire (i.e. $\rho = 1, \delta = 0$) comme dans le test de Kremers, Ericsson et Dolado [100]). Campos, Ericsson et Hendry comparent les propriétés des tests de cointégration dans chacun de ces deux cas, à l'aide de simulations de Monte Carlo.

Les simulations de Monte Carlo sont réalisées sur une taille d'estimation $T_e \leq T$.

- Si $T_e < T_0$, il n'y a pas de rupture sur l'échantillon servant à l'estimation,
- Si $T_0 \leq T_e < T_1$, il existe une rupture,
- Si $T_1 \leq T_e \leq T$, il existe deux ruptures qui peuvent être appréhendées comme une rupture temporaire.

Le seuil des tests est égal à 5%, pour des valeurs critiques calculées selon la méthode des surfaces de réponse de MacKinnon [121], identiques quel que soit le test utilisé. Les 10 000 simulations ont été effectuées en supposant que

$$\begin{aligned}\lambda &= 1 \\ \sigma_e^2 &= 1 \\ a &= 1 \text{ (facteur commun) ou } 0 \text{ (pas de facteur commun)} \\ b &= 0 \text{ (pas de cointégration) ou } -0.1 \text{ (cointégration)} \\ \delta &= 0 \text{ (pas de rupture) ou } 1 \text{ (rupture) voire } 3 \text{ (forte rupture)} \\ \rho &= 1 \text{ (intégration) ou } 0.8 \text{ (stationnarité)} \\ T_e &= 10, 11, \dots, T-1, T \\ T &= 100, T_0 = 25, T_1 = 75\end{aligned}$$

Leurs principaux résultats sont les suivants :

	Facteur commun	Type de variable	t_{DF}	t_{ECM}
Niveau	Oui	Intégrée sans rupture	5.0	3.5
	Oui	Stationnaire avec rupture	4.1	3.2
	Non	Intégrée sans rupture	4.8	3.3
	Non	Stationnaire avec rupture	4.3	3.1
	Non	Stationnaire avec forte rupture	4.5	3.1
Puissance	Oui	Intégrée sans rupture	22.8	[3.9, 19.2]
	Oui	Stationnaire avec rupture	22.2	[3.2, 18.3]
	Non	Intégrée sans rupture	14.6	[3.4, 40.4]
	Non	Stationnaire avec rupture	15.5	[3.2, 34.8]
	Non	Stationnaire avec forte rupture	2.9	[2.8, 98.1]

Unité : %

Le niveau et la puissance de t_{DF} sont estimées au point $T_e = 100$ uniquement. Les valeurs entre crochets de t_{ECM} indiquent l'intervalle sur T_e , les valeurs simples de t_{ECM} correspondent à la moyenne sur l'intervalle.

Source : Campos, Ericsson et Hendry (1996) [34] p. 213

- Quelles que soient la spécification de la variable et l'existence effective d'un facteur commun, les tests de Dickey-Fuller dans les modèles à la Engle et Granger possèdent les bonnes propriétés en terme de niveau. Le test sur le modèle à correction d'erreur a un niveau réel inférieur au niveau nominal, du fait de la prise en compte de valeurs critiques tabulées par la méthode de McKinnon.
- En cas de facteur commun, les tests t_{DF} ont une puissance plus importante que les tests t_{ECM} . Dès que la condition de facteur commun n'est plus valide, la puissance des tests t_{ECM} devient plus grande que celle des tests t_{DF} . Pour une rupture importante, la puissance du test t_{DF} peut devenir inférieure au seuil du test !

8.2. LES EFFETS D'UN CHANGEMENT STRUCTUREL SUR LES TESTS DE COINTÉGRATION

Prise en considération de séries ayant subi une modification du degré d'intégration Par souci de comparaison, nous avons analysé les propriétés des tests ADF sur des séries intégrées puis stationnaires ou stationnaires puis intégrées, telles qu'envisagées dans la partie précédente. Pour ce faire, nous avons dans un premier temps évalué les propriétés en terme de niveau réel des tests ADF pour des séries subissant une modification du degré d'intégration à la date k . 10 000 simulations de Monte Carlo sur des échantillons de taille 100 ont été réalisées en supposant que le paramètre auto-régressif était égal à 1 sur la partie intégrée et 0.8 sur la partie stationnaire. Le niveau nominal du test a été posé à 5 %.

<i>Niveau réel des tests ADF quand la série a subi une modification du degré d'intégration</i>									
Type du test ADF	Sans constante			Avec constante			Avec tendance		
Date de modification du degré d'intégration	25	50	75	25	50	75	25	50	75
Séries stationnaires puis intégrées : $\rho_a = 0.8, \rho_b = 1$									
$c_a = 0, c_b = 0$	10.37	25.40	54.52	3.56	9.07	27.48	5.56	10.82	22.36
$c_a = 0.5, c_b = 0$	6.76	8.75	11.80	4.25	10.67	29.14	6.53	11.55	25.08
$c_a = 0, c_b = 0.5$	0.00	0.04	6.65	0.00	0.05	1.91	3.10	0.12	2.23
Séries intégrées puis stationnaires : $\rho_a = 1, \rho_b = 0.8$									
$c_a = 0, c_b = 0$	71.88	35.38	16.03	33.97	11.56	4.64	26.94	10.76	3.45
$c_a = 0.5, c_b = 0$	14.82	0.26	0.00	3.90	0.04	0.00	4.23	0.12	0.01
$c_a = 0, c_b = 0.5$	16.39	11.86	9.12	28.15	10.71	3.93	22.81	9.61	3.25

Unité : pourcentage de rejection de l'hypothèse nulle d'existence de racine unitaire

c_a : constante sur la première sous-période, c_b : constante sur la seconde sous-période

Ainsi, en cas de rupture du degré d'intégration à la mi-échantillon, le test ADF de changement de structure avec constante estime le niveau réel du test à environ 10% lorsque la partie intégrée ne comporte pas de dérive, et conclut presque toujours à l'existence de racine unitaire en cas de dérive. En présence d'une dérive sur la partie intégrée, quelle que soit la date de rupture, le test conclut à l'existence d'une racine unitaire dans presque tous les cas. Le niveau réel du test est d'autant plus faible que la sous-période intégrée est longue.

Cette analyse nous motive à examiner les propriétés des tests de cointégration lorsque les variables prises en compte présentent une modification du degré d'intégration. Le processus générateur de données est un modèle à correction d'erreur

$$\Delta y_t = a\Delta x_t + b(y_{t-1} - \lambda x_{t-1}) + e_t$$

avec le processus marginal suivant

$$x_t = \begin{cases} \rho_a x_{t-1} + u_t, & t = 1, \dots, k \\ \rho_b x_{t-1} + u_t, & t = k + 1, \dots, T \end{cases}$$

Les 10 000 simulations sur un échantillon de taille 100 sont réalisées en prenant ⁵

$$\begin{aligned} \lambda &= 1 \\ \sigma_e^2 &= \sigma_u^2 = 1 \\ a &= 1 \text{ (facteur commun) ou } 0 \text{ (pas de facteur commun)} \\ b &= 0 \text{ (pas de cointégration) ou } -0.1 \text{ (cointégration)} \\ \{\rho_a, \rho_b\} &= \{0.8, 1\} \text{ (stationnaire puis intégré) ou } \{1, 0.8\} \text{ (intégré puis stationnaire)} \\ k &= 25, 50 \text{ ou } 75 \end{aligned}$$

⁵Nous avons tout d'abord simulé la variable x_t pour 130 observations, ainsi que la perturbation e_t . Nous avons ensuite calculé y_t en utilisant la formule récursive de type "autoregressive distributed lags "

$$y_{t+1} = y_t + a\Delta x_{t+1} + b(y_t - \lambda x_t) + e_{t+1}$$

pour une valeur initiale $y_0 = 0$. Les 30 premières observations de y_t et x_t ont été supprimées de manière à disposer d'échantillons de taille 100, peu influencés par les valeurs initiales.

8.2. LES EFFETS D'UN CHANGEMENT STRUCTUREL SUR LES TESTS DE COINTÉGRATION

<i>Niveau réel et puissance des tests t_{ADF} et t_{ECM} à 5% quand les séries ont subi une modification de leur degré d'intégration au cours de la période</i>									
VC à 5% = -3.40397	Niveau réel des tests			Puissance des tests					
	$b = 0$			$b = -0.1$			$b = -0.2$		
Date de modification du degré d'intégration	25	50	75	25	50	75	25	50	75
Séries stationnaires puis intégrées : $\rho_a = 0.8, \rho_b = 1$									
$a = 1$									
t_{CADF}	5.88	5.48	4.26	19.99	19.83	18.94	59.94	59.38	59.28
t_{ECM}	4.23	4.07	2.67	18.49	18.10	15.47	65.01	64.46	63.24
$a = 0$									
t_{CADF}	5.02	5.50	4.71	13.96	15.97	15.58	48.42	52.81	53.73
t_{ECM}	3.85	3.58	3.05	31.77	26.34	20.65	91.75	88.44	82.96
Séries intégrées puis stationnaires : $\rho_a = 1, \rho_b = 0.8$									
$a = 1$									
t_{CADF}	4.64	5.26	5.10	18.45	19.68	20.08	58.17	59.54	58.79
t_{ECM}	3.54	4.26	4.19	17.41	17.98	18.29	63.56	65.21	64.95
$a = 0$									
t_{CADF}	4.79	5.55	5.92	12.74	11.71	10.67	48.04	44.73	41.47
t_{ECM}	3.60	4.21	4.61	30.90	35.24	35.21	91.02	93.11	94.01

Unité : pourcentage de rejection de l'hypothèse nulle d'absence de cointégration

Les principaux résultats de ces simulations sont les suivants :

- Les résultats en terme de niveau réel et de puissance des tests t_{CADF} et t_{ECM} sont assez similaires quelle que soit la modification du degré d'intégration des séries intervenant dans la relation de cointégration (stationnaires puis intégrées ou intégrées puis stationnaires).
- Dans le cas de séries stationnaires puis intégrées ou intégrées puis stationnaires, les tests t_{CADF} et t_{ECM} possèdent de bonnes propriétés en terme de niveau : lorsque le niveau réel

est posé à 5%, le niveau réel ne s'en éloigne pas à plus d'un ou deux points. L'écart plus important du niveau réel du test t_{ECM} par rapport au niveau nominal s'explique en partie par le fait que la valeur critique utilisée est tabulée pour le test CADF.

- En cas de facteur commun ($a = 1$), pour une racine assez large ($b = -0.2$), la puissance des tests t_{CADF} est de l'ordre de 60%, soit environ 5 points de moins que les tests t_{ECM} . Néanmoins, les tests t_{CADF} ont une puissance plus importante que les tests t_{ECM} lorsque l'on se rapproche de l'hypothèse nulle d'absence de cointégration ($b = -0.1$).
- En l'absence de facteur commun ($a = 0$), la puissance des tests t_{CADF} diminue alors que la puissance des tests t_{ECM} s'accroît fortement. Dans le cas de séries intégrées puis stationnaires, la puissance des tests t_{ECM} est ainsi deux fois plus grande que celle des tests t_{CADF} .

Lorsque les séries intervenant au sein de la relation de cointégration sont soumises à un changement de structure (séries stationnaires avec rupture ou ayant subi une modification du degré d'intégration au cours de la période considérée), les propriétés des tests usuels de cointégration sont modifiées dans une faible mesure. Il n'en est pas de même lorsque le changement structurel a lieu non pas au sein des séries utilisées, mais dans la relation de cointégration. Les simulations de Monte Carlo effectuées par Gregory, Nason et Watt [67] indiquent que le test augmenté de Dickey-Fuller de cointégration (test "CADF") accepte trop souvent l'hypothèse nulle de non-cointégration quand il existe un changement de structure dans la relation de long terme. Ce constat a amené les économètres à construire des tests de cointégration prenant en compte l'existence éventuelle d'un changement structurel dans la relation de cointégration.

8.3 Les tests de changement structurel dans une relation de cointégration

Parmi les tests de changement de structure au sein d'une relation de cointégration, les apports majeurs sont les suivants :

- les tests de Gregory et Hansen de 1996 [65, 66] de cointégration avec rupture

– les tests de spécification de la relation de cointégration, qui peuvent être appréhendés comme des tests de cointégration :

- le test de Hansen [74, 75] en 1992
- le test de Quintos et Phillips [147] élaboré en 1993,
- ou encore le test de Kuo [101] publié en 1998.

Développons ces différents tests de manière à montrer leur complémentarité.

8.3.1 Les tests de Gregory et Hansen

Gregory et Hansen [65, 66] ont généralisé les tests de rupture aux tests de cointégration s'appuyant sur l'analyse de la stationnarité des résidus \hat{e}_t issus de l'estimation de la relation de long terme.

8.3.1.1 Hypothèses de test

De façon générale, les hypothèses de tests sont les suivantes ⁶ :

Test de Gregory et Hansen

$H_0 : \hat{e}_t \sim I(1)$ i.e. absence de relation de cointégration

$H_a : \hat{e}_t \sim I(0)$ i.e. existence d'une relation de cointégration

avec changement structurel à une date indéterminée $[\tau T]$

Différents types de rupture peuvent se produire au sein de la relation de cointégration ⁷.

1. Rupture sur le niveau :

$$y_t = c_1 + (c_2 - c_1) \cdot \mathbf{1}_{t > [\tau T]} + ax_t + e_t$$

⁶Soit T la taille de l'échantillon, et τ le pourcentage de l'échantillon à laquelle la rupture a lieu. La date de rupture correspond à la partie entière de τT , notée $[\tau T]$. La date de rupture est balayée sur une fenêtre de taille $\Pi = [0.15T, 0.85T]$ afin de se prémunir des effets de bord, conformément aux préconisations d'Andrews [6].

⁷Les modèles 3 et 5 ont été ajoutés par Lenormand [109] : les tests de Gregory et Hansen ne considèrent que les cas 1, 2, 4 et 6.

2. Rupture sur le niveau avec tendance :

$$y_t = c_1 + (c_2 - c_1) \cdot \mathbf{1}_{t > [\tau T]} + bt + ax_t + e_t$$

3. Rupture sur le niveau et la pente de la tendance :

$$y_t = c_1 + (c_2 - c_1) \cdot \mathbf{1}_{t > [\tau T]} + b_1 t + (b_2 - b_1) t + ax_t + e_t$$

4. Rupture sur le niveau et le vecteur cointégrant :

$$y_t = c_1 + (c_2 - c_1) \cdot \mathbf{1}_{t > [\tau T]} + a_1 x_t + (a_2 - a_1) \cdot (x_t)_{t > [\tau T]} + e_t$$

5. Rupture sur le niveau et le vecteur cointégrant, avec tendance :

$$y_t = c_1 + (c_2 - c_1) \cdot \mathbf{1}_{t > [\tau T]} + bt + a_1 x_t + (a_2 - a_1) \cdot (x_t)_{t > [\tau T]} + e_t$$

6. Rupture sur le niveau, la pente de la tendance et le vecteur cointégrant :

$$y_t = c_1 + (c_2 - c_1) \cdot \mathbf{1}_{t > [\tau T]} + b_1 t + (b_2 - b_1) t + a_1 x_t + (a_2 - a_1) \cdot (x_t)_{t > [\tau T]} + e_t$$

8.3.1.2 Statistiques de test

Pour chaque date de rupture possible $[\tau]$,

- La relation de cointégration avec rupture est estimée par les moindres carrés ordinaires dans les travaux de Gregory et Hansen, et par les moindres carrés totalement modifiés (méthode décrite page 317) dans la thèse de Lenormand [109].
- La statistique $t_{ADF}(\tau)$ ⁸ des tests de racine unitaire de Dickey et Fuller augmentés ainsi que les statistiques $Z_a(\tau)$ et $Z_t(\tau)$ des tests non paramétriques de Phillips et Ouliaris sont calculées à partir des résidus estimés de cette relation de long terme.

⁸La statistique t_{ADF} est calculée à partir de l'estimation sans constante ni tendance par les moindres carrés ordinaires : $\hat{e}_t = \rho \hat{e}_{t-1} + \varepsilon_t$, en ajoutant éventuellement des variables retardées.

8.3. LES TESTS DE CHANGEMENT STRUCTUREL DANS UNE RELATION DE COINTÉGRATION

Pour chacun de ces trois tests, la date de rupture et la statistique de test finale sont estimées en prenant la plus petite valeur de la statistique obtenue par balayage sur les dates de rupture éventuelles :

$$\begin{aligned}
 t_{ADF}^* &= \inf_{\tau \in [0.15, 0.85]} t_{ADF}(\tau) \\
 Z_a^* &= \inf_{\tau \in [0.15, 0.85]} Z_a(\tau) \\
 Z_t^* &= \inf_{\tau \in [0.15, 0.85]} Z_t(\tau)
 \end{aligned}$$

Si cette statistique finale est inférieure à la valeur critique adéquate pour le niveau de confiance choisi, l'hypothèse nulle d'absence de relation de cointégration est rejetée au profit de l'existence d'une relation de cointégration avec rupture. Les valeurs critiques à 5% obtenues par Lenormand [109] dans le cas d'une estimation de la relation par les moindres carrés totalement modifiés pour un échantillon de taille 120 sont les suivantes :

Valeurs critiques à 5% des tests t_{ADF}^* et Z_a^* de cointégration avec rupture pour les moindres carrés totalement modifiés

Type de rupture →	1	2	3	4	5	6
Nb de variables ↓						
1	-4.696	-5.140	-5.491	-5.036	-5.502	-5.721
2	-5.081	-5.469	-5.811	-5.689	-6.050	-6.289
3	-5.450	-5.807	-6.119	-6.224	-6.594	-6.752
4	-5.760	-6.072	-6.387	-6.724	-7.048	-7.180

Source : Lenormand [109]

Les tests de Gregory et Hansen ne donnent aucune preuve de l'existence d'un changement de structure dans la relation de cointégration, car les différentes hypothèses alternatives comportent un cas particulier : celui où il y a cointégration mais pas de rupture de la relation cointégrante. L'examen de la significativité associée aux variables indicatrices présente des dangers, dans la mesure où les statistiques de test habituelles n'ont généralement pas de distribution asymptotique standard sous l'hypothèse de cointégration, d'autant plus qu'en l'absence de changement

de structure, la date de rupture τ n'est pas identifiable.

La solution est d'étudier l'existence éventuelle d'un changement de régime en utilisant les tests de Hansen [74, 75], de Quintos et Phillips [147] ou encore de Kuo [101], qui s'intéressent à l'invariance dans le temps de la relation de cointégration, en posant comme hypothèse nulle l'existence de cointégration au sens de Engle et Granger. Ces trois tests peuvent être appréhendés comme des tests de spécification de modèles cointégrés au sens de Engle et Granger, tandis que les tests de Gregory et Hansen sont des tests de cointégration.

8.3.2 Les tests de spécification de la relation de cointégration

8.3.2.1 Premier test de spécification de la relation de cointégration : le test de Hansen

Supposons que grâce à un test de Gregory et Hansen, nous ayons montré l'existence d'une relation de cointégration et que l'on s'intéresse désormais à l'invariance dans le temps de cette relation. Nous devons mettre en oeuvre un test de spécification de la relation de cointégration : tests de Hansen, de Quintos et Phillips ou de Kuo.

Hansen s'intéresse à l'instabilité des paramètres dans des régressions contenant des processus intégrés à l'ordre 1 et a construit un test de changement de structure de la relation de cointégration (de type ratio du maximum de vraisemblance), suite à une estimation par les moindres carrés totalement modifiés de Phillips et Hansen [140]. L'instabilité des paramètres dans la relation de cointégration est incorporée en ne prenant pas en compte une relation de cointégration standard

$$y_t = c + a \cdot x_t + e_{1t}$$

8.3. LES TESTS DE CHANGEMENT STRUCTUREL DANS UNE RELATION DE COINTÉGRATION

mais en faisant dépendre les paramètres a et c du temps :

$$\begin{aligned} y_t &= c_t + a_t \cdot x_t + e_{1t} \\ &= \beta_t \cdot X_t + e_{1t} \\ \Delta x_t &= e_{2t} \\ e_{1t} &\sim I(0) \end{aligned}$$

Le processus aléatoire $e_t = \begin{pmatrix} e_{1t} \\ e_{2t} \end{pmatrix}$ a une moyenne nulle et une matrice de variance-covariance instantanée Σ . Soient les matrices suivantes :

$$\begin{aligned} \Omega &= \lim_{T \rightarrow \infty} E \left(\frac{1}{T} S_T S_T' \right) = \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{i=1}^T \sum_{j=1}^T E(e_i e_j') \\ \Lambda &= \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{j=2}^T \sum_{i=1}^t E(e_i e_j') \end{aligned}$$

et leur partition conformément au vecteur e_t :

$$\Omega = \begin{bmatrix} \Omega_{11} & \Omega_{12} \\ \Omega_{21} & \Omega_{22} \end{bmatrix} \text{ et } \Lambda = \begin{bmatrix} \Lambda_{11} & \Lambda_{12} \\ \Lambda_{21} & \Lambda_{22} \end{bmatrix}$$

Ecrivons la matrice de long terme de e_{1t} conditionnellement à e_{2t}

$$\Omega_{1.2} = \Omega_{11} - \Omega_{12} \Omega_{22}^{-1} \Omega_{21}$$

et la matrice de correction

$$\Lambda_{21}^+ = \Lambda_{21} - \Lambda_{22} \Omega_{22}^{-1} \Omega_{21}$$

Ces matrices sont estimées à l'aide d'un estimateur à noyau quadratique, pré-blanchi par un processus vectoriel d'ordre 1, pour une largeur de bande automatique. Cette estimation correspond à la première étape de la **méthode des moindres carrés "totalement modifiés"** de

Phillips et Hansen [140]. Les paramètres de la régression sont estimés dans une deuxième étape : les données sont d'abord corrigées de l'endogénéité

$$\begin{aligned} y_t^+ &= y_t - \hat{\Omega}_{12}\hat{\Omega}_{22}^{-1}\hat{e}_{2t} \\ e_{1t}^+ &= e_{1t} - \hat{\Omega}_{12}\hat{\Omega}_{22}^{-1}\hat{e}_{2t} \end{aligned}$$

puis de la corrélation temporelle, d'où

$$\begin{aligned} \hat{\beta}^+ &= \sum_{t=1}^T \left(X_t' y_t^+ - T \begin{pmatrix} 0 & \hat{\Lambda}_{21}^+ \end{pmatrix} \right) \left(\sum_{t=1}^T X_t X_t' \right)^{-1} \\ \text{et } \hat{e}_{1t}^+ &= y_t^+ - \hat{\beta}^+ X_t \end{aligned}$$

Définissons alors

$$\begin{aligned} \hat{s}_t &= X_t \cdot \hat{e}_{1t}^+ - \begin{pmatrix} 0 \\ \hat{\Lambda}_{21}^+ \end{pmatrix} \\ S_{T,\tau} &= \sum_{i=1}^{[\tau T]} \hat{s}_i \\ M_{T\tau} &= \sum_{i=1}^{[\tau T]} X_i X_i' \\ \text{et } V_{T,\tau} &= M_{T\tau} - M_{T\tau} M_{T\tau}^{-1} M_{T\tau} \end{aligned}$$

Hansen a construit quatre tests distincts. Présentons ces tests dans le cas d'une relation de cointégration sans tendance. En présence d'une tendance, elle aussi soumise à rupture sous les hypothèses alternatives, les statistiques de test demeurent inchangées ; seules les distributions asymptotiques sont modifiées.

1. Un premier test introduit une unique modification des paramètres à une date de rupture supposée connue. Si β_t est soumis à un unique changement de structure à une date $[\tau T]$,

où $\tau \in \Pi \subseteq [0, 1]$,

$$\begin{aligned}\beta_i &= \beta_1 \text{ si } i \leq [\tau T] \\ &= \beta_2 \text{ si } i > [\tau T]\end{aligned}$$

Les hypothèses de test sont les suivantes :

Test de Hansen n° 1

$H0$: $\{\beta'_1 = \beta'_2\}$, i.e. invariance des paramètres de la relation de cointégration

Ha : $\{\beta'_1 \neq \beta'_2\}$, τ connu

La statistique de test s'écrit :

$$\begin{aligned}F_{T,\tau} &= \text{vec}(S_T, \tau)' \left(\hat{\Omega}_{1.2} \otimes V_{T,\tau} \right)^{-1} \text{vec}(S_T, \tau) \\ &= \text{tr} \left\{ S'_{T,\tau} \cdot V_{T,\tau}^{-1} \cdot S_T, \tau \hat{\Omega}_{1.2} \right\}\end{aligned}$$

2. Le deuxième test relâche l'hypothèse selon laquelle la date du changement de structure est connue. Pour chaque $\tau \in \Pi$, la statistique du test 1 est calculée et la statistique finale du test 2 correspond à la plus grande valeur sur l'intervalle Π ⁹. La date du changement de structure correspond à la date où la statistique du test 1 atteint son maximum. Les hypothèses de tests sont les suivantes :

Test de Hansen n°2

$H0$: $\{\beta'_1 = \beta'_2\}$, i.e. invariance des paramètres de la relation de cointégration

Ha : $\{\beta'_1 \neq \beta'_2\}$, τ inconnu

⁹Comme pour le test de Gregory et Hansen, la date de rupture est balayée sur une fenêtre de taille $\Pi = [0.15T, 0.85T]$.

La statistique de test s'écrit :

$$SupF = \sup_{\tau \in \Pi} F_{T,\tau}$$

3. Dans ce troisième test, le paramètre β_t est supposé suivre une martingale, c'est-à-dire que $\beta'_t = \beta'_{t-1} + \varepsilon_t$, où $E(\varepsilon_t | I_{t-1}) = 0$ et $E(\varepsilon_t \varepsilon'_t) = \delta^2 G_t$. Sous l'hypothèse nulle d'invariance des paramètres, la variance des différences de martingale est nulle, d'où les hypothèses de test suivantes :

Test de Hansen n°3

$$\begin{aligned} H0 & : \{ \delta^2 = 0 \}, \text{ i.e. invariance des paramètres de la relation de cointégration} \\ Ha & : \left\{ \delta^2 > 0, G_t = \left(\hat{\Omega}_{1.2} \otimes V_{T,\tau} \right)^{-1} \right\}, \tau \in \Pi \end{aligned}$$

ce qui conduit à choisir pour statistique de test :

$$MeanF = \frac{\sum_{\tau \in \Pi} F_{T,\tau}}{\sum_{\tau \in \Pi} 1}$$

4. Enfin, le quatrième et dernier test de Hansen semble préférable aux trois tests précédents si le seul but du test de spécification est de vérifier la stabilité des paramètres de la relation de cointégration : il ne comporte pas le choix arbitraire de la zone de balayage Π . Il n'y a pas une rupture de la relation cointégrante mais évolution permanente de ses paramètres. Les hypothèses de test sont les suivantes :

Test de Hansen n°4

$$\begin{aligned} H0 & : \{ \delta^2 = 0 \}, \text{ i.e. invariance des paramètres de la relation de cointégration} \\ Ha & : \left\{ \delta^2 > 0, G_t = \left(\hat{\Omega}_{1.2} \otimes M_{T,1} \right)^{-1} \right\}, \\ & \text{i.e. évolution permanente des paramètres de la relation de cointégration} \end{aligned}$$

pour une statistique de test définie par :

$$L_C = tr \left\{ M_{T,1} \sum_{t=1}^T S_t \hat{\Omega}_{1,2}^{-1} S_t' \right\}$$

8.3.2.2 Deuxième test de spécification de la relation de cointégration : le test de Quintos et Phillips

Les tests de Quintos et Phillips [147] sont assez similaires au test n°4 de Hansen. Leur but est de vérifier la stabilité des coefficients dans la relation de cointégration, grâce à des statistiques obtenues en modifiant le test du multiplicateur de Lagrange. Détaillons le cas d'une relation de cointégration à trois variables. Les coefficients de la relation de cointégration, comme précédemment, peuvent varier au cours du temps :

$$\begin{aligned} y_t &= c + a_t x_t + b_t z_t + e_{1t} \\ &= c + \begin{pmatrix} a_t & b_t \end{pmatrix} \begin{pmatrix} x_t \\ z_t \end{pmatrix} + e_{1t} \\ &= c + A_t' \cdot X_t + e_{1t} \end{aligned}$$

où x_t et z_t sont intégrées à l'ordre 1 et vérifient $\Delta X_t = \Delta \begin{pmatrix} x_t \\ z_t \end{pmatrix} = \begin{pmatrix} e_{xt} \\ e_{zt} \end{pmatrix} = e_{2t}$, avec e_{2t} un processus I(0) donc stationnaire. Le vecteur de cointégration est supposé suivre une modélisation de type marche aléatoire :

$$\begin{aligned} A_t &= A_{t-1} + \varepsilon_t \\ \varepsilon_t &\sim i.i.N(0, \Sigma_\varepsilon) \end{aligned}$$

Cette condition peut se réécrire sous la forme

$$A_t = A_0 + \sum_{s=1}^t \varepsilon_s$$

et

$$\begin{aligned} y_t &= c + A'_0 \cdot x_t + \left(\sum_{s=1}^t \varepsilon_s \right)' x_t + e_{1t} \\ &= c + A'_0 \cdot x_t + \omega_t \end{aligned}$$

$$\text{avec } \omega_t \sim i.i.N(0, \Sigma_\varepsilon)$$

Le processus aléatoire $e_t = \begin{pmatrix} e_{1t} \\ e_{2t} \end{pmatrix}$ a une moyenne nulle et une matrice de variance-covariance instantanée Σ . La matrice

$$\Omega = \lim_{T \rightarrow \infty} E \left(\frac{1}{T} S_T S_T' \right) = \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{i=1}^T \sum_{j=1}^T E(e_i e_j')$$

est partitionnée conformément au vecteur e_t :

$$\Omega = \begin{bmatrix} \Omega_{11} & \Omega_{12} \\ \Omega_{21} & \Omega_{22} \end{bmatrix} \text{ avec } \Omega_{22} = \begin{bmatrix} \Omega_{xx} & \Omega_{xz} \\ \Omega_{zx} & \Omega_{zz} \end{bmatrix}$$

Le paramètre de la relation de cointégration est invariant au cours du temps si $\forall t, A_t = A_{t-1}$, ce qui revient à imposer la contrainte suivante : $\Sigma_\varepsilon = 0$. Si $\Sigma_\varepsilon > 0$, il existe un changement de structure des paramètres et le système ne peut être considéré comme une relation de cointégration : ω_t n'est pas stationnaire. Les hypothèses de test sont les suivantes :

Test de Phillips et Quintos

H_0 : $\{\Sigma_\varepsilon = 0\}$, i.e. stabilité des paramètres de la relation de cointégration

H_a : $\{\Sigma_\varepsilon > 0\}$, i.e. instabilité des paramètres de la relation de cointégration
(absence de cointégration sans rupture)

8.3. LES TESTS DE CHANGEMENT STRUCTUREL DANS UNE RELATION DE COINTÉGRATION

Pour construire la statistique de test, ω_t est réécrit sous la forme suivante :

$$\omega = D_X L \varepsilon + e_1 \text{ avec } L = \begin{bmatrix} I & 0 & \cdots & 0 \\ I & I & \ddots & \vdots \\ \vdots & & \ddots & 0 \\ I & \cdots & \cdots & I \end{bmatrix} \text{ et } D_X = \begin{bmatrix} X'_1 & 0 & \cdots & 0 \\ 0 & X'_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & X'_T \end{bmatrix},$$

le score est calculé sous l'hypothèse nulle de stabilité des paramètres de la relation de cointégration

$$s^* = \Sigma_{11} \omega' A \omega$$

et la statistique du multiplicateur de Lagrange est donnée par :

$$LM = \frac{s^*}{\text{var}(s^*)^{-\frac{1}{2}}} = \frac{\tilde{w}' D_X L L' D'_X \tilde{w}}{\tilde{\Sigma}_{11}}$$

où $\tilde{\omega}_t = y_t - \tilde{c} - \tilde{A}'_0 X_t$ est le résultat de l'estimation par la méthode du maximum de vraisemblance appliquée au modèle contraint par l'hypothèse nulle de stabilité des coefficients de la relation de cointégration. Dans le cas d'un terme déterministe H_t dans la relation de cointégration (i.e. $y_t = \tilde{c} + \tilde{A}'_t \cdot X_t + \Gamma H_t + e_{1t}$), $\tilde{\omega}_t = y_t - \tilde{c} - \tilde{A}'_0 X_t - \tilde{\Gamma} H_t$. Enfin, dans le cas de corrélations temporelles au sein de la relation de cointégration ou d'un biais d'endogénéité, cette relation de cointégration avec ou sans terme déterministe sera estimée à l'aide des moindres carrés "totalement modifiés" et la statistique devient, pour le résidu estimé \tilde{w}^+ :

$$LM = \frac{\tilde{w}^{+'} D_X L \left(I_T \otimes \hat{\Omega}_{22}^{-1} \right) L' D'_X \tilde{w}^+}{\hat{\Omega}_{1.2}}$$

Test de stabilité d'une partie des paramètres. Par rapport aux travaux de Hansen, le principal apport des travaux de Phillips et Quintos réside dans le fait que leur test permet de tester la stabilité de sous-vecteurs de la relation de cointégration, c'est-à-dire d'une partie seulement des paramètres. Détaillons le test dans le cas où la relation de cointégration comporte trois variables stochastiques et qu'un des deux paramètres associés aux régresseurs est soumis à

rupture :

$$\begin{aligned} y_t &= c + a_t x_t + b z_t + e_{1t} \\ \Delta X_t &= \Delta \begin{pmatrix} x_t \\ z_t \end{pmatrix} = \begin{pmatrix} e_{xt} \\ e_{zt} \end{pmatrix} = e_{2t} \\ \text{et } a_t &= a_{t-1} + \varepsilon_{xt} \end{aligned}$$

Comme précédemment,

$$a_t = a_0 + \sum_{s=1}^t \varepsilon_{xs}$$

et

$$\begin{aligned} y_t &= c + a_0 \cdot x_t + b z_t + \left(\sum_{s=1}^t \varepsilon_{xs} \right)' x_t + e_{1t} \\ &= c + a_0 \cdot x_t + b z_t + v_t \end{aligned}$$

L'instabilité de a_t demeure modélisée par la non-nullité de la variance $\sigma_{\varepsilon_x}^2$ associée à ε_{xt} , les hypothèses de test s'écrivent :

Test de Phillips et Quintos sur un sous-ensemble de paramètres

H_0 : $\{\sigma_{\varepsilon_x}^2 = 0\}$, i.e. stabilité de tous les paramètres de la relation de cointégration

H_a : $\{\sigma_{\varepsilon_x}^2 > 0\}$, i.e. instabilité du paramètre associé à x dans la relation de cointégration

La statistique de test est presque inchangée par rapport au cas où tous les paramètres sont soumis à rupture : ce qui concernait le vecteur $X_t = \begin{pmatrix} x_t \\ z_t \end{pmatrix}$ est remplacé par ce qui touche la seule variable x_t . La statistique de test s'écrit, pour le résidu \bar{w}^+ de l'estimation par les moindres carrés "totalement modifiés" de y_t sur la constante, x_t et z_t ,

$$LM = \frac{\bar{w}^{+'} D_x L \left(I_T \otimes \hat{\Omega}_{xx}^{-1} \right) L' D_x' \bar{w}^+}{\hat{\Omega}_{1,2}}$$

8.3.2.3 Troisième test de spécification de la relation de cointégration : le test de Kuo

Le test de Kuo [101] correspond à une extension des tests de Hansen à date de rupture inconnue : seule une partie des paramètres est soumise à rupture. La modification du test n°2 de Hansen permet de tester un changement brutal d'une partie des paramètres, contre une variation de ces paramètres sous forme de martingale pour l'extension du test n°3 de Hansen. La modification du test n°4 de Hansen (où les paramètres varient de façon continue) est identique à celle des travaux de Phillips et Quintos mais explique plus clairement la possibilité d'une modification de la constante au sein de la relation de cointégration.

La relation de cointégration sans rupture s'écrit, pour trois variables,

$$\begin{aligned} y_t &= a_1 + a_2x_t + a_3z_t + e_{1t} = A_tX_t + e_{1t} \\ x_t &= \pi_1 + \pi_2H_t + x_t^0, \quad x_t^0 = x_{t-1}^0 + e_{xt}, \quad t = 1, \dots, T \\ z_t &= \gamma_1 + \gamma_2H_t + z_t^0, \quad z_t^0 = z_{t-1}^0 + e_{zt} \end{aligned}$$

où $H_t = (t, t^2, \dots, t^p)$ et $e_t = \begin{pmatrix} e_{1t} \\ e_{2t} \end{pmatrix}$ avec $e_{2t} = \begin{pmatrix} e_{xt} \\ e_{zt} \end{pmatrix}$ est une séquence d'innovations stationnaires de moyenne nulle. La relation de cointégration est estimée sous l'hypothèse d'absence de rupture des paramètres, par la méthode des moindres carrés "totalement modifiés" de Phillips et Hansen évoqués à la page 317.

Deux types de coefficients peuvent varier au sein de la relation de cointégration :

1. Variation de la constante dans la relation de cointégration :

$$y_t = a_{1,t} + a_2x_t + a_3z_t + e_{yt}$$

2. Variation d'un paramètre associé à un régresseur à tendance stochastique :

$$y_t = a_1 + a_{2,t}x_t + a_3z_t + e_{yt}$$

Quel que soit le type de paramètre soumis à rupture à une date k inconnue, celui-ci s'écrit,

pour une unique modification des paramètres soumis à rupture :

$$\begin{aligned} a_{h,t} &= a_h, & t \leq k \\ &= a_h + \Delta, & t > k \end{aligned}$$

et le premier test de Kuo est défini par les hypothèses de tests suivantes :

Test de Kuo n°1

$H0$: $\{\Delta = 0\}$, i.e. stabilité des paramètres de la relation de cointégration

Ha : $\{\Delta \neq 0\}$, i.e. cointégration avec rupture à la date $k = [\tau T]$ inconnue

pour une statistique de test comparable à celle définie dans le test n°2 de Hansen :

$$SupF_h = \sup_{\tau \in \Pi} F_{h,T,\tau}$$

avec

$$F_{h,T,\tau} = \text{vec}(S_{h,T,\tau})' \left[\hat{\Omega}_{1.2} \otimes V_{h,T,\tau} \right]^{-1} \text{vec}(S_{h,T,\tau})$$

où

$$\begin{aligned} \hat{s}_{h,t} &= \hat{e}_{1t}^+ \text{ si } h = 1 \\ &= x_t \hat{e}_{1t}^+ - \hat{\Lambda}_{21}^+ \text{ si } h = 2 \\ S_{h,T,\tau} &= \sum_{i=1}^{[\tau T]} \hat{s}_{h,i} \\ V_{h,T,\tau} &= M_{h,T,\tau} - M_{h,T,\tau}^{\star} M_{T,1}^{-1} M_{h,T,\tau}^{\star'} \end{aligned}$$

et, si l'on appelle h_t la variable pour laquelle il y a rupture du paramètre,

$$M_{T,\tau} = \sum_{i=1}^{[\tau T]} X_i X_i', \quad M_{h,T,\tau} = \sum_{i=1}^{[\tau T]} h_i h_i', \quad \text{et } M_{h,T,\tau}^{\star} = \sum_{i=1}^{[\tau T]} h_i X_i'$$

En cas d'une modification continue de ce sous-vecteur des paramètres, la modélisation du

8.3. LES TESTS DE CHANGEMENT STRUCTUREL DANS UNE RELATION DE COINTÉGRATION

changement de structure est la suivante :

$$a_{h,t} = a_{h,t-1} + \varepsilon_t$$

avec $E(\varepsilon_t | I_{t-1}) = 0$ et $E(\varepsilon_t \varepsilon_t') = \delta^2 G_{h,t}$

où $G_{h,t}$ représente la structure de la covariance de la martingale $a_{h,t}$ qui détermine le choix entre les tests suivants :

Test de Kuo n°2

$$H0 : \{ \delta^2 = 0 \}, \text{ i.e. invariance des paramètres}$$

$$Ha : \left\{ \delta^2 > 0, G_t = \left(\hat{\Omega}_{1.2} \otimes V_{h,T,\tau} \right)^{-1} \right\}, \tau \in \Pi$$

pour une statistique de test

$$MeanF_h = \frac{\sum_{\tau \in \Pi} F_{h,T,\tau}}{\sum_{\tau \in \Pi} 1}$$

ou

Test de Kuo n°3 = test de Quintos et Phillips sur une partie des paramètres

$$H0 : \{ \delta^2 = 0 \}, \text{ i.e. invariance des paramètres}$$

$$Ha : \left\{ \delta^2 > 0, G_t = \left(\hat{\Omega}_{1.2} \otimes M_{h,T,1} \right)^{-1} \right\}, \tau \in \Pi$$

pour une statistique de test

$$L_{h,C} = tr \left\{ M_{h,T,1}^{-1} \sum_{t=1}^T S_t \hat{\Omega}_{1.2}^{-1} S_t' \right\}$$

8.3.3 Tests dans le cas d'un système de relations de cointégration

Dans le cas d'un système cointégré, estimé et testé par une procédure à la Johansen en l'absence de rupture, deux types de rupture peuvent être prises en compte :

- Le changement du nombre de relations de cointégration, c'est-à-dire du rang de l'espace

de cointégration, a été étudié par Hansen et Johansen [81] ou Quintos [146] en 1995. Quintos suppose la date de rupture connue, et sépare l'échantillon en sous-périodes. Sous l'hypothèse nulle, le rang associé à la matrice des paramètres est inchangé, que l'on soit en première sous-période, en seconde sous-période, ou sur l'ensemble de la période. En d'autres termes, les hypothèses de tests sont les suivantes :

Test de Quintos

H_0 : Cointégration sans rupture

H_a : Cointégration avec rupture

Les statistiques de test correspondent à une extension du test de la trace et du test de ratio de vraisemblance.

Hansen et Johansen supposent quant à eux une date de rupture inconnue et utilisent des méthodes récursives de façon à estimer la date de rupture. Ils suggèrent de prendre en compte les plus grandes valeurs propres λ_i issues des statistiques relatives au test de la trace, et d'effectuer un test de significativité des différences entre les estimations récursives des λ_i .

- En cas de changement des coefficients à l'intérieur des vecteurs cointégrants, le rang de l'espace de cointégration est inchangé par rapport à la procédure standard d'estimation à la Johansen, mais des conditions de normalisation doivent être posées pour conclure à un éventuel changement de structure. Ce type de test a été abordé par Seo en 1998 [153].

8.4 Inadaptation des tests de cointégration à notre problème

8.4.1 La non prise en compte de séries ayant subi une modification du degré d'intégration au cours de la période considérée

La première chose à constater est l'absence de travaux sur le comportement des tests usuels d'intégration comme de cointégration dans le cas de séries ayant subi une modification de leur degré d'intégration. Pourtant, le changement de persistance des séries a été envisagé, d'une

manière différente de la nôtre, par le biais des tests de Kim [97, 98].

Nous avons donc mené une analyse de simulations de Monte Carlo dans le cas de séries intégrées puis stationnaires ou stationnaires puis intégrées, pour différentes dates de rupture. Niveau réel et puissance des tests usuels de cointégration dans le cas de telles séries ne semblent pas dépendre de la place de la sous-période sur laquelle les séries sont intégrées, mais sont liés à la durée de cette période pendant laquelle la série est intégrée. Les tests standards d'intégration ont une forte tendance à conclure que des séries ayant subi une modification du degré d'intégration sont des séries intégrées. L'analyse de la cointégration, légitime dans le cas de séries intégrées, l'est tout autant pour des séries intégrées sur une sous-période et stationnaire sur l'autre sous-période.

Cette étape était en outre indispensable à la poursuite de notre travail, dans la mesure où nous allons rechercher des équilibres de long terme alors que de nombreuses séries ont été testées stationnaires puis intégrées dans le chapitre précédent.

8.4.2 L'absence de tests de changement du degré de cointégration

Les tests de cointégration les plus utilisés ont pour hypothèse nulle l'absence de cointégration, contre une hypothèse alternative d'existence d'une relation de cointégration stable (Engle et Granger [50], Phillips et Ouliaris [141], Kremers, Ericsson et Dolado [100], Zivot [167]...). Il existe en outre des tests possédant des hypothèses de test inverses, qui considèrent sous l'hypothèse nulle l'existence d'une relation de cointégration (Leybourne et McCabe [111], Park et Choi [132], Shin [154], Harris et Inder [82]...). Ces tests usuels ne prennent pas en compte l'éventualité d'une rupture.

Les tests de cointégration avec rupture de Gregory et Hansen [65, 66] considèrent les hypothèses de test suivantes : absence de cointégration contre cointégration avec rupture. Autrement dit, ils considèrent que les séries ne sont pas cointégrées sur l'ensemble de la période contre l'hypothèse alternative que les séries sont cointégrées sur chacune des sous-périodes considérées. Les tests de spécification de la relation de cointégration (Hansen [75], Quintos et Phillips [147], Kuo [101], Quintos [146], Seo [153], Hansen [79], Lütkepohl, Saikkonen et Trenkler [118]) considèrent des hypothèses de test un peu différentes : il s'agit de tester la cointégration sans rupture contre

une hypothèse alternative de cointégration avec rupture (ou de cointégration avec changement continu des paramètres).

La rupture prise en compte dans ces tests correspond à une modification de la relation de cointégration et non du degré de cointégration. Le test que nous désirons construire est un test ayant une hypothèse nulle comparable aux tests usuels de cointégration ainsi qu'aux tests de Gregory et Hansen (à savoir, l'absence de cointégration) contre une hypothèse alternative de cointégration sur une sous-période uniquement.

* * *

Les tests usuels de cointégration opposent existence d'une relation de cointégration et absence de celle-ci, mais ne prennent pas en compte l'éventualité de rupture.

La nature de la rupture a des conséquences différentes sur les conclusions des tests de cointégration standards :

- La présence d'une rupture des séries univariées intervenant dans la relation de cointégration a un faible impact sur le comportement des tests usuels de cointégration.

Nous avons mis en avant que la prise en compte dans la relation cointégrée de séries univariées soumises à une modification de leur degré d'intégration (mais testées intégrées sur l'ensemble de la période) n'invalide pas non plus les tests de cointégration usuels.

- Mais la présence d'une rupture de la relation de long terme peut conduire à des conclusions erronées lors de l'application des tests de cointégration standards : l'hypothèse de cointégration est fréquemment rejetée en présence de cointégration avec rupture.

Depuis une dizaine d'années s'est développée une littérature concernant les tests de cointégration avec rupture. La rupture prise en compte par ces tests correspond généralement à une modification de la relation de cointégration et non du degré de cointégration. Ce type de test est motivé et construit dans le chapitre suivant.

Chapitre 9

Tests de modification du degré de cointégration

La majorité des tests présentés au chapitre précédent s'intéressent à une modification de l'équilibre de long terme. En présence d'une rupture de la relation de cointégration, les tests usuels de cointégration ont tendance à conclure que les résidus sont intégrés, et qu'il n'y a pas d'équilibre de long terme. La solution généralement proposée revient à estimer la relation de cointégration en supposant l'existence d'une rupture (à une date connue ou inconnue), puis à tester l'hypothèse de stationnarité des résidus.

Nous proposons une autre méthode, qui découle directement des tests de modification d'intégration que nous avons construits dans la partie précédente. L'idée générale est la suivante : en l'absence de relation de cointégration, les résidus estimés sont non stationnaires, et nous pouvons appliquer nos tests de modification du degré d'intégration. Si ces tests concluent que le résidu est stationnaire sur une sous-période et intégré sur l'autre sous-période, nous pourrions dire qu'il existe un équilibre de long terme sur une sous-période.

Ce chapitre un peu technique a pour but de présenter les propriétés de ces tests de modification du degré de cointégration. Les hypothèses de test seront détaillées dans une première section ; le calcul des statistiques de test dans la deuxième. Pour permettre une utilisation opérationnelle de ces tests, les valeurs critiques seront présentées dans la troisième section. Enfin,

dans la quatrième section, nous établirons les propriétés de ces tests en terme de puissance et d'estimation de la date de rupture.

9.1 Motivation et hypothèses de test

Considérons deux séries intégrées à l'ordre 1 : y_t et x_t , et l'éventuelle relation de cointégration normalisée qui relie ces variables :

$$y_t = \hat{a}x_t + \hat{e}_t.$$

Cette relation de cointégration peut en outre comprendre une constante c ou une tendance t .

Dans les tests de cointégration classiques, les hypothèses de tests sont les suivantes :

$$H_0 : \hat{e}_t \sim I(1),$$

i.e. absence de cointégration pour l'ensemble de la période

$$H_a : \hat{e}_t \sim I(0),$$

i.e. cointégration entre y_t et x_t pour l'ensemble de la période

Il semble possible de considérer qu'un équilibre de long terme a été valide pendant une certaine période, puis ne l'a plus été, ou inversement. De la même façon, une variable a pu intervenir pendant un temps dans la relation et ne jouer plus aucun rôle ensuite, tandis que de nouveaux facteurs peuvent apparaître déterminants à partir d'une certaine période. Par exemple, on peut supposer que, du moins sur certaines lignes protégées, la vitesse de circulation dans Paris pénalise de moins en moins le transport en autobus à Paris du fait de la mise en place de couloirs protégés. A l'inverse, la qualité de l'information sur les lignes ou la propreté semblent des critères de plus en plus importants aux yeux des voyageurs.

Cette idée nous a conduit à construire des tests qui laissent la possibilité à un équilibre de long terme sur une sous-période, même si celui-ci n'est pas valide sur l'ensemble de la période. Les tests sont construits en supposant une date de rupture inconnue : ils déterminent eux-mêmes la date éventuelle d'apparition/disparition de la relation de cointégration. Les hypothèses de test

deviennent :

$$H_0 : \hat{e}_t \sim I(1),$$

i.e. absence de cointégration sur l'ensemble de la période

$$H_a : \hat{e}_t \sim I(0) - I(1) \text{ ou } I(1) - I(0),$$

i.e. cointégration entre y_t et x_t sur une des sous-périodes.

9.2 Estimation et statistiques de test

Dans un premier temps, la relation de cointégration éventuelle est estimée par les moindres carrés totalement modifiés. Elle permet d'obtenir les résidus \hat{e}_t , auxquels sera appliqué un test de modification du degré d'intégration.

Deux cas peuvent être considérés :

- Il y a eu une relation de cointégration pendant une première sous-période, et celle-ci n'est plus valide sur une deuxième sous-période
- Il n'y avait pas de relation de cointégration sur une première sous-période et un équilibre de long terme est apparu en deuxième sous-période.

Pour chacun de ces cas, un test supérieur de Wald et un test inférieur de Student peuvent être appliqués aux résidus de la relation testée, en contraignant le paramètre auto-régressif des résidus à un sur la période adéquate. Cette méthode en deux étapes (estimation de la relation de long terme puis application d'un test sur le degré d'intégration) est assez similaire à celle de Phillips et Ouliaris. La différence provient du fait que le test appliqué en deuxième étape n'est pas un test de racine unitaire mais un test de modification du degré d'intégration.

Notons

- $Y_{-i,k}$ le vecteur de taille $(T \times 1)$ composé des y_{t-i} de la ligne 1 à la ligne k puis de 0 de la ligne $t = k + 1$ jusqu'à la dernière ligne,
- $Y_{k,-i}$ le vecteur de taille $(T \times 1)$ composé de 0 jusqu'à la ligne k , puis des y_{t-i} de la ligne $t = k + 1$ à la ligne T .

Hypothèse nulle	Absence de cointégration sur l'ensemble de la période	
Hypothèse alternative	Absence puis présence de relation de cointégration	Présence puis absence de relation de cointégration
Estimation de la relation de cointégration	Avec constante : $y_t = \hat{c} + \hat{a}x_t + \hat{e}_t$. Avec constante et tendance : $y_t = \hat{c} + \hat{b}t + \hat{a}x_t + \hat{e}_t$.	
Modèle de régression	$\Delta \hat{e} = W\theta + X_{is}\beta_{is} + \varepsilon$ avec $W = (1, \Delta \hat{e}_{-i}, \Delta \hat{e}_{k,-i}), i = 1, \dots, k$ $X_{is} = ([1]_{t>k}, \hat{e}_{k,-1})$ $\beta'_{is} = (\Delta c_{is}, \Delta \rho_{is})$ où $\Delta \rho_{is} = \rho_b - 1$	$\Delta \hat{e} = W\theta + X_{si}\beta_{si} + \varepsilon$ avec $W = (1, \Delta \hat{e}_{-i}, \Delta \hat{e}_{-i,k}), i = 1, \dots, k$ $X_{is} = ([1]_{t \leq k}, \hat{e}_{-1,k})$ $\beta'_{is} = (\Delta c_{si}, \Delta \rho_{si})$ où $\Delta \rho_{is} = \rho_a - 1$
Test supérieur de Wald	★ <u>Hypothèses</u> : $H_0 : \beta_{is} = 0$ $H_a : \beta_{is} \neq 0$ ★ <u>Statistique</u> : $\sup_k Wald_T(k)$ où $Wald_T(k) = \frac{\Delta \hat{\beta}'_{is}(X'MX)\Delta \hat{\beta}_{is}}{\hat{\sigma}^2(k)}$ avec $M = I - W(W'W)^{-1}W'$ et $\hat{\sigma}^2(k) = \frac{S_T(k)}{T-1}$	★ <u>Hypothèses</u> : $H_0 : \beta_{is} = 0$ $H_a : \beta_{is} \neq 0$ ★ <u>Statistique</u> : $\sup_k Wald_T(k)$ où $Wald_T(k) = \frac{\Delta \hat{\beta}'_{si}(X'MX)\Delta \hat{\beta}_{si}}{\hat{\sigma}^2(k)}$ avec $M = I - W(W'W)^{-1}W'$ et $\hat{\sigma}^2(k) = \frac{S_T(k)}{T-1}$
Test inférieur de Student	★ <u>Hypothèses</u> : $H_0 : \Delta \rho_{is} = 0$ $H_a : \Delta \rho_{is} < 0$ ★ <u>Statistique</u> : $\inf_k Student_T(k)$ avec $Student_T(k) = \frac{\Delta \hat{\rho}_{is,k}}{\hat{\sigma}(\Delta \hat{\rho}_{is,k})}$	★ <u>Hypothèses</u> : $H_0 : \Delta \rho_{si} = 0$ $H_a : \Delta \rho_{si} < 0$ ★ <u>Statistique</u> : $\inf_k Student_T(k)$ avec $Student_T(k) = \frac{\Delta \hat{\rho}_{si,k}}{\hat{\sigma}(\Delta \hat{\rho}_{si,k})}$

9.3 Tables de valeurs critiques

Les tables de valeurs critiques ne dépendent pas de la modélisation de la série des résidus estimés, mais de la structure de l'équation de cointégration : absence éventuelle d'une constante c et d'une tendance t et nombre de variables exogènes hors c et t .

Pour calculer les valeurs critiques dans chacun des cas, nous avons effectué 10 000 simulations de Monte-Carlo sur des échantillons de taille 150. Après estimation des résidus de l'éventuelle relation de cointégration, 10% des observations sont exclues au début et à la fin de l'échantillon.

9.3.1 Processus générateur de données sous l'hypothèse nulle

Sous l'hypothèse nulle, il n'existe pas de relation de cointégration sur l'ensemble de la période. Dans les simulations, nous avons construit les $p - 1$ variables x_t apparaissant dans la relation de cointégration comme des processus intégrés à l'ordre 1 (plus précisément, des marches aléatoires). La $p^{\text{ième}}$ série, y_t , est construite en utilisant la formule

$$y_t = ax_t + c + e_t$$

si la relation de cointégration ne possède qu'une constante, ou, dans le cas d'une relation de cointégration avec constante et tendance

$$y_t = ax_t + bt + c + e_t$$

où e_t est une marche aléatoire :

$$e_t = \sum_{i=1}^t \varepsilon_i, \quad \varepsilon_i \sim BB(0, 1)$$

Les valeurs critiques des tests de modification du degré de cointégration pour une variable explicative (en plus de la variable expliquée et de la constante au sein de la relation de cointégration) sont presque identiques aux valeurs critiques obtenues pour les tests de modification du degré d'intégration. Les valeurs critiques associées aux tests de modification du degré de cointégration sont à peine plus élevées en valeur absolue, pour un écart de un à deux points seulement.

Elles augmentent en valeur absolue lorsque le nombre de variables présentes dans la relation de cointégration s'accroît ou en cas de tendance déterministe. Ces résultats sont similaires à ceux obtenus par Phillips et Ouliaris en cas de résidu intégré tout au long de la période.

Elles sont plus élevées dans le cas "absence de cointégration puis cointégration" que dans le cas "cointégration puis absence de cointégration", i.e. pour un résidu intégré puis stationnaire que pour un résidu stationnaire puis intégré. Ceci concorde avec les résultats de la partie précédente.

9.3.2 Valeurs critiques du test supérieur de Wald

Nb de variables (en plus de c)	$\hat{e}_t \sim I(1) - I(0)$				$\hat{e}_t \sim I(0) - I(1)$			
	1%	5%	10%	20%	1%	5%	10%	20%
1	23.092	18.816	16.630	14.361	19.889	15.664	13.620	11.565
2	24.657	19.378	17.275	14.870	21.751	17.067	14.877	12.599
3	25.580	20.531	18.380	15.796	23.790	18.424	16.190	13.802
4	27.398	22.150	19.760	17.012	25.265	20.324	17.952	15.392
5	29.543	23.766	21.159	18.350	27.911	22.366	19.786	17.030

Nb de variables (en plus de c et t)	$\hat{e}_t \sim I(1) - I(0)$				$\hat{e}_t \sim I(0) - I(1)$			
	1%	5%	10%	20%	1%	5%	10%	20%
1	24.160	19.649	17.319	14.931	22.144	17.062	14.852	12.552
2	25.777	20.836	18.428	15.827	24.714	18.827	16.347	13.794
3	26.790	22.052	19.773	16.945	25.750	20.468	17.948	15.305
4	28.977	23.784	21.295	18.370	27.487	22.110	19.571	17.020
5	31.472	25.891	23.128	19.926	30.030	24.107	21.524	18.486

9.3.3 Valeurs critiques du test inférieur de Student

Nb de variables (en plus de c)	$\hat{e}_t \sim I(1) - I(0)$				$\hat{e}_t \sim I(0) - I(1)$			
	1%	5%	10%	20%	1%	5%	10%	20%
1	-4.549	-4.041	-3.775	-3.483	-4.282	-3.742	-3.481	-3.171
2	-4.712	-4.140	-3.887	-3.580	-4.500	-3.960	-3.681	-3.368
3	-4.833	-4.303	-4.039	-3.724	-4.751	-4.174	-3.887	-3.564
4	-5.033	-4.481	-4.226	-3.897	-4.934	-4.406	-4.125	-3.789
5	-5.249	-4.676	-4.398	-4.078	-5.192	-4.624	-4.340	-4.009

Nb de variables (en plus de c et t)	$\hat{e}_t \sim I(1) - I(0)$				$\hat{e}_t \sim I(0) - I(1)$			
	1%	5%	10%	20%	1%	5%	10%	20%
1	-4.653	-4.104	-3.838	-3.537	-4.575	-3.956	-3.675	-3.369
2	-4.811	-4.291	-4.027	-3.699	-4.798	-4.221	-3.909	-3.580
3	-4.960	-4.475	-4.189	-3.862	-4.967	-4.409	-4.118	-3.783
4	-5.208	-4.679	-4.399	-4.059	-5.145	-4.612	-4.320	-4.011
5	-5.429	-4.898	-4.591	-4.256	-5.387	-4.825	-4.538	-4.188

9.4 Calculs de puissance

La puissance des tests est évaluée à l'aide de simulations de Monte-Carlo, en examinant le rôle de différents facteurs :

1. la valeur du paramètre auto-régressif dans la modélisation des résidus de l'éventuelle relation de cointégration pour la sous-période sur laquelle il y a cointégration,
2. la valeur de la constante dans la modélisation auto-régressive des résidus pour la sous-période sur laquelle il y aurait cointégration,
3. la date de la véritable date de rupture.

Pour chacun de ces facteurs, la puissance est calculée pour les tests supérieur de Wald et inférieur de Student dans le cas d'une relation de long terme avec tendance et/ou constante

$$y_t = c + x_t + e_t \text{ (avec constante)}$$

$$y_t = c + t + x_t + e_t \text{ (avec constante et tendance)}$$

avec

$$x_t \sim I(1) \text{ (marche aléatoire)}$$

$$e_t = c_1 + \rho_1 e_{t-1} + u_t \text{ si } t \leq k$$

$$= c_2 + \rho_2 e_{t-1} + u_t \text{ si } t > k$$

en supposant qu'il n'existe pas de cointégration sur la première ou la seconde sous-période (i.e. respectivement $\{\rho_1 = 1, |\rho_2| < 1\}$ ou $\{|\rho_2| < 1, \rho_1 = 1\}$).

Le niveau des tests est posé égal à 10%, la date de rupture est estimée en enlevant 10% des observations au début de l'échantillon et 10% des observations en fin d'échantillon. 2 500 simulations de Monte-Carlo sont réalisées pour chaque calcul de puissance. Les tables de puissance sont reportées dans l'annexe E à la page 643

9.4.1 Puissance des tests suivant la valeur du paramètre auto-régressif des résidus de la relation de long terme sur la période où il y a cointégration

La puissance des tests à 10% est calculée dans un premier temps en posant une constante nulle sur chacune des sous-périodes, avec une véritable date rupture à la moitié de l'échantillon. 2 500 simulations ont été réalisées. Dans les simulations suivantes, le paramètre auto-régressif sur la sous-période intégrée est posé égal à 1, tandis que le paramètre associé à l'autre sous-période varie entre -1 et 1 (exclus).

9.4.1.1 Cas d'un résidu de la relation de cointégration éventuelle intégré puis stationnaire : "Absence de cointégration puis cointégration"

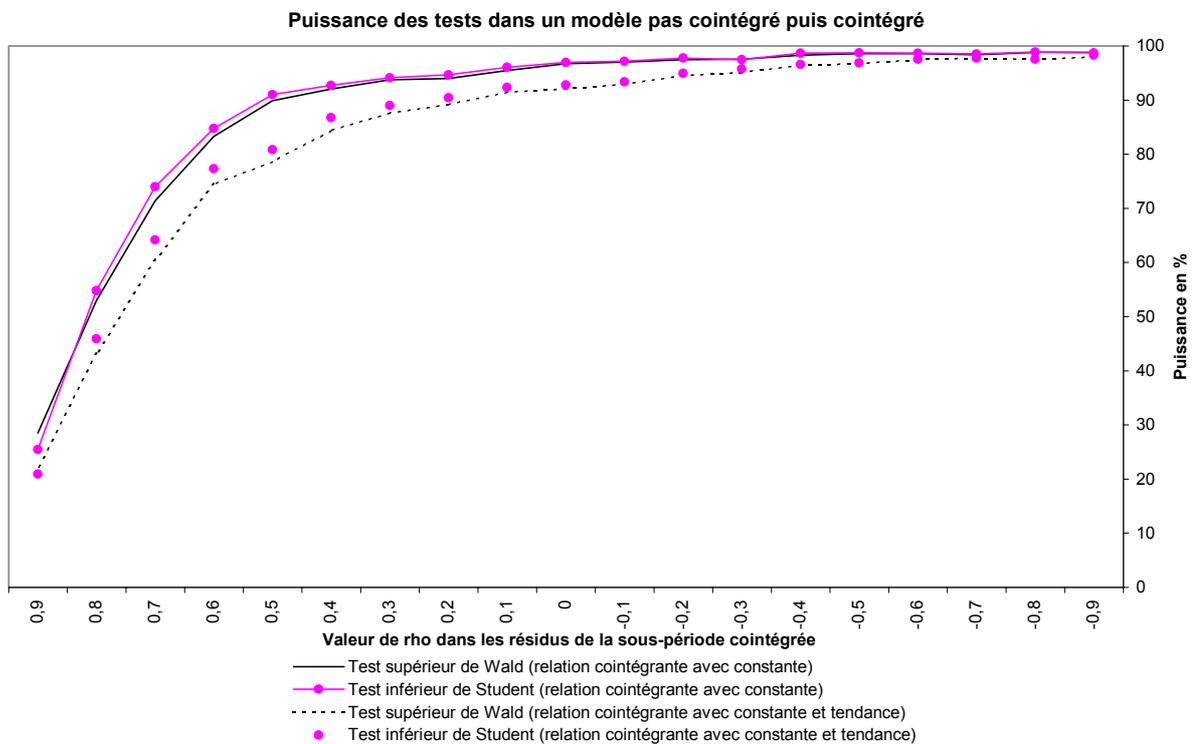
Posons $\rho_1 = 1$, $c_1 = c_2 = 0$ et faisons varier le paramètre ρ_2 pour une rupture à la date $k = 75$. Les résultats du test supérieur de Wald et du test inférieur de Student sont présentés sous forme de graphiques, les tableaux équivalents étant reportés à l'annexe E.2 page 645.

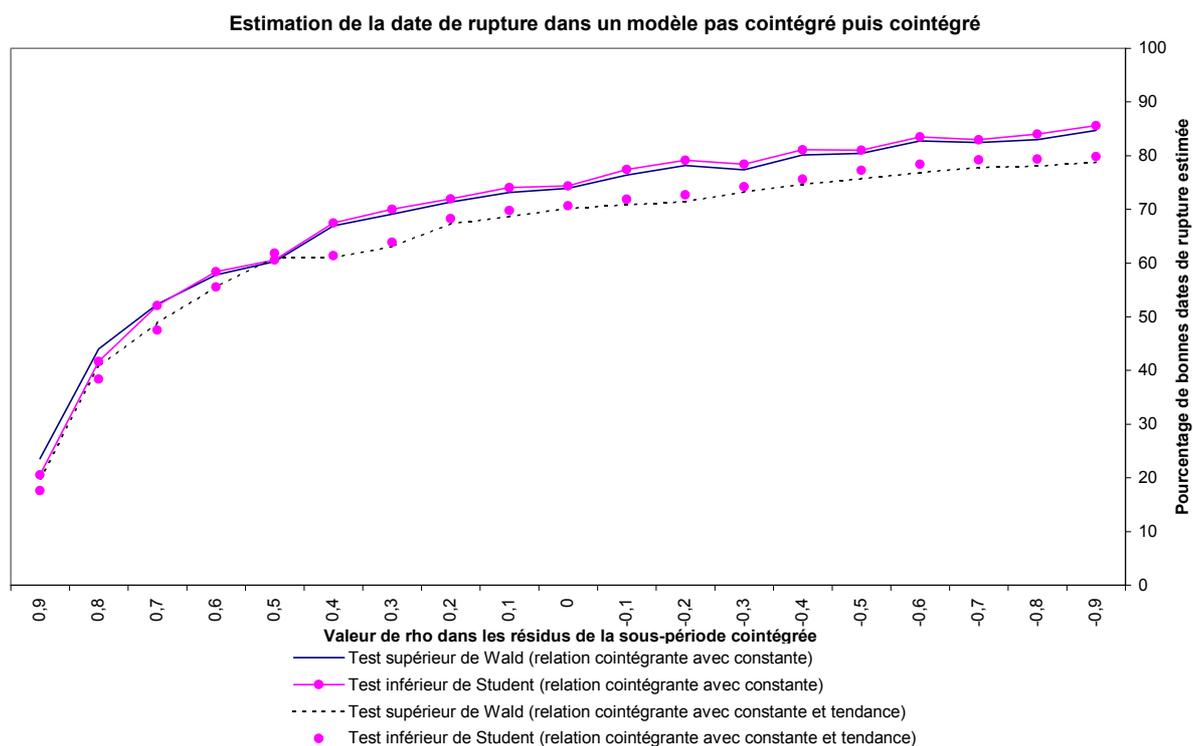
Les principaux faits marquants, en l'absence de constante sur chacune des sous périodes du résidu et pour une date de rupture égale à 75, sont les suivants :

- Dans le cas d'un modèle non cointégré puis cointégré, les puissances des tests supérieur de Wald et inférieur de Student sont équivalentes, pour une relation de cointégration comparable (avec constante ou constante et tendance). De même, l'estimation de la date de rupture est presque identique quel que soit le type de test choisi. Toutefois, comme c'était le cas pour le test de modification du degré d'intégration, le test supérieur de Wald a un peu plus de puissance lorsque ρ approche 1 (hypothèse nulle).
- La puissance des tests supérieur de Wald et inférieur de Student diminue au fur et à mesure que le paramètre associé au résidu de la période cointégré se rapproche de 1, pour atteindre le niveau nominal des tests. De même, la qualité de l'estimation de la date de rupture se dégrade au fur et à mesure que ρ_2 est proche de 1.
- Enfin, les puissances obtenues par les tests supérieur de Wald et inférieur de Student sous l'hypothèse alternative "absence de cointégration puis cointégration" sont inférieures à celles obtenues lors d'un test de modification du degré d'intégration (hypothèse alternative : "intégré puis stationnaire"). Pour $\rho_1 = 1$ et $\rho_2 = 0.8$, une puissance de 66.12% était obtenue pour le test supérieur de Wald (70.72% pour le test inférieur de Student) dans le cas d'un test ayant pour hypothèses "la série est intégrée sur l'ensemble de la période" contre "la série est intégrée puis stationnaire". Pour $\rho_1 = 1$ et $\rho_2 = 0.8$ dans le résidu de la relation de cointégration, le test supérieur de Wald de modification du degré de cointégration a une puissance de 53% si la relation de cointégration comporte uniquement une constante, et 43.20% dans le cas d'une relation de cointégration avec constante et tendance déterministe. Le test inférieur de Student a une puissance égale à 54.84% pour une relation de cointégration avec constante, et 45.92% en présence d'une tendance. Cette

perte de puissance s'explique par le fait que le test de degré d'intégration n'est pas appliqué sur la véritable série des résidus mais sur les résidus estimés.

- De même, la qualité de l'estimation de la date de rupture est un peu moins bonne que dans le cas des tests d'intégration ayant pour hypothèse alternative "la série est intégrée puis stationnaire". Le pourcentage de bonnes dates de rupture s'élevait à 45.25% pour le test supérieur de Wald et 42.03% pour le test inférieur de Student, contre respectivement 44% et 41.65% dans le cas de tests ayant pour hypothèse alternative "absence de cointégration puis cointégration" (relation de cointégration avec constante), et 40.71% et 38.42% (relation de cointégration avec constante et tendance).





9.4.1.2 Cas d'un résidu de la relation de cointégration éventuelle stationnaire puis intégré : "Cointégration puis absence de cointégration"

Posons $\rho_2 = 1$, $c_1 = c_2 = 0$ et faisons varier le paramètre ρ_1 pour une rupture à la date $k = 75$. Les résultats du test supérieur de Wald et du test inférieur de Student sont présentés sous forme de graphiques, les tableaux équivalent étant reportés à dans l'annexe E.2 page 651.

Les principaux faits marquants, en l'absence de constante sur chacune des sous-périodes du résidu et pour une date de rupture égale à 75, sont les suivants :

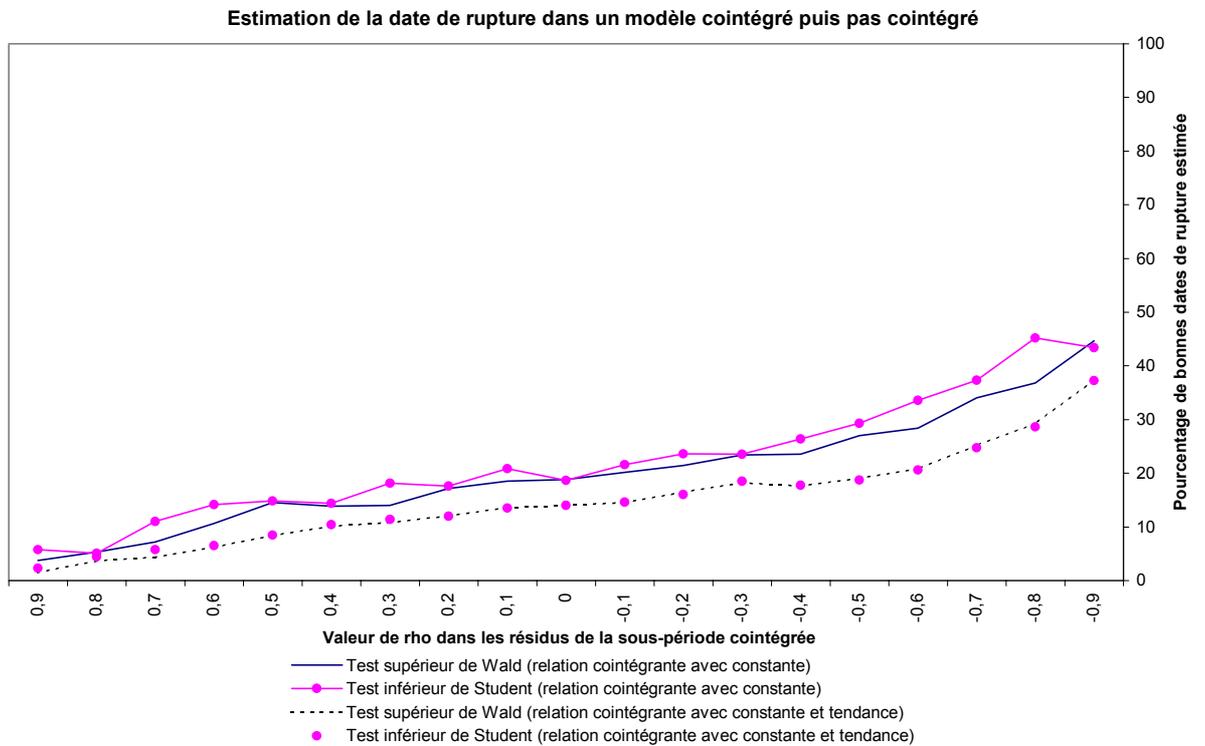
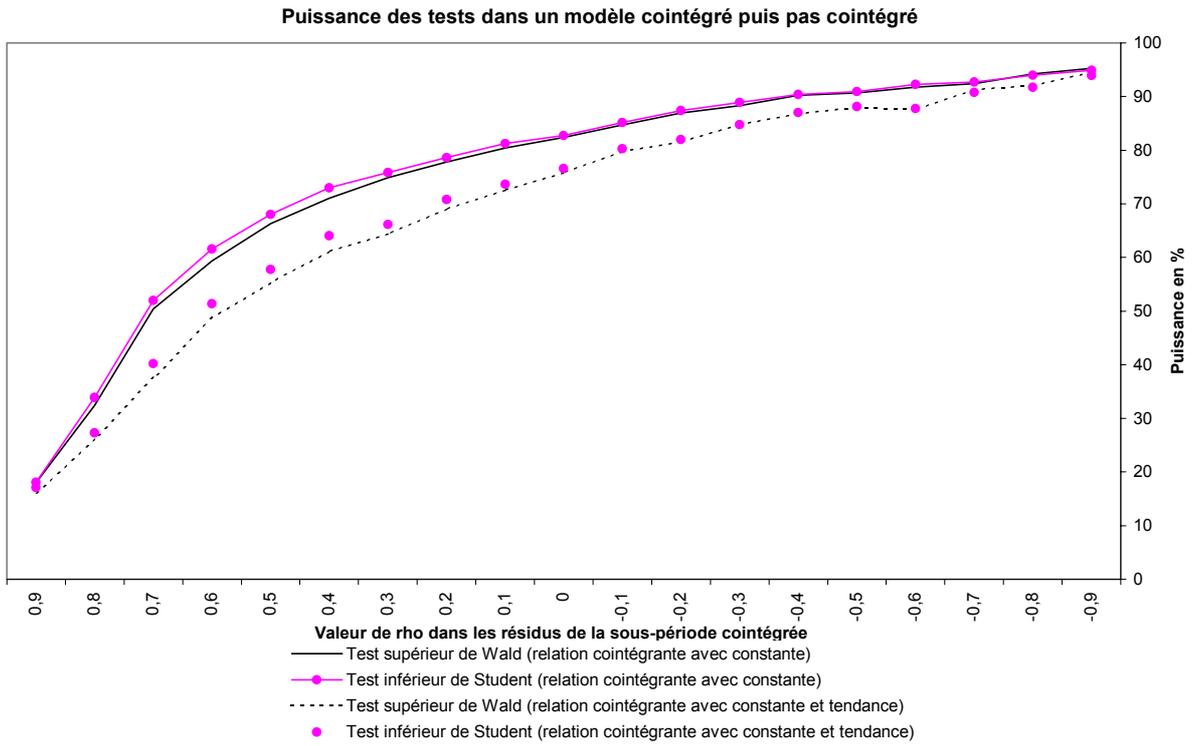
- Dans le cas d'un modèle cointégré puis non cointégré, comme c'était le cas pour un modèle non cointégré puis cointégré, les puissances des tests supérieur de Wald et inférieur de Student sont équivalentes, pour une relation de cointégration comparable (avec constante ou constante et tendance). De même, l'estimation de la date de rupture est presque identique quel que soit le type de test choisi.
- De la même façon que la puissance dans le cas de processus "stationnaires puis intégrés"

est plus faible que dans le cas de processus "intégrés puis stationnaires", la puissance des tests de modification du degré de cointégration est moins élevée pour un modèle "cointégré puis non cointégré" que pour un modèle "non cointégré puis cointégré". Ainsi, pour $\rho_1 = 0.8$, la puissance du test supérieur de Wald s'élève à 32.44% (relation de cointégration avec constante) et 26.16% (relation de cointégration avec constante et tendance) dans ce deuxième cas contre respectivement 53% et 43.20% pour l'hypothèse alternative "pas de cointégration puis cointégration".

- L'estimation de la date de rupture est altérée dans ce deuxième cas : seules 4.32% des bonnes dates de rupture sont estimées par le test supérieur de Wald pour $\rho_1 = 0.8$, $\rho_2 = 1$, et une relation de cointégration avec constante, contre 44% pour $\rho_1 = 1$, $\rho_2 = 0.8$. Dans le cas d'une relation de cointégration avec constante et tendance, le pourcentage de bonnes dates estimées tombe à 2.91% contre 40.75% sous l'hypothèse alternative précédent. Le test supérieur de Student ne donne pas de meilleurs résultats : 3.78% pour une relation de cointégration avec constante et 3.52% en cas de tendance déterministe.

Cette forte diminution de la qualité de l'estimation de la date de rupture avait été mise en avant lors de l'analyse de la puissance des tests de modification du degré d'intégration ayant pour hypothèse alternative "la série est stationnaire puis intégrée".

- La puissance des tests supérieurs de Wald et inférieur de Student, et la qualité de l'estimation de la date de rupture, augmentent au fur et à mesure que le paramètre associé au résidu de la période cointégrée s'éloigne de 1.



9.4.2 Puissance des tests suivant la valeur de la constante dans la modélisation auto-régressive des résidus pour la sous-période sur laquelle il y aurait cointégration

9.4.2.1 Cas d'un résidu de la relation de cointégration éventuelle intégré puis stationnaire : "Absence de cointégration puis cointégration".

Soient $\rho_1 = 1$ et $\rho_2 = 0.8$ pour une véritable date de rupture égale à 75. La constante dans le résidu est nulle sur la partie non cointégrée ($c_1 = 0$), nous faisons varier la constante sur la partie cointégrée.

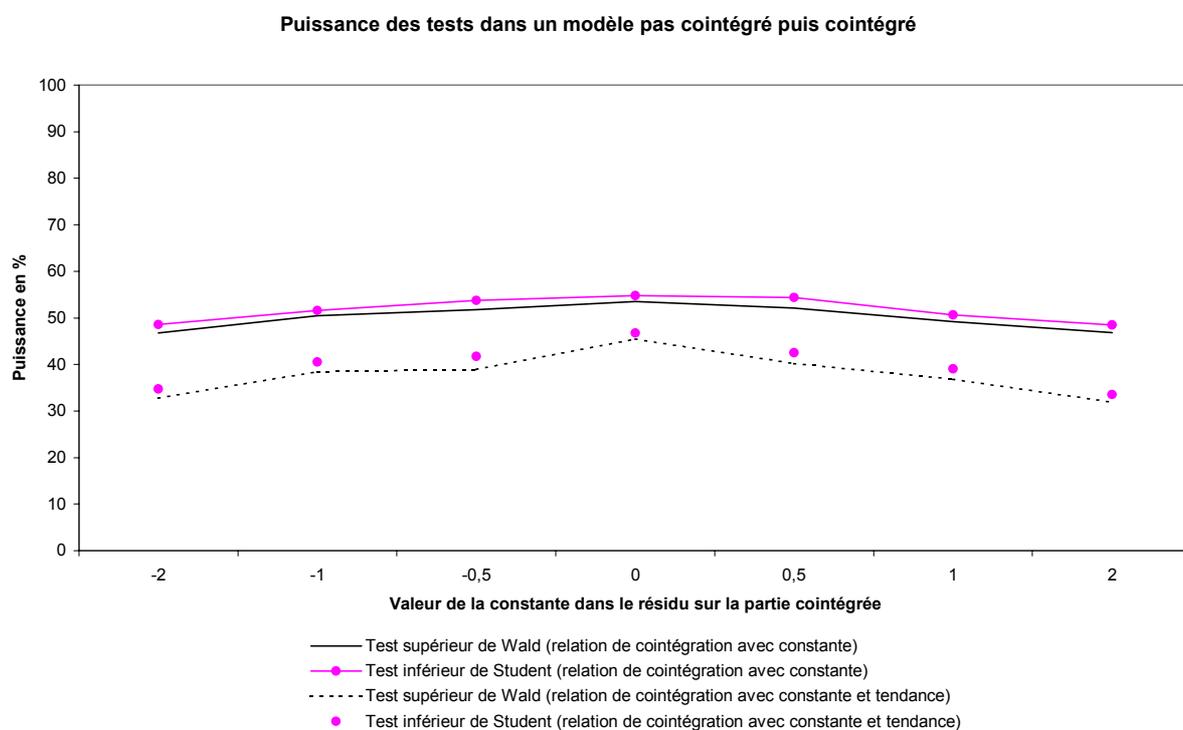
Les principales remarques sont les suivantes :

- La puissance des tests supérieur de Wald et inférieur de Student dans le cas d'une série intégrée puis stationnaire diminue légèrement quand la valeur absolue de la constante sur la partie stationnaire s'accroît. C'est l'importance en valeur absolue du changement de la constante qui est à prendre en compte, et non le signe de la variation : courbes de puissance et mesures de la qualité de l'estimation de la date de rupture sont symétriques par rapport à $c_2 = 0$. Le cas où la constante est nulle sur la partie stationnaire induit la puissance maximale des tests. Ainsi, dans le cas d'une relation de cointégration avec constante, la puissance du test supérieur de Wald atteint 53.52% lorsque $c_2 = 0$ contre environ 46.8% pour $|c_2| = 0$, la puissance du test inférieur de Student s'élève respectivement à 54.8% et environ 48.5%. En présence de tendance dans la relation de cointégration, le test supérieur de Wald conduit à des puissances égales à 45.44% et environ 32% (46.76% et environ 34% pour le test inférieur de Student.
- Dans le cas d'une série non cointégrée puis cointégrée, la qualité de l'estimation s'accroît pour chacun des tests quand la constante augmente en valeur absolue. Ainsi, si la constante sur la partie stationnaire du résidu est nulle, le test supérieur de Wald dans le cas d'une relation de cointégration avec constante estime la date de rupture correcte dans 45.74% des cas (43.87% pour le test inférieur de Student) contre 62 % dans le cas d'une constante égale à 2 (respectivement 60.23%). Une qualité similaire de l'estimation de la date de rupture était obtenue dans le cas d'une modification du degré d'intégration ayant pour hypothèse

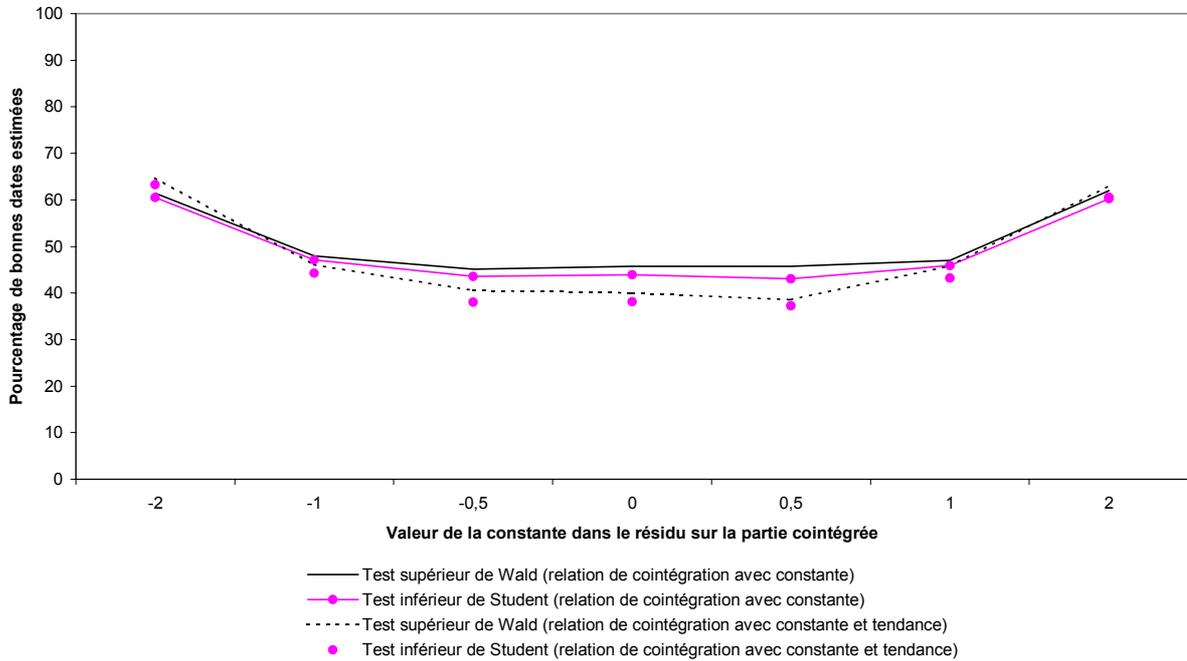
alternative "la série est intégrée puis stationnaire", à un point environ près. Dans le cas d'une relation de cointégration avec tendance, le pourcentage de bonnes dates estimées lorsque $c_2 = 0$ atteint 40.05% pour le test supérieur de Wald et 38.15% pour le test de Student (respectivement 62.99% et 60.55% pour $c_2 = 2$).

- Tests supérieur de Wald et inférieur de Student ont des résultats en terme de puissance et d'estimation de la date de rupture assez comparables.

Les tableaux de chiffres sont reportés dans l'annexe E.3 page 657.



Estimation de la date de rupture dans un modèle pas cointégré puis cointégré



9.4.2.2 Cas d'un résidu de la relation de cointégration éventuelle stationnaire puis intégré : "Cointégration puis absence de cointégration"

Posons $\rho_1 = 0.8$ et $\rho_2 = 1$ pour une véritable date de rupture égale à 75. La constante est nulle sur la partie non cointégrée ($c_2 = 0$), elle varie sur la période cointégrée.

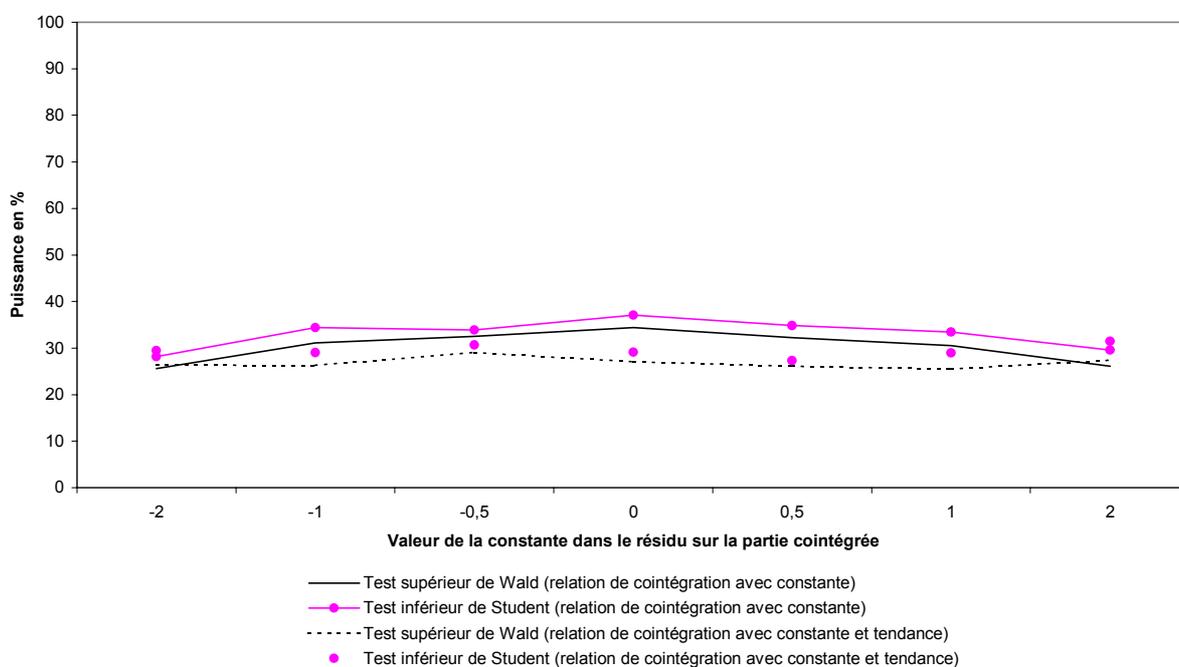
Les faits marquants associés aux tests ayant pour hypothèse alternative "cointégration puis absence de cointégration" se résument de la façon suivante :

- Comme pour le modèle non cointégré puis cointégré, la puissance des tests supérieur de Wald et inférieur de Student diminue quand la valeur absolue de la constante sur la partie stationnaire augmente, de façon symétrique autour de $c_1 = 0$, où la puissance est maximale.
- De même, la qualité de l'estimation s'accroît pour chacun des tests quand la constante augmente en valeur absolue.
- Le test inférieur de Student a une puissance comparable au test supérieur de Wald, même si elle se situe à des niveaux inférieurs à ceux obtenus dans le cas d'un modèle non cointégré

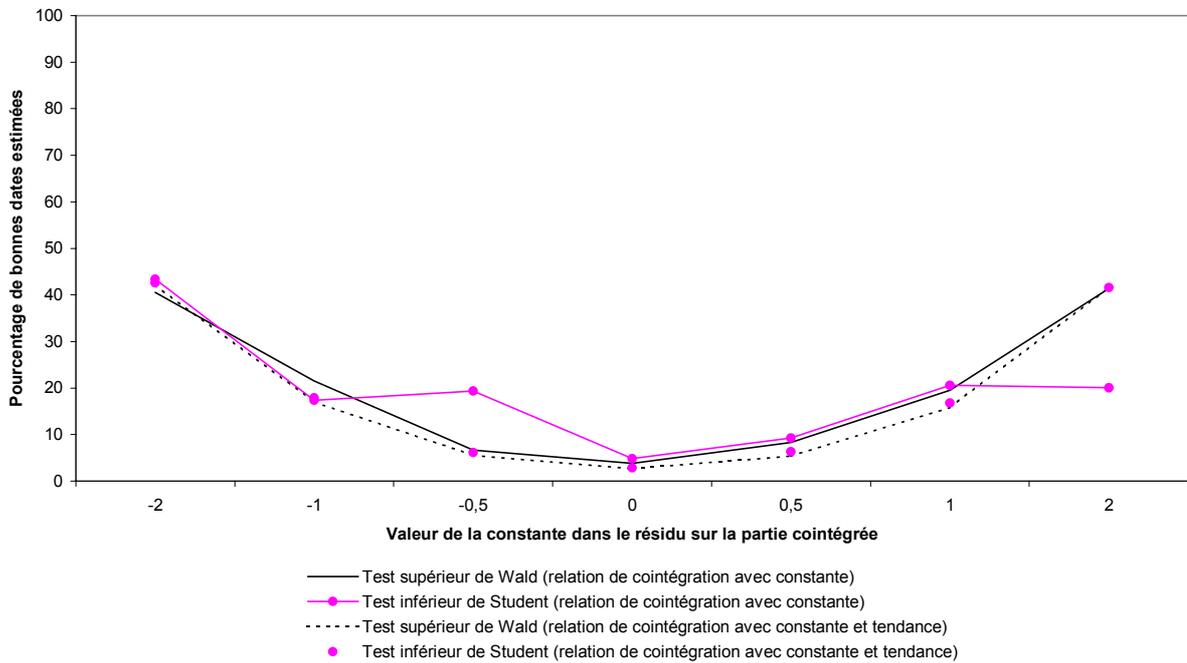
puis cointégré. Pour $c_1 = 2$, par exemple, la puissance du test supérieur de Wald s'élève à 26.08% (relation de cointégration avec constante) et 27.36% (relation de cointégration avec constante et tendance) et la puissance du test inférieur de Student à 29.56% et 31.44%. Par comparaison, dans le cas d'un modèle non cointégré puis cointégré, nous obtenions respectivement 46.84 et 31.88% pour le test supérieur de Wald, 48.48% et 33.56% pour le test inférieur de Student.

Les tableaux de chiffres relatifs au calcul de puissance sont repoussés dans l'annexe E.3 page 659.

Puissance des tests dans un modèle cointégré puis pas cointégré



Estimation de la date de rupture dans un modèle cointégré puis pas cointégré



9.4.3 Puissance des tests suivant la véritable date de rupture

9.4.3.1 Cas d'un résidu de la relation de cointégration éventuelle intégré puis stationnaire : "Absence de cointégration puis cointégration"

Posons $\rho_1 = 1, \rho_2 = 0.8, c_1 = c_2 = 0$ et calculons la puissance en faisant varier la date de rupture. Les principaux résultats sont les suivants :

- Dans le cas d'un modèle non cointégré puis cointégré, les tests supérieur de Wald et inférieur de Student ont une puissance plus importante quand la rupture a lieu au début de l'échantillon plutôt qu'à la fin. Ainsi, pour un processus généré par $\rho_1 = 1, \rho_2 = 0.8, c_1 = 0, c_2 = 1$, et une relation de cointégration avec constante, le test supérieur de Wald a une puissance de 68.4% quand la rupture se produit à la date 25, contre 40.88% pour une rupture à la date 125 (respectivement 70.80% et 33.56% pour le test inférieur de Student). Lorsque la relation de cointégration comporte en outre une tendance, le test supérieur de Wald a une puissance égale à 61.76% pour $k = 25$ et 33.88% pour $k = 125$ (respectivement

65.24% et 27.16% pour le test inférieur de Student).

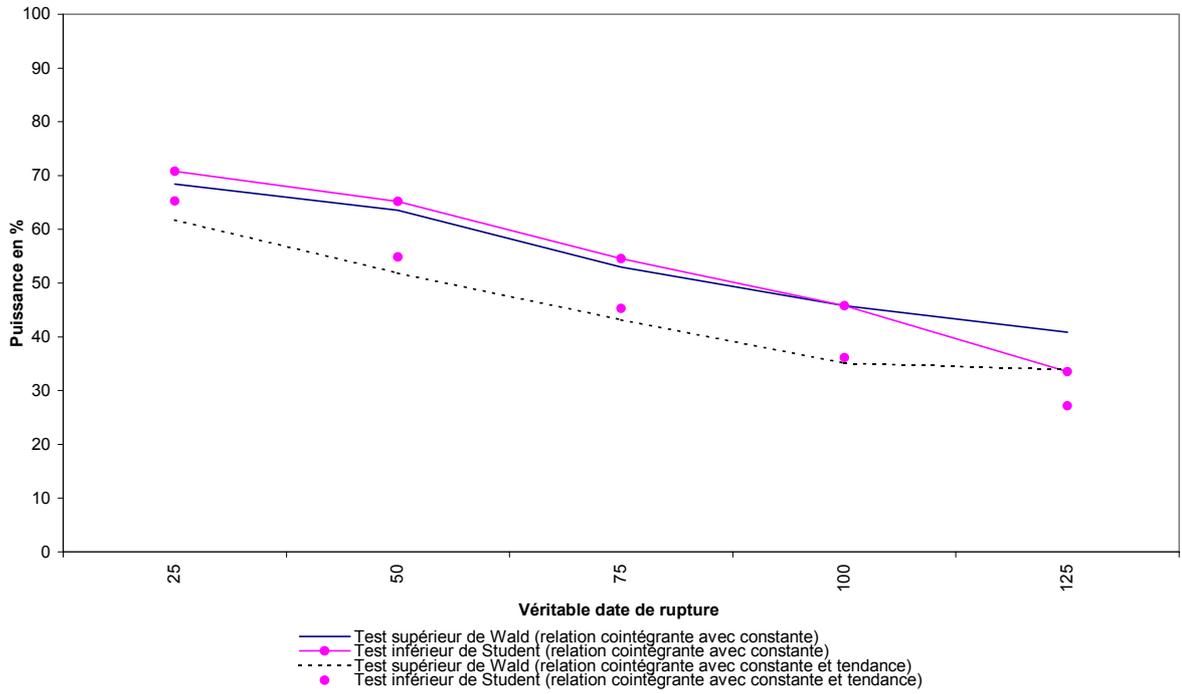
- Cette puissance est inférieure à celle obtenue pour les tests de modification du degré d'intégration ayant pour hypothèse alternative "la série est intégrée puis stationnaire". Nous avons mis en avant une puissance de 83.72% pour le test supérieur de Wald et 88.08% pour le test inférieur de Student ($k = 25$) contre respectivement 51.32% et 45.40% ($k = 125$).

- Au contraire, la qualité de l'estimation de la date de rupture est meilleure quand la rupture a lieu en fin d'échantillon. Pour une relation de cointégration avec constante, le test supérieur de Wald estime 27.60% des bonnes dates de rupture dans le cas d'une rupture à la date 25 contre 67.91% pour la date 125 (respectivement 25.93% contre 62.57% dans le cas d'un test inférieur de Student). Ces pourcentages de bonnes dates estimées sont inférieurs à ceux obtenus dans le cas de test de modification du degré d'intégration de sept points environ.

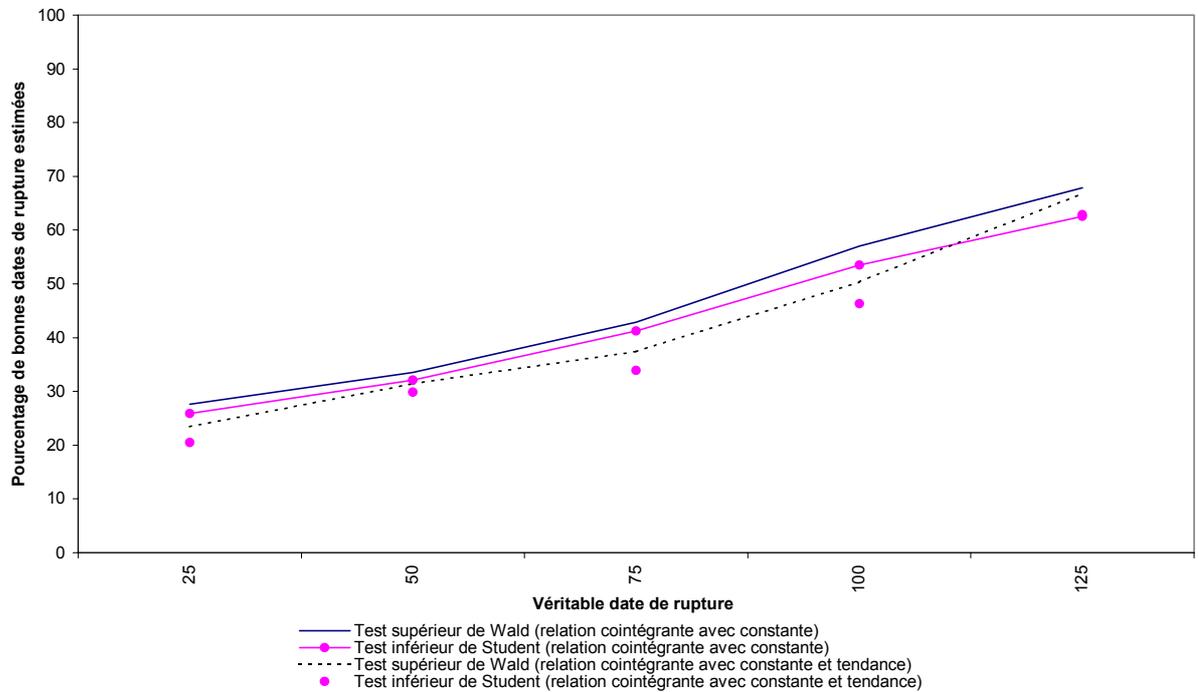
- Il y a relativement peu de différences entre les tests supérieur de Wald et inférieur de Student en termes de puissance et de qualité de l'estimation de la date de rupture.

Les calculs de puissance sont représentés de façon synthétique par les graphiques suivants. Les données chiffrées sont reportées dans l'annexe E.4 à la page 661.

Puissance des tests dans un modèle pas cointégré puis cointégré



Estimation de la date de rupture dans un modèle pas cointégré puis cointégré



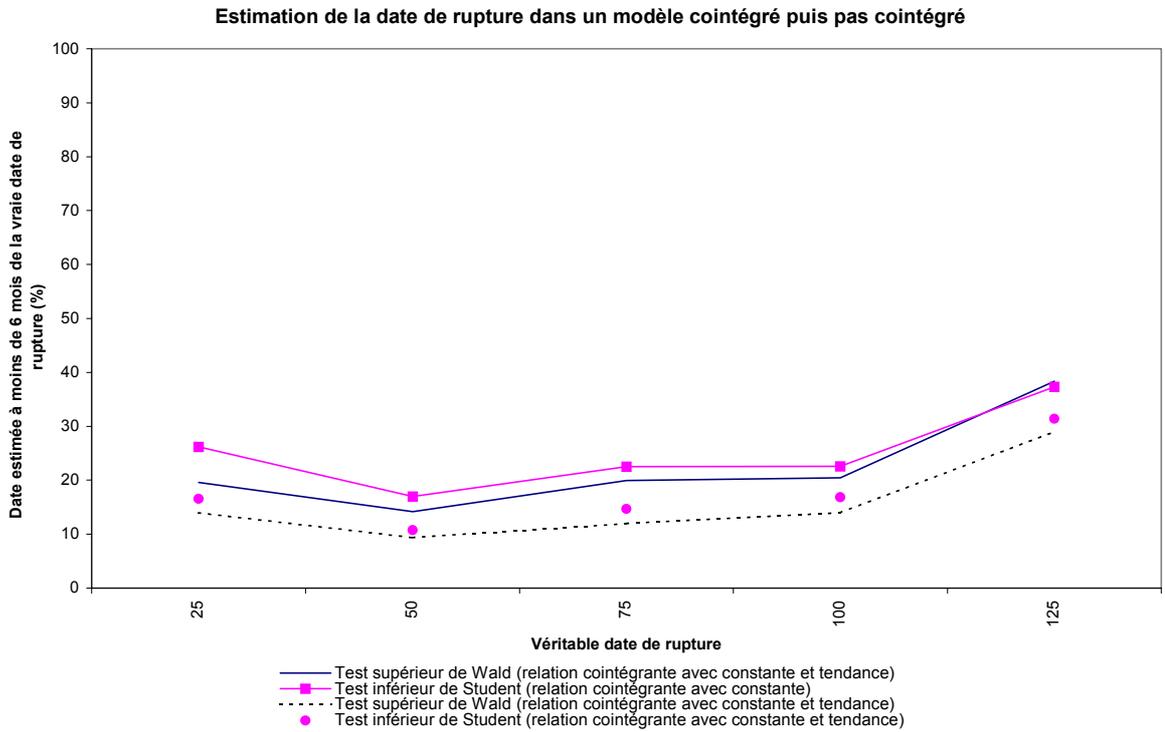
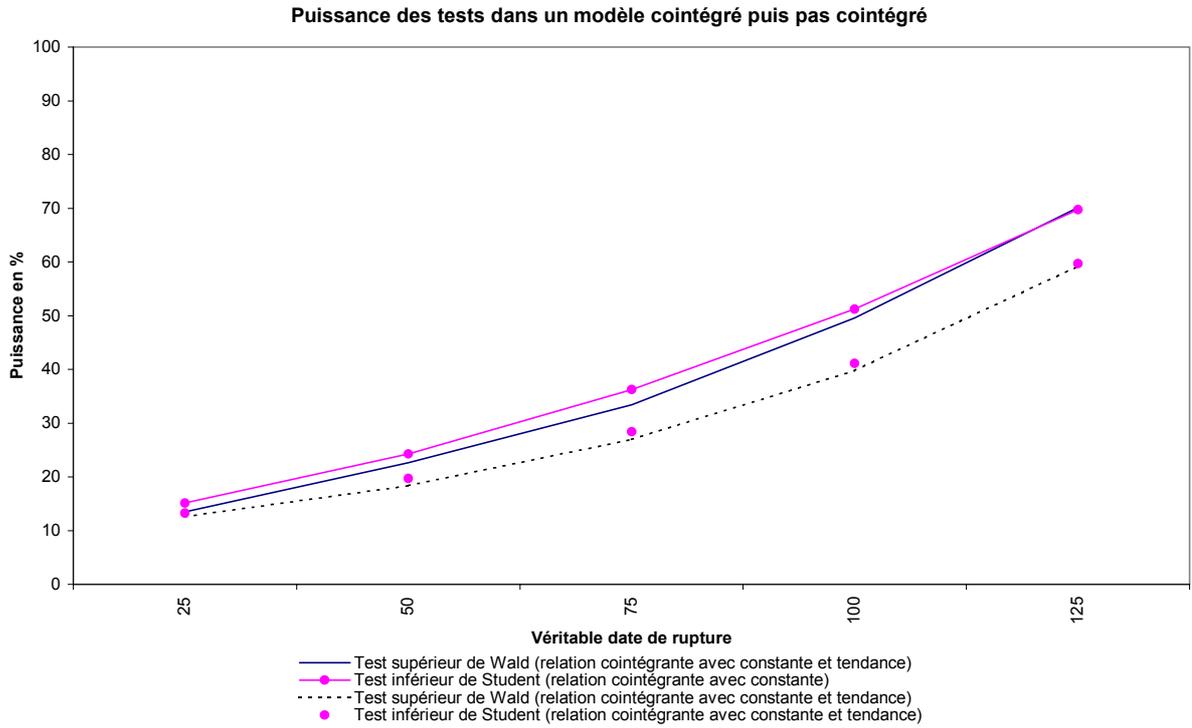
9.4.3.2 Cas d'un résidu de la relation de cointégration éventuelle stationnaire puis intégré : "Cointégration puis absence de cointégration"

Posons désormais $\rho_1 = 0.8, \rho_2 = 1, c_1 = 1, c_2 = 0$ et faisons varier la véritable date de rupture.

Dans le cas d'une série de résidus stationnaire puis intégrée, i.e. dans un modèle cointégré puis non cointégré, les faits marquants suivants peuvent être mis en évidence :

- À l'inverse d'une série de résidu intégrée puis stationnaire (absence de cointégration puis cointégration), la puissance des tests supérieur de Wald et inférieur de Student augmente au fur et à mesure que la rupture est tardive. Ainsi, pour un résidu généré par $\rho_1 = 0.8, \rho_2 = 1, c_1 = 1, c_2 = 0$, et une relation de cointégration avec constante, le test supérieur de Wald a une puissance de 13.48% quand la rupture se produit à la date 25, contre 70.12% pour une rupture à la date 125 (respectivement 15.12% et 69.76% pour le test inférieur de Student). Lorsque la relation de cointégration comporte une tendance, la puissance du test supérieur de Wald descend à 12.56% contre 13.28% pour le test inférieur de Student ($k = 25$), respectivement 59.28% et 59.72% pour $k = 125$.
- L'estimation de la véritable date de rupture est mauvaise quel que soit le test considéré, elle augmente au fur et à mesure que la date de rupture augmente dans le temps. Cet accroissement de la qualité d'estimation quand la rupture est tardive s'explique par le fait que les tests ont tendance, dans le cas d'une relation "cointégrée puis non cointégrée", à estimer la date de rupture par la dernière date de rupture envisageable.
- Puissance et qualité de l'estimation de la date de rupture sont altérées dans le cas "Cointégration puis absence de cointégration" par rapport au cas "Absence de cointégration puis cointégration".
- Ces éléments apparaissent déjà pour les tests de modification du degré d'intégration ayant pour hypothèse alternative la série est "stationnaire puis intégrée".

Les simulations de puissance sont résumées sous forme graphique, les tableaux de chiffres sont repoussés à l'annexe E.4 page 663.



* * *

Ce chapitre avait pour objet de construire des tests de modification du degré de cointégration, sous l'hypothèse nulle d'absence de relation cointégrée sur l'ensemble de la période considérée.

Deux types de tests ont été construits :

- un test ayant pour hypothèse alternative "Absence de cointégration puis cointégration", i.e. création en fin de période d'un équilibre de long terme qui n'existait pas auparavant,
- un test ayant pour hypothèse alternative "Cointégration puis absence de cointégration", i.e. disparition d'un équilibre de long terme qui était valide en début de période.

Pour chacune de ces alternatives, deux tests peuvent être envisagés :

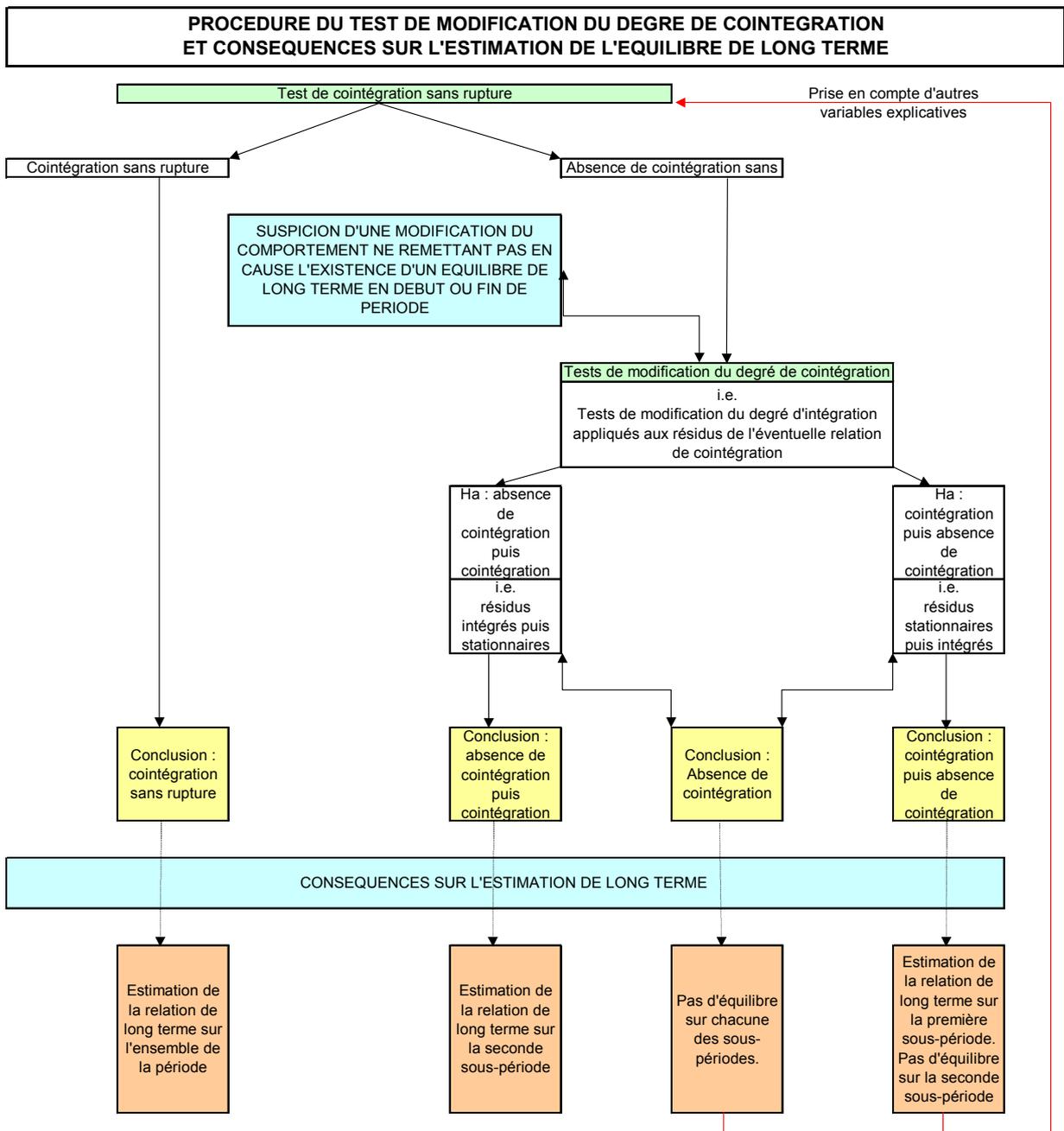
- le test supérieur de Wald, qui teste à la fois la modification de la constante et la modification du degré de cointégration,
- et un test inférieur de Student, qui permet une modification de la constante, mais ne s'intéresse qu'à la modification du degré de cointégration.

Ainsi, si le test supérieur de Wald conclut à une modification du degré de cointégration mais que cette conclusion n'est pas validée par le test inférieur de Student, il est probable que le niveau de l'équilibre a changé mais que cette variation ne constitue pas une modification du degré de cointégration au sens propre.

Pour tester uniquement la modification du degré de cointégration, nous conseillons d'appliquer aux résidus de la relation testée un test inférieur de Student.

Ces tests ont pris en compte la nécessité de corriger les auto-corrélations résiduelles. Les tables de valeurs critiques ont été tabulées.

Nous avons élaboré une procédure de test de modification du degré de cointégration, présentée à la page suivante. Cette procédure est valide quel que soit le type de test choisi (test supérieur de Wald ou test inférieur de Student). Suivant les conclusions du test, la relation de cointégration sera estimée sur tout ou partie de la période considérée.



Les calculs de puissance ont montré les résultats suivants, utiles lors de leurs applications :

- Le test inférieur de Student et le test supérieur de Wald ont des puissances similaires pour détecter la modification du degré de cointégration. Ils conduisent à des dates de rupture estimées assez semblables.
- Le cas "absence de cointégration puis cointégration" est mieux détecté que l'alternative "cointégration puis absence de cointégration". En d'autres termes, les tests sont plus à même de mettre en évidence l'apparition d'un équilibre que de montrer que l'équilibre valide en début de période ne l'est plus en fin de période. L'estimation de la date de rupture est aussi bien meilleure dans le premier cas. Dans le deuxième cas, les tests ont tendance à choisir la dernière date de rupture possible. Un examen judicieux des statistiques pour chaque date de rupture devra être effectué de façon à choisir pour date de rupture un extremum local antérieur.
- De façon générale, ces tests ont tendance à réduire la période non cointégrée : ils ont tendance à s'apercevoir assez tôt du passage à un régime d'équilibre de long terme, ou à se rendre compte tardivement que l'équilibre n'est plus valide.

La procédure présentée dans ce chapitre est utilisée dans le chapitre suivant pour mettre en évidence les équilibres de long terme. Les recommandations que nous avons faites y seront utilisées.

Chapitre 10

Recherche des équilibres de long terme et des dynamiques de court terme

Les tests construits dans le chapitre précédent sont appliqués aux relations entre une série de trafic (type de journée et de sous-réseau donné) et des variables décrivant l'environnement économique en Ile-de-France.

L'application de ces tests est justifiée car la majorité des séries prises en compte sont intégrées à l'ordre 1, sur tout ou partie de la période considérée (janvier 1990 - décembre 2002). En outre, certaines séries présentent une rupture de leur degré d'intégration ; cette rupture s'explique peut-être par une modification de l'environnement économique et des comportements (loisir, achat, travail...) des utilisateurs des réseaux de la RATP.

Dans la première section, est menée une analyse préliminaire de la corrélation et de la causalité entre les variables de trafic et les variables environnementales. Cette approche a pour but de réduire le nombre de variables explicatives, en ne sélectionnant que les plus adéquates.

Les tests de modification du degré de cointégration sont ensuite appliqués aux relations entre variables de trafic et variables décrivant l'environnement économiques francilien (deuxième section), lorsque les tests de cointégration rejettent l'existence d'une relation de cointégration.

Cette étape permet d'estimer les équilibres de long terme, qui se sont formés sur l'ensemble de la période ou une sous-période uniquement, et de les commenter.

Dans la troisième section sont estimées les dynamiques de court terme. Les élasticités à court terme sont comparées à celles de long terme.

10.1 Analyse préliminaire de la corrélation et de la causalité

10.1.1 Analyse des corrélations

La corrélation est une mesure de la relation entre deux ou plusieurs variables. Les coefficients de corrélation sont compris entre -1 (parfaite corrélation négative) et 1 (parfaite corrélation positive). La valeur 0 représente une absence de corrélation (indépendance).

La majorité des séries est intégrée à l'ordre 1 sur l'ensemble de la période ; le risque d'obtenir des corrélations fallacieuses est important. Un moyen d'éviter ce risque de corrélation fallacieuse réside dans le calcul des corrélations de court terme, c'est-à-dire sur données en différence, stationnaires. Ces corrélations, résumées dans le tableau suivant, laissent entrevoir la difficulté à modéliser le trafic à court terme du fait de leur faiblesse en valeur absolue.

Les variables pour lesquelles il existe au moins une sous-période stationnaire sont présentes dans le tableau, mais mises en italique et signalées d'une astérisque.

10.1.1.1 Corrélations à court terme

L'ensemble des corrélations de court terme est particulièrement faible en valeur absolue. Cette constatation implique qu'une modélisation ne prenant en compte que les variables en différence première serait très insatisfaisante.

CHAPITRE 10. RECHERCHE DES ÉQUILIBRES DE LONG TERME ET DES
DYNAMIQUES DE COURT TERME

	$\Delta \ln(\text{jomsa})$	$\Delta \ln(\text{jorersa})^*$	$\Delta \ln(\text{joparsa})$	$\Delta \ln(\text{jobansa})$
$\Delta \ln(\text{jomsa})$	1	-	-	-
$\Delta \ln(\text{jorersa})^*$	0.64	1	-	-
$\Delta \ln(\text{joparsa})$	0.30	0.16	1	-
$\Delta \ln(\text{jobansa})$	0.45	0.27	0.59	1
$\Delta \ln(\text{carbidf})^*$	0.00	-0.02	0.10	-0.02
$\Delta \ln(\text{circ75})$	-0.01	0.14	-0.12	0.01
$\Delta \ln(\text{circagar})$	0.05	0.06	0.20	0.06
Δcompet	0.06	0.09	0.15	-0.03
$\Delta \ln(\text{speed75})^*$	-0.36	-0.01	-0.39	-0.26
$\Delta \ln(\text{defm75})^*$	-0.04	-0.13	0.06	0.01
$\Delta \ln(\text{defmidf})^*$	-0.06	-0.19	0.08	-0.03
$\Delta \text{gdmagrps}^*$	0.19	0.15	0.00	-0.02
$\Delta \text{txhot75}$	0.21	0.31	-0.21	-0.07
$\Delta \text{txhotidf}^*$	0.31	0.36	-0.17	-0.02
$\Delta \ln(\text{totgl})^*$	-0.35	0.04	-0.53	-0.31

A court terme, les seules variables qui apparaissent jouer un rôle sur les variations des trafics du jour ouvrable sur chacun des sous-réseaux sont les variations de trafic sur les autres sous-réseaux. En particulier, il existe des liens entre les trafics du jour ouvrable sur le RER et sur le métro, ainsi qu'entre les bus de banlieue et bus de Paris, ou encore une évolution similaire des trafics du jour ouvrable sur le métro et sur le bus de banlieue. Ces corrélations indiquent les effets de correspondance entre les sous modes. Dans une mesure très amoindrie, la vitesse de circulation dans Paris ou le taux d'occupation des hôtels ont un léger impact à court terme.

L'existence éventuelle de relations de cointégration, et leur insertion le cas échéant dans un modèle à correction d'erreur, permettra d'améliorer l'estimation du modèle à court terme.

10.1.1.2 Corrélations à long terme

Il est possible de représenter les corrélations de long terme, c'est-à-dire sur variables en niveau. Le tableau suivant n'a qu'un but indicatif sur les variables économiques pouvant jouer un rôle à long terme. L'existence d'une véritable relation entre deux variables corrélées n'est validée qu'ultérieurement, avec les tests de cointégration.

	$ln(jomsa)$	$ln(jorersa)$	$ln(joparsa)$	$ln(jobansa)$
$ln(jomsa)$	1	-	-	-
$ln(jorersa)^*$	0.73	1	-	-
$ln(joparsa)$	0.06	0.50	1	-
$ln(jobansa)$	0.35	0.83	0.81	1
$ln(carbidf)^*$	-0.63	-0.42	-0.03	-0.20
$ln(circ75)$	-0.38	-0.76	-0.70	-0.90
$ln(circagar)$	0.16	0.68	0.75	0.87
compet	0.13	-0.49	-0.76	-0.82
$ln(speed75)^*$	-0.39	-0.63	-0.58	-0.73
$ln(defm75)^*$	-0.58	-0.83	-0.40	-0.71
$ln(defmidf)^*$	-0.76	-0.46	0.24	-0.07
$gdmagrps^*$	0.71	0.72	0.28	0.53
txhot75	0.40	0.52	0.38	0.48
$txhotidf^*$	0.47	0.58	0.39	0.52
$ln(totgl)^*$	0.43	0.34	-0.15	0.10

L'analyse des corrélations fait apparaître un lien entre les évolutions des différents trafics du jour ouvrable : les évolutions des trafics sur le métro et sur le RER sont liées de façon positive, de même que les trafics sur le RER et le bus de banlieue, ou les trafics bus de banlieue/bus de Paris. Le signe positif de ces corrélations signale la complémentarité des sous-réseaux de la RATP, et non une concurrence. L'analyse de ces corrélations précise certains aspects de cette complémentarité : s'il est relativement fréquent de combiner RER et métro, bus de banlieue et bus de Paris, ou RER et bus, il est moins courant de réaliser une correspondance entre le métro

et le bus de Paris.

La concurrence s'exerce face à l'automobile, comme l'indiquent les corrélations négatives des différents trafics avec les indicateurs de consommation de carburant en Ile-de-France, de circulation dans Paris, ou de vitesse de circulation dans Paris. Cette concurrence vise principalement le RER, et les bus de banlieue et de Paris. Le signe positif de la corrélation avec le volume de circulation dans la zone agglomérée sur les routes nationales et les autoroutes s'explique, peut-être, par une évolution similaire de la mobilité sur les transports collectifs et individuels, au cours de la période considérée. Enfin, il existe un impact du rapport entre le prix du ticket T et le prix du super sur le réseau des bus et, dans une moindre mesure, sur le trafic du jour ouvrable du RER.

La situation économique en Ile-de-France est elle aussi à mettre en lien avec les trafics du jour ouvrable sur les différents réseaux. Un accroissement du nombre de demandeurs d'emplois en fin de mois se traduit par une diminution des volumes de trafic. Cet effet peut résulter de deux causes : la diminution des trajets ayant pour motif le travail, mais aussi la diminution des parcours liés aux loisirs suite à une baisse de revenu, réelle ou redoutée. La consommation en Ile-de-France, prise en compte dans les modélisations par le volume des ventes dans les grands magasins en région parisienne (la majorité se trouvant à Paris), est en outre fortement liée au volume de trafic sur le réseau ferré. De même, le trafic de la RATP s'explique par les fluctuations de l'activité touristique (tourisme véritable ou tourisme "d'affaire"), introduite dans les modèles par le biais du taux d'occupation des hôtels à Paris ou dans l'ensemble de l'Ile-de-France.

Si l'analyse des corrélations est nécessaire pour appréhender les liens entre les trafics par sous-réseaux et les différentes variables économiques relatives à l'Ile-de-France, celle-ci n'est pas suffisante. De fortes corrélations peuvent être fallacieuses, i.e. ne reposer sur aucun fondement économique. L'analyse des corrélations doit s'accompagner d'une analyse de la causalité.

10.1.2 Analyse de la causalité au sens de Granger

Pour analyser la causalité, l'approche de Granger [62] est utilisée. " x cause y " si la valeur de y à une date t est mieux expliquée lorsque les valeurs passées de y et de x sont prises en compte, au lieu du seul passé de y . Dans ce cas là, x aide à prévoir la valeur de y , il apporte de

l'information. L'analyse de la causalité est expliquée plus en détail dans l'annexe F page 665.

Le test de causalité à la Granger réalisé inclut $p = 3$ retards. Il porte sur la causalité de court terme : le test n'étant valide que pour des données stationnaires, les données sont en différence première. L'hypothèse nulle "x ne cause pas y" est rejetée si la statistique de test excède la valeur critique à 5% d'un $F(p, T - 2p - 1) = F(3, 148)$, soit 2.67.

x	$\Delta \ln(\text{jomsa})$	$\Delta \ln(\text{jorersa})^*$	$\Delta \ln(\text{joparsa})$	$\Delta \ln(\text{jobansa})$
$\Delta \ln(\text{jomsa})$	-	0.89	0.09	0.12
$\Delta \ln(\text{jorersa})^*$	0.94	-	0.26	0.54
$\Delta \ln(\text{joparsa})$	1.94	2.43	-	3.31
$\Delta \ln(\text{jobansa})$	0.32	0.91	0.89	-
$\Delta \ln(\text{carbidf})^*$	3.69	1.22	0.53	1.71
$\Delta \ln(\text{circ75})$	1.24	0.90	0.25	0.38
$\Delta \ln(\text{circagar})$	1.75	0.49	0.20	0.24
Δcompet	1.58	2.18	2.69	2.02
$\Delta \ln(\text{speed75})^*$	4.00	1.50	0.06	0.66
$\Delta \ln(\text{defm75})^*$	3.86	5.69	0.70	1.11
$\Delta \ln(\text{defmidf})^*$	2.59	3.33	0.37	0.91
$\Delta \text{gdmagrps}^*$	0.72	0.19	0.92	0.77
$\Delta \text{txhot75}$	1.52	1.30	0.45	2.05
$\Delta \text{txhotidf}^*$	0.22	0.80	0.25	0.65
$\Delta \ln(\text{totgl})^*$	5.25	1.63	0.47	3.40

Les enseignements de cette analyse de la causalité à court terme sont les suivants :

- A court terme, les variables qui "causent" le trafic du jour ouvrable sur le métro sont la vitesse de circulation dans Paris, le nombre de demandes d'emplois en fin de mois, le volume de carburant vendu en Ile-de-France ou encore le nombre de voyageurs grandes lignes passant par une gare de Paris.
- Les variations de court terme du trafic du jour ouvrable sur le RER sont causées par le trafic du jour ouvrable sur le bus de Paris, l'évolution du nombre de demandes d'emplois

en fin de mois, et dans une moindre mesure par la variation du rapport entre prix du ticket T et prix du super.

- Pour le trafic du jour ouvrable sur les bus de Paris, les variations de court terme sont imputables à la variation du rapport entre prix du ticket T et prix du super uniquement.
- Enfin, le trafic du jour ouvrable sur le bus de Paris, le nombre de voyageurs grandes lignes passant par Paris, et dans une moindre mesure le rapport des prix du ticket T et du super ou encore l'évolution du taux d'occupation des hôtels, causent au sens de Granger le trafic du jour ouvrable sur le bus de banlieue.

Cette analyse de la causalité à court terme confirme en partie les résultats issus de l'analyse des corrélations. Elle permet de mettre en avant des facteurs explicatifs supplémentaires, notamment le rôle du rapport de prix entre le ticket T et le super, qui peut être considéré comme un indice de compétitivité des transports publics par rapport à la concurrence de l'automobile. Néanmoins, elle ne remet pas en question le sentiment qu'une modélisation des séries en différence première serait très insatisfaisante, et que la prise en compte de relations de cointégration, s'il en existe, ne peut qu'améliorer la qualité de l'estimation.

Les causalités de long terme, i.e. sur séries en niveau, ne peuvent être mises en lumière avec ce type de test, valide uniquement dans le cas de données stationnaires. Selon Granger [64], ces causalités entre variables pourront être établies lors de l'analyse de la cointégration : l'existence d'une relation de cointégration entre deux variables implique l'existence d'une causalité au sens de Granger dans au moins une direction (voir annexe F.3 page 668).

10.2 Application des tests de changement du degré de cointégration et modèles de trafic à long terme

Pour chaque trafic ont été recherchées les équations de cointégration sans rupture avec tendance et/ou constante en fonction des variables économiques suivantes : le nombre de demandes d'emplois en fin de mois en Ile-de-France ("*defmidf*"), l'indice de volume des ventes dans les grands magasins de la région parisienne ("*gdmagrp*") et le taux d'occupation des hôtels en Ile-de-France ("*txhotidf*"). Un graphique représente les résidus de cette éventuelle relation de

10.2. APPLICATION DES TESTS DE CHANGEMENT DU DEGRÉ DE COINTÉGRATION ET MODÈLES DE TRAFIC À LONG TERME

cointégration, tandis que les conclusions du test ADF de cointégration (test "CADF") sont détaillées dans un tableau ¹.

En cas de non-existence de relation de cointégration sans rupture, nous avons appliqué les tests de modification du degré de cointégration élaborés et évalués dans les chapitres précédents, qui ont pour hypothèse alternative l'absence de cointégration sur la période entière. Ces tests nous permettent de mettre en évidence que, s'il n'existe pas de relation de cointégration sur l'ensemble de la période, il existe une sous-période sur laquelle la relation de long terme était valide. En d'autres termes, l'absence de cointégration n'est vérifiée que sur une sous-période.

La date de rupture du degré de cointégration est évaluée, et conduit à l'estimation sur les deux sous-périodes dégagées. Il est alors possible d'affiner l'estimation sur la sous-période cointégrée et de mettre en évidence un nouvel équilibre sur la sous-période non cointégrée.

L'analyse du premier type de trafic, à savoir le trafic sur le jour ouvrable du métro, est présentée de façon détaillée de manière à préciser la lecture des tableaux et l'interprétation des graphiques. Les commentaires seront plus succincts lors de l'analyse des autres trafics.

10.2.1 Trafic du jour ouvrable sur le métro

10.2.1.1 Absence de modélisation sans rupture

Quelle que soit la modélisation choisie (tendance et/ou constante, prise en compte de tout ou partie des variables économiques), aucune relation relative au trafic du jour ouvrable sur le métro n'est cointégrée au seuil de 10%.

¹Le même travail a été réalisé en prenant en compte l'ensemble des séries économiques présentées en première partie et rappelées lors de l'analyse des corrélations et de la causalité.

Pour la majorité des trafics, les modélisations en fonction des variables *defmidf*, *txhotidf*, et/ou *gdmagr* ne permettent pas de mettre en évidence de relations de cointégration sans rupture. L'augmentation du nombre de variables explicatives ne conduit pas non plus à l'obtention de relations de cointégration sans rupture cohérentes avec la théorie économique, malgré un accroissement exponentiel du nombre de modèles testés. En outre, les conclusions des tests de modification du degré de cointégration étaient très proches des résultats présentés ci-après.

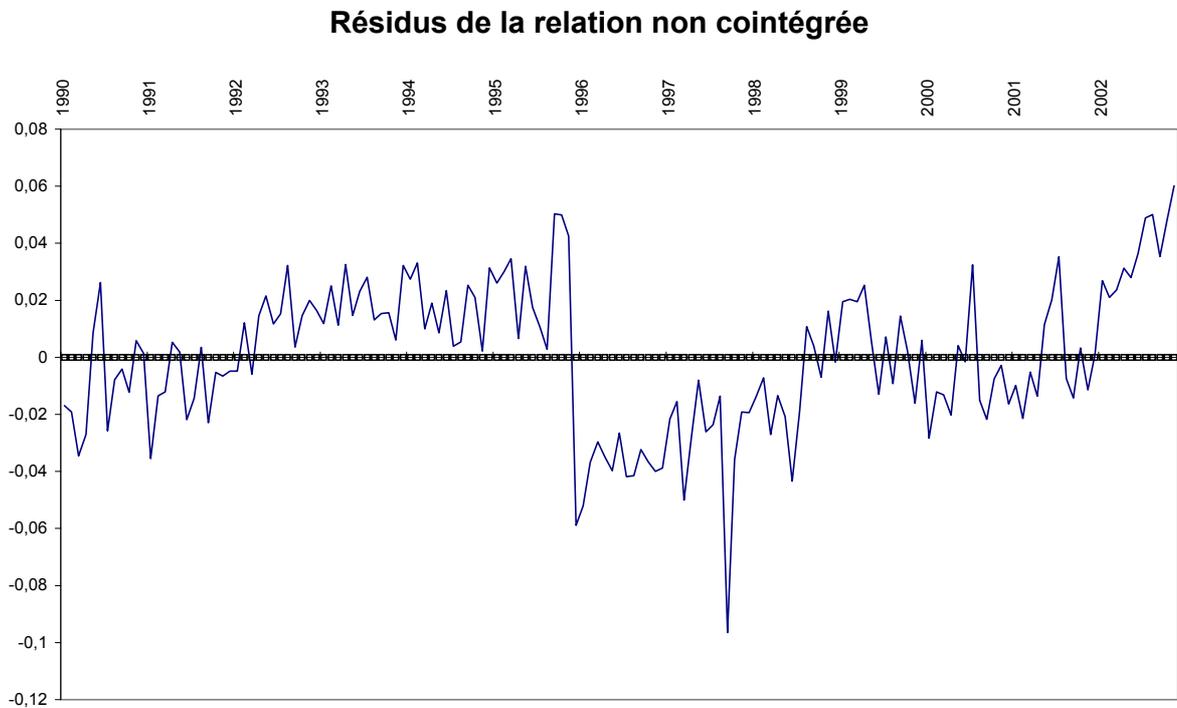
La seule exception notable provient de la modélisation des trafics pour lesquels il existe une relation de cointégration sans rupture (trafics du jour ouvrable sur le bus de Paris et sur les bus de banlieue).

CHAPITRE 10. RECHERCHE DES ÉQUILIBRES DE LONG TERME ET DES DYNAMIQUES DE COURT TERME

Considérons la relation suivante :

$$\begin{aligned} \ln(jomsa) = & - 0.1563 \ln(defmidf) + 0.0029 \ gdmagrps \\ & (-4.9244) \qquad \qquad \qquad (-4.1465) \\ & - 0.0374 \ txhotidf + 17.0043 \\ & (-0.6158) \qquad \qquad \qquad (37.2486) \end{aligned}$$

Les nombres entre parenthèses inscrits en dessous des paramètres estimés correspondent aux statistiques de Student. Les résidus de cette relation sont représentés dans le graphique suivant :



Cette relation n'est pas une relation de cointégration, les résidus ne sont pas stationnaires. Le test de cointégration appliqué correspond au test ADF de cointégration ayant pour statistique t_{CADF} , réalisé au seuil de 10%. Si la statistique est inférieure ou égale à la valeur critique $VC_{10\%}$, alors la relation est cointégrée sur la période considérée ("CO (0)"). Le paramètre auto-régressif

10.2. APPLICATION DES TESTS DE CHANGEMENT DU DEGRÉ DE COINTÉGRATION ET MODÈLES DE TRAFIC À LONG TERME

ρ des résidus de la relation de cointégration est significativement inférieur à 1.

Test de cointégration sans rupture et sans tendance

ln(<i>jomsa</i>)	Retards	Rho	Statistique	VC 10%	Conclusion
<i>t</i> _{CADF}	2	0.8033	-2.8277	-3.9243	Absence de cointégration

En l'absence de relation de cointégration sur l'ensemble de la période, nos tests de modification du degré de cointégration peuvent être appliqués à la relation considérée.

10.2.1.2 Application des tests de modification du degré de cointégration au cours de la période considérée au trafic du jour ouvrable sur le métro

Les tests de modification du degré de cointégration sont appliqués à la relation non cointégrée :

Tests de modification du degré de cointégration

ln(<i>jomsa</i>)	Statistique	VC 10%	VC 20%	Date	Conclusion
<i>Absence de cointégration puis cointégration</i>					
Wald supérieur	11.9008	18.380	15.796	novembre 1995	CI(1)-CI(1)
Student inférieur	-3.3819	-4.039	-3.724	décembre 1995	CI(1)-CI(1)
<i>Cointégration puis absence de cointégration</i>					
Wald supérieur	18.6887	16.190	13.802	avril 2001	CI(0)-CI(1)
Student inférieur	-4.1942	-3.887	-3.564	mai 2001	CI(0)-CI(1)

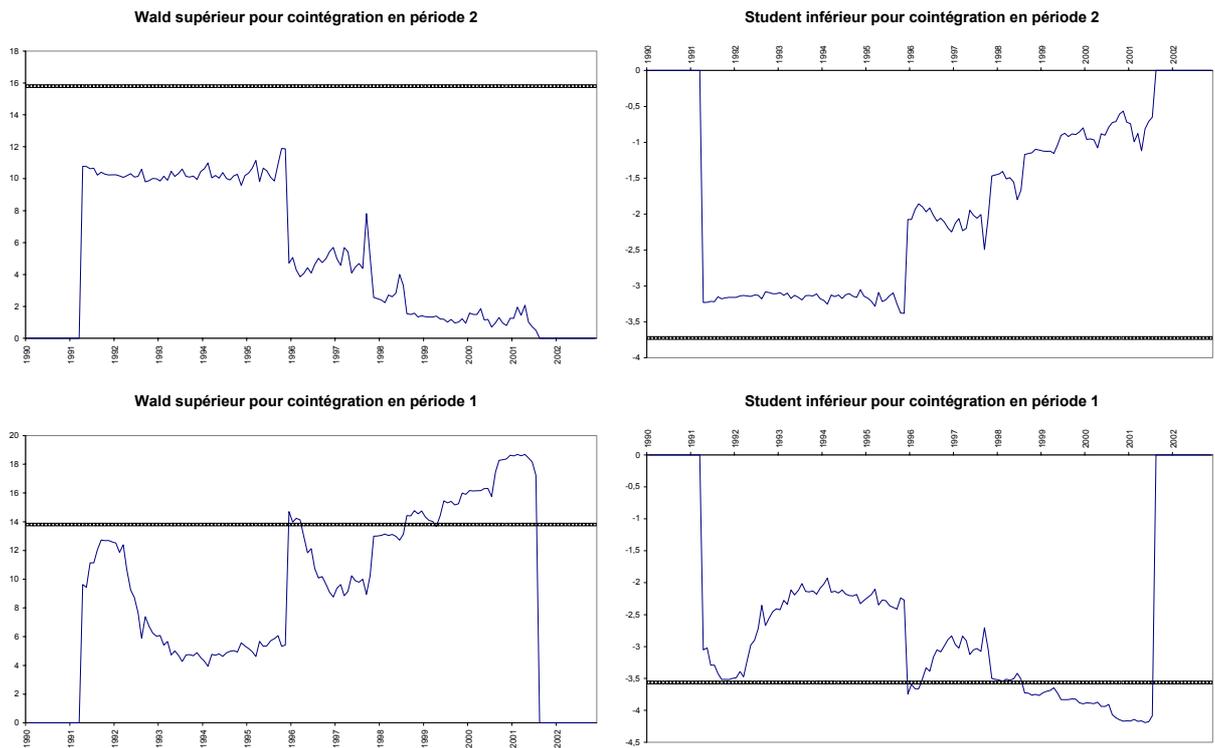
Ils concluent que la relation non cointégrée correspond à un équilibre sur une première sous-période. Mais la relation ne permet pas d'expliquer de façon satisfaisante ce qui se passe en seconde sous-période.

Les calculs de puissance réalisés dans le chapitre précédent ont montré que l'estimation de la date de rupture était peu satisfaisante lorsque le test rejetait l'hypothèse nulle d'absence de cointégration sur l'ensemble de la période au profit de l'alternative "absence de cointégration sur une deuxième sous-période". Cette faiblesse du test amène à une analyse plus fine des

CHAPITRE 10. RECHERCHE DES ÉQUILIBRES DE LONG TERME ET DES DYNAMIQUES DE COURT TERME

statistiques, analyse qui peut être réalisée graphiquement.

Nous avons représenté graphiquement les statistiques des tests de Wald et de Student pour toutes les dates de rupture éventuelles (i.e. en écartant 10% des observations au début de l'échantillon et 10% des observations à la fin de l'échantillon), d'abord dans le cas de l'hypothèse alternative "la relation n'est pas cointégrée puis cointégrée", puis pour "la relation est cointégrée puis non cointégrée". La droite en gros pointillés correspond à la valeur critique associée au seuil de 20%.



Les graphiques associés aux tests de modification du degré de cointégration (hypothèse alternative : "relation cointégrée puis non cointégrée") mettent en évidence un optimum local en **décembre 1995**. Cette date de rupture est la date de rupture retenue dans la mesure où les calculs de puissance ont montré que ces tests ont tendance à s'apercevoir trop tard du changement de régime. En outre, lors de l'examen des relations non cointégrées au seuil de 10% entre le trafic du jour ouvrable sur le métro et la variable de ventes dans les grands magasins ou le taux d'occupation des hôtels en Ile-de-France, les tests de modification du degré de cointégration au seuil de 5% indiquent le passage d'un régime cointégré à un régime non cointégré en décembre

10.2. APPLICATION DES TESTS DE CHANGEMENT DU DEGRÉ DE COINTÉGRATION ET MODÈLES DE TRAFIC À LONG TERME

1995.

S'il n'existe pas de relation de cointégration sans rupture au seuil de 10%, nos tests de modification du degré de cointégration indiquent que l'absence d'équilibre n'est en réalité vérifiée que sur la deuxième sous-période. Ceci nous amène à considérer un équilibre avec rupture en décembre 1995.

10.2.1.3 Modélisation du trafic du jour ouvrable sur le métro pour chacune des sous-périodes

La modélisation avec rupture est obtenue en scindant l'échantillon en deux sous-périodes à la date de rupture, ici en considérant la première sous-période "janvier 1990 - décembre 1995", et la deuxième sous-période "janvier 1995 - décembre 2002".

Pour chaque sous-période, l'équation est estimée à l'aide des moindres carrés "totalement modifiés", et peut comprendre une tendance et/ou une constante et tout ou partie des variables explicatives. La modélisation choisie répond aux critères suivants :

- les élasticités sont significativement différentes de 0,
- la modélisation possède un critère BIC minimal,
- la relation est cointégrée au seuil de 10%.

ln(<i>jomsa</i>)	Valeurs des coefficients					Test CADF			
	ln(defmidf)	gdmagrp	txhotidf	c	t	ρ	t_{CADF}	VC _{5%}	Concl.
Avant	-0.0874 (-5.5117)	0.0007 (2.9699)	0.0751 (2.0109)	16.2392 (69.6877)	n.s.	0.2631	-5.1588	-4.2249	CO(0)
Après	-0.1131 (-4.0746)	n.s.	0.2099 (3.1029)	16.4762 (43.6521)	0.0014 (9.9034)	0.4861	-4.2500	-4.2429	CO(0)

Entre les deux sous-périodes, l'élasticité du trafic du jour ouvrable sur le métro au nombre de demandes d'emplois s'est accrue en valeur absolue, tandis que l'élasticité par rapport à la consommation dans les grands magasins a chuté. Cette évolution traduit une spécificité française du comportement de consommation. En période de chômage, la consommation des ménages

correspond au principal facteur de la faible croissance économique. En période de reprise économique, une partie des individus qui recommencent à travailler réemprunte les transports collectifs pour se rendre sur leur lieu de travail, mais aussi pour des activités de loisirs telles que le cinéma, le lèche-vitrine, les visites aux amis...

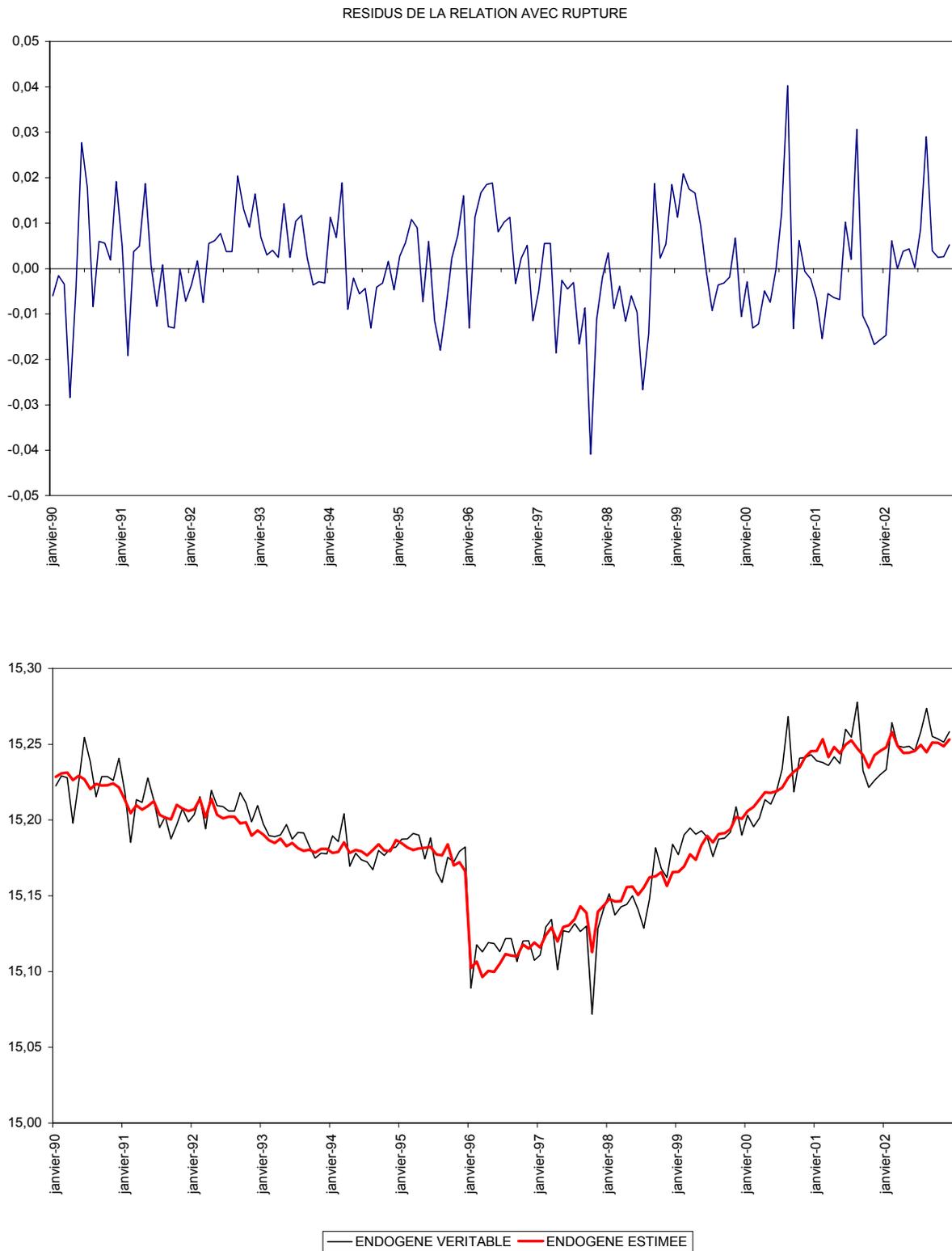
L'élasticité du trafic du jour ouvrable sur le métro au taux d'occupation dans les hôtels a fortement augmenté entre les deux sous-périodes. La deuxième sous-période se caractérise par une forte augmentation du tourisme dans la région Ile-de-France : la France était la 4^{ème} destination touristique internationale en 1990, elle est aujourd'hui la première destination mondiale.

La présence d'une tendance dans la relation cointégrée de seconde-période, alors que la relation sur la période entière n'en comporte pas, explique l'absence de relation de cointégration sans rupture sur l'ensemble de la période. La tendance à l'augmentation du trafic du jour ouvrable sur le métro à partir de la fin des grèves de l'hiver 1995, en dehors du retour des touristes ou de la baisse du chômage, provient des facteurs suivants :

- la prise en compte par la RATP de la nécessité croissante de mesurer et d'améliorer la qualité de service sur ces réseaux,
- l'amélioration de l'offre de transport par la RATP par de successifs renforcements de l'offre durant les flancs de pointe puis le week-end,
- la mise en place de l'alarme sociale en 1996 qui a fortement réduit les mouvements de grèves,
- la création de titres de transport (Imagine R) ou la mise en valeur de titres existants (carte Intégrale en particulier) qui permettent une utilisation illimitée des sous-réseaux RATP pendant une année, et fidélisent la clientèle,
- ou encore l'effet de vitrine provoqué par la création de la ligne 14 du métro.

Les résidus de cette relation cointégrée par morceaux sont représentés dans le graphique suivant :

10.2. APPLICATION DES TESTS DE CHANGEMENT DU DEGRÉ DE COINTÉGRATION ET MODÈLES DE TRAFIC À LONG TERME



10.2.2 Trafic du jour ouvrable sur le RER

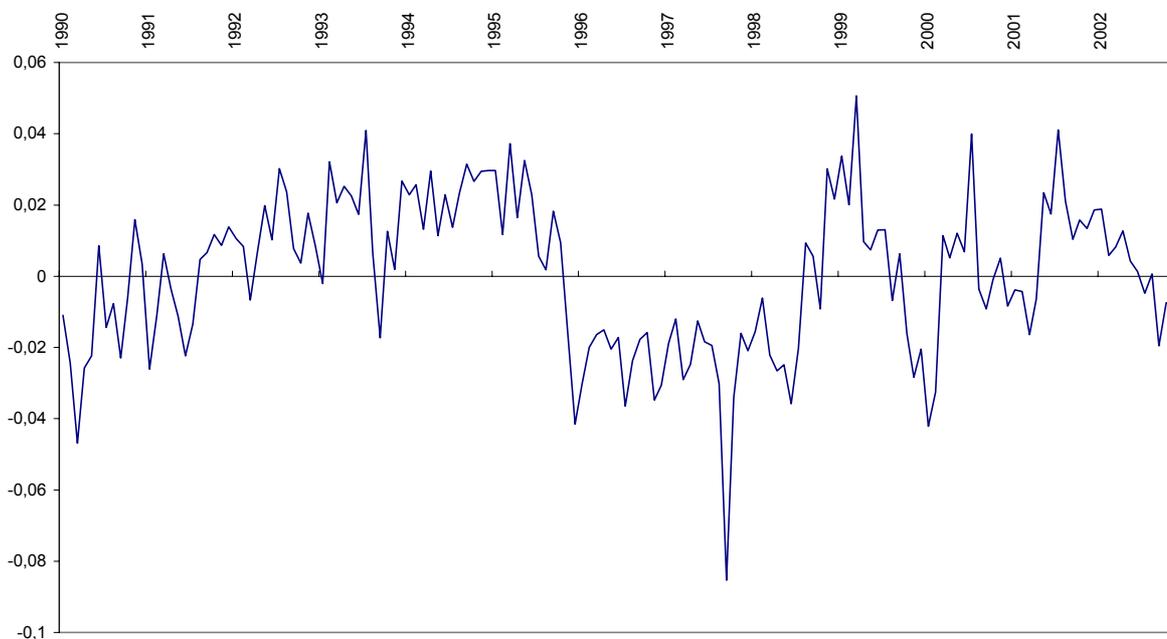
Quelle que soit la modélisation choisie (tendance et/ou constante, prise en compte de tout ou partie des variables économiques), aucune relation relative au trafic du jour ouvrable sur le RER n'est cointégrée au seuil de 10%.

10.2.2.1 Absence de relation sans rupture

La relation suivante possède le critère BIC minimal parmi toutes les relations non cointégrées :

$$\begin{aligned} \ln(joversa) = & - 0.1743 \ln(defmidf) + 0.0015 gdmagr p \\ & (-5.1061) \qquad \qquad \qquad (3.0664) \\ & + 16.1743 + 0.0008 t \\ & (33.5546) \qquad (12.8680) \end{aligned}$$

Résidus de la relation non cointégrée



10.2. APPLICATION DES TESTS DE CHANGEMENT DU DEGRÉ DE COINTÉGRATION ET MODÈLES DE TRAFIC À LONG TERME

Cette relation n'est pas une relation de cointégration :

Test de cointégration sans rupture et sans tendance

$\ln(jorersa)$	Retards	Rho	Statistique	VC 5%	VC 10%	Conclusion
t_{CADF}	2	0.7816	-3.3864	-4.2429	-3.9267	Absence de cointégration

10.2.2.2 Application des tests de modification du degré de cointégration au cours de la période considérée au trafic du jour ouvrable sur le RER

Les tests de modification du degré de cointégration sont appliqués à cette relation non cointégrée :

Tests de modification du degré de cointégration

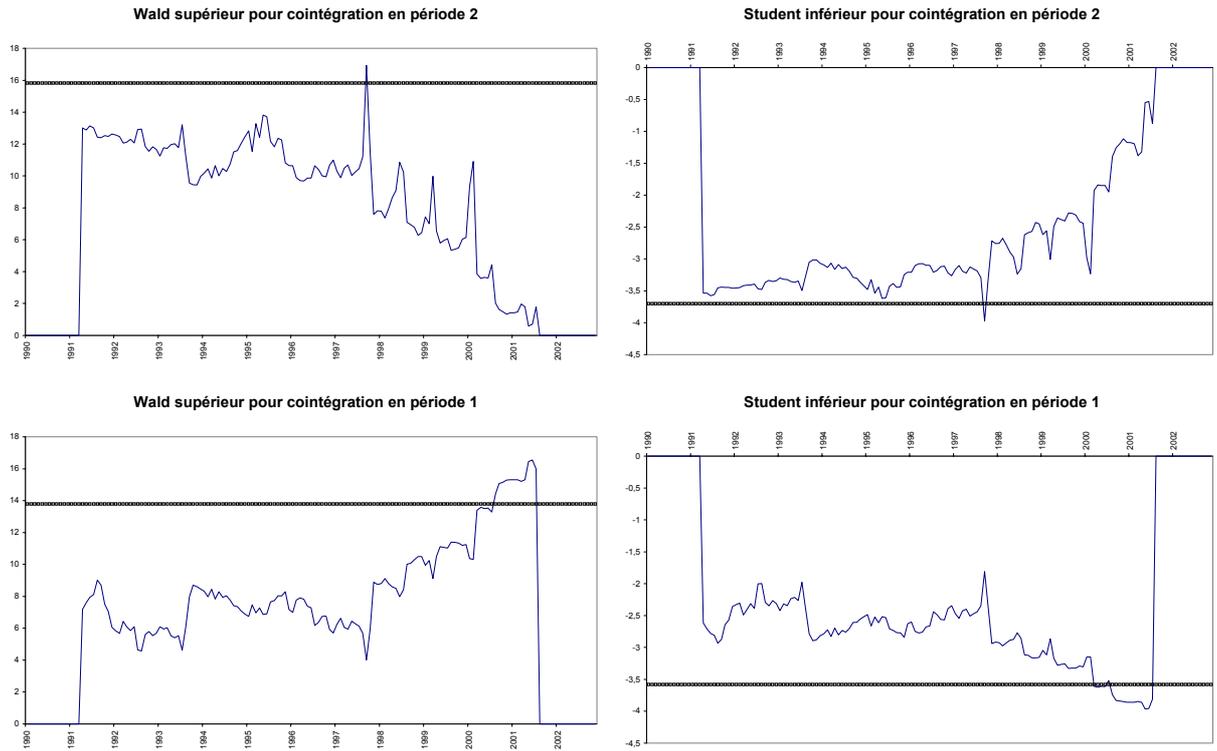
$\ln(jorersa)$	Statistique	VC 10%	VC 20%	Date	Conclusion
Absence de cointégration puis cointégration					
Wald supérieur	16.9352	18.428	15.827	octobre 1997	CI(1)-CI(0)
Student inférieur	-3.9744	-4.027	-3.699	octobre 1997	CI(1)-CI(0)
Cointégration puis absence de cointégration					
Wald supérieur	16.5485	16.347	13.794	juin 2001	CI(0)-CI(1)
Student inférieur	-3.9651	-3.909	-3.580	mai 2001	CI(0)-CI(1)

Les tests de modification du degré de cointégration concluent à la fois que la relation est "cointégrée puis non cointégrée" et "non cointégrée puis cointégrée".

La relation est cointégrée par morceaux, avec une date de rupture probable en **octobre 1997**.

Cette date correspond à l'extrémum des statistiques des tests ayant pour alternative "absence puis présence de cointégration", alors que les tests ayant l'hypothèse alternative inverse ont tendance à estimer une date de rupture postérieure à la véritable date de rupture.

CHAPITRE 10. RECHERCHE DES ÉQUILIBRES DE LONG TERME ET DES DYNAMIQUES DE COURT TERME



10.2.2.3 Modélisation avec rupture du trafic du jour ouvrable sur le RER

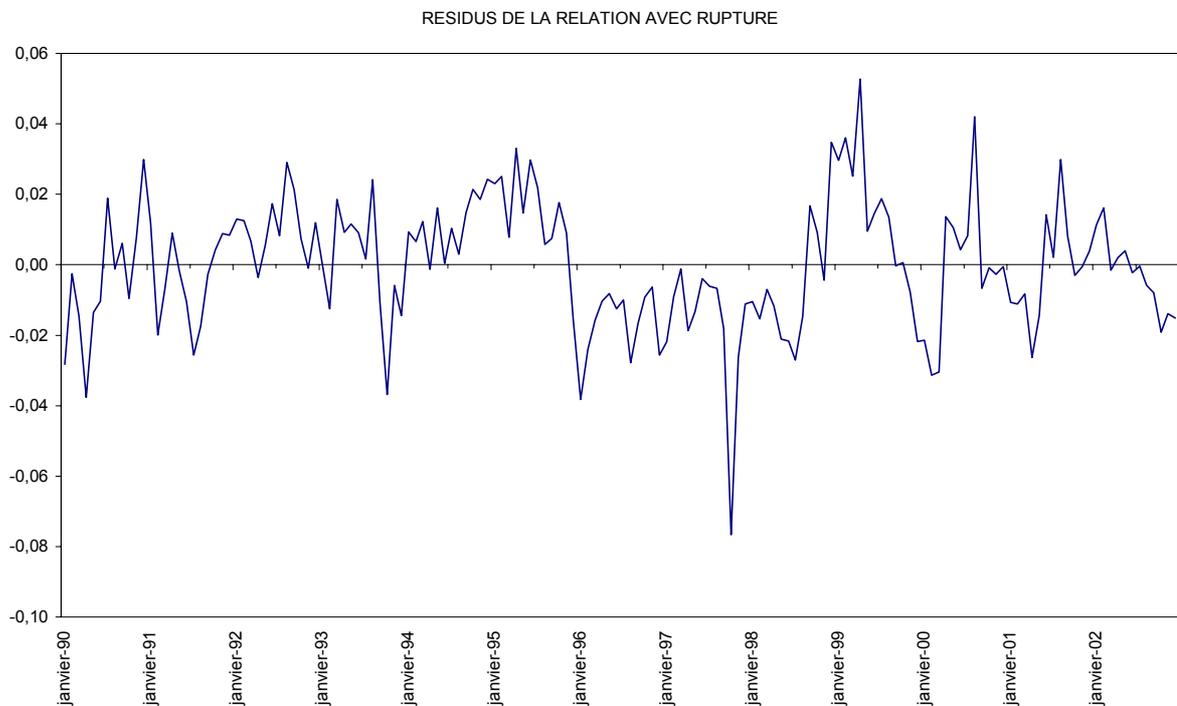
$\ln(jorersa)$	Valeurs des coefficients				Test CADF			
	$\ln(\text{defmidf})$	$gdmagrp$	c	t	ρ	t_{CADF}	$VC_{10\%}$	Concl.
Oct. 1997								
Avant	n.s.	0.0012 (3.3106)	13.9374 (411.7491)	n.s.	0.6301	-3.8646	-3.0890	CO(0)
Après	-0.2592 (-5.8355)	n.s.	17.5236 (29.6294)	0.0010 (3.9062)	0.6222	-3.6139	-3.6062	CO(0)

La rupture de la modélisation du trafic du jour ouvrable sur le RER se justifie par des facteurs similaires à ceux observés sur le métro. Le maintien de la consommation pendant les périodes de faible croissance économique puis la reprise des activités de loisirs en lien avec la diminution du chômage expliquent l'évolution des élasticités du trafic aux variables d'environnement.

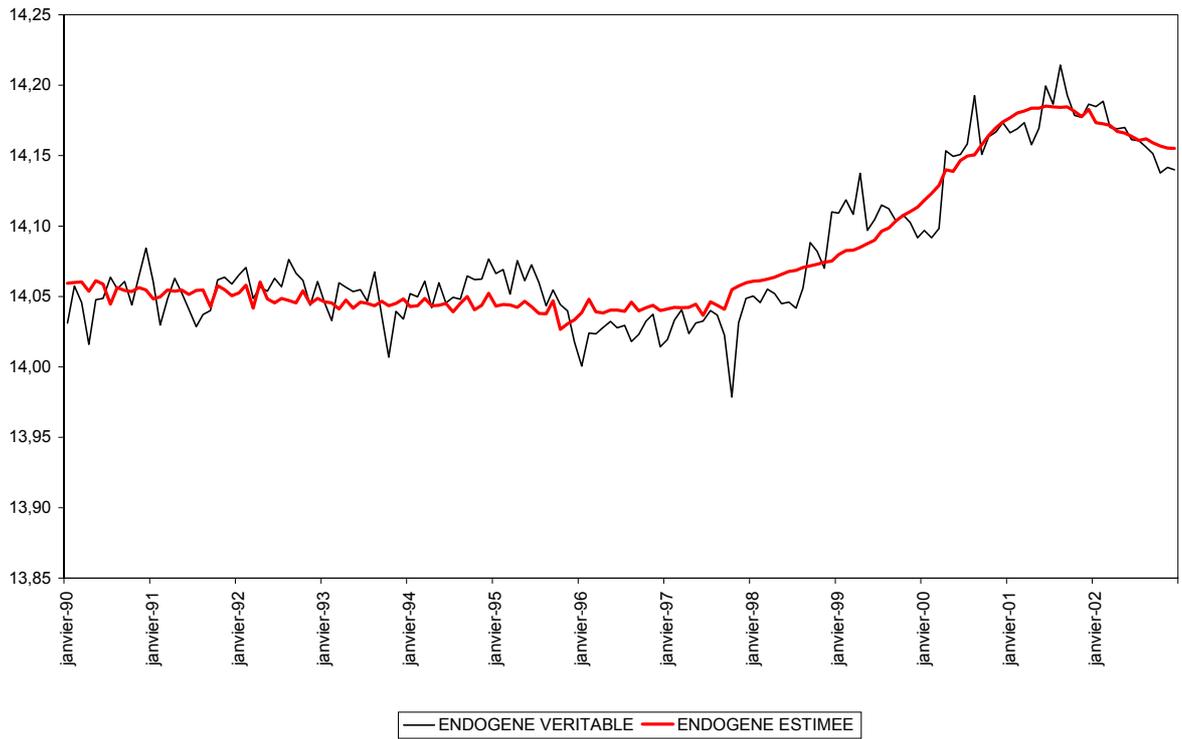
10.2. APPLICATION DES TESTS DE CHANGEMENT DU DEGRÉ DE COINTÉGRATION ET MODÈLES DE TRAFIC À LONG TERME

La rupture a eu lieu de façon plus tardive sur le RER que sur le métro : octobre 1997 contre décembre 1995. L'impact des grèves et attentats sur le trafic du jour ouvrable a été moins important sur le RER (baisse de -3.0% en moyenne annuelle entre 1995 et 1996) que sur le métro (-6.5%). La reprise a été plus lente (augmentation du trafic du jour ouvrable de $+0.7\%$ en moyenne annuelle sur le RER contre $+0.3\%$ sur le métro). La plus faible variabilité du trafic du jour ouvrable sur le réseau s'explique par la part plus importante des abonnements : sur le RER, 72.7% du trafic est réalisé à l'aide d'abonnements, contre 65.8% sur le métro. L'importance des abonnements limite la chute du trafic du jour ouvrable, par une fidélisation de la clientèle.

En seconde sous-période, le trafic du jour ouvrable sur le RER se caractérise par une élasticité à la variable de demandes d'emplois en fin de mois plus importante que sur le métro. Le RER correspond au mode de transport privilégié des banlieusards éloignés qui viennent travailler à Paris et n'utilisent pas leur voiture du fait des embouteillages et des difficultés à stationner dans Paris.



CHAPITRE 10. RECHERCHE DES ÉQUILIBRES DE LONG TERME ET DES DYNAMIQUES DE COURT TERME



10.2.3 Trafic du jour ouvrable sur le réseau des bus de Paris

Les relations du trafic du jour ouvrable sur le réseau des bus de Paris en fonction d'une tendance et/ou constante et de tout ou partie du sous-ensemble des variables économiques sont des relations de cointégration au seuil de 10%. La meilleure au sens du critère BIC est la suivante :

$$\begin{aligned}
 \ln(joparsa) = & 0.0858 \ln(defmidf) + 0.0007 gdmagrp \\
 & (4.0885) \qquad\qquad\qquad (2.1947) \\
 & + 0.0783 txhotidf + 12.6823 + 0.0004 t \\
 & (1.6640) \qquad\qquad\qquad (41.8173) \qquad (8.3249)
 \end{aligned}$$

Test de cointégration sans rupture avec tendance

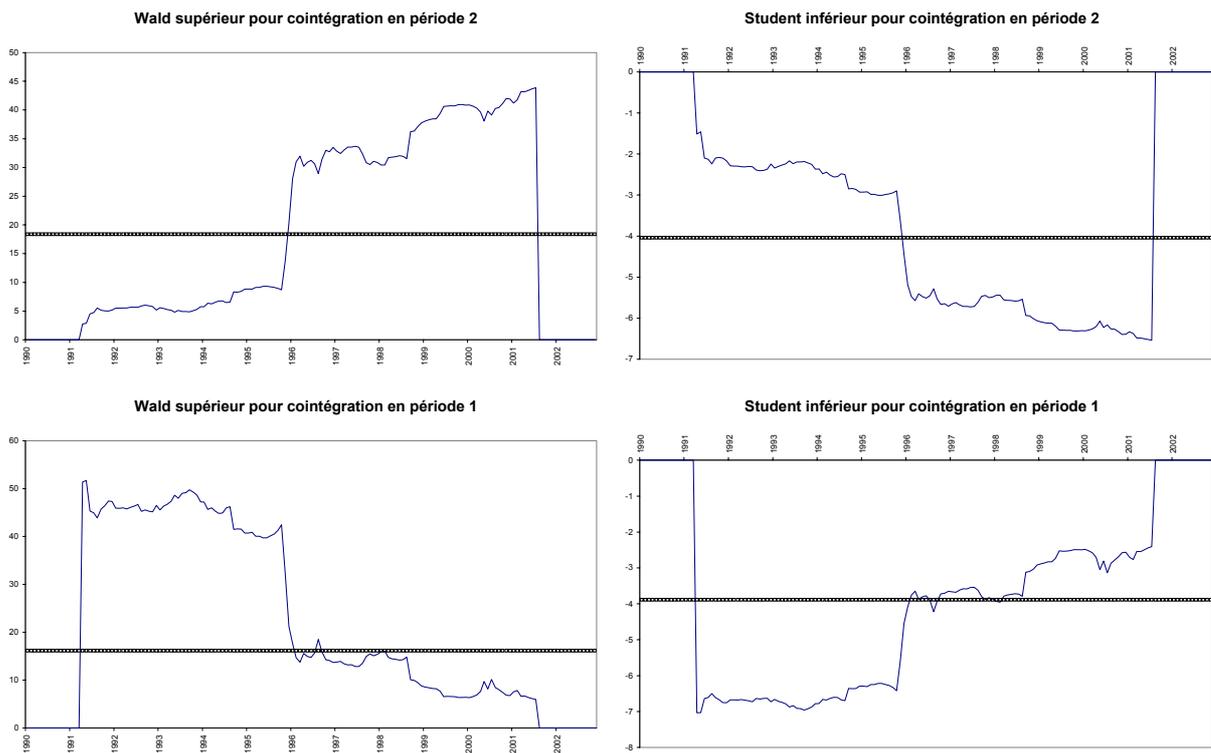
$\ln(joparsa)$	Retards	Rho	Statistique	VC 5%	VC 10%	Conclusion
t_{CADF}	1	0.3027	-7.1247	-4.5672	-4.2717	Cointégration

10.2. APPLICATION DES TESTS DE CHANGEMENT DU DEGRÉ DE COINTÉGRATION ET MODÈLES DE TRAFIC À LONG TERME

Remarque : Appliquer les tests de modification de cointégration dans ce cadre cointégré conduit à mettre en avant l'existence de cointégration sur une très longue sous-période non contrainte.

Tests de modification du degré de cointégration

$\ln(joparsa)$	Statistique	VC 10%	VC 20%	Date	Conclusion
Absence de cointégration puis cointégration					
Wald supérieur	51.6668	19.773	15.796	juin 1991	CI(1)-CI(0)
Student inférieur	-7.0364	-4.189	-3.862	mai 1991	CI(1)-CI(0)
Cointégration puis absence de cointégration					
Wald supérieur	43.8965	17.948	15.305	juillet 2001	CI(0)-CI(1)
Student inférieur	-6.5341	-4.118	-3.783	juillet 2001	CI(0)-CI(1)



Graphiquement, néanmoins, les tests de modification du degré de cointégration mettent en avant l'existence d'une forme de rupture en 1995 : les tests ayant pour hypothèse nulle "absence

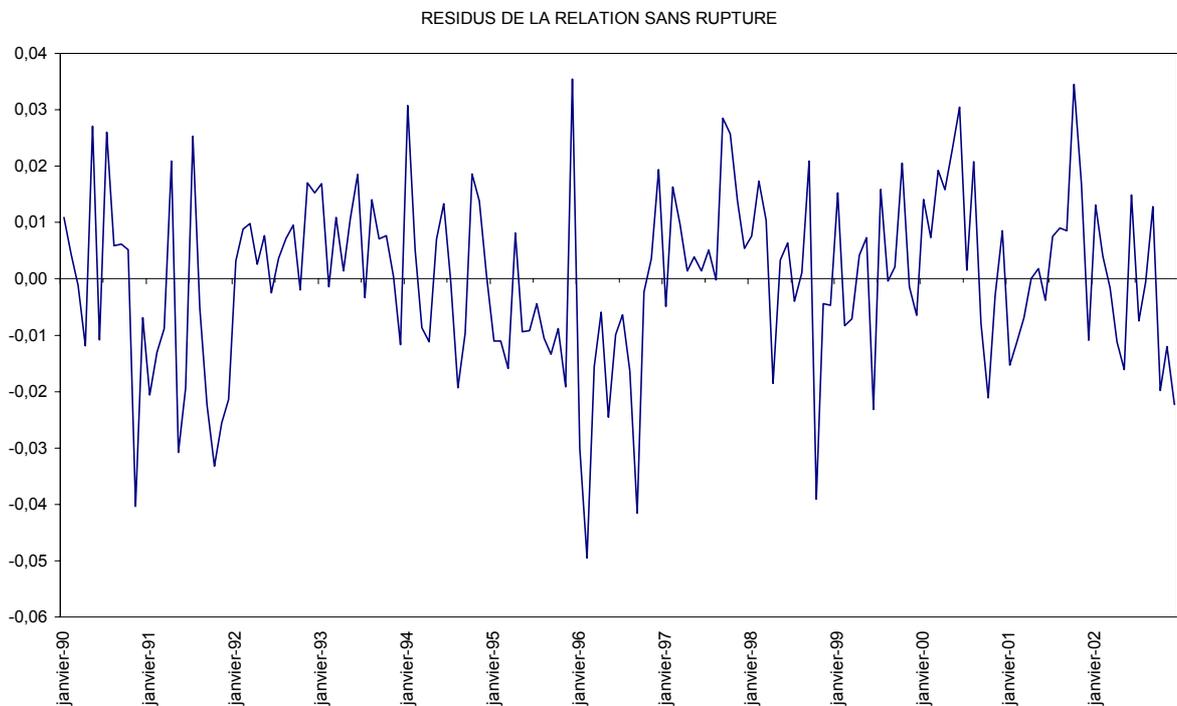
CHAPITRE 10. RECHERCHE DES ÉQUILIBRES DE LONG TERME ET DES
DYNAMIQUES DE COURT TERME

de cointégration puis cointégration” comme ceux ayant pour alternative ”cointégration puis absence de cointégration” sont proches de la valeur critique jusqu’à la date de **décembre 1995**. Ces graphiques suggèrent que si l’introduction d’une rupture n’est pas nécessaire pour obtenir une équation de long terme (cointégrée), celle-ci permettra sans doute d’obtenir une meilleure qualité de la régression.

Une autre façon d’obtenir une meilleure modélisation du trafic du jour ouvrable sur le bus de banlieue consiste à prendre en compte un plus grand nombre de variables explicatives dans la relation de cointégration.

Equation retenue :

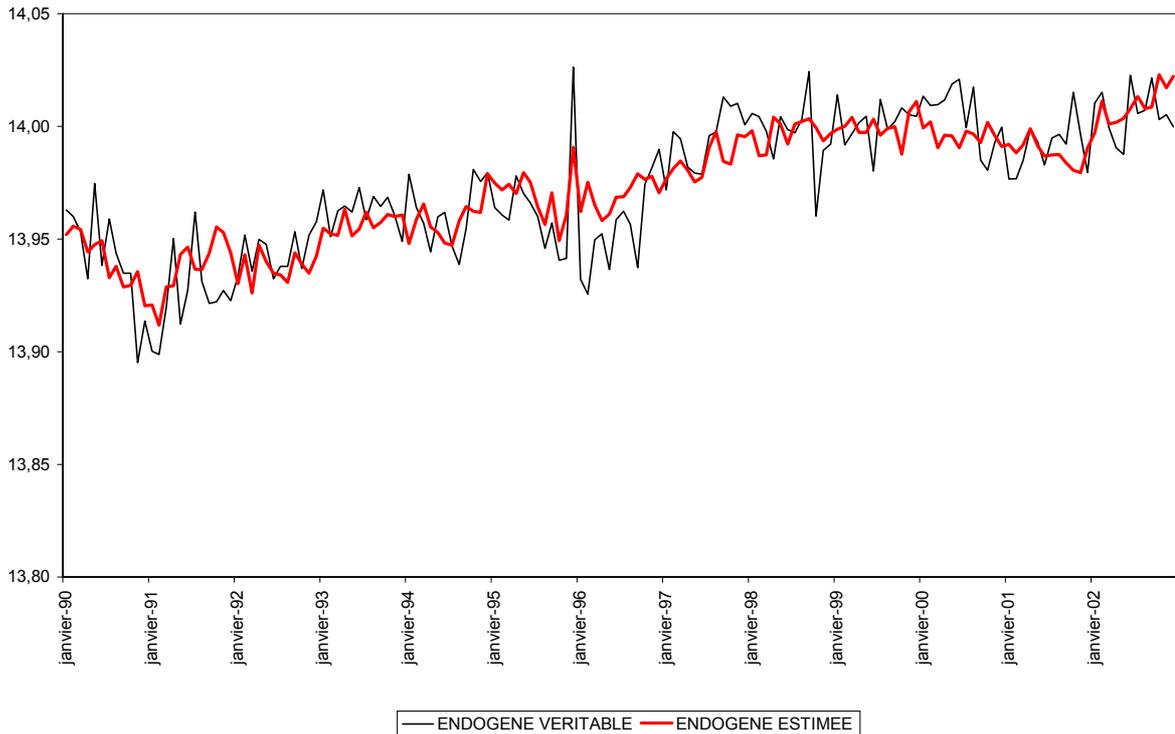
$$\begin{aligned} \ln(joparsa) = & - 0.2116 \text{ } circ75 - 0.0775 \text{ } \ln(speed75) + 0.1605 \text{ } \ln(defmidf) \\ & (-4.3731) \qquad \qquad (-1.8059) \qquad \qquad (9.6471) \\ & + 0.0012 \text{ } gdmagr + 0.0914 \text{ } txhotidf + 13.5304 \\ & (4.1499) \qquad \qquad (2.8172) \qquad \qquad (36.6813) \end{aligned}$$



10.2. APPLICATION DES TESTS DE CHANGEMENT DU DEGRÉ DE COINTÉGRATION ET MODÈLES DE TRAFIC À LONG TERME

Test de cointégration sans rupture sans tendance

$\ln(joparsa)$	Retards	Rho	Statistique	VC 10%	Conclusion
t_{CADF}	2	0.3728	-5.2238	-4.5501	Cointégration



Le trafic du jour ouvrable sur le réseau des autobus de Paris n'a pas subi de rupture aussi importante que celle sur le réseau ferré. Les attentats dans le milieu des années 1990 ont moins pénalisé ce sous-réseau, qui a bénéficié d'un report des modes souterrains.

Sur l'ensemble de la période, le trafic du réseau des bus de Paris est pénalisé par les conditions de circulation dans Paris. La consommation dans les grands magasins de la région parisienne est un facteur explicatif de même que le taux d'occupation des hôtels. Une faible proportion des touristes utilise les autobus de Paris, du fait de la complexité du sous-réseau par rapport à celui du métro. Néanmoins, les autobus peuvent attirer une clientèle qui ressent un sentiment d'insécurité dans le métro souterrain ou désire se déplacer en découvrant un peu la ville. D'après

une étude d'avril 1998, la clientèle des Français de province se déplace essentiellement en taxi, bus, métro et au moyen de la marche à pied. Ainsi, environ 2% des provinciaux utilisent le bus lorsqu'ils se rendent à Paris pour des raisons autres que purement professionnelles. Ces touristes viennent principalement pour visiter leurs amis (57% des motivations principales) et pour le shopping (seconde raison la plus évoquée).

L'élément le plus frappant de cette modélisation réside dans le signe positif de l'élasticité du trafic du jour ouvrable sur le réseau des autobus de banlieue. L'accroissement du chômage entraîne une augmentation du trafic observé sur les bus. Plusieurs pistes d'explication peuvent être lancées. La première consiste à considérer qu'il existe un report de la clientèle d'un réseau fermé vers un réseau plus ouvert, dans lequel le non paiement est moins visible. La deuxième explication réside dans le fait que la lutte contre la fraude est facilitée lors de périodes de reprise économique. Mais l'explication la plus probable réside dans le comportement même des utilisateurs des autobus de Paris. Selon une étude de septembre 2003 [61], "le bus de Paris est un mode dont l'utilisation est souvent circonscrite aux moments exemptés de contraintes de temps", "un luxe que l'on peut s'offrir quand on a le temps". En outre, un client du bus de Paris emprunte en moyenne un faible nombre de lignes (2 à 3 sur les 58 existantes à l'époque), pour des trajets connus et habituels.

10.2.4 Trafic du jour ouvrable sur le réseau des bus de banlieue

Quelle que soit le type de modélisation du trafic du jour ouvrable sur le réseau des bus de banlieue avec tendance, la relation obtenue est une relation de cointégration au seuil de 10%. La meilleure relation de cointégration en terme de critère BIC est la suivante :

$$\begin{aligned} \ln(\text{jobansa}) = & - 0.0741 \ln(\text{defmidf}) + 0.0015 \text{ gdmagrp} \\ & (-3.6056) \qquad \qquad \qquad (4.5399) \\ & + 15.0792 + 0.0018 t \\ & (51.7825) \qquad (37.3373) \end{aligned}$$

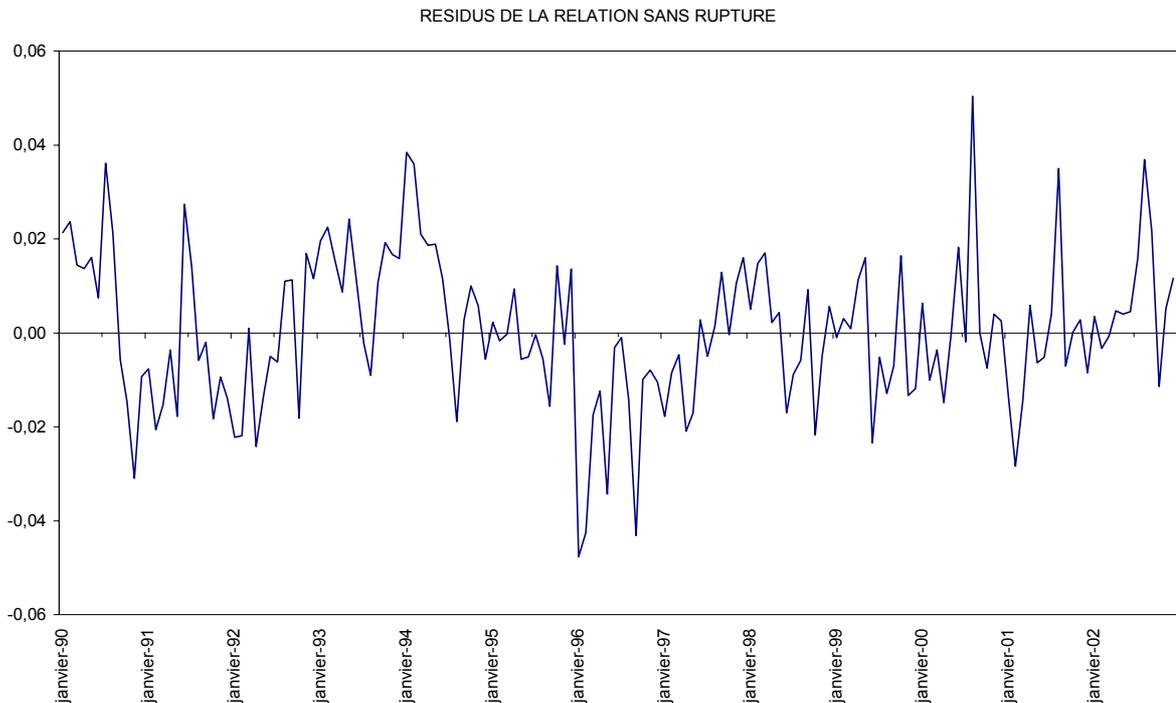
10.2. APPLICATION DES TESTS DE CHANGEMENT DU DEGRÉ DE COINTÉGRATION ET MODÈLES DE TRAFIC À LONG TERME

Cette relation est une relation de cointégration :

Test de cointégration sans rupture avec tendance

$\ln(jobansa)$	Retards	Rho	Statistique	VC 5%	VC 10%	Conclusion
t_{CADF}	1	0.4920	-5.9758	-4.2429	-3.9267	Cointégration

Elle demeure la meilleure relation en termes de critère BIC lorsque l'ensemble des variables économiques est pris en compte. L'absence de rupture de la relation traduit l'amélioration constante du sous-réseau des autobus de banlieue.



Remarque : Appliquer les tests de modification de cointégration dans ce cadre cointégré conduit à mettre en avant l'existence de cointégration sur une très longue sous-période non

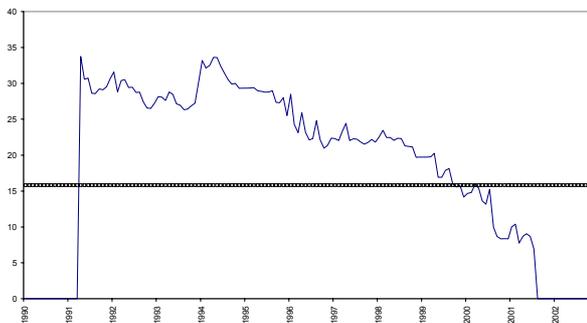
CHAPITRE 10. RECHERCHE DES ÉQUILIBRES DE LONG TERME ET DES DYNAMIQUES DE COURT TERME

contrainte.

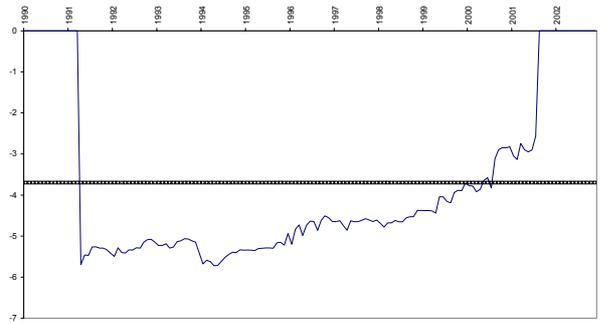
Test de modification du degré de cointégration

ln (<i>jobansa</i>)	Statistique	VC 10%	VC 20%	Date	Conclusion
Absence de cointégration puis cointégration					
Wald supérieur	33.7408	18.428	15.827	mai 1991	CI(1)-CI(0)
Student inférieur	-5.7209	-4.027	-3.699	mai 1994	CI(1)-CI(0)
Cointégration puis absence de cointégration					
Wald supérieur	128.5954	16.347	13.794	juillet 2001	CI(0)-CI(1)
Student inférieur	-5.2674	-3.909	-3.580	août 2001	CI(0)-CI(1)

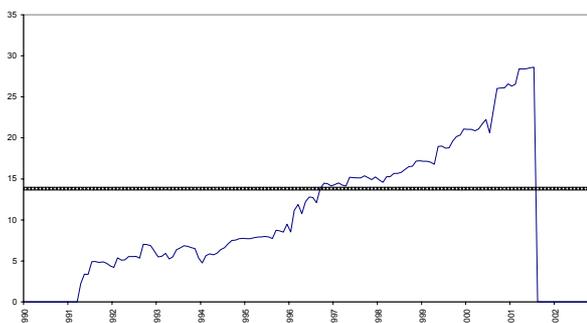
Wald supérieur pour cointégration en période 2



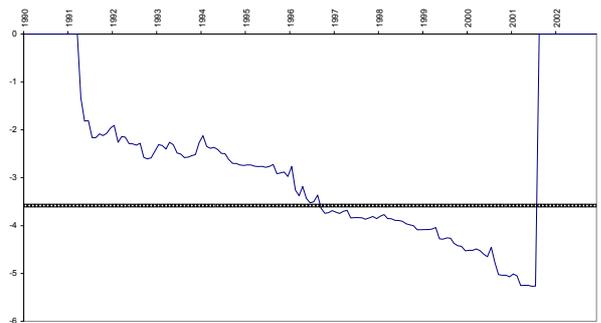
Student inférieur pour cointégration en période 2



Wald supérieur pour cointégration en période 1



Student inférieur pour cointégration en période 1

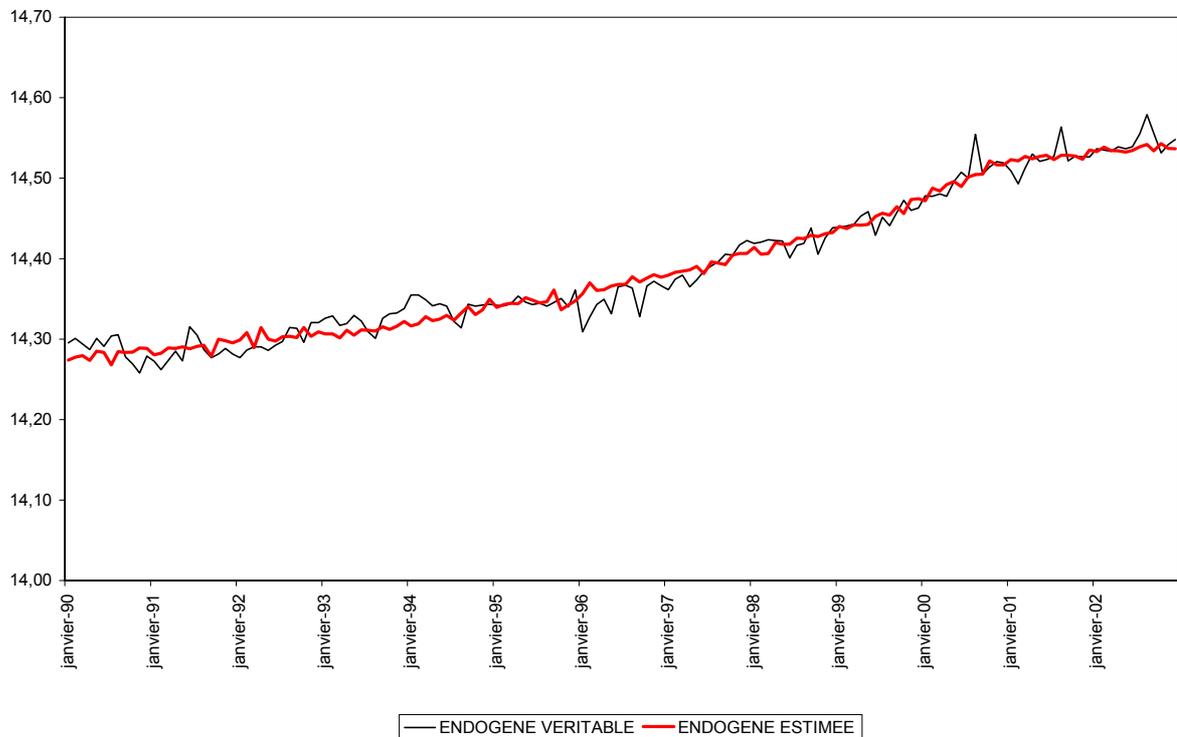


Les tests de modification du degré de cointégration concluent que la relation est à la fois "cointégrée puis non cointégrée" et "non cointégrée puis cointégrée", ce qui revient à considérer que la relation est cointégrée sur l'ensemble de la période, conformément aux résultats du test de cointégration.

10.2. APPLICATION DES TESTS DE CHANGEMENT DU DEGRÉ DE COINTÉGRATION ET MODÈLES DE TRAFIC À LONG TERME

Contrairement au sous-réseau des autobus de Paris, le trafic du jour ouvrable sur le sous-réseau des autobus de banlieue possède une élasticité négative aux demandes d'emplois en fin de mois en Ile-de-France, comme sur les sous-réseaux ferrés sur période entière. Cet élément souligne la complémentarité entre bus de banlieue et réseau ferré, les bus de banlieue étant en grande partie utilisés comme mode de rabattement aux stations de métro et gares de RER lors des trajets domicile-travail, ou pour les achats dans les grands magasins.

L'élasticité du trafic du jour ouvrable sur les autobus de banlieue aux données de tourisme est non significative. Ce mode de transport n'est pas le mode de transport favori des touristes qui résident en banlieue. Les Français de province qui résident en banlieue lors de leur séjour en banlieue privilégient l'automobile pour leurs déplacements.



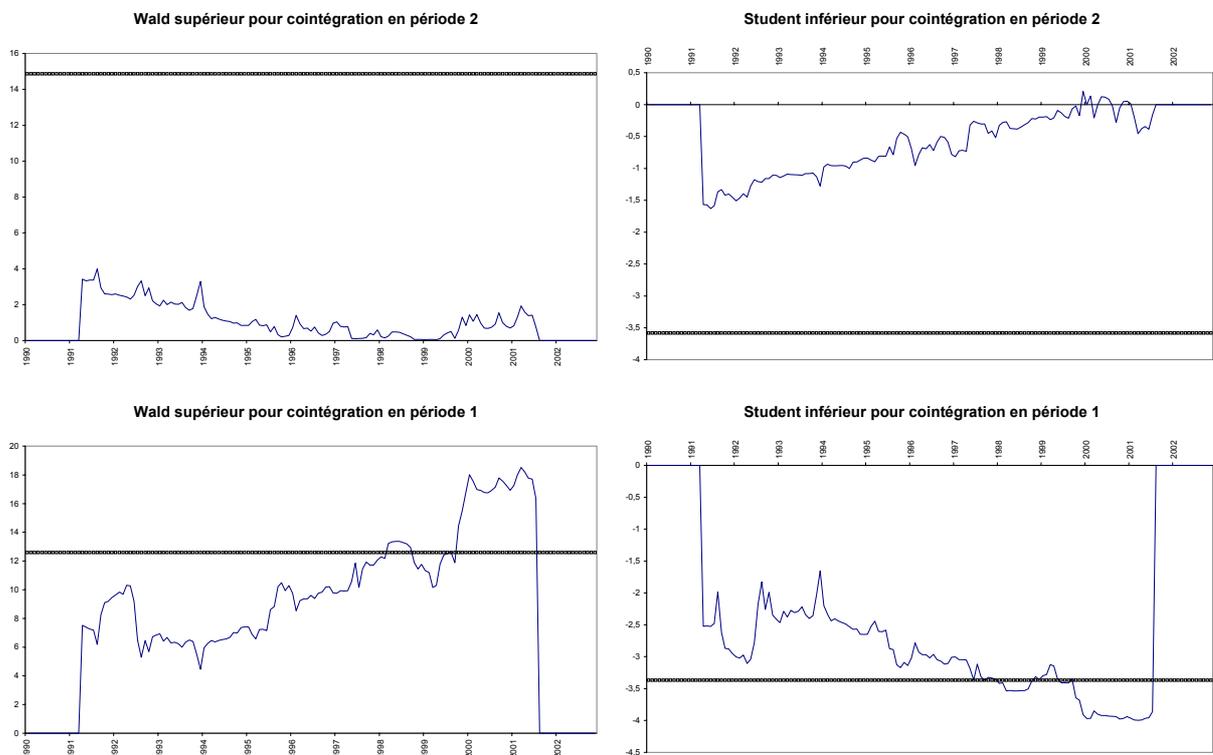
10.2. APPLICATION DES TESTS DE CHANGEMENT DU DEGRÉ DE COINTÉGRATION ET MODÈLES DE TRAFIC À LONG TERME

La relation n'est pas cointégrée, nous lui appliquons les tests de modification du degré de cointégration.

Tests de modification du degré de cointégration

$\ln(smsa)$	Statistique	VC 10%	VC 20%	Date	Conclusion
Absence de cointégration puis cointégration					
Wald supérieur	3.9953	17.275	14.870	septembre 1991	CI(1)-CI(1)
Student inférieur	-1.6300	-3.887	-3.580	juillet 1991	CI(1)-CI(1)
Cointégration puis absence de cointégration					
Wald supérieur	18.5241	14.877	12.599	mars 2001	CI(0)-CI(1)
Student inférieur	-3.9945	-3.681	-3.368	mars 2001	CI(0)-CI(1)

Les tests de modification du degré de cointégration ayant pour hypothèse alternative "cointégration puis absence de cointégration" s'aperçoivent avec retard du changement de régime.



CHAPITRE 10. RECHERCHE DES ÉQUILIBRES DE LONG TERME ET DES
DYNAMIQUES DE COURT TERME

La date de rupture retenue correspond au premier extrêum tel que la statistique du test de Student est inférieure à la valeur critique à 20%, soit **février 1998**.

10.2.5.2 Modélisation du trafic du samedi sur le métro pour chacune des sous-périodes

ln(<i>smsa</i>)	Valeurs des coefficients					Test CADF			
	ln(defmidf)	gdmagrp	txhotidf	c	t	ρ	t_{CADF}	VC _{10%}	Concl.
Fév. 1998									
Avant	0.3556 (4.9992)	0.0041 (5.3272)	0.4650 (5.3167)	9.3271 (9.5886)	-0.0013 (-4.2663)	0.1963	-8.0162	-4.2717	CO(0)
Après		0.0020 (2.3369)	0.2747 (1.5773)	14.3056 (99.2953)	0.0022 (7.2488)	0.2887	-5.5499	-3.9267	CO(0)

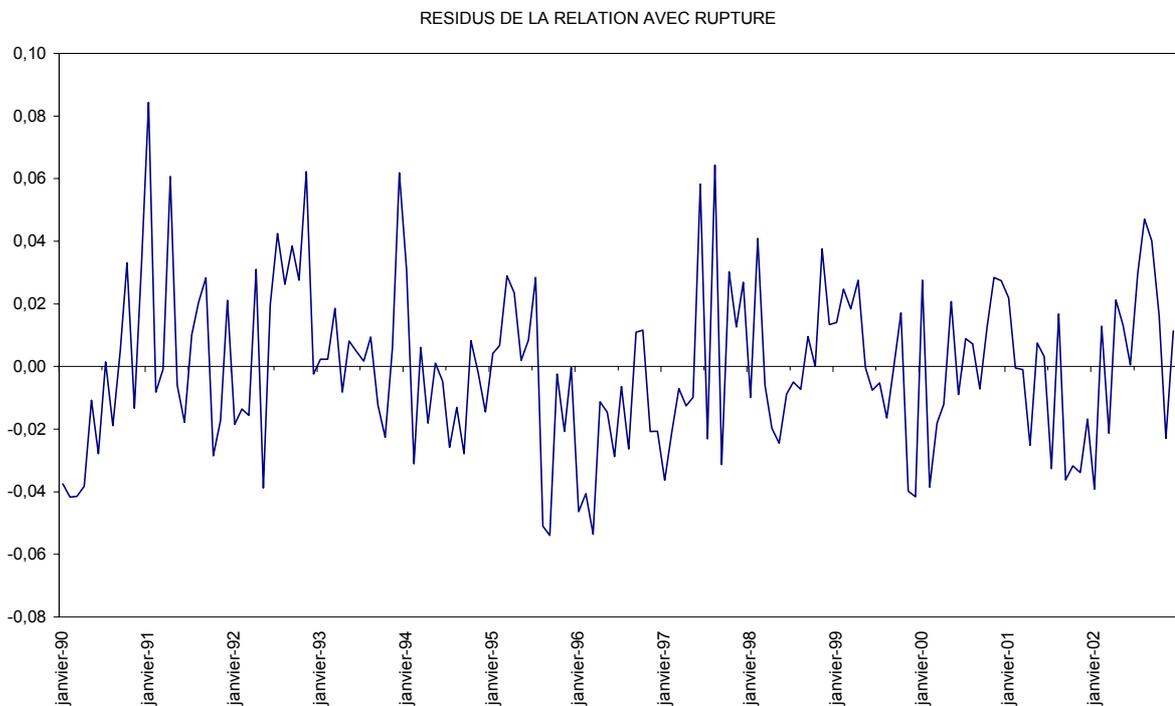
L'année 1998 correspond à l'année de plus forte progression de l'activité du commerce de détail et de l'artisanat commercial depuis 1986 au niveau national. La forte consommation des ménages bénéficie aux commerces en régulière expansion, mais aussi à l'ameublement, après 7 années de recul, ou encore au commerce automobile. Cette modification du type d'achat, ainsi que le développement de l'achat en ligne depuis 1999, expliquent en partie la baisse de l'élasticité du trafic du métro à la variable "*gdmagrp*". La principale explication de la diminution de l'élasticité entre les deux sous-périodes réside dans la forte reprise de la consommation. Lorsque les consommateurs se rendent dans les magasins, ils ont tendance à acheter : les déplacements dans les commerces qui se soldent par l'absence d'achat diminuent.

Les années 1997 et 1998 marquent une rupture en terme de tourisme pour la ville de Paris. En 1997 ont eu lieu à Paris les Journées Mondiales de la Jeunesse ; en 1998 a été organisée la Coupe du Monde de football dans l'Ile-de-France. Ces deux éléments ont attiré une clientèle touristique différente des années précédentes, en assurant la promotion de Paris comme destination touristique. L'attraction de Paris a été renforcée les années suivantes par les opérations "Paris Plage" à partir de 2002, concept qui a été repris dans de nombreuses villes étrangères par la suite, ou encore la création de la "Nuit Blanche".

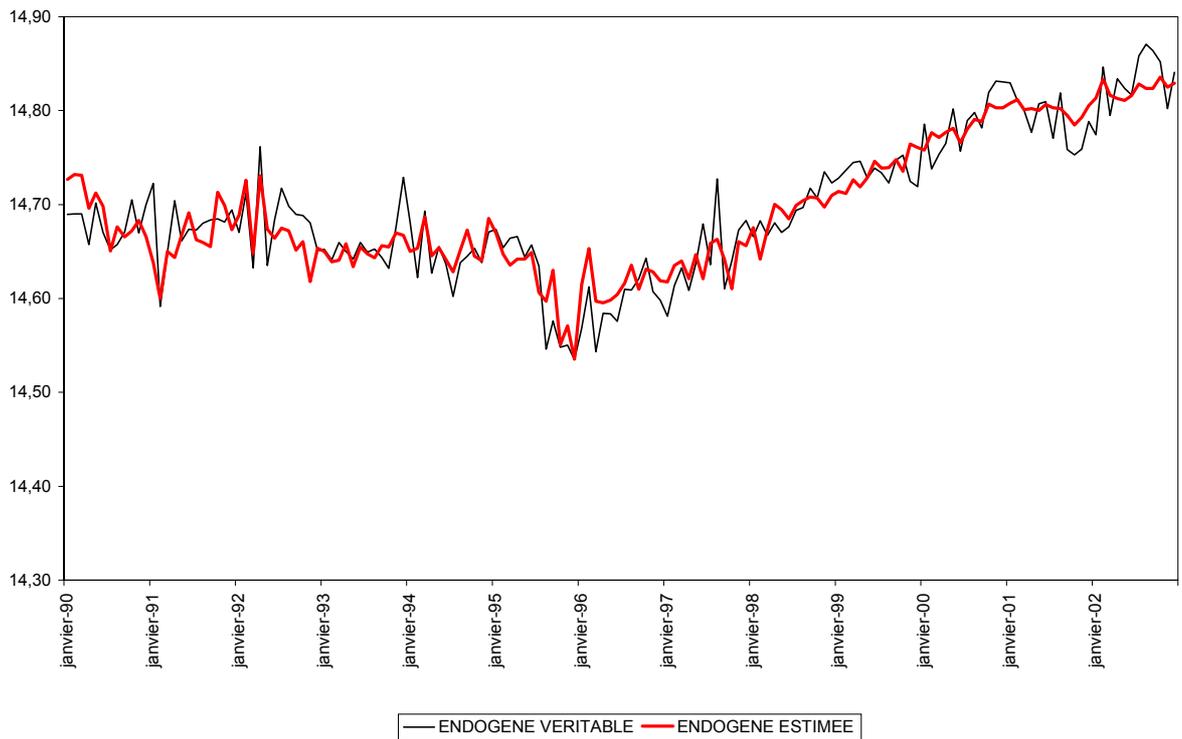
10.2. APPLICATION DES TESTS DE CHANGEMENT DU DEGRÉ DE COINTÉGRATION ET MODÈLES DE TRAFIC À LONG TERME

La Coupe du Monde de football a eu en outre un impact fort sur les loisirs des Français, et des Franciliens en particulier. Les personnes ne se sont pas seulement déplacées pour suivre les matchs dans les stades, elles se sont aussi rendues dans des cafés, ou chez des amis pour suivre la progression des équipes. La liesse provoquée par chaque victoire a été partagée collectivement et a donné lieu à des manifestations spontanées, sur les Champs-Élysées par exemple. Pour de nombreux observateurs, cet événement a réveillé l'envie de "sortir" de chez soi. La création de deux formules d'abonnement "illimité" donnant "libre accès" aux salles de cinéma illustre cet essor des loisirs.

Cette rupture dans les comportements a été accompagnée de la création du titre Imagine R pour les étudiants et scolaires au cours de l'année 1998. Ce titre leur donne droit à un nombre illimité de voyages durant la durée de validité, il permet en outre de se déplacer sans limite de zone durant les week-ends, et, à partir de l'année 2000, durant toutes les petites vacances scolaires.



CHAPITRE 10. RECHERCHE DES ÉQUILIBRES DE LONG TERME ET DES DYNAMIQUES DE COURT TERME



10.2.6 Trafic du dimanche/ jour férié sur le métro

10.2.6.1 Absence de modélisation sans rupture

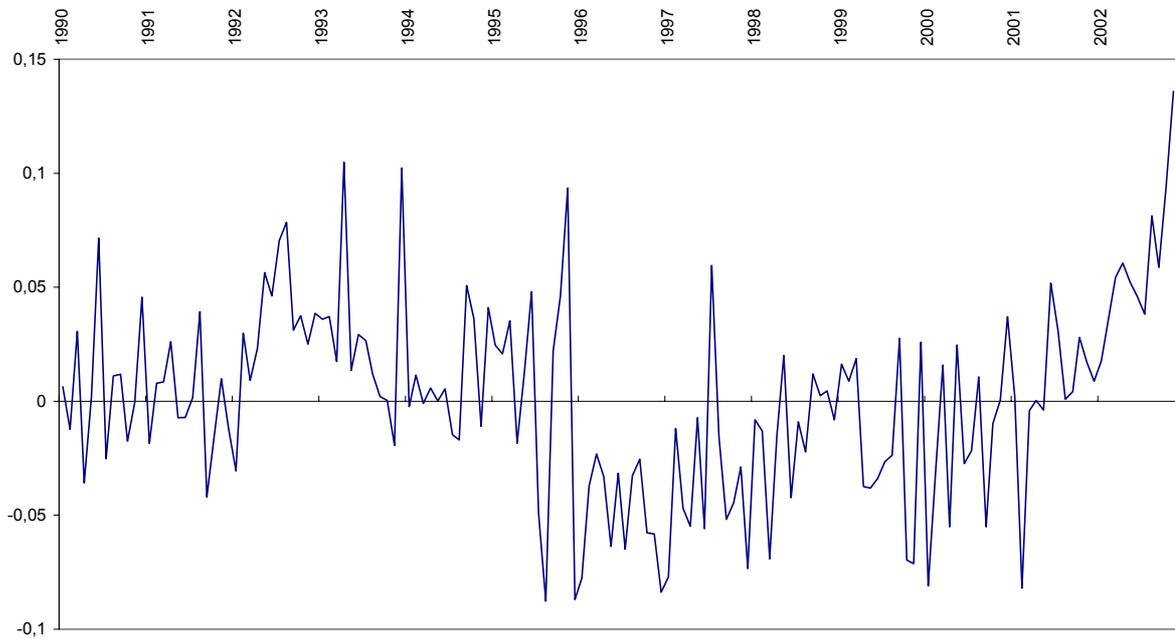
Quelle que soit la modélisation choisie, aucune relation relative au trafic du dimanche/jour férié sur le métro n'est cointégrée au seuil de 10%. Considérons la relation suivante :

$$\ln(dfmsa) = 0.0072 \text{ } gdmagr\text{p} + 0.1519 \text{ } txhotidf + 13.3569 + 0.0011 \text{ } t$$

(8.8475)
(1.7763)
(205.9530)
(8.8049)

10.2. APPLICATION DES TESTS DE CHANGEMENT DU DEGRÉ DE COINTÉGRATION ET MODÈLES DE TRAFIC À LONG TERME

Résidus de la relation non cointégrée



Test de cointégration sans rupture et avec tendance

$\ln(dfmsa)$	Retards	Rho	Statistique	VC 5%	VC 10%	Conclusion
t_{CADF}	3	0.7650	-2.2886	-4.2429	-3.9267	Absence de cointégration

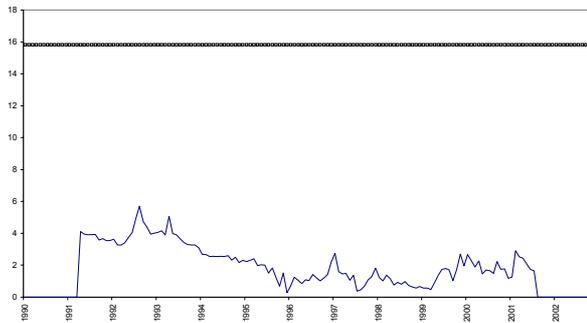
Cette relation n'est pas une relation de cointégration, nous lui appliquons les tests de modification du degré de cointégration.

CHAPITRE 10. RECHERCHE DES ÉQUILIBRES DE LONG TERME ET DES
DYNAMIQUES DE COURT TERME

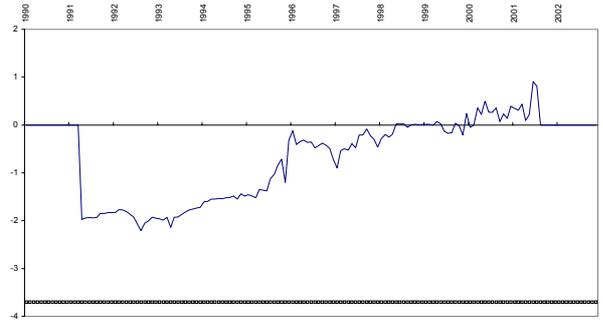
Tests de modification du degré de cointégration

$\ln(dfmsa)$	Statistique	VC 10%	VC 20%	Date	Conclusion
Absence de cointégration puis cointégration					
Wald supérieur	5.6968	18.428	15.827	septembre 1992	CI(1)-CI(1)
Student inférieur	-2.2067	-4.027	-3.699	septembre 1992	CI(1)-CI(1)
Cointégration puis absence de cointégration					
Wald supérieur	14.9581	16.347	13.794	février 2001	CI(0)-CI(1)
Student inférieur	-3.4765	-3.909	-3.580	décembre 1995	CI(1)-CI(1)

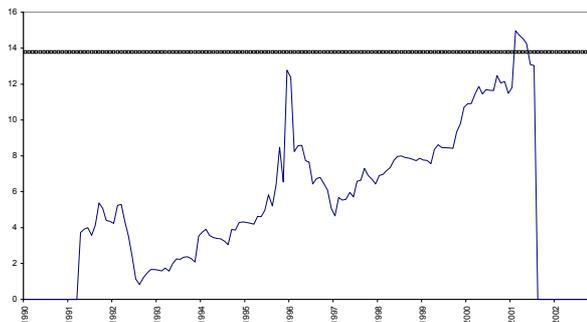
Wald supérieur pour cointégration en période 2



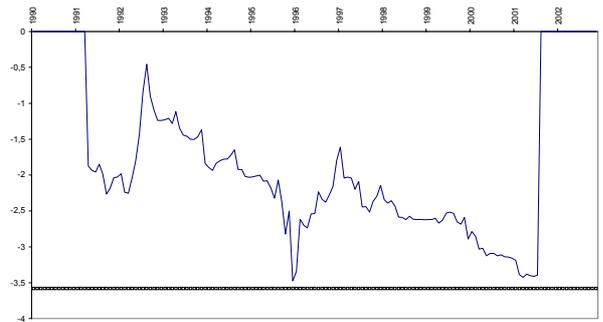
Student inférieur pour cointégration en période 2



Wald supérieur pour cointégration en période 1



Student inférieur pour cointégration en période 1



Les tests de modification du degré de cointégration ayant pour hypothèse alternative "cointégration puis absence de cointégration" estiment la véritable date de rupture de façon médiocre, mais indiquent que s'il n'existe pas de relation de cointégration sans rupture au seuil de 10%, il existe un équilibre sur une première sous-période. La date de rupture retenue correspond à

10.2. APPLICATION DES TESTS DE CHANGEMENT DU DEGRÉ DE COINTÉGRATION ET MODÈLES DE TRAFIC À LONG TERME

l'extrémum local de la statistique de Wald, au mois de **décembre 1995**, comme c'était le cas pour le trafic du jour ouvrable.

10.2.6.2 Modélisation du trafic du dimanche/jour férié sur le métro pour chacune des sous-périodes

$\ln(dfmsa)$	Valeurs des coefficients				Test CADF			
Déc. 1995	gdmagrp	txhotidf	c	t	ρ	t_{CADF}	$VC_{10\%}$	Concl.
Avant	0.0014 (1.6251)	0.1854 (2.6928)	13.9307 (206.0218)	n.s.	0.4112	-4.0628	-3.5292	CO(0)
Après	0.0010 (1.1976)	0.3340 (3.1987)	13.7923 (164.7750)	0.0036 (15.7145)	0.1690	-7.5585	-3.9267	CO(0)

La date de rupture de la modélisation du trafic du dimanche/jour férié sur le métro est similaire à celle observée pour le trafic du jour ouvrable sur ce même sous-réseau : décembre 1995. Elle est plus précoce que sur le samedi.

Une explication de la rupture réside dans l'assouplissement de la loi sur l'ouverture des magasins le dimanche à Paris. Selon le code du travail (article L-221-5, relatif à l'emploi salarié), "le repos hebdomadaire doit être donné le dimanche". Une dérogation est accordée de plein droit à ceux qui contribuent à la continuité de la vie économique et sociale (restauration, spectacles, musées...) et aux commerces à proximité des marchés alimentaires. Les commerces de détail ont la possibilité d'ouvrir 5 dimanches par an (Noël, soldes...)

Pour les autres secteurs, il existe des dérogations possibles :

- s'il y a risque de préjudice au public : pharmacie, transport, loisirs...
- s'il y a atteinte au fonctionnement de l'établissement (contraintes de production par exemple),
- ou dans les zones touristiques (actuellement 6 zones autorisées à Paris), pour les activités de loisirs ou de culture.

CHAPITRE 10. RECHERCHE DES ÉQUILIBRES DE LONG TERME ET DES DYNAMIQUES DE COURT TERME

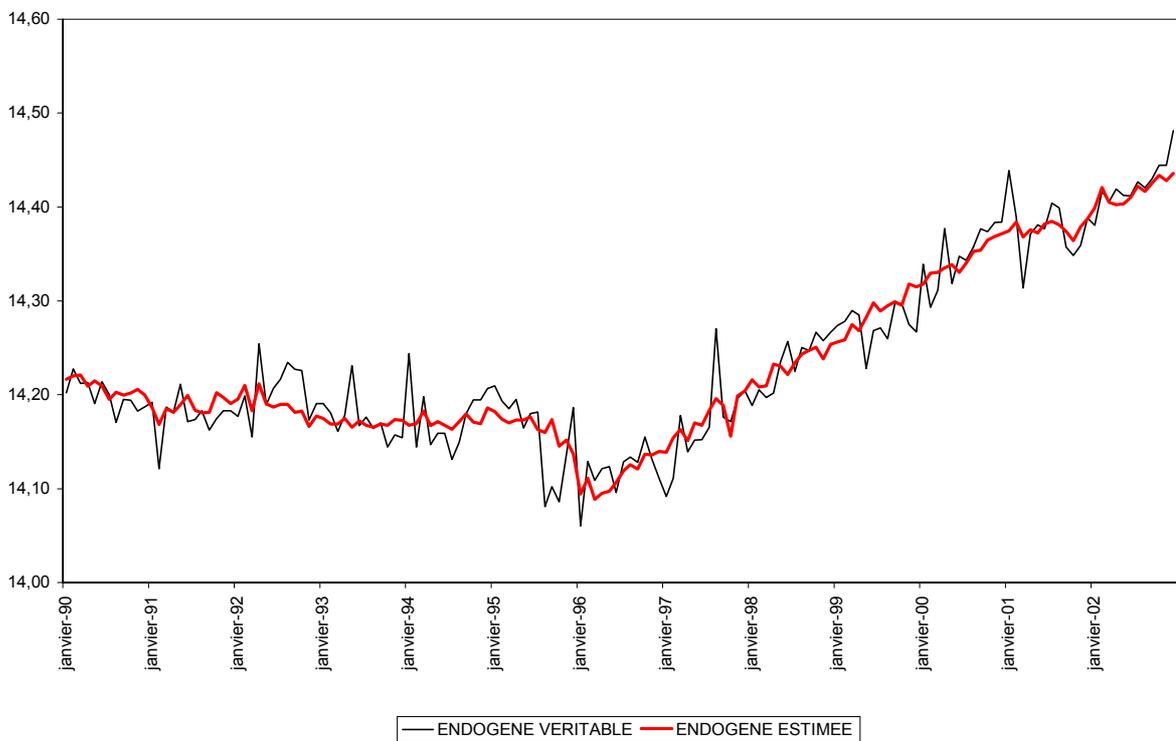
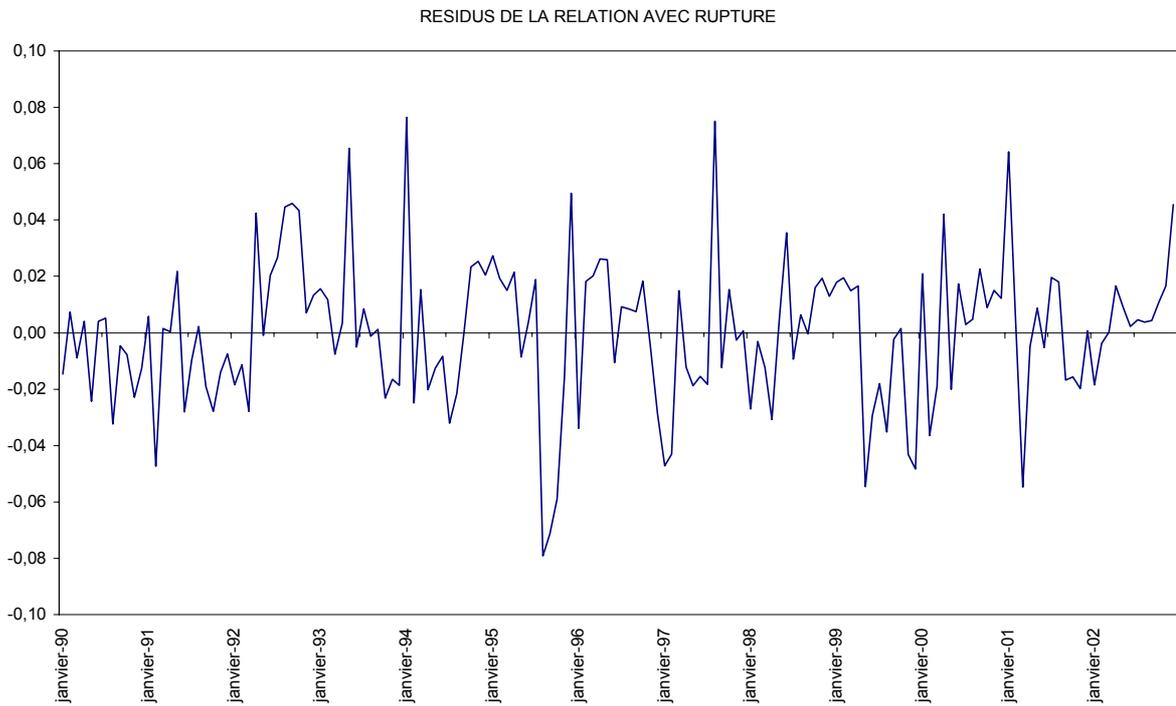
En 1994, cinq zones ont obtenu une dérogation leur permettant d'ouvrir le dimanche. D'après la loi, tout magasin situé dans ces zones touristiques, dont l'activité est en rapport avec les loisirs, la culture, le sport ou la détente, peut obtenir l'autorisation d'ouvrir le dimanche. Cette exigence est assez lâche : l'achat d'un téléphone portable est considéré par la Préfecture comme une "activité ludique", car il s'en vend "des millions de toutes les formes et de toutes les couleurs". Il suffit parfois d'ajouter quelques produits culturels (livres, disques...) dans sa boutique de prêt-à-porter pour obtenir une autorisation. La sixième et dernière zone autorisée à ouvrir le dimanche à Paris est le boulevard Saint-Germain, depuis le 20 septembre 2000.

Seuls quelques grands magasins (Virgin sur les Champs-Élysées par exemple) ouvrent le dimanche, mais les ouvertures sauvages des petits commerces sont fréquentes.

Selon le président du Comité Francs-Bourgeois, "des clients anglais, belges, suisses ou italiens prennent exprès le TGV pour faire leurs emplettes rue des Francs-Bourgeois le dimanche après-midi!". Le site de la Mairie de Paris, mais aussi de nombreux guides touristiques donnent des indications sur les rues commerçantes ouvertes le dimanche. Le dimanche est devenu un jour privilégié pour les animations de rue.

Les éléments évoqués pour la modélisation du trafic du samedi sur le métro ont aussi joué pour le trafic du dimanche ou jour férié. L'accroissement des loisirs observés à la suite de la coupe du monde de football, le développement de la fréquentation du cinéma suite à la création d'abonnements permettant l'accès illimité aux salles de deux distributeurs ou la création de la carte Imagine R ont eu un impact positif sur la fréquentation dominicale des transports collectifs. Un dernier élément favorable à l'essor des déplacements dans Paris lors des week-ends réside dans la récente gratuité d'un grand nombre de musées parisiens, ou le développement des tarifs réduits le dimanche.

10.2. APPLICATION DES TESTS DE CHANGEMENT DU DEGRÉ DE COINTÉGRATION ET MODÈLES DE TRAFIC À LONG TERME



10.2.7 Trafic du samedi sur le RER

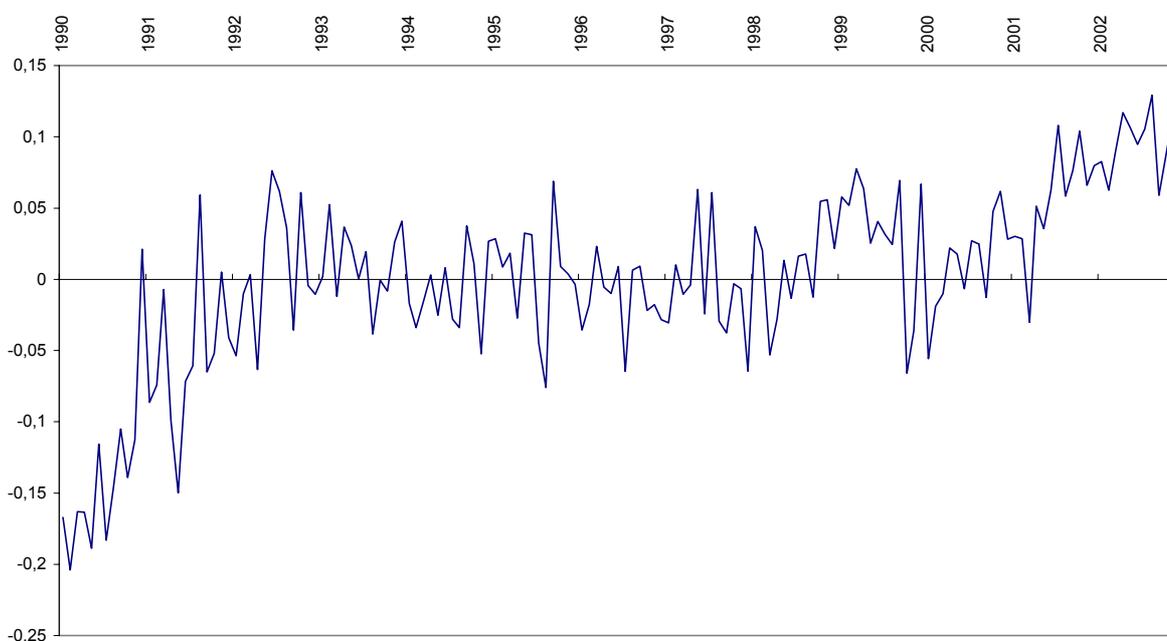
10.2.7.1 Absence de relation de cointégration sans rupture

Quelle que soit la modélisation choisie, aucune relation relative au trafic du samedi sur le RER n'est cointégrée au seuil de 10%. Considérons la relation suivante :

$$\ln(srersa) = 0.0091 \text{ } gdmagrps + 0.1541 \text{ } txhotidf + 12.3775$$

$$(7.0871) \qquad (0.9094) \qquad (114.4259)$$

Résidus de la relation non cointégrée



Cette relation n'est pas cointégrée

Test de cointégration sans rupture et sans tendance

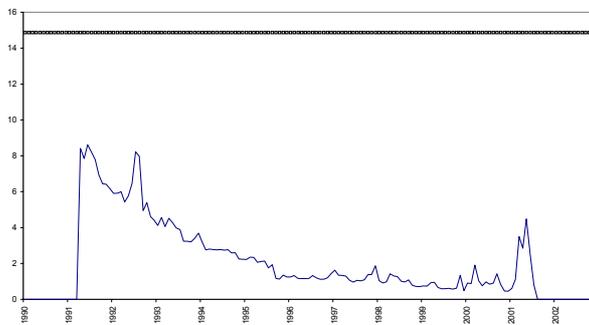
$\ln(srersa)$	Retards	Rho	Statistique	VC 5%	VC 10%	Conclusion
t_{CADF}	3	0.8346	-2.9079	-3.8378	-3.5292	Absence de cointégration

10.2. APPLICATION DES TESTS DE CHANGEMENT DU DEGRÉ DE COINTÉGRATION ET MODÈLES DE TRAFIC À LONG TERME

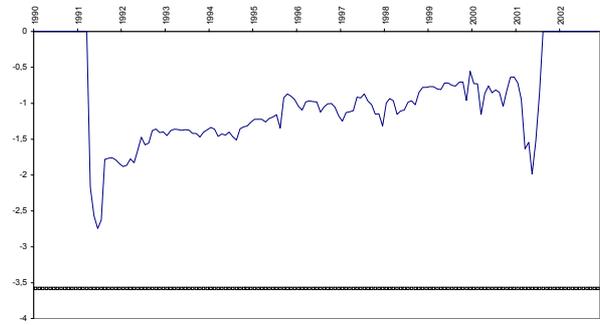
Tests de modification du degré de cointégration

$\ln(srersa)$	Statistique	VC 10%	VC 20%	Date	Conclusion
Absence de cointégration puis cointégration					
Wald supérieur	8.6110	17.275	14.870	juillet 1991	CI(1)-CI(1)
Student inférieur	-2.7436	-3.8870	-3.580	juillet 1991	CI(1)-CI(1)
Cointégration puis absence de cointégration					
Wald supérieur	16.2540	14.877	12.599	mars 2001	CI(0)-CI(1)
Student inférieur	-3.8551	-3.6810	-3.368	mars 2001	CI(0)-CI(1)

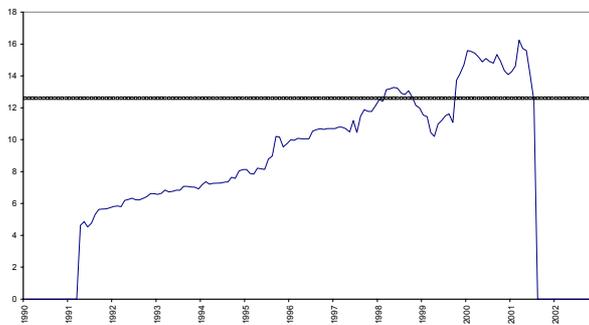
Wald supérieur pour cointégration en période 2



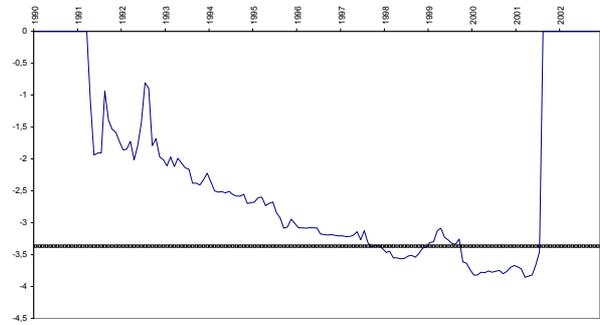
Student inférieur pour cointégration en période 2



Wald supérieur pour cointégration en période 1



Student inférieur pour cointégration en période 1



Les tests de modification du degré de cointégration ayant pour hypothèse alternative "cointégration puis absence de cointégration" ont tendance à s'apercevoir avec retard du changement de régime. La date de rupture retenue correspond au premier minimum local tel que la statistique du test de Student est inférieure à la valeur critique à 20%, soit **juin 1998**. Cette date de rupture est proche de celle observée sur le métro pour le même type de journée.

10.2.7.2 Modélisation du trafic du samedi sur le RER pour chacune des sous-périodes

La meilleure relation de cointégration pour chaque sous-période dégagée est résumée dans le tableau suivant :

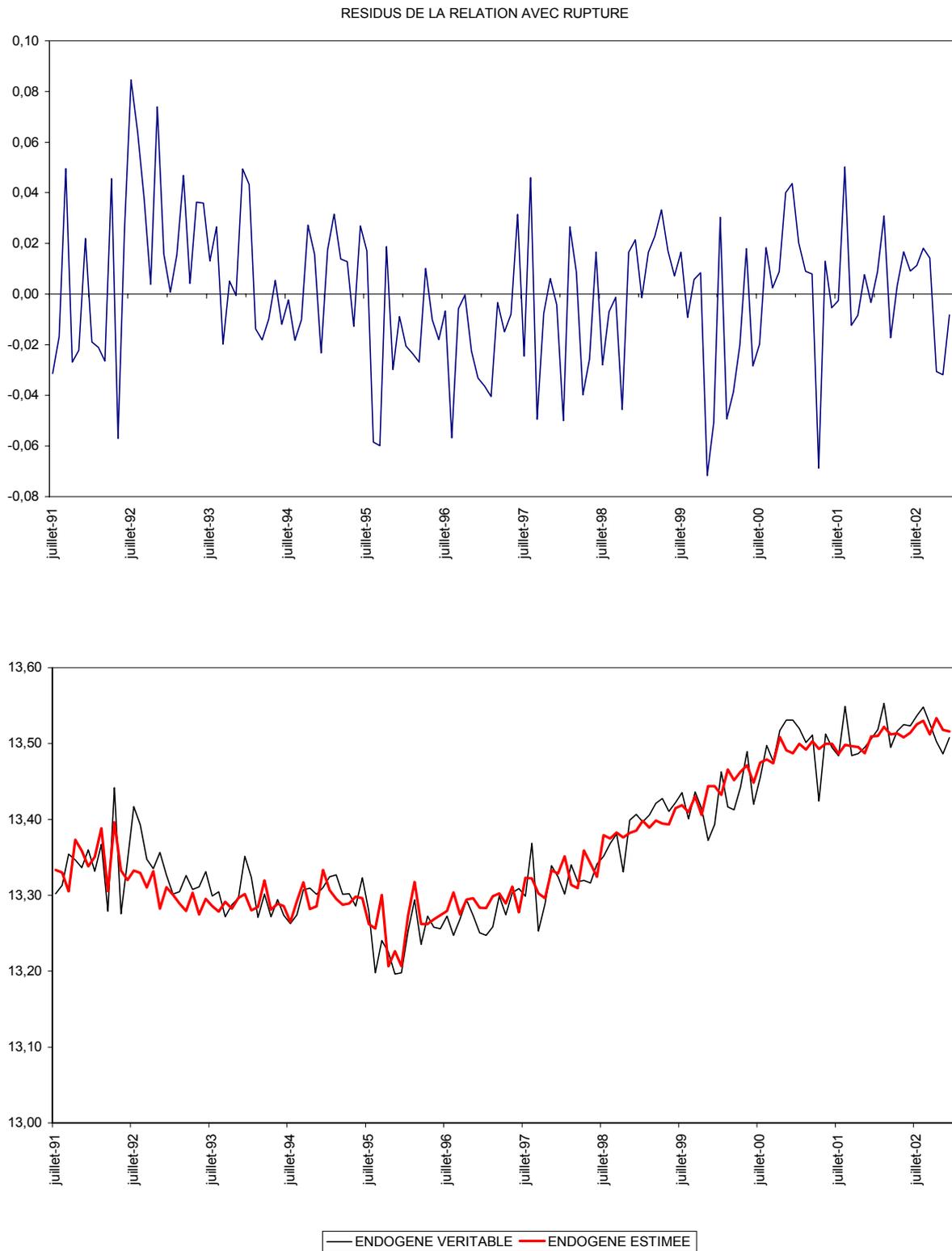
ln(<i>sresa</i>)	Valeurs des coefficients				Test CADF			
	<i>gdmagrps</i>	<i>txhotidf</i>	<i>c</i>	<i>t</i>	ρ	t_{CADF}	VC _{10%}	Concl.
Juin 1998								
Avant	0.0051 (6.9163)	0.3231 (3.7871)	12.6370 (189.2668)	n.s.	0.1602	-7.7401	-3.5292	CO(0)
Après	0.0037 (3.0513)	n.s.	13.0221 (113.0191)	0.0024 (7.3769)	0.1576	-6.2126	-3.6062	CO(0)

Comme sur le métro, le trafic du samedi sur le RER a connu une modification du niveau et de la pente de la tendance, à mettre en lien avec l'accroissement des loisirs dans un contexte de reprise économique et de coupe du monde de football. La création de l'abonnement Imagine R a lieu au cours de cette année là, et a favorisé par la suite les déplacements des scolaires et étudiants au cours des week-ends.

L'élasticité du trafic du samedi à la variable de volume des ventes dans les grands magasins de la région parisienne est presque deux fois plus importante sur le RER que sur le métro. Cette différence s'explique par la localisation des grands magasins les plus fréquentés : Printemps et Galeries Lafayette (RER A, gare d'"Auber"), FNAC du Forum des Halles (gare de "Châtelet" desservis par les RER A, B et D), Virgin sur les Champs-Élysées (gare Etoile-Charles de Gaulle), centre commercial des Quatre-Temps (la Défense) ...

L'ouverture au public de la ligne 14 du métro (Météor) a provoqué un report des touristes du RER vers le métro au cours du week-end. Au delà de la nouveauté de cette ligne entièrement automatique, les temps d'attente y sont moins élevés que sur le RER. La non significativité de la variable de tourisme pour ce type de trafic s'explique en outre par la faiblesse du nombre de sites touristiques desservis par rapport au métro, en particulier la ligne 1.

10.2. APPLICATION DES TESTS DE CHANGEMENT DU DEGRÉ DE COINTÉGRATION ET MODÈLES DE TRAFIC À LONG TERME



10.2.8 Trafic du dimanche/jour férié sur le RER

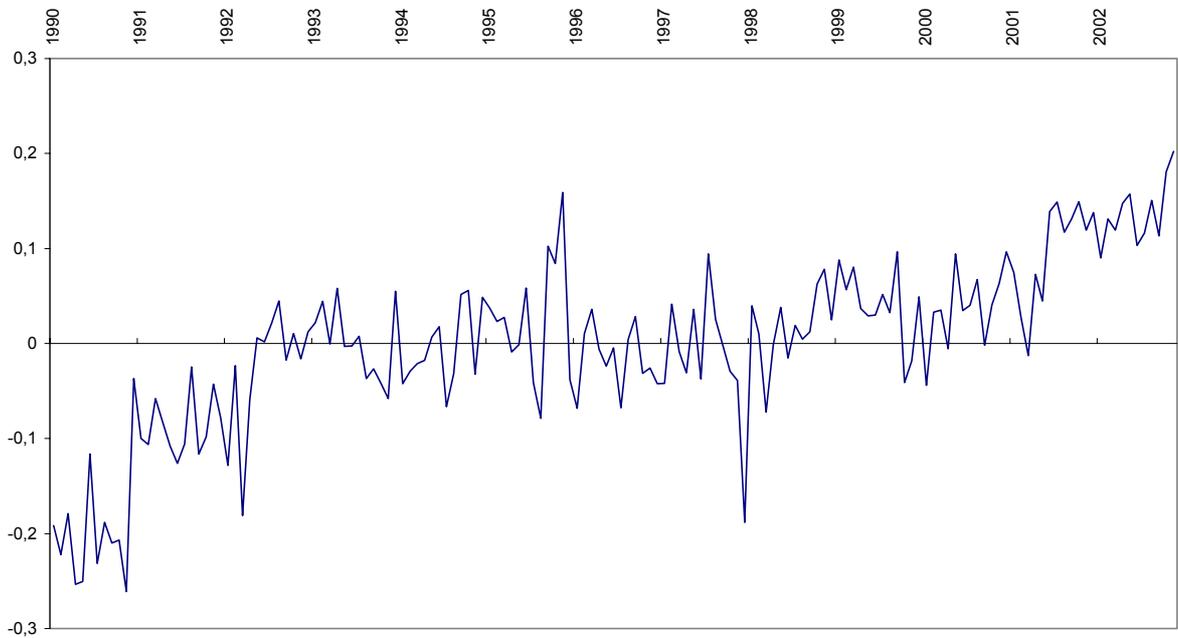
10.2.8.1 Absence de relation sans rupture

Il n'existe pas de relation de cointégration sur la période. Considérons la meilleure relation en terme de BIC :

$$\ln(df\text{rersa}) = 0.0092 \text{ } gdmagr\text{p} + 0.2129 \text{ } txhotid\text{f} + 11.8128$$

(4.8847) (0.9703) (72.3256)

Résidus de la relation non cointégrée



Cette relation n'est pas cointégrée :

Test de cointégration sans rupture et sans tendance

$\ln(df\text{rersa})$	Retards	Rho	Statistique	VC 5%	VC 10%	Conclusion
t_{CADF}	3	0.8876	-2.0617	-3.8378	-3.5292	Absence de cointégration

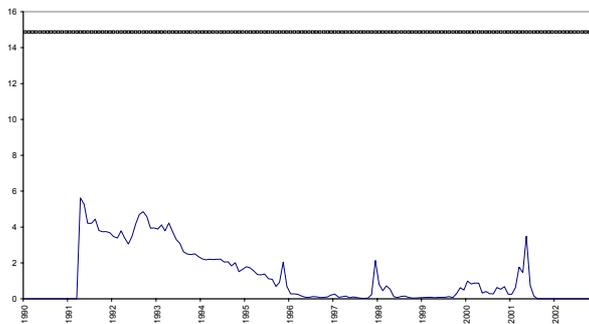
10.2. APPLICATION DES TESTS DE CHANGEMENT DU DEGRÉ DE COINTÉGRATION ET MODÈLES DE TRAFIC À LONG TERME

Les tests concluent que cette relation exprime un équilibre sur une première sous-période :

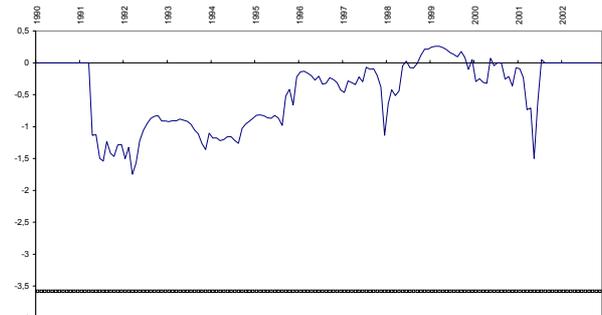
Tests de modification du degré de cointégration

$\ln(df\text{rersa})$	Statistique	VC 10%	VC 20%	Date	Conclusion
Absence de cointégration puis cointégration					
Wald supérieur	5.6161	17.275	14.870	mai 1991	CI(1)-CI(1)
Student inférieur	-1.7476	-3.887	-3.580	avril 1992	CI(1)-CI(1)
Cointégration puis absence de cointégration					
Wald supérieur	13.4641	14.877	12.599	janvier 2000	CI(0)-CI(1)
Student inférieur	-3.4260	-3.6810	-3.368	janvier 2000	CI(0)-CI(1)

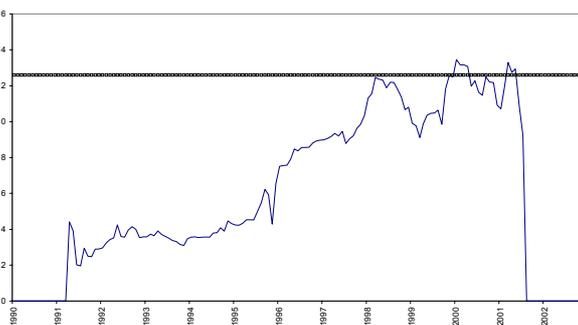
Wald supérieur pour cointégration en période 2



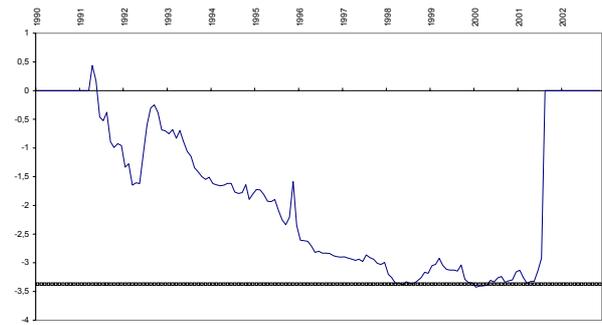
Student inférieur pour cointégration en période 2



Wald supérieur pour cointégration en période 1



Student inférieur pour cointégration en période 1



Les tests de modification du degré de cointégration ayant pour hypothèse alternative "cointégration puis absence de cointégration" ont tendance à s'apercevoir avec retard du changement de régime. La date de rupture retenue correspond à la première date pour laquelle la statistique du test de Student est un minimum local inférieur à la valeur critique à 20%, soit **juin 1998**.

10.2.8.2 Modélisation du trafic du dimanche/jour férié sur le RER pour chacune des sous-périodes

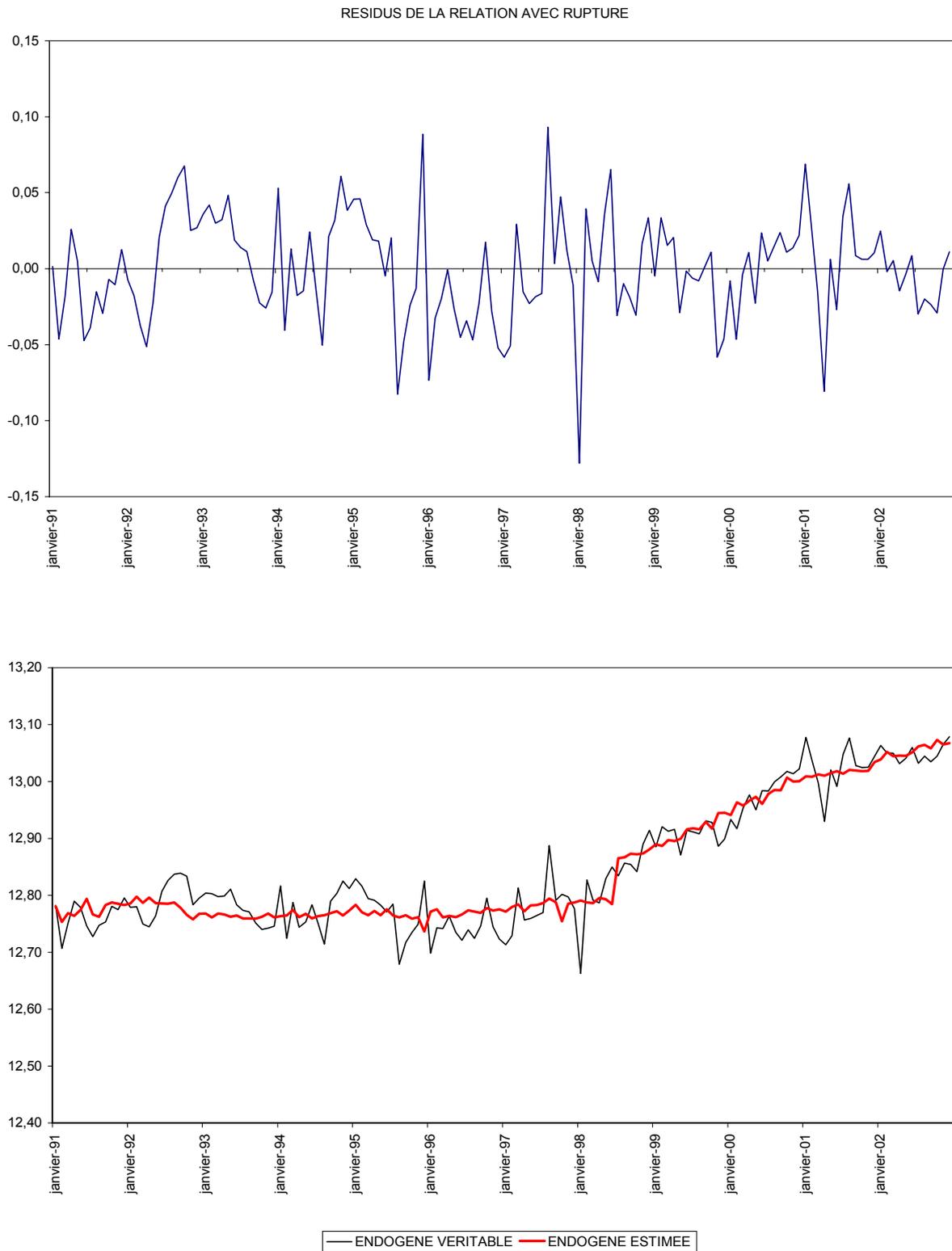
La meilleure relation de cointégration pour chaque sous-période dégagée est résumée dans le tableau suivant :

$\ln(df\text{rersa})$	Valeurs des coefficients				Test CADF			
Juin 1998	$gdmagrp$	$txhotidf$	c	t	ρ	t_{CADF}	$VC_{10\%}$	Concl.
Avant	n.s.	0.2610	12.6065	n.s.	0.4170	-4.5978	-3.0890	CO(0)
		(2.4995)	(190.4153)					
Après	0.0026		12.6147	0.0036	0.2894	-5,4205	-3,6062	CO(0)
	(2.2316)		(114.2653)	(11.2196)				

Le trafic du dimanche sur le RER a connu une rupture identique à celle observée le samedi sur le même sous-réseau. Ce type de trafic est en forte augmentation : aucune modélisation ne possède une tendance aussi élevée. Le premier élément à relier à l'augmentation de la fréquentation des transports collectifs le dimanche résulte de la création de la carte Imagine R qui permet aux scolaires et étudiants d'effectuer un nombre illimité de voyages, sans limite de zone le week-end. La création de ce titre répond à la mobilité croissante lors des week-ends, à mettre en lien avec le développement des loisirs en dehors du domicile à la suite de la coupe du monde de football et de la reprise économique. L'ouverture des commerces le dimanche dans cinq zones touristiques n'a pas été un évènement d'ampleur suffisante pour développer le transport dominical par le RER, mais il contribue à la tendance positive. De même, les abonnements de cinéma et la récente gratuité de certains musées parisiens ont contribué à accroître le nombre de déplacements le dimanche.

Cette augmentation du trafic du dimanche sur le RER est en partie compensée par le report de clientèle sur la ligne 14 du métro, comme l'indique la forte diminution de l'élasticité de ce type de trafic à la variable de tourisme prise en compte.

10.2. APPLICATION DES TESTS DE CHANGEMENT DU DEGRÉ DE COINTÉGRATION ET MODÈLES DE TRAFIC À LONG TERME



10.3 Dynamiques de court terme

Les modèles de long terme ayant été estimés dans la section précédente, éventuellement en prenant deux sous-périodes, nous calculons désormais les élasticités de court terme.

Ces équilibres de court terme sont estimés par les moindres carrés "totalement modifiés" sur la deuxième sous-période uniquement. Parmi les différentes modélisations possibles à court terme, l'équation est choisie en minimisant le critère d'information BIC, et en prenant au moins une variable économique explicative.

10.3.1 Trafic du jour ouvrable sur le métro

Force de rappel issue de la modélisation de long terme

$$z_t = \ln(jomsa)_t + 0.1131 \ln(defmidf)_t - 0.2099 txhotidf_t - 16.4762 - 0.0014 t \quad \text{si } t > \text{déc95}$$

Dynamique de court terme

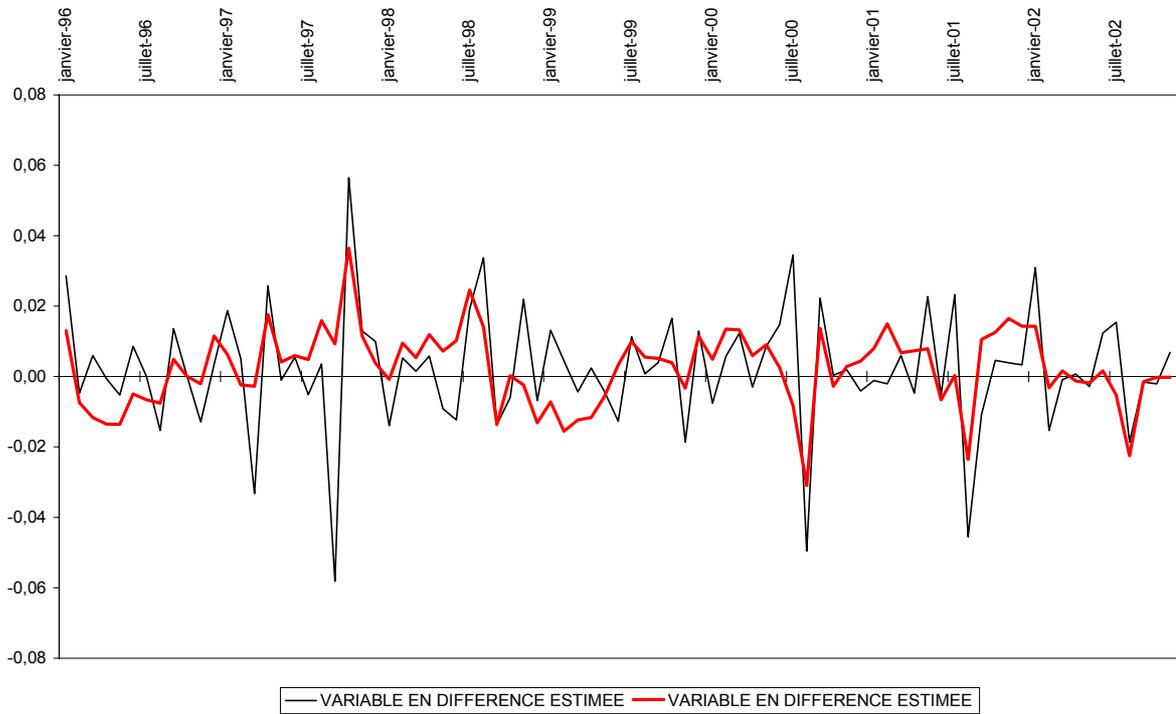
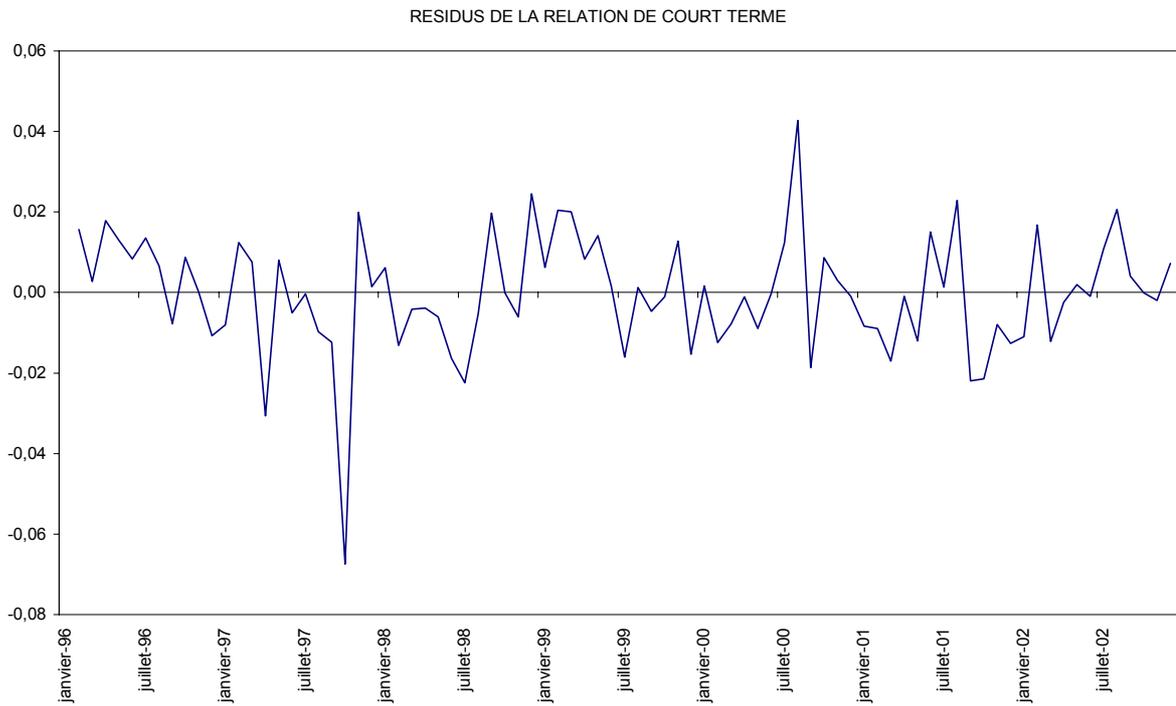
$$\Delta \ln(jomsa)_t = -0.0245 \Delta \ln(defmidf)_t + 0.0021 - 0.8384 z_{t-1}$$

(-0.2767) (1.1893) (-6.7841)

La principale explication des variations de court terme du trafic du jour ouvrable sur le métro correspond à la force de rappel vers la trajectoire de long terme. La forte significativité de cette force de rappel constitue une nouvelle preuve de l'existence d'une relation de cointégration. La constante, non significative, est un signe de la bonne qualité de l'estimation de l'équilibre de long terme.

Les variations des variables économiques jouent un rôle non significatif sur les variations du trafic considéré. La meilleure relation de court terme indique une très faible élasticité négative aux demandes d'emplois. La montée du chômage a un impact à long terme, mais joue relativement peu à court terme : une variation isolée du chômage n'a pas d'impact sur le trafic du jour ouvrable sur le métro.

10.3. DYNAMIQUES DE COURT TERME



10.3.2 Trafic du jour ouvrable sur le RER

Force de rappel issue de la modélisation de long terme

$$z_t = \ln(jorersa)_t + 0.2592 \ln(defmidf)_t \quad \text{si } t > \text{sept97} \\ -175236 - 0.0010t$$

Dynamique de court terme

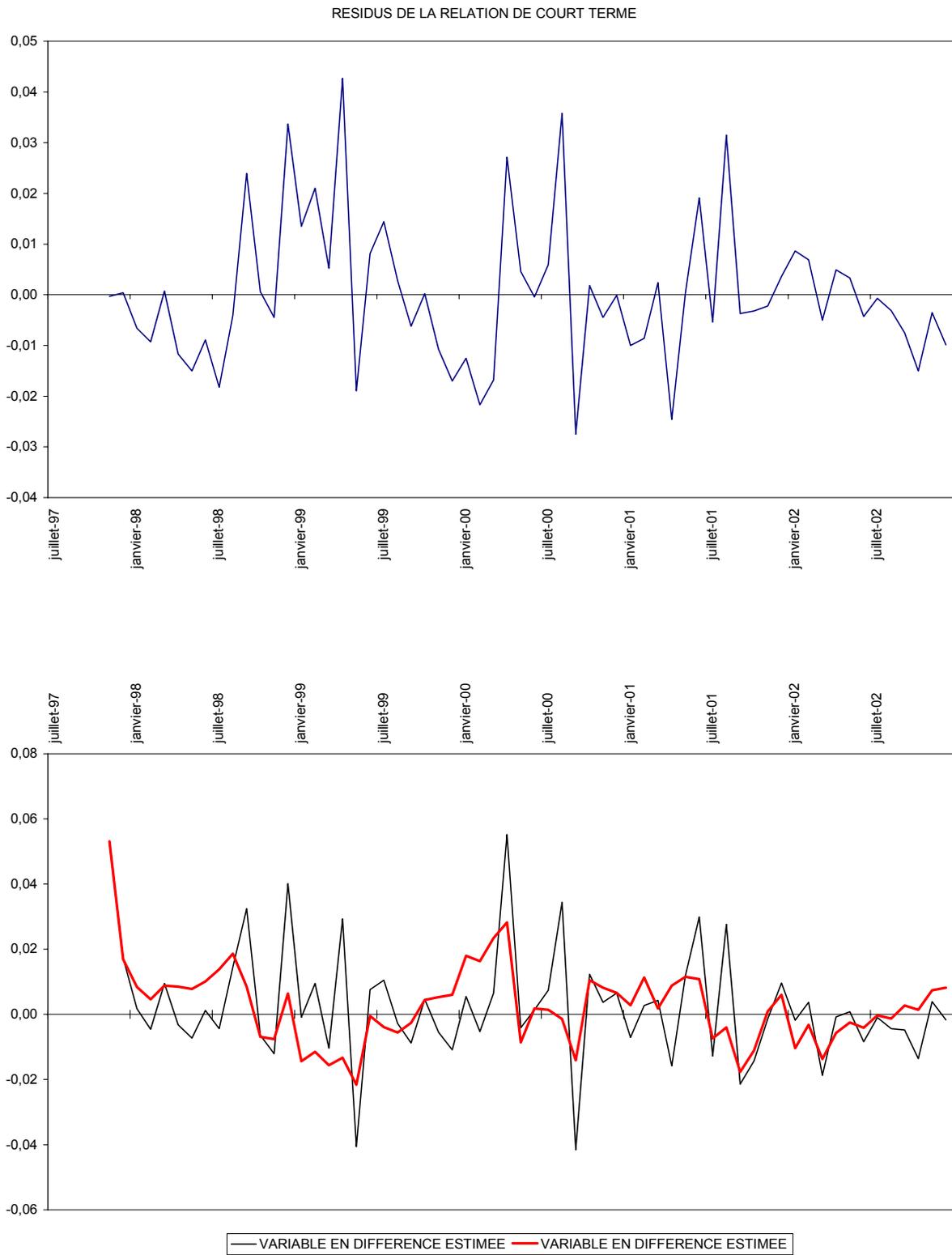
$$\Delta \ln(jorersa)_t = -0.2939 \Delta \ln(defmidf)_t + 0.0871 \Delta txhotidf_t \\ (-2.1147) \quad (1.3221) \\ + 0.0011 - 0.5238 z_{t-1} \\ (0.6327) \quad (-6.2475)$$

Contrairement au métro, une variation isolée des demandes d'emplois en fin de mois en Ile-de-France a un impact significatif sur le trafic du jour ouvrable sur le RER. L'effet est rémanent : les élasticités du trafic à cette variable sont identiques à court et à long terme. La modélisation met en lumière la prépondérance des liaisons domicile-travail sur ce sous-réseau.

Une variation isolée du taux d'occupation des hôtels en Ile-de-France entraîne une faible variation positive du trafic du jour ouvrable sur le RER. Cette élasticité traduit sans doute le fait que de nombreux touristes iront visiter le parc d'attraction Eurodisneyland Paris au cours de leur séjour, les Champs-Élysées ou la Défense (pour le tourisme d'affaires par exemple) s'ils viennent pour la première fois en Ile-de-France.

L'importance de la force de rappel à la trajectoire de long terme retardée ainsi que la non-significativité de la constante dans la relation de court terme signalent la qualité de l'estimation de la relation de cointégration. La dynamique de court-terme est ramenée vers la trajectoire d'équilibre.

10.3. DYNAMIQUES DE COURT TERME



10.3.3 Trafic du jour ouvrable sur le réseau des bus de Paris

Force de rappel issue de la modélisation de long terme La modélisation de long terme résulte de l'estimation sur la période entière :

$$z_t = \ln(joparsa)_t + 0.2116 \text{ circ75} + \\ 0.0775 \ln(speed75)_t - 0.1605 \ln(defmidf)_t \\ + 0.0012 \text{ gdmagrpt} - 0.0914 \text{ txhotidf}_t + 13.5304$$

Dynamique de court terme

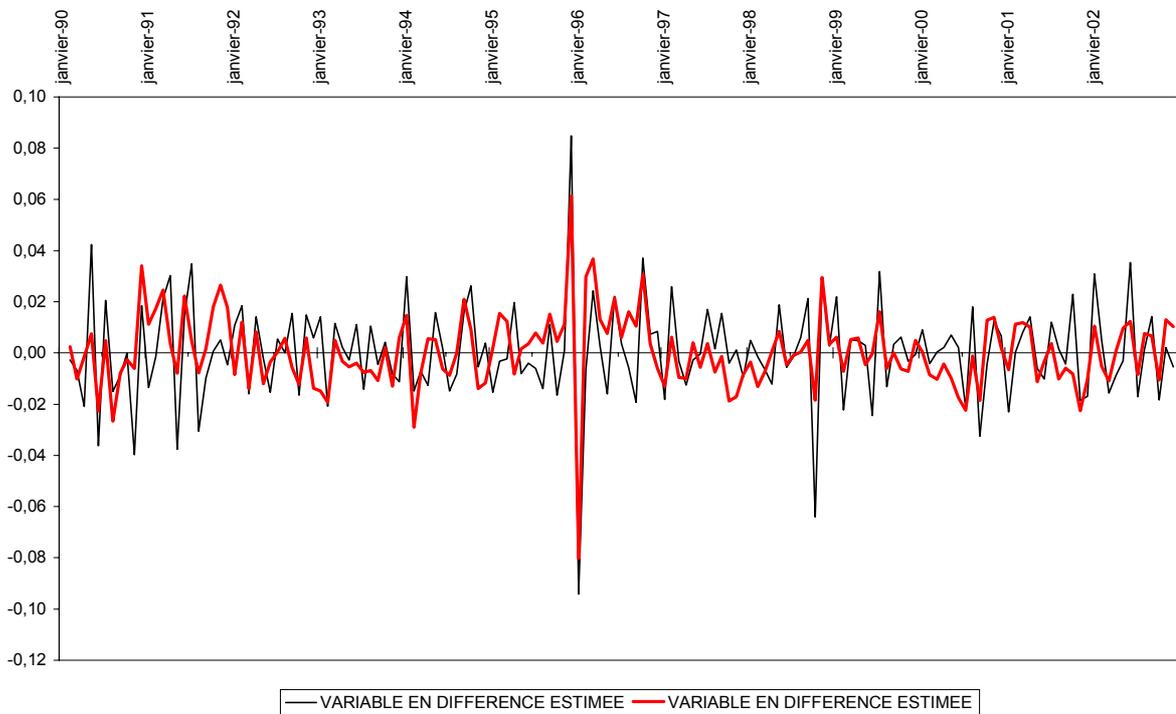
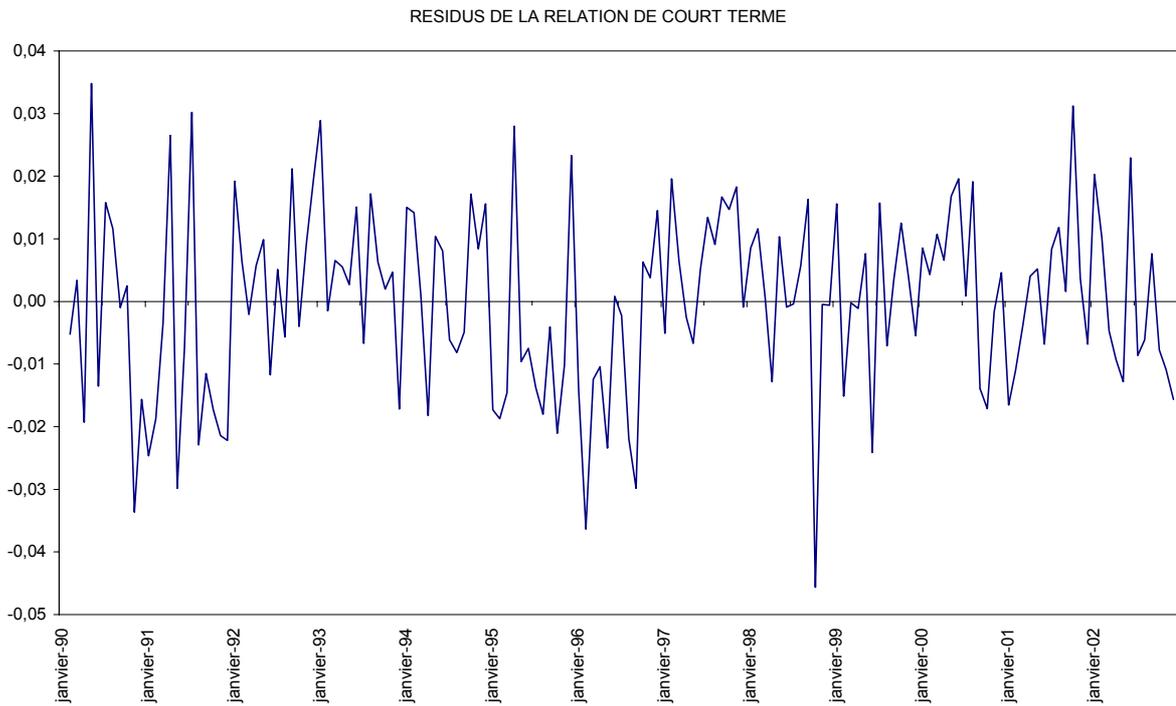
$$\Delta \ln(joparsa)_t = - 0.1556 \Delta \ln(speed75)_t + 0.0003 \Delta \text{gdmagrpt} \\ (-3.0466) \quad (1.5030) \\ + 0.0005 - 0.7719 z_{t-1} \\ (0.3940) \quad (-8.7450)$$

La faiblesse de la constante dans la régression de court terme ainsi que la forte significativité de la force de rappel indiquent que la relation de cointégration est correctement spécifiée. La dynamique de court terme est fortement ramenée vers la trajectoire de long terme.

Le nombre de variables explicatives à court terme est plus faible qu'à long terme. Le facteur principal des évolutions à court terme correspond à la vitesse de circulation dans Paris. Les embouteillages jouent un rôle plus important à court terme qu'à long terme : si les encombrements dissuadent la clientèle de prendre le bus, ils finissent par s'y habituer. A l'inverse, un accroissement de la vitesse de circulation (et donc de la vitesse des bus pour les portions hors couloirs de bus) est un facteur de croissance positive à court terme, mais ne modifie pas de façon durable le comportement des usagers.

Les achats dans les grands magasins de la région parisienne constituent un autre facteur explicatif des variations à court terme du trafic du jour ouvrable sur les autobus de Paris. En période de soldes, ou de promotion, par exemple, une partie de la clientèle utilise les autobus de Paris pour se rendre de leur lieu de travail aux grands magasins.

10.3. DYNAMIQUES DE COURT TERME



10.3.4 Trafic du jour ouvrable sur le réseau des bus de banlieue

Force de rappel issue de la modélisation de long terme La modélisation de long terme résulte de l'estimation sur la période entière :

$$z_t = \ln(\text{jobansa})_t + 0.07401 \ln(\text{defmidf}) - 0.0015 \text{gdmagrpt} \\ -15.0792 - 0.0018 t$$

Dynamique de court terme

$$\Delta \ln(\text{joparsa})_t = 0.0002 \Delta \text{gdmagrpt}_t + 0.0018 - 0.6124 z_{t-1} \\ (1.3642) \qquad (1.7363) \qquad (-8.1873)$$

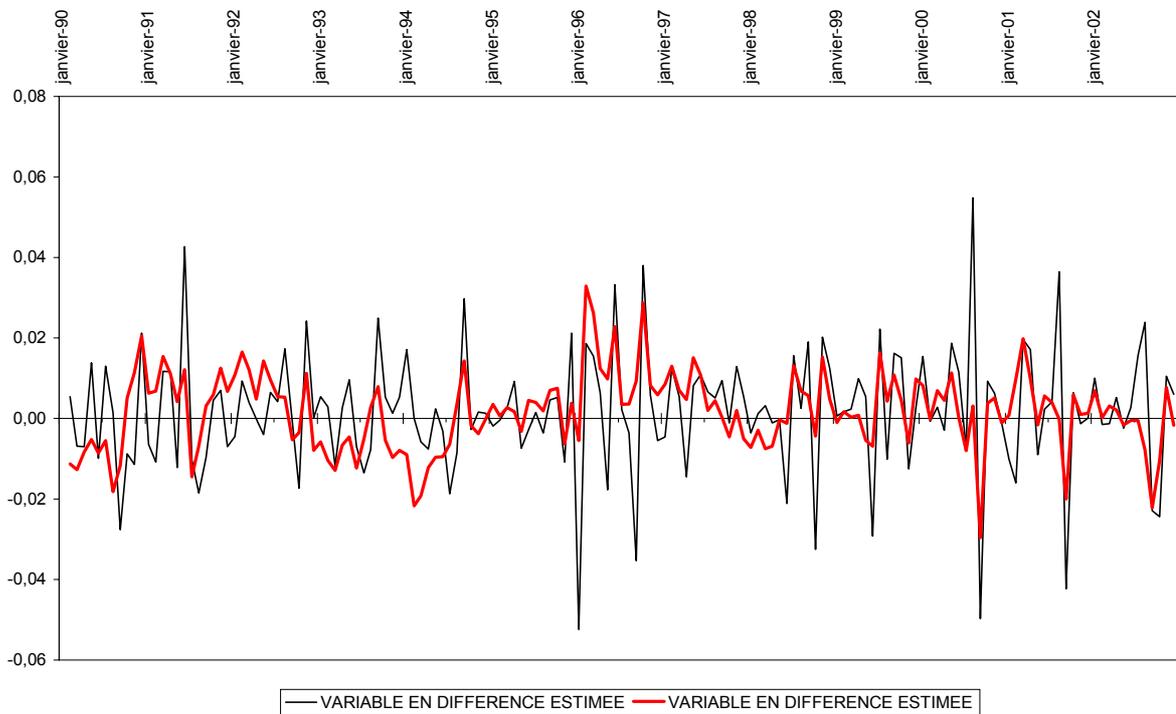
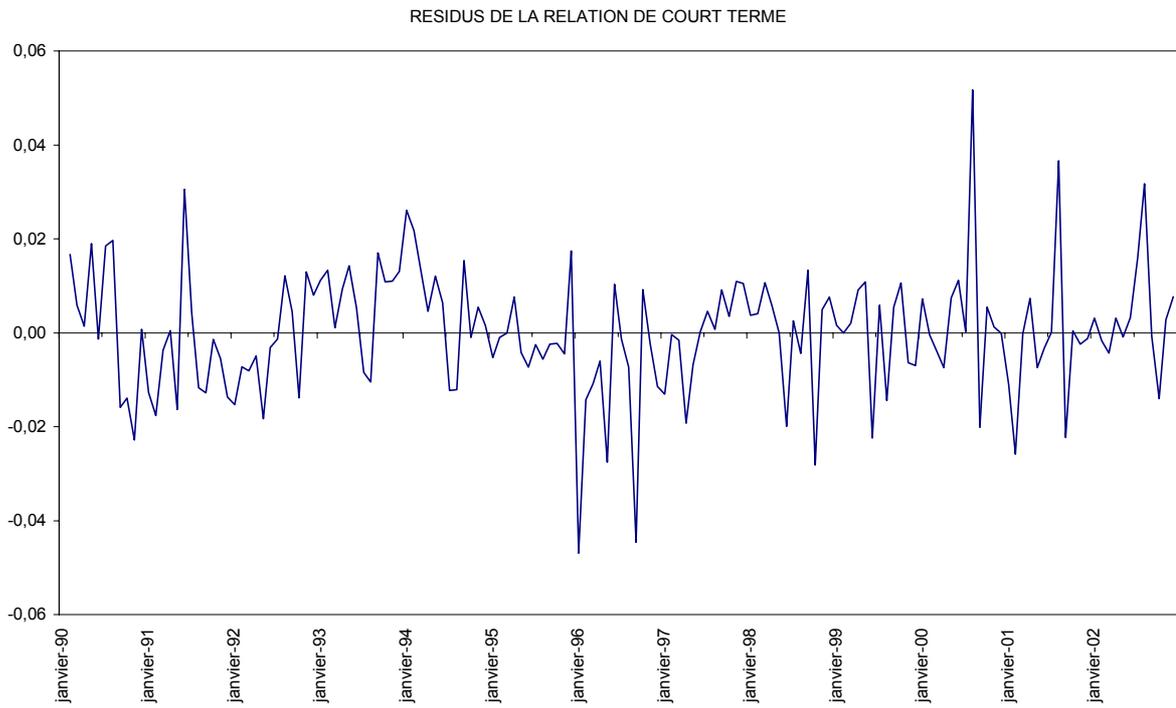
A court terme, seule la force de rappel à la trajectoire de long terme est une variable significative. Les variations de court terme du trafic du jour ouvrable sur le réseau des autobus de banlieue ont pour objectif principal de se rapprocher de l'équilibre de long terme. Encore une fois, la constante n'est pas significative.

Nous avons néanmoins décidé au départ de rechercher les relations de court terme en supposant la prise en compte d'au moins une variable représentative de l'économie francilienne.

La variable économique qui explique les variations ponctuelles de ce type de trafic correspond au volume de ventes dans les grands magasins de la région parisienne.

Cette élasticité positive est plus faible à court terme qu'à long terme. Il s'agit sans doute d'une mesure de l'afflux supplémentaire de voyageurs au moment des soldes ou des opérations spéciales (par exemple les 3J aux Galeries Lafayette, ou les 4S de la Samaritaine).

10.3. DYNAMIQUES DE COURT TERME



10.3.5 Trafic du samedi sur le métro

Force de rappel issue de la modélisation de long terme

$$z_t = \ln(smsa)_t - 0.0020 \text{ } gdmagrpt - 0.2747 \text{ } txhotidf_t - 14.3056 - 0.0022 \text{ } t \quad \text{si } t > \text{ } fev98$$

Dynamique de court terme

$$\begin{aligned} \Delta \ln(smsa)_t = & \quad 0.0014 \quad \Delta gdmagrpt + \quad 0.2258 \quad \Delta txhotidf_t \\ & (1.4946) \qquad \qquad \qquad (2.1830) \\ & - \quad 0.1359 \quad \text{ } compet + \quad 0.0032 \quad - \quad 0.7563 \quad z_{t-1} \\ & (-1.1076) \qquad \qquad \qquad (1.0884) \qquad \qquad (-5.5850) \end{aligned}$$

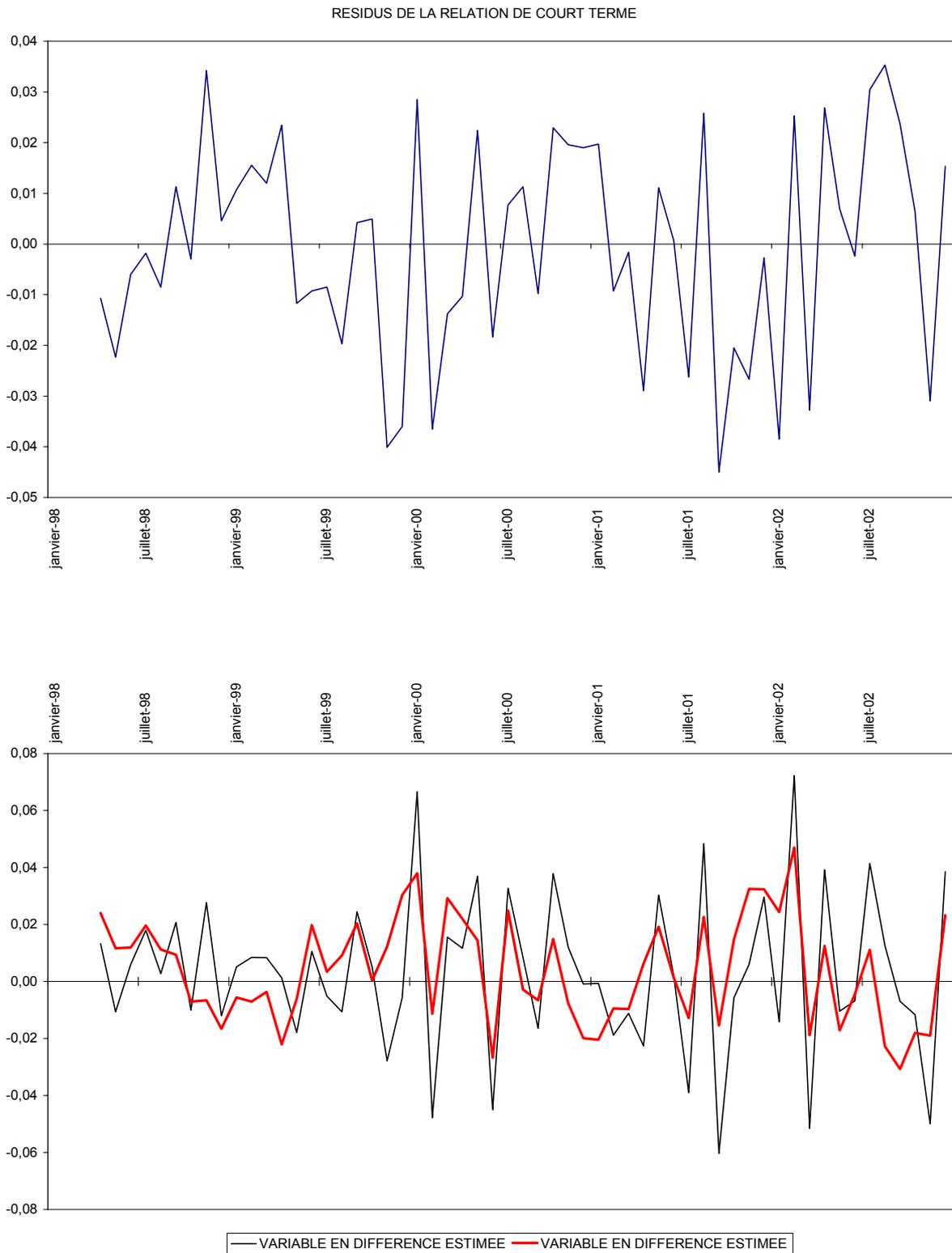
Plusieurs variables économiques expliquent les variations à court terme du trafic sur le métro au cours des samedis. La principale variable explicative à court terme correspond au taux d'occupation des hôtels en Ile-de-France, dont l'élasticité à court terme est presque aussi importante que l'élasticité de long terme. Un afflux ponctuel de touristes en Ile-de-France provoque un accroissement immédiat de trafic sur le métro, qui correspond au mode de transport collectif urbain favori dans Paris (par rapport au bus et au RER). En particulier, la ligne 1 du métro dessert les principaux sites touristiques.

L'activité dans les grands magasins parisiens a aussi un impact à court terme sur le trafic du métro, plus marqué le samedi que le jour ouvrable.

En outre, l'indice de "compétitivité" des transports collectifs urbains par rapport à l'automobile apparaît pour la première fois dans nos modélisations. Cet indice correspond au rapport du prix du ticket T au prix du super. Cette variable a un impact ponctuel, et assez faible, sur les déplacements non obligés (achats et loisirs), déplacements plus fréquents les jours chômés tels que le samedi.

Enfin, la force de rappel correspond une fois de plus à la variable la plus significative de la relation de court terme. La principale dynamique à court terme vise à retourner à l'équilibre de long terme.

10.3. DYNAMIQUES DE COURT TERME



10.3.6 Trafic du dimanche/jour férié sur le métro

Force de rappel issue de la modélisation de long terme

$$z_t = \ln(dfmsa)_t - 0.0010 \ gdmagrpt - 0.3340 \ txhotidf_t - 13.7923 - 0.0036 \ t \quad \text{si } t > fev98$$

Dynamique de court terme

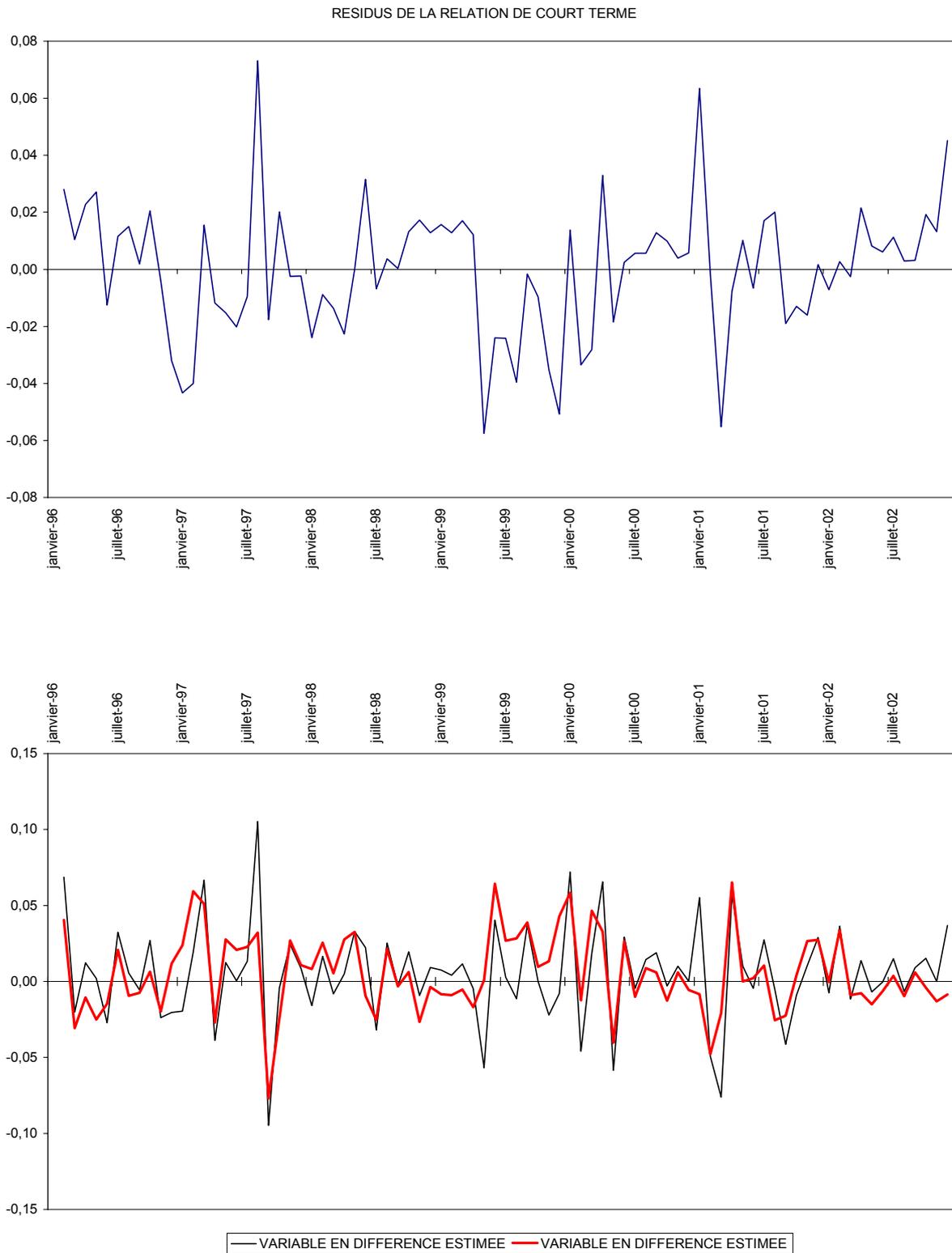
$$\begin{aligned} \Delta \ln(dfmsa)_t = & - 0.2975 \ \Delta \ln(defmidf)_t + 0.3069 \ \Delta txhotidf_t \\ & (-1.3612) \qquad \qquad \qquad (3.8723) \\ & + 0.0041 \ - \ 0.9641 \ z_{t-1} \\ & (1.3931) \qquad \qquad (-9.8638) \end{aligned}$$

A court terme, les variations de trafic sur le métro au cours des dimanches ou jours fériés s'expliquent en partie par les variations du nombre de demandes d'emplois en fin de mois. En période de reprise sur le marché de l'emploi, la fréquentation sur le métro augmente immédiatement pour ce jour non travaillé. Plusieurs facteurs concourent à cette élasticité positive. La première explication correspond à une reprise du "moral" des consommateurs, et de la consommation de loisirs. La deuxième explication possible réside dans le fait que pour de nombreuses personnes qui retrouvent un travail, leurs déplacements domicile-travail seront effectués à l'aide d'un abonnement de transport (carte Orange ou Intégrale), remboursé à moitié par leur employeur, et dont ils n'avaient pas forcément le besoin au cours de leur période non travaillée. Le fait d'avoir déjà un titre de transport favorise la prise des transports en commun pour des déplacements hors jour ouvrable.

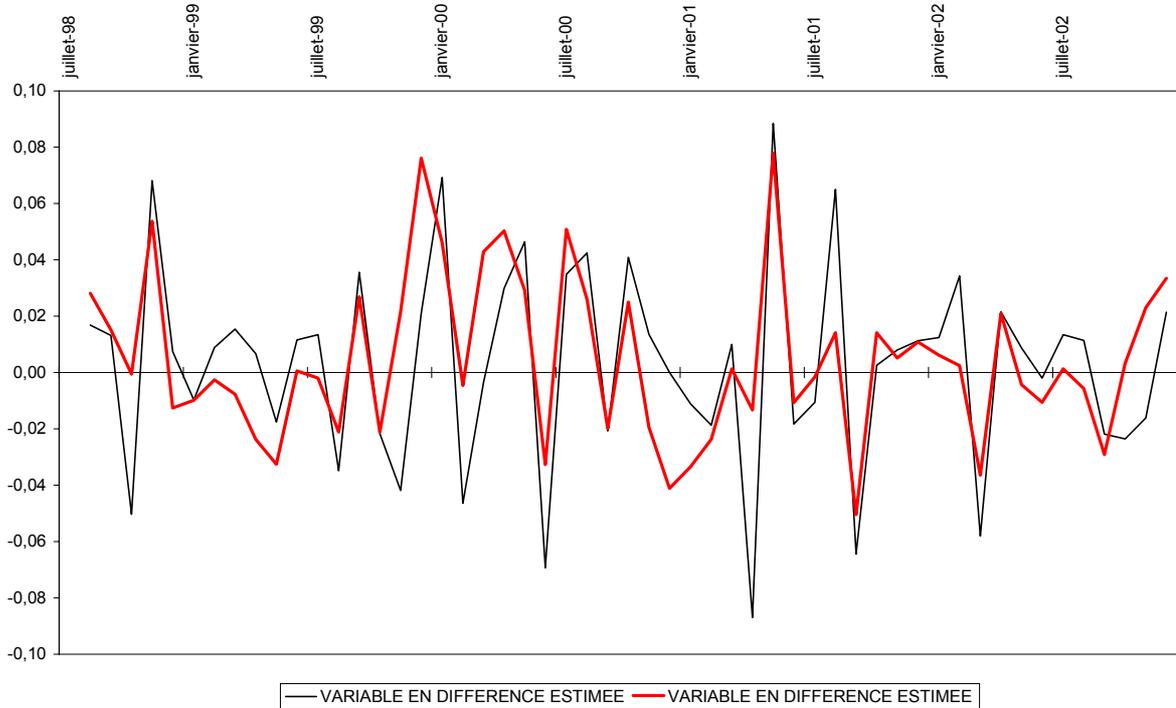
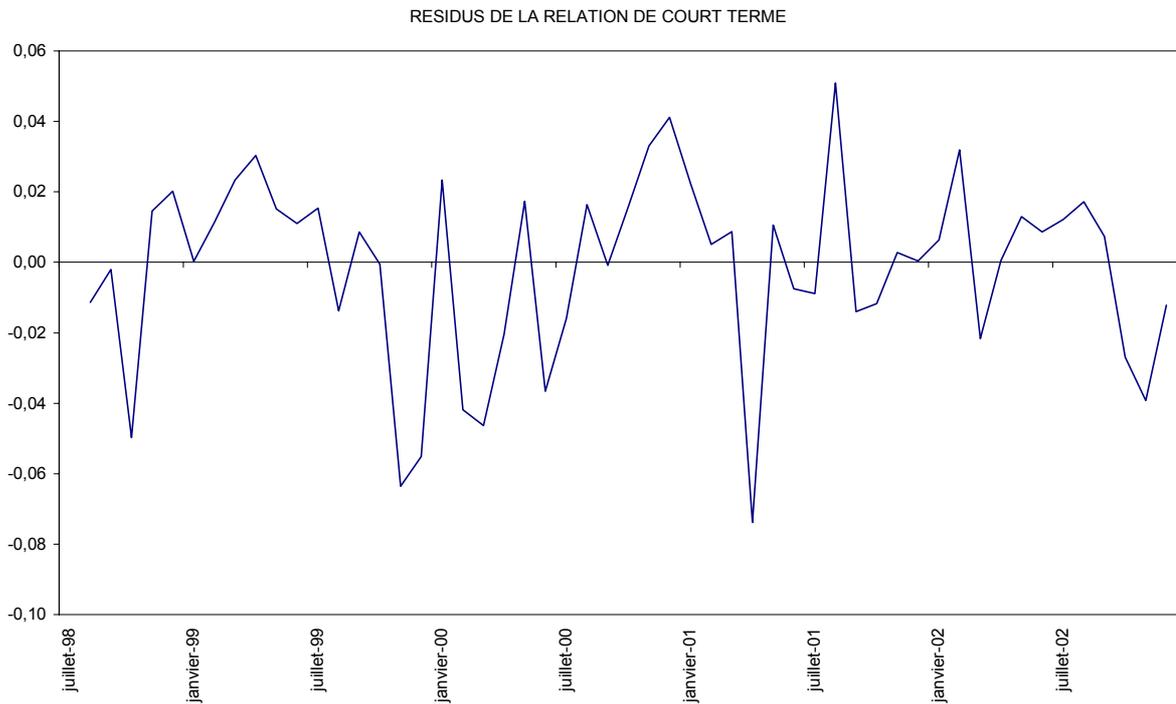
La modélisation du trafic du dimanche/jour férié sur le métro possède une élasticité positive significative au taux d'occupation des hôtels en Ile-de-France, presque aussi importante à court terme qu'à long terme.

Enfin, la significativité de la force de rappel confirme l'existence de la relation de cointégration représentant l'équilibre de long terme.

10.3. DYNAMIQUES DE COURT TERME



10.3. DYNAMIQUES DE COURT TERME



10.3.8 Trafic du dimanche/jour férié sur le RER

Force de rappel issue de la modélisation de long terme

$$z_t = \ln(dfrrersa)_t + 0.0026 gdmagrpt - 12.6147 - 0.0036t \quad \text{si } t > \text{fév}98$$

Dynamique de court terme

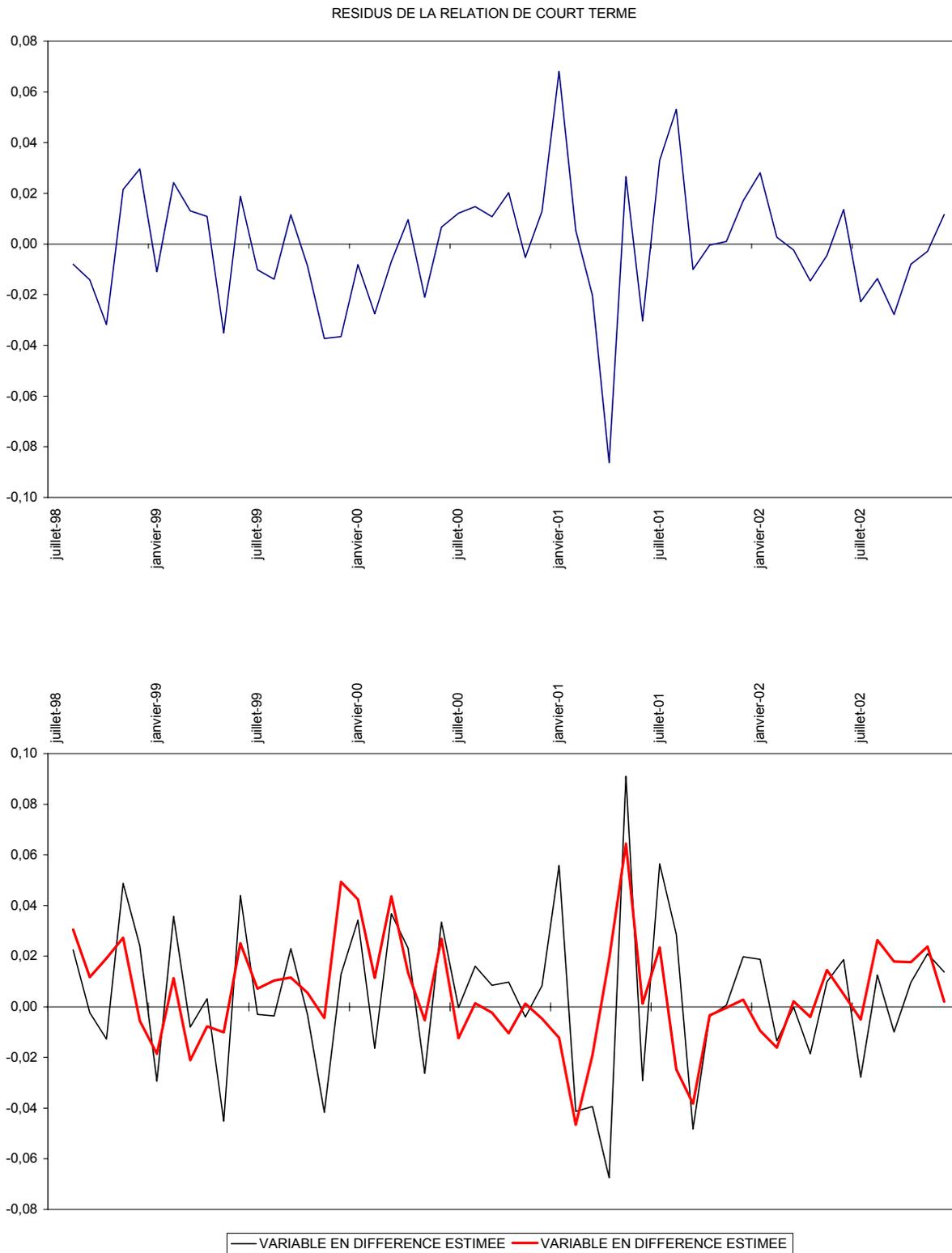
$$\Delta \ln(dfrrersa)_t = - \begin{matrix} 0.1872 \\ (-1.0523) \end{matrix} \Delta \ln(defmidf)_t + \begin{matrix} 0.0044 \\ (1.5348) \end{matrix} - \begin{matrix} 0.7599 \\ (-7.6284) \end{matrix} z_{t-1}$$

Comme c'était le cas sur le métro, la variable explicative des variations de court terme du trafic du dimanche ou jour férié sur le RER correspond aux demandes d'emplois en fin de mois en Ile-de-France. Bien que ce type de jour ne soit pas un jour ouvrable et que la majorité de la population francilienne ne travaille pas, la situation de l'emploi a un impact sur la fréquentation dominicale des sous-réseaux ferrés. La possession d'un abonnement de trafic est généralement associée à l'activité professionnelle (ou scolaire) et aux déplacements pendulaires qu'elle occasionne. L'achat d'un titre de la gamme Orange est en outre favorisé par le remboursement à la hauteur de 50% par l'employeur. Le fait de détenir un titre de transport modifie la décision individuelle du choix du mode de transport utilisé lors du déplacement. Il peut même favoriser le déplacement dans la mesure où celui-ci apparaît comme un transport gratuit !

L'élasticité aux demandes d'emplois en Ile-de-France est moins importante sur ce sous-réseau que sur le métro. Cette différence s'explique tout d'abord par la fréquence moins importante des trains sur le RER. En outre, le temps de transport sur le RER est généralement plus long que sur le métro, surtout pour les banlieusards. Le temps de transport sur ce sous-réseau est accru lors des week-ends ; alors que certains trains ne s'arrêtent pas dans certaines gares durant les heures de pointe des jours ouvrables, toutes les gares sont desservies le dimanche. Ces deux éléments conduisent à une augmentation du temps moyen de transport par rapport au reste de la semaine.

La force de rappel, très significative, confirme que l'on a bien une relation de cointégration.

10.3. DYNAMIQUES DE COURT TERME



Sur le réseau routier (bus de Paris et bus de banlieue), un jour ouvrable, un équilibre de long terme existe lorsque l'ensemble de la période janvier 1990 - décembre 2002 est prise en compte. Ce n'est pas le cas sur le réseau ferré, quel que soit le sous-réseau ou le type de journée considéré. L'application de nos tests de modification de cointégration a permis de montrer que pour tous les trafics journaliers, s'il n'existe pas d'équilibre sur l'ensemble de la période, il existe au moins un équilibre sur une sous-période :

- Sur le métro, les grèves de la fin de l'année 1995 ont modifié les comportements des utilisateurs de ce sous-réseau au cours d'un jour ouvrable ou d'un dimanche/jour férié. La création d'un équilibre a été plus tardive pour le trafic du samedi (début 1998).
- Sur le RER, il existe une relation d'équilibre à partir de la fin de l'année 1997 pour le jour ouvrable, et de juin 1998 pour chaque journée de week-end.

Cette analyse permet de déterminer les principaux facteurs de demande de transport en commun sur les sous-réseaux de la RATP, et de quantifier les impacts. Parmi les sous-réseaux de la RATP, les touristes privilégient le métro pour leurs déplacements : les élasticités du trafic du métro au taux d'occupation des hôtels en Ile-de-France sont comprises entre 0.21 et 0.33 selon le type de journée, alors qu'elles sont peu (souvent pas) significatives pour les autres types de trafic. La situation du chômage en Ile-de-France, insérée dans les modèles grâce à la variable de demandes d'emplois en fin de mois, joue un rôle dépressif sur le métro, le RER et le bus de banlieue, mais épargne le réseau des autobus de Paris, où la clientèle est différente. Enfin, l'essor de la consommation ou du lèche-vitrine, notamment avec l'ouverture de magasins pendant les dimanches, draine une importante clientèle sur le réseau ferré au cours des weeks-ends.

Conclusion à la troisième partie

Le caractère intégré à l'ordre 1 des séries sur tout ou partie de la période considérée, et les liens entre variables de trafic et variables environnementales, amènent à supposer l'existence de relations de cointégration. La présence d'une modification du processus de mémoire dans les séries (voir partie II) nous a conduit à dresser une synthèse des travaux existants sur la cointégration en présence de rupture. Les tests existants supposent une modification de la relation de cointégration mais négligent une possible modification du degré de cointégration (chapitre 8).

Pour pallier ce manque, nous avons construit des tests de modification du degré de cointégration qui s'appliquent dès que la relation testée n'est pas une relation cointégrée (chapitre 9). La construction de ces tests de modification du degré de cointégration correspond à une extension naturelle des tests de modification du degré d'intégration présentés au chapitre 6.

Si les tests d'intégration concluent que le résidu de cette relation non cointégrée est stationnaire sur une sous-période et intégrée sur l'autre, il n'existe pas d'équilibre de long terme sur l'ensemble de la période, mais il existe un équilibre sur une sous-période.

La construction de ces tests (hypothèses testées et calcul des statistiques de test) a été précisée, et les valeurs critiques ont été tabulées de manière à permettre un usage opérationnel de ces tests. En outre, les propriétés en terme de puissance et de qualité de l'estimation de la date de rupture ont été évaluées, et ont conduit à l'élaboration d'une procédure de test et d'estimation de la relation de long terme, éventuellement sur une seule sous-période.

Ces tests de modification du degré de cointégration ont ensuite été appliqués aux relations entre une variable de trafic et les variables décrivant l'environnement économique en Ile-de-

CHAPITRE 10. RECHERCHE DES ÉQUILIBRES DE LONG TERME ET DES DYNAMIQUES DE COURT TERME

France. Cette approche a permis de mettre en lumière des équilibres, valides sur une seconde sous-période uniquement pour la plupart des trafics. Seuls les trafics du jour ouvrable sur les sous-réseaux de bus (Paris et banlieue) possèdent un équilibre de long terme sur l'ensemble de la période, ce qui traduit le caractère progressif des éventuelles modifications des comportements des utilisateurs du réseau routier.

A long terme, les variations du trafic du métro sont expliquées par l'évolution du tourisme ; la situation sur le marché de l'emploi influence le niveau observé au cours des jours ouvrables, la consommation des ménages influe sur le trafic observé les week-ends. Sur le RER, l'impact du tourisme sur le trafic est moins important, les variations de trafic dépendant en grande partie de la consommation des ménages. Le prix du transport collectif (par rapport au prix de l'essence) n'apparaît pas comme un déterminant essentiel de l'utilisation des transports en commun, ce qui souligne l'existence d'une demande captive (car ne possédant pas de voiture par exemple) ou acquise aux transports collectifs (pour des raisons écologiques par exemple).

Les ajustements de court terme ont pour but principal de se ramener vers la trajectoire de long terme.

Quatrième partie

Prévisions

Introduction à la quatrième partie

Trois types de modélisation ont été élaborées dans cette thèse : une modélisation auto-régressive univariée avec racine unitaire sur une sous-période adéquate (chapitre 7) et, au sein de l'analyse de la cointégration, un équilibre de long terme ainsi qu'une dynamique de court terme (chapitre 10). Il est possible que la modélisation à retenir pour la prévision dépende de l'horizon de prévision considéré.

En l'absence de modification de l'environnement, la modélisation en fonction du passé de la variable ("modélisation auto-régressive") conduit à de bonnes prévisions. Elle sert à la fois à réaliser une première prévision des données de trafic, et à prévoir l'évolution des variables économiques qui apparaissent dans les modèles cointégrés de trafic. Elle est d'utilisation aisée et immédiatement opérationnelle. Toutefois, ce type de prévision a pour défaut majeur de ne pas permettre d'anticiper les retournements, en cas de modification des comportements économiques par exemple.

Pour prendre en compte les mutations économiques ou les changements conjoncturels, des variables économiques explicatives ont été ajoutées aux modèles de trafic, au sein d'une analyse de la cointégration. L'inconvénient de la prévision à partir d'un modèle cointégré réside dans le fait qu'en plus de prévoir les séries de trafic, les variables explicatives elles-mêmes doivent être modélisées. L'erreur de prévision comporte alors deux sources : l'erreur de modélisation du trafic par rapport aux variables explicatives d'une part, et l'erreur de la modélisation des variables explicatives d'autre part.

Le premier objectif de cette partie vise à répondre à la question suivante : "quelle modélisation utiliser pour prévoir?".

La partie IV sera décomposée en 2 chapitres :

- Le premier chapitre de la partie IV (chapitre 11) correspond à un exercice de rétro-prévision. Les modèles estimés sur la période janvier 1990 - décembre 2002 (ou sur une seconde sous-période uniquement) sont utilisés pour "prévoir" l'année 2003, dont les résultats sont d'ores et déjà disponibles. Grâce à des critères de qualité des rétro-prévisions (par rapport aux véritables données de trafic), il est possible de choisir un modèle de prévision pour chaque trafic, éventuellement selon l'horizon de prévision considéré. Le travail de rétro-prévision permet aussi de mesurer, pour une année écoulée, l'impact des différents facteurs économiques sur les variations de trafic.
- Enfin, le deuxième chapitre de cette partie (chapitre 12) est consacré aux prévisions réelles de trafic. Les estimations des modèles sont remises à jour avec les dernières données disponibles. Les prévisions ponctuelles ainsi que des intervalles de confiance "calibrés" seront calculés ou simulés à des horizons divers. Cette méthode ne s'avère toutefois pas totalement satisfaisante : les prévisions des variables environnementales sont issues de leur modélisation auto-régressive, i.e. fonction des valeurs passées, alors qu'un changement conjoncturel est annoncé pour les périodes à venir. Pour prendre en compte ce retournement conjoncturel éventuel, des scénarios seront introduits au cours de ce chapitre.

Chapitre 11

Des modèles cointégrés sur une sous-période plus adaptés aux prévisions

L'ensemble des modèles présentés jusqu'à présent ont été estimés sur une période prenant fin au mois de décembre 2002. Ces modèles de trafic pour un sous-réseau et un type de journée peuvent servir à prévoir les différents trafics pour les mois ultérieurs, en particulier l'année 2003. Les véritables données de l'année 2003 n'ont pas été prises en compte dans l'estimation, mais sont d'ores et déjà disponibles, permettant un travail de rétro-prévision.

Le début du chapitre sera consacré à l'analyse du trafic et de l'environnement en Ile-de-France au cours de l'année 2003, de manière à situer le cadre de l'étude de la qualité des rétro-prévisions. Dans la deuxième section, nous examinerons la qualité des prévisions issues des différentes modélisations pour chaque type de trafic, à l'aide de critères de qualité de prévision. Cette étape est un moyen de valider les différentes modélisations proposées dans la thèse dans la mesure où elles conduisent à des résultats satisfaisants par rapport à la situation réelle. Enfin, dans la troisième section, une synthèse de l'analyse des rétro-prévisions sera élaborée, de manière à évaluer l'apport de nos modélisations par rapport aux méthodes couramment employées au département Commercial de la RATP, et à obtenir des recommandations lors des exercices de

prévision.

11.1 Le bilan de l'année 2003

La fin de l'année 2002 était caractérisée par une montée des incertitudes, dans un contexte conjoncturel défavorable : stagnation de la croissance dans les pays de la zone euro, ralentissement de l'économie américaine, crise boursière, tensions géopolitiques avec risque de guerre en Irak. Le Centre d'Observation Economique de la Chambre de Commerce et d'Industrie de Paris prévoyait que la France se sortirait mieux de cet ensemble de phénomènes que ses partenaires européens. Les experts de l'INSEE tablaient sur une croissance lente au niveau national (+0.3% par trimestre). La progression de la consommation des ménages devait ralentir, du fait de l'érosion de leur pouvoir d'achat suite à la poursuite de la montée du chômage et du regain d'inflation induit par la hausse du pétrole.

La politique de restriction de l'usage de l'automobile mise en oeuvre par la mairie de Paris, ainsi que le renforcement de l'offre de service sur les différents réseaux de la RATP correspondaient à des éléments pouvant compenser en partie l'effet négatif sur le trafic de la RATP de la dégradation de la situation économique.

Le pessimisme de ces prévisions au niveau national n'a pas été infirmé par l'évolution des indicateurs économiques en Ile-de-France au cours de l'année 2003.

11.1.1 Evolution de la situation économique en Ile-de-France au cours de l'année 2003

L'année 2003 correspond à une année de décélération de la croissance économique dans les pays de la zone euro (taux de croissance égal à +0.5%). En France, le taux de croissance du PIB s'est élevé à seulement +0.2% : l'économie française n'a pas été en mesure de s'affranchir des effets du repli du cycle mondial. La dégradation économique générale n'a pas épargné la région Ile-de-France : elle aurait connu la récession avec -0.3% de croissance selon les estimations du Centre d'Observation Economique de la Chambre de Commerce et d'Industrie de Paris¹.

¹Les chiffres et analyses de ce paragraphe sont issus en grande partie des informations disponible sur les sites du CROCIS et du COE-CCIP.

CHAPITRE 11. DES MODÈLES COINTÉGRÉS SUR UNE SOUS-PÉRIODE PLUS ADAPTÉS AUX PRÉVISIONS

La diminution de la demande globale, notamment étrangère, ainsi que la forte appréciation de l'euro face au dollar, ont pénalisé le commerce extérieur de la région. Les exportations franciliennes ont subi un recul continu au cours de l'année (-8% au premier trimestre, -2.5% au quatrième trimestre), tandis que les importations se contractaient (-4.8% en évolution annuelle). La production industrielle a en outre souffert des multiples jours fériés et mouvements sociaux au cours du mois de mai.

La consommation des ménages jouait le rôle de moteur de la croissance depuis 1996. L'année 2003 est marquée par un essoufflement de cette consommation : le volume des ventes dans les grands magasins a régulièrement diminué, à l'exception des périodes de rentrée scolaire et de fêtes de fin d'année. La limitation des dépenses s'explique par la baisse de moral suite à l'accroissement du chômage, mais aussi par l'augmentation perçue de l'inflation.

La morosité de la conjoncture a lourdement pesé sur le marché du travail en Ile-de-France : le taux de chômage francilien a rejoint le taux de chômage national (9.7% de la population active à la fin du quatrième trimestre 2003), alors qu'il est traditionnellement inférieur à celui-ci. Le nombre de demandeurs d'emplois s'est accru régulièrement au cours de l'année, à un rythme plus rapide qu'en France (+6% au cours du premier trimestre contre +3% sur la France entière). L'emploi industriel a reculé dans les entreprises industrielles, mais aussi dans le secteur tertiaire (domaines de l'ingénierie informatique ou du conseil pour les affaires et la gestion), phénomène nouveau dans la région. En une année, l'Ile-de-France a perdu près de 55 000 emplois salariés (soit -1.4%, perte régionale la plus marquée).

En outre, 2003 a été une année particulièrement difficile pour le tourisme. Le ralentissement économique mondial, l'épidémie de SRAS², l'appréciation rapide de l'euro face au dollar, les tensions géopolitiques telles que le conflit irakien expliquent la diminution du nombre de touristes (surtout américains et asiatiques) en Ile-de-France. La canicule au cours de l'été a découragé les touristes français qui auraient souhaité effectuer un séjour en Ile-de-France.

En termes de conjoncture immobilière, enfin, la crise redoutée n'a pas eu lieu. Depuis deux ans, la tendance est au déplacement des ménages de Paris vers les communes limitrophes, de

²Le sigle "SRAS" signifie "Syndrome Respiratoire Aigu Sévère", une maladie apparue initialement sous forme de foyers dans les régions d'Asie du Sud-Est. Cette maladie infectieuse transmissible représente une menace pour la santé publique au niveau international.

la Petite Couronne vers la Grande Couronne, et des villes résidentielles vers les communes les moins favorisées.

En fin d'année, quelques signes d'embellie sont apparus en Ile-de-France : après le mois d'octobre 2003, les chefs d'entreprises reprennent progressivement confiance dans la conjoncture, bien que la prudence et l'attentisme demeurent de mise.

11.1.2 Evolution des trafics de la RATP en 2003

Avec plus de 2.6 milliards de voyages réalisés en 2003, le trafic brut de la RATP a diminué de -2% par rapport à 2002. Cette évolution est en rupture avec la progression régulière du trafic de la RATP observée depuis 1996. L'incidence défavorable de la morosité de la conjoncture économique, conjuguée à divers éléments pénalisants, expliquent la diminution de l'usage des transports en commun. En particulier, le conflit en Irak ou encore la canicule lors de l'été 2003 ont eu un impact négatif sur l'utilisation des transports collectifs, par le biais de la baisse de la fréquentation touristique. Le second trimestre a été pénalisé par les arrêts de travail observés dans les entreprises de transport en commun ou dans la fonction publique. En fin d'année, la reprise progressive de la fréquentation a permis de retrouver un niveau mensuel de trafic comparable à celui observé en fin 2002.

Les sous-réseaux ont été affectés de façon différente par la désaffection de la clientèle :

- La diminution du trafic a été plus sensible sur les lignes des autobus de Paris (-3.3%) et sur le métro (-2.8%) que sur le RER (-1.2%) et le Réseau Régional de Surface (-0.9%), le trafic a par ailleurs augmenté légèrement sur le réseau des autobus de banlieue (+0.6%).
- Le trafic du réseau ferré a diminué dès le mois de février pour atteindre en août son niveau le plus bas. A l'automne, le trafic se redresse pour retrouver en fin d'année le niveau atteint en 2002. La chute de trafic sur le réseau des autobus (Paris + banlieue) a été plus tardive (juin), du fait d'un report des sous-réseaux souterrains vers le réseau de surface : les cinq premiers mois de l'année 2003 sont caractérisés par une progression du trafic. Les problèmes de circulation en fin d'année, mais aussi la baisse de la fréquentation touristique ont pénalisé les autobus de Paris, qui constituent le seul sous-réseau à ne pas avoir bénéficié d'une reprise de trafic au quatrième trimestre.

- Le trafic du jour ouvrable sur l'ensemble des sous-réseaux a été plus affecté (-0.5%) que le trafic des week-ends (+0.1%).

11.2 Les rétro-prévisions de trafic sur l'année 2003

Pour chaque type de trafic (par type de journée et par sous-réseau), trois rétro-prévisions ont été effectuées, pour l'ensemble des mois de l'année 2003 :

- les rétro-prévisions à partir de la modélisation auto-régressive,
- les rétro-prévisions à partir de l'équation de long terme et des véritables données décrivant l'environnement économique en Ile-de-France,
- les rétro-prévisions à partir de l'équation de court-terme et de ces mêmes données environnementales.

Dans le premier paragraphe, la construction des prévisions ponctuelles, i.e. des trajectoires prévues, sera détaillée. Les trois types de modélisation de chaque trafic (modélisation auto-régressive, équilibre de long-terme, dynamique de court terme) seront abordés, en supposant que les variables explicatives sont connues (pour les "rétro-prévisions") ou prévues à l'aide d'une équation auto-régressive ("prévisions").

Ces différentes prévisions ont été comparées aux valeurs de trafic effectivement observées, grâce à des indicateurs de mesure de la qualité de la prévision, présentés dans le deuxième paragraphe.

Les résultats seront détaillés pour les rétro-prévisions du trafic du jour ouvrable sur le métro. L'analyse sera plus succincte pour les trafics relatifs à d'autres types de journée et/ou d'autres sous-réseaux.

11.2.1 Construction des prévisions ponctuelles

11.2.1.1 Prévisions d'une variable auto-régressive intégrée à l'ordre 1

Les prévisions auto-régressives ont deux utilités :

- Les résultats sur les données de trafic donnent une première idée de l'évolution future de ces variables,

- Les résultats sur les données décrivant l'environnement économique en Ile-de-France sont utilisés dans le cadre de prévisions calculées grâce aux modélisations issues de l'analyse de la cointégration.

La majorité des séries log-linéarisées utilisées sont intégrées à l'ordre 1 en seconde sous-période, ou sur la période entière ³. La modélisation auto-régressive a été réalisée sur la série log-linéarisée en différence première (stationnaire), en choisissant le nombre de retards p par une procédure de vraisemblance pénalisée BIC.

Pour faciliter les notations, posons

$$y_t = \ln(x_t)$$

La modélisation auto-régressive d'ordre p de la variable en différence première s'écrit :

$$\Delta y_t = c + a_1 \Delta y_{t-1} + a_2 \Delta y_{t-2} + \dots + a_p \Delta y_{t-p} + \varepsilon_t$$

où ε_t est un processus d'innovation de moyenne nulle et de variance σ_ε^2 . Le processus Δy_t est stationnaire.

Le modèle se réécrit en niveau

$$\begin{aligned} y_t &= c + (a_1 + 1) y_{t-1} + (a_2 - a_1) y_{t-2} + \dots + (a_p - a_{p-1}) y_{t-p} - a_p y_{t-p-1} + \varepsilon_t \\ &= c + \sum_{i=1}^{p+1} (a_i - a_{i-1}) y_{t-i} + \varepsilon_t \end{aligned}$$

où

$$a_0 = -1 \text{ et } a_{p+1} = 0$$

³Les seules exceptions sont les suivantes :

- la vente de carburant en Ile-de-France est stationnaire en seconde sous-période,
- le trafic grandes lignes dans les gares de Paris ainsi que le taux d'occupation des hôtels dans Paris sont stationnaires autour d'un terme déterministe ayant subi une rupture.

Ces séries n'ont pas été utilisées lors des modélisations.

soit

$$y_t = c + \sum_{i=1}^{p+1} \phi_i y_{t-i} + \varepsilon_t \quad (11.1)$$

A h pas dans le futur, la réalisation de ce processus est égale à

$$y_{t+h} = c + \sum_{i=1}^{p+1} \phi_i y_{t+h-i} + \varepsilon_{t+h} \quad (11.2)$$

et la meilleure prévision à h pas dans le futur est donnée par

$$y_{AR,t}^P(h) = c + \sum_{i=1}^{p+1} \phi_i y_{AR,t}^P(h-i) \quad (11.3)$$

où $y_{AR,t}^P(h-i) = y_{t+h-i}$ si $h \leq i$

La prévision à h pas associée à la série initiale x_t est retrouvée par la simple transformation exponentielle :

$$x_{AR,t}^P(h) = \exp y_{AR,t}^P(h)$$

11.2.1.2 Prévisions à l'aide de l'équation de long terme issue de l'analyse de la cointégration

La prévision à l'aide de l'équation de long terme L'équation de long terme avec tendance issue de l'analyse de la cointégration se présente sous la forme suivante :

$$y_{LT,t} = c_{LT} + \delta_{LT} \cdot t + w_t \beta_{LT} + \varepsilon_{LT,t}$$

où c_{LT} représente la constante dans l'équation de long terme, δ_{LT} correspond au paramètre associé à la tendance, et β_{LT} est l'ensemble des paramètres associés aux variables de long terme w_t . Les variables y_t et w_t sont des variables log-linéarisées et intégrées à l'ordre 1.

A h pas dans le futur, la réalisation de ce processus est égale à

$$y_{LT,t+h} = c_{LT} + \delta_{LT}(t+h) + w_{t+h}\beta_{LT} + \varepsilon_{LT,t+h} \quad (11.4)$$

et la meilleure prévision à h pas dans le futur est donnée par

$$y_{LT,t}^P(h) = c_{LT} + \delta_{LT}(t+h) + w_t^P(h)\beta_{LT} \quad (11.5)$$

ce qui suppose de prévoir les variables explicatives w_t dans le futur.

Les modèles ont été réalisés sur des séries transformées par le logarithme népérien. Pour retrouver les véritables séries de trafic, la transformation exponentielle est appliquée :

$$x_{LT,t}(h) = \exp(y_{LT,t}(h))$$

Rétro-prévisions grâce à l'équation de long terme lorsque les valeurs des variables explicatives sont connues Dans le cadre de rétro-prévisions, les variables $w_t(h)$ sont connues et égales à w_{t+h} . Un premier travail consiste à évaluer la qualité des prévisions des séries de trafic $y_{LT,t}$ à h pas dans le futur lorsque les véritables valeurs des variables explicatives sont connues.

$$y_{LT,t}^P(h) = c_{LT} + \delta_{LT}(t+h) + w_{t+h}\beta_{LT} \quad (11.6)$$

Prévisions grâce à l'équation de long terme lorsque les variables explicatives sont prévues grâce à leur modélisation auto-régressive En situation réelle, les véritables valeurs des variables explicatives ne sont pas observées et doivent être prévues. Les prévisions auto-régressives des variables sont utilisées à cette effet. A long terme, le trafic pour un type de journée et un sous-réseau à l'horizon h est obtenu en calculant

$$y_{LT,AR,t}(h) = c_{LT} + \delta_{LT}(t+h) + w_{AR,t}^P(h)\beta_{LT} \quad (11.7)$$

où les prévisions des variables explicatives sont réalisées à partir de la modélisation auto-régressive en différence réécrite sous la forme :

$$w_{AR,t}^P(h) = c + \sum_{i=1}^{p+1} \phi_i w_{AR,t}^P(h-i) \quad (11.8)$$

où $w_{AR,t}^P(h-i) = w_{t+h-i}$ si $h \leq i$

11.2.1.3 Rétro-prévisions à l'aide de l'équation de court terme (i.e. du modèle à correction d'erreur)

La prévision à l'aide de l'équation de court terme (modèle à correction d'erreur)

L'équation de court terme issue de l'analyse de la cointégration se présente sous la forme suivante :

$$\Delta y_{CT,t} = c_{CT} + \Delta w_t \beta_{CT} + \alpha z_{t-1} + \varepsilon_{CT,t}$$

avec

$$\begin{aligned} z_t &= y_t - c_{LT} - \delta_{LT} \cdot t - w_t \beta_{LT} \\ &= y_t - y_{LT,t} = \varepsilon_{LT,t} \end{aligned}$$

où z_{t-1} correspond à la force de rappel (erreur de prévision à long terme à la date précédente), c_{CT} est la constante de la dynamique de court terme. Cette équation se réécrit en niveau :

$$y_{CT,t} = y_{CT,t-1} + c_{CT} + \Delta w_t \beta_{CT} + \alpha z_{t-1} + \varepsilon_{CT,t}$$

A l'horizon h , cette équation s'écrit

$$y_{CT,t+h} = y_{CT,t+h-1} + c_{CT} + \Delta w_{t+h} \beta_{CT} + \alpha z_{t+h-1} + \varepsilon_{CT,t+h} \quad (11.9)$$

et la meilleure prévision à h pas est donnée par

$$y_{CT,t}^P(h) = y_{CT,t}^P(h-1) + c_{CT} + \Delta w_t^P(h) \beta_{CT} + \alpha \mathbf{E}[z_t(h-1)]$$

soit

$$\begin{aligned} y_{CT,t}^P(1) &= y_t + c_{CT} + \Delta w_t^P(1) \beta_{CT} + \alpha z_t && \text{pour } h = 1 \\ y_{CT,t}^P(h) &= y_{CT,t}^P(h-1) + c_{CT} + \Delta w_t^P(h) \beta_{CT} + \alpha \left[y_{CT,t}^P(h-1) - y_{LT,t}^P(h-1) \right] && \text{pour } h \geq 2 \end{aligned} \quad (11.10)$$

Rétro-prévisions grâce à l'équation de court terme lorsque les valeurs des variables explicatives sont connues Lorsque les variables utilisées dans la modélisation cointégrée sont connues, il n'est pas nécessaire de prévoir les variables explicatives à la date $t+h$. Supposons que les valeurs de trafic ne sont pas observées, de même que les erreurs de long terme futures.

Le modèle de prévision est :

$$\begin{aligned} y_{CT,t}^P(1) &= y_t + c_{CT} + \Delta w_{t+1} \beta_{CT} + \alpha z_t \\ y_{CT,t}^P(h) &= y_{CT,t}^P(h-1) + c_{CT} + \Delta w_{t+h} \beta_{CT} + \alpha \left[y_{CT,t}^P(h-1) - y_{LT,t}^P(h-1) \right] \text{ pour } h \geq 2 \end{aligned} \quad (11.11)$$

Prévisions grâce à l'équation de court terme lorsque les valeurs des variables explicatives sont issues des prévisions auto-régressives Le modèle de prévision est :

$$\begin{aligned} y_{CT,AR,t}^P(1) &= y_t + c_{CT} + \Delta w_{AR,t}^P(1) \beta_{CT} + \alpha z_t \\ y_{CT,AR,t}^P(h) &= y_{CT,t}^P(h-1) + c_{CT} + \Delta w_{AR,t}^P(h) \beta_{CT} + \alpha \left[y_{CT,AR,t}^P(h-1) - y_{LT,AR,t}^P(h-1) \right] \text{ pour } h \geq 2 \end{aligned} \quad (11.12)$$

où $w_{AR,t}(h)$ correspond à la modélisation auto-régressive des variables w_t à l'horizon h :

$$\begin{aligned} w_{AR,t}^P(h) &= c + \sum_{i=1}^{p+1} \phi_i w_{AR,t}^P(h-i) \\ \text{où } w_{AR,t}^P(h-i) &= w_{t+h-i} \text{ si } h \leq i \end{aligned}$$

11.2.2 Critères de comparaison de la qualité des rétro-prévisions

Plusieurs critères doivent être pris en compte lors de l'évaluation de la qualité des rétro-prévisions :

- L'absence de biais systématique.

En moyenne, les erreurs de prévision doivent être nulles, i.e. le niveau moyen de la série estimée doit être égal au niveau moyen observé. Les erreurs de prévision à chaque date se compensent sur l'ensemble de la période considérée.

- La variance des erreurs de prévisions doit être minimale. Si deux modèles ont une erreur moyenne nulle, le meilleur modèle est celui qui possède des erreurs de prévision proches de 0 à chaque horizon de prévision, i.e. celui pour lequel les valeurs prévues sont concentrées autour des véritables valeurs observées.
- La covariance entre la véritable série et la série estimée doit être la plus grande possible en valeur absolue. Ce dernier critère mesure la capacité du modèle de prévision à anticiper les retournements de la série.

Soient

- $\{y_h\}_{h=t+1, \dots, t+h_{\max}}$ la série des véritables données observées pour toute date future $h = t + 1, \dots, t + h_{\max}$, où h_{\max} correspond à l'horizon futur maximal prévu,
- $\{y_h^P\}_{h=t+1, \dots, t+h_{\max}}$ la série des données prévues pour tout horizon futur $h = t + 1, \dots, t + h_{\max}$,
- $\{e_h\}_{h=t+1, \dots, t+h_{\max}}$ la série des erreurs de prévision pour tout horizon futur $h = t + 1, \dots, t + h_{\max}$, avec $e_h = y_h^P - y_h$.

11.2.2.1 La racine de l'erreur quadratique moyenne

La racine carrée de l'erreur quadratique moyenne (ou "Root Mean Squared Error" en langue anglaise), est définie par :

$$RMSE = \sqrt{\frac{1}{h_{\max}} \sum_{h=t+1}^{t+h_{\max}} (y_h^P - y_h)^2} \quad (11.13)$$

Ce critère a pour inconvénient majeur l'existence d'un facteur d'échelle, du fait de l'absence de normalisation : s'il permet de comparer plusieurs modèles de prévision pour un type précis de trafic, cet outil ne pourra pas servir à la comparaison des prévisions pour différents types de trafic. En outre, selon Christoffersen et Diebold (1997 [40]), ce critère valorise mal la prévision à long terme dans le cadre d'un système cointégré.

11.2.2.2 La racine de l'erreur quadratique moyenne normalisée

Pour obtenir un critère normalisé, nous avons modifié la statistique de la façon suivante :

$$\begin{aligned} RMSE\% &= \sqrt{\frac{1}{h_{\max}} \sum_{h=t+1}^{t+h_{\max}} \left(\frac{y_h^P - y_h}{y_h} * 100 \right)^2} \\ &= \sqrt{\frac{1}{h_{\max}} \sum_{h=t+1}^{t+h_{\max}} \left(\frac{e_h}{y_h} * 100 \right)^2} \end{aligned} \quad (11.14)$$

11.2.2.3 Le coefficient d'inégalité de Theil

Autre critère de qualité de la prévision indépendant de l'échelle de la série, le coefficient d'inégalité de Theil est calculé de la façon suivante :

$$Theil = \frac{\sqrt{\frac{1}{h_{\max}} \sum_{h=t+1}^{t+h_{\max}} (y_h^P - y_h)^2}}{\sqrt{\frac{1}{h_{\max}} \sum_{h=t+1}^{t+h_{\max}} y_h^P + \sqrt{\frac{1}{h_{\max}} \sum_{h=t+1}^{t+h_{\max}} y_h}}} \quad (11.15)$$

Ce coefficient est compris entre 0 et 1 ; plus il est proche de 0, meilleure est la qualité de la prévision.

11.2.2.4 La décomposition de l'erreur quadratique moyenne

L'erreur quadratique moyenne de la prévision se décompose, selon Murphy [124], de la façon suivante :

$$\frac{1}{h_{\max}} \sum_{h=t+1}^{t+h_{\max}} (y_h^P - y_h)^2 = (\bar{y}^P - \bar{y})^2 + (\sigma_y^P - \sigma_y)^2 + 2(1-r) \sigma_y^P \sigma_y$$

CHAPITRE 11. DES MODÈLES COINTÉGRÉS SUR UNE SOUS-PÉRIODE PLUS
ADAPTÉS AUX PRÉVISIONS

où \bar{y}^P et σ_y^P (respectivement \bar{y} et σ_y) correspondent à la moyenne et à l'écart type de la variable prévue $\{y_h^P\}$ (respectivement de la véritable variable dans le futur $\{y_h\}$), et r est le coefficient de corrélation entre les séries $\{y_h^P\}$ et $\{y_h\}$, pour $h = t + 1, \dots, t + h_{\max}$.

L'erreur quadratique moyenne s'explique par trois types d'erreur, en proportion différente suivant les modèles :

- La proportion du biais

$$\frac{(\bar{y}^P - \bar{y})^2}{\frac{1}{h_{\max}} \sum_{h=t+1}^{t+h_{\max}} (y_h^P - y_h)^2} \quad (11.16)$$

est une indication de la distance entre la valeur moyenne des prévisions et la valeur moyenne des réalisations futures.

- La proportion de la variance

$$\frac{(\sigma_y^P - \sigma_y)^2}{\frac{1}{h_{\max}} \sum_{h=t+1}^{t+h_{\max}} (y_h^P - y_h)^2} \quad (11.17)$$

est un indicateur mesurant la différence de dispersion entre les prévisions et les réalisations futures effectivement observées autour de leur moyenne respective.

- La proportion de la covariance

$$\frac{2(1-r)\sigma_y^P\sigma_y}{\frac{1}{h_{\max}} \sum_{h=t+1}^{t+h_{\max}} (y_h^P - y_h)^2} \quad (11.18)$$

mesure l'importance de la part de l'erreur de prévision non systématique (i.e. non prise en compte par les indicateurs précédents).

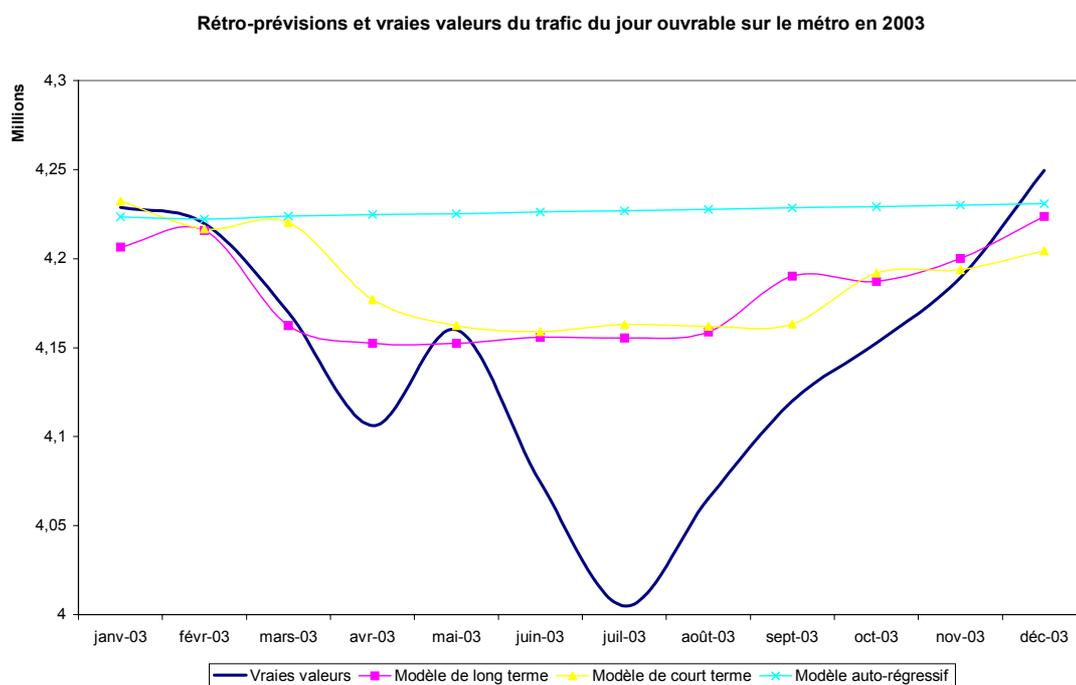
Une prévision de qualité est caractérisée par une faible proportion du biais et de la variance de l'erreur de prévision.

11.2.3 Trafic du jour ouvrable sur le métro

11.2.3.1 Rétro-prévisions effectuées en fin 2002 pour l'ensemble de l'année 2003

Supposons que l'on soit à la fin du mois de décembre 2002, mais que l'on dispose des valeurs réelles que prendront l'ensemble des variables économiques décrivant l'environnement économique en Ile-de-France en 2003. L'application de ces véritables valeurs des variables explicatives, ainsi que le niveau de trafic observé sur le métro pour un jour ouvrable en décembre 2002, conduit à trois types de rétro-prévisions pour l'ensemble de l'année 2003. Ces rétro-prévisions sont encore en logarithme, elles sont immédiatement retransformées par la fonction exponentielle.

Les niveaux de ces rétro-prévisions, ainsi que les véritables valeurs du trafic du jour ouvrable sur le métro (connues à la fin 2003) sont représentés dans le graphique suivant.



Pour comparer les différentes rétro-prévisions aux valeurs de trafic réellement observées, la représentation graphique n'est pas toujours suffisante, et conduit à examiner les critères de mesure de la qualité des rétro-prévisions présentés précédemment.

Trois types de comparaison ont été effectués pour chacune des rétro-prévisions réalisées en se plaçant en décembre 2002 :

CHAPITRE 11. DES MODÈLES COINTÉGRÉS SUR UNE SOUS-PÉRIODE PLUS
ADAPTÉS AUX PRÉVISIONS

- en calculant les critères pour les 12 mois de l’année 2003,
- en ne prenant en compte dans le calcul que les 6 premiers mois (janvier 2003 - juin 2003),
- et en ne tenant compte que des 6 derniers mois de l’année (juillet 2003 - décembre 2003).

Cette distinction entre le premier et le second semestre a pour but de vérifier si le choix du meilleur modèle de prévision dépend de l’horizon considéré. A priori, le modèle auto-régressif ou le modèle à correction d’erreur devraient se montrer plus performants à court terme (*a priori* au 1^{er} semestre), tandis que les prévisions à long terme devraient être améliorées par l’utilisation de la relation de long terme (au second semestre par exemple).

Le calcul des critères de qualité de chaque rétro-prévision a été effectué sur la série re-transformée par la fonction exponentielle, et non sur les rétro-prévisions directement issues des modélisations qui utilisent des variables log-linéarisées. Ce mode de calcul est défavorable aux modèles issus de l’analyse de la cointégration dans la mesure où les prévisions atypiques ne sont pas lissées, alors que la probabilité d’avoir des prévisions extrêmes est plus importante pour ce type de modélisation que pour les modèles auto-régressifs.

<i>jomsa</i>	Période entière			6 premiers mois			6 derniers mois		
	LT	CT	AR	LT	CT	AR	LT	CT	AR
Critère de qualité									
RMSE%	0.008	0.010	0.019	0.003	0.008	0.015	0.013	0.012	0.023
Proportion biais	29%	37%	55%	11%	45%	52%	44%	33%	57%
Proportion variance	52%	46%	41%	56%	30%	44%	54%	62%	43%
Proportion covariance	19%	17%	4%	33%	25%	4%	2%	5%	0%
Theil	0.008	0.015	0.013	0.005	0.006	0.010	0.010	0.018	0.015

LT : Long terme, CT : Court terme (à correction d’erreur), AR : Auto-régressive

En novembre 2002, le département Commercial de la RATP avait prévu que le trafic du jour ouvrable sur le métro s’élèverait chaque mois à environ 4.23 millions de voyageurs, soit un niveau comparable aux prévisions issues de la modélisation auto-régressive.

Sur l’ensemble de l’année 2003, les rétro-prévisions issues de la modélisation de long terme (avec valeurs effectivement observées des variables économiques) sont de meilleure qualité que

les rétro-prévisions calculées pour un autre type de modélisation. La RMSE% ou le critère de Theil sont plus proches de 0, de même que la part du biais dans l'erreur quadratique moyenne, tandis que la part de l'erreur non systématique ("proportion covariance") est élevée par rapport aux modèles alternatifs. Sur les six premiers mois de l'année comme sur les 6 derniers mois, le constat est identique.

11.2.3.2 Rétro-prévisions effectuées en fin août 2003 pour le dernier trimestre 2003

Les trois premiers trimestres de l'année 2003 ont été marqués par des aléas importants (mouvements sociaux dans les transports ou la fonction publique puis canicule). Juillet et août correspondent aux mois de plus faible trafic sur le métro au cours de l'année 2003, même après correction des effets calendaires. A la fin de l'été se posait la question de l'éventualité et de la rapidité de la reprise de la fréquentation.

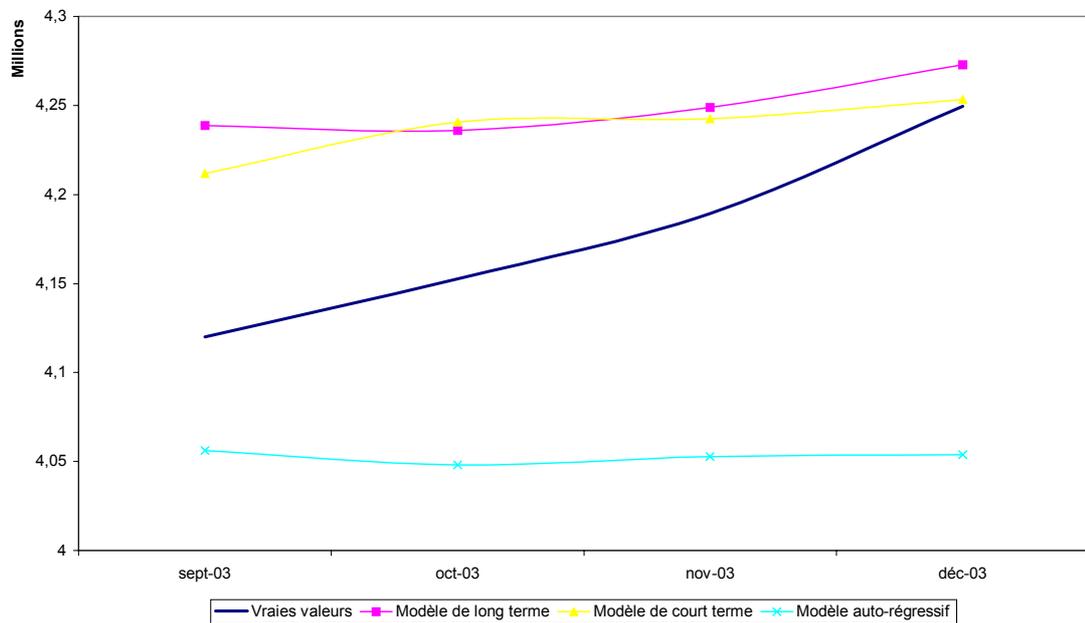
L'exercice de rétro-prévisions a été reconduit en se plaçant à la fin du mois d'août.

- La rétro-prévision calculée à partir du modèle de long terme n'est pas modifiée par la connaissance de nouvelles données de trafic : aucune variable de trafic retardée n'intervient dans l'équation de long terme.
- Les rétro-prévisions associées à la modélisation auto-régressive ainsi qu'à la modélisation de court terme (à correction d'erreur) sont ajustées en fonction des dernières données de trafic disponibles.

Ces nouvelles rétro-prévisions, ainsi que les valeurs de trafic effectivement observées sont représentées dans le graphique suivant.

CHAPITRE 11. DES MODÈLES COINTÉGRÉS SUR UNE SOUS-PÉRIODE PLUS ADAPTÉS AUX PRÉVISIONS

Rétro-prévisions et vraies valeurs du trafic du jour ouvrable sur le métro (dernier trimestre 2003)



Les critères de qualité de ces rétro-prévisions ont été synthétisés dans le tableau suivant. A titre d'information ont été ajoutées les valeurs de ces critères si le calcul des rétro-prévisions de long terme et de court terme avait été effectué à partir des valeurs des exogènes prévues par la modélisation auto-régressive (modélisations "LTAR" et "CTAR").

Le modèle de long terme ("LT") demeure le meilleur modèle de prévision par rapport aux modèles de court terme, à correction d'erreur, ("CT") et auto-régressif ("AR"). La RMSE% et le critère de Theil sont inférieurs, alors que la part de l'erreur non systématique dans l'erreur quadratique moyenne est la plus élevée.

11.2. LES RÉTRO-PRÉVISIONS DE TRAFIC SUR L'ANNÉE 2003

<i>jomsa</i>	Prévisions du dernier trimestre 2003				
	LT	CT	AR	LTAR	CTAR
RMSE%	0.017	0.014	0.030	0.002	0.009
Proportion biais	76%	68%	84%	2%	56%
Proportion variance	22%	28%	14%	98%	0%
Proportion covariance	2%	4%	2%	0%	4%
Theil	0.010	0.010	0.017	0.005	0.006

LT : Long terme, CT : Court terme (à correction d'erreur), AR : Auto-régressive

LTAR : Long terme avec explicatives prévues par modélisation AR

CTAR : Court terme avec explicatives prévues par modélisation AR

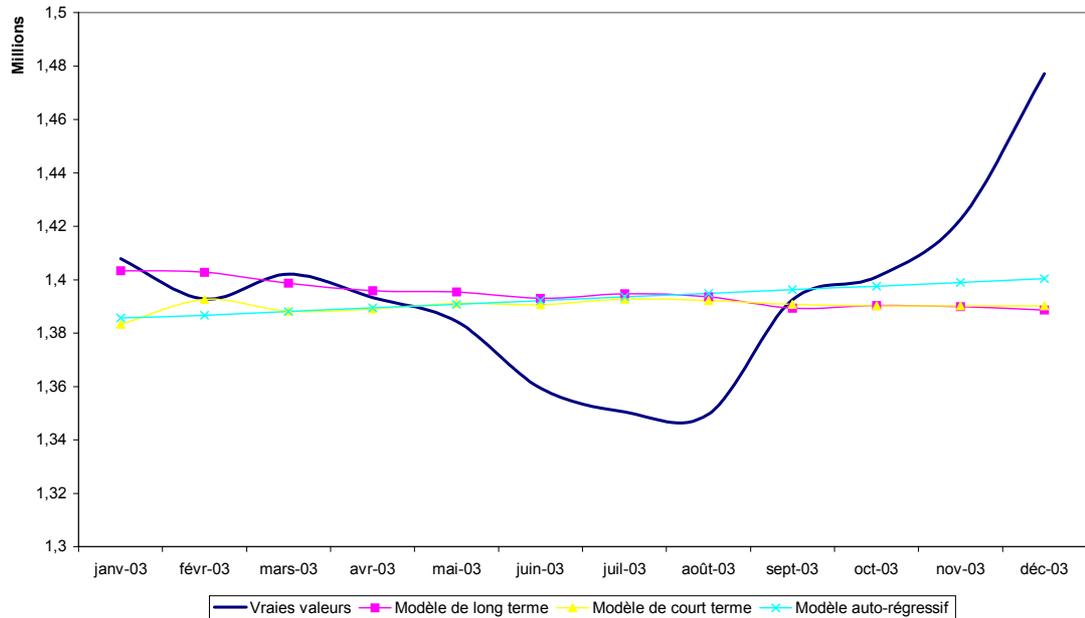
Les performances du modèle de long terme avec variables explicatives prévues par leur modélisation auto-régressive ne remettent pas en question le choix du modèle de long terme comme modèle privilégié de prévision.

Le fait que la RMSE% et le coefficient d'inégalité de Theil sont inférieurs dans le cas de variables explicatives prévues (par rapport aux valeurs effectivement observées) indique que les anticipations concernant la situation économique sont aussi importantes que la véritable conjoncture économique dans la demande de transport sur le métro.

11.2.4 Trafic du jour ouvrable sur le RER

11.2.4.1 Rétro-prévisions effectuées en fin 2002 pour l'ensemble de l'année 2003

Rétro-prévisions et vraies valeurs du trafic du jour ouvrable sur le RER en 2003



<i>joersa</i>	Période entière			6 premiers mois			6 derniers mois		
	LT	CT	AR	LT	CT	AR	LT	CT	AR
Critère de qualité	LT	CT	AR	LT	CT	AR	LT	CT	AR
RMSE%	0.000	0.004	0.002	0.006	0.001	0.001	0.007	0.007	0.002
Proportion biais	0%	2%	0%	25%	0%	0%	2%	2%	0%
Proportion variance	68%	80%	80%	63%	52%	57%	80%	90%	100%
Proportion covariance	32%	18%	20%	12%	48%	43%	17%	8%	0%
Theil	0.012	0.013	0.012	0.006	0.007	0.007	0.018	0.018	0.016

LT : Long terme, CT : Court terme (à correction d'erreur), AR : Auto-régressive

Notre modélisation auto-régressive permet de mieux prévoir le trafic du jour ouvrable sur le RER que la méthode utilisée par le département Commercial de la RATP, qui conduisait à une prévision de 1.43 millions de voyageurs chaque jour ouvrable moyen mensuel.

Sur l'ensemble de l'année, les modélisations auto-régressive et de long terme donnent des résultats assez comparables en termes de RMSE% et de coefficient d'inégalité de Theil. La

décomposition de l'erreur quadratique moyenne conduit à privilégier le modèle de long terme : absence de biais systématique, proportion de l'erreur due à la variance plus faible et part de l'erreur non systématique plus importante.

Sur les six premiers mois de l'année, les critères de qualité des rétro-prévisions amènent à hésiter entre les modélisations de court terme et la modélisation auto-régressive. Le critère RMSE% conduit à privilégier le modèle auto-régressif, tandis que le coefficient d'inégalité de Theil amène à prendre en compte le modèle à correction d'erreur, voire la modélisation de long terme.

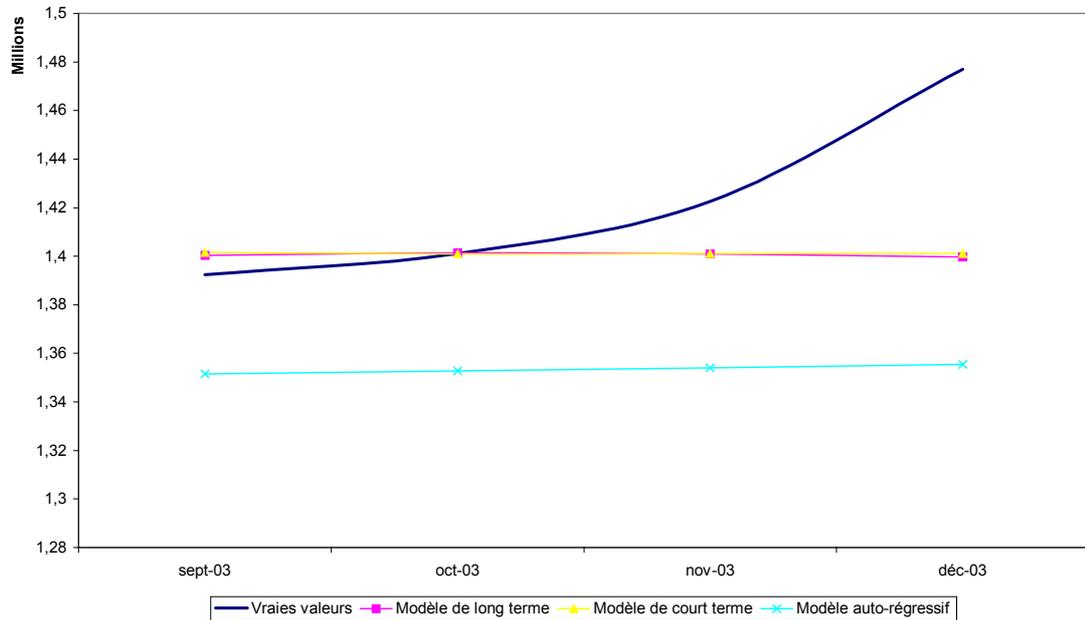
Sur les six derniers mois, enfin, les critères généraux poussent à choisir le modèle auto-régressif, qui ne présente pas de biais systématique. Toutefois, la décomposition de l'erreur quadratique moyenne indique que cette absence de biais systématique cache des erreurs de prévisions plus importantes à chaque date que les autres modélisations. Le modèle de long terme, avec un biais systématique très faible et une erreur quadratique moyenne imputable en plus grande partie à l'erreur non systématique, semble un choix plus adéquat.

11.2.4.2 Rétro-prévisions effectuées en fin août 2003 pour le dernier trimestre 2003

Le calcul des rétro-prévisions a été réactualisé à la fin du mois d'août (compris).

CHAPITRE 11. DES MODÈLES COINTÉGRÉS SUR UNE SOUS-PÉRIODE PLUS ADAPTÉS AUX PRÉVISIONS

Rétro-prévisions et vraies valeurs du trafic du jour ouvrable sur le RER (dernier trimestre 2003)



<i>jorersa</i>	Prévisions du dernier trimestre 2003				
Critère de qualité	LT	CT	AR	LTAR	CTAR
RMSE%	0.017	0.016	0.050	0.021	0.036
Proportion biais	26%	25%	79%	38%	80%
Proportion variance	69%	74%	21%	62%	14%
Proportion covariance	5%	1%	0%	0%	6%
Theil	0.016	0.016	0.028	0.017	0.020

LT : Long terme, CT : Court terme (à correction d'erreur), AR : Auto-régressive

LTAR : Long terme avec explicatives prévues par modélisation AR

CTAR : Court terme avec explicatives prévues par modélisation AR

Les modèles issus de l'analyse de la cointégration ressortent clairement comme les meilleurs modèles de prévision après un début d'année perturbé.

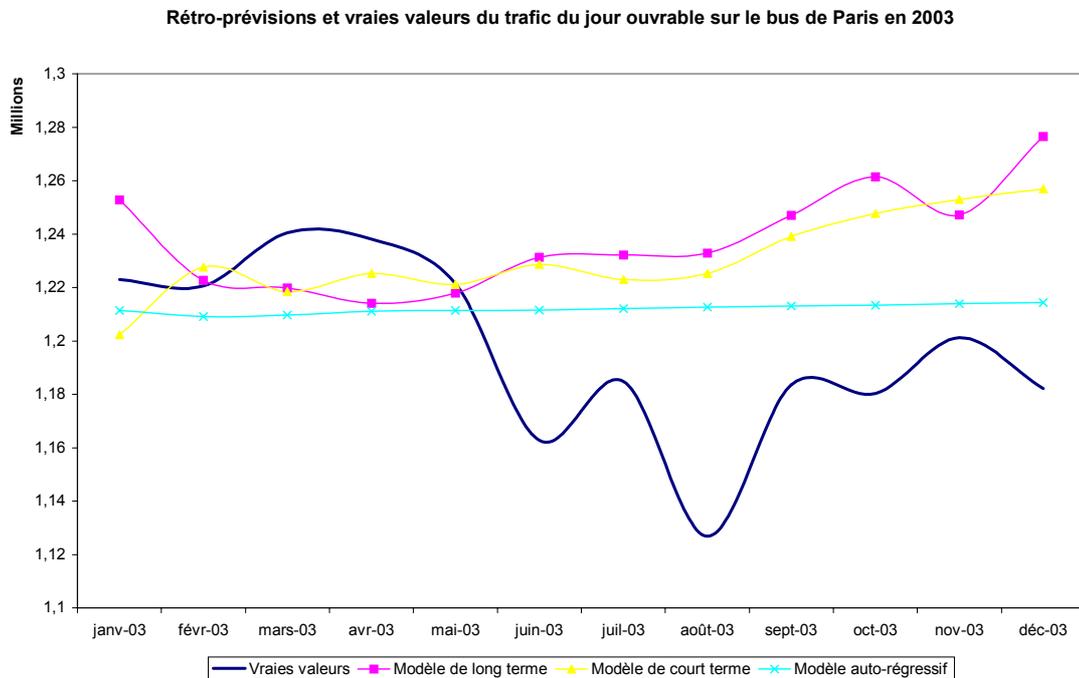
Le biais systématique est moins important en proportion que pour le modèle auto-régressif, tandis que la part de l'erreur non systématique dans l'erreur quadratique moyenne est accrue. RMSE% et coefficient de Theil confirment le choix de ces modèles.

La prise en compte de données d'environnement prévues grâce à leur modélisation AR (au lieu des valeurs observées) ne remet pas en question le choix du modèle de long terme.

11.2.5 Trafic du jour ouvrable sur le réseau des autobus de Paris

Les modélisations du trafic du jour ouvrable sur le réseau des autobus de Paris ont pour particularité d'avoir été estimées sans rupture. La série est intégrée sur l'ensemble de la période, et il existe une relation de cointégration sur l'ensemble de la période considérée (janvier 1990 - décembre 2002). La relation de court terme est elle aussi estimée sur période entière. Il en sera de même pour les rétro-prévisions associées au trafic du jour ouvrable sur le réseau des autobus de banlieue.

11.2.5.1 Rétro-prévisions effectuées en fin 2002 pour l'ensemble de l'année 2003



CHAPITRE 11. DES MODÈLES COINTÉGRÉS SUR UNE SOUS-PÉRIODE PLUS
ADAPTÉS AUX PRÉVISIONS

<i>joparsa</i>	Période entière			6 premiers mois			6 derniers mois		
Critère de qualité	LT	CT	AR	LT	CT	AR	LT	CT	AR
RMSE%	0.033	0.027	0.012	0.007	0.002	0.006	0.062	0.054	0.031
Proportion biais	46%	40%	16%	6%	0%	6%	90%	90%	68%
Proportion variance	6%	11%	72%	15%	32%	85%	1%	3%	31%
Proportion covariance	48%	49%	12%	79%	68%	9%	9%	7%	1%
Theil	0.025	0.022	0.016	0.014	0.014	0.012	0.032	0.028	0.019

LT : Long terme, CT : Court terme (à correction d'erreur), AR : Auto-régressive

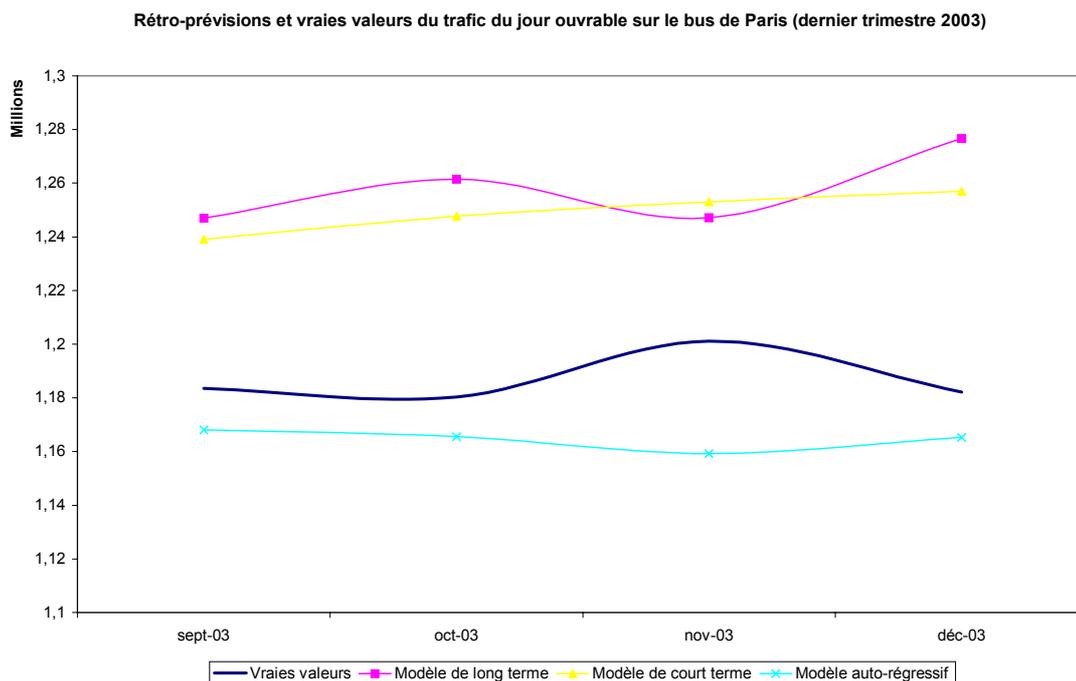
Pour le trafic du jour ouvrable sur le réseau des autobus de Paris, les critères de mesure de la qualité des différents types de rétro-prévisions (RMSE% et coefficient d'inégalité de Theil) indiquent que le modèle le plus satisfaisant pour l'ensemble de l'année 2003 correspond à la modélisation auto-régressive. Néanmoins, l'analyse de la décomposition de l'erreur quadratique moyenne montre que si le biais systématique est inférieur, la variance de l'erreur de prévision est plus importante par rapport aux modélisations issues de l'analyse de la cointégration.

Pour le premier semestre 2003, les trois modèles possèdent des qualités de prévision assez équivalentes au sens de la RMSE% et du coefficient de Theil. Les modélisations issues de l'analyse de la cointégration se distinguent par une proportion de l'erreur non systématique plus importante par rapport à la modélisation auto-régressive. En particulier, la modélisation de court terme correspond à la modélisation la moins biaisée.

Au deuxième semestre, toutes les rétro-prévisions sont biaisées. La modélisation auto-régressive conduit à un biais systématique moins important.

Rétro-prévisions effectuées en fin août 2003 pour le dernier trimestre 2003

11.2. LES RÉTRO-PRÉVISIONS DE TRAFIC SUR L'ANNÉE 2003



La RMSE% ainsi que le critère de Theil (tableau ci-dessous) indiquent que, pour le trafic du jour ouvrable sur le réseau des autobus de Paris, la meilleure prévision réalisée à la fin des vacances d'été pour le dernier trimestre est obtenue à l'aide de la modélisation auto-régressive. Le modèle à correction d'erreur a néanmoins pour avantage de conduire à une erreur systématique moins importante.

<i>joparsa</i>	Prévisions du dernier trimestre 2003				
	LT	CT	AR	LTAR	CTAR
Critère de qualité					
RMSE%	0.060	0.053	0.019	0.043	0.013
Proportion biais	75%	97%	74%	73%	9%
Proportion variance	6%	0%	5%	15%	66%
Proportion covariance	19%	3%	21%	12%	25%
Theil	0.030	0.015	0.011	0.022	0.022

LT : Long terme, CT : Court terme (à correction d'erreur), AR : Auto-régressive

LTAR : Long terme avec explicatives prévues par modélisation AR

CTAR : Court terme avec explicatives prévues par modélisation AR

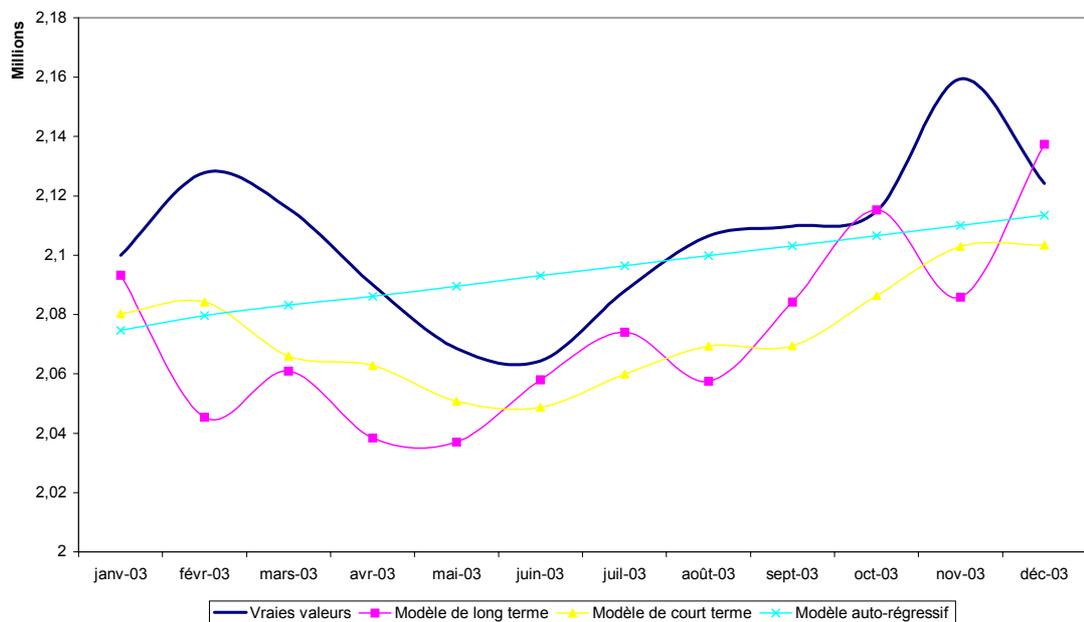
11.2.6 Trafic du jour ouvrable sur le réseau des autobus de banlieue

Les modélisations du trafic du jour ouvrable sur le réseau des autobus de banlieue ont été estimées sans rupture. La série est intégrée sur l'ensemble de la période, et il existe une relation de cointégration sur l'ensemble de la période considérée (janvier 1990 - décembre 2002). La relation de court terme est elle aussi estimée sur période entière.

11.2.6.1 Rétro-prévisions effectuées en fin 2002 pour l'ensemble de l'année 2003

Les différentes rétro-prévisions du trafic du jour ouvrable sur le réseau des autobus de banlieue ont été calculées en décembre 2002 pour l'ensemble de l'année 2003, en supposant l'ensemble des variables exogènes connues. Elles sont représentées dans le graphique suivant, ainsi que les véritables valeurs de trafic observées.

Rétro-prévisions et vraies valeurs du trafic du jour ouvrable sur le bus de banlieue en 2003



11.2. LES RÉTRO-PRÉVISIONS DE TRAFIC SUR L'ANNÉE 2003

<i>jobansa</i>	Période entière			6 premiers mois			6 derniers mois		
Critère de qualité	LT	CT	AR	LT	CT	AR	LT	CT	AR
RMSE%	0.015	0.015	0.005	0.019	0.014	0.005	0.012	0.017	0.006
Proportion biais	52%	85%	17%	63%	80%	10%	37%	89%	28%
Proportion variance	1%	6%	27%	1%	11%	33%	2%	2%	59%
Proportion covariance	47%	9%	56%	36%	9%	57%	61%	9%	13%
Theil	0.010	0.008	0.006	0.011	0.008	0.008	0.009	0.009	0.005

LT : Long terme, CT : Court terme (à correction d'erreur), AR : Auto-régressive

Sur le réseau des autobus de banlieue, sur l'ensemble de l'année, la meilleure rétro-prévision est obtenue grâce au modèle auto-régressif (conclusions de la comparaison des RMSE% et des coefficients d'inégalité de Theil associés à chacun des modèles alternatifs).

Il en est de même lorsque l'analyse est menée sur chaque semestre, pour le seul critère RMSE%. L'examen des coefficients de Theil conduit à considérer que la qualité des modèles de court terme et auto-régressif est très similaire, à l'exception du biais, légèrement plus important pour le modèle à correction d'erreur. Enfin, le modèle de long terme présente l'avantage de présenter une dispersion presque identique à celle observée en réalité.

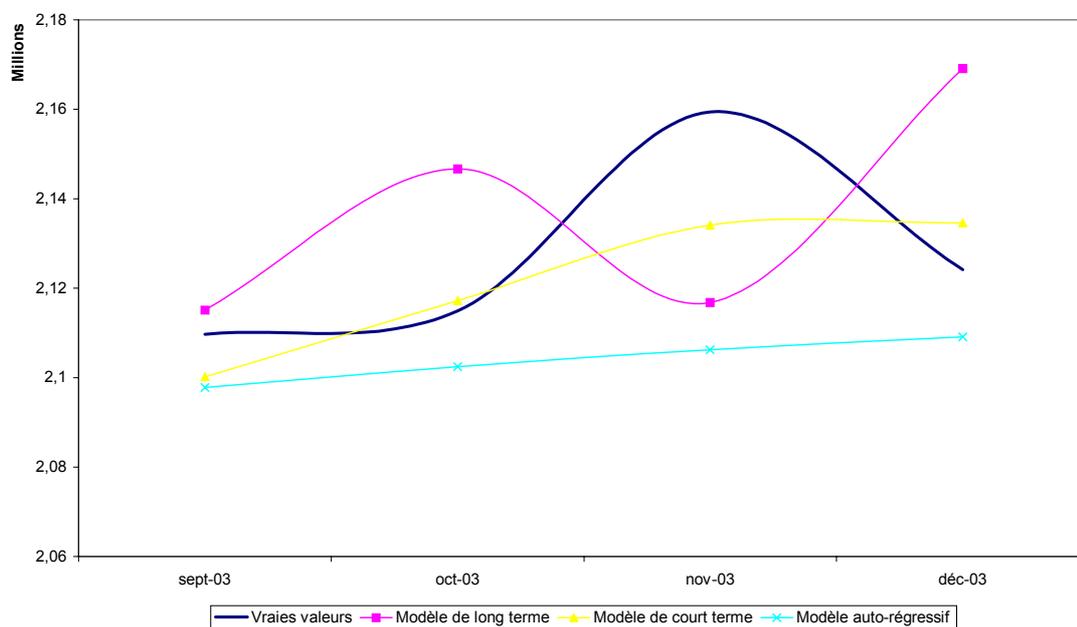
11.2.6.2 Rétro-prévisions effectuées en fin août 2003 pour le dernier trimestre 2003

Lorsque les rétro-prévisions sont effectuées en se plaçant à la fin du mois d'août pour le dernier trimestre, le meilleur modèle au sens des critères RMSE% et Theil correspond au modèle à correction d'erreur. Lorsque les variables environnementales ne sont plus observées mais prévues en fonction de leur variations passées, ce modèle demeure le meilleur modèle au sens du critère d'inégalité de Theil.

Toutefois, la modélisation de long terme est tout à fait judicieuse : sur cette sous-période, elle conduit à une erreur systématique bien moins importante lorsque les vraies valeurs des variables décrivant l'environnement économique francilien sont connues.

CHAPITRE 11. DES MODÈLES COINTÉGRÉS SUR UNE SOUS-PÉRIODE PLUS ADAPTÉS AUX PRÉVISIONS

Rétroprévisions et vraies valeurs du trafic du jour ouvrable sur le bus de banlieue (dernier trimestre 2003)



<i>jobansa</i>	Prévisions du dernier trimestre 2003				
Critère de qualité	LT	CT	AR	LTAR	CTAR
RMSE%	0.005	0.003	0.011	0.024	0.017
Proportion biais	6%	11%	57%	86%	60%
Proportion variance	1%	13%	33%	10%	6%
Proportion covariance	93%	76%	10%	4%	35%
Theil	0.008	0.004	0.009	0.013	0.011

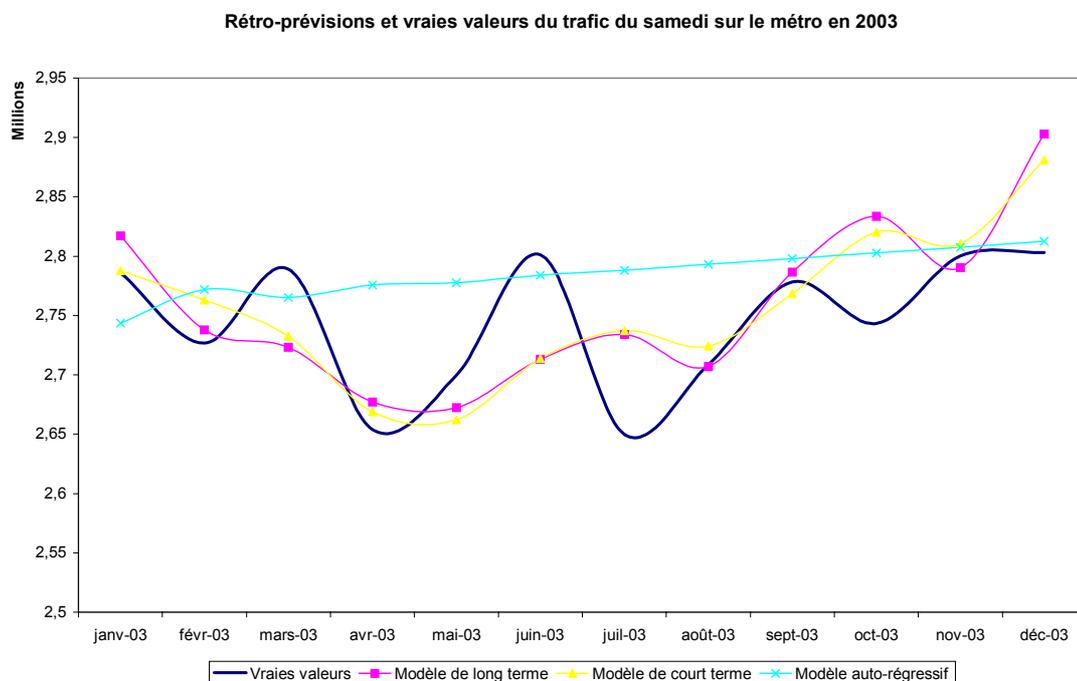
LT : Long terme, CT : Court terme (à correction d'erreur), AR : Auto-régressive

LTAR : Long terme avec explicatives prévues par modélisation AR

CTAR : Court terme avec explicatives prévues par modélisation AR

11.2.7 Trafic du samedi sur le métro

11.2.7.1 Rétro-prévisions effectuées en fin 2002 pour l'ensemble de l'année 2003



<i>smsa</i>	Période entière			6 premiers mois			6 derniers mois		
	LT	CT	AR	LT	CT	AR	LT	CT	AR
Critère de qualité									
RMSE%	0.005	0.014	0.014	0.007	0.008	0.009	0.016	0.016	0.019
Proportion biais	5%	4%	32%	13%	17%	15%	44%	51%	51%
Proportion variance	4%	1%	28%	1%	3%	40%	2%	0%	47%
Proportion covariance	91%	95%	40%	85%	80%	45%	54%	49%	2%
Theil	0.010	0.010	0.013	0.009	0.009	0.013	0.012	0.011	0.013

LT : Long terme, CT : Court terme (à correction d'erreur), AR : Auto-régressive

Sur l'ensemble de l'année, les meilleures rétro-prévisions relatives au trafic du samedi sur le métro ont été obtenues par la modélisation de long terme. Cette modélisation présente une RMSE% et un critère de Theil inférieurs aux autres modélisations. L'analyse de la décomposition de l'erreur quadratique moyenne indique en outre que la part du biais systématique ou de la

CHAPITRE 11. DES MODÈLES COINTÉGRÉS SUR UNE SOUS-PÉRIODE PLUS ADAPTÉS AUX PRÉVISIONS

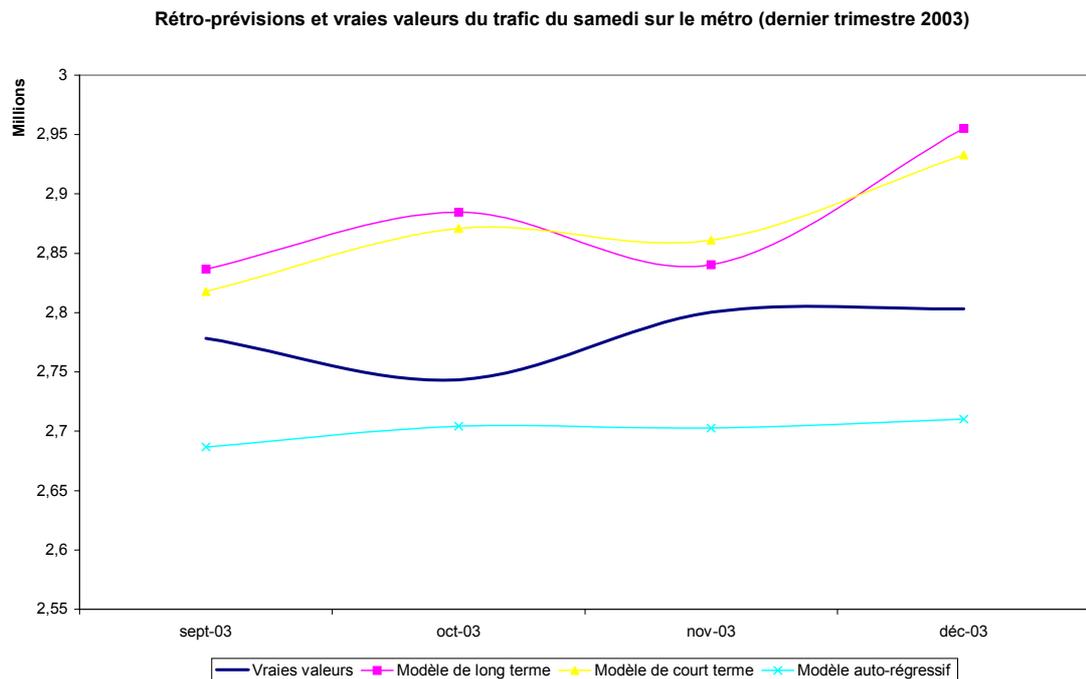
variance de l'erreur estimée sont bien plus faibles. La majeure partie de l'erreur quadratique moyenne correspond à une forme non systématique d'erreur.

Sur les six premiers mois de l'année, les modèles issus de l'analyse de la cointégration améliorent la qualité des rétro-prévisions par rapport au modèle de auto-régressif. Ces modélisations minimisent la RMSE% et le coefficient d'inégalité de Theil, ainsi que la part de l'erreur systématique dans l'erreur quadratique moyenne.

Enfin, sur les six derniers mois de l'année, les modèles de long terme et de court terme redeviennent les modèles à choisir pour effectuer les prévisions du trafic du samedi sur le RER.

11.2.7.2 Rétro-prévisions effectuées en fin août 2003 pour le dernier trimestre 2003

Comme pour les autres types de trafic, nous nous sommes placés à la fin de l'été 2003, après les mouvements sociaux et la canicule, pour prévoir la fin de l'année. Les meilleures rétro-prévisions ont été réalisées à l'aide du modèle auto-régressif, à peine moins biaisé.



Si les variables explicatives avaient été prévues à l'aide de leur modélisation auto-régressive au lieu d'être connues, le modèle de long terme aurait réalisé de meilleures performances que le

modèle à correction d'erreur, en diminuant le biais par exemple.

<i>smsa</i>	Prévisions du dernier trimestre 2003				
	LT	CT	AR	LTAR	CTAR
Critère de qualité					
RMSE%	0.035	0.032	0.029	0.015	0.016
Proportion biais	75%	79%	90%	73%	71%
Proportion variance	6%	4%	4%	15%	3%
Proportion covariance	19%	17%	6%	12%	26%
Theil	0.019	0.018	0.015	0.008	0.010

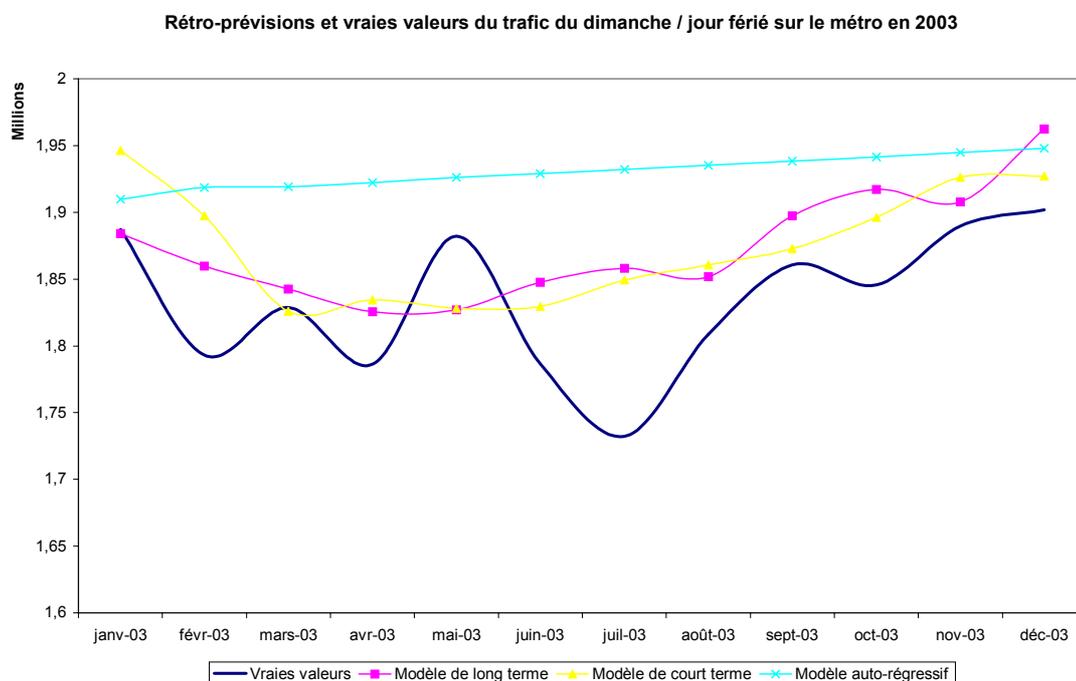
LT : Long terme, CT : Court terme (à correction d'erreur), AR : Auto-régressive

LTAR : Long terme avec explicatives prévues par modélisation AR

CTAR : Court terme avec explicatives prévues par modélisation AR

11.2.8 Trafic du dimanche/jour férié sur le métro

11.2.8.1 Rétro-prévisions effectuées en fin 2002 pour l'ensemble de l'année 2003



CHAPITRE 11. DES MODÈLES COINTÉGRÉS SUR UNE SOUS-PÉRIODE PLUS
ADAPTÉS AUX PRÉVISIONS

<i>dfmsa</i>	Période entière			6 premiers mois			6 derniers mois		
Critère de qualité	LT	CT	AR	LT	CT	AR	LT	CT	AR
RMSE%	0.021	0.022	0.052	0.011	0.018	0.051	0.032	0.026	0.054
Proportion biais	44%	45%	78%	17%	27%	77%	72%	64%	76%
Proportion variance	4%	3%	14%	25%	0%	15%	9%	22%	24%
Proportion covariance	52%	52%	8%	58%	73%	8%	19%	14%	0%
Theil	0.016	0.016	0.029	0.013	0.017	0.028	0.018	0.016	0.030

LT : Long terme, CT : Court terme (à correction d'erreur), AR : Auto-régressive

Pour le trafic du dimanche ou jour férié sur le métro, les meilleurs modèles sur l'ensemble de l'année correspondent aux modèles issus de l'analyse de la cointégration (choix d'après la RMSE% et le coefficient de Theil). Ces modèles apportent une plus grande information en matière de dispersion du trafic. Le modèle auto-régressif a systématiquement surestimé le trafic, même dans les premiers mois, moins perturbés.

Sur les six premiers mois de l'année 2003, les modélisations issues de l'analyse de la cointégration réalisent de meilleures performances prédictives : moins biaisés, ils sont moins entachés d'erreur systématique.

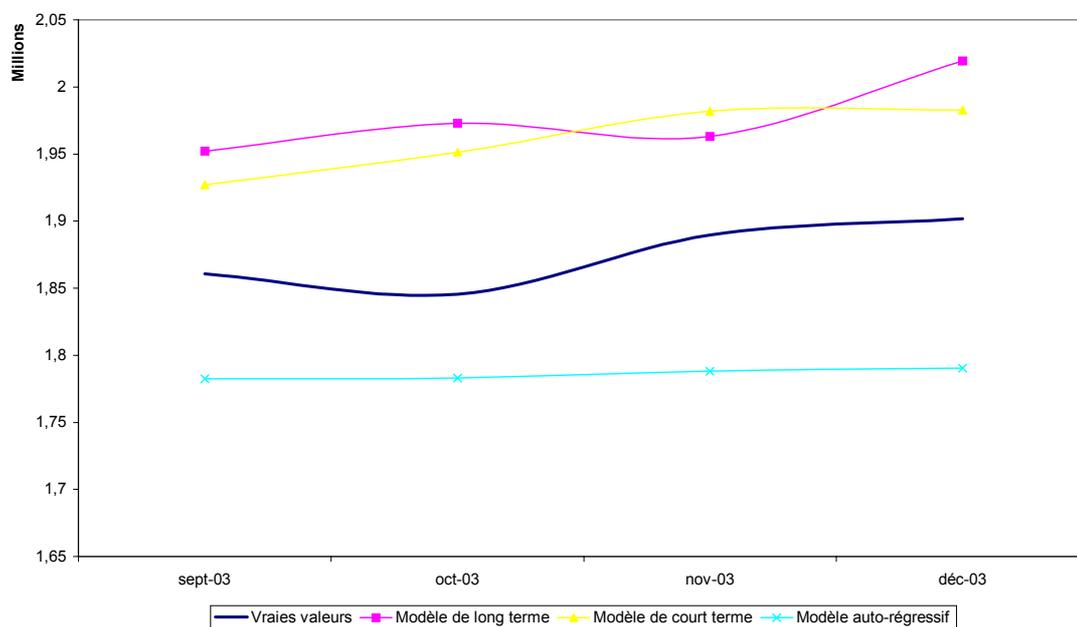
Enfin, sur les 6 derniers mois de l'année, RMSE% et coefficient de Theil poussent à privilégier le choix du modèle à correction d'erreur. Si ce modèle entraîne des rétro-prévisions moins biaisées que les autres modélisations, il possède toutefois une erreur systématique plus importante dans l'erreur quadratique moyenne que le modèle de long terme.

11.2.8.2 Rétro-prévisions effectuées en fin août 2003 pour le dernier trimestre 2003

Sur le dernier trimestre 2003, le meilleur modèle (lorsque les variables d'environnement sont supposées connues) correspond au modèle auto-régressif ou à correction d'erreur, conformément aux critères RMSE% et Theil présentés dans le tableau ci-dessous. Toutefois, l'analyse de la décomposition de l'erreur quadratique moyenne incite à choisir le modèle de long terme qui, pour un critère de Theil équivalent, dispose d'une part de covariance plus importante dans l'erreur quadratique moyenne.

11.2. LES RÉTRO-PRÉVISIONS DE TRAFIC SUR L'ANNÉE 2003

Rétro-prévisions et vraies valeurs du trafic du dimanche / jour férié sur le métro (dernier trimestre 2003)



<i>dfmsa</i>	Prévisions du dernier trimestre 2003				
	LT	CT	AR	LTAR	CTAR
Critère de qualité					
RMSE%	0.055	0.046	0.047	0.001	0.008
Proportion biais	95%	96%	94%	1%	21%
Proportion variance	0%	0%	6%	74%	9%
Proportion covariance	5%	4%	0%	25%	71%
Theil	0.027	0.023	0.025	0.005	0.009

LT : Long terme, CT : Court terme (à correction d'erreur), AR : Auto-régressive

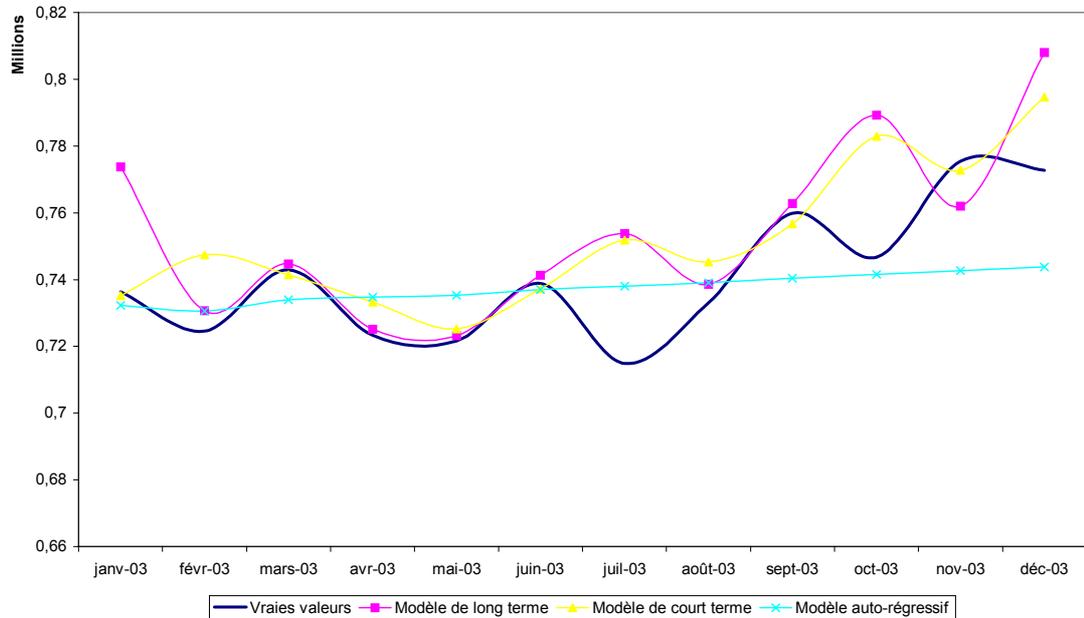
LTAR : Long terme avec explicatives prévues par modélisation AR

CTAR : Court terme avec explicatives prévues par modélisation AR

11.2.9 Trafic du samedi sur le RER

11.2.9.1 Rétro-prévisions effectuées en fin 2002 pour l'ensemble de l'année 2003

Rétro-prévisions et vraies valeurs du trafic du samedi sur le RER en 2003



<i>sversa</i>	Période entière			6 premiers mois			6 derniers mois		
	LT	CT	AR	LT	CT	AR	LT	CT	AR
Critère de qualité	LT	CT	AR	LT	CT	AR	LT	CT	AR
RMSE%	0.018	0.015	0.005	0.012	0.007	0.004	0.024	0.022	0.013
Proportion biais	34%	35%	4%	26%	24%	8%	39%	47%	16%
Proportion variance	7%	1%	80%	33%	2%	53%	0%	3%	82%
Proportion covariance	59%	64%	16%	40%	74%	39%	61%	50%	2%
Theil	0.015	0.013	0.012	0.011	0.008	0.006	0.019	0.016	0.016

LT : Long terme, CT : Court terme (à correction d'erreur), AR : Auto-régressive

Sur l'ensemble de l'année, pour le trafic du samedi sur le RER, le critère de Theil et la RMSE% concluent que la meilleure rétro-prévision est obtenue à l'aide de la modélisation auto-régressive, alors que la représentation graphique laisse penser que la modélisation de long terme permet d'atteindre de meilleurs résultats. L'analyse de la décomposition de l'erreur quadratique moyenne montre que les modélisations issues de l'analyse de la cointégration conduisent à une

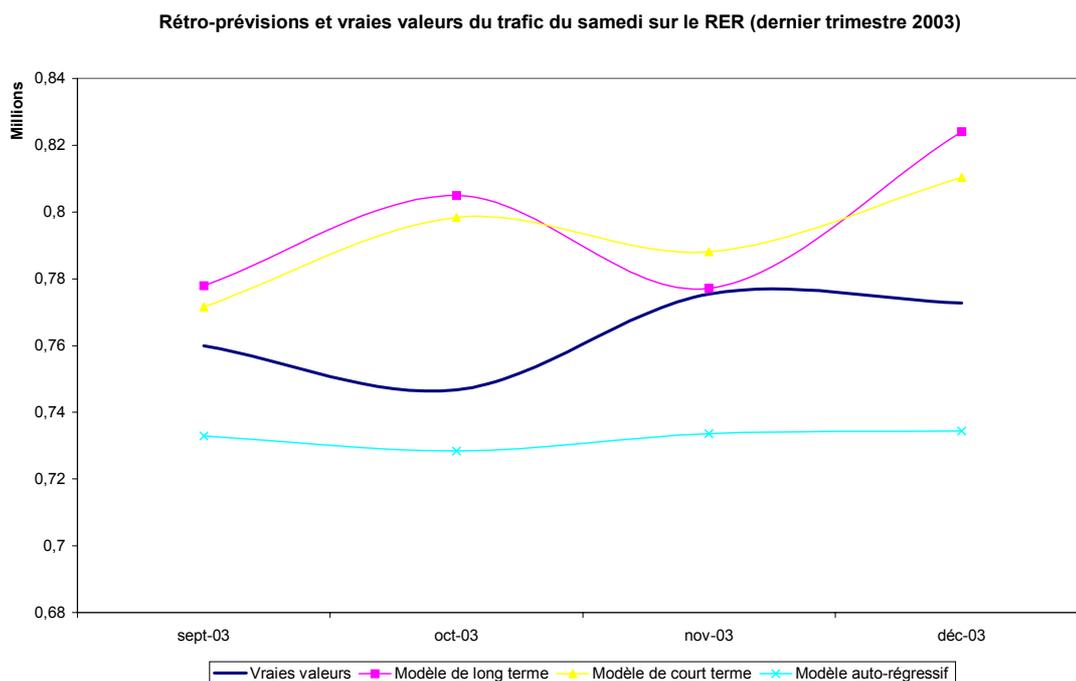
erreur systématique bien moins importante en proportion.

Les modèles de long terme et de court terme possèdent des erreurs de rétro-prévisions plus faibles en valeur absolue, mais ont tendance à surestimer très légèrement le trafic du samedi sur le RER, ce qui les pénalise par rapport à la modélisation auto-régressive. L'analyse est identique lorsque les rétro-prévisions sont évaluées sur le premier et le second semestre séparément.

11.2.9.2 Rétro-prévisions effectuées en fin août 2003 pour le dernier trimestre 2003

A la fin de l'été 2003, la meilleure rétro-prévision pour le dernier trimestre 2003 a été réalisée à l'aide de la modélisation à correction d'erreur : moins biaisée que les autres modélisations, la part de l'erreur non systématique dans l'ensemble de l'erreur quadratique moyenne est plus importante.

Les différents types de rétro-prévision sont représentés dans le graphique suivant :



Lorsque les véritables variables environnementales ne sont pas disponibles, le modèle de long terme conduit à des prévisions moins biaisées, quoique plus dispersées autour des véritables

CHAPITRE 11. DES MODÈLES COINTÉGRÉS SUR UNE SOUS-PÉRIODE PLUS
ADAPTÉS AUX PRÉVISIONS

valeurs de trafic. La part de l'erreur systématique dans l'erreur quadratique moyenne est plus faible.

<i>sversa</i>	Prévisions du dernier trimestre 2003				
	LT	CT	AR	LTAR	CTAR
RMSE%	0.042	0.037	0.041	0.016	0.019
Proportion biais	59%	68%	89%	47%	43%
Proportion variance	5%	1%	10%	38%	1%
Proportion covariance	36%	31%	1%	15%	56%
Theil	0.027	0.021	0.022	0.011	0.019

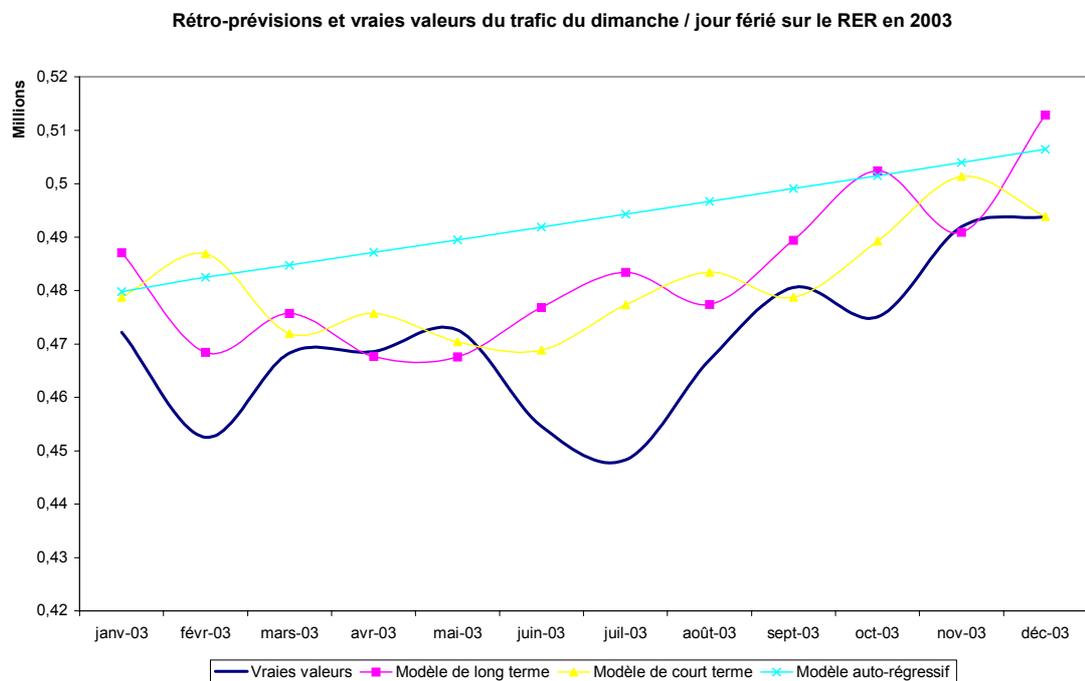
LT : Long terme, CT : Court terme (à correction d'erreur), AR : Auto-régressive

LTAR : Long terme avec explicatives prévues par modélisation AR

CTAR : Court terme avec explicatives prévues par modélisation AR

11.2.10 Trafic du dimanche/jour férié sur le RER

11.2.10.1 Rétro-prévisions effectuées en fin 2002 pour l'ensemble de l'année 2003



11.2. LES RÉTRO-PRÉVISIONS DE TRAFIC SUR L'ANNÉE 2003

<i>df</i> <i>re</i> <i>rsa</i>	Période entière			6 premiers mois			6 derniers mois		
Critère de qualité	LT	CT	AR	LT	CT	AR	LT	CT	AR
RMSE%	0.027	0.023	0.048	0.019	0.022	0.045	0.034	0.023	0.050
Proportion biais	53%	47%	80%	43%	41%	80%	61%	49%	78%
Proportion variance	0%	7%	5%	1%	2%	3%	4%	23%	21%
Proportion covariance	47%	46%	15%	57%	57%	17%	35%	28%	1%
Theil	0.018	0.017	0.026	0.015	0.018	0.025	0.021	0.016	0.028

LT : Long terme, CT : Court terme (à correction d'erreur), AR : Auto-régressive

Selon les critères RMSE% et Theil, le meilleur modèle sur l'ensemble de la période correspond au modèle à correction d'erreur. Le biais systématique de cette modélisation est, proportionnellement à l'erreur quadratique moyenne, moins élevé par rapport aux autres modélisations. Toutefois, le modèle de long terme correspond à un modèle très utile à la prévision, dans la mesure où malgré un biais plus important, la dispersion des rétro-prévisions autour des vraies valeurs de trafic du dimanche/jour férié sur le RER est très faible. Le modèle de long terme permet de construire les rétro-prévisions qui possèdent une erreur systématique (proportionnellement à l'erreur quadratique moyenne) plus faible que les modélisations alternatives.

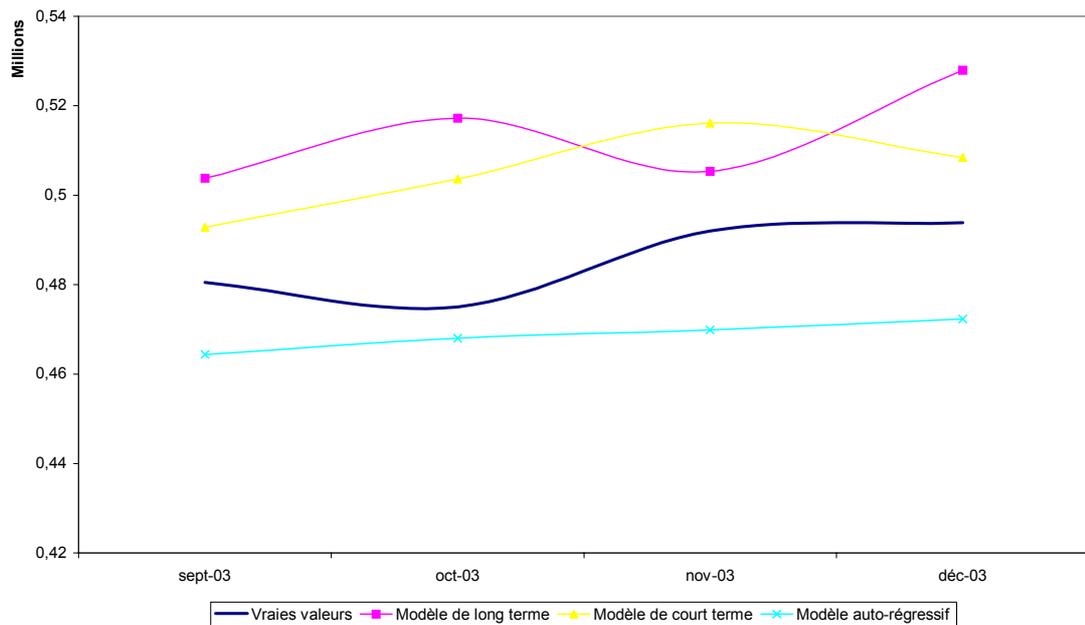
Pour le premier semestre, les rétro-prévisions ayant la meilleure qualité au sens des critères utilisés sont celles qui ont été élaborées à partir de cette même équation de long terme. RMSE%, coefficient de Theil et part de l'erreur systématique sont plus faibles que dans les autres modèles.

Sur la fin de l'année, cependant, le choix qui découle de l'examen de la RMSE% et du coefficient de Theil se porte sur le modèle à correction d'erreur. Ce choix s'explique par la faiblesse du biais relatif, alors que la part de l'erreur systématique est particulièrement importante. Malgré un biais plus important, le modèle qui semble le mieux prévoir correspond au modèle de long terme, qui possède une erreur non systématique plus importante en proportion.

11.2.10.2 Rétro-prévisions effectuées en fin août 2003 pour le dernier trimestre 2003

Les rétro-prévisions ont été recalculées à la fin du mois d'août, pour tenir compte des baisses de trafic observées les mois précédents du fait des mouvements sociaux et de la canicule.

Rétro-prévisions et vraies valeurs du trafic du dimanche / jour férié sur le RER (dernier trimestre 2003)



<i>dfrersa</i>	Prévisions du dernier trimestre 2003				
	LT	CT	AR	LTAR	CTAR
Critère de qualité					
RMSE%	0.058	0.041	0.035	0.001	0.010
Proportion biais	83%	86%	85%	0%	25%
Proportion variance	1%	0%	10%	77%	0%
Proportion covariance	16%	13%	5%	23%	75%
Theil	0.030	0.022	0.019	0.007	0.010

LT : Long terme, CT : Court terme (à correction d'erreur), AR : Auto-régressive

LTAR : Long terme avec explicatives prévues par modélisation AR

CTAR : Court terme avec explicatives prévues par modélisation AR

Cette nouvelle rétro-prévision révèle les mêmes propriétés que l'analyse sur les rétro-prévisions effectuées en fin décembre 2002. Les modèles à correction d'erreur et auto-régressif correspondent aux modèles les meilleurs au sens des critères de la RMSE% et du coefficient de Theil, du fait de la moins grande importance du biais systématique. Toutefois, le modèle de long terme permet d'obtenir des rétro-prévisions moins entachées d'erreur systématique.

Lorsque les valeurs des variables explicatives sont inconnues et qu'elles doivent être prévues grâce à leur modélisation auto-régressive, le modèle de long terme fait mieux que ses alternatives en terme de biais et de variance de l'erreur de prévision.

11.3 Apports des rétro-prévisions

11.3.1 Recommandations pour le choix du modèle de prévision

L'année 2003 est une année atypique, du fait des perturbations (grèves et canicule notamment), ce qui altère la qualité de nos rétro-prévisions par rapport à ce que nous aurions obtenu si les chocs externes avaient été moins nombreux. Toutefois, l'exercice de rétro-prévision avant les perturbations, puis après, permet d'effectuer quelques recommandations pour le choix du modèle de prévision associé à chaque trafic.

Dans le tableau suivant, les meilleurs modèles issus de l'analyse précédente ont été synthétisés.

CHAPITRE 11. DES MODÈLES COINTÉGRÉS SUR UNE SOUS-PÉRIODE PLUS
ADAPTÉS AUX PRÉVISIONS

Choix du modèle de prévision selon la situation				
	Avant des perturbations non prévues			Ap. les perturbations
	Année 2003	1 ^{er} semestre	2 ^{ème} semestre	4 ^{ème} trimestre 2003
JO métro	Long terme	Long terme	Long terme	Long terme
JO RER	Long terme	Court terme	Long terme	Long terme
JO bus Paris	Auto-régressif	Court terme	Auto-régressif	Auto-régressif
JO bus banlieue	Auto-régressif	Auto-régressif	Auto-régressif	Court terme
Samedi métro	Long terme	Court terme	Long terme	Court terme
Dimanche métro	Court terme	Court Long } terme	Court terme	Long terme
Samedi RER	Long terme	Long terme	Long terme	Court terme
Dimanche RER	Court terme	Long terme	Court Long } terme	Court terme

A l'exception du réseau routier (autobus de Paris et de banlieue), les modèles issus de l'analyse de la cointégration permettent de réaliser des prévisions de meilleure qualité que les modèles auto-régressifs. Ajouter des variables économiques au modèle de prévision améliore les critères de qualité de la prévision par rapport à la simple prise en compte des variations de trafic passées.

Ce tableau montre de façon forte l'intérêt du travail de la thèse, en particulier l'apport des tests de modification du degré de cointégration :

- Seuls les trafics des autobus de Paris et de banlieue pouvaient être estimés sur l'ensemble de la période (janvier 1990 - décembre 2002), que la modélisation utilisée soit auto-régressive ou issue de l'analyse de la cointégration. Pour ces trafics sans rupture, la modélisation auto-régressive semble une modélisation suffisante.
- Pour les autres types de trafics, il n'existait pas de relation de cointégration sans rupture sur l'ensemble de la période. L'application des tests de modification du degré de cointégration a permis de montrer que l'absence d'équilibre de long terme depuis janvier 1990 masquait la création d'un équilibre en seconde sous-période, la date de rupture étant estimée à l'aide de ce même test. Toutes ces modélisations sur seconde sous-période, qui n'auraient pu être

mises en avant sans l'application de nos tests, réalisent des prévisions de meilleure qualité que les modèles auto-régressifs.

De façon générale, en l'absence de perturbations exogènes, le modèle à privilégier correspond au modèle de long terme lorsqu'il s'agit de prévoir à plus de 6 mois. Le modèle à correction d'erreur a pour principal avantage de diminuer le biais des prévisions à court terme et particulièrement après un choc, mais il n'est pas fortement explicatif. Lors de l'estimation des paramètres de la relation de court terme, nous avons montré que l'élément le plus significatif correspondait à la force de rappel vers l'équilibre de long terme. Le fait d'avoir une variable de trafic retardée dans l'équation permet de corriger une partie du biais après une perturbation.

11.3.2 Impact de l'estimation des paramètres de la modélisation auto-régressive sur une sous-période uniquement

Les modélisations auto-régressives sont utilisées pour prévoir les différents types de trafic, ainsi que pour prévoir les variables économiques qui interviennent dans l'analyse de la cointégration.

L'application des tests de modification du degré d'intégration à l'ensemble des séries de trafic et des séries décrivant l'environnement économique francilien a montré que certaines variables n'étaient pas intégrées à l'ordre 1 sur l'ensemble de la période, mais stationnaires sur une première sous-période et intégrées ensuite. Lorsque les tests ont conclu à une telle rupture, les paramètres auto-régressifs ont été estimés sur la seconde sous-période uniquement. Le but de ce paragraphe est d'évaluer le gain en terme de prévision d'une telle troncature :

- pour les rétro-prévisions de trafic à partir de la modélisation auto-régressive,
- lors de l'utilisation des modélisations auto-régressives des variables d'environnement dans les rétro-prévisions effectuées par les équations de long terme et de court terme.

11.3.2.1 Modélisation auto-régressive des trafics sur sous-période et période entière

Les tests de changements du degré d'intégration ont conclu que trois séries de trafic n'étaient pas intégrées sur l'ensemble de la période (janvier 1990 - décembre 2002), mais sur une seconde sous-période uniquement :

- le trafic du jour ouvrable sur le RER (rupture en novembre 1993),
- le trafic du samedi sur le métro (rupture en mars 1994),
- le trafic du dimanche ou jour férié sur le RER (rupture en février 1998).

Dans les graphiques et tableaux associés à chacun de ces trois trafics, la modélisation auto-régressive prise en compte correspond à celle estimée sur la seconde sous-période uniquement. Nous avons effectué le même travail de rétro-prévision sur le modèle auto-régressif en utilisant les paramètres issus de l'estimation sur la période entière. En comparant la qualité de ces deux rétro-prévisions par rapport aux véritables données, nous avons pu évaluer les gains éventuels de la prise en compte de cette seule seconde sous-période lors de l'estimation des paramètres auto-régressifs.

Pour l'ensemble de l'année 2003, la prise en compte des paramètres estimés sur la seconde sous-période ne détériore pas les résultats des rétro-prévisions de ces différents trafics.

- Si les paramètres de la modélisation auto-régressive du jour ouvrable sur le RER avaient été estimés depuis janvier 1990, l'erreur normalisée quadratique moyenne (RMSE%) se serait élevée à 0.004 sur l'ensemble de l'année 2003, 0.002 pour les six premiers mois et 0.006 pour les six derniers mois (contre, respectivement 0.002, 0.001 et 0.002 pour des paramètres estimés à partir de décembre 1993).
- Pour le trafic du samedi sur le métro, si les paramètres avaient été estimés sur l'ensemble de la période, les rétro-prévisions issues de la modélisation auto-régressive seraient restées de moins bonne qualité que celles provenant de la modélisation de long terme. En outre, la proportion de la variance dans l'erreur quadratique moyenne aurait augmenté à 63% pour l'ensemble de l'année, et 76% pour les six premiers mois et les six derniers mois (contre, respectivement, 28%, 40% et 47%).

- Pour le trafic du dimanche ou jour férié sur le RER, la proportion de la variance dans l'erreur quadratique moyenne est plus élevée lorsque les paramètres sont estimés depuis janvier 1990, pour une erreur non-systématique inchangée. La part du biais est moins importante, et le critère de Theil recommande de privilégier les paramètres estimés sur l'ensemble de la période. Toutefois, ce constat statistique est difficile à interpréter, du fait de la spécificité de l'année 2003 : la prise en compte de l'ensemble de la période lors de l'estimation aurait conduit à des rétro-prévisions plus pessimistes, et des éléments externes ont contribué à diminuer le trafic réalisé pendant 2003 (canicule et grèves notamment).

Le gain de qualité des prévisions pour l'ensemble de l'année 2003 lorsque les paramètres auto-régressifs sont estimés sur seconde sous-période est indiscutable pour le modèle de trafic du jour ouvrable sur le RER. Pour les modélisations du trafic du samedi sur le métro et du dimanche/jour férié sur le RER, le gain est moins évident. Si la part de l'erreur non systématique diminue dans l'erreur quadratique moyenne, la part du biais systématique est plus élevée. L'explication réside dans le fait que les rétro-prévisions les plus pessimistes pour l'année 2003 sont avantagées par l'atypicité de l'année 2003.

Les rétro-prévisions effectués après les perturbations, i.e. à la fin du mois d'août 2003 pour le dernier trimestre 2003, rétablissent l'avantage des modélisations sur seconde sous-période.

- Pour le trafic du jour ouvrable sur le RER, les rétro-prévisions associées à cette estimation sur seconde sous-période sont améliorées à la marge par rapport à celles obtenues à partir de l'estimation sur l'ensemble de la période ($RMSE\% = 0.051$ et $Theil = 0.029$ contre respectivement 0.050 et 0.028) alors qu'un moins grand nombre de données a été nécessaire à l'estimation.
- L'utilisation des paramètres issus de l'estimation sur la période entière (i.e. depuis janvier 1990) n'améliore pas les performances de ce type de modélisation pour le trafic du samedi sur le métro : les proportions du biais et de la variance dans l'erreur de prévision auraient été plus importantes.
- Prendre en compte l'ensemble de la période pour l'estimation des paramètres auto-régressifs aurait conduit à de moins bons résultats de rétro-prévisions du trafic du dimanche ou jour

férié sur le RER ($RMSE\% = 0.045$ et $Theil = 0.025$), dans un contexte de retour à la normale du trafic.

11.3.2.2 Modélisation auto-régressive des variables environnementales dans les équations de court terme et de long terme.

Les rétro-prévisions issues de l'analyse de la cointégration à la Engle et Granger (équation de long terme et équation de court terme) représentées graphiquement dans ce chapitre ont été calculées en prenant en compte les valeurs observées des variables décrivant l'environnement économique francilien.

Nous avons indiqué quelques résultats concernant les valeurs de trafic qui auraient été prévues en l'absence de ces données, i.e. en se plaçant réellement en situation de prévision. Les prévisions "LTAR" et "CTAR" présentées dans les tableaux de critère de mesure des rétro-prévisions effectuées à la fin de l'été pour le dernier trimestre de l'année ont été calculées à partir de prévisions des séries d'environnement. Ces prévisions ont été réalisées à l'aide des modélisations auto-régressives présentées dans le chapitre 7. Pour certaines variables, les paramètres auto-régressifs ont été estimés sur une seconde sous-période uniquement, conformément aux conclusions des tests de modification du degré d'intégration.

A titre de comparaison, nous avons réalisé les prévisions de ces variables économiques en utilisant les paramètres estimés sur l'ensemble de la période, puis les avons introduits dans les prévisions de long terme et de court terme. Les prévisions obtenues de cette façon sont de qualité à peine inférieure aux prévisions issues des modèles LTAR et CTAR où la modélisation auto-régressive des explicatives a été estimée sur une sous-période.

Estimer les paramètres de la modélisation auto-régressive sur une seconde sous-période uniquement, lorsque nos tests de modification du degré d'intégration le suggèrent, n'améliore pas beaucoup la qualité des prévisions.

Le principal avantage issu de cette troncature est un gain d'efficacité : les prévisions sont aussi bien réalisées en utilisant moins de données passées.

11.3.3 L'amélioration des explications des variations de trafic passées

11.3.3.1 Les enseignements tirés de la comparaison entre rétro-prévisions issues de l'analyse de la cointégration et véritables données de trafic

Les modèles issus de l'analyse de la cointégration correspondent aux meilleurs modèles de rétro-prévision en particulier pour les trafics du jour ouvrable sur les sous-réseaux ferrés. L'écart entre les rétro-prévisions réalisées à la fin de l'année 2002 pour l'ensemble de l'année 2003 et les valeurs réelles du trafic du jour ouvrable en 2003 mettent en valeur des explications des variations observées.

Le tableau ci-dessus résume les principaux écarts (supérieurs à 2% du trafic réel) entre le niveau de trafic réellement observé et ces rétro-prévisions :

Périodes des principaux écarts entre valeurs réelles et rétro-prévisions			
Jour & sous-réseau	Jour ouvrable	Samedi	Dimanche/Férié
Métro	⊖ : avril, juin-août	⊖ : juill., oct., déc. ⊕ : juin	⊖ : fév., avr., juin-août ⊕ : mai
RER	⊖ : juin-août ⊕ : nov.-déc.	⊖ : janv., juill., oct., déc.	⊖ : fév., juin-août, oct.
Bus de Paris	⊖ : juin-déc.	-	-
Bus de banlieue	⊕ : fév., nov	-	-

⊕ : valeur réelle \gg rétro-prévision, ⊖ : valeur réelle \ll rétro-prévision, de 2%

Sur les sous-réseaux ferrés comme sur le sous-réseau des autobus de Paris, le trafic du jour ouvrable a systématiquement été sur-estimé pour la période juin-août. L'impact sur les trafics de la diminution observée du nombre de touristes au cours de ces mois-là a été pris en compte dans les équations issues de l'analyse de la cointégration, par le biais de la variable du taux d'occupation des hôtels en Ile-de-France. Les journées de grèves à la RATP ou à la SNCF (en cas de correspondance avec le réseau ferré RATP) ont été écartées du calcul du jour ouvrable. En outre, effets calendaires et saisonniers ont été écartés.

CHAPITRE 11. DES MODÈLES COINTÉGRÉS SUR UNE SOUS-PÉRIODE PLUS ADAPTÉS AUX PRÉVISIONS

L'écart entre les valeurs de trafic réellement observées et les rétro-prévisions par type de journée et de sous-réseau mettent en lumière de nouveaux facteurs explicatifs des variations passées des trafics :

- En février 2003 se tenait le Forum social européen à Florence, et près de 15 millions de citoyens dans le monde manifestaient contre la guerre en Irak. Ce contexte perturbé explique la forte baisse de trafic au cours des week-ends des mois de février et avril. L'imminence de guerre ou la menace d'attentats bénéficient généralement au trafic des autobus.
- L'importance des grèves dans les transports en commun (SNCF et RATP) et dans la fonction publique au cours du mois de juin a entraîné une diminution de la fréquentation des transports collectifs. Même si le trafic du jour ouvrable est corrigé des journées de grèves dans les transports en commun, il est possible que le climat social et la crainte de nouvelles perturbations (grèves des transports collectifs ou manifestations sur la voie publique qui perturbent la circulation des autobus) aient dissuadé les utilisateurs de se déplacer. Les nombreuses grèves dans l'enseignement ont de plus eu pour conséquence de diminuer fortement les déplacements des scolaires et étudiants. L'analyse est identique pour la sous-estimation de trafic du jour ouvrable sur le métro au mois d'avril.

En outre, le beau temps au cours du mois de juin a sans doute favorisé les courts séjours des Franciliens dans le reste de la France. A l'inverse, ces conditions climatiques très favorables ont permis l'augmentation des déplacements lors des week-ends.

- En juillet et surtout août, la canicule a limité les déplacements des usagers des transports (transports en commun mais aussi voiture) : le temps passé dans les musées, magasins et autres lieux climatisés a fortement augmenté. En plus de la baisse du nombre de touristes, il existe une diminution de la mobilité de ces touristes, ainsi que des Franciliens. Cette baisse de mobilité se retrouve au cours des journées de week-ends sur les sous-réseaux ferrés.
- Sur la fin de l'année, le trafic du jour ouvrable se redresse plus vite sur le RER que le léger redressement de la situation économique et sociale ne le laissait présager. Ce redressement n'est pas observé au cours des journées de week-ends : il semble y avoir eu report des

déplacements du week-end vers les jours de semaine.

11.3.3.2 La rétro-prévision des trafics mensuels sur les sous-réseaux

Au cours du chapitre 3, nous avons décrit la méthode de calcul du trafic du jour ouvrable moyen pour chaque sous-réseau à partir des données journalières de l'ensemble du mois. Lors de ce calcul, pour un mois donné, les journées perturbées étaient écartées, et un équivalent de jours ouvrables était calculé (le "coefficient calendaire"). Le trafic du jour ouvrable pour un mois et un sous-réseau donné était ensuite désaisonnalisé à l'aide de la méthode du Census X-11, qui calcule des "coefficients saisonniers".

Ce chapitre a permis d'obtenir des rétro-prévisions du jour ouvrable moyen pour les sous-réseaux suivants : le métro (rétro-prévision associée à "*jomsa*"), le RER ("*joversa*"), le sous-réseau des autobus de Paris ("*joparsa*") ainsi que le sous-réseau des autobus de banlieue ("*jobansa*"). Il est possible de calculer les rétro-prévisions des trafics mensuels corrigés des grèves en utilisant la formule suivante :

$$\begin{aligned} \text{trafic prévu le mois } m &= \text{trafic du jour ouvrable prévu} \\ &\quad * \text{coefficient calendaire pour le mois } m \\ &\quad * \text{coefficient saisonnier associé au mois } m \\ &\quad (+ \text{ impact prévu des modifications de l'offre}) \end{aligned}$$

Il est souhaitable d'ajouter à ces rétro-prévisions l'impact prévu des modifications de l'offre pour l'année prévue. Par exemple, à la mi-décembre, la ligne 13 devait être (et a été) prolongée jusqu'à Saint-Lazare, ce qui devait entraîner un accroissement du trafic du métro de 50 000 voyageurs ce mois-là, et devait diminuer le trafic du RER de 100 000 personnes. En outre étaient prévues des modifications de l'offre sur le réseau des autobus de banlieue.

- Dans le cadre d'un exercice de rétro-prévision, les coefficients calendaires et saisonniers sont connus pour la période étudiée, *a posteriori*.
- Lorsque de véritables prévisions sont réalisées, les coefficients calendaires correspondent

aux coefficients calendaires calculés *a priori* (voir page 136 dans le paragraphe 3.3.4.1), et les coefficients saisonniers correspondent à ceux de la dernière année disponible.

La comparaison entre le trafic mensuel corrigé des perturbations et les rétro-prévisions de ce trafic mensuel peuvent amener des conclusions différentes en terme de choix du meilleur modèle de prévision. L'application des coefficients de passage modifie le poids des mois dans le trafic annuel : il ne s'agit pas seulement de bien prévoir l'ensemble des mois de l'année prévue, il est indispensable que les mois ayant le plus fort coefficient de passage soient prévus au mieux.

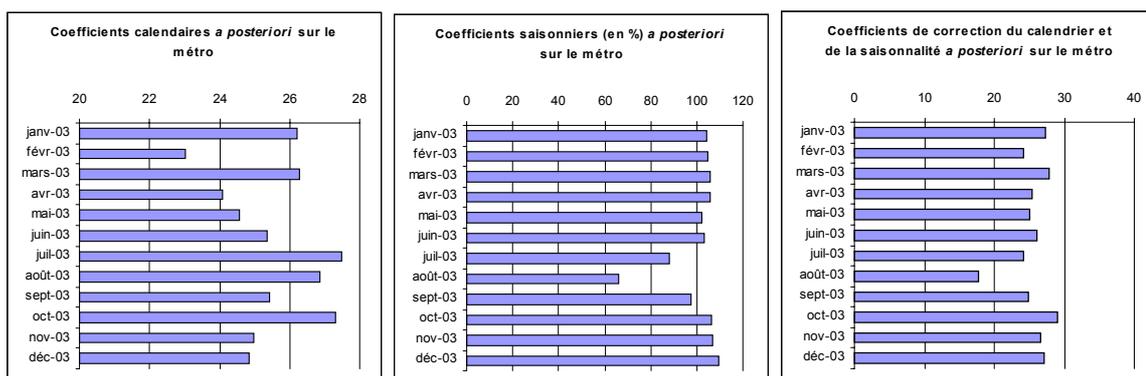
Nous verrons dans la suite de cette sous-section que le choix du modèle de prévision n'est pas remis en cause par la prise en compte des rétro-prévisions de trafic mensuel corrigé des perturbations au lieu du trafic du jour ouvrable. Cette analyse sera détaillée pour les rétro-prévisions du trafic mensuel corrigé des perturbations sur le métro, et plus succincte pour les autres types de sous-réseaux.

11.3.3.3 Rétro-prévisions du trafic mensuel corrigé des perturbations sur le métro

Dans les graphiques suivants, nous avons représenté les différents coefficients (*a posteriori*) de passage entre le trafic du jour ouvrable moyen désaisonnalisé (i.e. corrigé des perturbations, du calendrier et des effets saisonniers) et le trafic mensuel corrigé uniquement des perturbations. Le premier graphique correspond aux coefficients calendaires, le deuxième correspond aux coefficients saisonniers, tandis que le troisième correspond au produit des deux précédents coefficients.

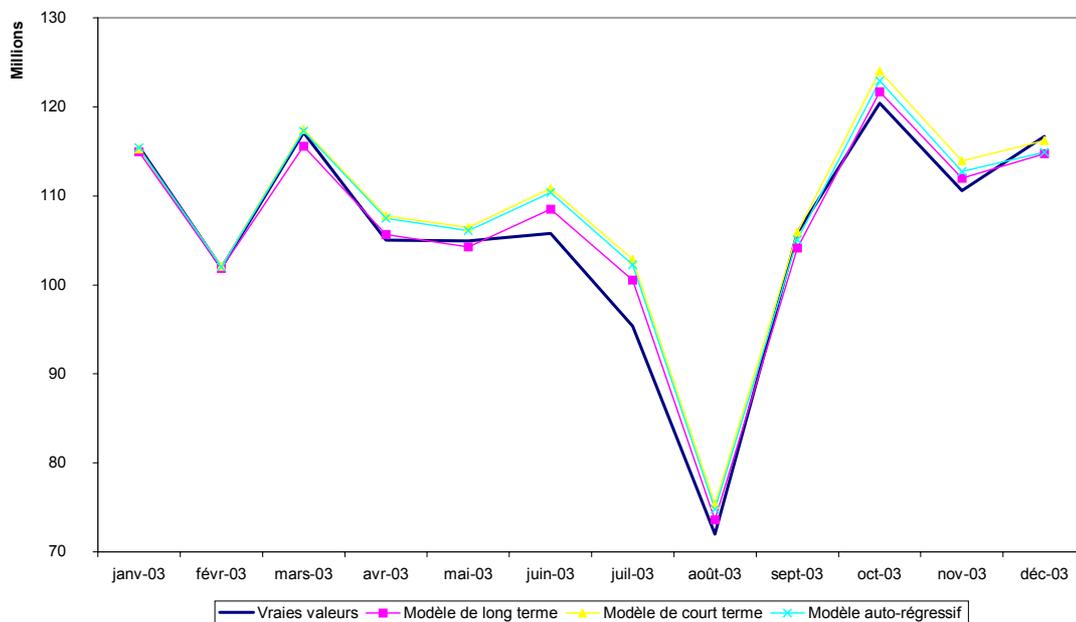
Les coefficients présentés dans le troisième graphique permettent de réintroduire la saisonnalité et le nombre équivalent de jours ouvrables au cours du mois considéré.

11.3. APPORTS DES RÉTRO-PRÉVISIONS



Le trafic mensuel rétro-prévu correspond, pour chaque mois, au produit du trafic du jour ouvrable rétro-prévu et du coefficient correcteur du calendrier et de la saisonnalité. Au mois de décembre 2003, 50 000 voyages ont été ajoutés à ces valeurs prévues (estimation par le réseau ferré) : ces voyages gagnés correspondent à l'impact au cours de ce mois là du prolongement de la ligne 14 à Saint-Lazare. Dans le graphique suivant, nous avons représenté le trafic mensuel corrigé des perturbations observé, ainsi que les rétro-prévisions de ces trafics mensuels à l'aide des trois modélisations utilisées.

Trafics mensuels corrigés des perturbations et rétro-prévus sur le métro



CHAPITRE 11. DES MODÈLES COINTÉGRÉS SUR UNE SOUS-PÉRIODE PLUS
ADAPTÉS AUX PRÉVISIONS

Par rapport aux rétro-prévisions du jour ouvrable sur le métro, les erreurs des rétro-prévisions au niveau du trafic du mois semblent moins importantes (de façon relative). La moins grande dispersion autour des valeurs observées de trafic s'explique d'une part par un effet d'échelle, et d'autre part par le fait que la plus grande part de la volatilité du trafic mensuel est expliqué par les effets saisonniers et calendaires.

Le tableau suivant indique quelques propriétés de l'erreur de prévision "normée", i.e. de la différence entre la rétro-prévision et la véritable valeur observée, rapportée au trafic observé. Ce calcul permet la comparaison entre les rétro-prévisions de trafic au niveau du jour ouvrable et au niveau du mois, qui n'ont pas la même échelle, mais aussi avec les autres types de trafic.

METRO	Propriétés de l'erreur de prévision normée e_t suivant le mode de calcul de la rétro-prévision								
	Type de rétro-prévision	Trafic du jour ouvrable			Trafic mensuel avec coefficients a priori			Trafic mensuel avec coefficients a posteriori	
Modélisation	LT	CT	AR	LT	CT	AR	LT	CT	AR
Moyenne de e_t	+0.9%	+1.4%	+2.0%	+0.0%	+0.2%	+1.2%	+0.6%	+1.7%	+1.7%
Ecart-type de e_t	0.013	0.014	0.018	0.016	0.022	0.022	0.021	0.026	0.025

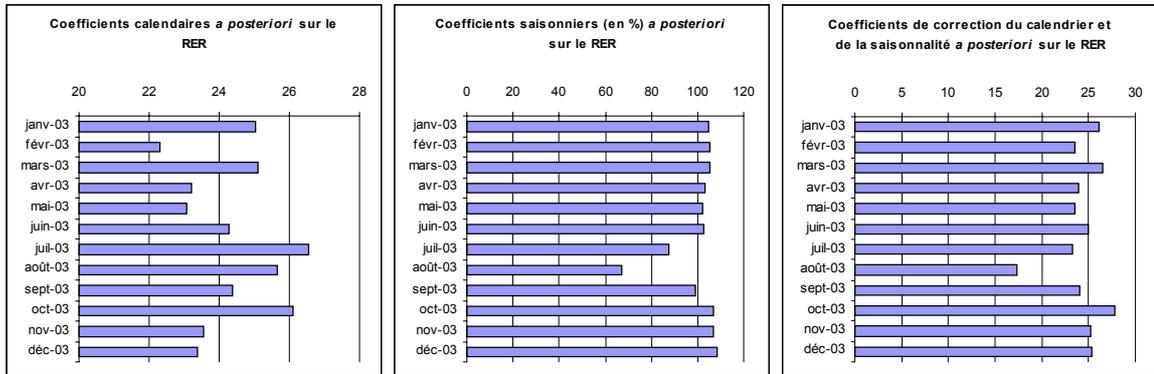
LT : Long terme, CT : Court terme (à correction d'erreur), AR : Auto-régressive

Par rapport aux rétro-prévisions du trafic du jour ouvrable, les rétro-prévisions du trafic mensuel sur le métro possèdent un biais systématique moins important. La différence entre rétro-prévision du trafic du jour ouvrable sur le métro et la véritable valeur observée était maximale au mois de juillet 2003⁴. Au niveau annuel, cette erreur de prévision est plus négligeable dans la mesure où le mois d'août contribue à moins de 6% du trafic annuel brut.

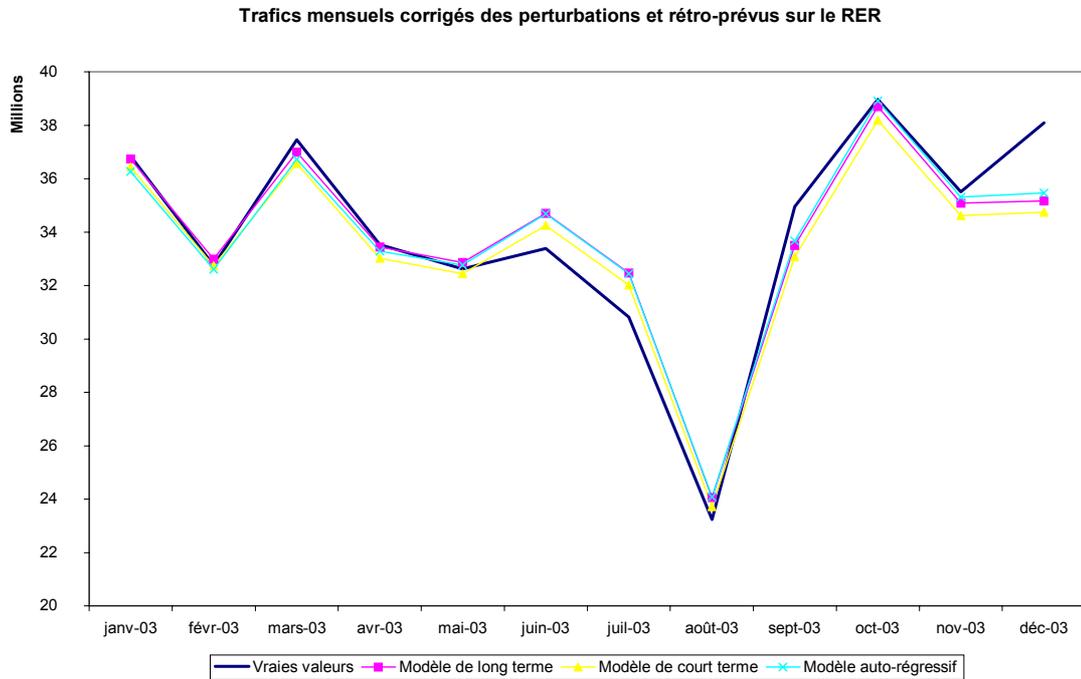
Enfin, que les coefficients de passage soient connus ou estimés, le modèle de long terme demeure le meilleur modèle de prévision : l'erreur de prévision est plus faible en moyenne et moins dispersée.

⁴Le trafic du jour ouvrable sur le métro, en juillet 2003, était surestimé de +3.76% par rapport aux valeurs réellement observées.

11.3.3.4 Rétro-prévisions du trafic mensuel corrigé des perturbations sur le RER



Au produit de la rétro-prévision du trafic du jour ouvrable par les coefficients de passage au niveau mensuel, nous avons ôté 100 000 voyages au cours du mois de décembre 2003, du fait du prolongement de Météor à Saint-Lazare (estimation du réseau ferré).



CHAPITRE 11. DES MODÈLES COINTÉGRÉS SUR UNE SOUS-PÉRIODE PLUS ADAPTÉS AUX PRÉVISIONS

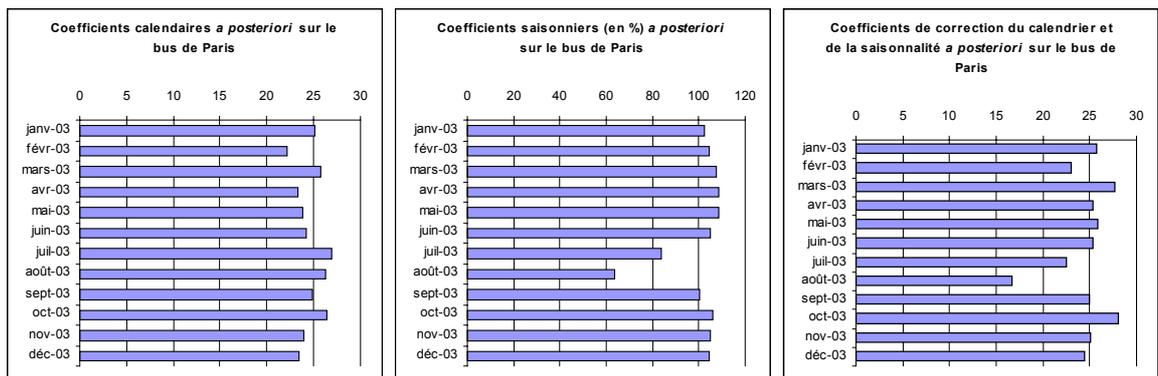
RER	Propriétés de l'erreur de prévision normée e_t suivant le mode de calcul de la rétro-prévision								
	Type de rétro-prévision	Trafic du jour ouvrable			Trafic mensuel avec coefficients a priori			Trafic mensuel avec coefficients a posteriori	
Modélisation	LT	CT	AR	LT	CT	AR	LT	CT	AR
Moyenne de e_t	+0.1%	-0.3%	-0.1%	-0.8%	-1.1%	-0.9%	-0.1%	-0.4%	-0.2%
Ecart-type de e_t	0.03	0.03	0.02	0.03	0.03	0.03	0.03	0.03	0.03

LT : Long terme, CT : Court terme (à correction d'erreur), AR : Auto-régressive

Que l'impact prévu de l'ouverture de Météor à Saint-Lazare soit pris en compte ou non dans les prévisions, et que les coefficients de passage soient calculés *a priori* ou *a posteriori*, le modèle de long terme demeure le meilleur modèle de prévision, avec un biais systématique plus faible pour une erreur de prévision aussi concentrée autour de 0.

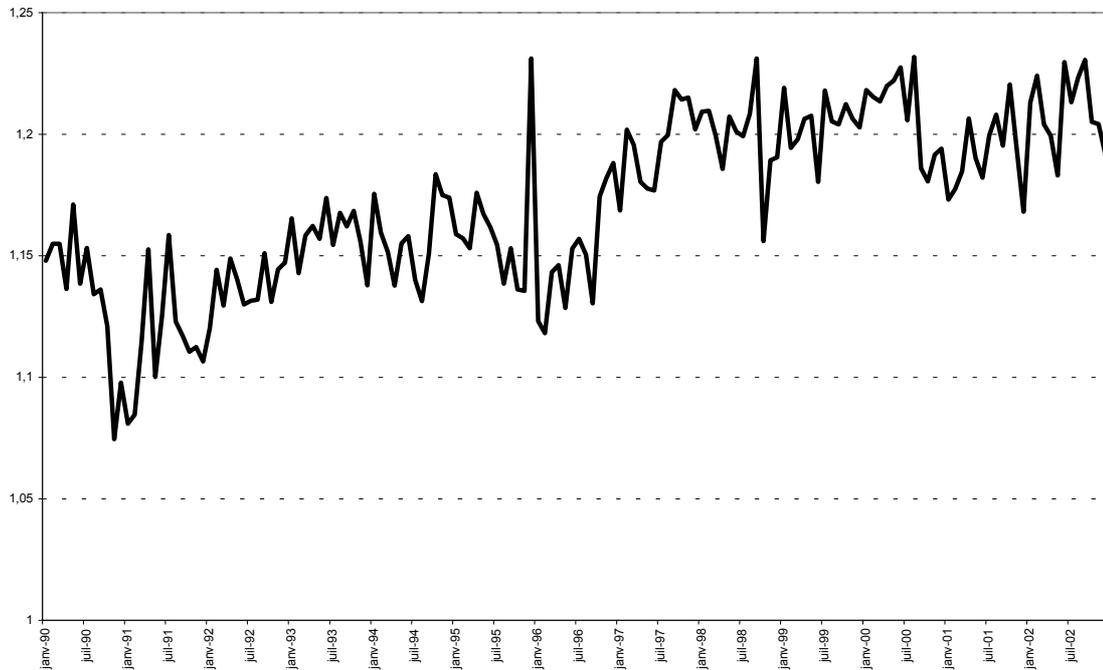
Contrairement au métro, le passage des rétro-prévisions du trafic du jour ouvrable au trafic mensuel corrigé des grèves n'améliore pas réellement l'erreur de prévision normée : l'erreur de prévision du jour ouvrable sur le RER la plus importante se produit en décembre 2003. La légère embellie économique pousse les gens à emprunter les transports en commun, en particulier le RER, dès le mois d'octobre. Ces mois correspondent à des mois de trafic plein : l'erreur de prévision de ce type de jour ouvrable n'est pas gommée par les poids des différents mois.

11.3.3.5 Rétro-prévisions du trafic mensuel corrigé des perturbations sur le réseau des autobus de Paris



En 2003, aucun impact de modification de l'offre ou de lutte contre la fraude n'a été utilisé

pour les rétro-prévisions du trafic sur le réseau des autobus de Paris. La rétro-prévision à un niveau mensuel et non simplement journalier correspond chaque mois au produit de la prévision du trafic du jour ouvrable et du coefficient de passage au mois.



AUTOBUS DE PARIS	Propriétés de l'erreur de prévision normée e_t suivant le mode de calcul de la rétro-prévision								
	Type de rétro-prévision	Trafic du jour ouvrable			Trafic mensuel avec coefficients a priori			Trafic mensuel avec coefficients a posteriori	
Modélisation	LT	CT	AR	LT	CT	AR	LT	CT	AR
Moyenne de e_t	+3.5%	+2.9%	+1.3%	+4.3%	+3.7%	+2.1%	+4.3%	+3.7%	+2.1%
Ecart-type de e_t	0.04	0.04	0.03	0.04	0.04	0.03	0.04	0.04	0.04

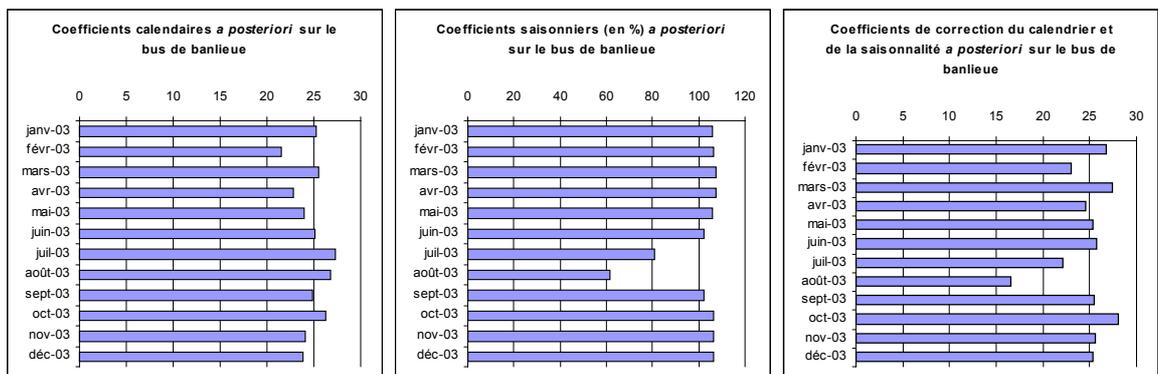
LT : Long terme, CT : Court terme (à correction d'erreur), AR : Auto-régressive

Prendre en compte l'ensemble du trafic observé au cours du mois au lieu de celui observé en moyenne au cours d'un jour ouvrable introduit plus d'erreur de prévision. Sur le trafic du jour ouvrable, la principale erreur de prévision a lieu au cours du mois d'octobre, mois qui a un fort poids saisonnier et calendaire. Le fait que le trafic ne remonte pas en fin d'année, alors que les

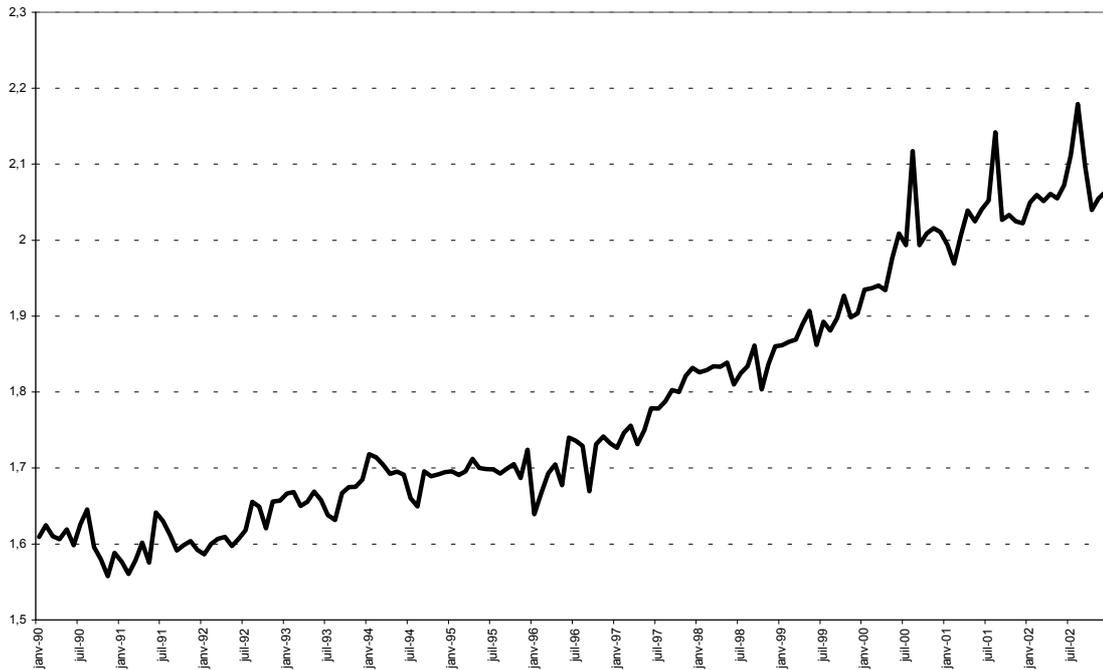
CHAPITRE 11. DES MODÈLES COINTÉGRÉS SUR UNE SOUS-PÉRIODE PLUS ADAPTÉS AUX PRÉVISIONS

conditions économiques sont réunies pour que la reprise ait lieu, signale soit un effet rémanent des grèves, soit un problème dans le comptage des données, soit un impact de la lutte contre la fraude.

11.3.3.6 Rétro-prévisions du trafic mensuel corrigé des perturbations sur le réseau des autobus de banlieue



Le produit de la rétro-prévision du trafic du jour ouvrable par les coefficients de passage au niveau mensuel a été augmenté par des prévisions faites par le réseau routier, et concernant l'augmentation des voyageurs induite par l'accroissement de l'offre par avenant sur le réseau des autobus de banlieue.



AUTOBUS DE BANLIEUE	Propriétés de l'erreur de prévision normée e_t suivant le mode de calcul de la rétro-prévision								
	Trafic du jour ouvrable			Trafic mensuel avec coefficients a priori			Trafic mensuel avec coefficients a posteriori		
Type de rétro-prévision									
Modélisation	LT	CT	AR	LT	CT	AR	LT	CT	AR
Moyenne de e_t	-1.5%	-1.5%	-0.5%	-1.8%	-1.8%	-0.8%	-1.5%	-1.5%	-0.5%
Ecart-type de e_t	0.01	0.01	0.01	0.03	0.02	0.02	0.03	0.02	0.02

LT : Long terme, CT : Court terme (à correction d'erreur), AR : Auto-régressive

La prise en compte des rétro-prévisions a un niveau mensuel ou pour un journée ouvrable type de ce mois ne modifie pas la qualité des rétro-prévisions pour ce type de sous-réseau. Aucune prévision ne s'éloigne beaucoup de la valeur observée mois ni sur le jour ouvrable ni sur le mois entier.

* * *

CHAPITRE 11. DES MODÈLES COINTÉGRÉS SUR UNE SOUS-PÉRIODE PLUS ADAPTÉS AUX PRÉVISIONS

Dans ce chapitre, les modélisations auto-régressive, de long terme et de court terme, ont été utilisées pour prévoir les trafics de l'année 2003, dont les mois n'ont pas été utilisés au cours de l'estimation. Cette analyse de rétro-prévision est un moyen simple de voir que nos modèles sont correctement calibrés et conduisent à des prédictions satisfaisantes, même pour une année fortement perturbée comme le fut l'année 2003 (grèves et canicule). Elle nous apprend en outre que les modèles issus de l'analyse de la cointégration, i.e. les modèles de court et de long terme, expliquent de façon plus satisfaisante les variations de trafic observées sur les sous-réseaux ferrés au cours de l'année considérée. Ces modèles en fonction de variables d'environnement permettent de mettre en évidence certains phénomènes ne dépendant pas de la conjoncture économique, comme les variations de mobilité au cours de la canicule, l'existence d'effets rémanents des grèves ou l'impact d'une politique de lutte contre la fraude.

Ces divers éléments justifient le bien fondé de nos modélisations en fonction de variables économiques et non seulement en fonction des valeurs passées du trafic. De ce fait, ils permettent d'établir l'intérêt de notre travail de construction de tests de modification du degré de cointégration, sans lesquels aucune relation de cointégration n'aurait été trouvée de façon automatique. Seuls les trafics du réseau de surface sont mieux prévus grâce à leur modélisation auto-régressive. Ceci provient du fait que ce sont les seuls modèles qui ont été estimés sans rupture.

Enfin, ce chapitre montre aussi l'apport des tests de modification du degré d'intégration : pour les séries intégrées sur la seconde sous-partie uniquement alors qu'elles auraient été testées intégrées sur l'ensemble de la période, les rétro-prévisions sont améliorées ou inchangées en prenant les paramètres estimés sur cette seconde sous-période uniquement. Obtenir un résultat identique en utilisant un moins grande nombre de données apporte un gain de parcimonie.

Chapitre 12

Prévisions à l'horizon 2004 - 2005

Dans le chapitre précédent, les modèles conduisaient à des rétro-prévisions de qualité pour l'année 2003. Il s'agit désormais d'effectuer des prévisions de trafic pour les mois de 2004 et 2005, en introduisant la possibilité de scénarios dans les prévisions et en évaluant le risque associé à cette prévision ponctuelle, par l'évaluation d'intervalles de prévision si possible réduits.

Pour effectuer les prévisions à partir de 2004, les modèles seront mis à jour en prenant en compte les données d'environnement et de trafic observées au cours de l'année 2003 (première section).

Il est possible de modifier les équations auto-régressives estimées de façon à introduire différents scénarios d'évolution de la variable endogène ou des variables exogènes (deuxième section).

La méthode de Monte Carlo "calibrée", dans la troisième section, permettra le calcul d'intervalles de confiance simulés.

La quatrième section concerne l'application de ces méthodes aux équations de trafic de la RATP par type de journée et de sous-réseau. Les prévisions ponctuelles seront calculées pour chacun de ces trafics, en supposant un scénario d'évolution des données plausible, de même que les intervalles de confiance simulés. Nous présenterons en outre les prévisions de trafic par sous-réseau pour les années 2004 et 2005 pour l'ensemble du mois et non plus seulement pour un type de journée. Ce type de données est utilisé lors de l'exercice de prévision des objectifs budgétaires au sein du département Commercial de la RATP.

12.1 Mise à jour des estimations des modèles

Les données d'environnement en Ile-de-France et de trafic ont été mises à jour jusqu'à la fin 2003, désaisonnalisées, puis log-linéarisées si les séries ne sont pas exprimées en terme de ratio ou d'indice. Toutes ces séries conservent un degré d'intégration égal à 1 sur l'ensemble de la période.

Les équations (auto-régressives, long terme, court terme) de chaque trafic ont été réestimées en prenant des variables identiques à celles prises lors de l'estimation jusqu'à la fin 2002. Le degré de cointégration des modélisations de long terme a été vérifié.

Dans cette première section sont présentés les résultats de ces nouvelles estimations et des tests de cointégration. Les éventuelles modifications des élasticités sont expliquées.

12.1.1 Trafic du jour ouvrable sur le métro

Le tableau suivant présente le résultat des nouvelles estimations, pour chacun des modèles :

Modèle	Début	Equation
AR	01/90	$\Delta \ln(jomsa)_t = 0.0003 - 0.3795 \Delta \ln(jomsa)_{t-1}$ $(0.2179) \quad (-4.9220)$ $- 0.2091 \Delta \ln(jomsa)_{t-2}$ (-2.7111)
LT	01/96	$\ln(jomsa)_t = -0.1234 \ln(defmidf)_t + 0.2805 txhotidf_t$ $(-5.6760) \quad (4.7559)$ $+ 16.5689 + 0.0013 t$ $(53.3545) \quad (14.7100)$ <p>Relation de cointégration : $t_{CADF} = -4.2053 \leq VC_{5\%} = -4.2249$</p>
CT	01/96	$\Delta \ln(jomsa)_t = 0.0869 \Delta \ln(defmidf)_t - 0.7560 z_{t-1} + 0.0015$ $(0.9719) \quad (-6.1932) \quad (1.0767)$

Dans ce tableau figure les estimations de chacune des modélisations :

- "AR" pour le modèle auto-régressif,
- "LT" pour le modèle de long terme,
- et "CT" pour le modèle de court terme (à correction d'erreur).

Les indications suivantes ont été ajoutées :

- Les résultats du test de cointégration CADF (conclusion du test, statistique de test t_{CADF} et valeur critique à $\alpha\%$ notée $VC_{\alpha\%}$) sont indiqués juste en dessous de l'équation de long terme.
- Chaque modèle est estimé sur période entière ou la seconde sous-période uniquement, conformément aux estimations réalisées dans les chapitres précédents. La date de début de l'estimation pour chaque modélisation est précisée dans le tableau synthétique.
- Les statistiques de Student sont notées entre parenthèses en dessous de chaque paramètre estimé.

L'équation de long terme est peu modifiée par la mise à jour des données. L'élasticité du trafic du jour ouvrable sur le métro au tourisme a légèrement augmenté : durant l'année 2003, les touristes à Paris ont moins voyagé sur le métro du fait de la canicule. A court terme, seule la force de rappel demeure significative.

12.1.2 Trafic du jour ouvrable sur le RER

Les réestimations des modélisations du trafic du jour ouvrable sur le RER jusqu'à la fin 2003 sont présentées dans le tableau suivant :

12.1. MISE À JOUR DES ESTIMATIONS DES MODÈLES

Modèle	Début	Equation
AR	12/93	$\Delta \ln(jorersa)_t = 0.0017 - 0.2299 \Delta \ln(jorersa)_{t-1}$ <p style="text-align: center;">(1.1590) (-2.5097)</p>
LT	11/97	$\ln(jorersa)_t = -0.1275 \ln(defmidf)_t + 0.0015 gdmgrp_t$ <p style="text-align: center;">(-3.6432) (3.0006)</p> $+15.6266 + 0.0011 t$ <p style="text-align: center;">(32.3452) (6.5895)</p> <p>Relation cointégrée : $t_{CADF} = -3.9352 \geq VC_{10\%} = -3.9267$</p>
CT	11/97	$\Delta \ln(jorersa)_t = -0.1400 \Delta \ln(defmidf)_t + 0.0847 \Delta txhotidf_t$ <p style="text-align: center;">(-1.0105) (0.9488)</p> $-0.4647 z_{t-1} + 0.0013$ <p style="text-align: center;">(-6.3569) (0.7629)</p>

Sur le RER, pour l'estimation jusqu'à la fin de l'année 2002, le critère BIC amenait à ne prendre en compte qu'une seule variable environnementale dans l'équation de long terme : les demandes d'emplois en fin de mois en Ile-de-France. Les tests de cointégration CADF associés à l'estimation jusqu'à fin 2003 avec cette seule variable ne permettent pas de rejeter l'hypothèse d'absence de cointégration.

Pour retrouver un équilibre de long terme non fallacieux, nous avons ajouté la variable décrivant le volume des ventes dans les grands magasins de la région parisienne.

L'impact négatif sur le trafic de l'augmentation du chômage s'ajoute à l'effet dépressif d'une diminution de la consommation des ménages.

A court terme, seule la force de rappel est significative. Rappelons que cette forte significativité est une preuve de la présence de cointégration dans la relation de long terme présentée.

12.1.3 Trafic du jour ouvrable sur le réseau des autobus de Paris

Modèle	Début	Equation
AR	01/90	$\Delta \ln(joparsa)_t = 0.0006 \quad -0.6062 \quad \Delta \ln(joparsa)_{t-1}$ $(0.4134) \quad (-7.8689)$ $-0.3927 \quad \Delta \ln(joparsa)_{t-2} \quad -0.1956 \quad \Delta \ln(joparsa)_{t-2}$ $(-4.6139) \quad \quad \quad (-2.5349)$
LT	01/90	$\ln(joparsa)_t = -0.1332 \quad \ln(circ75)_t \quad -0.0753 \quad \ln(speed75)_t$ $(-1.9002) \quad \quad \quad (-1.4677)$ $+0.1597 \quad \ln(defmidf)_t + 0.0012 \quad gdmagrpt$ $(6.6445) \quad \quad \quad (2.2499)$ $+0.1392 \quad txhotidf_t + 12.8973$ $(2.8252) \quad \quad \quad (23.3269)$ <p>Relation de cointégration : $tCADF = -5.5133 \leq VC_{5\%} = -4.8852$</p>
CT	01/90	$\Delta \ln(joparsa)_t = -0.1479 \quad \Delta \ln(speed75)_t + 0.0753 \quad \Delta gdmagrpt$ $(-2.9437) \quad \quad \quad (1.0406)$ $-0.6517 \quad z_{t-1} + 0.0003$ $(-8.6600) \quad \quad \quad (0.2576)$

Sur le réseau des autobus de Paris, les nouvelles estimations conduisent à des élasticités identiques à celles obtenues au cours de l'estimation avec un échantillon d'observation se terminant en décembre 2002. Deux exceptions toutefois :

- L'élasticité du trafic du bus de Paris au volume de circulation dans Paris a diminué en valeur absolue. Cette diminution traduit l'impact de l'augmentation des lignes protégées du reste de la circulation par des couloirs de bus.
- L'élasticité du trafic par rapport au taux d'occupation des hôtels en Ile-de-France a augmenté, comme c'était le cas pour le trafic du jour ouvrable sur le métro. En plus de la baisse du nombre de touristes durant la canicule, ces touristes se sont déplacés moins

fréquemment.

A court terme, l'élasticité du trafic au volume des ventes dans les grands magasins de la région parisienne s'est fortement accrue par rapport à l'estimation jusqu'à la fin 2002. Cet accroissement résulte du même phénomène de diminution des déplacements : touristes et franciliens ont passé de nombreuses heures dans les différents magasins climatisés, sans acheter à chaque fois.

12.1.4 Trafic du jour ouvrable sur le réseau des autobus de banlieue

Les résultats des estimations sur la période janvier 1990 - décembre 2003 conduisent aux équations suivantes :

Modèle	Début	Equation
AR	01/90	$\Delta \ln (jobansa)_t = 0.0028 - 0.4225 \Delta \ln (jobansa)_{t-1}$ <p style="text-align: center;">(2.5359) (-5.5101)</p> $- 0.2249 \Delta \ln (jobansa)_{t-2}$ <p style="text-align: center;">(-2.9187)</p>
LT	01/90	$\ln (jobansa)_t = -0.0736 \ln (defmidf)_t + 0.0015 gdmagrpt$ <p style="text-align: center;">(-3.8826) (4.6648)</p> $+ 15.0748 + 0.0019 t$ <p style="text-align: center;">(56.7157) (39.2696)</p> <p>Relation de cointégration : $t_{CADF} = -5.8998 \leq VC_{5\%} = -4.2429$</p>
CT	01/90	$\Delta \ln (jobansa)_t = 0.0001 \Delta gdmagrpt - 0.5948 z_{t-1} + 0.0017$ <p style="text-align: center;">(0.5003) (-8.1998) (1.7608)</p>

Sur le réseau des autobus de banlieue, les élasticités du trafic au différentes variables décrivant l'environnement économique ne sont pas modifiées par la mise à jour des données jusqu'à la fin de l'année 2003.

12.1.5 Trafic du samedi sur le métro

Modèle	Début	Equation
AR	04/94	$\Delta \ln(sm\text{sa})_t = 0.0025 - 0.4893 \Delta \ln(sm\text{sa})_{t-1}$ <p style="text-align: center;">(0.9977) (-6.1487)</p>
LT	03/98	$\ln(sm\text{sa})_t = 0.0012 \text{ gdmagr}p_t + 0.4066 \text{ txhotidf}_t$ <p style="text-align: center;">(2.5024) (2.9564)</p> $+ 14.2804 + 0.0023 t$ <p style="text-align: center;">(164.07721) (11.7189)</p> <p>Relation cointégrée : $t_{CADF} = -6.8671 \leq VC_{5\%} = -4.2429$</p>
CT	03/98	$\Delta \ln(jom\text{sa})_t = 0.0008 \Delta \text{gdmagr}p_t + 0.1612 \Delta \text{txhotidf}_t$ <p style="text-align: center;">(1.8380) (1.8957)</p> $- 0.1125 \Delta \text{compet}_t - 0.8366 z_{t-1} + 0.0015$ <p style="text-align: center;">(-1.0960) (-6.2709) (0.5297)</p>

Les équations du trafic du samedi sur le métro ont été peu modifiées par la mise à jour de l'échantillon des données jusqu'à la fin de l'année 2003, à long terme comme à court terme.

- A long terme,
 - L'élasticité du trafic du samedi sur le métro au volume de ventes dans les grands magasins de la région parisienne a légèrement diminué. Cette évolution des paramètres diffère des autres types de trafic où la prise en compte de 2003 conduisait à un accroissement de l'élasticité au ventes dans les grands magasins. Elle peut signifier que les utilisateurs du métro ont reporté leurs achats du samedi à un autre jour de la semaine (jour ouvrable ou dimanche), i.e. qu'ils ont réduit leurs déplacements pour ce motif le samedi.
 - L'élasticité de ce type de trafic au taux d'occupation des hôtels en Ile-de-France s'est accrue, comme pour les autres types de trafic.
- A court terme,

- La variation du volume des ventes dans les grands magasins de la région parisienne engendre une variation plus faible de trafic que dans l'estimation jusqu'à fin 2002.
- Une variation du taux d'occupation des hôtels en Ile-de-France a un impact moins élevé sur le trafic du samedi sur le métro.

12.1.6 Trafic du dimanche/jour férié sur le métro

Modèle	Début	Equation
AR	01/90	$\Delta \ln(dfmsa)_t = 0.0027 - 0.5807 \Delta \ln(dfmsa)_{t-1}$ <p style="text-align: center;">(1.1152) (-7.6689)</p> $- 0.2690 \Delta \ln(dfmsa)_{t-2}$ <p style="text-align: center;">(-3.5537)</p>
LT	01/96	$\ln(dfmsa)_t = 0.0015 gdmagrpt + 0.3708 txhotidf_t$ <p style="text-align: center;">(2.7847) (3.9921)</p> $+ 13.7262 + 0.0033 t$ <p style="text-align: center;">(239.6318) (22.4955)</p> <p>Relation cointégrée : $t_{CADF} = -10.9410 \leq VC_{5\%} = -4.2429$</p>
CT	01/96	$\Delta \ln(dfmsa)_t = -0.0833 \Delta \ln(defmidf)_t + 0.3149 \Delta txhotidf_t$ <p style="text-align: center;">(-0.3984) (5.9776)</p> $- 0.9248 z_{t-1} + 0.0039$ <p style="text-align: center;">(-9.6479) (1.5225)</p>

Pour le dimanche ou jour férié sur le métro, la réestimation des paramètres conduit à des résultats similaires à ceux issus de l'estimation des équations à partir des données jusqu'à fin 2002.

A long terme, l'élasticité par rapport au volume des ventes dans les grands magasins ou au taux d'occupation des hôtels en Ile-de-France a légèrement augmenté, pour des raisons évoquées précédemment.

A court terme, les demandes d'emplois en fin de mois en Ile-de-France, peu significatives dans l'estimation jusqu'à fin 2002, semblent ne jouer désormais qu'un rôle marginal.

12.1.7 Trafic du samedi sur le RER

Modèle	Début	Equation
AR	01/90	$\Delta \ln(srersa)_t = 0.0036 - 0.6263 \Delta \ln(srersa)_{t-1}$ <p style="text-align: center;">(1.3398) (-8.6160)</p> $- 0.3724 \Delta \ln(srersa)_{t-2}$ <p style="text-align: center;">(-5.1108)</p>
LT	07/98	$\ln(srersa)_t = 0.0036 gdmagrpt + 13.0363 + 0.0023 t$ <p style="text-align: center;">(5.9519) (221.2285) (12.4983)</p> <p>Relation de cointégration : $t_{CADF} = -7.2266 \leq VC5\% = -3.9157$</p>
CT	07/98	$\Delta \ln(srersa)_t = 0.0017 \Delta gdmagrpt - 0.9526 z_{t-1} + 0.0017$ <p style="text-align: center;">(3.1629) (-6.9805) (0.5550)</p>

Les élasticités du trafic du samedi sur le RER au volume des ventes dans les grands magasins de la région parisienne ne se sont pas modifiées lorsque les données de l'année 2003 ont été ajoutées à la période d'estimation. Seule une légère diminution de l'élasticité à court terme est perceptible.

12.1.8 Trafic du dimanche/jour férié sur le RER

Modèle	Début	Equation
AR	03/98	$\Delta \ln(dfersa)_t = 0.0057 - 0.2642 \Delta \ln(dfersa)_{t-1}$ <p style="text-align: center;">(1.6564) (-2.2751)</p>
LT	07/98	$\ln(dfersa)_t = 0.0032 gdmagrpt + 12.5596 + 0.0033 t$ <p style="text-align: center;">(4.8226) (196.9543) (16.6676)</p> <p>Relation de cointégration : $t_{CADF} = -6.4140 \leq VC5\% = -3.9157$</p>
CT	07/98	$\Delta \ln(dfersa)_t = 0.0185 \Delta \ln(defmidf)_t - 0.7908 z_{t-1} + 0.0003$ <p style="text-align: center;">(0.0957) (-7.8997) (1.4403)</p>

La réestimation des paramètres en prenant en compte les données de l'année 2003 ne permet pas de conclure à la modification des élasticités du trafic du dimanche ou jour férié au volume des ventes dans les grands magasins de la région parisienne. Seule la force de rappel demeure significative à court terme.

Pour chaque type de trafic, les paramètres de long terme et de court terme sont modifiés de façon marginale par l'ajout des données de 2003. Le paramètre associé à la tendance et la constante sont stables, de même que la force de rappel dans l'équation de court terme. Les modifications mineures constatées concernent les paramètres associés à la variable de tourisme, du fait de la baisse de mobilité constatée au cours de la canicule.

12.1.9 Données décrivant l'environnement économique en Ile-de-France

Les estimations auto-régressives des données environnementales ont été mises à jour en prenant en compte les données jusqu'à la fin 2003.

Début	Equation auto-régressive			
janv. 1990	$\Delta \ln(circ75)_t = -$	0.0038	$-$	$0.3435 \Delta \ln(circ75)_{t-1}$
		(-2.3665)		(-3.0139)
mars 1996	$\Delta \ln(speed75)_t = -$	0.0006	$-$	$0.5344 \Delta \ln(speed75)_{t-1} - 0.2523 \Delta \ln(speed75)_{t-2}$
		(-0.1730)		(-7.0402) (-3.3371)
janv. 1992	$\Delta \ln(defmidf)_t = +$	0.0003	$-$	$0.0610 \Delta \ln(defmidf)_{t-1} + 0.3777 \Delta \ln(defmidf)_{t-2}$
		(0.2567)		(-0.5210) (3.5908)
			$+ 0.5459 \Delta \ln(defmidf)_{t-3}$	
			(4.6627)	
mai 1992	$\Delta gdmagrpt = +$	0.1572	$-$	$0.6975 \Delta gdmagrpt_{t-1} - 0.3813 \Delta gdmagrpt_{t-2}$
		(0.4733)		(-8.7015) (-4.6815)
janv. 1998	$\Delta txhotidf_t = -$	0.0000	$-$	$0.2825 \Delta txhotidf_{t-1}$
		(-0.0040)		(-2.4311)
janv. 1990	$\Delta compet_t = -$	0.0009	$+$	$0.0236 \Delta compet_{t-1}$
		(-0.3089)		(0.1968)

Les séries sont intégrées sur l'ensemble de la période ou sur une seconde sous-période uniquement (voir chapitre 7). La date de début de période est issue de nos tests de modification du degré d'intégration.

Les estimations auto-régressives permettent d'estimer le taux de variation des variables économiques sur l'ensemble de la période d'estimation. Ces taux d'évolution passés interviennent dans les prévisions : les variations futures de chaque variable seront identiques aux variations passées dans les modélisations auto-régressives. Il est toutefois possible d'introduire des scénarios d'évolution différents.

12.2 L'introduction de scénarios

12.2.1 Prévisions auto-régressives avec scénarios

Les prévisions auto-régressives supposent que l'évolution future de la variable est identique à l'évolution observée par le passé. Pour prendre en compte une modification du taux de variation τ de la série, le terme constant dans la modélisation auto-régressive en log-niveau doit être modifié.

Considérons par exemple un processus auto-régressif en différence d'ordre 1 (i.e. un processus $ARI(p)$) :

$$\Delta \ln(x_t) = c + a_1 \Delta \ln(x_{t-1}) + a_2 \Delta \ln(x_{t-2}) \dots + a_p \Delta \ln(x_{t-p}) + \varepsilon_t$$

où ε_t est un processus d'innovation de moyenne nulle et de variance σ_ε^2 .

Du fait de la propriété suivante de la transformation logarithmique

$$\Delta \ln(x_t) = \ln(x_t) - \ln(x_{t-1}) = \ln\left(\frac{x_t}{x_{t-1}}\right) \tag{12.1}$$

le processus $ARI(p)$ se réécrit sous la forme suivante :

$$\ln\left(\frac{x_t}{x_{t-1}}\right) = c + a_1 \ln\left(\frac{x_{t-1}}{x_{t-2}}\right) + a_2 \ln\left(\frac{x_{t-2}}{x_{t-3}}\right) \dots + a_p \ln\left(\frac{x_{t-p}}{x_{t-p-1}}\right) + \varepsilon_t \quad (12.2)$$

Or

$$\ln\left(\frac{x_t}{x_{t-1}}\right) = \ln\left(\frac{x_t - x_{t-1} + x_{t-1}}{x_{t-1}}\right) = \ln\left(\frac{x_t - x_{t-1}}{x_{t-1}} + 1\right) \quad (12.3)$$

et, au voisinage de 0,

$$\ln\left(\frac{x_t}{x_{t-1}}\right) \approx \frac{x_t - x_{t-1}}{x_{t-1}} = \tau \text{ (taux de croissance de la série } x_t) \quad (12.4)$$

Le processus auto-régressif en log-différence se réécrit

$$\tau = c + a_1\tau + a_2\tau + \dots + a_p\tau \quad (12.5)$$

et le taux de croissance par rapport aux paramètres de la modélisation a pour écriture

$$\tau = \frac{c}{1 - a_1 - a_2 \dots - a_p} \quad (12.6)$$

L'introduction d'un scénario se traduit par une modification de la constante du modèle auto-régressif :

$$\Delta y_t = c_{\text{Scénario}} + a_1\Delta y_{t-1} + a_2\Delta y_{t-2} + \dots + a_p\Delta y_{t-p} + \varepsilon_t \quad (12.7)$$

où

$$c_{\text{Scénario}} = \tau(1 - a_1 - a_2 - \dots - a_p) \quad (12.8)$$

Cet coefficient correspond à un taux de variation mensuel moyen, sur la période considérée. Pour passer d'une évolution mensuelle $\tau = \tau_m$ à un taux de variation annuel moyen τ_a , la

formule suivante est utilisée :

$$\tau_a = (1 + \tau_m)^{12} - 1 \quad (12.9)$$

12.2.2 Evolution des variables dans le passé prise en compte dans le scénario implicite

Les estimations auto-régressives précédentes permettent de calculer les taux de variation moyen des variables décrivant le trafic ou l'environnement en Ile-de-France au cours de la période précédente. Ce scénario implicite est supposé se reproduire dans le futur.

Selon ce scénario implicite, les évolutions présentées dans le tableau ci-dessous sont les suivantes :

- Pour les données de trafic suivant le sous-réseau et le type de journée :
 - Relative stabilité du taux de croissance du trafic du jour ouvrable sur le métro et le réseau des autobus de Paris,
 - Accroissement plus marqué sur le RER et surtout sur le réseau des bus de banlieue,
 - Croissance du trafic des week-ends, plus importante le dimanche, et sur le réseau du RER.

- Pour les données décrivant l'environnement en Ile-de-France :
 - Baisse des nuisances liées à l'utilisation de la voiture dans Paris (baisse du volume de circulation mais aussi de la vitesse), dans un contexte de renchérissement du prix de l'essence par rapport au prix des transports collectifs urbains,
 - Situation économique peu favorable (augmentation du nombre de demandes d'emplois en fin de mois en Ile-de-France et stagnation du taux d'occupation des hôtels) soutenue par la consommation des ménages.

EVOLUTION DES DONNEES SIMILAIRE A CELLE OBSERVEE DANS LE PASSE
UTILISEE DANS LES MODELISATIONS AUTO-REGRESSIVES

Taux d'accroissement moyen estimé	Mensuel	Annuel
Données de trafic		
Jour ouvrable sur le métro	+0.02%	+0.21%
Jour ouvrable sur le RER	+0.09%	+1.06%
Jour ouvrable sur les bus de Paris	+0.02%	+0.30%
Jour ouvrable sur les bus de banlieue	+0.17%	+2.03%
Samedi sur le métro	+0.09%	+1.12%
Dimanche/férié sur le métro	+0.25%	+3.05%
Samedi sur le RER	+0.18%	+2.16%
Dimanche/férié sur le RER	+0.23%	+2.75%
Données d'environnement		
Volume de circulation dans Paris	-0.29%	-3.38%
Vitesse de circulation dans Paris	-0.03%	-0.38%
Demandes d'emplois en fin de mois (IdF*)	+0.25%	+2.99%
Volume des ventes dans les grands magasins (IdF*)	+0.08%	+0.96%
Taux d'occupation des hôtels en IdF*	-0.00%	-0.01%
Rapport entre le prix du ticket T et de l'essence	-0.09%	-1.07%

* IdF : Ile-de-France

12.2.3 Eléments de perspective en Ile-de-France à fin 2003

En décembre 2003, l'INSEE prévoyait une accélération de la croissance en 2004¹ de +2% sur l'ensemble de l'année au niveau national (+1.7% dans le budget 2004). Le taux de chômage aurait du demeurer sous la barre des 10%, et les effets négatifs d'une éventuelle augmentation de l'euro par rapport au dollar auraient du être compensés par la reprise de la demande mondiale.

Le schéma de reprise au niveau national est le suivant : la reprise du commerce extérieur

¹Source : "Le 18h.com" (édition informatique quotidienne du journal L'Expansion), n°1186, daté du 17 décembre 2003, disponible sur www.l'expansion.com/art/225272377.html.

(Etats-Unis, Japon) entraînerait une croissance de la production industrielle, qui stimulerait le marché de l'emploi, en particulier intérimaire, dès le début de l'année. L'emploi industriel diminuerait de façon plus lente. Cette hausse de l'emploi, ainsi que le départ en retraite programmé des personnes ayant commencé à travailler jeunes, contribuerait à une stabilisation du taux de chômage. Le moral des Français devrait en être amélioré, et conduire à une diminution de leur taux d'épargne malgré la baisse de leur pouvoir d'achat depuis deux ans. La reprise de la consommation des ménages constitue le socle de la croissance, à côté de la "vigueur retrouvée" des entreprises.

Pour la seule Ile-de-France², les chefs d'entreprise franciliens demeurent prudents en termes de prévisions économiques pour 2004, particulièrement dans les domaines de l'emploi et de l'investissement. Alors que règne un certain optimisme dans le secteur tertiaire, le pessimisme est important dans le secteur industriel où les délocalisations se multiplient. De nombreux aléas peuvent peser sur la reprise éventuelle :

- Alors que la reprise au niveau national devrait reposer sur une croissance du commerce extérieur, les chefs d'entreprises industrielles prévoient une nouvelle baisse de leurs ventes. Le niveau de l'euro face au dollar jouera un rôle prépondérant dans l'accroissement attendu des exportations.
- La reprise en Ile-de-France devrait plutôt s'appuyer sur un redémarrage de la consommation, après une année 2003 décevante. Toutefois, le moral des ménages demeure à un niveau bas, entravé par la hausse du chômage et l'inflation perçue.
- Les nouvelles embauches devraient être peu nombreuses en 2004 puisque 90% des entreprises prévoient une stagnation ou une baisse de leurs effectifs. Seules quelques créations devraient être enregistrées dans le secteur tertiaire, après une forte dégradation de l'emploi salarié au cours de l'année 2003. Selon une étude de Natexis Banques Populaires, une

²Source : Enquête Eurochambres du CCIP - Crocis de janvier 2004. Cette enquête est réalisée dans 26 pays européens, soit plus de 100 000 entreprises interrogées.

Les résultats concernant l'Ile-de-France sont obtenus grâce à une enquête téléphonique auprès de 200 PME et 101 grands établissements de plus de 500 salariés. 26% des établissements relèvent du secteur industriel, 74% du commerce ou des services. Les résultats ont été pondérés par taille et secteur.

croissance de 2% au minimum serait nécessaire pour faire baisser de façon durable le taux de chômage en Ile-de-France.

- En termes d'investissements, les conditions financières sont favorables (taux d'intérêt stabilisé à un bas niveau) mais l'attentisme demeure. La croissance des investissements serait pourtant le moteur d'une véritable croissance, et non d'une "reprise de rattrapage" soutenue exclusivement par la demande extérieure et la consommation des ménages.

12.2.4 Un scénario d'évolution des données plus plausible

L'année 2003 est une année particulière (grèves dans les transports et la fonction publique, canicule). Il est plausible de supposer que l'année 2004 sera une année de rattrapage, d'autant plus que la reprise économique est attendue.

Pour quantifier ce scénario, nous avons choisi de supposer que les variables de trafic et les données décrivant l'environnement en Ile-de-France auront en 2004 et 2005 un taux de croissance identique à celui observé en 1997, date de la reprise économique précédente.

Par rapport au scénario implicite, ce scénario d'évolution des variables décrit un environnement dans lequel les données de trafic par type de journée et de sous-réseau ont un taux d'accroissement plus important. Seul le RER a une évolution moins marquée (+0.40% par an contre +1.06% dans le scénario implicite) : la reprise n'engendre pas immédiatement une diminution du nombre de demandeurs d'emplois.

En termes d'environnement économique en Ile-de-France, ce scénario suppose une diminution moins importante du volume de circulation de Paris, tandis que la vitesse de circulation devrait décroître plus fortement. Ce scénario est plausible : s'il existe une volonté de diminuer le volume de circulation dans Paris d'ici 2005, la circulation a déjà fortement baissé les années précédentes (baisse supérieure à -4.5% entre 2003 et 2002), alors que la vitesse de circulation a décliné à un rythme moins important malgré l'augmentation des contrôles. La consommation et le tourisme suivent un schéma plus optimiste que dans le scénario implicite : le volume des ventes dans les grands magasins augmente de façon plus importante (+2.02 points chaque année contre +0.91 point pour le scénario implicite), tandis que le tourisme redémarre en Ile-de-France (le taux de croissance du taux d'occupation des hôtels franciliens est positif alors qu'il stagnait dans

le scénario implicite). L'effet de rattrapage par rapport à une année marquée par les tensions géopolitiques et la canicule est pris en compte. Enfin, le rapport entre le prix du ticket T et le prix de l'essence décroît moins fortement : le processus de hausse du prix de l'essence du fait des tensions internationales et de la montée du cours de l'euro devrait prendre fin.

EVOLUTION SIMILAIRE A CELLE OBSERVEE EN 1997

Taux d'accroissement moyen estimé	Mensuel	Annuel
Données de trafic		
Jour ouvrable sur le métro	+0.06%	+0.75%
Jour ouvrable sur le RER	+0.03%	+0.40%
Jour ouvrable sur les bus de Paris	+0.32%	+3.96%
Jour ouvrable sur les bus de banlieue	+0.33%	+4.07%
Samedi sur le métro	+0.39%	+4.72%
Dimanche/férié sur le métro	+0.39%	+4.83%
Samedi sur le RER	+0.27%	+3.24%
Dimanche/férié sur le RER	+0.33%	+4.02%
Données d'environnement		
Volume de circulation dans Paris	-0.03%	-0.37%
Vitesse de circulation dans Paris	-0.18%	-2.08%
Demandes d'emplois en fin de mois (IdF*)	+0.26%	+3.15%
Volume des ventes dans les grands magasins (IdF*)	+0.17%	+2.06%
Taux d'occupation des hôtels en IdF*	+0.33%	+4.01%
Rapport entre le prix du ticket T et de l'essence	-0.02%	-0.23%

* IdF : Ile-de-France

Le calcul d'une prévision ponctuelle ne permet pas de prendre en compte le risque associé à cette prévision. Dans la section suivante, nous montrerons une façon, novatrice, de construire des intervalles de confiance, à l'aide de simulations de Monte Carlo "calibrées".

12.3 Les intervalles de prévision

L'obtention de prévisions ponctuelles est un préalable indispensable mais non suffisant à tout travail de prévision. Aucune information n'est fournie sur la robustesse des conclusions tirées d'une seule réalisation (la prévision ponctuelle), i.e. sur le risque associé à ces prévisions ponctuelles. Pour évaluer ce risque, d'autant plus important que l'horizon de prévision est éloigné, des intervalles de confiance des prévisions doivent être construits.

La méthode de calcul des intervalles de confiance asymptotique repose sur l'hypothèse de normalité des résidus, et se place dans un cadre asymptotique, i.e. à long terme. Une autre façon de procéder réside dans la prise en compte d'intervalles de confiance obtenus à partir de simulations de Monte Carlo. Cette méthode est novatrice dans la mesure où elle suppose des perturbations futures "calibrées", i.e. possédant une moyenne et une variance similaires à celles observées dans le passé.

Cette section présente les étapes du calcul des intervalles simulés pour l'ensemble des modélisations (auto-régressive, de long terme et de court terme).

12.3.1 Modèle auto-régressif avec racine unitaire

D'après les résultats des tests de racine unitaire et de nos tests de modification du degré d'intégration, chacune des séries est intégrée à l'ordre 1, sur l'ensemble de la période ou sur une sous-période uniquement. Les modélisations auto-régressives ont été effectuées sur les variables en différence première, sur tout ou partie de la période.

Les perturbations de chacune des variables sont supposées indépendantes les unes des autres, ce qui permet de les estimer et de les prévoir une par une.

Première étape : Estimation des paramètres auto-régressifs Chaque série i s'écrit sous la forme

$$\Delta y_{i,t} = c_i + a_{i,1}\Delta y_{i,t-1} + a_{i,2}\Delta y_{i,t-2} + \dots + a_{i,p_i}\Delta y_{i,t-p_i} + \varepsilon_{i,t} \quad (12.10)$$

avec $y_{i,t} = \ln(x_{i,t})$

L'estimation par les moindres carrés ordinaires des paramètres de cette dynamique auto-régressive permet d'obtenir, pour chaque variable explicative i , des estimations convergentes des paramètres $\hat{c}_i, \hat{a}_{i,1}, \dots, \hat{a}_{i,p_i}$ où le nombre de retards p_i est choisi à l'aide d'un critère BIC.

Ces estimations sont détaillées au début du chapitre pour les données de trafic comme pour les données décrivant l'environnement économique en Ile-de-France.

Deuxième étape : Génération d'une variable auto-régressive à l'horizon h L'estimation de l'équation (12.10) permet de calculer la moyenne et la variance empiriques des résidus estimés $\hat{\varepsilon}_i$, notées respectivement $\bar{\varepsilon}_i$ et $\hat{\sigma}_{\hat{\varepsilon}_i}^2$. Sous l'hypothèse de normalité de chacun de ces résidus, une trajectoire des résidus futurs $\check{\varepsilon}_{i,t}^{(m)}$ avec $t > T$ (date de fin d'échantillon) peut être simulée par une loi normale de moyenne $\bar{\varepsilon}_i$ et de variance $\hat{\sigma}_{\hat{\varepsilon}_i}^2$. La trajectoire future des perturbations est "calibrée".

Pour cette trajectoire m et cette variable i , la meilleure prévision linéaire en différence à l'horizon h , notée $\Delta \hat{y}_{i,t}^{(m)}$, est obtenue à l'aide des paramètres $\hat{c}_i, \hat{a}_{i,1}, \dots, \hat{a}_{i,p_i}$ estimés à l'étape précédente et des résidus futurs simulés calibrés $(\check{\varepsilon}_{i,T+1}^{(m)}, \dots, \check{\varepsilon}_{i,T+h}^{(m)})$:

$$\Delta \hat{y}_{i,t}^{(m)} = \hat{c}_i + \hat{a}_{i,1} \Delta \hat{y}_{i,t-1}^{(m)} + \hat{a}_{i,2} \Delta \hat{y}_{i,t-2}^{(m)} + \dots + \hat{a}_{i,p_i} \Delta \hat{y}_{i,t-p_i}^{(m)} + \check{\varepsilon}_{i,t}^{(m)}$$

Pour obtenir les variables explicatives en log-niveau à l'horizon h , la variable en niveau retardée est ajoutée à la variable retardée :

$$\hat{y}_{i,t+h}^{(m)} = \Delta \hat{y}_{i,t+h}^{(m)} + \hat{y}_{i,t+h-1}^{(m)}$$

Un nombre $nmont = 10\,000$ de simulations est réalisé, de façon à obtenir $nmont$ trajectoires possibles de la variable y_{it} . A chaque horizon h , pour la variable i , on dispose de $nmont$ valeurs simulées par le modèle auto-régressif : $y_{i,t+h}^{(1)}, \dots, y_{i,t+h}^{(m)}, \dots, y_{i,t+h}^{(nmont)}$.

Troisième étape : Construction des intervalles de prévision pour une variable auto-régressive La méthode des "percentiles" est utilisée pour construire les intervalles de prévision : l'intervalle de confiance à $(100 - \alpha)\%$ à un horizon h est donné par la limite inférieure $\hat{\theta}_{INF}$, telle que $\frac{\alpha}{2}\%$ des points de la distribution générée à l'horizon h sont en dessous de cette

limite, et la limite supérieure $\hat{\theta}_{SUP}$ telle que $\frac{\alpha}{2}\%$ des points sont au dessus de cette borne.

12.3.2 Modèle de long terme

Quatrième étape : Estimation des paramètres de la relation de long terme Le processus cointégré (sur période entière ou sous-période uniquement) à la date t s'écrit, pour chaque type de trafic j :

$$y_{j,LT,t} = c_{j,LT} + \delta_{j,LT} \cdot t + w_{j,t} \cdot \beta_{j,LT} + \varepsilon_{j,LT,t}$$

L'application de la méthode des moindres carrés "totalement modifiés" de Phillips et Hansen [140] à la relation cointégrée permet d'obtenir pour chaque type de trafic j les estimations convergentes des paramètres $\check{c}_{j,LT}$, $\check{\delta}_{j,LT}$, et $\check{\beta}_{j,LT}$. Ces paramètres estimés sont présentés en début de chapitre.

En outre, cette estimation permet le calcul des résidus de long terme estimés $\check{\varepsilon}_{j,LT,t}$.

Cinquième étape : Génération d'une variable par la modélisation de long terme à l'horizon h Soient $\bar{\varepsilon}_{j,LT}$ la moyenne empirique des résidus de long terme estimée et $\hat{\sigma}_{\check{\varepsilon}_{j,LT}}^2$ leur variance empirique. Sous l'hypothèse de normalité des résidus de la relation de long terme, une trajectoire m des résidus futurs de long terme, notée $\check{\varepsilon}_{j,LT,t}^{(m)}$, peut être simulée à l'aide d'une loi $\mathcal{N}\left(\bar{\varepsilon}_{j,LT}, \hat{\sigma}_{\check{\varepsilon}_{j,LT}}^2\right)$.

La meilleure prévision de long terme de la variable endogène à l'horizon h , pour une trajectoire m , est obtenue par l'équation suivante :

$$\hat{y}_{j,LT,t+h}^{(m)} = \hat{c}_{j,LT} + \hat{\delta}_{j,LT} (t+h) + \hat{w}_{j,t+h}^{(m)} \cdot \hat{\beta}_{j,LT} + \check{\varepsilon}_{j,LT,t}^{(m)}$$

dans laquelle la matrice $\hat{w}_{j,t+h}^{(m)}$ est constituée des explicatives simulées à la troisième étape.

A chaque horizon h , pour la variable j , on dispose de $nmont$ valeurs simulées par la relation de long terme : $y_{j,LT,t+h}^{(1)}, \dots, y_{j,LT,t+h}^{(m)}, \dots, y_{j,LT,t+h}^{(nmont)}$.

Sixième étape : Construction des intervalles de prévision par la relation de long terme La méthode des "percentiles" est appliquée, à chaque horizon h et à chaque variable j ,

à l'ensemble des simulations $y_{j,LT,t+h}^{(1)}, \dots, y_{j,LT,t+h}^{(m)}, \dots, y_{j,LT,t+h}^{(nmont)}$.

12.3.3 Modèle de court terme (à correction d'erreur)

Septième étape : Estimation des paramètres de la relation de court terme La relation de court terme s'écrit, pour chaque trafic j , sous la forme :

$$\begin{aligned} \Delta y_{j,CT,t} &= c_{j,CT} + \Delta w_{j,t} \cdot \beta_{j,CT} + \alpha_j z_{j,t-1} + \varepsilon_{j,CT,t} \\ \text{avec } z_{j,t-1} &= \varepsilon_{j,LT,t-1} \end{aligned}$$

Les estimations des paramètres pour chaque trafic j , notées $\hat{c}_{j,CT}$, $\hat{\beta}_{j,CT}$, et $\hat{\alpha}_j$, ainsi que des résidus $\hat{\varepsilon}_{j,CT,t}$, ont été effectués par la méthode des moindres carrés "totalement modifiés" et sont présentés en début de chapitre.

Huitième étape : Génération d'une variable par la modélisation de court terme à l'horizon h Notons $\bar{\varepsilon}_{j,CT}$ la moyenne empirique des résidus de long terme estimée et $\hat{\sigma}_{\hat{\varepsilon}_{j,CT}}^2$ leur variance empirique. Sous l'hypothèse de normalité des résidus de la relation de long terme, une trajectoire m des résidus futurs de long terme, notée $\check{\varepsilon}_{j,CT,t}^{(m)}$, peut être simulée à l'aide d'une loi $\mathcal{N}(\bar{\varepsilon}_{j,CT}, \hat{\sigma}_{\hat{\varepsilon}_{j,CT}}^2)$.

Pour une trajectoire m des explicatives de la matrice $\Delta \hat{w}_{j,t+h}^{(m)}$ et des résidus de long terme, la meilleure prévision de court terme s'écrit sous la forme :

$$y_{j,CT,t+h}^{(m)} = y_{j,CT,t+h-1}^{(m)} + \hat{c}_{j,CT} + \Delta \hat{w}_{j,t+h}^{(m)} \hat{\beta}_{j,CT} + \check{\varepsilon}_{j,CT,t+h}^{(m)}$$

La répétition de ces simulations permet d'obtenir, à chaque horizon h et pour chaque trafic j , les simulations de court terme $y_{j,CT,t+h}^{(1)}, \dots, y_{j,CT,t+h}^{(m)}, \dots, y_{j,CT,t+h}^{(nmont)}$.

Neuvième et dernière étape : Construction des intervalles de prévision pour la relation de court terme La méthode des "percentiles" est appliquée, à chaque horizon h et à chaque variable j , à l'ensemble des simulations $y_{j,CT,t+h}^{(1)}, \dots, y_{j,CT,t+h}^{(m)}, \dots, y_{j,CT,t+h}^{(nmont)}$.

12.4 Prévisions des trafics RATP par type de journée et de sous-réseau

Pour chaque type de journée, chaque sous-réseau, et chaque type de modèle, les formules de prévision ont été appliquées en se plaçant à la fin de l'année 2003. La première sous-section synthétise les résultats sous forme de graphiques et de tableaux. Ces prévisions, ainsi que l'utilisation des intervalles de confiance simulés, seront analysés dans la deuxième sous-section. Enfin, les prévisions de trafic par sous-réseau seront présentées pour l'ensemble du mois et non plus seulement pour un type de journée.

12.4.1 Résultats des prévisions pour chaque type de trafic journalier

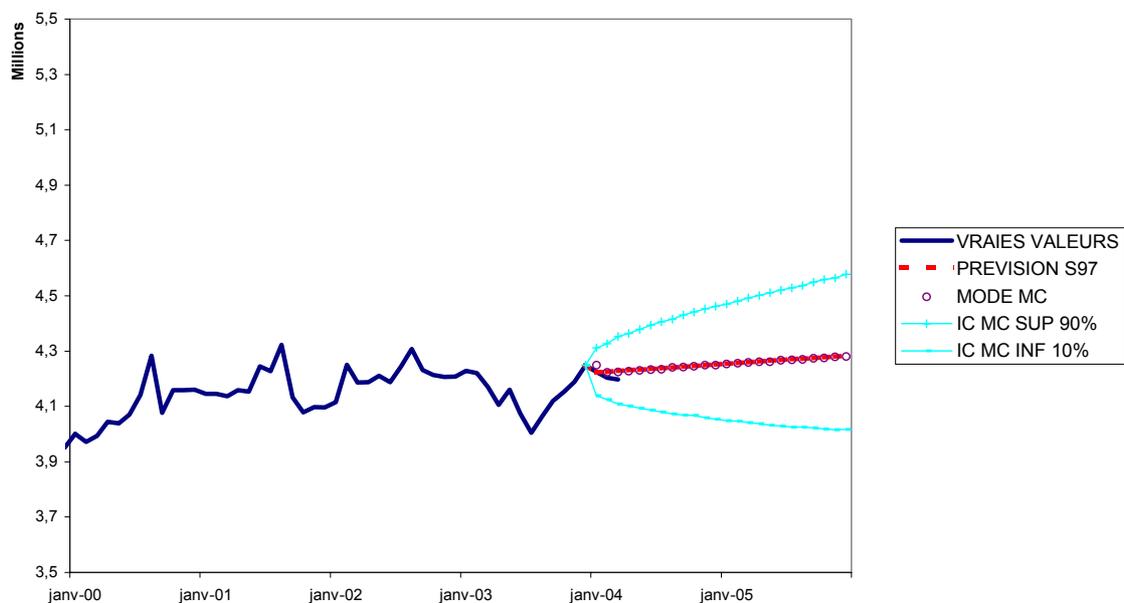
Le travail a été réalisé sur les données log-linéarisées. Pour faciliter la lecture des graphiques et tableaux, les résultats ont été transformés par la fonction exponentielle. Sur chaque graphique sont reportés, pour l'ensemble des horizons h futurs :

- la prévision ponctuelle, appelée "PREVISION S97", calculée en supposant que les variables économiques auront une évolution similaire à celle observée en 1997,
- le mode des 10 000 simulations de Monte Carlo ("MODE MC") qui doit correspondre à la prévision ponctuelle si les simulations ont été correctement calibrées et si les bruits sont bien symétriques conformément à la loi normale,
- ainsi que les bornes supérieures et inférieures de l'intervalle de confiance à 80% obtenu grâce à 10 000 simulations de Monte Carlo : "IC MC SUP 90%" et "IC MC INF 10%" telles qu'il existe une chance sur 10 que la valeur future soit au dessus de IC MC SUP 90%, et une chance sur 10 qu'elle soit au dessous de IC MC INF 10%.

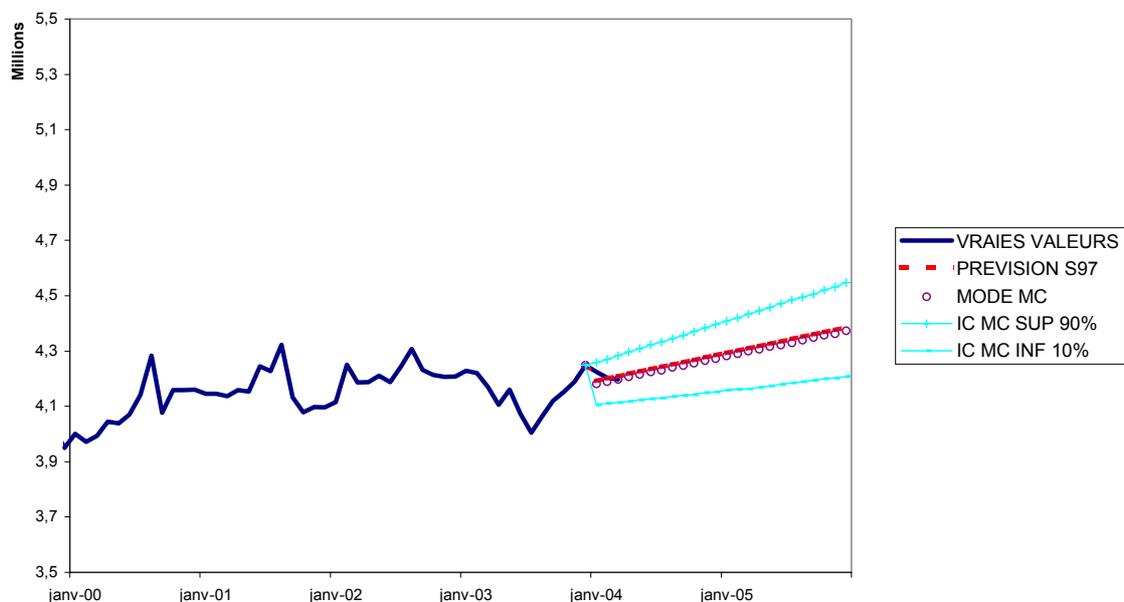
En outre, les données disponibles pour le début de l'année 2004 (non prises en compte dans les calculs de prévision) ont été reportées sur les graphiques *a posteriori*, sous le nom de "VRAIES VALEURS". Un tableau indique ensuite les principales caractéristiques des prévisions.

12.4.1.1 Trafic du jour ouvrable sur le métro

Modèle auto-régressif et simulations de Monte-Carlo (MC)

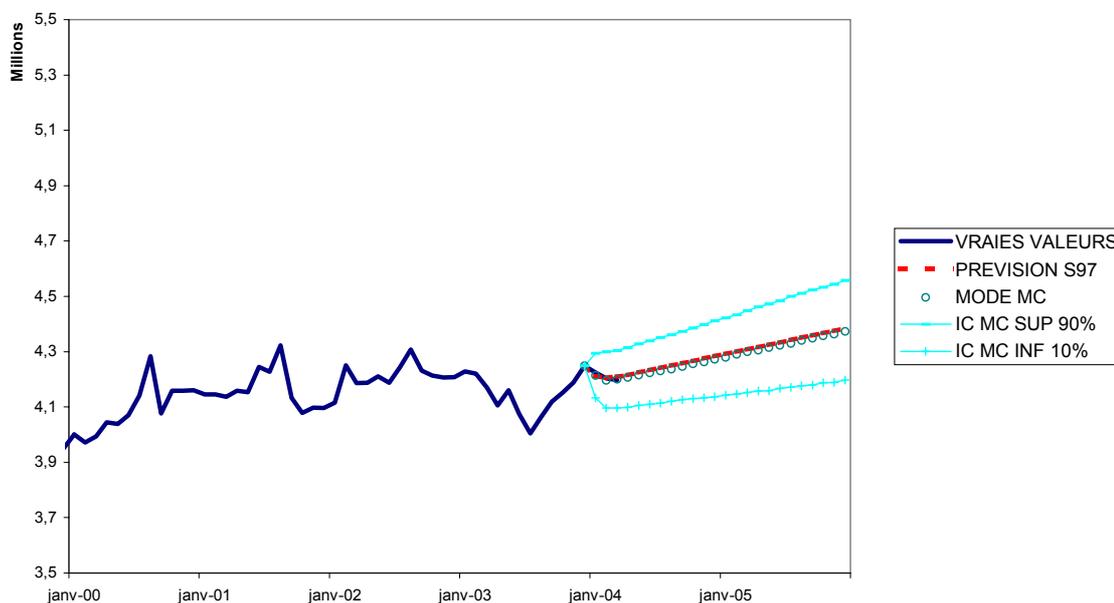


Modèle de long terme et simulations de Monte-Carlo (MC)



12.4. PRÉVISIONS DES TRAFICS RATP PAR TYPE DE JOURNÉE ET DE SOUS-RÉSEAU

Modèle de court terme et simulations de Monte-Carlo (MC)



Intervalles de confiance à 80% suivant leur mode de calcul et l'horizon considéré (en millions)

Trafic du Jour Ouvrable sur le métro	Auto-Régressif	Long Terme	Court Terme
6 mois	4.09 - 4.39	4.13 - 4.32	4.11 - 4.34
12 mois	4.05 - 4.46	4.15 - 4.40	4.14 - 4.41
18 mois	4.03 - 4.52	4.18 - 4.47	4.17 - 4.48
24 mois	4.02 - 4.58	4.21 - 4.55	4.20 - 4.56

Etendues de chaque intervalle à 80% (en millions)

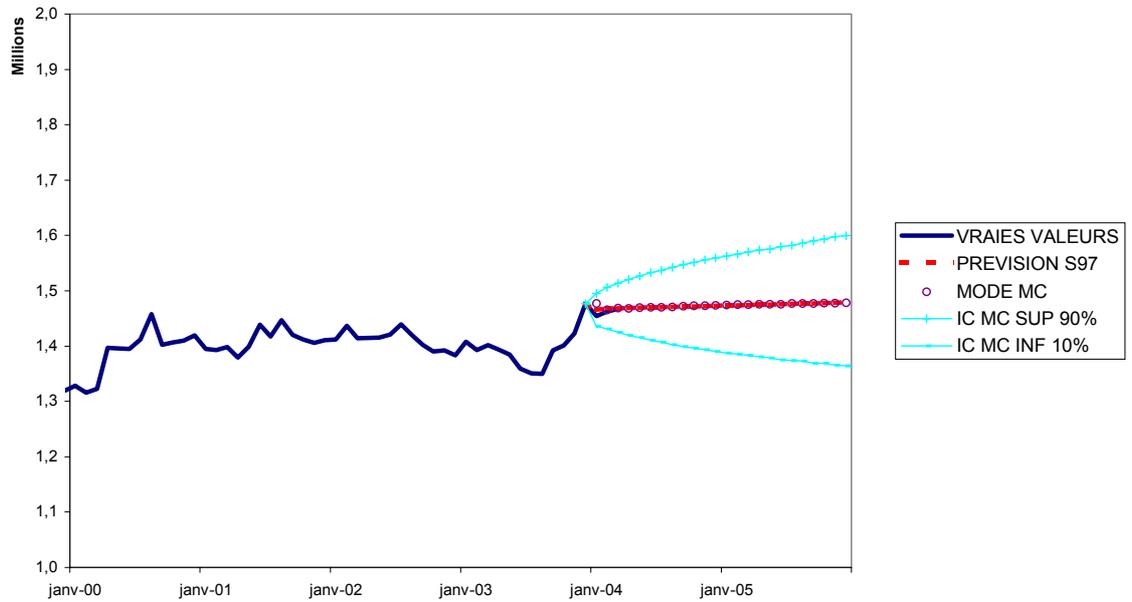
6 mois	0.31	0.20	0.23
12 mois	0.41	0.24	0.27
18 mois	0.49	0.29	0.32
24 mois	0.56	0.34	0.36

Prévision moyenne de ce type de trafic journalier sur l'année (en millions)

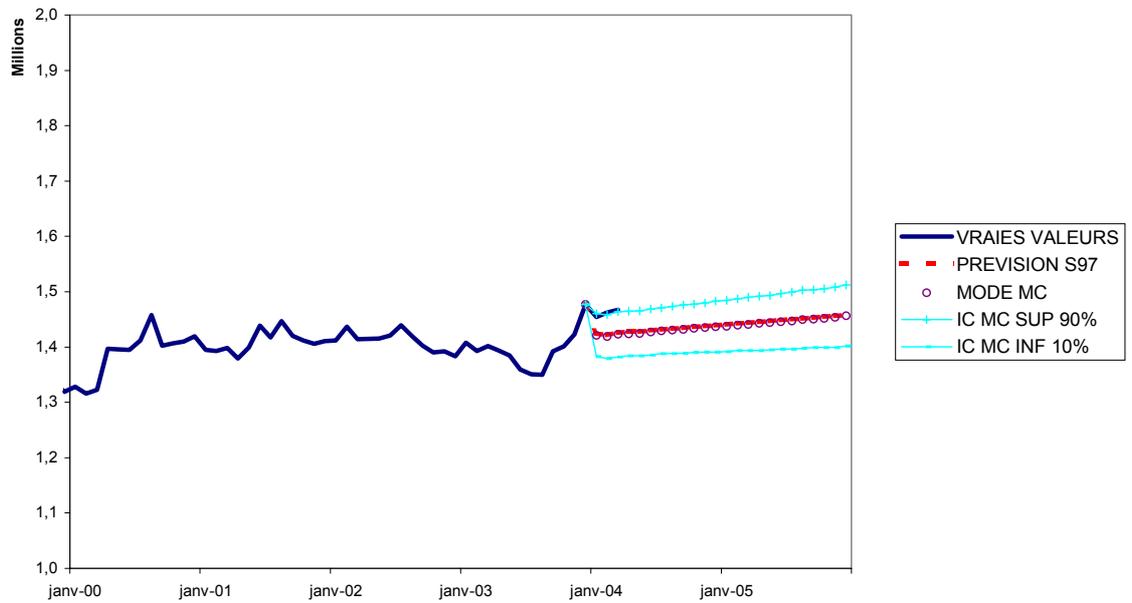
2004	4.24	4.23	4.23
2005	4.27	4.33	4.33

12.4.1.2 Trafic du jour ouvrable sur le RER

Modèle auto-régressif et simulations de Monte-Carlo (MC)

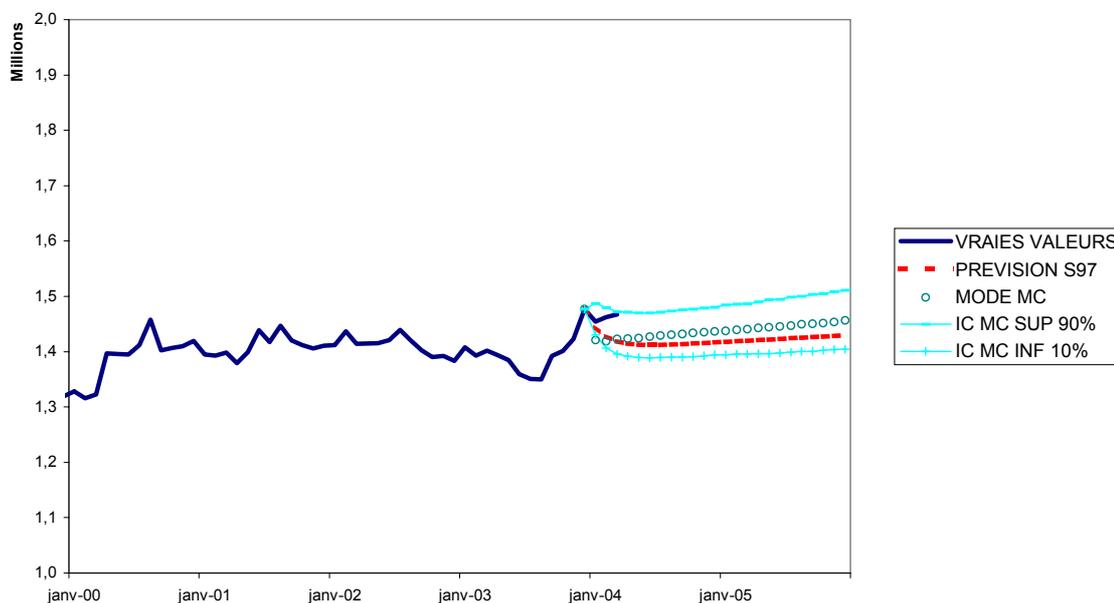


Modèle de long terme et simulations de Monte-Carlo (MC)



12.4. PRÉVISIONS DES TRAFICS RATP PAR TYPE DE JOURNÉE ET DE SOUS-RÉSEAU

Modèle de court terme et simulations de Monte-Carlo (MC)



Intervalles de confiance à 80% suivant leur mode de calcul et l'horizon considéré (en millions)

Trafic du Jour ouvrable sur le RER	Auto-Régressif	Long Terme	Court Terme
6 mois	1.41 - 1.53	1.39 - 1.47	1.39 - 1.47
12 mois	1.39 - 1.56	1.39 - 1.48	1.39 - 1.48
18 mois	1.38 - 1.58	1.40 - 1.50	1.40 - 1.49
24 mois	1.36 - 1.60	1.40 - 1.51	1.40 - 1.51

Etendues de chaque intervalle à 80% (en millions)

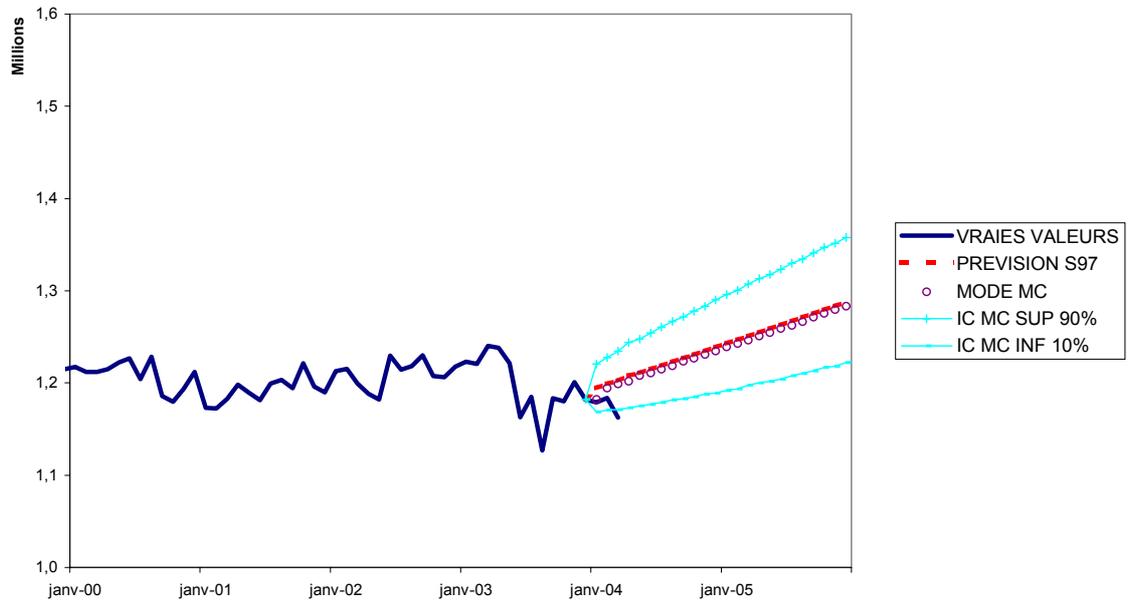
6 mois	0.12	0.08	0.08
12 mois	0.17	0.09	0.09
18 mois	0.20	0.10	0.10
24 mois	0.24	0.11	0.11

Prévision moyenne de ce type de trafic journalier sur l'année (en millions)

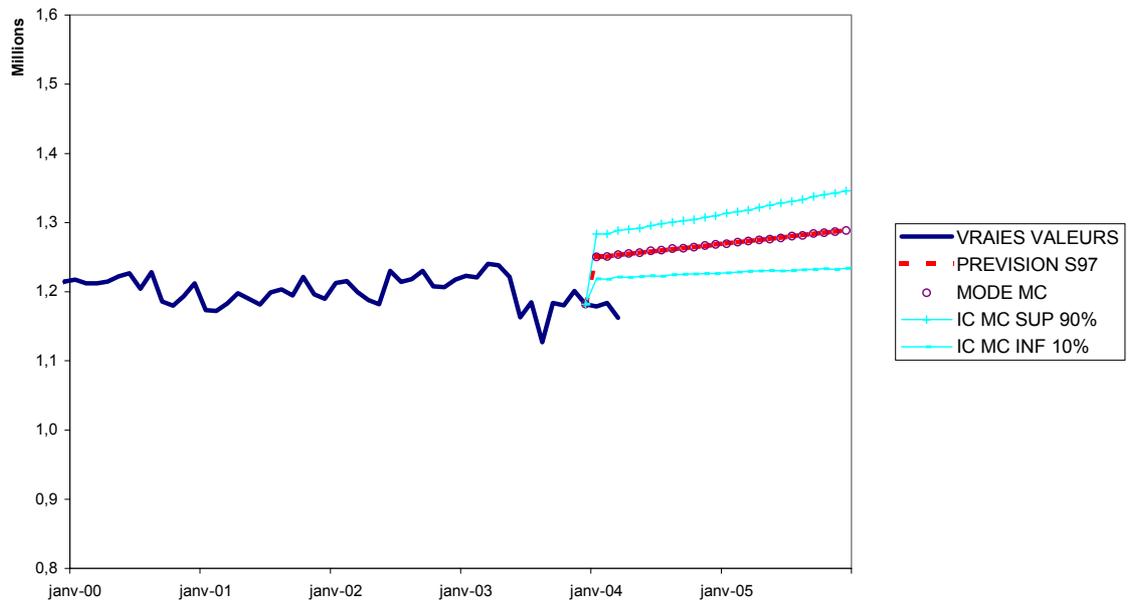
2004	1.47	1.43	1.44
2005	1.48	1.45	1.45

12.4.1.3 Trafic du jour ouvrable sur le réseau des autobus de Paris

Modèle auto-régressif et simulations de Monte-Carlo (MC)

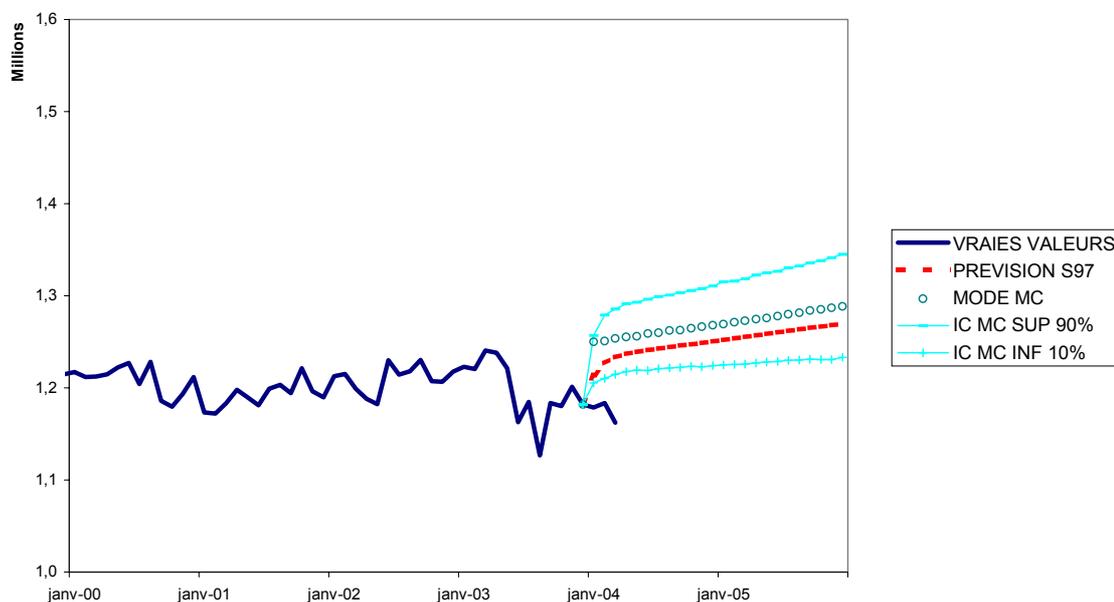


Modèle de long terme et simulations de Monte-Carlo (MC)



12.4. PRÉVISIONS DES TRAFICS RATP PAR TYPE DE JOURNÉE ET DE SOUS-RÉSEAU

Modèle de court terme et simulations de Monte-Carlo (MC)



Intervalles de confiance à 80% suivant leur mode de calcul et l'horizon considéré (en millions)

Trafic du jour ouvrable sur les bus de Paris	Auto-Régressif	Long Terme	Court Terme
6 mois	1.18 - 1.25	1.22 - 1.30	1.22 - 1.30
12 mois	1.19 - 1.29	1.23 - 1.31	1.23 - 1.31
18 mois	1.20 - 1.32	1.23 - 1.33	1.23 - 1.33
24 mois	1.22 - 1.36	1.23 - 1.35	1.23 - 1.34

Etendues de chaque intervalle à 80% (en millions)

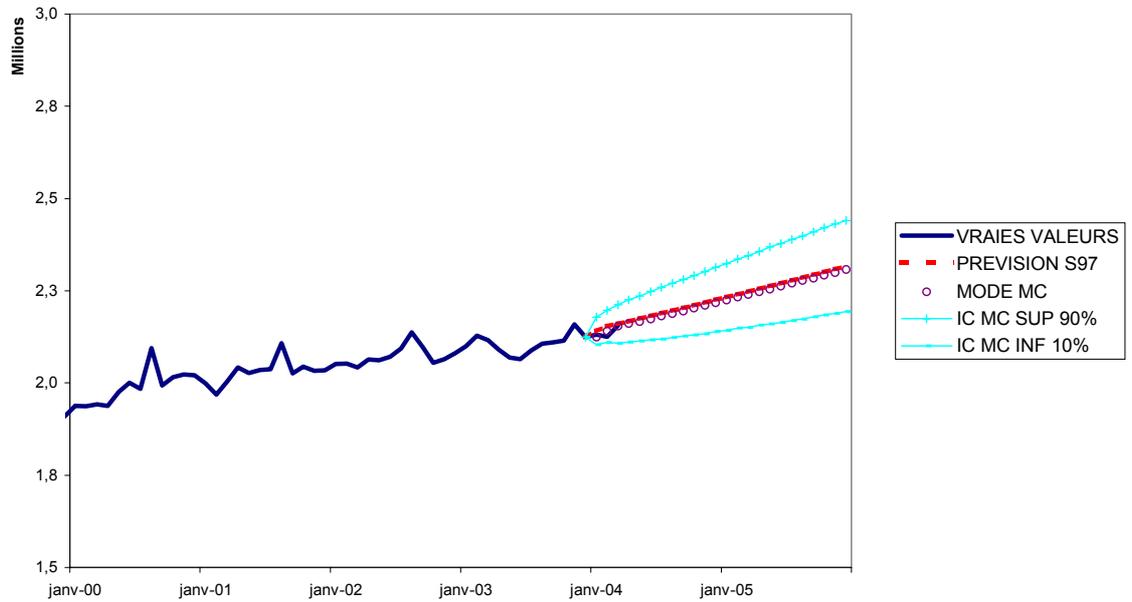
6 mois	0.08	0.07	0.08
12 mois	0.10	0.08	0.09
18 mois	0.12	0.10	0.10
24 mois	0.14	0.11	0.11

Prévision moyenne de ce type de trafic journalier sur l'année (en millions)

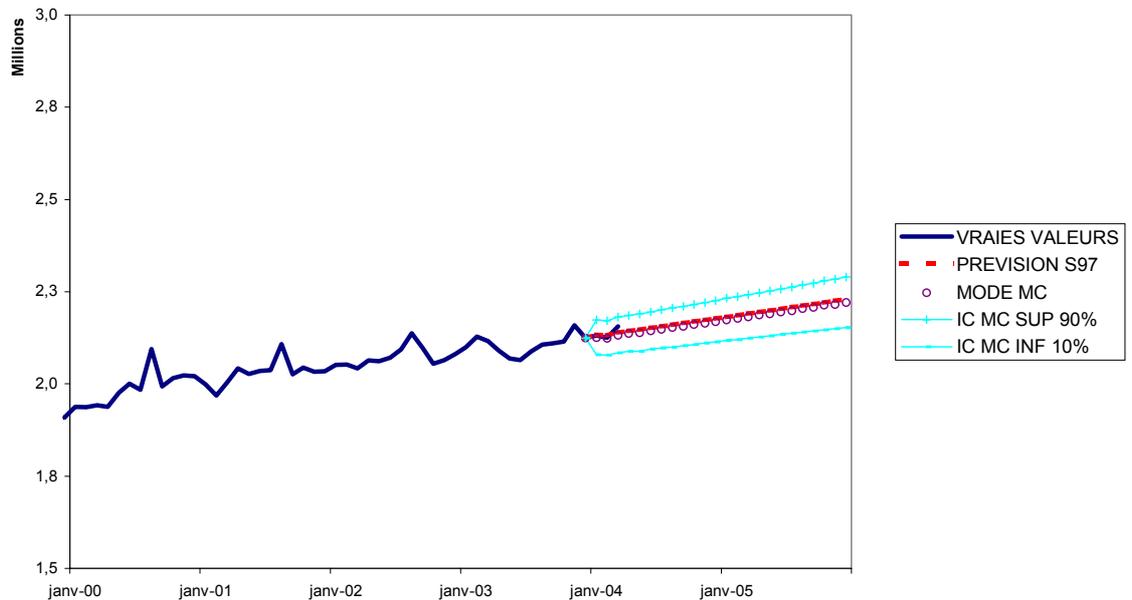
2004	1.22	1.26	1.26
2005	1.27	1.28	1.28

12.4.1.4 Trafic du jour ouvrable sur le réseau des autobus de banlieue

Modèle auto-régressif et simulations de Monte-Carlo (MC)

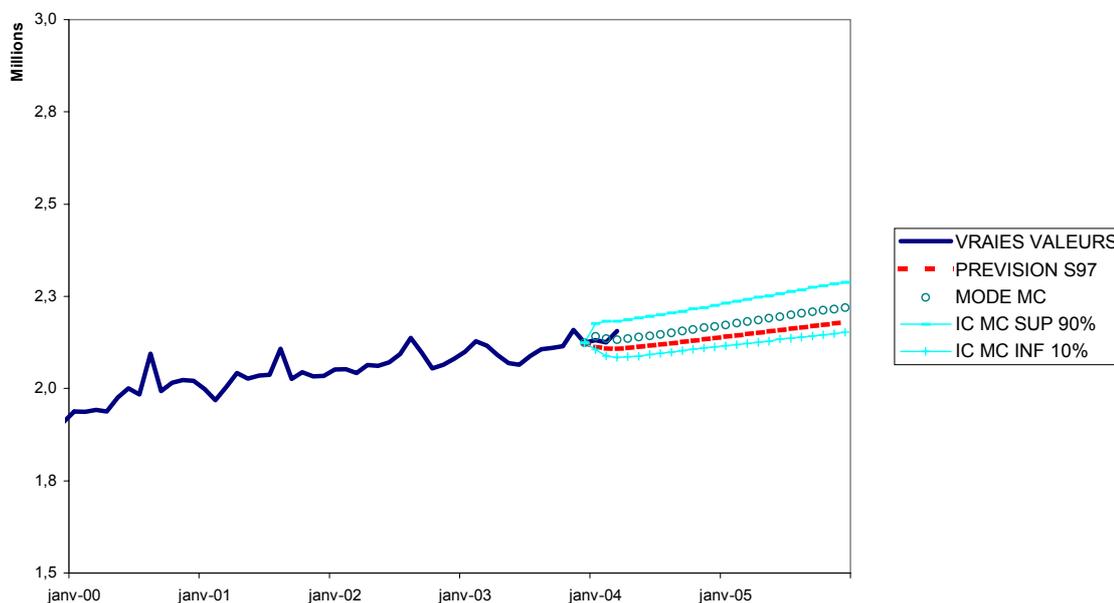


Modèle de long terme et simulations de Monte-Carlo (MC)



12.4. PRÉVISIONS DES TRAFICS RATP PAR TYPE DE JOURNÉE ET DE SOUS-RÉSEAU

Modèle de court terme et simulations de Monte-Carlo (MC)



Intervalles de confiance à 80% suivant leur mode de calcul et l'horizon considéré (en millions)

Trafic du jour ouvrable sur les bus de banlieue	Auto-Régressif	Long Terme	Court Terme
6 mois	2.12 - 2.25	2.09 - 2.19	2.09 - 2.20
12 mois	2.14 - 2.31	2.11 - 2.23	2.06 - 2.31
18 mois	2.16 - 2.38	2.13 - 2.26	2.11 - 2.23
24 mois	2.19 - 2.44	2.15 - 2.29	2.13 - 2.26

Etendues de chaque intervalle à 80% (en millions)

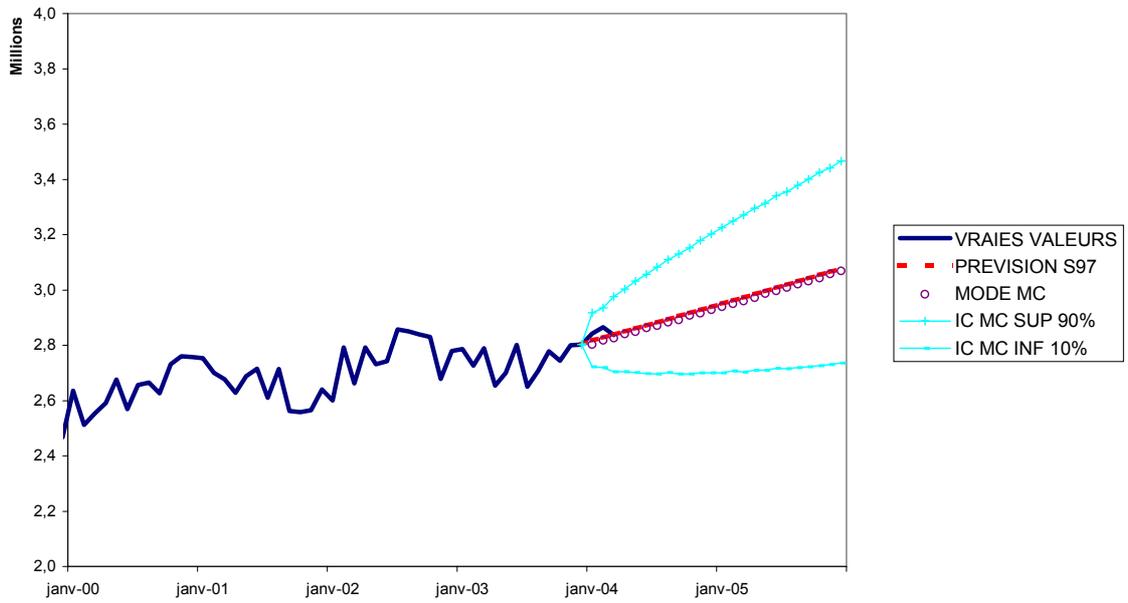
6 mois	0.13	0.10	0.10
12 mois	0.17	0.11	0.11
18 mois	0.21	0.12	0.12
24 mois	0.25	0.14	0.14

Prévision moyenne de ce type de trafic journalier sur l'année (en millions)

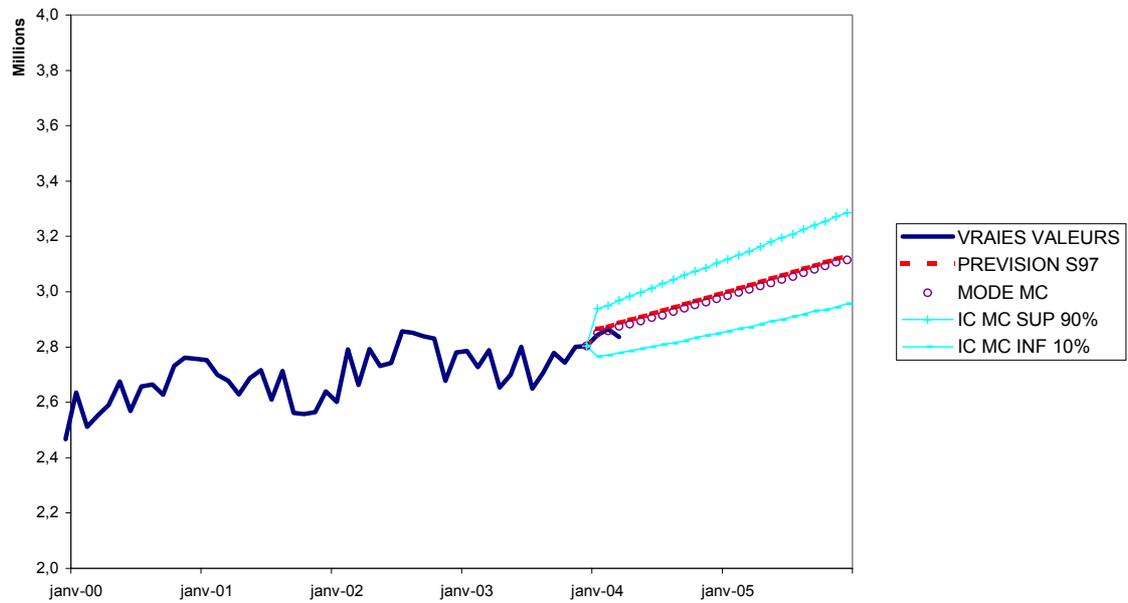
2004	2.18	2.15	2.15
2005	2.27	2.20	2.20

12.4.1.5 Trafic du samedi sur le métro

Modèle auto-régressif et simulations de Monte-Carlo (MC)

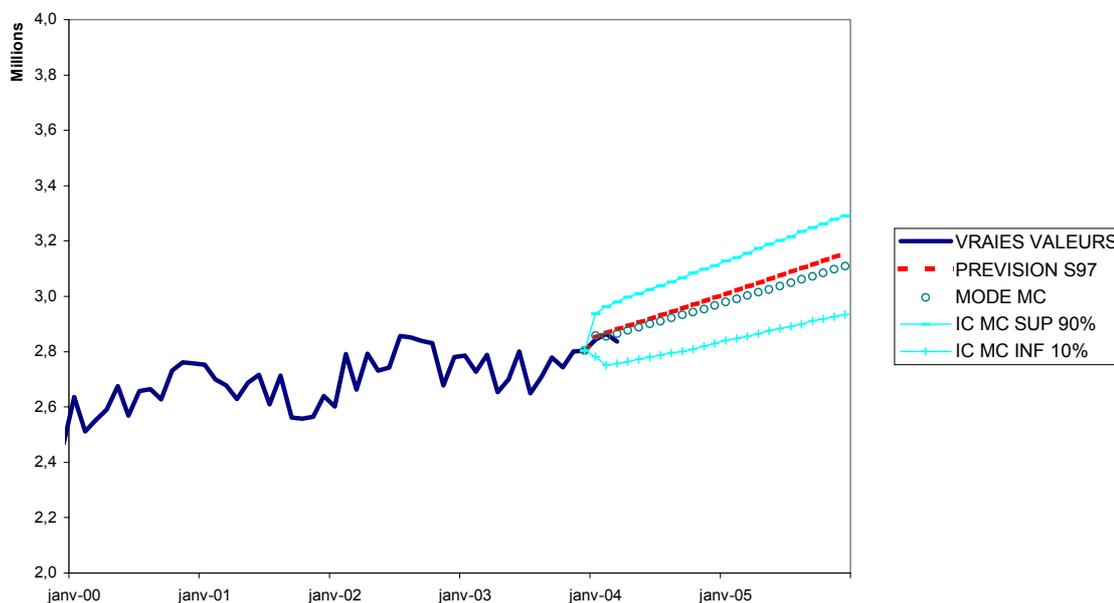


Modèle de long terme et simulations de Monte-Carlo (MC)



12.4. PRÉVISIONS DES TRAFICS RATP PAR TYPE DE JOURNÉE ET DE SOUS-RÉSEAU

Modèle de court terme et simulations de Monte-Carlo (MC)



Intervalles de confiance à 80% suivant leur mode de calcul et l'horizon considéré (en millions)

Trafic du samedi sur le métro	Auto-Régressif	Long Terme	Court Terme
6 mois	2.70 - 3.06	2.80 - 3.01	2.78 - 3.02
12 mois	2.70 - 3.20	2.85 - 3.10	2.83 - 3.11
18 mois	2.72 - 3.34	2.90 - 3.20	2.88 - 3.20
24 mois	2.74 - 3.47	2.96 - 3.28	2.94 - 3.29

Etendues de chaque intervalle à 80% (en millions)

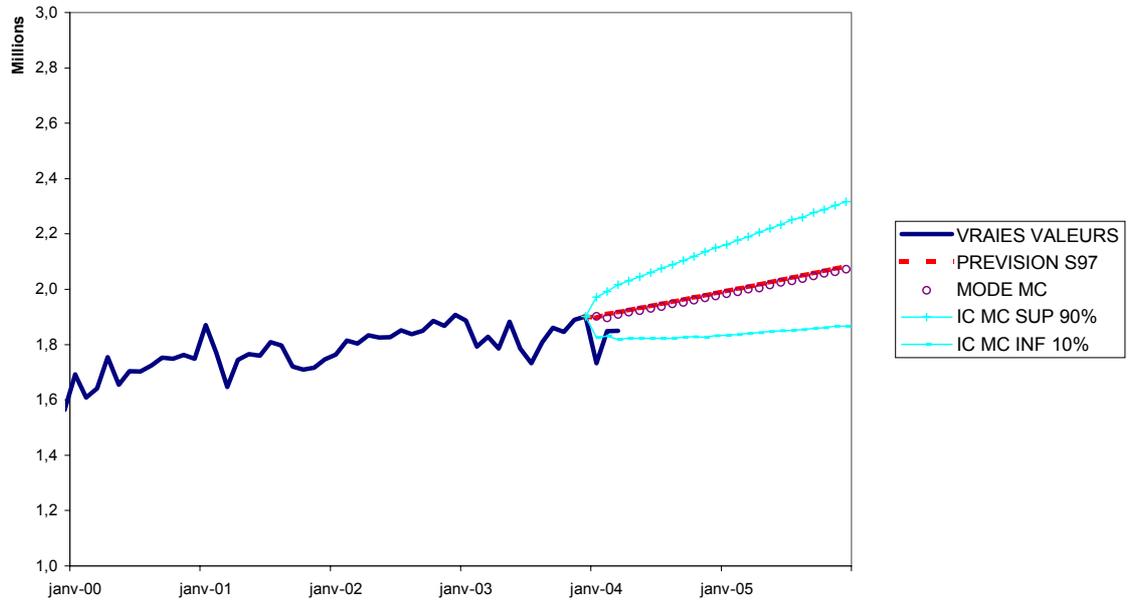
6 mois	0.36	0.21	0.24
12 mois	0.50	0.25	0.28
18 mois	0.62	0.30	0.32
24 mois	0.73	0.33	0.35

Prévision moyenne de ce type de trafic journalier sur l'année (en millions)

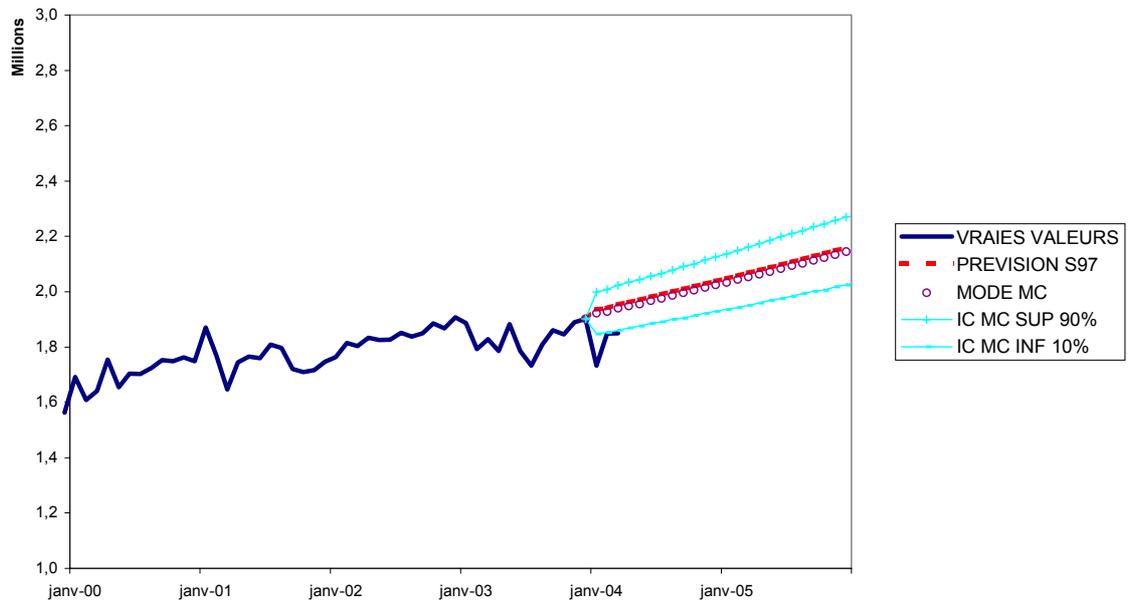
2004	2.88	2.91	2.91
2005	3.02	3.05	3.04

12.4.1.6 Trafic du dimanche/jour férié sur le métro

Modèle auto-régressif et simulations de Monte-Carlo (MC)

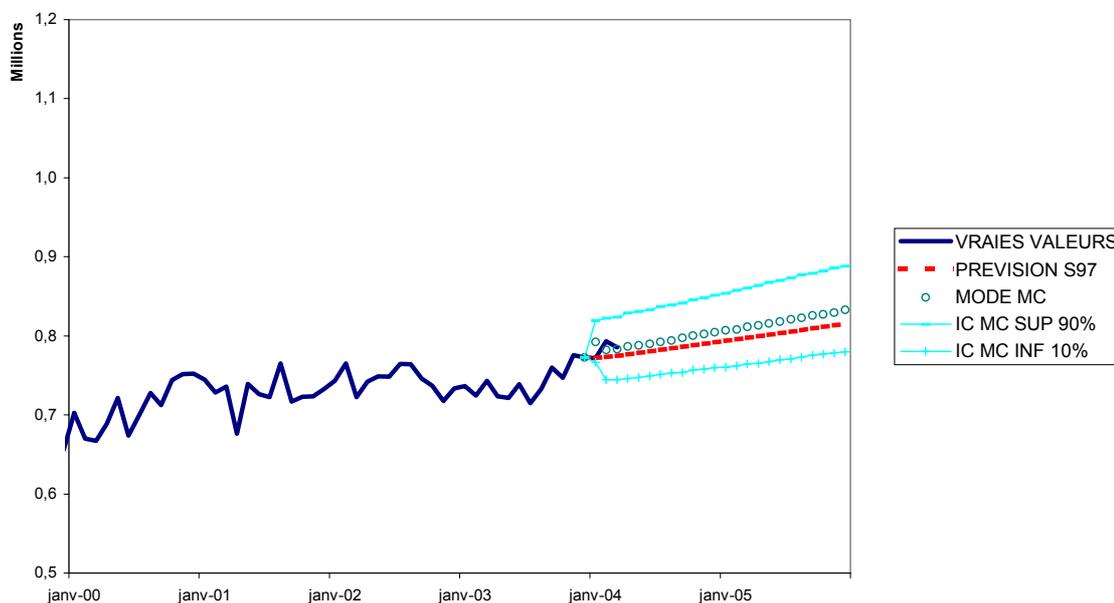


Modèle de long terme et simulations de Monte-Carlo (MC)



12.4. PRÉVISIONS DES TRAFICS RATP PAR TYPE DE JOURNÉE ET DE SOUS-RÉSEAU

Modèle de court terme et simulations de Monte-Carlo (MC)



Intervalles de confiance à 80% suivant leur mode de calcul et l'horizon considéré (en millions)

Trafic du dimanche / jour férié sur le métro	Auto-Régressif	Long Terme	Court Terme
6 mois	1.82 - 2.06	1.89 - 2.06	1.87 - 2.07
12 mois	1.83 - 2.15	1.93 - 2.13	1.91 - 2.14
18 mois	1.85 - 2.23	1.98 - 2.20	1.96 - 2.21
24 mois	1.87 - 2.32	2.02 - 2.27	2.01 - 2.28

Etendues de chaque intervalle à 80% (en millions)

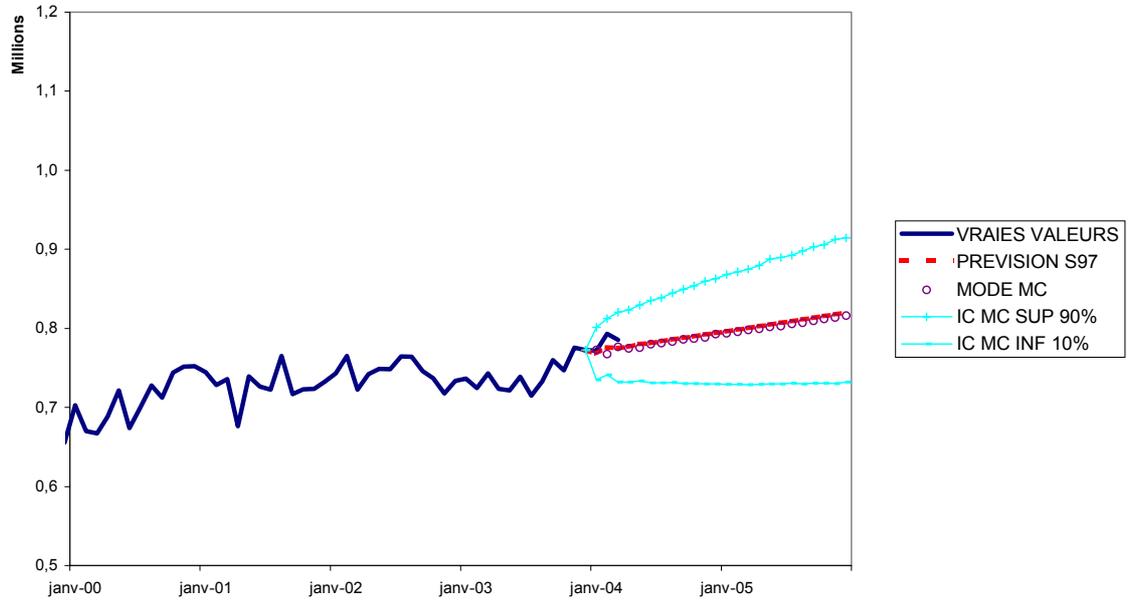
6 mois	0.24	0.17	0.21
12 mois	0.32	0.20	0.23
18 mois	0.38	0.22	0.25
24 mois	0.45	0.24	0.27

Prévision moyenne de ce type de trafic journalier sur l'année (en millions)

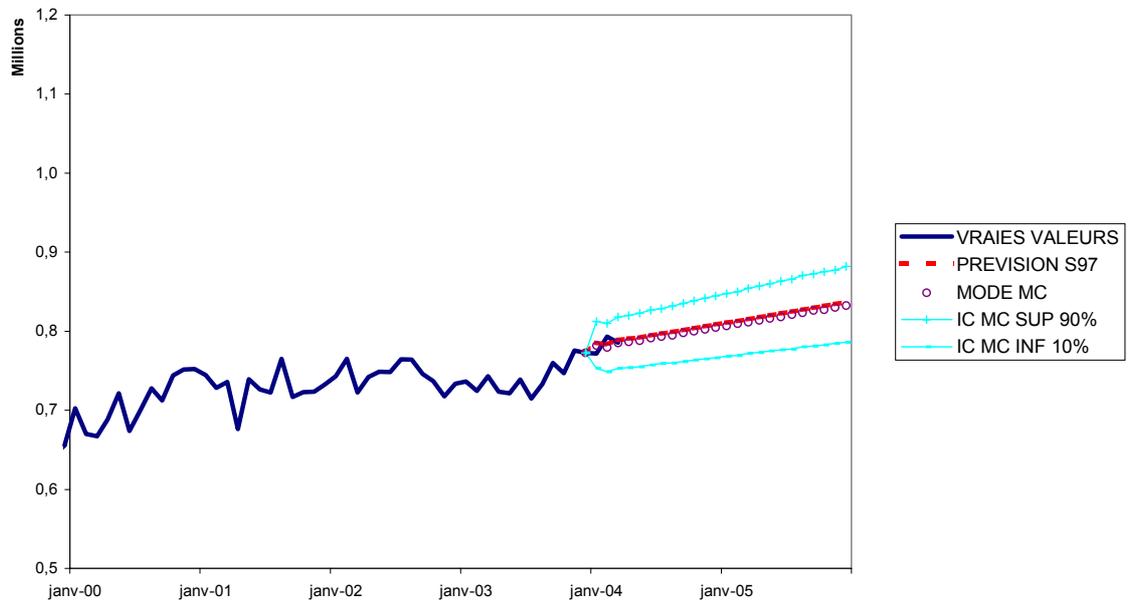
2004	1.94	1.97	1.97
2005	2.03	2.09	2.09

12.4.1.7 Trafic du samedi sur le RER

Modèle auto-régressif et simulations de Monte-Carlo (MC)

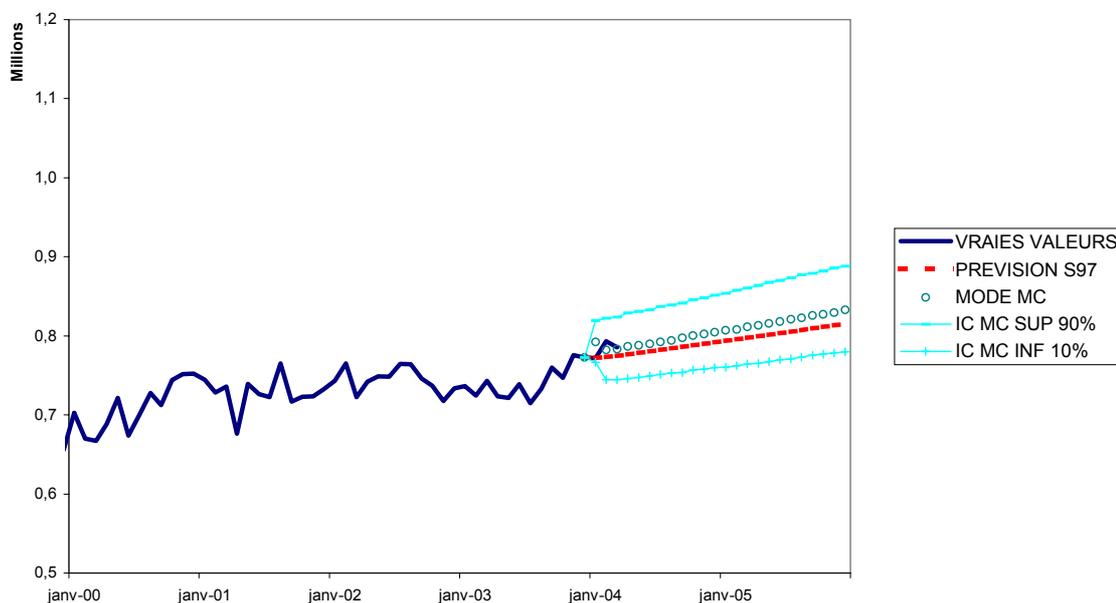


Modèle de long terme et simulations de Monte-Carlo (MC)



12.4. PRÉVISIONS DES TRAFICS RATP PAR TYPE DE JOURNÉE ET DE SOUS-RÉSEAU

Modèle de court terme et simulations de Monte-Carlo (MC)



Intervalles de confiance à 80% suivant leur mode de calcul et l'horizon considéré (en millions)

Trafic du samedi sur le RER	Auto-Régressif	Long Terme	Court Terme
6 mois	0.73 - 0.83	0.76 - 0.83	0.75 - 0.83
12 mois	0.73 - 0.86	0.77 - 0.84	0.76 - 0.85
18 mois	0.73 - 0.89	0.78 - 0.86	0.77 - 0.87
24 mois	0.73 - 0.91	0.79 - 0.88	0.78 - 0.89

Etendues de chaque intervalle à 80% (en millions)

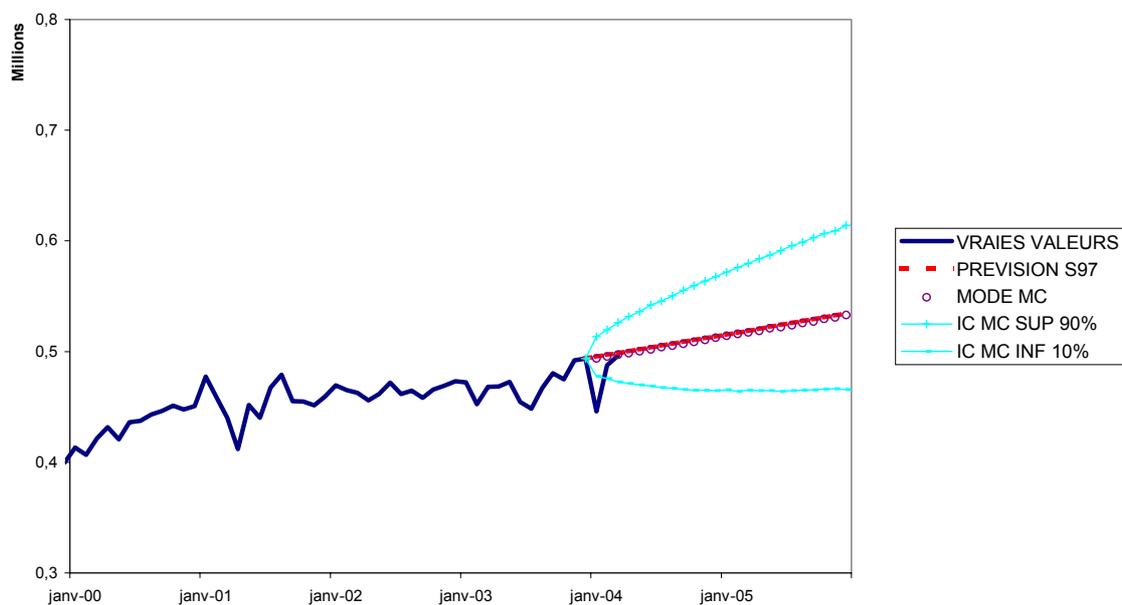
6 mois	0.10	0.07	0.08
12 mois	0.13	0.08	0.09
18 mois	0.16	0.09	0.10
24 mois	0.18	0.10	0.11

Prévision moyenne de ce type de trafic journalier sur l'année (en millions)

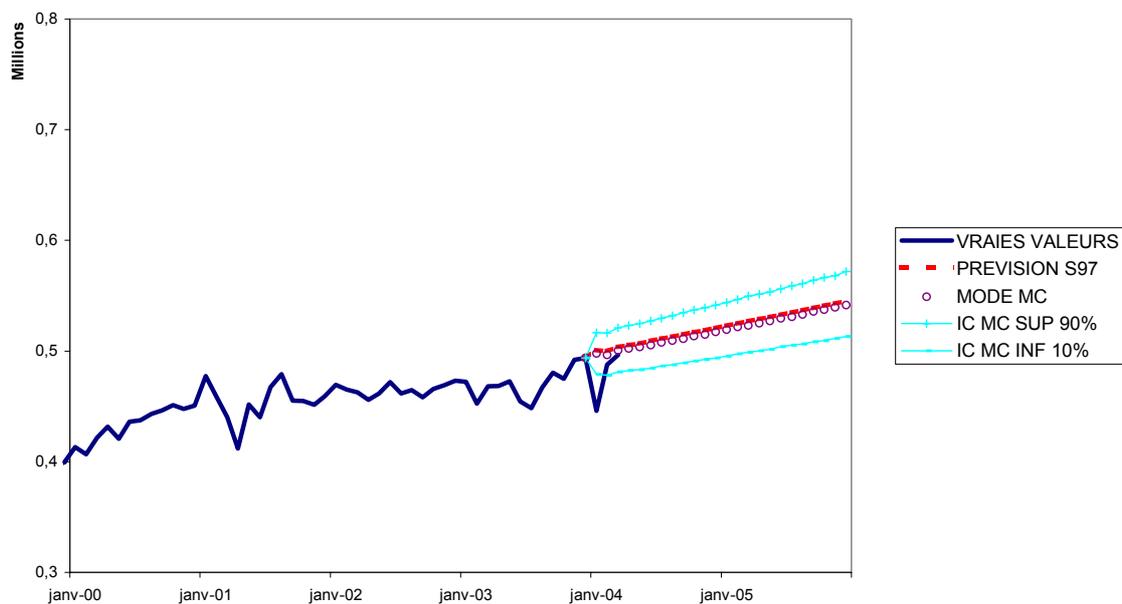
2004	0.78	0.79	0.79
2005	0.81	0.82	0.82

12.4.1.8 Trafic du dimanche/jour férié sur le RER

Modèle auto-régressif et simulations de Monte-Carlo (MC)

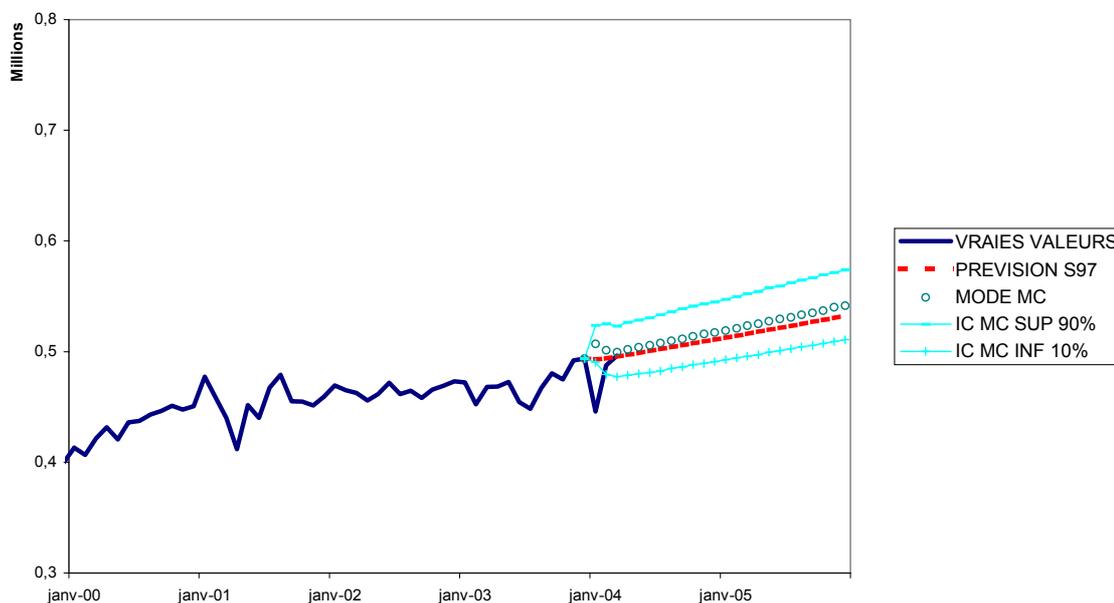


Modèle de long terme et simulations de Monte-Carlo (MC)



12.4. PRÉVISIONS DES TRAFICS RATP PAR TYPE DE JOURNÉE ET DE SOUS-RÉSEAU

Modèle de court terme et simulations de Monte-Carlo (MC)



Intervalles de confiance à 80% suivant leur mode de calcul et l'horizon considéré (en millions)

Trafic du dimanche / jour férié sur le RER	Auto-Régressif	Long Terme	Court Terme
6 mois	0.47 - 0.54	0.48 - 0.53	0.48 - 0.53
12 mois	0.46 - 0.57	0.49 - 0.54	0.49 - 0.54
18 mois	0.46 - 0.59	0.50 - 0.56	0.50 - 0.56
24 mois	0.47 - 0.61	0.51 - 0.57	0.51 - 0.57

Etendues de chaque intervalle à 80% (en millions)

6 mois	0.07	0.04	0.05
12 mois	0.10	0.05	0.05
18 mois	0.13	0.05	0.06
24 mois	0.15	0.06	0.06

Prévision moyenne de ce type de trafic journalier sur l'année (en millions)

2004	0.50	0.51	0.51
2005	0.52	0.53	0.53

12.4.2 Evolution future des différents trafics journaliers

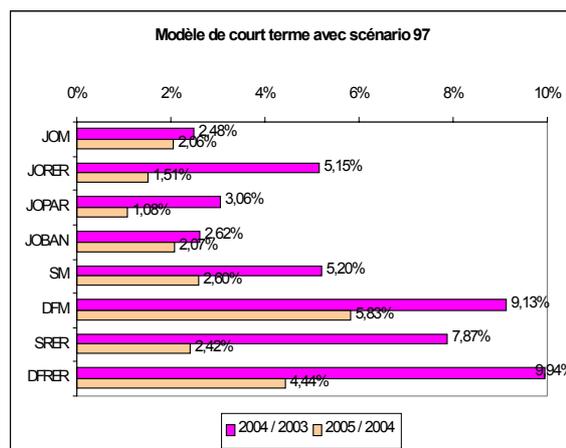
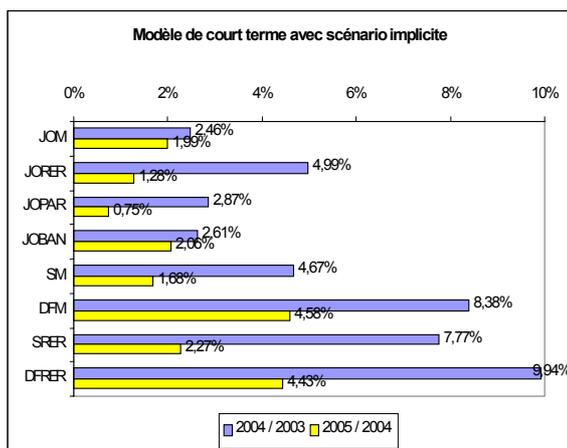
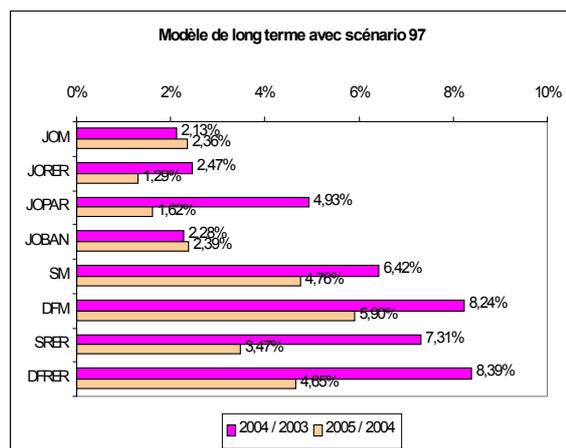
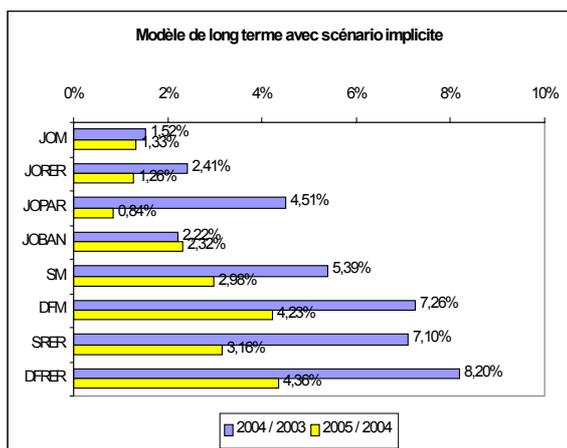
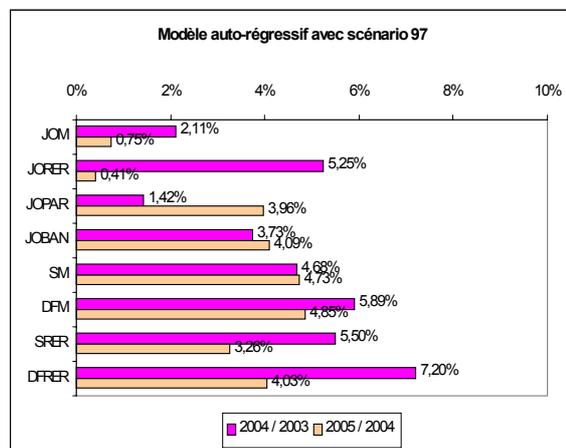
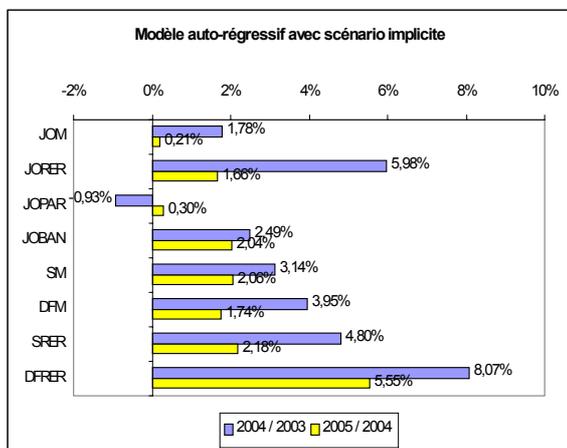
Dans cette sous-section, les évolutions des variables de trafic prévues en 2004 et 2005 sous l'hypothèse d'un scénario d'évolution semblable à celui observé en 1997, sont commentées. L'analyse porte sur les prévisions avec ce scénario, ou en prenant en compte le schéma implicite d'évolution des variables (évolution similaire à celle observée dans le passé).

La prise en compte d'une évolution des variables similaire à celle observée en 1997 amène à prévoir des trafics en augmentation en 2004 et 2005, quel que soit le type de journée considérée ou le type de sous-réseau. De façon générale, la tendance à l'augmentation des trafics journaliers sur les sous-réseaux n'est pas remise en cause lorsque les variables sont supposées évoluer comme au cours de l'ensemble des mois précédents. Seul le modèle auto-régressif avec scénario implicite conduit à une prévision moyenne du trafic du jour ouvrable sur le réseau des bus de Paris inférieure au trafic moyen observé en 2003.

L'ampleur des variations est différente selon le scénario considéré, comme le montrent les graphiques suivants. Ces graphiques comparent, pour chaque type de trafic, la prévision moyenne pour 2004 à la valeur moyenne observée en 2003, ainsi que la variation entre la prévision pour 2005 par rapport à la prévision de 2004. Dans la colonne de droite, les prévisions ont été réalisées en supposant que les variables économiques subiront dans le futur des variations moyennes identiques à celles observées dans le passé. Dans la colonne de gauche sont représentées les variations 2004/2003 et 2005/2004 sous l'hypothèse d'une évolution semblable à celle qui a eu lieu en 1997. Cette seconde colonne précise les graphiques et tableaux de la précédente sous-section.

12.4. PRÉVISIONS DES TRAFICS RATP PAR TYPE DE JOURNÉE ET DE SOUS-RÉSEAU

Evolution des trafics journaliers moyens prévus en 2004 et 2005 par rapport au trafic moyen observé en 2003



Les variations entre la prévision moyenne 2004 et la valeur moyenne observée en 2003 sont plus importantes lorsque le modèle à correction d'erreur est utilisé : ce type de modèle n'est pas le mieux adapté pour prévoir jusqu'à 12 périodes. Les graphiques ont été présentés dans un souci d'exhaustivité mais le reste de l'analyse exclura ce type de modélisation.

Rappelons que ces évolutions ne prennent en compte ni les modifications du calendrier ni les effets saisonniers mais se rapportent à la tendance du trafic journalier sur le réseau considéré. En outre, au cours de l'année 2004, devrait se produire un effet de rattrapage par rapport à 2003, qui fut une année mauvaise en termes de tourisme ou d'emploi.

L'amélioration de l'attractivité des réseaux La hausse prévue du trafic sur chacun des sous-réseaux peut en partie être justifiée par la poursuite de l'amélioration de l'attractivité des transports en commun, passant par la meilleure adaptation du service offert aux besoins de la clientèle. La contractualisation entre la RATP et le STIF³ fixe des règles concernant l'évolution de l'offre et des normes de service (confort, régularité, propreté... voir chapitre 1) mais la RATP reste force de proposition. En outre, l'entreprise a entrepris depuis plusieurs années une politique volontariste de certification de ses activités, notamment du service de ses lignes de transport. Le développement du "Service Attentionné", i.e. de la nouvelle relation de service entre la RATP et ses clients dans les différents espaces de transport (véhicules, stations de métro, gares de RER, points d'arrêt d'autobus, espaces de vente...) traduit cette volonté d'améliorer la qualité de service. Associée à l'utilisation des nouvelles technologies (passe Navigo par exemple), elle constitue un outil de fidélisation de la clientèle et de lutte contre la fraude. Cette nouvelle relation de service est mise en oeuvre sur chacun des sous-réseaux.

L'attractivité des transports collectifs urbains est renforcée par la politique de lutte contre l'utilisation trop importante de la voiture. La mairie de Paris a pour volonté d'élargir l'espace dédié aux transports en commun, aux piétons et aux cyclistes, en particulier sur les axes rouges qui drainent beaucoup de trafic automobile. Le STIF prévoit d'améliorer les parkings relais aux abords des gares de RER en Ile-de-France.

Prévisions par type de jour sur le réseau ferré

³STIF : Syndicat des Transports en Ile-de-France

Evolutions futures sur le métro Sur le métro, le trafic du jour ouvrable devrait s'accroître en 2004 et 2005.

- Selon la modélisation auto-régressive, si l'évolution constatée depuis janvier 1990 se poursuit, ce trafic devrait s'établir en 2004 à 4.223 millions de voyages au cours d'un jour ouvrable moyen, soit une augmentation de +1.78% par rapport à la moyenne constatée en 2003. En cas de croissance plus marquée, comme ce fut le cas lors de la reprise économique en 1997, ce trafic pourrait s'élever à 4.237 millions de voyages (+2.11%).

En 2005, les taux de croissance par rapport à l'année 2004 sont égaux à +0.21% pour le scénario implicite (voir page 496) et +0.75% dans le cas du scénario de reprise (voir page 499).

- D'après l'analyse des rétro-prévisions, la modélisation de long terme correspond à la modélisation la plus adaptée pour prévoir le trafic du jour ouvrable sur le métro. Les prévisions réalisées à partir de l'équation cointégrée conduisent à des prévisions légèrement différentes. Si l'environnement économique en Ile-de-France continue à évoluer comme par le passé, le trafic journalier moyen en 2004 devrait se monter à 4.213 millions de voyages (+1.52% par rapport à 2003) et pourrait augmenter jusqu'à 4.238 millions de voyages en cas de reprise de la croissance (+2.13%).

Si l'embellie économique se poursuivait avec la même force en 2005, le trafic du métro dépasserait les valeurs de la fin 2002, année exceptionnelle en terme de trafic sur le métro.

Le trafic des week-ends devrait connaître un accroissement plus marqué : entre +3.14% et +5.39% en 2004 pour le samedi selon le scénario implicite (entre +4.58% et +6.42% en cas de reprise), et plus encore au cours des dimanches et jours fériés.

L'évolution sur le métro devrait en outre bénéficier des renforcements de service des lignes 2, 3, 7, 9 et 12 avant la fin de l'année 2005. Il est en outre possible que le service soit prolongé d'une heure sur toutes les lignes de métro les vendredis soirs et samedis soirs.

Evolutions futures sur le RER Le trafic sur le RER devrait augmenter après 2003, aussi bien au cours des jours ouvrables que des samedis et dimanches/jours fériés.

- La prévision auto-régressive du trafic du jour ouvrable (1.47 millions de voyages en moyenne pour 2004) est plus élevée que celle obtenue à l'aide de l'équation de long terme (1.43 millions de voyages) et ce, quel que soit le scénario utilisé. En 2004, le modèle auto-régressif prévoit une hausse de ce trafic du jour ouvrable supérieure à +5.25%, tandis que la modélisation de long terme considère un accroissement d'environ +2.44%. Ce taux d'accroissement, plus vraisemblable empiriquement, est en outre plus fiable dans la mesure où le modèle de long terme correspond au modèle à privilégier d'après l'analyse des rétro-prévisions effectuée au chapitre précédent. Il permettrait d'atteindre des niveaux de trafic moyen du jour ouvrable sur le RER supérieurs à ceux observés en 2001 et 2002.
- Au cours des week-ends, la modélisation auto-régressive conduit pour chacun des deux scénarios et chaque type de jour (samedi ou dimanche/férié) à des prévisions plus faibles que celles issues des relations de long terme. L'analyse des rétro-prévisions a pourtant indiqué que la modélisation cointégrée conduisait à des prévisions plus fiables que la modélisation auto-régressive. En 2004, 794 000 voyages devraient être effectués le samedi sur le RER si l'évolution de l'environnement économique francilien demeure inchangée par rapport à l'évolution passée moyenne (+7.10% par rapport à 2003) et 509 000 le dimanche/jour férié (+8.20% par rapport à 2003). Ces forts taux de croissance sont envisageables : des évolutions similaires ont eu lieu en 1999 et 2000. Ils devraient toutefois diminuer en 2005 : +3.16% (par rapport à la prévision 2004) pour le samedi et +4.36% pour le dimanche, même en cas de reprise économique.

Cette hausse devrait être facilitée par la mise en service de 12 rames supplémentaires sur le RER A d'ici à 2005. Le RER B devrait quant à lui bénéficier dans les années à venir d'un renforcement de la ligne de 25% en heures de pointe et de 50% en heures creuses, et, avant la fin 2005, d'une amélioration de la ponctualité, de la fiabilité, de la sécurité...

Prévisions sur le réseau de surface

Les points d'appui du développement du réseau de surface Sur le réseau routier, deux types d'éléments structurels peuvent entraîner une hausse des trafics de la RATP :

- **Une volonté politique de développer la part des transports collectifs.** Dans le nouveau contrat STIF - RATP pour la période 2004-2007, une orientation majeure consiste en la poursuite de la mise en place du réseau de surface "Mobilien" défini par le Plan de Déplacement Urbain en Ile-de France (PDUIF). Ce réseau sera constitué à terme d'environ 150 lignes de bus, dont 70 exploitées par la RATP, reliées à des pôles d'échanges de voyageurs. L'ensemble de ces lignes doit être aménagé avec des couloirs en site propre, i.e. séparés du reste de la circulation sur la totalité de leur itinéraire. Une deuxième orientation du contrat STIF - RATP réside dans la mise en oeuvre des opérations prévues dans le contrat de Plan Etat - Région (CPER) pour la période 2000-2006 : construction/prolongement de lignes de tramway⁴, projets de bus en site propre, grands pôles d'échanges intermodaux. Ces orientations ont pour but principal de mieux répondre aux besoins de déplacements de banlieue à banlieue.
- **Une politique d'amélioration et d'adaptation du service offert à la clientèle RATP.** En plus de l'amélioration contractuelle des conditions de transport de la clientèle, de la politique de certification au sein de l'entreprise et de la mise en place du "Service Attentionné", le réseau de surface bénéficie de la mise en oeuvre de la "Bus Attitude" et de la "Tram Attitude". Cette démarche a pour but de restaurer les règles de paiement des titres de transport et de comportement à l'intérieur des autobus et des tramways. Expérimentée sur une dizaine de lignes pilotes dès octobre 2003, elle devrait être généralisée à l'ensemble du réseau de surface au cours de l'année 2004, et permettre une réduction drastique de la fraude (de 16% à 8% sur les autobus).

Evolutions futures sur le réseau des autobus de Paris Le trafic du jour ouvrable sur les autobus de Paris correspond à une exception : la modélisation auto-régressive associée au scénario implicite (i.e. évolution annuelle moyenne du trafic du jour ouvrable égale à +0.30%) conduit à une prévision en 2004 inférieure de -0.93% à la valeur moyenne de trafic observée un jour ouvrable en 2003. Le modèle de long terme considère que la valeur moyenne qui sera atteinte, en moyenne, au cours de l'année 2004, devrait avoisiner 1.25 millions de voyages par

⁴Les modifications de l'offre sur le tramway prévues dans le CPER ont été détaillées au chapitre 3 de la thèse.

journée ouvrable (scénario implicite), et être bien supérieure aux trafics moyens observés depuis 10 ans. Le modèle auto-régressif associé au scénario de reprise économique prévoit une croissance de +1.42% en 2004, et, si l'embellie économique se poursuit, à +3.96% en 2005.

Ces taux de croissances positifs peuvent s'expliquer par les éléments suivants :

- un renforcement de l'offre de service : en décembre 2003, huit nouvelles lignes d'autobus de Paris ont été renforcées, de façon à répondre à l'accroissement de la mobilité des Franciliens le soir, le week-end, et en début de soirée, mais aussi à optimiser l'offre de services sur les axes de bus bénéficiant de couloirs sécurisés, et à améliorer les relations entre Paris et les communes limitrophes. L'impact de ce renforcement sur le trafic de l'ensemble de l'année 2004 devrait s'élever à 3 millions de voyages supplémentaires sur les bus de Paris.
- une amélioration de l'offre de service. D'ici à décembre 2005, 17 lignes de bus devraient être aménagées selon les critères du Mobilien. Ces aménagements permettent d'accroître la vitesse commerciale et surtout la régularité du temps de transport sur ces lignes.

Les taux de croissance sont inférieurs à ceux observés sur les autres sous-réseaux lorsque l'on suit les recommandations de la modélisation auto-régressive. Cette modélisation prend en compte la chute brutale de trafic du jour ouvrable sur les autobus de Paris au cours de l'année 2003, alors que cette baisse n'a pas été expliquée. Il est possible qu'un problème de mesure du trafic ait lieu ou que la lutte contre la fraude porte déjà ses fruits. La modélisation de long terme amène à des prévisions supérieures, qui d'une certaine façon prolongent la tendance observée jusqu'au milieu de l'année 2003.

Evolutions futures sur le réseau des autobus de banlieue Sur le réseau des autobus de banlieue, le trafic du jour ouvrable devrait atteindre en 2004 et 2005 des niveaux plus élevés que ceux observés en moyenne depuis une dizaine d'années. Le modèle le plus pessimiste prévoit plus de 2.15 millions de voyages au cours d'un jour ouvrable de 2004, soit une hausse de +2.22% par rapport à 2003. En cas de reprise économique, le modèle auto-régressif indique que ce nombre de voyages pourrait s'élever à 2.195 millions en 2004 (+3.73% par rapport à 2003) et 2.274 millions en 2005 (+4.09% par rapport à la prévision 2004). Les prévisions réalisées à l'aide de la relation de long terme sont moins optimistes dans ce cadre de reprise économique (+2.28%

12.4. PRÉVISIONS DES TRAFICS RATP PAR TYPE DE JOURNÉE ET DE SOUS-RÉSEAU

en 2004, +2.39% en 2005, par rapport à l'année précédente).

Divers éléments permettent de justifier l'accroissement du trafic du jour ouvrable sur le réseau des autobus de banlieue :

- Le trafic du tramway, en plein essor, est comptabilisé avec les lignes de banlieue.
- Ce sous-réseau bénéficie en outre d'une amélioration constante du service offert. En 2003, par exemple, la ligne T1 a été prolongée de Bobigny à Noisy-le-Sec (5 nouvelles stations), tandis que 5 nouvelles lignes de banlieue (dont 4 de services urbains) ont été créées et qu'une quinzaine de lignes étaient prolongées ou modifiées (pour tenir compte du prolongement du tramway par exemple).

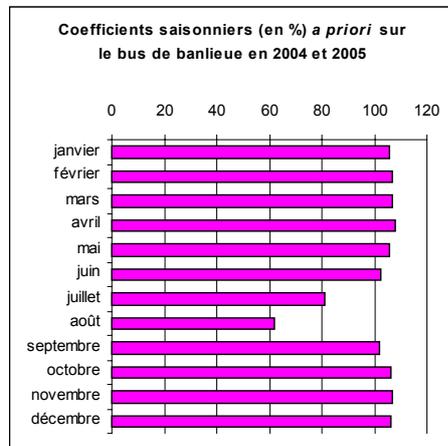
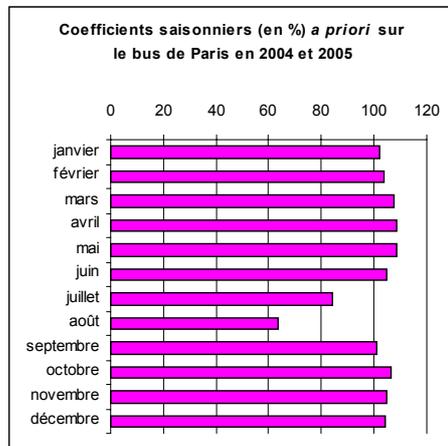
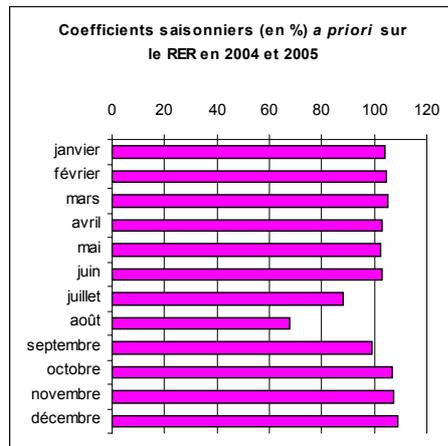
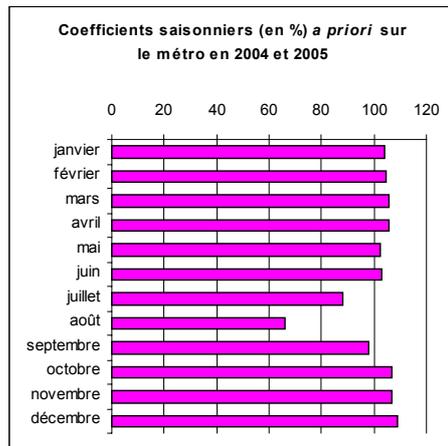
La prévision ponctuelle ne permet pas de prendre en compte le risque associé aux aléas futurs. Cette information est fournie par les intervalles de confiance simulés.

Si l'étude des variations de trafic par type de journée et de sous-réseau est porteuse d'information, les objectifs budgétaires sont évalués en termes de trafic mensuel sur chacun des sous-réseaux. La sous-section suivante présente les résultats des prévisions pour l'ensemble du mois, de même que les intervalles de confiance. La qualité de ces prévisions est évaluée par rapport aux premières valeurs observées pour 2004 et aux prévisions budgétaires réalisées antérieurement par la RATP.

12.4.3 Les prévisions sur chaque sous-réseau pour les années 2004 et 2005

A l'aide des coefficients (*a priori*) de passage entre le trafic du jour ouvrable moyen désaisonnalisé, il est possible de calculer trafic mensuel corrigé uniquement des perturbations, pour chaque type de sous-réseau. Ces coefficients de passage corrigent à la fois la modification du calendrier (i.e. la calendarité) et le poids relatifs des différents mois (i.e. la saisonnalité).

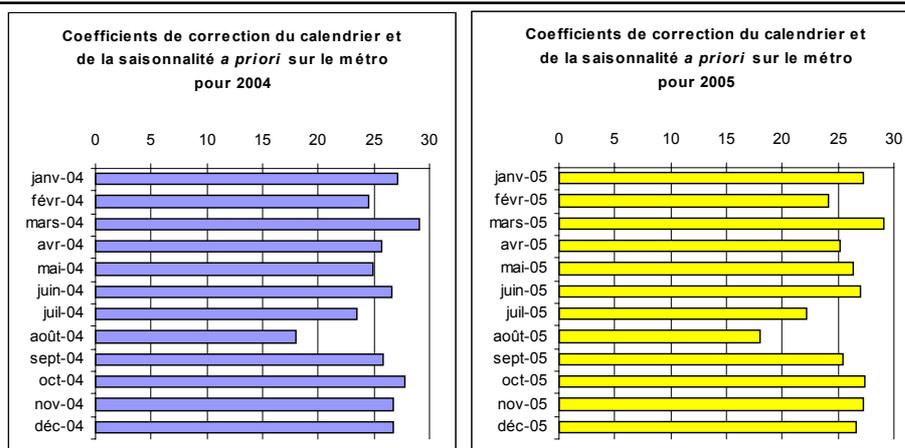
Les coefficients présentés ci-dessous, pour chaque sous-réseau, corrigent les effets saisonniers, en affectant en particulier un poids plus faible aux trafics observés durant les grandes vacances scolaires.



12.4.3.1 Le trafic prévu sur le métro en 2004 et 2005

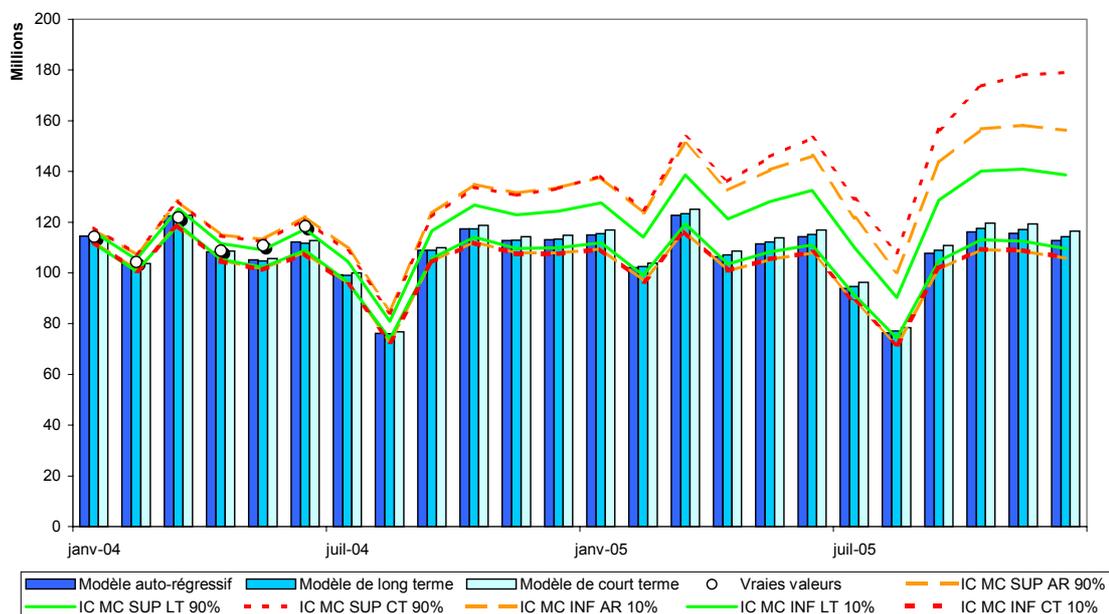
Les coefficients de passage du trafic du jour ouvrable au trafic mensuel en 2004 et 2005 sont présentés dans les graphiques ci-dessous.

12.4. PRÉVISIONS DES TRAFICS RATP PAR TYPE DE JOURNÉE ET DE SOUS-RÉSEAU



Les prévisions calculées à partir des différents modèles de prévision (AR, de LT et de CT) avec scénario et les intervalles de confiance à 80% simulés associés, ont été multipliés par ces coefficients de passage et corrigés de l'impact de Météor et de la politique de lutte contre la fraude. Les prévisions mensuelles sont représentées par des histogrammes, les intervalles de confiance simulés par les courbes de couleur, les valeurs mensuelles brutes par un rond blanc.

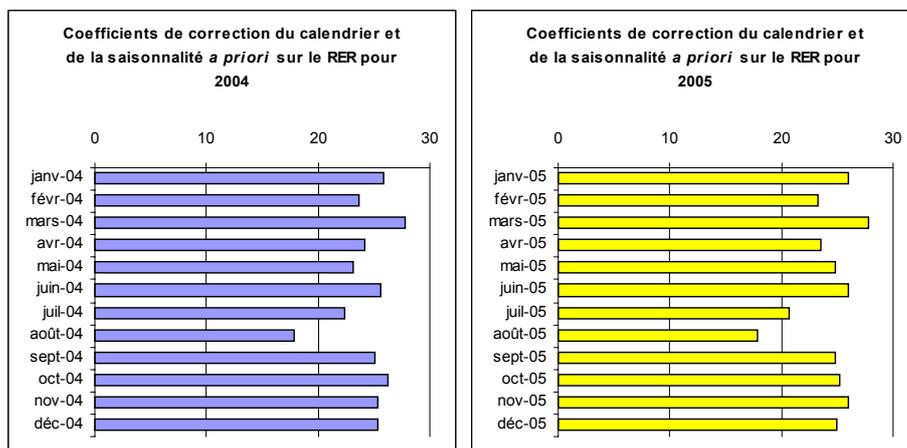
Trafics mensuels prévus sur le métro avec scénario



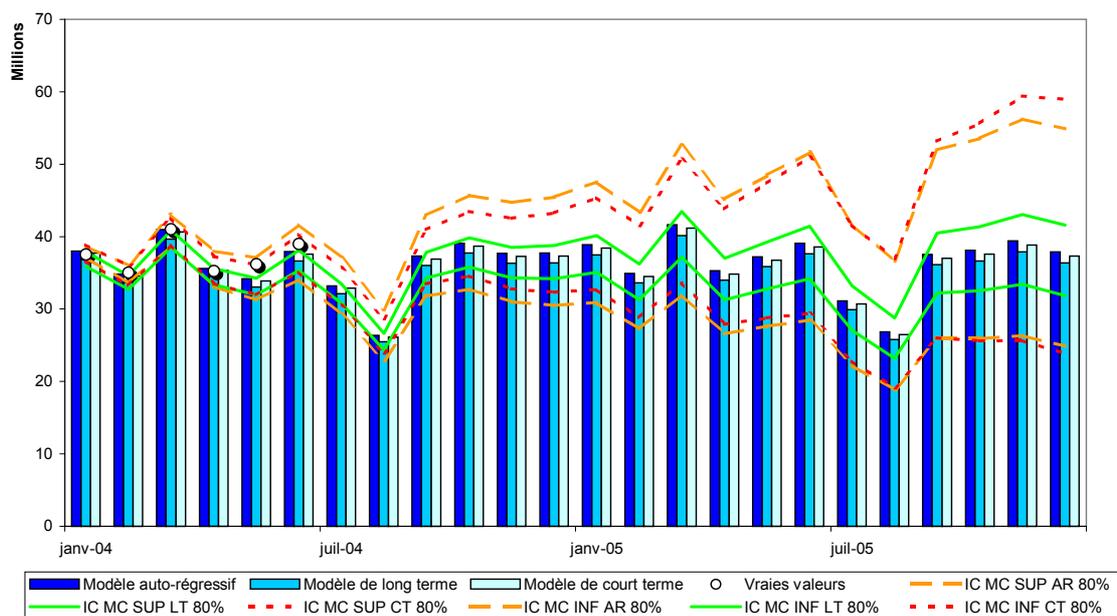
Sur les 6 premiers mois de l'année (en cumul), les valeurs observées sont supérieures de +2.5% aux prévisions de la RATP, de +1.2% à nos prévisions auto-régressives, de +1.6% à nos

prévisions de long terme et de +1.1% aux prévisions de court terme. Le point de sortie en juin 2004 a en outre mieux été prévu par nos modélisations.

12.4.3.2 Le trafic prévu sur le RER en 2004 et 2005



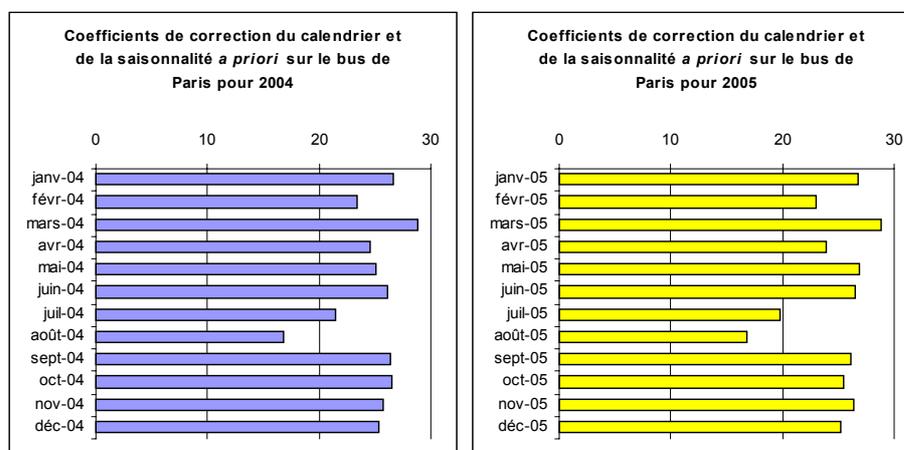
Trafics mensuels prévus sur le RER avec scénario



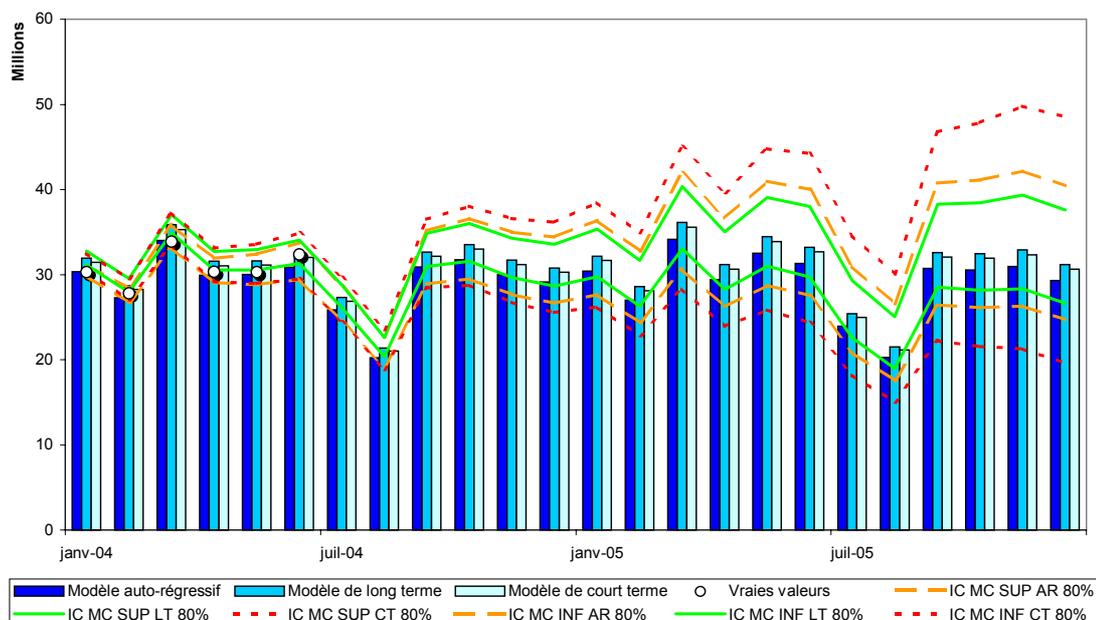
Sur le RER, le trafic cumulé observé entre janvier et juin 2004 est supérieur aux prévisions de la RATP de +7.8%, de +2.4% à nos prévisions AR, de +5.4% à nos prévisions de LT et de +2.7% à nos prévisions de CT. Tous nos modèles prévoient mieux le point de sortie de juin 2004.

12.4. PRÉVISIONS DES TRAFICS RATP PAR TYPE DE JOURNÉE ET DE SOUS-RÉSEAU

12.4.3.3 Le trafic prévu sur le réseau des autobus de Paris en 2004 et 2005

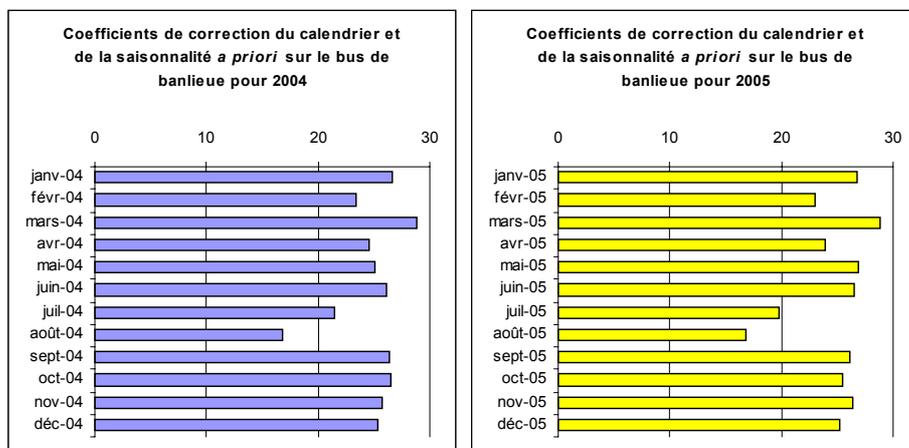


Trafics mensuels prévus sur le réseau des autobus de Paris avec scénario

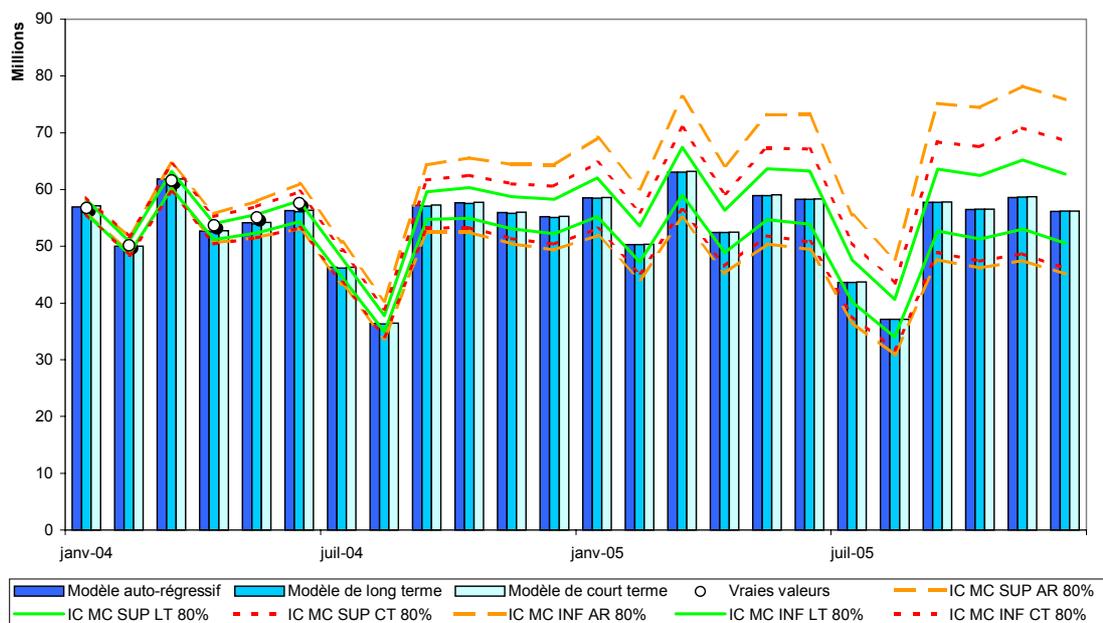


Sur le sous-réseau des autobus de Paris, le trafic cumulé observé entre janvier et juin 2004 est inférieur aux prévisions de la RATP de -0.6% , de $+0.1\%$ à nos prévisions AR, de -3.8% à nos prévisions de LT et de -2.2% à nos prévisions de CT. Tous nos modèles prévoient mieux le point de sortie de juin 2004 ($+2.9\%$ pour la prévision AR, -0.6% pour la prévision de LT et $+1.2\%$ pour la prévision de CT, contre $+2.7\%$ pour la prévision RATP).

12.4.3.4 Le trafic prévu sur le réseau des autobus de banlieue en 2004 et 2005



Traffic mensuels prévus sur le réseau des autobus de banlieue avec scénario



Sur le réseau des autobus de banlieue, les prévisions du premier semestre effectuées par la RATP sont supérieures aux valeurs observées de +2.5%, contre +0.4% pour le modèle AR, +1.4% pour le modèle de LT et +1.0% pour le modèle de CT. La prévision du point de sortie du mois de juin 2004 s'écarte de la valeur observée de +4.5% pour le modèle RATP, +1.4% pour le modèle AR, +2.7% pour le modèle de LT et +2.4% pour le modèle de CT.

12.4. PRÉVISIONS DES TRAFICS RATP PAR TYPE DE JOURNÉE ET DE SOUS-RÉSEAU

Sur l'ensemble des réseaux, les prévisions réalisées à partir de nos modélisations se sont avérées plus proches des valeurs effectivement observées *a posteriori* au premier semestre de l'année 2004. L'utilisation des méthodes présentées dans cette thèse apporte un gain dans la prévision des trafics par sous-réseau.

* * *

Ce chapitre avait pour but d'élaborer des prévisions pour les années 2004 et 2005, en permettant l'introduction d'un scénario alternatif d'évolution des variables économiques décrivant l'environnement en Ile-de-France.

Si les variables économiques poursuivent leurs évolutions passées, et en l'absence de choc majeur, les années futures devraient être marquées par une augmentation continue des trafics observés sur les sous-réseaux. Cette progression des trafics de la RATP s'explique en partie par l'amélioration constante de l'offre de service, qui apparaît dans les modèles sous la forme d'une tendance positive dans les modélisations de long terme, et d'une constante positive dans les équations auto-régressives en différence. Le trafic des week-ends sur le réseau ferré devrait continuer à s'accroître fortement, en particulier les dimanches et jours fériés.

Si une reprise économique de même ampleur que celle observée en 1997 avait lieu, les progressions de trafics seraient encore plus fortes, et pourraient conduire à des niveaux de trafic plus élevés que ceux observés depuis 1990.

Ces prévisions confirment les tendances moyennes observées depuis quelques années (à l'exception de l'année récessive 2003). Elles sont validées par les données de trafic récoltées *a posteriori*, et conduisent à de meilleurs résultats que les prévisions réalisées avec l'ancienne méthode de la RATP.

Pour chaque projection, les intervalles de prévision ont été simulés, grâce à l'utilisation d'une méthode particulière de répliquions de Monte Carlo.

Conclusion à la quatrième partie

Les parties II et III s'intéressaient à l'élaboration de modèles explicatifs de trafic en fonction du seul passé des variables de trafic ou en prenant en compte l'évolution des variables décrivant l'environnement économique en Ile-de-France. Ces modélisations corrigeaient l'existence d'une éventuelle rupture, affectant les phénomènes de mémoire des chocs passés, ou l'existence d'équilibre entre les variables de trafic et les variables environnementales. Ils avaient été estimés sur la période janvier 1990 - décembre 2002, ou sur une seconde sous-période uniquement en cas de rupture.

La partie IV valide en quelque sorte ces modélisations, en les appliquant pour prévoir les trafics futurs, et en évaluant la qualité des prévisions réalisées.

Dans le chapitre 11, a été effectué un exercice de rétro-prévisions à partir de ces modélisations. Les différents trafics par type de journée et de sous-réseau ont été "prévus" pour chaque mois de l'année 2003, en supposant connues les valeurs des variables décrivant l'activité économique francilienne. Les critères de qualité de la prévision ont permis de montrer que, pour la majorité des trafics et des horizons considérés, les modèles utilisant des variables économiques réalisaient de meilleures prévisions que les modèles auto-régressifs ou la simple projection de la tendance (méthode utilisée à la RATP). Ce chapitre a en outre permis de montrer que les modélisations ne servaient pas seulement à prévoir, mais aussi, lorsque les données d'environnement économique étaient connues, à mieux expliquer les variations passées de trafic. L'exercice de rétro-prévision permet non seulement de valider la modélisation utilisée, mais elle permet en outre d'isoler les impacts sur les trafics de la conjoncture économique des effets de rémanence des grèves ou de modification de la mobilité des utilisateurs.

12.4. PRÉVISIONS DES TRAFICS RATP PAR TYPE DE JOURNÉE ET DE SOUS-RÉSEAU

Le dernier chapitre de cette partie a été consacré à l'élaboration de prévisions réelles, pour les années 2004 et 2005. Les modèles d'estimation ont été remis à jour jusqu'à la fin de l'année 2003, sans que les paramètres estimés ne varient beaucoup par rapport à l'estimation jusqu'à fin 2002. Les prévisions de trafic ont été calculées jusqu'en décembre 2005, en supposant que les variables conservaient une évolution similaire à celle observée par le passé ou en posant un scénario optimiste de reprise économique d'une ampleur équivalente à celle observée en 1997. La prise en compte de l'incertitude des prévisions a été obtenue à l'aide des intervalles de confiance simulés calibrés. Les données par type de journée ont été ramenées à leur niveau mensuel, de façon à être cohérentes avec le type de données nécessaires à la construction des budgets. Les prévisions de trafic à l'aide de nos modélisations sont plus proches des valeurs effectivement observées depuis que ne le sont les prévisions réalisées à l'aide de l'ancienne méthode de la RATP. Les trafics par sous-réseaux devraient croître dans les années à venir.

Conclusion générale

Dans cette de thèse, nous nous sommes intéressés à l'élaboration de modèles opérationnels d'analyse et de prévision des trafics mensuels de la RATP, sur ses différents sous-réseaux : le métro, le RER, les bus de Paris ainsi que les bus de banlieue. L'analyse a été réalisée en prenant en compte des trafics journaliers types (trafic du jour ouvrable, trafic du samedi, trafic du dimanche/jour férié).

Trois types de modélisation ont été envisagés pour expliquer et prévoir les trafics : une modélisation auto-régressive, fonction du seul passé de chacune des variables, ainsi que deux analyses issues de la cointégration (modèle de long terme et dynamique de court terme, appelée aussi modèle à correction d'erreur), qui prennent en compte des variables décrivant l'environnement économique en Ile-de-France.

Pour chaque type de trafic, ces modélisations ont été utilisées pour analyser les variations de trafic au cours de l'année 2003 et prévoir les trafics en 2004 et 2005. Les prévisions ont été réalisées en prenant en compte une évolution des variables économiques similaire à leur tendance passée. Nous avons en outre considéré que les variables décrivant l'environnement économique en Ile-de-France retrouveraient en 2004 des taux d'évolution semblables à ceux observés lors de la précédente reprise économique, en 1997. Le risque associé à chaque prévision ponctuelle a été pris en compte par la construction d'intervalles de confiance simulés. Les résultats ont été comparés aux valeurs effectivement observées au cours de l'année 2003 et du premier semestre 2004.

L'existence de rupture éventuelle a été prise en compte tout au long de ce travail, grâce à nos tests de modification du degré d'intégration et de cointégration. Ces tests permettent de tester l'hypothèse d'une rupture du degré d'intégration ou de cointégration et estiment la date de rupture. La période est alors scindée en deux sous-périodes, et l'estimation des paramètres est réalisée sur la seconde sous-période uniquement.

L'ensemble de ce travail correspond à une **réponse à la demande de la RATP**, qui désirait améliorer ses méthodes d'analyse et de prévision des trafics sur les sous-réseaux. Il a permis de formaliser des méthodes employées couramment dans l'unité Tarification Ventes Résultats du département Commercial de la RATP et de constituer une base de travail sur laquelle l'unité pourra s'appuyer lors de ses analyses sur longue période. En outre, la classification des stations de métro ainsi que des gares de RER en fonction des motifs à l'origine des déplacements et des modes de rabattement a permis d'enrichir le rapport annuel édité par l'entité. Enfin, le caractère opérationnel de nos modélisations a été accru par la construction de fichiers permettant le calcul des rétro-prévisions et des prévisions pour différents types de scénarios. Des macros ont été élaborées sous Visual Basic afin de permettre la mise à jour des paramètres de chacun des trafics étudiés, ou l'étude d'autres types de trafic.

D'un **point de vue économétrique**, la principale avancée de la thèse réside dans la construction de nos tests de modification du degré d'intégration. L'analyse préliminaire de nos données de trafic ou d'environnement en Ile-de-France invalidait à la fois l'existence d'une racine unitaire et l'absence de celle-ci sur deux sous-périodes⁵. Certaines séries semblaient avoir subi une modification de leur degré d'intégration, i.e. être stationnaire sur une sous-période et intégrée sur l'autre sous-période. Une synthèse des tests de changement de structure existants a montré l'absence de tests de modification du degré d'intégration ayant pour hypothèse nulle la présence de racine unitaire sur l'ensemble de la période. Nous avons donc construit de tels tests (hypothèses et statistiques de tests, calcul des valeurs critiques) en supposant que la date

⁵Les notions de "non-stationnarité", "d'intégration" ou de présence de "racine unitaire" formalisent le fait de garder en mémoire les chocs passés, de façon pérenne.

A l'inverse, en présence de "stationnarité", l'impact des chocs passés s'atténue au fur et à mesure.

de rupture était inconnue, et évalué leurs propriétés en termes de puissance. Les simulations de Monte Carlo réalisées ont montré que les hypothèses alternatives de modification du degré d'intégration étaient particulièrement bien détectées, et que la date de rupture était estimée avec une grande précision dans le cas d'une série intégrée puis stationnaire. Une procédure de test a été élaborée : suivant les conclusions du test d'intégration standard puis, éventuellement, des tests de modification du degré d'intégration, l'estimation auto-régressive est effectuée sur tout ou partie de la période, sur la variable en niveau ou bien en différence⁶. Ces tests ont ensuite été appliqués aux séries de trafic et d'environnement économique dans la région Ile-de-France. Pour chacune de ces séries, détaillées dans la première partie, les conclusions des tests de modification du degré d'intégration ont été comparées aux conclusions des tests de Perron. Cette comparaison a montré la pertinence de nos tests en présence d'une modification du degré d'intégration. La prise en compte de la seconde sous-période uniquement lors de l'estimation des paramètres permet en outre de prévoir aussi bien (voire mieux) les différents trafics : nos tests permettent d'obtenir un gain de parcimonie.

Deuxième apport économétrique, la construction de nos tests de modification du degré de cointégration correspond à une suite logique : les tests de modification du degré d'intégration sont appliqués aux résidus d'une relation non cointégrée. Après avoir vérifié que ce type de tests n'avait pas été développé dans la littérature existante, les hypothèses de tests et les statistiques ont été détaillées en prenant en compte deux types d'hypothèses alternatives ("Absence puis présence de cointégration" ou "Présence puis absence de cointégration) et une date de rupture inconnue. Des simulations de Monte Carlo ont permis de tabuler les valeurs critiques mais aussi de calculer la puissance de ces tests et d'évaluer la qualité de l'estimation de la date de rupture pour différents cas de figure (place de la véritable date de rupture, modification éventuelle de la constante entre les deux sous-périodes, valeurs du paramètre auto-régressif lorsque la relation testée n'est pas cointégrée...). Nous pensons que ce type de test permet de dater l'apparition ou la disparition d'un équilibre de long terme. L'application de ces tests à des relations entre un type de trafic et les données environnementales a permis de mettre en lumière des équilibres,

⁶En présence d'une racine unitaire, les estimations auto-régressives sont réalisées sur les séries en différence. Cette différenciation n'est pas nécessaire pour les processus stationnaires.

valides sur une seconde sous-période uniquement pour la plupart des trafics. Sans nos tests de modification du degré de cointégration, nous n'aurions pu modéliser les variables de trafic du réseau ferré en fonction de l'environnement économique francilien et améliorer autant la qualité des prévisions réalisées.

Un troisième apport de cette thèse réside dans la construction d'intervalles de confiance à l'aide de simulations de Monte Carlo "calibrées". Cette façon de procéder est innovante dans la mesure où les bruits futurs sont simulés (mais calibrés grâce à l'évolution passée des variables). Les méthodes de simulation traditionnelles simulent des bruits dans le passé, estiment la relation et effectuent des prévisions ponctuelles. Notre méthode nous semble à la fois plus rapide et plus performante.

Ces deux types de tests comme la simulation calibrée des intervalles de confiance ne concernent pas uniquement notre cas d'étude : elles sont transposables à tout autre secteur concernant des séries non stationnaires.

D'un **point de vue économique**, ce travail a permis de mieux comprendre chacune des séries de trafic et d'environnement économique en Ile-de-France, grâce aux tests d'intégration et à nos tests de modification du degré d'intégration.

De façon générale, sur chacun des sous-réseaux et pour tout type de journées, les utilisateurs des réseaux gardent en mémoire les chocs passés, qu'ils soient positifs (amélioration de l'offre de service, création d'abonnements attractifs...) ou négatifs (grèves ou attentats par exemple). Des éléments souvent exogènes à la seule activité de transporteur public expliquent les modifications du comportement des consommateurs.

Par exemple, la modification du degré d'intégration de la série de trafic sur le RER un jour ouvrable est en partie imputable à la récession économique de 1993, à l'augmentation du chômage,... Le trafic du dimanche/jour férié sur le RER a subi une rupture en 1998, après l'ouverture du prolongement de la ligne B du RER à la "Plaine Stade de France" et juste avant la création de la carte Imagine R et l'essor des trafics les journées de week-ends.

Cette analyse a montré l'importance de prendre en compte la conjoncture économique dans les explications de trafic.

Nous avons pu quantifier les liens étroits entre les trafics de la RATP et les variables économiques décrivant la situation économique en Ile-de-France.

Cette approche a permis de mettre en lumière des équilibres, valides sur une seconde sous-période uniquement pour la plupart des trafics.

Seuls les trafics du jour ouvrable sur les sous-réseaux de bus (Paris et banlieue) possèdent un équilibre de long terme sur l'ensemble de la période, ce qui traduit le caractère progressif des éventuelles modifications des comportements des utilisateurs du réseau routier.

Les trafics sur le métro, pour le jour ouvrable et le dimanche ou jour férié, se sont modifiés dès les grèves de novembre-décembre 1995, la rupture ayant eu lieu deux années après pour la journée du samedi. Le trafic du RER a subi une rupture plus tardive : fin 1997 pour le trafic du jour ouvrable, et durant l'été 1998. Ces modifications plus tardives s'expliquent par la modification de comportement induite par la reprise économique, surtout en terme de consommation, la forte augmentation de l'attractivité de la région parisienne comme destination touristique (Journées Mondiales de la Jeunesse, coupe du monde de football...) et la création de la carte Imagine R.

Les sous-réseaux sont affectés de manière différente par l'évolution conjoncturelle.

A long terme, les variations du trafic du métro sont expliquées par l'évolution du tourisme ; la situation sur le marché de l'emploi influence le niveau observé au cours des jours ouvrables, la consommation des ménages influe sur le trafic observé les week-ends.

Sur le RER, l'impact du tourisme sur le trafic est moins important, les variations de trafic dépendent en grande partie de la consommation des ménages.

Il en est de même sur les autobus de banlieue, qui constituent un mode de rabattement sur les gares de RER.

Sur le sous-réseau des autobus de Paris, enfin, les conditions de circulation ont un impact négatif sur le trafic mais l'augmentation du nombre de lignes protégées devrait diminuer le poids de cette variable au fur et à mesure. Evolution du chômage, de la consommation et du tourisme expliquent les variations du trafic.

Le prix du transport collectif (par rapport au prix de l'essence) n'apparaît pas comme un déterminant essentiel de l'utilisation des transports en commun, ce qui souligne l'existence d'une demande captive (car ne possédant pas de voiture par exemple) ou acquise aux transports collectifs (pour des raisons écologiques par exemple). Les ajustements de court terme ont pour but principal de se ramener vers la trajectoire de long terme.

La prise en compte de variables économiques relatives à la seule Ile-de-France est apparue comme un élément primordial pour la modélisation : les variations de la conjoncture au niveau régional sont relativement indépendantes de celles observées au niveau de la France entière. Elles dépendent plus de la situation internationale, qui possède une influence forte sur le tourisme en Ile-de-France ou encore l'activité dans les grands magasins de la région parisienne⁷.

Plusieurs axes de **prolongement** à la thèse peuvent être envisagés.

Le premier axe de prolongement correspond à un élargissement de l'approche économique, avec une prise en compte de données économiques supplémentaires.

Il serait intéressant de prendre en compte dans les modélisations des variables décrivant l'immobilier en Ile-de-France (immobilier résidentiel et immobilier de bureau) : l'implantation des habitats et des bureaux influence fortement le mode de transport choisi. Ce type de données n'est malheureusement pas disponible. Un travail de construction d'indices d'attractivité de Paris en termes de situation immobilière (habitation ou bureau), et son intégration dans les modélisations, permettrait sans doute d'améliorer les prévisions de trafic, surtout à long terme.

L'offre de service a été traitée de manière assez superficielle : elle apparaît dans les modélisations sous la forme d'une tendance positive, et explique certaines des ruptures que nous avons mises en lumière. Par exemple, sur le trafic des week-ends, la mise en place de la carte Imagine R en 1998 ou l'augmentation de l'offre de service au cours des samedis, dimanches et jours fériés, expliquent sans doute la formation d'un nouvel équilibre. Nous avons essayé de prendre en compte dans les modélisations les indicateurs de qualité de service. Cette démarche n'a pas permis d'apporter d'éclairage nouveau dans la mesure où ils ne reflétaient pas réellement les préoccupations

⁷25% du chiffre d'affaires en volume des grands magasins en Ile-de-France est réalisé par des touristes.

principales : l'indicateur de régularité sera affecté de la même façon par une attente égale à 4 minutes en heure de pointe que par une attente supérieure à une heure. Le ressenti des voyageurs sera pourtant bien différent. De nouveaux indicateurs allant dans ce sens ont été mis en place dans le nouveau contrat entre la RATP et le STIF (Syndicat des Transports d'Ile-de-France), ils concernent les "situations inacceptables" sur chaque ligne.

Un second approfondissement, au sein de l'entreprise, pourra consister dans la modélisation et la prévision des trafics de la nouvelle décomposition du réseau des autobus.

Jusqu'en 2003, le réseau des autobus distinguait les autobus de Paris et les autobus de banlieue. Afin de suivre les impacts de l'amélioration de l'offre de service sur le réseau de surface (en particulier la mise en place du "Mobilien" défini dans le cadre du PDU⁸), une nouvelle décomposition de ce réseau a été établie. Elle distingue les TCSP (Transports en Commun en Site Propre, comprenant les lignes de tramway et le Trans Val de Marne), les lignes "PDU" et hors "PDU".

Lorsque ces séries posséderont un historique suffisamment long, il sera intéressant d'appliquer le même type de modélisation que celles développées dans la thèse, afin de mettre en lien les évolutions de trafic constatées et l'environnement économique. D'ici là, des modélisations de type SARIMA demeureront plus adaptées

Un deuxième axe de prolongement concerne l'économétrie, et les extensions possibles à nos tests de modification du degré d'intégration et de cointégration.

Les tests de modification du degré d'intégration pourraient être développés en prenant en compte des données de panel, i.e. en segmentant les différents trafics. Par exemple, la classification des stations de métro et des gares de RER permettrait une segmentation des différents trafics en fonction de leurs motifs de déplacement. Il serait en outre intéressant d'analyser les

⁸Le Plan de Déplacements Urbains de l'Ile-de-France vise à réduire la place de la voiture particulière grâce à une amélioration des transports en commun. Le projet Mobilien consiste en une amélioration de la qualité de 150 lignes de bus à fort trafic, ainsi que de 150 pôles (gares de RER ou stations de métro) qu'elles desservent. La RATP est impliquée dans ce projet sur 70 lignes de bus et 90 pôles, situés principalement à Paris et en Petite Couronne.

Pour les lignes de bus, ces améliorations concernent l'augmentation de la vitesse et de la régularité (avec les couloirs de bus par exemple), la modernisation des bus pour en accroître le confort, une adaptation des horaires de service aux demandes des voyageurs ainsi qu'une augmentation de l'information en temps réel aux points d'arrêt.

trafics ligne par ligne, voire par station, afin de mettre en lumière les spécificités éventuelles des comportements des agents à un niveau fin.

L'analyse des données de panel suppose de disposer à la fois d'un grand nombre d'individus et d'observations. Cette conjonction permet de se rapprocher des résultats théoriques, le plus souvent asymptotiques dans un cadre non stationnaire. L'extension de nos tests de modification du degré d'intégration aux données de panel, et leur application, serait d'autant plus intéressante que la RATP dispose d'un grand nombre de données, surtout pour le réseau ferré. Par exemple, le métro comporte 14 lignes (sans compter les fourches) et 296 stations et les trafics mensuels par station ont été conservés sous forme informatique depuis plus de 15 années.

Enfin, une modélisation multivariée permettrait d'étudier le niveau de concurrence ou de complémentarité entre les différents sous-réseaux de la RATP. Nous avons en particulier souligné que les évolutions du bus de banlieue étaient fortement corrélées à celles observées sur le métro et le RER, car ils constituent un mode de rabattement au réseau ferré important.

Cette analyse pourrait en outre prendre en compte les trafics de la SNCF en Ile-de-France.

L'interdépendance des variables assurerait une meilleure cohérence des estimations. Se poserait toutefois le problème de l'interprétation des relations de cointégration et des dates de rupture.

Bibliographie

- [1] AIRPARIF (2003), "Life "Resolution" bilan final", 4 pages, disponible sur www.airparif.asso.fr/airparif/pdf/3LIFE_syn.pdf
- [2] ALTISSIMO F. et V. CORRADI (1999), "Strong rules for detecting the number of breaks in a time series", Internet : <http://www.bc.edu/RePEc/es2000/0574.pdf>
- [3] AMEMIYA, T. (1985), "Advanced Econometric Theory", Blackwell.
- [4] AMSLER C. et LEE J. (1995), "An LM test for a unit root in the presence of a structural change", *Econometric Theory* 11, 359 - 368
- [5] ANDREWS D.W.K. (1991), "Heterostedasticity and autocorrelation consistent covariance matrix estimation", *Econometrica* 59, 817 - 858 ou internet : <http://cowles.econ.yale.edu/P/cp/p07b/p0780.pdf>
- [6] ANDREWS D.W.K. (1993), "Tests for parameter instability and structural change with unknown change point", *Econometrica* 61(4), 821 - 856 ou Internet : <http://cowles.econ.yale.edu/P/cp/p08a/p0845.pdf>
- [7] ANDREWS D.W.K. , et W. PLOBERGER (1994), "Optimal tests when a nuisance parameter is present only under the alternative", *Econometrica* 62(6) ou Internet : <http://cowles.econ.yale.edu/P/cp/p08b/p0879.pdf>
- [8] ANDREWS D.W.K. , I. LEE et W. PLOBERGER (1996), "Optimal changepoint tests for normal linear regression", *Journal of Business and Economic Statistics* 10, 271 - 287
- [9] ARBIC A. et M. GAUDRY(1996) , "On the demand for and value of monthly transit passes within the framework of the DEMTEC transit ridership demand model", 6 pages, AJD-21 disponible sur <http://www.ajd.umontreal.ca/source-pdf/ajd-21.pdf>

-
- [10] ASHLEY R., C. W. J. GRANGER et R. SCHMALENSEE (1980), "Advertising and aggregate consumption : an analysis of causality", *Econometrica*, 48, 1149 - 1167
- [11] AUDENIS C., P. BISCOURP et N. RIEDINGER (avril 2003), "Le prix des carburants est plus sensible à une hausse qu'à une baisse du brut", *Revue Economie et Statistiques de l'INSEE* n°359 - 360, disponible sur www.insee.fr/fr/ffc/tabpdf/ES359H.pdf
- [12] BAI J. (1994), "Least squares estimation of a shift in linear processes", *Journal of Time Series Analysis* 15(5), 453 - 472
- [13] BAI J. (1997), "Estimation of a change point in multiple regression models", *Review of Economics and Statistics* 79(4), 551 - 563
- [14] BAI J. (1997), "Estimating multiple breaks one at a time", *Econometric Theory* 13(3), 315 - 352
- [15] BAI J. (1998), "A note on spurious regression", *Econometric Theory* 14(5), 663 - 669
- [16] BAI J. (1999), "Likelihood ratio tests for multiple structural changes", *Journal of Econometrics* (91), 299 - 323
- [17] BAI J., R.L. LUMSDAINE et J.H. STOCK (1998), "Testing for and dating common breaks in multivariate time series", *Review of Economic Studies* 64(3), 395 - 432
- [18] BAI J. et P. PERRON (1998), "Estimating and testing linear models with multiple structural changes", *Econometrica* 66(1), 47 - 78
- [19] BAI J. et P. PERRON (1998), "Computation and analysis of multiple structural change models", Internet : <http://www.crde.umontreal.ca/cahiers/0998-cah.pdf> ou *Cahier n°0998 du Centre de Recherche en Développement et Economique, Université de Montréal*
- [20] BAI J. et P. PERRON (2001), "Multiple structural change models : a simulation analysis", Internet : <http://econ.bu.edu/perron/papers/break.pdf>, à venir dans *Journal of Applied Econometrics*.
- [21] BANERJEE A. (1995), "Dynamic specification and testing for unit roots and cointegration", in *K.D. Hoover (eds), "Macroeconometrics", Kluwer, Boston, chapter 12*

-
- [22] BANERJEE A., R.L. LUMSDAINE et J.H. STOCK (1992), "Recursive and sequential tests of the unit-root and trend-break hypotheses : Theory and International Evidence", *Journal of Business and Economic Statistics* 10, 271 - 287
- [23] BAUMOL W., J. PANZAR et R. WILLIG (1982), "*Contestable markets and the theory of industry structure*", Harcourt Brace Jonavitch, New York.
- [24] BEN AKIVA M. et S. LERMAN (1985), "Discrete choice analysis : theory and application to to travel demand", *Cambridge, MIT Press*
- [25] BERTRAND F. et L. JALUZOT (2003), "Plus de la moitié des déplacements des Franciliens pour leurs affaires personnelles et loisirs", *Ile-de-France "A la page", INSEE, n°231*
- [26] BIGOT J.F. et D. GOUX (2003), " Enquête sur l'emploi du premier trimestre 2003", *INSEE Première n°909*
- [27] BONNAFOUS A. (1972), "La logique de l'investigation économétrique", *Thèse pour le Doctorat es Sciences Economiques, Université Lumière 2, Lyon*
- [28] BONNAFOUS A., P. BONNEL, D. BOUF, L. CLEMENT, O. KLEIN, F. LEURENT, J.P. NICOLAS M. LE NIR et A ? PERL (1995), " A review of current modelling practice in France", *Rapport pour le compte du Transport Research Laboratory du Ministère des Transports britannique. Lyon, Laboratoire d'Economie des Transports, Ecole Nationale des Travaux Publics de l'Etat, Université Lumière Lyon II, INRETS, 1995, 24 pages*
- [29] BOSWIJK H.P. (1992), "Cointegration, identification and exogeneity", *Ph.D. Dissertation, Tinbergen Institute, n°37, Tibergen Institute Research Series*
- [30] BOSWIJK H.P. (1994), "Testing for unstable root in conditionnal and structural error correction models", *Journal of Econometrics, 63, 37 - 60*
- [31] BOULIN J.Y. et C. du TERTRE (2001), "L'impact de la RTT sur les usages du temps : conséquences pour les loisirs et le tourisme", *Rapport de l'IRIS - Université de Paris Dauphine pour le Commissariat Général au plan et la Direction du tourisme.*
- [32] BUSETTI F. et A.M.R. TAYLOR (2002), "Tests of stationarity against a change in persistence", *University of Birmingham Department of Economics Discussion Paper n° 01-13*, ou Internet : <http://www.staff.city.ac.uk/~giourga/BusettiTaylor.PDF>

-
- [33] "Calgary Transit Ridership Forecasting" (1994), *Calgary Transit, Service Planning and Community Relations, Corporate Economics Finance Department*
- [34] CAMPOS J., N.R. ERICSSON et D.F. HENDRY (1996), "Cointegration tests in the presence of structural breaks", *Journal of Econometrics*, 70, 187-220
- [35] CHEUNG Y. W. et K. S. LAI (1993), "Finite-sample sizes of Johansen's likelihood ratio tests for cointegration", *Oxford Bulletin of Economics and Statistics*, 55, 313-328
- [36] CHONG T. (1995), "Partial parameter consistency in a misspecified structural change model", *Economic Letters* 49 (4), 351 - 357
- [37] CHOW G.C. (1960), "Tests of equality between sets of coefficients in two linear regressions", *Econometrica* 28, 591 - 605
- [38] CHRISTIANO L.J. (1992), "Searching for a break in GNP", *Journal of Business and Economic Statistics* 10 (3), 427 - 431
- [39] CLARK T. et M. McCRACKEN (2002), "Forecast-based model selection in the presence of structural breaks", Internet : <http://www.kc.frb.org/publicat/reswkpap/pdf/rwp02-05.pdf>
- [40] CHRISTOFFERSEN P.H. et F.X. DIEBOLD (1997), "Cointegration and long-horizon forecasting", *NBER Technical Working Paper 217*
- [41] Comité des Constructeurs Français d'Automobiles (CCFA), sur www.ccfa.fr
- [42] COUR DES COMPTES (1999) "Rapport Public 1999", *Chapitre VIII " Entreprises Publiques", Section 29 "Eole et Météor (1989-1999)*, disponible sur www.ccomptes.fr/cour-des-comptes/publications/rapports/rp1999/rp1999_37.html
- [43] DAGENAIS M. et M. GAUDRY (1979), "The DOGIT model", *Transportation Research B* 13, 2, 105 - 112
- [44] DAGENAIS M., M. GAUDRY et T.C. LIEM, "Urban travel demand : the impact of Box-Cox Transformation with non spherical residuals errors", *Transportation Research B* 21, 443 - 477

-
- [45] DICKEY D.A. et W.A. FULLER (1979), "Distribution of the estimators for time series regressions with a unit root", *Journal of the American Statistical Association* 74, 427 - 431
- [46] DICKEY D.A. et W.A. FULLER (1981), "The likelihood ratio statistics for autoregressive time series with unit root", *Journal of the American Statistical Association*, 49 (4), 1057 - 1072
- [47] Direction Générale de l'Energie et des Matières Premières du Ministère de l'Economie, des Finances et de l'Industrie, Observatoire de l'Energie (Mars 2003), "Consommation de carburants des voitures particulières en France 1990 -2001", disponible sur www.industrie.gouv.fr/energie/statisti/pdf/consom_carbu_01.pdf
- [48] Direction Régionale de l'Industrie, de la Recherche et de l'environnement (décembre 2001), "Le pétrole en Ile de France. De la ressource aux consommateurs", disponible sur www.ile-de-france.drirc.gouv.fr/energie/b-petrole/b-petrole-complete.pdf
- [49] DUGUIT L. (1928), "Libertés publiques ", *Editions Cujas, Collection Traité de droit constitutionnel, Paris, Edition 1981.*
- [50] ENGLE R.F. et C.W.J. GRANGER (1987), "Co-integration and error correction : representation, estimation and testing", *Econometrica*, 55, 251-276
- [51] ETTEMA D.F. et H.J.P TIMMERMANS (1997), "Activity-based approaches to travel analysis", *Pergamon, Elsevier Science Ltd, Oxford.*
- [52] GABRIEL V. J., M. SOLA et Z. PSARADAKIS (2001), "Residual-based tests for cointegration and multiple regime shifts", *NIPE Working Paper n° 7/2002, Universidade do Minho*, disponible sur l'internet : <http://www.eeg.uminho.pt/economia/nipe/vgabriel/mscoint.pdf>
- [53] GARCIA R. et P. PERRON (1996), "An analysis of the real interest rate under regime shifts", *Review of Economics and Statistics* 78(1), 111 - 125 ou Internet : <http://www.cirano.qc.ca/pdf/publication/95s-05.pdf>

-
- [54] GAUDRY M. (1973), "The Demand for public transit in Montream and its implications for transportation planning and cost-benefit analysis", Ph.D. Thesis, Economics Department ; Princeton University and University Microfilms, Ann Harbor, Michigan, 188 pages
- [55] GAUDRY M. (1975), "An Aggregate time-series analysis of urban transit demand : the Montreal case", *Transportation research* 9(4), 249 - 258
- [56] GAUDRY M. (1978), "Seemingly unrelated static and dynamic urban travel demands", *Transportation Research* 12(3), 195 - 211
- [57] GAUDRY M. (1981), "The INVERSE POWER TRANSFORMATION LOGIT and DO-GIT mode choice models", *Transportation Research B* 15(2), 97 - 103
- [58] GAUDRY M. (1985), "Modèles de demandes agrégés et désagrégés à forme variable : résultats sur Montréal et Paris", *Transports n°304*, 288 - 293
- [59] GAUDRY M. et T. LIEM (1998), "QDF : a Quasi-Direct Format Used to combine Total and mode choice results to obtain modal elasticities and diversion rates", *Centre de Recherche sur les Transports, Université de Montréal, Working paper n°982*
- [60] GAUDRY M. et M. WILLS (1978), "Estimating the functional form of travel demand models", *Transportation Research* 12, 257 - 289
- [61] GEORGE S. (2003), "Le bus de Paris : comment améliorer l'offre selon des critères quantitatifs", *RATP Département Commercial Unité Marketing - Connaissance clientèle et préconisations*
- [62] GRANGER C. W. J. (1969), "Investigating causal relations by econometric models and cross-spectral methods", *Econometrica* 37, 428 - 438
- [63] GRANGER C. W. J. (1980), "Testing for causality : a personal viewpoint", *Journal of Economic Dynamics and Control*, 2, 329 - 352
- [64] GRANGER C. W. J. (1988), "Some recent developments in a concept of causality", *Journal of Econometrics*, 39, 199 - 211
- [65] GREGORY A.W. et B.E. HANSEN (1996), "Residual-based tests for cointegration in models with regime shifts", *Journal of econometrics*, 70(2), 99-126

-
- [66] GREGORY A.W. et B.E. HANSEN (1996), "Tests for cointegration in models with regime shifts", *Oxford Bulletin of Economics and Statistics*, 58, 555-560
- [67] GREGORY A.W., J.M. NASON et P.G. WATT (1996), "Testing for structural breaks in cointegrated relationships", *Journal of Econometrics*, 71, 321-341
- [68] GRIMM C. et R. HARRIS (1983), "Structural economics of the U.S. rail freight industry : concepts, evidence, and merger policy implications", *Transportation Research*, 17(4), 271-281
- [69] HAO K. (1996), "Testing for structural change in cointegrated regression models : some comparisons and generalizations", *Econometric Reviews*, 15, 401-429
- [70] HALL A. (1994), "Testing for a unit root in time series with pretest data-based model selection", *Journal of Business and Economic Statistics*, 12, 461-470
- [71] HALL S. G., Z. PSARADAKIS et M. SOLA (1997), "Cointegration and changes in regime : the japanese consumption function", *Journal of Applied Econometrics*, 12, 151-168
- [72] HANSEN B. (1990), "Lagrange Multiplier tests for parameter instability in non-linear models", *unpublished*, <http://www.ssc.wisc.edu/~bhansen/papers/LMtests.pdf>
- [73] HANSEN B. (1991), "A Comparison of tests for parameter instability : an examination of local power", *unpublished*, <http://www.ssc.wisc.edu/~bhansen/papers/comparison.pdf>
- [74] HANSEN B. (1992), "Testing for parameter instability in linear models", *Journal of Policy Modelling* 14(4), 517 - 533
- [75] HANSEN B. (1992), "Tests for parameter instability in regression with I(1) Processes", *Journal of Business and Economic Statistics* 10(3), 321 - 335
- [76] HANSEN B. (1997), "Approximate Asymptotical P Values for Structural - Change Tests", *Journal of Business and Economic Statistics* 15(1), 60 - 67
- [77] HANSEN B. (1998), "Testing For Structural Change in Conditionnal Models", *Journal of Econometrics* 97, 93 - 115
- [78] HANSEN B. (2001), "The New Econometrics of Structural Change : Dating Breaks in U.S Labor Productivity", *Journal of Economic Perspectives* 15(4), 117 - 128

-
- [79] HANSEN P.R. (1999), "Structural breaks in the cointegrated vector autoregressive model", Internet : [http ://fmwww.bc.edu/RePEc/es2000/1240.pdf](http://fmwww.bc.edu/RePEc/es2000/1240.pdf)
- [80] HANSEN P.R. (2000), "Structural breaks in the cointegrated vector autoregressive model", *Working Paper 00-20, Brown University*
- [81] HANSEN P.R. et S. JOHANSEN (1988), "Workbook on cointegration", Advanced texts in econometrics, Oxford University Press
- [82] HARRIS D. et B. INDER (1994), "A test of the null of cointegration", in C. Hargreaves (ed.) "Non-stationary time series analysis and cointegration", *Oxford University Press, Oxford, 133 - 152*
- [83] HENDRY D.F. et A.J. NEALE (1991), "A Monte Carlo study of the effects of structural breaks on tests for unit root", dans "Economic structural change", P. Hackland A, westlung (eds), *Springer Verlag, New York, 95-119*
- [84] HERAN F. , S. PERRIOLAT et A. PAQUES (juin 1999), "Recherche documentaire sur la monétarisation de certains effets externes des transports en milieu urbain : effets de coupure, effets sur l'affectation des espaces publics, effets sur les paysages", *contribution au PREDIT II, Groupe Thématique "Recherches Stratégiques", Axe 8 : "Monétarisation des effets externes des transports"*.
- [85] HE C. , T. TERÄSVIRTA et A. GONZALEZ (2002), "Testing parameter constancy in stationary vector autoregressive models against continuous change", *SSE/ EFI Working Paper Series in Economics and Finance, n°507*
- [86] HUANG T.J. et P. WALTERS (1994), "Transit Use in Edmonton : an econometric analysis", *Canadian Journal of Urban Research 3 :2, 134 - 148*
- [87] Institut d'Aménagement et d'Urbanisme de la Région Ile de France (IAURIF) (2000), Atlas "Paris et l'agglomération centrale dans l'espace régional", disponible sur www.iaurif.org/fr/savoirfaire/etudesenligne/paris_espace_regional/fiches.pdf.
- [88] INFOPOLIS 2 Consortium (1998-2000) Deliverables 3 : "Needs of travellers : an analysis based on the study of their tasks and activities", [http ://www.ul.ie/~infopolis/index.html](http://www.ul.ie/~infopolis/index.html)

-
- [89] JANSON J.O. (1998), "An analysis of the rail transport system" in *Report of the hundred and seventh round table on transport economics held in Paris on 26th -27th march 1998 on the following topics : "User charges for railway infrastructure", European conference of ministers of transport*, 135-181.
- [90] JOHANSEN S. J. (1988), "Statistical analysis of cointegration vectors", *Journal of Economic Dynamics and Control*, 12, 121-254
- [91] JOHANSEN S. J. (1991), "Estimation and Hypothesis testing of cointegration vectors in gaussian vector autoregressive models", *Econometrica*, 59, 1551-1580
- [92] JOHANSEN S. J. et K. Juselius (1990), "Maximum likelihood estimation and inference on cointegration with applications to the demand for money", *Oxford Bulletin of Economics and Statistics*, 52, 169-210
- [93] JOHANSEN S. , R. MOSCONI et B. NIELSEN (2000), "Cointegration analysis in the presence of structural breaks in the deterministic trend", ou Internet : [http ://www.blackwellpublishers.co.uk/ectj/PDF/johansen/ectj049.pdf](http://www.blackwellpublishers.co.uk/ectj/PDF/johansen/ectj049.pdf)
- [94] KAPETANIOS G. (2002), "Unit root testing against the alternative hypothesis of up to m Structural Breaks", *Queen Mary University of London, Working Paper n° 469, ISSN 1473-0278*
- [95] KAUFMANN V. , C. JEMELIN et J.-M. GUIDEZ (2001), "Automobile et modes de vie urbains : quel degré de liberté?", *Documentation Française, recherche réalisée dans le cadre du PREDIT*.
- [96] KEELER T. E. (1983), "Railroads, freight and public policy", *Brookings Institution, Washington D. C.*
- [97] KIM J.Y. (2000), "Detection of change in persistence of a linear time series", *Journal of Econometrics* 95, 97 - 116
- [98] KIM J.-Y. et R.B. AMADOR (2002), "Corrigendum to" Detection of change in persistence of a linear time series"[J. Econom. 95 (2000)]", *Journal of Econometrics* 109 (2), 389 - 392

-
- [99] KIM T. H. , LEYBOURNE S. J. et NEWBOLD P. (2000), "Spurious rejection by Perron tests in the presence of a break", *Oxford Bulletin of Economics and Statistics* 62, 433 - 444
- [100] KREMERS J.J. , N.R. ERICSSON et J.J. DOLADO (1992), "The Power of co-integration tests", *Oxford Bulletin of Economics and Statistics*, 54, 325 - 348
- [101] KUO B. S. (1998), " Test for partial instability in regressions with I(1) processes", *Journal of Econometrics*, 86, 337-368
- [102] KWIATKOWSKI A. , P.C.B. PHILLIPS , P. SCHMIDT et Y. SHIN (1992), "Testing the null hypothesis of stationarity against the alternative of a unit Root : how sure are we that economic time series have a unit root ?", *Journal of Econometrics* 54, 159 - 178
- [103] LANCASTER K.J. (1966), "A New approach to consumer theory", *Journal of Political Economy* 74, 132 - 157
- [104] LANNE M. et H. LÜTKEPOHL (2001), "Unit root tests for time series with level shifts : a comparison of different proposals", *Sonderforschungsbereich 373 / Humboldt Universität Berlin* ou Internet : <http://netec.mcc.ac.uk/WoPEc/data/Papers//wophumbsf2001-5.html>
- [105] LANNE M. , H. LÜTKEPOHL et P. SAIKKONEN (1999), "Comparison of Unit root Tests for Time Series with Level Shifts", *Sonderforschungsbereich 373 88/99, Humboldt Universitaet Berlin*, Internet : <http://netec.mcc.ac.uk/WoPEc/data/Papers//wophumbsf1999-88.html>
- [106] LANNE M. , H. LÜTKEPOHL et P. SAIKKONEN (2001), "Unit root tests in the Presence of Innovational Outliers", *Sonderforschungsbereich 373 / Humboldt Universitaet Berlin* ou Internet : <http://netec.mcc.ac.uk/WoPEc/data/Papers//wophumbsf2001-82.html>
- [107] LANNE M. , H. LÜTKEPOHL et P. SAIKKONEN (2001), "Test procedure for Unit Roots in Time Series with Level Shifts at Unknown Time", *Sonderforschungsbereich 373 39/01, Humboldt Universitaet Berlin* ou Internet : <http://netec.mcc.ac.uk/WoPEc/data/Papers//wophumbsf2001-39.html>

-
- [108] LEE J., HUANG C. J. et Y. SHIN (1997), "On stationary tests in the presence of structural breaks", *Economic Letters*, 165 - 172
- [109] LENORMAND A. (2002), "Prévision dans les modèles cointégrés avec rupture : Application à la demande de transport terrestre de marchandises et de voyageurs", *Thèse pour le doctorat es Sciences économiques, Université Paris I Panthéon-Sorbonne*
- [110] LEVASSEUR S. et V. VICAIRE (avril 2003), "Les transferts d'établissements sur un territoire : l'exemple du Val de Marne", *INSEE Ile-de-France "à la page", pour le compte du Centre d'Observation pour l'action du Val-de-Marne, convention n°36/02*
- [111] LEYBOURNE S.J. et B.P.M. McCABE (1993), "A Simple test for cointegration", *Oxford Bulletin of Economics and Statistics*, 55, 97-103
- [112] LEYBOURNE S. J. , MILLS T. C. et P. NEWBOLD (1998), "Spurious rejections by Dickey-Fuller tests in the presence of a break under the null", *Journal of Econometrics* 87, 191 - 203
- [113] LEYBOURNE S.J. et P. NEWBOLD (2000), "Behavior of Dickey-Fuller t-tests when there is a break under the alternative hypothesis", *Econometric Theory* 16, 779 - 789
- [114] LEYBOURNE S.J. , NEWBOLD P. et VOUGAS D. (1998), "Unit Root and Smooth Transition", *Journal of Time Series Analysis* 19, 83 - 93
- [115] LIEBENSTEIN H. (1966), "Allocative efficiency versus X-efficiency", *American Economic Review* 56, 392-415.
- [116] LUMSDAINE R.L. et D.H. PAPELL (1997), "Multiple trend breaks and the unit-root hypothesis", *Review of Economics and Statistics* 79, 212 - 218
- [117] LÜTKEPOHL H. , C. MÜLLER et P. SAIKKONEN (1999), "Unit root tests for time series with a structural break when the break point is known", *Discussion Paper, SFB 373, Humboldt-Universität zu Berlin* ou Internet : <http://netec.mcc.ac.uk/WoPEc/data/Papers//wophumbsf1999-33.html>
- [118] LÜTKEPOHL H. , P. SAIKKONEN et C. TRENKLER (2001), "Testing for the cointegrated rank of a VAR process with level shifts at unknown time", *63-2001, Humboldt Universität zu Berlin*

-
- [119] MADDALA G.S. et I. KIM (1988), "Unit roots, cointegration and structural change", *Cambridge University Press*.
- [120] McCABE B.P.M. , S.J . LEYBOURNE et Y. SHIN (1997), "A Parametric approach to testing for the null of cointegration", *Journal of Time Series Analysis*, 18 (4), 395-413
- [121] MACKINNON J.G. (1991), "Critical Values for Cointegration Tests", *Chapter 13 in Long-run Economic Relationships : Readings in Cointegration*, edited by R.F.Engle and C.W.J. Granger, Oxford University Press.
- [122] MERLIN P. (2001), "Densité urbaine, transport et qualité de vie", *Transport Public International*.
- [123] MONTANES A. et REYES M. (1998), "Effect of a shift in the trend function on Dickey - Fuller unit root tests", *Econometric Theory* 14, 355 - 367
- [124] MURPHY A. H. (1995), "The coefficients of correlation and determination as measures of performance in forecast verification", *Weather Forecasting* 10, 681-688.
- [125] NASH et PRESTON (1993), "The policy debate in Great Britain" *in report of the nineteenth round table on transport economics held in Paris on 4th -5th february 1993 on the following topic : "Privatisation of Railways"*, *European Conference of ministers of transport*, 85-107, 110-119.
- [126] NELSON C.R. et C.I. PLOSSER (1982), "Trends and random walks in macroeconomic time series", *Journal of Monetary Economics* 10 (2), 139 - 162
- [127] NEWEY W.K. et K.D. WEST (1987), "A simple positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix", *Econometrica*, 55, 703 - 708
- [128] NEWEY W.K. et K.D. WEST (1994), "Automatic lag selection in covariance matrix estimation", *Review of Economic Studies*, 61, 631 - 654
- [129] OSTERWALD et LENUM (1992), "A Note with fractiles of the asymptotic distribution of the maximum likelihood cointegration rank test statistics : four cases", *Oxford Bulletin of Economics and Statistics*, 54, 461-472

-
- [130] OUM T.H. et C. YU (1994), "Economic efficiency of railways and implications for public policy : a comparative study of the OECD countries' railways", *Journal of Transport Economics and Policy* 28(2), 121-138.
- [131] PARK J.Y. (1990), "Testing for cointegration through variable addition", *Studies in Econometric Theory. Fromby and Rhodes, eds., pp. 107-133. New York : JAI Press, 1990.*
- [132] PARK J.Y. et B. CHOI (1988), "A New approach for testing for a unit root", *CAE Working Paper n° 88-23, Cornell University*
- [133] PARK J.Y et SUNG J. (1994), "Testing for unit root in models with structural change", *Econometric Theory* 10, 917 - 936
- [134] PERRON P. (1989), "The great crash, the oil price shock, and the unit root hypothesis", *Econometrica* 57, 1361 - 1401
- [135] PERRON P. (1990), "Testing for a unit root in a time series with a changing mean", *Journal of Business and Economic Statistics* 8, 153 - 162
- [136] PERRON P. (1997), "Further Evidence on Breaking Trend Functions in Macroeconomic Variables", *Journal of Econometrics* 80, 355 - 385
- [137] PERRON P. et T.J. VOGELSANG (1992), "Nonstationarity and Level shifts with an application to Purchasing Power Parity", *Journal Business and Economic Statistics* 10, 301 - 320
- [138] PHILLIPS P.C.B. (1986), "Understanding spurious regressions in Econometrics", *Journal of Econometrics*, 33, 311 - 340
- [139] PHILLIPS P.C.B. (1995), "Fully modified least square and vector autoregression", *Econometrica*, 63(5), 1023- 1078
- [140] PHILLIPS P.C.B. et B.E. HANSEN (1990), "Statistical Inference in Instrumental Variables Regression with I(1) Processes", *Review of Economic Studies* 57, 99 - 125
- [141] PHILLIPS P.C.B. et S. OULIARIS (1990), "Asymptotic Properties of residual-based tests for cointegration", *Econometrica*, 58, 165-193
- [142] PHILLIPS P.C.B. et P. PERRON (1988), "Testing for a unit root in time series regression", *Biometrika*, 75, 335 - 346

-
- [143] PIGOU A. (1932), "*The Economics of welfare*", fourth edition, Macmillan, London.
- [144] Programme national de Recherche Et D'Innovation dans les Transports terrestres (2002-2007), à l'initiative des ministères chargés de la recherche, des transports, de l'environnement et de l'industrie, de l'Ademe et l'Anvar, consultable sur internet : www.predit.prd.fr
- [145] QUANDT R.E. (1960), "Tests of the hypothesis that a linear regression system obeys two separate regimes", *Journal of American Statistical Association* 55, 320 - 330
- [146] QUINTOS C.E. (1995), "Sustainability of the deficit process with structural shifts", *Journal of Business and Economic Statistics*, 13, 409-417
- [147] QUINTOS C.E. et P.C.B. PHILLIPS (1993), "Parameter constancy in cointegrating regressions", *Empirical Economics*, 18, 675-706
- [148] Rapport du Sénat (2002) pour le projet de loi de finances pour 2003, adopté par l'Assemblée Nationale. Tome XVIII : "*Transports terrestres*", Deuxième partie, "*La SNCF et RFF*", chapitre 1er, "*Les Trafics de la SNCF*", disponible sur www.senat.fr/rap/a02-070-18/a02-070-182.html
- [149] SAID E. and D.A. DICKEY (1984), "Testing for Unit Roots in Autoregressive Moving Average Models of Unknown Order", *Biometrika*, 71, 599 - 607.
- [150] SAIKKONEN P. et H. LÜTKEPOHL (1999), "Testing for Unit Roots in Time Series With Level Shifts", *Discussion Paper SBF 373, Humboldt - Universität zu Berlin*
- [151] SCHWERT G.W. (1989), "Test for unit root : a Monte Carlo investigation", *Journal of Business and Economic Statistics*, 7, 147 - 159
- [152] Secrétariat d'Etat au Tourisme (2001), " La demande touristique en espace urbain", disponible sur www.tourisme.gouv.fr/stat-conj/espaceurbain.htm
- [153] SEO B. (1998), " Tests for structural change in cointegrated systems", *Econometric Theory*, 14, 222-259
- [154] SHIN Y. (1994), "A Residual-based test of the null of cointegration against the alternative of no-cointegration", *Econometric Theory*, 10, 91 - 115
- [155] SHIN S. (2001), "Covariate Unit Root Test with a Structural Change", *mimeo, The Ohio State University*

-
- [156] Syndicat des Transports d'Ile de France, sur www.stif-idf.fr
- [157] STOCK J.H. et M.W. WATSON (1996), "Evidence and Structural Instability in Macroeconomic Time Series Relations", *Journal of Business and Economic Statistics* 14(1), 11 - 30
- [158] TYE W.B. (1990), "The theory of contestable markets. Applications to regulatory and antitrust problems in the rail industry", *Greenwood, New York*.
- [159] VARIAN H.R. (1993), "Introduction à la microéconomie", *traduit de l'anglais par Bernard Thiry, Ouvertures Economiques, série Prémises, 3ème édition, De Boeck Université*
- [160] VIARD J. (2001), "Les premières conséquences de la RTT sur le temps libre, la vie familiale et le commerce", *rapport pour le Ministère de l'Emploi et de la Solidarité*
- [161] VOGELSANG T.J. (1997), "Wald type tests for detecting breaks in the trend function of a dynamic time series", *Econometric Theory* 13, 818 - 849
- [162] VOGELSANG T.J. & P. PERRON (1998), "Additional Tests for a Unit Root Allowing for a Break in the Trend Function at an Unknown Time", *International Economic Review* 39, 1073 - 1100.
- [163] WILLIAMSON O.E. (1967), "Hierarchical control and optimum firm size", *Journal of political economy* 75(2).
- [164] WOLD H. (1938) (deuxième édition en 1954), "A study in the analysis of stationary time series", *Uppsala, Sweden : Almqvist and Wicksell*.
- [165] WHITE H. (1980), "A heteroskedasticity - consistent covariance matrix estimator and a direct test for heteroskedasticity", *Econometrica* 48, 817 - 838
- [166] ZHANG C., S. MERCIER et P. BONNEL (2001), "Définition d'une typologie de la mobilité adaptée à la modélisation de la demande", *ISIS, LET, Rapport pour la DRAST, programme PREDIT*
- [167] ZIVOT E. (1994), "Single equation conditionnal error correction model based tests for cointegration", *Discussion paper 94 - 12, Department of Economics, University of Washington, Seattle*

- [168] ZIVOT E. et ANDREWS D.W.K. (1992), "Further Evidence on the great crash, the oil price shock, and the unit root hypothesis", *Journal of Business and Economic Statistics* 10, 251 - 270

Sites internet utilisés :

- www.amtuir.org : Musée des transports urbains interurbains et ruraux
- www.cci.fr/crocis : Chambre de Commerce et d'Industrie de Paris, Centre Régional d'Observation du Commerce, de l'Industrie et des Services
- www.iaurif.org : Institut d'Aménagement et d'Urbanisme de la Région Ile de France
- www.insee.fr : Institut National de la Statistique et des Etudes Economiques
- www.lexpansion.com : site internet du magazine mensuel "L'expansion". La rubrique quotidienne, "Le 18h.com" apporte des informations journalières sur l'économie, les entreprises ou les technologies

ANNEXES

Annexe A

Tests de racine unitaire sur période entière ou sous-période éventuelle pour l'ensemble des 8 trafics

Nous avons appliqué les deux tests usuels de racine unitaire sur la période entière (janvier 1990 - décembre 2002) : un test de Dickey - Fuller augmenté avec constante, dans lequel le nombre de retards est calculé à l'aide du critère BIC et un test de Kwiatkowski, Phillips, Schmidt et Shin, pour différentes largeurs de fenêtre. Ces deux tests ont ensuite été réalisés sur deux sous-périodes éventuelles : du début de l'échantillon aux grèves de décembre 1995, et de janvier 1996 à décembre 2002.

Rappelons les valeurs critiques à 5% :

- -2.9124 pour le test ADF si l'échantillon est de taille supérieure à 100, -2.9370 pour un échantillon de taille inférieure,
- 0.4572 pour le test KPSS

Trafic d'un jour ouvrable sur le métro

Période considérée	1990-2002	1990-1995	1996-2002
TEST ADF			
ρ estimé	0.9667	0.8718	0.9701
t_{ADF}	-1.0142	-1.6481	-0.8199
TEST KPSS			
Largeur de la fenêtre	Statistique		
3	0.9187	1.5608	2.1223
6	0.5460	0.9846	1.2626
9	0.3955	0.7514	0.9166

Sur l'ensemble de la période janvier 1990 - décembre 2002, le test ADF accepte l'hypothèse de racine unitaire au seuil de 5%. Le test KPSS ne permet pas de conclure, les résultats dépendent de la largeur de fenêtre considérée.

Par contre, quand les tests sont appliqués sur chacune des sous-périodes éventuelles, les deux tests apportent des conclusions identiques : sur chaque sous-période existe une racine unitaire.

Trafic d'un jour ouvrable sur le RER

Période considérée	1990-2002	1990-1995	1996-2002
TEST ADF			
ρ estimé	0.9709	0.4690	0.9530
t_{ADF}	-1.0658	-3.2614	-1.3267
TEST KPSS			
Largeur de la fenêtre	Statistique		
3	2.4552	0.0683	2.0223
6	1.4533	0.0605	1.2057
9	1.0503	0.0619	0.8777

ANNEXE A. TESTS DE RACINE UNITAIRE SUR PÉRIODE ENTIÈRE OU
SOUS-PÉRIODE ÉVENTUELLE POUR L'ENSEMBLE DES 8 TRAFICS

Sur période entière, les tests ADF et KPSS concluent tous deux à l'existence d'une racine unitaire. L'analyse par sous-période indique que la série de trafic d'un jour ouvrable sur le RER serait stationnaire sur la première sous-période et intégrée ensuite. Il semble y avoir eu changement de degré d'intégration.

Trafic d'un jour ouvrable sur le réseau des bus de Paris

Période considérée	1990-2002	1990-1995	1996-2002
TEST ADF			
ρ estimé	0.9240	0.7741	0.7159
t_{ADF}	-1.5476	-1.8407	-3.0820
TEST KPSS			
Largeur de la fenêtre	Statistique		
3	3.1350	0.9535	0.9157
6	1.8914	0.6550	0.6074
9	1.3791	0.5208	0.4870

Le test ADF, comme le test KPSS, conclut à la présence d'une racine unitaire sur l'ensemble de la période considérée. Les résultats sont plus difficiles à analyser lors de l'étude par sous période :

- Pour la période janvier 1990 - décembre 1995, le test ADF à 5% ne permet pas de rejeter l'hypothèse de racine unitaire, malgré un paramètre auto-régressif assez faible. Le test KPSS ne rejette l'hypothèse de stationnarité que pour des largeurs de fenêtre assez grandes.
- Sur la période post grève de 1995, le test ADF rejette au seuil de 5% l'hypothèse d'une racine unitaire. Les conclusions du test KPSS sur cette seconde sous-période dépendent de la largeur de fenêtre considérée et coïncident pour une largeur de fenêtre assez grande. Néanmoins, dans la mesure où le test ADF, pourtant peu puissant, rejette la racine unitaire et où le paramètre auto-régressif est très faible, cette série semble stationnaire.

L'existence d'une modification de l'ordre d'intégration de la série ne peut être exclue.

Trafic d'un jour ouvrable sur le réseau des bus de banlieue

Période considérée	1990-2002	1990-1995	1996-2002
TEST ADF			
Nombre de retards	4	2	3
ρ estimé	1.0053	0.9591	0.9662
constante	n.s.	n.s.	n.s.
t_{ADF}	0.3540	-0.7583	-1.1529
TEST KPSS			
Largeur de la fenêtre	Statistique		
3	3.7300	1.6452	2.1472
6	2.1779	1.0022	1.2829
9	1.5579	0.7420	0.9370

Sur la période entière, comme sur les deux sous-périodes éventuelles, le test ADF ne permet pas de rejeter l'hypothèse d'une racine unitaire. Le test KPSS confirme ce résultat, pour toute largeur de fenêtre. Il ne semble pas y avoir eu de modification du degré de la série de trafic un jour ouvrable sur le réseau du bus de banlieue.

Trafic d'un samedi sur le métro

Période considérée	1990-2002	1990-1995	1996-2002
TEST ADF			
ρ estimé	0.9827	0.9552	0.9484
t_{ADF}	-0.4918	-0.2747	-1.3271
TEST KPSS			
Largeur de la fenêtre	Statistique		
3	2.2083	0.9466	2.0666
6	1.3170	0.7008	1.2493
9	0.9595	0.6166	0.9234

ANNEXE A. TESTS DE RACINE UNITAIRE SUR PÉRIODE ENTIÈRE OU
SOUS-PÉRIODE ÉVENTUELLE POUR L'ENSEMBLE DES 8 TRAFICS

Sur période entière comme sur chaque sous-période éventuelle, les tests ADF et KPSS montrent que la série de trafic d'un samedi sur le métro possède une racine unitaire.

Trafic d'un samedi sur le RER

Période considérée	1990-2002	1990-1995	1996-2002
TEST ADF			
ρ estimé	0.9813	0.8410	0.9611
t_{ADF}	-0.5821	-0.1590	-1.0062
TEST KPSS			
Largeur de la fenêtre	Statistique		
3	2.6677	0.3349	2.1109
6	1.5776	0.2293	1.2544
9	1.1406	0.1874	0.9123

Quelle que soit la période, ou sous-période, considérée, le test ADF au seuil de 5% ne permet jamais de refuser l'hypothèse nulle de racine unitaire pour le trafic d'un samedi sur le RER. Le test KPSS confirme l'existence de cette racine unitaire sur la période entière ainsi que la deuxième sous-période, mais accepte l'hypothèse de stationnarité sur la première sous-période, caractérisée par un paramètre auto-régressif assez faible.

Il semble y avoir eu un changement de l'ordre d'intégration de la série au cours de la période d'étude.

Trafic d'un jour dimanche/jour férié sur le métro Sur l'ensemble de la période janvier 1990 - décembre 2002, les deux tests concluent à l'existence d'une racine unitaire. Par contre, le trafic d'un dimanche/jour férié sur le métro semble avoir changé d'ordre d'intégration au cours du temps : il serait stationnaire avant les grèves, et posséderait une racine unitaire après.

Période considérée	1990-2002	1990-1995	1996-2002
TEST ADF			
ρ estimé	1.0338	0.4748	1.0017
t_{ADF}	1.2347	-3.2144	0.0557
TEST KPSS			
Largeur de la fenêtre	Statistique		
3	2.6442	0.4706	2.1683
6	1.5633	0.3970	1.2956
9	1.1272	0.3922	0.9454

Trafic d'un jour dimanche/jour férié sur le RER

Période considérée	1990-2002	1990-1995	1996-2002
TEST ADF			
ρ estimé	0.9965	0.7419	0.9823
t_{ADF}	-0.1362	-2.5520	-0.4963
TEST KPSS			
Largeur de la fenêtre	Statistique		
3	3.0564	0.6746	2.1534
6	1.8051	0.4610	1.2804
9	1.3001	0.3718	0.9284

La série de trafic sur le RER un dimanche/jour férié possède une racine unitaire sur l'ensemble de la période considérée. D'après l'analyse par sous-période, elle serait stationnaire puis intégrée. Les tests ADF et KPSS apportent les mêmes conclusions.

Annexe B

Propriétés des tests de changement de structure dans un modèle auto-régressif stationnaire de taille finie

L'analyse de la littérature existante a montré que l'analyse des propriétés des tests de changement de structure dans un cadre auto-régressif avait été peu étudiée en économétrie. Le but de cette annexe est d'analyser les propriétés statistiques du test supérieur de Wald dans le cas stationnaire, notamment sa capacité à repérer l'existence d'une rupture et la qualité de l'estimation de la date de rupture, pour une taille d'échantillon finie. Nous avons ensuite cherché un raffinement en supposant que le modélisateur possédait une intuition sur le signe de la différence entre les paramètres auto-régressifs des premières et secondes sous-périodes éventuelles. Ceci nous a conduit à construire des "tests extrêmes de Student".

Pour connaître le comportement des tests de changement de structure (avec une unique rupture ayant lieu à une date inconnue) dans le cas stationnaire, nous avons réalisé des simulations dans le cas de modèles univariés auto-régressifs sans constante, pour des échantillons de taille finie (150 observations, taille des échantillons que nous aurons à traiter empiriquement).

B.1 Modèle et estimation

Le modèle initial est posé sous la forme

$$\begin{aligned} y_t &= y'_{t-1}\rho_1 + u_t, & t = 1, \dots, k_0 \\ y_t &= y'_{t-1}\rho_2 + u_t, & t = k_0 + 1, \dots, T \end{aligned}$$

avec y_t la variable dépendante, et u_t une perturbation inobservable. Dans les simulations, nous supposons que $u_t \sim i.i.\mathcal{N}(0, 1)$. Les paramètres ρ_1 et ρ_2 sont inconnus (avec $\rho_1 \neq \rho_2$), ainsi que la véritable date de rupture k_0 .

Le modèle se réécrit sous la forme

$$\begin{aligned} y_t &= y'_{t-1}\rho_1 + u_t, & t = 1, \dots, k_0 \\ y_t &= y'_{t-1}\rho_1 + y'_{t-1}(\rho_2 - \rho_1) + u_t = y'_{t-1}\rho_1 + y'_{t-1}d\rho + u_t, & t = k_0 + 1, \dots, T \end{aligned}$$

soit, sous forme matricielle :

$$Y = Y_{-1}\rho_1 + (Y_{k_0})_{-1}d\rho + U$$

$$\text{où } (Y_{k_0})_{-1} = \begin{pmatrix} 0 & \dots & 0 & y_{k_0} & \dots & y_{T-1} \end{pmatrix}'.$$

Cette équation est estimée au moyen des moindres carrés ordinaires.

Notons :

- $S_T(k)$ la somme des carrés résiduels de Y sur Y_{-1} et $(Y_k)_{-1}$, c'est-à-dire la somme des carrés résiduels issue du modèle avec rupture à la date k ,
- et $S_T(0)$ la somme des carrés résiduels de Y sur Y_{-1} , c'est-à-dire la somme des carrés résiduels issue du modèle sans rupture.

B.2 Les tests utilisés

Rappelons les différents tests (tests de Wald et de Student quand la date de rupture éventuelle est inconnue) dont on désire connaître les propriétés, en particulier les hypothèses de test, les

ANNEXE B. PROPRIÉTÉS DES TESTS DE CHANGEMENT DE STRUCTURE DANS
UN MODÈLE AUTO-RÉGRESSIF STATIONNAIRE DE TAILLE FINIE

statistiques utilisées et leur éventuelle distribution asymptotique.

Test supérieur de Wald : Ce test permet d'établir si une rupture existe et, si tel est le cas, la date de cette rupture.

Hypothèses de test et statistique de test Les hypothèses de test sont les suivantes :

- Hypothèse nulle : "Absence de rupture",
- Hypothèse alternative : "Existence d'une rupture à une date inconnue".

Proposition 1 *Le point de rupture estimé est défini par*

$$\hat{k} = \arg \min_{1 \leq k \leq T} S_T(k)$$

ce qui correspond à l'application d'un test supérieur de Wald, qui a pour statistique de test :

$$\sup_{k \in [\pi T, (1-\pi)T]} Wald_T(k), \pi \in \left(0, \frac{1}{2}\right)$$

$$\text{où } Wald_T(k) = \frac{d\hat{\rho}' \left((Y_k)'_{-1} M (Y_k)_{-1} \right) d\hat{\rho}}{\hat{\sigma}^2(k)}$$

$$\text{avec } M = I - Y_{-1} (Y_{-1}' Y_{-1})^{-1} Y_{-1}' \text{ et } \hat{\sigma}^2(k) = \frac{S_T(k)}{T-1}$$

Distribution asymptotique de la statistique de test Selon Andrews (1993 [6]), on dispose du théorème suivant, pour $\pi = \frac{k}{T}$:

Theorem 2 *Sous les hypothèses précédentes,*

$$Wald_T(\cdot) \Rightarrow Q(\cdot)$$

$$\sup_{\pi \in \Pi} Wald_T(\cdot) \rightarrow \sup_{\pi \in \Pi} Q(\cdot)$$

$$\text{avec } Q(\cdot) = \frac{(B(\pi) - \pi B(1))' (B(\pi) - \pi B(1))}{\pi(1-\pi)}$$

et B est un mouvement brownien standard sur $\Pi \subset [0, 1]$.

Le test supérieur de Wald exposé précédemment a pour hypothèse alternative l'existence d'une rupture, sans présupposer sur quelle sous-période le paramètre auto-régressif est le plus fort. Si le modélisateur possède une intuition sur le signe de la différence entre le paramètre de la seconde période et celui de la première période, mais qu'il ne connaît pas la date de rupture éventuelle, il peut choisir d'appliquer un test équivalent construit à partir de tests de Student.

Tests "extrêmes" de Student Pour une simulation précise, pour chaque date de rupture k possible, la statistique de Student est définie par :

$$Student_T(k) = \frac{d\hat{\rho}_k}{\hat{\sigma}(\hat{\rho}_k)}$$

Les tests de Wald sont des tests bilatéraux qui maximisent une norme (donc les statistiques de test sont toujours positives). Au contraire, le test de Student est un test unilatéral dans lequel le signe de la statistique est important.

L'hypothèse nulle est identique aux tests précédents : $\{d\rho = 0\}$, mais, au lieu de l'hypothèse alternative l'hypothèse bilatérale $\{Ha : d\rho \neq 0\}$, considérons

- soit $\{Ha : d\rho > 0\}$, la date de rupture étant inconnue,

Dans ce premier cas, la statistique de test correspond à la plus grande valeur positive parmi les statistiques de Student associées aux dates de rupture k possibles. Nous parlerons par la suite de "test de Student unilatéral droit".

- soit $\{Ha : d\rho < 0\}$, la date de changement de structurant étant inconnue,

Dans ce cas, seules les statistiques de Student négatives sont prises en considération, et la statistique de test est égale à la statistique de Student qui possède la plus grande valeur absolue, c'est-à-dire la plus petite statistique de Student. Ce test est appelé "test de Student unilatéral gauche".

Test unilatéral droit Si l'hypothèse alternative est $\{Ha : d\rho > 0\}$, la statistique de test est

$$\sup_{k \in [\pi T, (1-\pi)T]} Student_T(k) \text{ avec } Student_T(k) = \frac{d\hat{\rho}_k}{\hat{\sigma}(\hat{\rho}_k)}$$

Test unilatéral gauche Si l'hypothèse alternative est $\{Ha : d\rho < 0\}$, la statistique de test est

$$\inf_{k \in [\pi T, (1-\pi)T]} Student_T(k) \text{ avec } Student_T(k) = \frac{d\hat{\rho}_k}{\hat{\sigma}(\hat{\rho}_k)}$$

B.3 Valeurs critiques

Pour appliquer ces différents tests dans des échantillons de taille finie, nous avons construit les tables de valeurs critiques en appliquant la méthode de Monte-Carlo (10 000 simulations) à des échantillons de taille finie : $T = 150$ (un peu plus de 12 années, taille empirique des échantillons traités). Le processus générateur de données est un processus auto-régressif d'ordre 1. Sous l'hypothèse nulle, le paramètre auto-régressif est stable et posé à $\rho = 0.8$. Par ailleurs, nous avons enlevé 20% des observations au début et 20% à la fin, c'est-à-dire que le balayage a lieu sur l'intervalle 31 – 120.

Les simulations ont été réalisées au moyen du logiciel GAUSS, et, pour avoir pour chaque type de test le même ensemble de trajectoires simulées, nous avons initialisé la congruence.

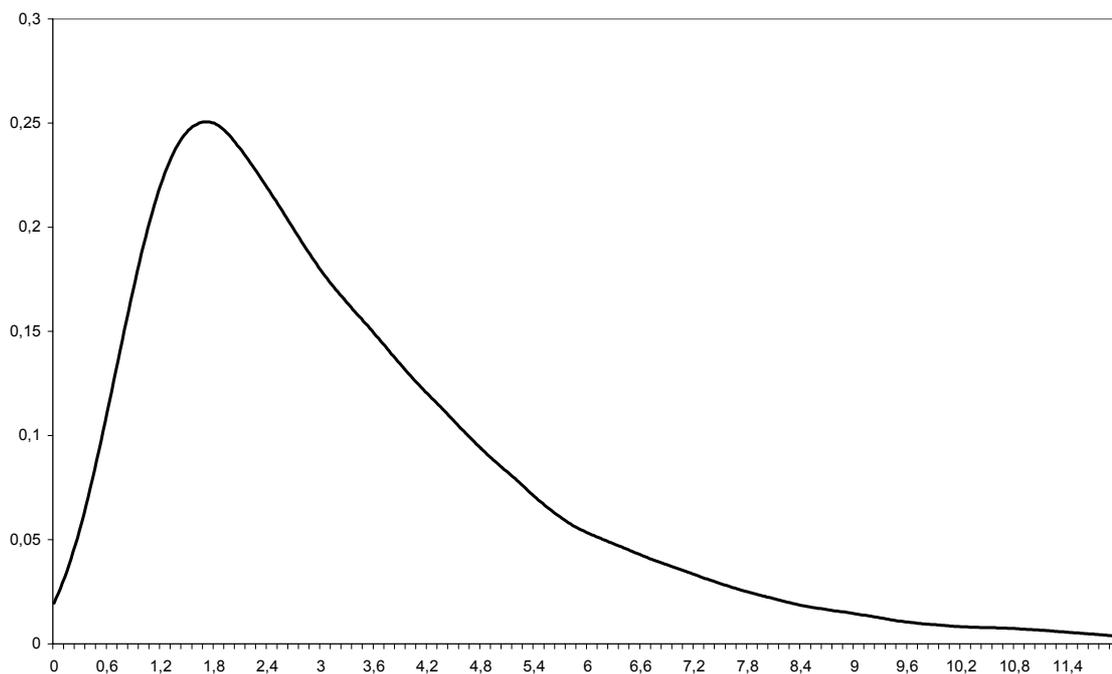
Le tableau des valeurs critiques dans un cadre stationnaire est le suivant :

$\rho < 1$	Valeurs critiques		
Niveau	Wald	Student droit	Student gauche
90%	6.4416	2.0983	-2.2977
91%	6.6674	2.1436	-2.3557
92%	6.9240	2.2024	-2.4180
93%	7.2409	2.2588	-2.4910
94%	7.5568	2.3228	-2.5501
95%	7.9205	2.4115	-2.6249
96%	8.4670	2.4988	-2.7218
97%	9.0774	2.6224	-2.8328
98%	10.0610	2.7697	-2.9800
99%	11.3474	3.0408	-3.2182

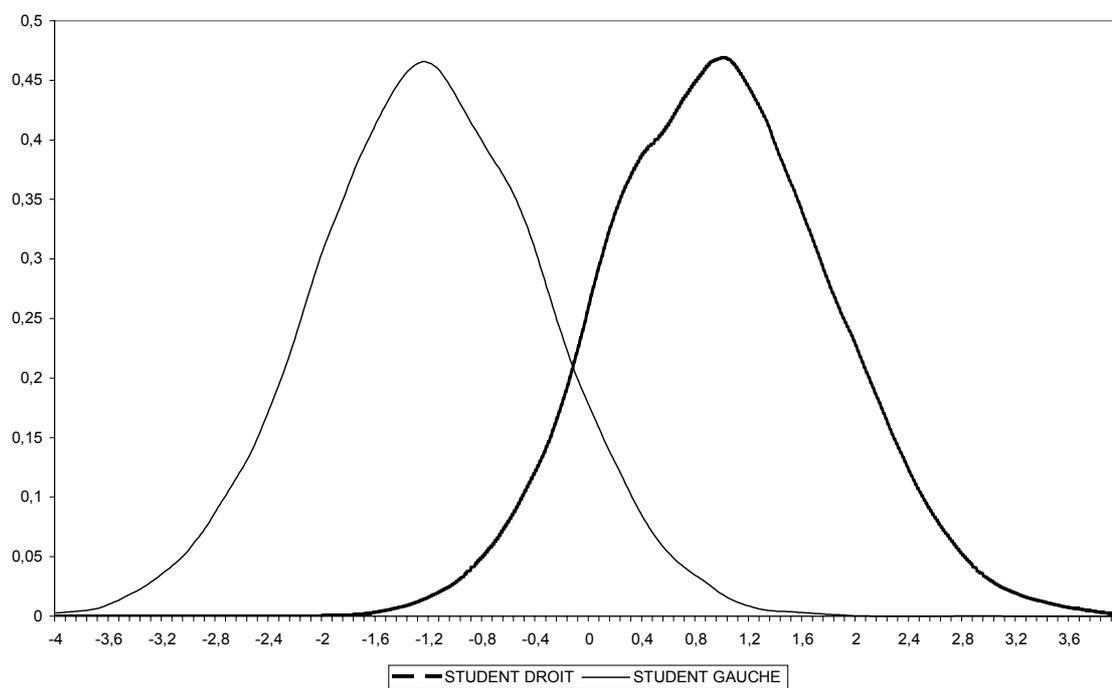
Les valeurs critiques que nous avons obtenues pour le test supérieur de Wald sont comparables à celles d'Andrews (1993 [6]) pour un modèle avec une seule variable soumise à rupture, cette rupture ne pouvant avoir lieu dans les 20% de données qui se situent au début de l'échantillon ni dans les 20% de données qui se situent à la fin de l'échantillon.

En outre, les valeurs critiques obtenues sont indépendantes de la valeur du paramètre auto-régressif.

ANNEXE B. PROPRIÉTÉS DES TESTS DE CHANGEMENT DE STRUCTURE DANS
UN MODÈLE AUTO-RÉGRESSIF STATIONNAIRE DE TAILLE FINIE



Densité de probabilité du test supérieur de Wald dans un cadre auto-régressif stationnaire.



Densité de probabilité des tests extrêmes de Student dans un cadre auto-régressif stationnaire.

B.4 Puissance des différents tests et estimation de la date de rupture par rapport à la vraie date

Evaluons la capacité de ces tests à détecter que l'hypothèse alternative (ici, l'existence d'une rupture) est vraie, en envisageant les deux cas suivants :

- $\{\rho_1 = 0.6 \text{ et } \rho_2 = 0.8\}$
- et $\{\rho_1 = 0.8 \text{ et } \rho_2 = 0.6\}$.

Lorsqu'un test rejette l'hypothèse d'absence de rupture, la date de rupture est importante. Pour évaluer cette qualité des estimations des dates de rupture si l'absence de rupture a été rejetée, nous avons représenté la différence entre la date de rupture estimée par le test et la véritable date de rupture. Quand cette différence est positive, le modèle s'aperçoit avec "retard" du changement de régime ; quand cette différence est négative, le modèle "anticipe" la rupture, c'est-à-dire considère trop tôt que le régime a changé.

10 000 simulations de Monte-Carlo sur des échantillons de taille 150 ont été réalisées. Pour chaque simulation, la véritable date de rupture a été tirée au hasard. La date de rupture est estimée en supposant qu'elle n'a lieu ni dans les 20% des observations situées au début de l'échantillon, ni dans les 20% des dernières observations.

B.4.0.5 Puissance des tests et estimations des dates de rupture dans le cas $\{\rho_1 = 0.6, \text{ et } \rho_2 = 0.8\}$

Test supérieur de Wald Dans le cas stationnaire avec rupture $\{\rho_1 = 0.6 \text{ et } \rho_2 = 0.8\}$, la puissance du test supérieur de Wald pour un niveau de confiance de 5% est égale à **20%**. Autrement dit, le test supérieur de Wald ne repère l'existence de la rupture que dans un cas sur cinq. En cas de rejet de l'hypothèse d'absence de rupture, l'estimation des dates de rupture est de qualité médiocre : si le véritable modèle est généré avec $\{\rho_1 = 0.6 \text{ et } \rho_2 = 0.8\}$, la probabilité d'estimer la bonne date de rupture est inférieure à 5%. En outre, l'erreur systématique est assez importante : le test supérieur de Wald anticipe en moyenne la date de rupture de 1 à 2 mois.

ANNEXE B. PROPRIÉTÉS DES TESTS DE CHANGEMENT DE STRUCTURE DANS
UN MODÈLE AUTO-RÉGRESSIF STATIONNAIRE DE TAILLE FINIE

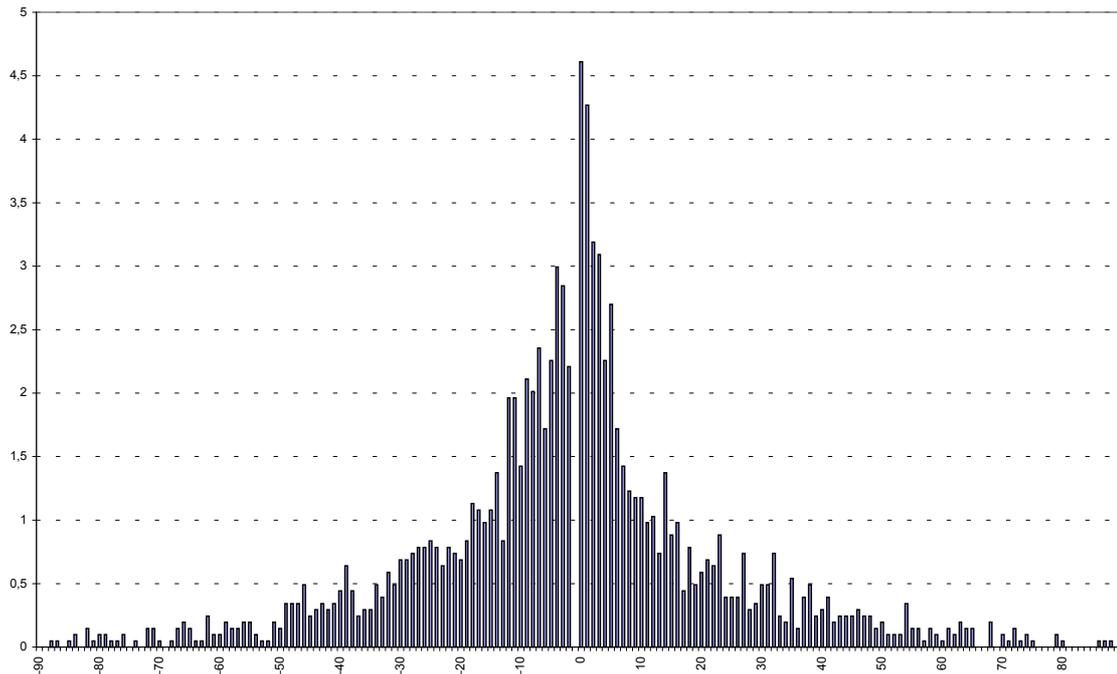
Test extrême de Student unilatéral droit La puissance du test de Student unilatéral droit est égale à près de **35%** pour un seuil de confiance de 5% quand l'hypothèse alternative correspond à $\{d\rho > 0\}$ et que le véritable modèle est généré avec $\{\rho_1 = 0.6 \text{ et } \rho_2 = 0.8\}$. La bonne date de rupture n'est estimée qu'une fois sur 20 et il faut prendre un intervalle de confiance égal à deux ans (plus ou moins une année autour de la date de rupture estimée) pour avoir une chance sur deux que la véritable date de rupture se situe dans cet intervalle. En outre, le test de Student droit a tendance à anticiper la véritable date de rupture d'un mois environ.

De façon plus précise, les résultats de puissance sont les suivants :

$\rho_1 = 0.6, \rho_2 = 0.8$	Wald		Student droit	
Puissance à 5%	20%		35%	
Bonne date estimée	4,61%		4.74%	
Erreur d'estimation de la date de rupture par rapport à la vraie date				
	Anticipation	Retard	Anticipation	Retard
\leq à 6 mois	12%	17.2%	11.6%	16.5%
\leq à 9 mois	18.5%	21.1%	17.7%	20.8%
\leq à 12 mois	23.8%	24.2%	23.1%	24.1%
Intervalles de confiance centrés autour de la date de rupture estimée				
Largeur de l'intervalle	Seuil de confiance		Seuil de confiance	
6 mois	20.2%		20.2%	
12 mois	33.9%		32.8%	
18 mois	44.2%		43.2%	
24 mois	52.7%		52%	

Les histogrammes des différences entre la date de rupture estimée et la véritable date de rupture sont :

B.4. PUISSANCE DES DIFFÉRENTS TESTS ET ESTIMATION DE LA DATE DE RUPTURE PAR RAPPORT À LA VRAIE DATE



Différence entre la date de rupture estimée et la véritable date de rupture, quand le test de Wald rejette l'hypothèse de stabilité et que le vrai modèle est tel que $\{\rho_1 = 0.6 \text{ et } \rho_2 = 0.8\}$.

Puissance des tests et estimations des dates de rupture dans le cas $\{\rho_1 = 0.8, \text{ et } \rho_2 = 0.6\}$

Test supérieur de Wald Dans le cas stationnaire avec rupture, la puissance du test supérieur de Wald pour un niveau de confiance de 5% est égale à **26%**. Si le véritable modèle comporte une rupture, le test supérieur de Wald repère l'existence de cette rupture dans plus d'un cas sur quatre, contre une fois sur cinq dans le modèle précédent. Quand le test de Wald écarte une série (i.e. qu'on ne peut conclure à la stabilité de cette série) et estime les dates de ruptures, la bonne date de rupture n'est jamais estimée, il vaut mieux prendre la date qui précède la date estimée, et qui correspond à la véritable date de rupture dans 8% des cas. Une fois sur quatre, la date de rupture estimée est située dans un intervalle de confiance de 6 mois (plus ou moins 3 mois de la véritable date de rupture), et 72% des dates estimées sont éloignées de la véritable date de rupture de deux années au maximum, en plus ou en moins. Le test a

ANNEXE B. PROPRIÉTÉS DES TESTS DE CHANGEMENT DE STRUCTURE DANS
UN MODÈLE AUTO-RÉGRESSIF STATIONNAIRE DE TAILLE FINIE

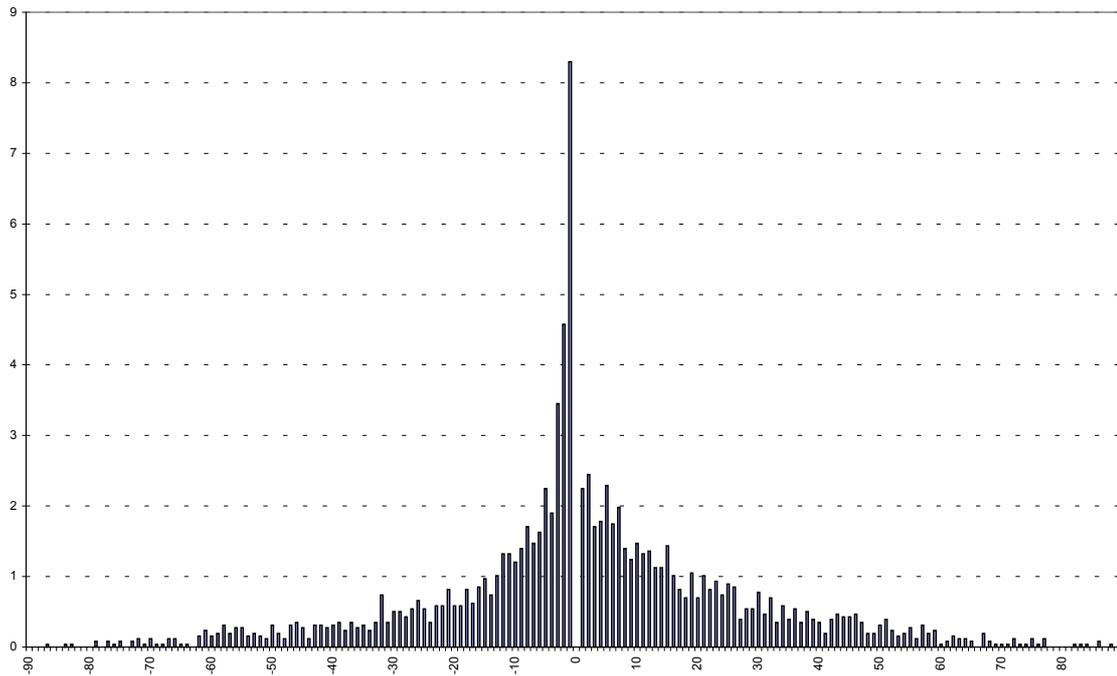
tendance à s'apercevoir avec retard du changement de régime.

Test extrême de Student unilatéral gauche Pour réaliser un test extrême de Student unilatéral gauche à $\alpha\%$, l'hypothèse nulle de stabilité du paramètre auto-régressif est rejetée lorsque la statistique de test est inférieure à la valeur critique. Quand le test a pour hypothèse alternative $\{d\rho < 0\}$, et que le vrai modèle est tel que $\{\rho_1 = 0.8 \text{ et } \rho_2 = 0.6\}$, la puissance au seuil de 5% est égale à **33%**. Les dates de rupture sont estimées de façon insatisfaisante : seule la moitié de ces dates estimées est éloignée de la véritable date de rupture de plus ou moins un an. Le test s'aperçoit tardivement du changement de régime.

Les résultats sont synthétisés dans le tableau suivant :

$\rho_1 = 0.8, \rho_2 = 0.6$	Wald		Student gauche	
Puissance à 5%	26%		33%	
Bonne date estimée	0%		0%	
Erreur d'estimation de la date de rupture par rapport à la vraie date				
	Anticipation	Retard	Anticipation	Retard
\leq à 6 mois	22.1%	12.2%	22.2%	12.3%
\leq à 9 mois	26.7%	16.8%	26.8%	17.2%
\leq à 12 mois	30.5%	21%	30.4%	21.4%
Intervalles de confiance centrés autour de la date de rupture estimée				
Largeur de l'intervalle	Seuil de confiance		Seuil de confiance	
6 mois	22.7%		22.7%	
12 mois	34.3%		34.5%	
18 mois	43.5%		44%	
24 mois	51.5%		51.8%	

B.4. PUISSANCE DES DIFFÉRENTS TESTS ET ESTIMATION DE LA DATE DE RUPTURE PAR RAPPORT À LA VRAIE DATE



Différence entre la date de rupture estimée et la véritable date de rupture, quand le test de Wald rejette l'hypothèse de stabilité et que le vrai modèle est tel que $\{\rho_1 = 0.8 \text{ et } \rho_2 = 0.6\}$.

Annexe C

Tables de puissance des tests de changement du degré d'intégration

C.1 Processus générateur de données et calculs de puissance

Processus générateur de données

2 500 simulations de Monte-Carlo ont été réalisées sur un échantillon de taille 150 en supposant un processus générateur sous l'hypothèse alternative s'écrivant sous la forme :

$$\begin{aligned}y_t &= c_1 + \rho_1 y_{t-1} + u_t \text{ si } t \leq k \\ &= c_2 + \rho_2 y_{t-1} + u_t \text{ si } t > k \\ u_t &\sim \text{bruit blanc}\end{aligned}$$

Pour chaque type de simulations sont indiquées les valeurs de la date de rupture k , des paramètres auto-régressifs sur la première sous-période (ρ_1) et la deuxième sous-période (ρ_2), des constantes relatives à chaque sous-période (c_1 et c_2).

Le niveau des tests est posé égal à 10%, la date de rupture est estimée en enlevant 10% des observations au début de l'échantillon et 10% des observations en fin d'échantillon.

Les calculs de puissance sont effectués

- pour une série "intégrée puis stationnaire",

C.2. PUISSANCE DES TESTS SUIVANT LA VALEUR DU PARAMÈTRE AUTO-RÉGRESSIF SUR LA SOUS-PÉRIODE STATIONNAIRE

– pour une série "stationnaire puis intégrée".

Lecture des tableaux de puissance

Les tableaux de puissance comportent les renseignements suivants :

- la puissance du test en %, c'est-à-dire le pourcentage de fois où notre test rejette à raison l'hypothèse d'absence de changement du degré d'intégration alors que le processus générateur de données suppose l'existence d'un tel changement du degré d'intégration,
- le pourcentage de bonnes dates estimées lorsque le test rejette l'absence de changement du degré d'intégration qui représente la probabilité que la date de rupture estimée \hat{k} soit égale à la véritable date de rupture k ,
- les intervalles de confiance à 3 et 6 mois (respectivement "IC à ± 3 mois" et "IC à ± 6 mois") autour de la date de rupture k , qui correspondent à la probabilité que la date de rupture estimée \hat{k} soit comprise entre $k - 3$ et $k + 3$ pour l'intervalle de confiance à 3 mois et $k - 6$ et $k + 6$ pour l'intervalle de confiance à 6 mois.

C.2 Puissance des tests suivant la valeur du paramètre auto-régressif sur la sous-période stationnaire

Ha : "Série intégrée puis stationnaire" :

$$\left\{ k = 75, \rho_1 = 1, c_1 = 0, c_2 = 0 \right\}$$

ANNEXE C. TABLES DE PUISSANCE DES TESTS DE CHANGEMENT DU DEGRÉ
D'INTÉGRATION

Test supérieur de Wald dans le cas intégré puis stationnaire

ρ_2	Puissance	Bonne date estimée	IC à ± 3 mois	IC à ± 6 mois
-0.9	100	93.44	97.72	99.20
-0.8	100	92.60	97.28	99.00
-0.7	100	92.24	96.84	98.68
-0.6	100	91.64	96.40	98.40
-0.5	100	90.96	96.00	98.08
-0.4	100	90.16	95.60	97.84
-0.3	100	89.20	94.84	97.28
-0.2	100	87.92	93.88	96.40
-0.1	100	86.32	92.48	95.56
0	100	84.60	91.08	94.68
0.1	100	82.72	89.40	93.32
0.2	100	80.68	87.52	91.84
0.3	100	77.32	84.56	89.36
0.4	99.96	73.87	81.63	86.47
0.5	99.56	68.58	77.90	82.80
0.6	97.24	63.31	73.96	79.19
0.7	87.68	57.03	68.75	74.64
0.8	65.80	49.18	63.53	70.21
0.9	35.52	28.60	54.73	63.51

C.2. PUISSANCE DES TESTS SUIVANT LA VALEUR DU PARAMÈTRE
AUTO-RÉGRESSIF SUR LA SOUS-PÉRIODE STATIONNAIRE

Test inférieur de Student dans le cas intégré puis stationnaire

ρ_2	Puissance	Bonne date estimée	IC à ± 3 mois	IC à ± 6 mois
-0.9	100	93.44	97.76	99.28
-0.8	100	92.60	97.32	99.12
-0.7	100	92.20	96.92	98.84
-0.6	100	91.80	96.48	98.52
-0.5	100	90.88	95.84	98.08
-0.4	100	90.28	95.68	97.88
-0.3	100	89.12	94.84	97.28
-0.2	100	88.08	94.08	96.68
-0.1	100	86.52	92.60	95.64
0	100	84.76	91.36	94.80
0.1	100	82.68	89.56	93.64
0.2	100	80.68	87.68	92.00
0.3	100	77.88	85.28	90.12
0.4	99.96	74.07	82.27	87.15
0.5	99.76	68.57	78.43	83.92
0.6	98.80	62.63	74.25	80.00
0.7	91.60	55.11	68.04	74.28
0.8	70.72	42.03	63.41	70.87
0.9	33.60	24.05	45.83	55.83

ANNEXE C. TABLES DE PUISSANCE DES TESTS DE CHANGEMENT DU DEGRÉ
D'INTÉGRATION

A l'approche de l'hypothèse nulle

Test supérieur de Wald dans le cas intégré puis stationnaire

ρ_2	Puissance	Bonne date estimée	IC à ± 3 mois	IC à ± 6 mois
0.91	32.08	26.19	53.24	61.60
0.92	28.48	23.03	49.44	58.29
0.93	25.60	20.47	46.25	55.31
0.94	22.16	17.51	41.52	51.27
0.95	19.84	14.32	36.69	46.98
0.96	16.92	9.69	30.50	39.95
0.97	14.44	6.37	24.10	33.79
0.98	12.60	3.18	16.51	25.40
0.99	12.16	1.32	11.84	16.78

Test inférieur de Student dans le cas intégré puis stationnaire

ρ_2	Puissance	Bonne date estimée	IC à ± 3 mois	IC à ± 6 mois
0.91	30.12	20.05	43.56	52.99
0.92	26.40	18.49	39.70	49.85
0.93	22.44	14.44	34.94	45.28
0.94	18.72	9.40	27.56	37.61
0.95	15.16	6.07	20.85	29.82
0.96	13.44	4.76	15.77	24.40
0.97	11.40	1.75	10.53	17.19
0.98	10.44	0.77	6.90	13.03
0.99	10.28	0.78	7.00	11.28

C.2. PUISSANCE DES TESTS SUIVANT LA VALEUR DU PARAMÈTRE
AUTO-RÉGRESSIF SUR LA SOUS-PÉRIODE STATIONNAIRE

Ha : "Série stationnaire puis intégrée" :

$$\left\{ k = 75, \rho_2 = 1, c_1 = 0, c_2 = 0 \right\}$$

Test supérieur de Wald dans le cas stationnaire puis intégré

ρ_1	Puissance	Bonne date estimée	IC à ± 3 mois	IC à ± 6 mois
-0.9	100	74.96	93.36	97.96
-0.8	100	64.64	90.68	97.24
-0.7	100	57.12	87.12	95.96
-0.6	100	50.72	83.60	94.36
-0.5	100	45.80	79.84	92.56
-0.4	100	41.08	77.76	91.16
-0.3	100	36.16	74.52	89.24
-0.2	100	32.16	72.04	87.80
-0.1	100	28.56	68.12	84.76
0	100	24.56	64.00	81.00
0.1	100	21.20	59.72	77.20
0.2	100	17.64	55.16	72.32
0.3	100	14.00	50.44	67.68
0.4	99.88	11.53	45.86	62.75
0.5	98.96	8.61	40.26	57.28
0.6	94.64	5.88	31.83	48.23
0.7	79.32	4.94	25.11	39.64
0.8	51.04	2.74	17.09	27.82
0.9	20.92	1.15	10.13	16.25

ANNEXE C. TABLES DE PUISSANCE DES TESTS DE CHANGEMENT DU DEGRÉ
D'INTÉGRATION

Test inférieur de Student dans le cas stationnaire puis intégré

ρ_1	Puissance	Bonne date estimée	IC à ± 3 mois	IC à ± 6 mois
-0.9	100	74.68	93.24	98.04
-0.8	100	65.04	90.60	97.28
-0.7	100	56.68	86.84	96.00
-0.6	100	50.52	83.72	94.40
-0.5	100	45.76	79.80	92.56
-0.4	100	41.56	77.72	91.44
-0.3	100	37.24	74.28	89.24
-0.2	100	32.60	71.60	87.40
-0.1	100	29.72	68.88	84.84
0	100	25.56	64.48	81.20
0.1	100	22.28	60.40	77.64
0.2	100	18.84	55.68	72.92
0.3	100	15.52	50.72	68.40
0.4	100	12.72	46.24	63.00
0.5	99.56	10.17	40.34	57.57
0.6	97.28	7.48	33.35	49.96
0.7	86.48	5.18	26.09	40.43
0.8	59.48	3.56	18.09	30.40
0.9	25.80	1.86	11.32	17.05

C.2. PUISSANCE DES TESTS SUIVANT LA VALEUR DU PARAMÈTRE
AUTO-RÉGRESSIF SUR LA SOUS-PÉRIODE STATIONNAIRE

A l'approche de l'hypothèse nulle

Test supérieur de Wald dans le cas stationnaire puis intégré

ρ_1	Puissance	Bonne date estimée	IC à ± 3 mois	IC à ± 6 mois
0.91	19.36	0.83	7.44	14.26
0.92	17.68	0.68	6.33	11.99
0.93	15.92	0.25	5.53	10.30
0.94	14.40	0.28	4.17	8.89
0.95	12.72	0.31	3.46	8.17
0.96	11.56	0.35	3.46	7.61
0.97	10.84	0.74	2.95	7.01
0.98	10.92	0.73	4.76	8.06
0.99	10.12	1.58	4.35	7.51

Test inférieur de Student dans le cas stationnaire puis intégré

ρ_1	Puissance	Bonne date estimée	IC à ± 3 mois	IC à ± 6 mois
0.91	22.40	0.89	9.11	16.61
0.92	20.32	0.79	7.68	14.57
0.93	18.92	0.63	7.19	14.37
0.94	17.00	0.24	5.88	12.00
0.95	15.24	0.26	4.46	10.23
0.96	14.00	0.57	4.29	9.71
0.97	13.12	0.92	4.88	9.15
0.98	11.92	1.01	5.03	10.74
0.99	11.32	0.35	6.36	10.95

**C.3 Puissance des tests suivant la valeur de la constante sur la
sous-période stationnaire**

Ha : "Série intégrée puis stationnaire"

$$\left\{ k = 75, \rho_1 = 1, \rho_2 = 0.8, c_1 = 0 \right\}$$

Test supérieur de Wald dans le cas intégré puis stationnaire

c_2	Puissance	Bonne date estimée	IC à ± 3 mois	IC à ± 6 mois
-2	79.00	62.79	81.67	84.35
-1	69.52	49.94	72.15	76.52
-0.5	68.48	44.16	67.47	72.96
0	66.12	45.25	67.94	73.26
0.5	67.28	46.91	68.96	74.08
1	70.60	51.90	72.97	77.96
2	78.52	63.88	81.46	84.56

Test inférieur de Student dans le cas intégré puis stationnaire

c_2	Puissance	Bonne date estimée	IC à ± 3 mois	IC à ± 6 mois
-2	84.08	60.75	80.21	83.78
-1	75.96	46.76	69.14	74.25
-0.5	74.12	41.82	65.08	70.75
0	70.72	42.03	65.33	70.64
0.5	72.68	44.47	67.03	72.59
1	76.12	49.50	71.05	76.56
2	83.76	61.51	79.42	82.86

C.3. PUISSANCE DES TESTS SUIVANT LA VALEUR DE LA CONSTANTE SUR LA SOUS-PÉRIODE STATIONNAIRE

Ha : "Série stationnaire puis intégrée"

$$\left\{ k = 75, \rho_1 = 0.8, \rho_2 = 1, c_2 = 0 \right\}$$

Test supérieur de Wald dans le cas stationnaire puis intégré

c_1	Puissance	Bonne date estimée	IC à ± 3 mois	IC à ± 6 mois
-2	70.80	3.79	19.43	29.77
-1	55.48	3.24	17.59	28.19
-0.5	50.56	3.40	17.17	27.21
0	48.68	2.71	16.60	28.35
0.5	50.52	3.01	16.15	25.97
1	54.92	2.84	16.24	26.00
2	72.40	3.20	17.96	28.51

Test inférieur de Student dans le cas stationnaire puis intégré

c_1	Puissance	Bonne date estimée	IC à ± 3 mois	IC à ± 6 mois
-2	78.12	2.92	21.61	33.69
-1	62.60	2.75	20.19	32.46
-0.5	58.24	2.68	19.51	31.46
0	57.20	3.71	17.55	29.86
0.5	59.64	2.48	19.05	30.92
1	63.56	2.90	19.76	31.34
2	78.76	3.25	20.37	32.50

C.4 Puissance des tests suivant la véritable date de rupture

Ha : "Série intégrée puis stationnaire"

$$\left\{ \rho_1 = 1, \rho_2 = 0.8, c_1 = 0, c_2 = 0 \right\}$$

Test supérieur de Wald dans le cas intégré puis stationnaire

<i>k</i>	Puissance	Bonne date estimée	IC à ± 3 mois	IC à ± 6 mois
25	83.72	34.07	64.50	76.16
50	79.36	41.78	64.92	71.17
75	70.60	51.90	72.97	77.96
100	59.40	60.81	80.40	83.30
125	51.32	72.10	87.22	90.26

Test inférieur de Student dans le cas intégré puis stationnaire

<i>k</i>	Puissance	Bonne date estimée	IC à ± 3 mois	IC à ± 6 mois
25	88.08	33.11	63.17	75.25
50	84.64	39.08	62.62	69.19
75	76.12	49.50	71.05	76.56
100	61.48	58.36	78.01	81.72
125	45.40	69.60	84.58	88.72

Ha : "Série stationnaire puis intégrée"

$$\left\{ \rho_1 = 0.8, \rho_2 = 1, c_1 = 0, c_2 = 0 \right\}$$

Test supérieur de Wald dans le cas stationnaire puis intégré

k	Puissance	Bonne date estimée	IC à ± 3 mois	IC à ± 6 mois
25	17.48	2.75	14.87	26.31
50	33.92	1.18	13.44	21.46
75	54.92	2.84	16.24	26.00
100	74.12	3.78	19.43	30.01
125	86.12	4.65	27.31	44.54

Test inférieur de Student dans le cas stationnaire puis intégré

k	Puissance	Bonne date estimée	IC à ± 3 mois	IC à ± 6 mois
25	21.84	2.02	17.58	28.94
50	41.00	2.73	15.90	27.12
75	63.56	2.90	19.76	31.34
100	81.32	3.64	20.95	33.79
125	90.28	4.43	29.91	49.23

Annexe D

Application des tests de Perron et de nos tests de modification du degré d'intégration

Selon Perron, la plupart des séries macro-économiques peuvent être modélisées sous la forme de fluctuations stationnaires autour d'une tendance déterministe dès lors que l'on autorise la présence d'une rupture dans le niveau ou dans la pente de la tendance. Deux tests sont construits :

- un test à date de rupture connue (test de 1989 [134])
- un test à date de rupture inconnue (test de 1997 [136]).

Le choix du nombre de retards servant à blanchir le bruit est présenté dans le troisième paragraphe.

D.1 Test pour une date de rupture k connue

D.1.1 Hypothèses de test

- Hypothèse nulle : la série temporelle possède une racine unitaire et une dérive pouvant ne pas être nulle.
- Hypothèse alternative : le modèle est stationnaire en tendance.

Sous l'hypothèse nulle comme sous l'hypothèse alternative, il existe un changement dans le niveau ou dans la pente de la tendance déterministe. Les tests de racine unitaire standards ayant pour hypothèse alternative la "stationnarité en tendance" peuvent conclure au non-rejet de l'hypothèse de racine unitaire si le véritable processus générateur de données est stationnaire autour d'une tendance soumise à une rupture.

D.1.2 Modèles pris en considération

Sous l'hypothèse nulle de racine unitaire, avec

$$\begin{aligned} A(L) e_t &= B(L) v_t \\ v_t &\sim i.i.d. (0, \sigma^2) \\ \text{i.e. } e_t &\sim ARMA(p, q) \end{aligned}$$

trois types de modèles peuvent être pris en compte :

- Modèle A : changement exogène dans le niveau de la série ("crash model") :

$$y_t = c + y_{t-1} + c_k \cdot [\mathbf{1}]_{t=k+1} + e_t$$

- Modèle B : changement exogène dans le taux de croissance

$$y_t = c_1 + y_{t-1} + (c_2 - c_1) \cdot [\mathbf{1}]_{t \geq k+1} + e_t$$

- Modèle C : changement exogène du niveau et du taux de croissance de la série

$$y_t = c_1 + y_{t-1} + c_k \cdot [\mathbf{1}]_{t=k+1} + (c_2 - c_1) \cdot [\mathbf{1}]_{t \geq k+1} + e_t$$

Les alternatives respectives de chacun de ces modèles sont :

- Modèle A' : modèle stationnaire en tendance avec changement de niveau

$$y_t = c_1 + \beta t + (c_2 - c_1) \cdot [\mathbf{1}]_{t \geq k+1} + e_t$$

- Modèle B' : modèle stationnaire en tendance avec changement du taux de croissance sans changement de niveau

$$y_t = c + \beta_1 t + (\beta_2 - \beta_1) \cdot (t - k) + e_t$$

- Modèle C' : modèle stationnaire en tendance avec changement du taux de croissance et du niveau

$$y_t = c_1 + \beta_1 t + (c_2 - c_1) \cdot [\mathbf{1}]_{t \geq k+1} + (\beta_2 - \beta_1) \cdot [\mathbf{t}]_{t \geq k+1} + e_t$$

Perron étend la stratégie de Dickey-Fuller de façon à assurer une procédure de test convergente malgré la rupture de la tendance. Soient $\{\hat{e}_t^i\}$, $i = A, B, C$ les résidus de la régression de y_t suivant le modèle A', B', ou C'. Soit $\hat{\rho}^i$ l'estimateur des moindres carrés ordinaires issus de la régression

$$\hat{e}_t^i = \hat{\rho}^i \hat{e}_{t-1}^i + u_t \quad i = A, B, C \quad t = 1, \dots, T$$

Perron dérive les distributions asymptotiques du biais normalisé $T(\hat{\rho}^i - 1)$ et de la t -statistique $t_{\hat{\rho}^i}$ sous l'hypothèse de racine unitaire. La distribution asymptotique dépend de la place de la rupture dans l'échantillon et de fonctionnelles de mouvements browniens. Si la rupture a lieu à la première date de l'échantillon ou à la dernière date, les valeurs tabulées par Dickey-Fuller sont retrouvées.

Pour prendre en compte une éventuelle corrélation temporelle, il est possible d'ajouter des variables en différence retardées comme dans le test ADF

$$\hat{e}_t^i = \hat{\rho}^i \hat{e}_{t-1}^i + \sum_{j=1}^p \gamma_j \Delta^i \hat{e}_{t-j}^i + u_t, \quad i = A, B, C, \quad t = 1, \dots, T.$$

Par la suite, pour un nombre de variables retardées égal à p , et pour une date k , on note

$$t_{\hat{\rho}}(i, k, p)$$

la t – *statistique* associée à $\rho = 1$ pour le modèle $i = A, B, C$

$$t_{\Delta c}(A, k, p)$$

la t – *statistique* associée à $\Delta c = (c_2 - c_1)$ pour le modèle A , et

$$t_{\rho, \Delta \beta}(i, k, p), i = B, C$$

la t – *statistique* associée au changement de la pente $\Delta \beta = (\beta_2 - \beta_1)$ pour les modèles B et C.

D.2 Test pour une date de rupture inconnue

En 1997, Perron [136] reprend son idée de 1989 selon laquelle la plupart des séries macroéconomiques peuvent être modélisées sous la forme de fluctuations stationnaires autour d’une tendance déterministe dès lors que l’on autorise la présence d’une rupture dans le niveau ou dans la pente de la tendance. Mais il suppose en plus que la date de rupture est inconnue, considère différentes méthodes pour sélectionner la date de rupture et le nombre de variables en différence retardées, et écrit les distributions des différentes statistiques de test.

Pour réaliser les tests $i = A, B, C$ et estimer simultanément les dates de ruptures associées, deux méthodes peuvent être envisagées, pour p fixé :

- Première méthode. La date de rupture est choisie de façon à minimiser la t – *statistique* associée à $\rho = 1$, i.e. la statistique de test est égale à

$$t_{\rho}^*(i) = \min_{k \in (p+1, T)} t_{\hat{\rho}}(i, k, p)$$

La distribution des statistiques $t_{\rho}^*(A)$ et $t_{\rho}^*(B)$ a été étudiée par Zivot et Andrews [168] en 1992, dans le cas où les dates de rupture possibles sont supposées appartenir à un sous-ensemble qui exclut les valeurs en début et en fin d’échantillon.

- Deuxième méthode (cas d’un choc négatif, ponctuel ou rémanent).
 - Dans le cas du modèle A, la date de rupture est choisie de façon à minimiser la

t – statistique associée au paramètre de changement de la constante $\Delta c = (c_2 - c_1)$

$$k^* = \arg \min_{k \in (p+1, T)} t_{\widehat{\Delta c}}(A, k, p)$$

$$t_{\rho, \Delta c}^*(A) = t_{\widehat{\rho}}(A, k^*, p)$$

- Dans le cas d'un modèle B ou C , la date de rupture correspond à la date qui minimise t – statistique associée au changement de la pente $\Delta\beta$ et la statistique est notée

$$\begin{aligned} k^* &= \arg \min_{k \in (p+1, T)} t_{\widehat{\Delta\beta}}(i, k, p), i = B, C \\ t_{\rho, \Delta\beta}^*(A) &= t_{\widehat{\rho}}(i, k^*, p), i = B, C \end{aligned}$$

En l'absence d'hypothèse sur le signe du choc, ce n'est pas le minimum qui est pris en compte, mais le maximum en valeur absolue de $t_{\widehat{\Delta c}}(A, k, p)$ ou $t_{\widehat{\Delta\beta}}(i, k, p)$, et les statistiques sont alors notées

$$t_{\rho, |\Delta c|}^*(A) \text{ et } t_{\rho, |\Delta\beta|}^*(i), i = B, C$$

D.2.1 Sélection du nombre p de variables en différences retardées

Selon Perron, les méthodes associées aux critères d'information sont trop parcimonieuses et peuvent conduire à des distortions en terme de niveau et/ou des pertes de puissance. Il propose d'utiliser à la place deux méthodes alternatives :

- Une stratégie à la p max : p^* est calculé de façon à ce que le coefficient du $p^{*ième}$ retard soit significatif alors que le $(p^* + 1)^{ième}$ ne le soit pas. p max retards sont pris en compte au départ, et le dernier retard est enlevé de l'équation chaque fois que celui-ci n'est pas significatif. Cette procédure est notée " $t - sig$ ".
- Une stratégie à la Fisher. Il s'agit de tester si des retards additionnels sont significatifs de façon jointe en utilisant un test de Fisher. Pour un retard maximal $p - \max$, le test de $p - \max$ contre $p \max - 1$ retards est effectué, et ainsi de suite. Cette procédure est notée

”*F – sig*”.

D.2.2 Valeurs critiques

Les valeurs critiques tabulées par Perron [136] dans le cas d’un échantillon de taille finie ou de taille infinie, sont résumées dans le tableau suivant. Elles correspondent à un modèle dont la date de rupture a été obtenue en minimisant la *t – statistique* associée à $\rho = 1$ (méthode 1), et où le nombre de retards a été choisi à l’aide de la procédure *t – sig* décrite précédemment.

	Taille	1%	2.5%	5%	10%
Modèle A	100	-5.70	-5.36	-5.10	-4.82
	∞	-5.41	-5.02	-4.80	-4.58
Modèle C	150	-5.28	-4.96	-4.65	-4.38
	∞	-4.91	-4.62	-4.36	-4.07

Source : Perron [136]

D.3 Application des tests de Perron aux données de la RATP

Les tests de Perron à date de rupture inconnue sont appliqués, au seuil de 5%, aux séries de trafic de la RATP. Après un bref résumé relatif à chaque ensemble de séries, les résultats détaillés des tests (A) et (C) de Perron sont indiqués sous forme de tableaux. Des tests de racine unitaire sont appliqués aux sous-périodes dégagées, afin de vérifier les conclusions des tests de Perron (présence ou absence de racine unitaire, existence d’un changement de la tendance) et mettre en valeur d’éventuelles lacunes. Ainsi, pour chaque sous-période éventuelle sont appliqués un test ADF avec constante, un test ADF avec constante et tendance (nombre de retards choisi à l’aide d’un critère BIC) ainsi qu’un test KPSS (largeur de fenêtre égale à 9).

D.3.1 Mise en garde préliminaire

La faiblesse éventuelle des coefficients ρ estimés, alors même que les tests concluent à la présence de racine unitaire, ne doit pas surprendre : ils sont liés entre autres à la faiblesse du nombre d'observations dans les séries. Des simulations de Monte-Carlo (50 000 répliques réalisées par nos soins) ont montré que dans le cas d'une marche aléatoire sans dérive de 150 observations, le ρ estimé pouvait atteindre 0.67 dans le cas d'un test ADF avec constante, et 0.56 dans le cas d'un test ADF avec constante et tendance. Pour 75 observations, l'estimation est encore plus mauvaise et le ρ estimé peut descendre jusqu'à 0.22 pour le test ADF avec constante, et ... 0.04 (!) dans le test avec constante et tendance.

En présence d'une dérive, le test ADF avec constante estime avec plus de précision le paramètre auto-régressif, tandis que le test ADF avec constante et tendance conduit à une sous-estimation de ce paramètre. Ces résultats apparaissent dans le tableau synthétique suivant :

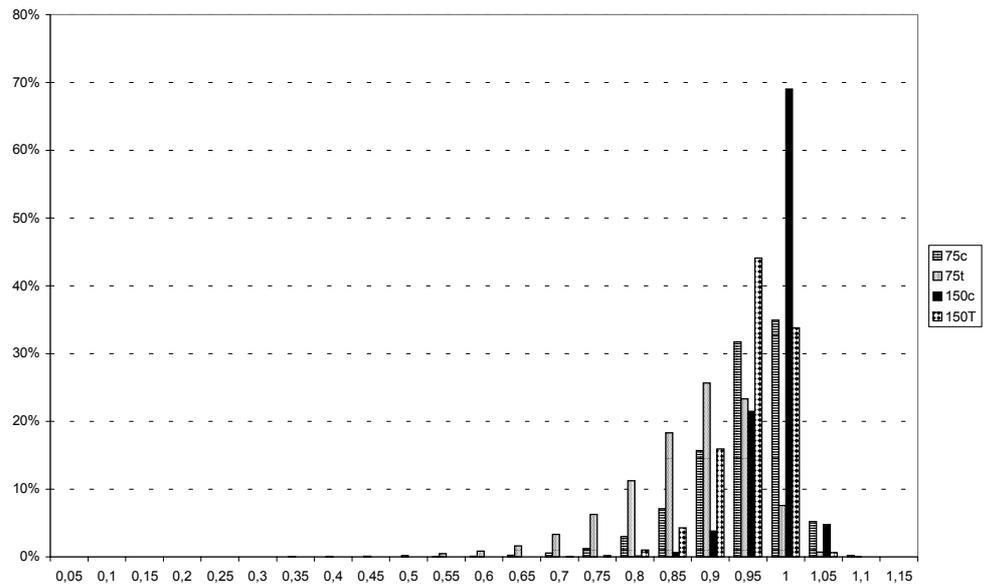
D.3. APPLICATION DES TESTS DE PERRON AUX DONNÉES DE LA RATP

ρ estimé	PGD : marche aléatoire				PGD : marche aléatoire avec dérive			
	75		150		75		150	
Nombre d'obs.	c	c, t	c	c, t	c	c, t	c	c, t
ADF	c	c, t	c	c, t	c	c, t	c	c, t
$\rho \leq 0$	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
$0 < \rho \leq 0.1$	0.000	0.002	0.000	0.000	0.000	0.000	0.000	0.000
$0.1 < \rho \leq 0.2$	0.000	0.006	0.000	0.000	0.000	0.006	0.000	0.000
$0.2 < \rho \leq 0.3$	0.004	0.024	0.000	0.000	0.000	0.020	0.000	0.000
$0.3 < \rho \leq 0.4$	0.008	0.104	0.000	0.000	0.000	0.094	0.000	0.000
$0.4 < \rho \leq 0.5$	0.034	0.294	0.000	0.000	0.000	0.324	0.000	0.000
$0.5 < \rho \leq 0.6$	0.130	1.330	0.000	0.004	0.000	1.310	0.000	0.000
$0.6 < \rho \leq 0.7$	0.788	5.022	0.004	0.062	0.000	5.062	0.000	0.052
$0.7 < \rho \leq 0.8$	4.186	17.524	0.140	1.208	0.000	17.544	0.000	1.246
$0.8 < \rho \leq 0.9$	22.770	43.958	4.506	20.216	0.000	44.062	0.000	19.986
$0.9 < \rho \leq 1$	66.660	30.924	90.674	77.900	53.758	30.766	52.842	78.086
$1 < \rho$	5.420	0.812	4.776	0.610	46.242	0.812	47.158	0.630
Minimum	0.215	0.039	0.672	0.559	0.984	0.107	0.995	0.605
Maximum	1.124	1.095	1.048	1.042	1.020	1.125	1.005	1.055
Moyenne	0.923	0.849	0.963	0.928	1.000	0.848	1.000	0.928

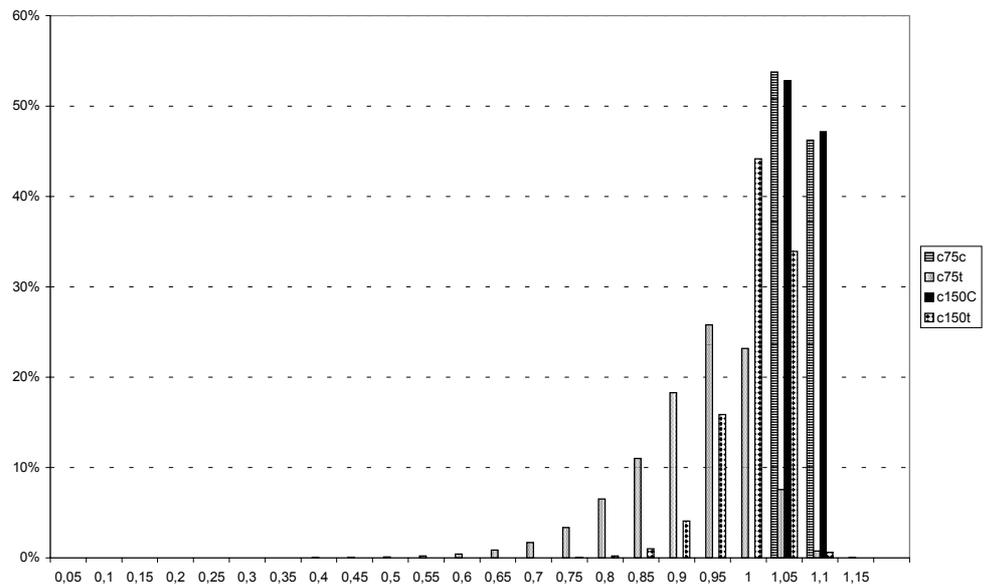
Unité : % , PGD : Processus Générateur de données

On obtient les représentations graphiques suivantes :

ANNEXE D. APPLICATION DES TESTS DE PERRON ET DE NOS TESTS DE
MODIFICATION DU DEGRÉ D'INTÉGRATION



Histogramme des ρ estimés par les tests ADF avec constante (c) ou constante et tendance (t) pour deux tailles d'échantillon : 75 et 150. Le PGD est une marche aléatoire (50 000 simulations de Monte-Carlo).



Histogramme des ρ estimés par les tests ADF avec constante (c) ou constante et tendance (t) pour deux tailles d'échantillon : 75 et 150. Le PGD est une marche aléatoire avec constante (50 000 simulations de Monte-Carlo).

D.3.2 Application aux séries de trafic

Trafic du jour ouvrable sur le métro

Test de Perron Le tableau ci-dessous résume l'application des tests de Perron à la série considérée, pour deux hypothèses alternatives différentes : la série est stationnaire autour d'une constante avec rupture (test A), et stationnaire d'une constante et d'une pente avec rupture (test C).

$\ln(jomsa)$	Rupture sur le niveau (A)	Rupture sur le niveau et la pente (C)
ρ estimé	0.8705	0.3126
Statistique	-3.0144	-9.8953
Date de rupture	juillet 1998	novembre 1995
Conclusion	I(1)-I(1)	I(0)-I(0)

Pour chacune de ces hypothèses alternatives, le test de Perron estime la date de rupture de la tendance et/ou de la constante, et indique le degré d'intégration de la série lorsque cette rupture est prise en compte dans l'estimation. Ici, la série est stationnaire avec rupture de la pente et du niveau en novembre 1995. Si la statistique est supérieure à la valeur critique définie page 598, la présence d'une racine unitaire est rejetée.

Tests standards de racine unitaire sur les sous-périodes dégagées Pour chacune de ces dates de rupture dégagées par les tests A et C de Perron, la période est scindée en deux sous-périodes, auxquelles sont appliqués les tests standards de stationnarité (test ADF avec tendance et/ou constante et test KPSS). Le tableau permet de vérifier les résultats des tests de Perron en termes de degré d'intégration.

ANNEXE D. APPLICATION DES TESTS DE PERRON ET DE NOS TESTS DE
MODIFICATION DU DEGRÉ D'INTÉGRATION

Date de rupture	$\ln(jomsa)$	juillet 1998		novembre 1995	
		Avant	Après	Avant	Après
Test ADF avec constante					
	ρ estimé	0.9539	0.8739	0.8907	0.9803
	t_{ADF}	-1.0460	-1.7531	-1.2589	-0.5977
	VC 5%	-2.9152	-2.9370	-2.9370	-2.9152
	Conclusion	I(1)	I(1)	I(1)	I(1)
Test ADF avec constante et tendance					
	ρ estimé	0.6915	0.4664	0.3028	0.6188
	t_{ADF}	-2.7922	-2.8100	-2.8355	-0.3812
	VC 5%	-3.4611	-3.4615	-3.4615	-3.4611
	Conclusion	I(1)	I(1)	I(1)	I(1)
KPSS					
	Statistique	1.0015	0.6106	0.7562	0.9115
	VC 5%	0.4572	0.4572	0.4572	0.4572
	Conclusion	I(1)	I(1)	I(1)	I(1)

Ici, il existe une racine unitaire quelle que soit la partition ou le test envisagé. La conclusion du test C de Perron (stationnarité autour d'une tendance ruptée) est invalidée.

Trafic du jour ouvrable sur le RER

Tests de Perron

$\ln(jorersa)$	Rupture sur le niveau (A)	Rupture sur le niveau et la pente (C)
ρ estimé	0.7929	0.6535
Statistique	-3.9595	-5.5610
Date de rupture	octobre 1998	juillet 1998
Conclusion	I(1)-I(1)	I(0)-I(0)

Tests standards de racine unitaire sur les sous-périodes éventuelles

Date de rupture	$\ln(jorersa)$	octobre 1998		juillet 1998	
		Avant	Après	Avant	Après
Test ADF avec constante					
	ρ estimé	0.7018	0.8996	0.6901	0.8557
	t_{ADF}	-3.0496	-1.3252	-3.2270	-2.1207
	VC 5%	-2.9152	-2.9493	-2.9152	-2.9370
	Conclusion	I(0)	I(1)	I(0)	I(1)
Test ADF avec constante et tendance					
	ρ estimé	0.6628	0.9040	0.5968	0.8636
	t_{ADF}	-3.1054	-0.7677	-3.7458	-0.1364
	VC 5%	-3.4611	-3.5449	-3.4611	-3.4615
	Conclusion	I(1) ^{n.s.}	I(1) ^{n.s.}	I(0) ^{n.s.}	I(1) ^{n.s.}
KPSS					
	Statistique	0.2919	0.4186	0.4303	0.4855
	VC 5%	0.4572	0.4572	0.4572	0.4572
	Conclusion	I(0)	I(0)	I(0)	I(1)

ANNEXE D. APPLICATION DES TESTS DE PERRON ET DE NOS TESTS DE
MODIFICATION DU DEGRÉ D'INTÉGRATION

Trafic du jour ouvrable sur le bus de Paris

Tests de Perron

ln (<i>joparsa</i>)	Rupture sur le niveau	Rupture sur le niveau
	(A)	et la pente (C)
ρ estimé	0.4967	0.4824
Statistique	-5.0502	-5.2833
Date de rupture	juillet 2000	août 1996
Conclusion	$I(0)-I(0)^{\text{à } 10\%}$	$I(0)-I(0)$

Tests standards de racine unitaire sur les sous-périodes éventuelles

Date de rupture	ln (<i>joparsa</i>)	juillet 2000		août 1996	
		Avant	Après	Avant	Après
Test ADF avec constante					
	ρ estimé	0.9486	0.2153	0.7117	0.4106
	t_{ADF}	-0.9025	-3.3897	-2.3733	-3.9620
	VC 5%	-2.9124	-2.9370	-2.9152	-2.9152
	Conclusion	I(1)	I(0)	I(1)	I(0)
Test ADF avec constante et tendance					
	ρ estimé	0.5498	-0.5002	0.5332	0.3777
	t_{ADF}	-3.5509	-5.9769	-3.2375	-3.9399
	VC 5%	-3.4542	-3.4615	-3.4611	-3.4611
	Conclusion	I(0)	I(0)	I(1)	$I(0)^{n.s.}$
KPSS					
	Statistique	1.1351	0.3316	0.4584	0.2679
	VC 5%	0.4572	0.4572	0.4572	0.4572
	Conclusion	I(1)	I(0)	I(1)	I(0)

Trafic du jour ouvrable sur le bus de banlieue

Tests de Perron

$\ln(jobansa)$	Rupture sur le niveau (A)	Rupture sur le niveau et la pente (C)
ρ estimé	0.6689	0.4468
Statistique	-4.5105	-7.6188
Date de rupture	avril 1994	novembre 1995
Conclusion	I(1)-I(1)	I(0)-I(0)

Tests standards de racine unitaire sur les sous-périodes éventuelles

Date de rupture	$\ln(jobansa)$	avril 1994		novembre 1995	
		Avant	Après	Avant	Après
Test ADF avec constante					
	ρ estimé	0.9485	0.9940	0.9448	0.9641
	t_{ADF}	-0.6006	-0.2987	-1.0190	-1.4263
	VC 5%	-2.9370	-2.9152	-2.9370	-2.9152
	Conclusion	I(1)	I(1)	I(1)	I(1)
Test ADF avec constante et tendance					
	ρ estimé	0.6820	0.6902	0.6554	0.4962
	t_{ADF}	-2.4641	-2.6996	-3.0435	-2.6961
	VC 5%	-3.4615	-3.4611	-3.4615	-3.4611
	Conclusion	I(1)	I(1)	I(1)	I(1)
KPSS					
	Statistique	0.5156	1.1210	0.7192	0.9460
	VC 5%	0.4572	0.4572	0.4572	0.4572
	Conclusion	I(1)	I(1)	I(1)	I(1)

ANNEXE D. APPLICATION DES TESTS DE PERRON ET DE NOS TESTS DE
MODIFICATION DU DEGRÉ D'INTÉGRATION

Trafic du samedi sur le métro

Tests de Perron

$\ln(smsa)$	Rupture sur le niveau (A)	Rupture sur le niveau et la pente (C)
ρ estimé	0.7998	0.3117
Statistique	-3.4168	-6.8433
Date de rupture	décembre 1993	juin 1995
Conclusion	I(1)-I(1)	I(0)-I(0)

Tests standards de racine unitaire sur les sous-périodes éventuelles

Date de rupture	$\ln(smsa)$	décembre 1993		juin 1995	
		Avant	Après	Avant	Après
Test ADF avec constante					
	ρ estimé	0.0260	0.9811	0.2139	0.9577
	t_{ADF}	-4.2322	-0.5741	-4.5436	-1.2775
	VC 5%	-2.9370	-2.9152	-2.9370	-2.9152
	Conclusion	I(0)	I(1)	I(0)	I(1)
Test ADF avec constante et tendance					
	ρ estimé	-0.0634	0.8052	0.0039	0.5974
	t_{ADF}	-4.2793	-2.7267	-5.2875	-2.8946
	VC 5%	-3.4615	-3.4611	-3.4615	-3.4611
	Conclusion	I(0) ^{n.s.}	I(1)	I(0)	I(1)
KPSS					
	Statistique	0.1883	1.0318	0.4693	0.9730
	VC 5%	0.4572	0.4572	0.4572	0.4572
	Conclusion	I(0)	I(1)	I(1)	I(1)

Trafic du dimanche/jour férié sur le métro

Tests de Perron

$\ln(dfmsa)$	Rupture sur le niveau (A)	Rupture sur le niveau et la pente (C)
ρ estimé	0.8523	0.1802
Statistique	-2.7278	-10.5921
Date de rupture	janvier 1994	juin 1995
Conclusion	I(1)-I(1)	I(0)-I(0)

Tests standards de racine unitaire sur les sous-périodes éventuelles

Date de rupture	$\ln(dfmsa)$	janvier 1994		juin 1995	
		Avant	Après	Avant	Après
Test ADF avec constante					
	ρ estimé	0.3662	1.0201	0.4446	1.0018
	t_{ADF}	-3.3706	0.6831	-2.9949	0.0628
	VC 5%	-2.9370	-2.9152	-2.9370	-2.9152
	Conclusion	I(0)	I(1)	I(0)	I(1)
Test ADF avec constante et tendance					
	ρ estimé	0.3564	0.8583	0.4158	0.4976
	t_{ADF}	-3.2404	-1.8961	-2.9882	-2.6258
	VC 5%	-3.4615	-3.4611	-3.4615	-3.4611
	Conclusion	I(1) ^{n.s.}	I(1)	I(1) ^{n.s.}	I(1)
KPSS					
	Statistique	0.1169	1.0736	0.2048	0.9923
	VC 5%	0.4572	0.4572	0.4572	0.4572
	Conclusion	I(0)	I(1)	I(0)	I(1)

ANNEXE D. APPLICATION DES TESTS DE PERRON ET DE NOS TESTS DE
MODIFICATION DU DEGRÉ D'INTÉGRATION

Trafic du samedi sur le RER

Tests de Perron

$\ln(srersa)$	Rupture sur le niveau (A)	Rupture sur le niveau et la pente (C)
ρ estimé	0.7675	0.5988
Statistique	-3.2871	-4.2697
Date de rupture	décembre 1993	mai 1995
Conclusion	I(1)-I(1)	I(1)-I(1)

Tests standards de racine unitaire sur les sous-périodes éventuelles

Date de rupture	$\ln(srersa)$	décembre 1993		mai 1995	
		Avant	Après	Avant	Après
Test ADF avec constante					
	ρ estimé	0.6930	0.9850	0.6800	0.9617
	t_{ADF}	-1.8984	-0.4921	-2.3230	-1.2017
	VC 5%	-2.9370	-2.9152	-2.9370	-2.9152
	Conclusion	I(1)	I(1)	I(1)	I(1)
Test ADF avec constante et tendance					
	ρ estimé	0.5973	0.8020	0.6796	0.5510
	t_{ADF}	-2.0210	-2.4690	-2.2076	-2.5159
	VC 5%	-3.4615	-3.4611	-3.4615	-3.4611
	Conclusion	I(1) ^{n.s.}	I(1)	I(1) ^{n.s.}	I(1)
KPSS					
	Statistique	0.3212	1.0704	0.2001	0.9781
	VC 5%	0.4572	0.4572	0.4572	0.4572
	Conclusion	I(0)	I(1)	I(0)	I(1)

Trafic du dimanche/jour férié sur le RER

Tests de Perron

$\ln(df\text{rersa})$	Rupture sur le niveau (A)	Rupture sur le niveau et la pente (C)
ρ estimé	0.7388	0.5571
Statistique	-4.1306	-5.3450
Date de rupture	janvier 1995	juin 1995
Conclusion	I(1)-I(1)	I(0)-I(0)

Tests standards de racine unitaire sur les sous-périodes éventuelles

Date de rupture	$\ln(df\text{rersa})$	janvier 1995		juin 1995	
		Avant	Après	Avant	Après
Test ADF avec constante					
	ρ estimé	0.8159	0.9924	0.7985	0.9779
	t_{ADF}	-1.7411	-0.2166	-2.1096	-0.6542
	VC 5%	-2.9370	-2.9152	-2.9370	-2.9152
	Conclusion	I(1)	I(1)	I(1)	I(1)
Test ADF avec constante et tendance					
	ρ estimé	0.7000	0.5331	0.6978	0.4946
	t_{ADF}	-2.2779	-3.8985	-2.3850	-3.3401
	VC 5%	-3.4615	-3.4611	-3.4615	-3.4611
	Conclusion	I(1) ^{n.s.}	I(0)	I(1) ^{n.s.}	I(1)
KPSS					
	Statistique	0.4198	0.9967	0.4324	0.9800
	VC 5%	0.4572	0.4572	0.4572	0.4572
	Conclusion	I(0)	I(1)	I(0)	I(1)

D.4 Application des tests de Perron et de nos tests de modification du degré d'intégration aux séries économiques

Carburants en Ile de France

Tests de Perron

ln (<i>carbidf</i>)	Rupture sur le niveau (A)	Rupture sur le niveau et la pente (C)
ρ estimé	0.5843	0.6105
Statistique	-4.2424	-4.6022
Date de rupture	juillet 1997	septembre 1997
Conclusion	I(1)-I(1)	I(0)-I(0)

Pour la série de ventes de carburant en Ile-de-France, une simple rupture du niveau de la série ne permet pas au test de Perron de rejeter l'existence de racine unitaire sur l'ensemble de la période (date estimée : juillet 1997), une rupture du niveau et de la pente de la tendance en septembre 1997 amène à conclure à la stationnarité de la série autour de cette tendance flexible. Néanmoins, les tests de racine unitaire refusent l'hypothèse de stationnarité sur chacune des sous-périodes dégagées : pour ces partitions la série semble intégrée puis stationnaire.

D.4. APPLICATION DES TESTS DE PERRON ET DE NOS TESTS DE MODIFICATION
DU DEGRÉ D'INTÉGRATION AUX SÉRIES ÉCONOMIQUES

Tests standards de racine unitaire sur les sous-périodes éventuelles

Date de rupture	$\ln(carbidf)$	juillet 1997		septembre 1997	
		Avant	Après	Avant	Après
Test ADF avec constante					
	ρ estimé	0.4385	0.9137	0.9574	0.3090
	t_{ADF}	-2.0434	-1.1574	-0.6727	-4.7170
	VC 5%	-2.9370	-2.9152	-2.9152	-2.9370
	Conclusion	I(1)	I(1)	I(1)	I(0)
Test ADF avec constante et tendance					
	ρ estimé	0.0135	0.5608	0.8129	0.1117
	t_{ADF}	-2.6526	-3.2963	-1.9152	-5.5259
	VC 5%	-3.4615	-3.4611	-3.4611	-3.4615
	Conclusion	I(1) ^{n.s.}	I(1)	I(1) ^{n.s.}	I(0)
KPSS					
	Statistique	0.4353	0.8622	0.7173	0.5924
	VC 5%	0.4572	0.4572	0.4572	0.4572
	Conclusion	I(0)	I(1)	I(1)	I(1)

Application des tests de modification du degré d'intégration

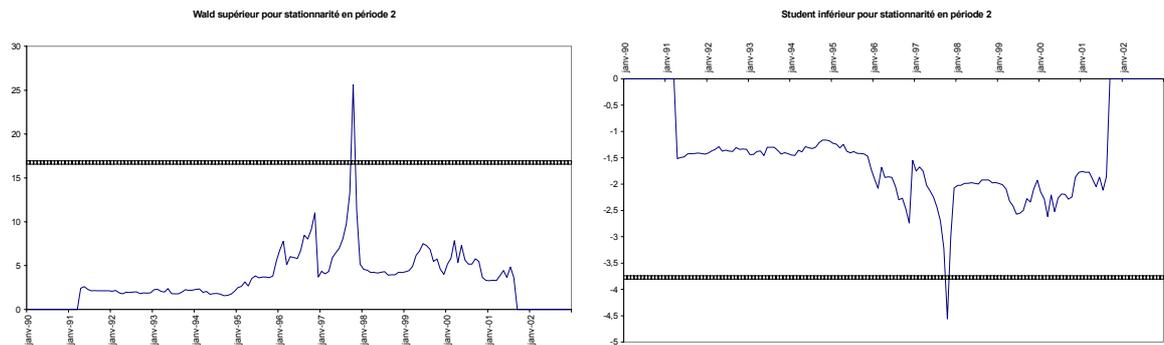
$\ln(carbidf)$	Statistique	VC 10%	Date	Conclusion
Cas intégré puis stationnaire				
Wald supérieur	25.6101	16.7610	octobre 1997	I(1)-I(0)
Student inférieur	-4.5569	-3.7730	octobre 1997	I(1)-I(0)
Cas stationnaire puis intégré				
Wald supérieur	7.7390	12.9260	septembre 1992	I(1)-I(1)
Student inférieur	-2.7390	-3.2940	septembre 1992	I(1)-I(1)

Nos tests dans le cas "stationnaire puis intégré" estiment une date de rupture proche de

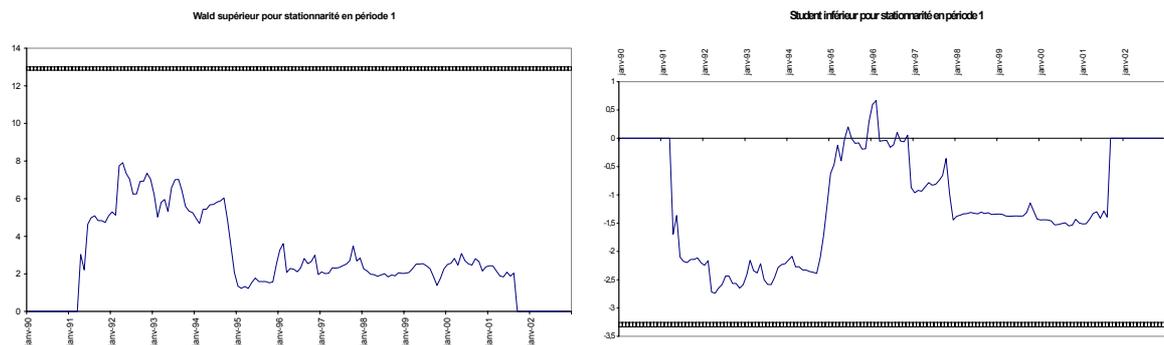
ANNEXE D. APPLICATION DES TESTS DE PERRON ET DE NOS TESTS DE MODIFICATION DU DEGRÉ D'INTÉGRATION

celles issues des tests de Perron. Le maximum du test supérieur de Wald et le minimum du test inférieur de Student sont obtenus au mois d'octobre 1997. Il y a modification du degré de stationnarité au cours de la période considérée.

Cas intégré puis stationnaire



Cas stationnaire puis intégré



Modélisation auto-régressive de la vente de carburants en Ile-de-France Sur la deuxième sous-période, postérieure à octobre 1997 et stationnaire, l'équation suivante est retenue :

Equation retenue :

$$\ln(carbidf)_t = 7.9277 + 0.3998 \ln(carbidf)_{t-1}$$

(5.5006) (3.6644)

$$\hat{\sigma} = 0.032921$$

D.4. APPLICATION DES TESTS DE PERRON ET DE NOS TESTS DE MODIFICATION
DU DEGRÉ D'INTÉGRATION AUX SÉRIES ÉCONOMIQUES

Circulation à Paris

Tests de Perron

$\ln(circ75)$	Rupture sur le niveau (A)	Rupture sur le niveau et la pente (C)
ρ estimé	0.6973	0.6486
Statistique	-4.3745	-4.7201
Date de rupture	octobre 1999	juillet 1998
Conclusion	I(1)-I(1)	I(0)-I(0)

Tests standards de racine unitaire sur les sous-périodes éventuelles

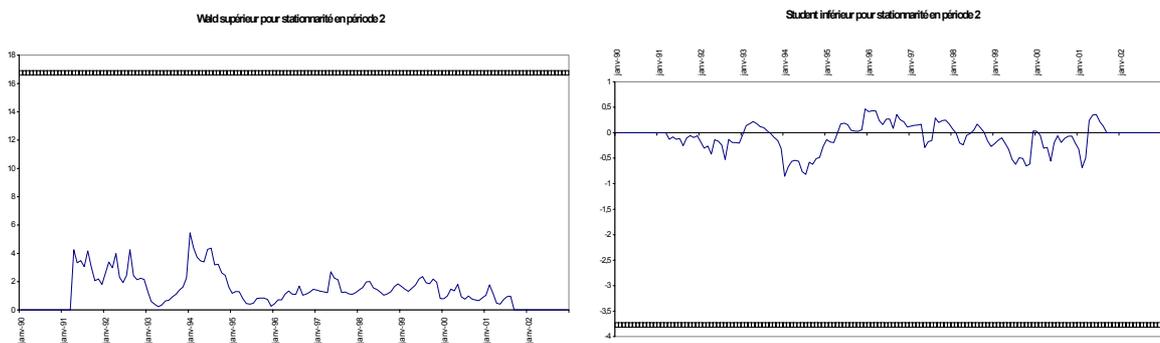
Date de rupture	$\ln(circ75)$	octobre 1999		juillet 1998	
		Avant	Après	Avant	Après
Test ADF avec constante					
	ρ estimé	0.9160	0.9890	0.8792	0.9938
	t_{ADF}	-1.3490	-0.1177	-1.5509	-0.1227
	VC 5%	-2.9152	-2.9370	-2.9152	-2.9370
	Conclusion	I(1)	I(1)	I(1)	I(1)
Test ADF avec constante et tendance					
	ρ estimé	0.7175	-0.5173	0.7220	-0.2374
	t_{ADF}	-3.0803	-3.6835	-2.7986	-4.0203
	VC 5%	-3.4611	-3.4615	-3.4611	-3.4615
	Conclusion	I(1)	I(0)	I(1)	I(0)
KPSS					
	Statistique	0.8779	0.5103	0.6380	0.6408
	VC 5%	0.4572	0.4572	0.4572	0.4572
	Conclusion	I(1)	I(1)	I(1)	I(1)

ANNEXE D. APPLICATION DES TESTS DE PERRON ET DE NOS TESTS DE
MODIFICATION DU DEGRÉ D'INTÉGRATION

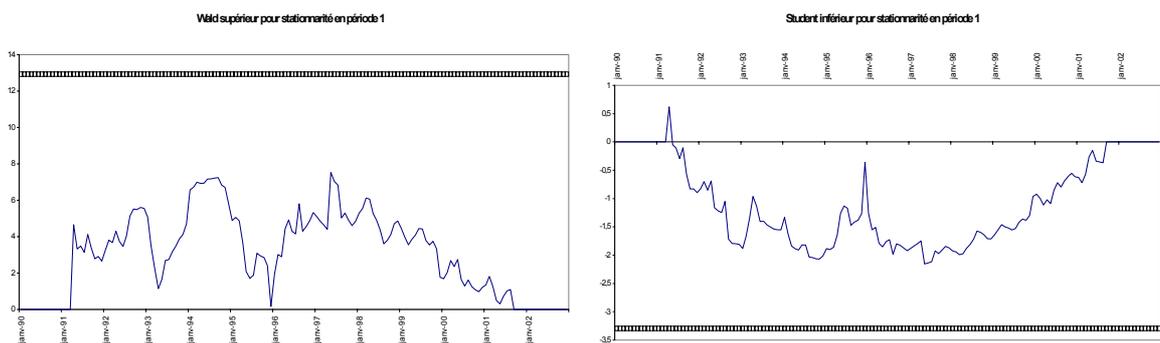
Application des tests de modification du degré d'intégration

$\ln(circ75)$	Statistique	VC 10%	Date	Conclusion
Cas intégré puis stationnaire				
Wald supérieur	5.4540	16.7610	janvier 1994	I(1)-I(1)
Student inférieur	-0.8536	-3.7730	janvier 1994	I(1)-I(1)
Cas stationnaire puis intégré				
Wald supérieur	7.5263	12.9260	mai 1997	I(1)-I(1)
Student inférieur	-2.1534	-3.2940	mai 1997	I(1)-I(1)

Cas intégré puis stationnaire



Cas stationnaire puis intégré



Le test de Perron avec rupture de niveau ne permet pas de rejeter l'hypothèse de racine unitaire pour la série relative au volume de circulation dans Paris (date estimée : octobre 1999)

D.4. APPLICATION DES TESTS DE PERRON ET DE NOS TESTS DE MODIFICATION
DU DEGRÉ D'INTÉGRATION AUX SÉRIES ÉCONOMIQUES

tandis que le test avec rupture du niveau et de la pente de la tendance a pour conclusion la stationnarité de la série avec rupture en juillet 1998. Le test augmenté de Dickey-Fuller confirme la stationnarité sur la seconde sous-période, mais ne remet pas en question l'hypothèse de racine unitaire sur la première sous-période. Nos tests ne détectent pas de modification du degré d'intégration. Ils estiment une date de rupture éventuelle plus précoce.

Equation retenue :

$$\begin{aligned} \Delta \ln(\text{circ75})_t = & - 0.0022 - 0.5424 \Delta \ln(\text{circ75})_{t-1} \\ & (-1.3303) \quad (-6.6873) \\ & - 0.2921 \Delta \ln(\text{circ75})_{t-2} - 0.1795 \Delta \ln(\text{circ75})_{t-3} \\ & (-3.2670) \quad (-2.1993) \end{aligned}$$

Vitesse de circulation à Paris

Tests de Perron

$\ln(\text{speed75})$	Rupture sur le niveau (A)	Rupture sur le niveau et la pente (C)
ρ estimé	0.1450	0.1891
Statistique	-11.0063	-9.9779
Date de rupture	novembre 1991	février 1992
Conclusion	I(0)-I(0)	I(0)-I(0)

Le test de Perron avec rupture du niveau de la série comme le test avec rupture du niveau et de la pente de la tendance concluent à la stationnarité de la série relative à la vitesse de circulation dans Paris autour de ce niveau et de cette tendance flexibles, pour des dates de rupture égales respectivement à novembre 1991 et février 1992.

ANNEXE D. APPLICATION DES TESTS DE PERRON ET DE NOS TESTS DE
MODIFICATION DU DEGRÉ D'INTÉGRATION

Tests standards de racine unitaire sur les sous-périodes éventuelles

Date de rupture	$\ln(\text{speed75})$	novembre 1991		février 1992	
		Avant	Après	Avant	Après
Test ADF avec constante					
	ρ estimé	0.6386	0.9457	0.7460	0.9269
	t_{ADF}	-2.1317	-1.0813	-0.9509	-1.3964
	VC 5%	-2.9493	-2.9124	-2.9370	-2.9124
	Conclusion	I(1)	I(1)	I(1)	I(1)
Test ADF avec constante et tendance					
	ρ estimé	0.6688	-0.1978	0.6578	-0.1206
	t_{ADF}	-1.9016	-5.4032	-1.2136	-5.6384
	VC 5%	-3.5449	-3.4542	-3.4615	-3.4542
	Conclusion	I(1) ^{n.s.}	I(0)	I(1) ^{n.s.}	I(0)
KPSS					
	Statistique	0.4144	0.7914	0.4717	0.7784
	VC 5%	0.4572	0.46572	0.4572	0.4572
	Conclusion	I(0)	I(1)	I(1)	I(1)

Application des tests de modification du degré d'intégration

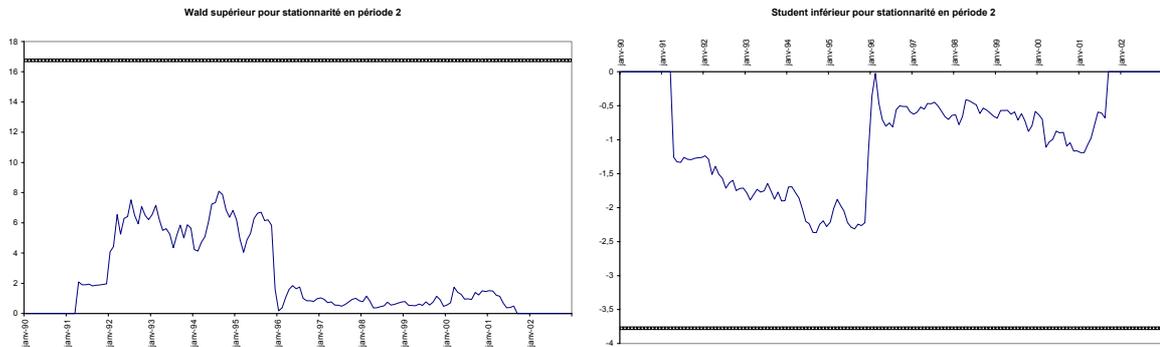
$\ln(\text{speed75})$	Statistique	VC 10%	Date	Conclusion
Cas intégré puis stationnaire				
Wald supérieur	8.0875	16.7610	août 1994	I(1)-I(1)
Student inférieur	-2.3673	-3.7730	septembre 1995	I(1)-I(1)
Cas stationnaire puis intégré				
Wald supérieur	14.3047	12.9260	avril 1996	I(0)-I(1)
Student inférieur	-3.6319	-3.2940	janvier 1996	I(0)-I(1)

Nos tests concluent que la série n'est pas stationnaire sur chaque sous-période, mais station-

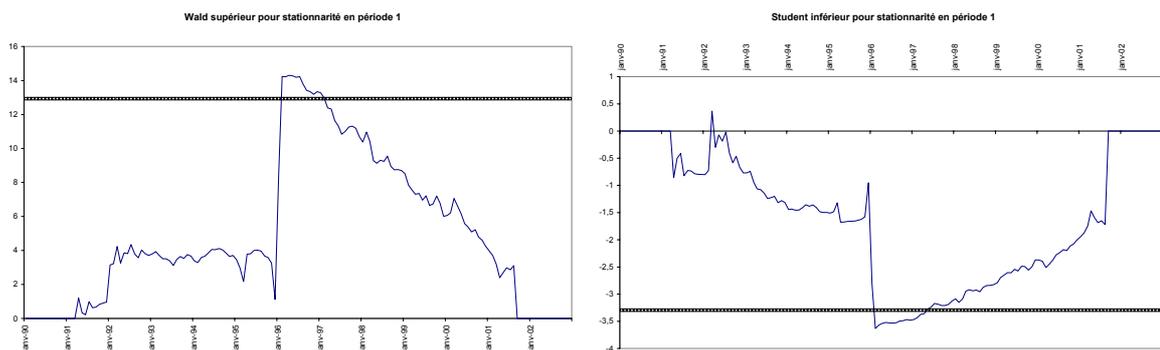
D.4. APPLICATION DES TESTS DE PERRON ET DE NOS TESTS DE MODIFICATION DU DEGRÉ D'INTÉGRATION AUX SÉRIES ÉCONOMIQUES

naire jusqu'en janvier 1996 puis intégrée. Les graphiques montrent une rupture particulièrement nette.

Cas intégré puis stationnaire



Cas stationnaire puis intégré



Modélisation auto-régressive de la vitesse de circulation dans Paris Deux modélisations auto-régressives de la vitesse de circulation en différence sont réalisées :

- sur période entière

$$\begin{aligned} \Delta \ln(\text{speed75})_t &= -0.0006 - 0.5478 \Delta \ln(\text{speed75})_{t-1} - 0.2507 \Delta \ln(\text{speed75})_{t-2} \\ &\quad (-0.1729) \quad (-6.9395) \quad (-3.1886) \\ \hat{\sigma} &= 0.043302 \end{aligned}$$

ANNEXE D. APPLICATION DES TESTS DE PERRON ET DE NOS TESTS DE
MODIFICATION DU DEGRÉ D'INTÉGRATION

– sur la deuxième sous-période, avril 1996 - décembre 2002

Equation retenue :

$$\Delta \ln(\text{speed75})_t = -0.0020 - 0.1568 \Delta \ln(\text{speed75})_{t-1} - 0.0643 \Delta \ln(\text{speed75})_{t-2}$$

(-0.8904)
(-2.4535)
(-1.2470)

$$\hat{\sigma} = 0.020876$$

Volume de circulation sur le réseau aggloméré

Tests de Perron

$\ln(\text{circag})$	Rupture sur le niveau (A)	Rupture sur le niveau et la pente (C)
ρ estimé	0.2331	0.1748
Statistique	-9.7119	-10.2545
Date de rupture	septembre 1991	juin 1993
Conclusion	I(0)-I(0)	I(0)-I(0)

D'après les tests de Perron, le volume de circulation sur le réseau aggloméré (routes nationales et autoroutes) est stationnaire, avec rupture en niveau au mois de septembre 1991, ou rupture du niveau et de la pente de la tendance en juin 1993.

Au contraire, les tests de racine unitaire sur les sous-périodes dégagées ne permettent pas de remettre en cause l'hypothèse de racine unitaire.

D.4. APPLICATION DES TESTS DE PERRON ET DE NOS TESTS DE MODIFICATION
DU DEGRÉ D'INTÉGRATION AUX SÉRIES ÉCONOMIQUES

Tests standards de racine unitaire sur les sous-périodes éventuelles

Date de rupture	$\ln(circag)$	septembre 1991		juin 1993	
		Avant	Après	Avant	Après
Test ADF avec constante					
	ρ estimé	0.1719	0.9184	0.9578	0.9217
	t_{ADF}	-2.6121	-1.5891	-0.3413	-1.2849
	VC 5%	-2.9493	-2.9124	-2.9379	-2.9152
	Conclusion	I(1)	I(1)	I(1)	I(1)
Test ADF avec constante et tendance					
	ρ estimé	-0.0271	0.3817	-0.2831	0.3234
	t_{ADF}	-2.8801	-4.5581	-3.3105	-0.6766
	VC 5%	-3.5449	-3.4542	-3.4615	-3.4611
	Conclusion	I(1) ^{n.s.}	I(0)	I(1)	I(1)
KPSS					
	Statistique	0.4546	0.7938	0.5273	1.1959
	VC 5%	0.4572	0.4572	0.4572	0.4572
	Conclusion	I(0)	I(1)	I(1)	I(1)

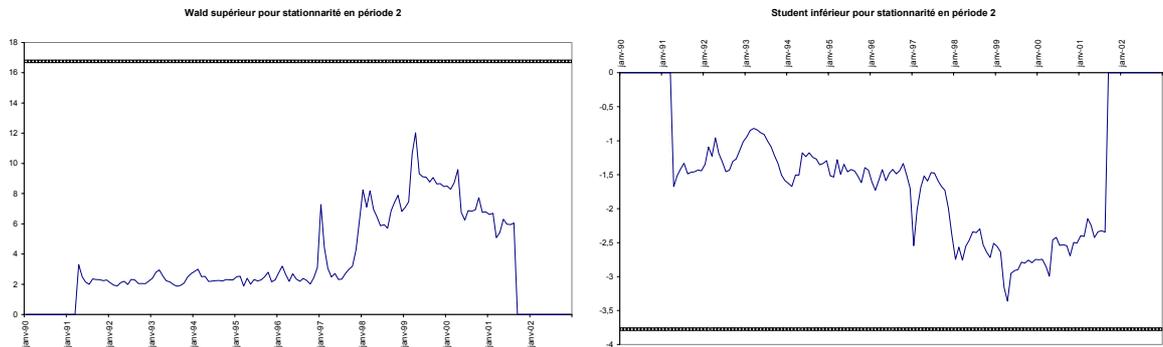
Application des tests de modification du degré d'intégration

$\ln(circagar)$	Statistique	VC 10%	Date	Conclusion
Cas intégré puis stationnaire				
Wald supérieur	12.0336	16.7610	avril 1999	I(1)-I(1)
Student inférieur	-3.3616	-3.7730	avril 1999	I(1)-I(1)
Cas stationnaire puis intégré				
Wald supérieur	3.9502	12.9260	mars 1993	I(1)-I(1)
Student inférieur	-1.5069	-3.2940	janvier 1997	I(1)-I(1)

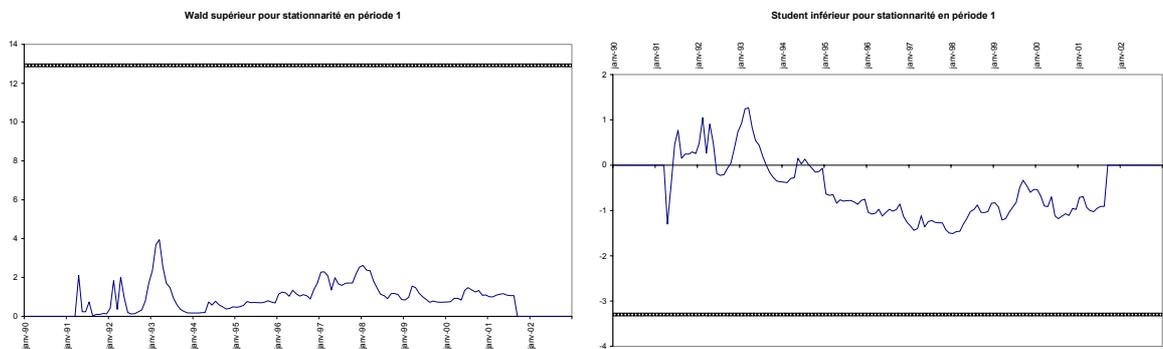
Nos tests confirment le caractère intégré de la série tout au long de la période considérée.

ANNEXE D. APPLICATION DES TESTS DE PERRON ET DE NOS TESTS DE MODIFICATION DU DEGRÉ D'INTÉGRATION

Cas intégré puis stationnaire



Cas stationnaire puis intégré



Modélisation auto-régressive du volume de circulation sur le réseau aggloméré Sur l'ensemble de la période, le volume de circulation sur le réseau aggloméré conduit à l'équation auto-régressive en différence

Equation retenue :

$$\begin{aligned} \Delta \ln(circagar)_t = & - 0.022 - 0.5424 \Delta \ln(circagar)_{t-1} \\ & (-1.3303) \quad (-6.6873) \\ & - 0.2921 \Delta \ln(circagar)_{t-2} - 0.1795 \Delta \ln(circagar)_{t-3} \\ & (-3.2670) \quad (-2.1993) \end{aligned}$$

D.4. APPLICATION DES TESTS DE PERRON ET DE NOS TESTS DE MODIFICATION
DU DEGRÉ D'INTÉGRATION AUX SÉRIES ÉCONOMIQUES

Demandes d'emplois en fin de mois à Paris

Tests de Perron

$\ln(defm75)$	Rupture sur le niveau (A)	Rupture sur le niveau et la pente (C)
ρ estimé	0.8539	0.8969
Statistique	-5.2209	-3.7671
Date de rupture	décembre 2000	novembre 1997
Conclusion	I(0)-I(0)	I(1)-I(1)

Tests standards de racine unitaire sur les sous-périodes éventuelles

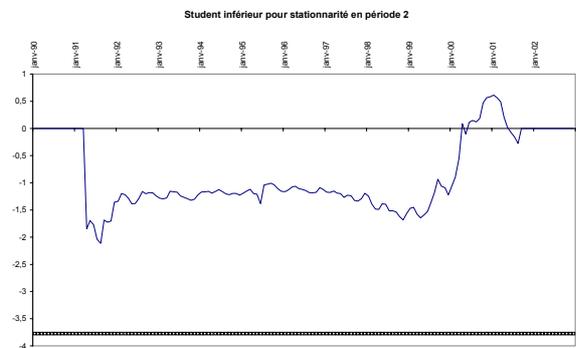
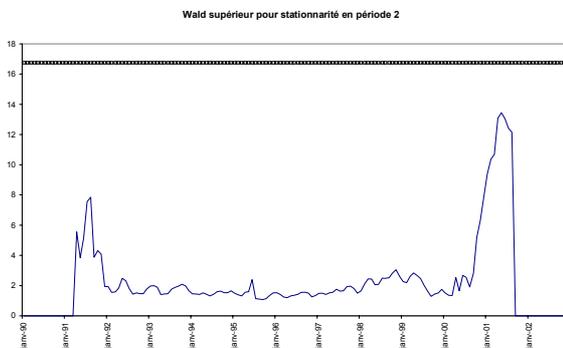
Date de rupture	$\ln(defm75)$	décembre 2000		novembre 1997	
		Avant	Après	Avant	Après
Test ADF avec constante					
	ρ estimé	0.8957	0.9498	0.8928	0.9666
	t_{ADF}	-2.9186	-1.3219	-2.8985	-2.1238
	VC 5%	-2.9152	-2.9370	-2.9152	-2.9370
	Conclusion	I(0)	I(1)	I(1)	I(1)
Test ADF avec constante et tendance					
	ρ estimé	0.8969	0.9235	0.8937	0.9628
	t_{ADF}	-2.8896	-1.8121	-2.8335	-1.3540
	VC 5%	-3.4611	-3.4615	-3.4611	-3.4615
	Conclusion	I(1) ^{n.s.}	I(1) ^{n.s.}	I(1) ^{n.s.}	I(1) ^{n.s.}
KPSS					
	Statistique	0.0899	0.2102	0.1036	0.4055
	VC 5%	0.4572	0.4572	0.4572	0.4572
	Conclusion	I(0)	I(0)	I(0)	I(0)

ANNEXE D. APPLICATION DES TESTS DE PERRON ET DE NOS TESTS DE
MODIFICATION DU DEGRÉ D'INTÉGRATION

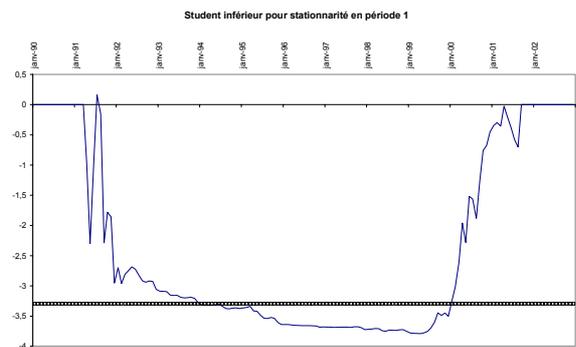
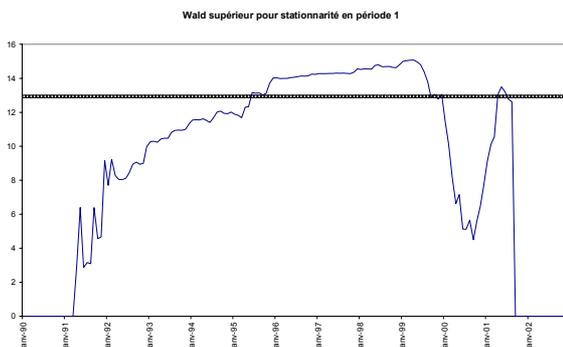
Application des tests de modification du degré d'intégration

$\ln(defm75)$	Statistique	VC 10%	Date	Conclusion
Cas intégré puis stationnaire				
Wald supérieur	13.4445	16.7610	mai 2001	I(1)-I(1)
Student inférieur	-2.1148	-3.7730	août 1991	I(1)-I(1)
Cas stationnaire puis intégré				
Wald supérieur	15.0726	12.9260	avril 1999	I(0)-I(1)
Student inférieur	-3.7902	-3.2940	avril 1999	I(0)-I(1)

Cas intégré puis stationnaire



Cas stationnaire puis intégré



Comparaison avec les tests de Perron Le test de Perron avec modification du niveau de la série de demandes d'emplois en fin de mois à Paris, pour une date de rupture estimée en décembre 2000, a pour conclusion que la série est stationnaire autour de ce niveau mobile. Le

D.4. APPLICATION DES TESTS DE PERRON ET DE NOS TESTS DE MODIFICATION DU DEGRÉ D'INTÉGRATION AUX SÉRIES ÉCONOMIQUES

test avec rupture sur le niveau et la pente de la tendance conclut que la série est intégrée sur l'ensemble de la période, pour une rupture (non significative pour la tendance) en novembre 1997. Les tests de racine unitaire sur les deux partitions dégagées indiquent que la série est stationnaire puis intégrée, voire stationnaire sur les deux sous-périodes.

Nos tests confirment l'hypothèse d'une série stationnaire puis intégrée. La date de rupture estimée correspond au mois d'avril 1999, mais toute date entre juillet 1995 et septembre 1999 pourrait être admise (période correspondant au plateau au delà des valeurs critiques sur les graphiques associés au cas stationnaire puis intégré).

Modélisation auto-régressive des demandes en fin de mois à Paris Sur l'ensemble de la période

$$\begin{aligned} \Delta \ln(defm75)_t &= 0.0874 \Delta \ln(defm75)_{t-1} + 0.2842 \Delta \ln(defm75)_{t-2} \\ &\quad (1.0876) \qquad\qquad\qquad (3.6749) \\ &+ 0.1880 \Delta \ln(defm75)_{t-3} \\ &\quad (2.3202) \\ \hat{\sigma} &= 0.011892 \end{aligned}$$

Les tests d'intégration concluent à l'existence d'une rupture en avril 1999. L'estimation auto-régressive en différence sur la variable en différence conduit à l'équation suivante :

Equation retenue :

$$\begin{aligned} \Delta \ln(defm75)_t &= 0.0006 - 0.1231 \Delta \ln(defm75)_{t-1} \\ &\quad (0.3466) \qquad\quad (-0.9402) \\ &+ 0.4133 \Delta \ln(defm75)_{t-2} + 0.5730 \Delta \ln(defm75)_{t-3} \\ &\quad (3.5609) \qquad\qquad\qquad (4.2841) \\ \hat{\sigma} &= 0.011216 \end{aligned}$$

ANNEXE D. APPLICATION DES TESTS DE PERRON ET DE NOS TESTS DE
MODIFICATION DU DEGRÉ D'INTÉGRATION

Demandes d'emploi en fin de mois en Ile de France

Tests de Perron

$\ln(defmidf)$	Rupture sur le niveau (A)	Rupture sur le niveau et la pente (C)
ρ estimé	0.9595	0.9471
Statistique	-4.7256	-4.0180
Date de rupture	mars 1999	mars 1996
Conclusion	I(0)-I(0)	I(1)-I(1)

Tests standards de racine unitaire sur les sous-périodes éventuelles

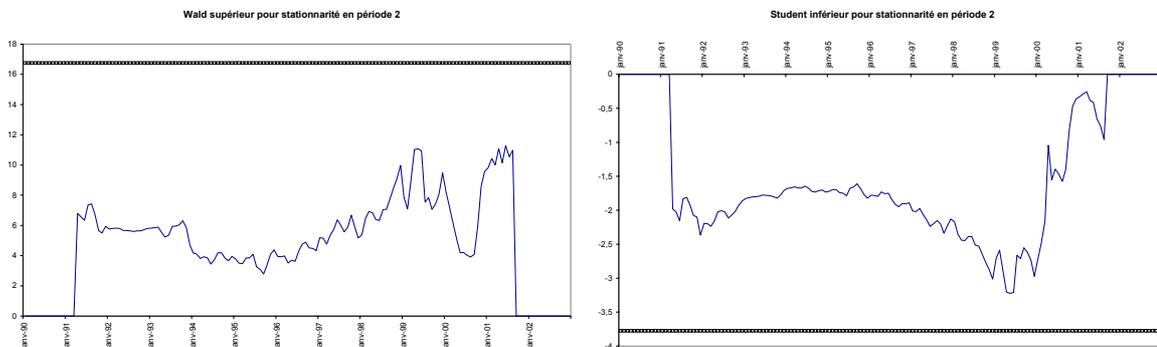
Date de rupture	$\ln(defmidf)$	mars 1999		mars 1996	
		Avant	Après	Avant	Après
Test ADF avec constante					
	ρ estimé	0.9829	0.9495	0.9846	0.9796
	t_{ADF}	-2.8317	-2.0762	-2.0413	-2.3975
	VC 5%	-2.9152	-2.9370	-2.9152	-2.9152
	Conclusion	I(1)	I(1)	I(1)	I(1)
Test ADF avec constante et tendance					
	ρ estimé	0.9839	0.9622	0.9695	0.9591
	t_{ADF}	-1.3658	-1.2899	-1.3660	-2.4135
	VC 5%	-3.4611	-3.4615	-3.4611	-3.4611
	Conclusion	I(1) ^{n.s.}	I(1) ^{n.s.}	I(1) ^{n.s.}	I(1) ^{n.s.}
KPSS					
	Statistique	0.9895	0.2550	0.7797	0.6760
	VC 5%	0.4572	0.4572	0.4572	0.4572
	Conclusion	I(1)	I(0)	I(1)	I(1)

D.4. APPLICATION DES TESTS DE PERRON ET DE NOS TESTS DE MODIFICATION DU DEGRÉ D'INTÉGRATION AUX SÉRIES ÉCONOMIQUES

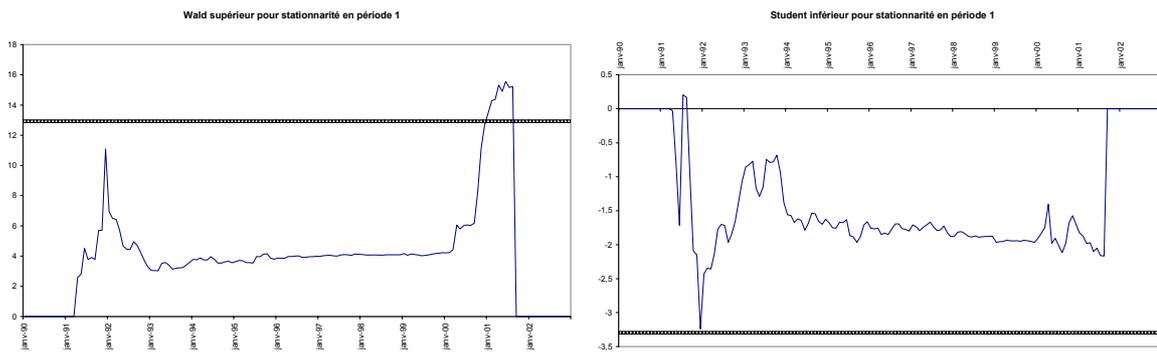
Application des tests de modification du degré d'intégration

ln(<i>defmidf</i>)	Statistique	VC 10%	Date	Conclusion
Cas intégré puis stationnaire				
Wald supérieur	11.2691	16.7610	juin 2001	I(1)-I(1)
Student inférieur	-3.2223	-3.7730	mai 1999	I(1)-I(1)
Cas stationnaire puis intégré				
Wald supérieur	15.5518	12.9260	juin 2001	I(0)-I(1)
Student inférieur	-3.2390	-3.2940	décembre 1991	I(0)-I(1) à 20%

Cas intégré puis stationnaire



Cas stationnaire puis intégré



Comparaison avec les tests de Perron Pour la série de demandes d'emplois en fin de mois en Ile-de-France, le test de Perron avec rupture sur le niveau de la série conclut à la stationnarité

ANNEXE D. APPLICATION DES TESTS DE PERRON ET DE NOS TESTS DE
MODIFICATION DU DEGRÉ D'INTÉGRATION

de la série, pour une rupture estimée en mars 1999. Le test de Perron avec rupture du niveau et de la pente de la tendance ne permet pas de rejeter l'hypothèse de racine unitaire, et estime la date de rupture en mars 1996. L'application des tests de racine unitaire aux sous-périodes dégagées confirme l'hypothèse de racine unitaire, exception faite de la sous-période avril 1999 - décembre 2002 pour laquelle le test KPSS accepte l'hypothèse de stationnarité.

Nos tests repèrent correctement la date à laquelle la série aurait pu être testée "intégrée puis stationnaire" (mai 1999), mais concluent que la variation du paramètre auto-régressif entre les deux sous-périodes est insuffisante pour considérer qu'il existe une modification du degré d'intégration de la série.

Le test inférieur de Student ayant pour hypothèse alternative "la série est stationnaire puis intégrée" conclut à la modification du degré d'intégration au début de la période (décembre 1991), juste avant la forte montée du nombre de demandeurs d'emplois en Ile-de-France.

Modélisation auto-régressive des demandes d'emploi en fin de mois en Ile-de-France

Sur la période entière :

$$\begin{aligned} \Delta \ln(defmidf)_t &= 0.0005 + 0.0962 \Delta \ln(defmidf)_{t-1} \\ &\quad (0.7081) \quad (1.2542) \\ &+ 0.3344 \Delta \ln(defmidf)_{t-2} + 0.3599 \Delta \ln(defmidf)_{t-3} \\ &\quad (4.6382) \quad (4.6853) \\ \hat{\sigma} &= 0.008190 \end{aligned}$$

Sur la sous-période janvier 1992 - décembre 2002,

Equation retenue :

$$\begin{aligned} \Delta \ln(defmidf)_t &= 0.0003 + 0.0501 \Delta \ln(defmidf)_{t-1} \\ &\quad (0.5094) \quad (0.6370) \\ &+ 0.3448 \Delta \ln(defmidf)_{t-2} + 0.4173 \Delta \ln(defmidf)_{t-3} \\ &\quad (4.7119) \quad (5.2716) \\ \hat{\sigma} &= 0.007670 \end{aligned}$$

D.4. APPLICATION DES TESTS DE PERRON ET DE NOS TESTS DE MODIFICATION
DU DEGRÉ D'INTÉGRATION AUX SÉRIES ÉCONOMIQUES

Volume des ventes dans les grands magasins de la région parisienne

Tests de Perron

<i>gdmagrp</i>	Rupture sur le niveau (A)	Rupture sur le niveau et la pente (C)
ρ estimé	0.8141	0.7235
Statistique	-3.2795	-2.7688
Date de rupture	août 1997	novembre 1994
Conclusion	I(1)-I(1)	I(1)-I(1)

Tests standards de racine unitaire sur les sous-périodes éventuelles

Date de rupture	<i>gdmagrp</i>	août 1997		novembre 1994	
		Avant	Après	Avant	Après
Test ADF avec constante					
	ρ estimé	0.8036	0.9003	0.9032	0.9586
	t_{ADF}	-2.2109	-1.3770	-0.7465	-0.8850
	VC 5%	-2.9152	-2.9370	-2.9370	-2.9152
	Conclusion	I(1)	I(1)	I(1)	I(1)
Test ADF avec constante et tendance					
	ρ estimé	0.3758	0.9595	0.2305	0.7874
	t_{ADF}	-3.0393	-0.3320	-1.6062	-1.8079
	VC 5%	-3.4611	-3.4615	-3.4615	-3.4611
	Conclusion	I(1)	I(1)	I(1) ^{n.s.}	I(1) ^{n.s.}
KPSS					
	Statistique	0.9125	0.5273	0.6828	0.9055
	VC 5%	0.4572	0.4572	0.4572	0.4572
	Conclusion	I(1)	I(1)	I(1)	I(1)

ANNEXE D. APPLICATION DES TESTS DE PERRON ET DE NOS TESTS DE
MODIFICATION DU DEGRÉ D'INTÉGRATION

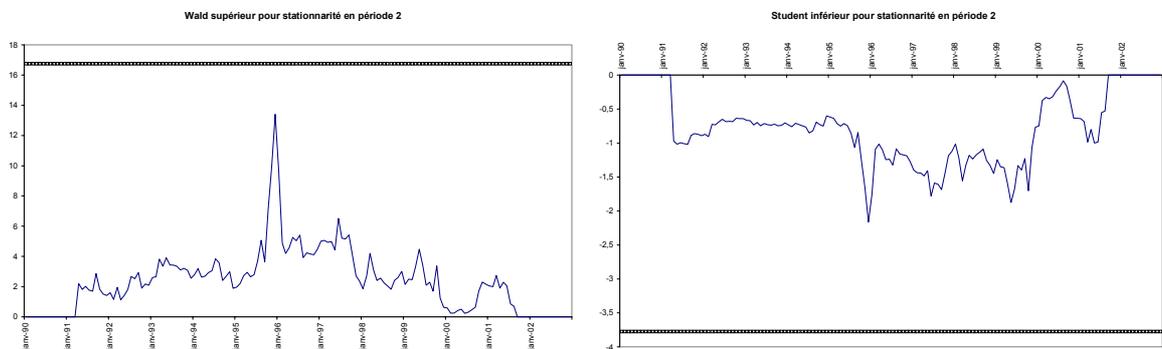
Application des tests de modification du degré d'intégration

<i>gdmagrp</i>	Statistique	VC 10%	Date	Conclusion
Cas intégré puis stationnaire				
Wald supérieur	13.4091	16.7610	décembre 1995	I(1)-I(1)
Student inférieur	-2.1626	-3.7730	décembre 1995	I(1)-I(1)
Cas stationnaire puis intégré				
Wald supérieur	17.1667	12.9260	avril 1992	I(0)-I(1)
Student inférieur	-3.9717	-3.2940	avril 1992	I(0)-I(1)

Comparaison avec les tests de Perron L'application des tests de Perron à la série d'indices des ventes en volume dans les grands magasins de la région parisienne conclut au caractère intégré de la série. La date estimée par le test avec rupture du niveau correspond au mois d'août 1997, tandis que le test avec rupture sur le niveau et la pente de la tendance estime la date de rupture en novembre 1994.

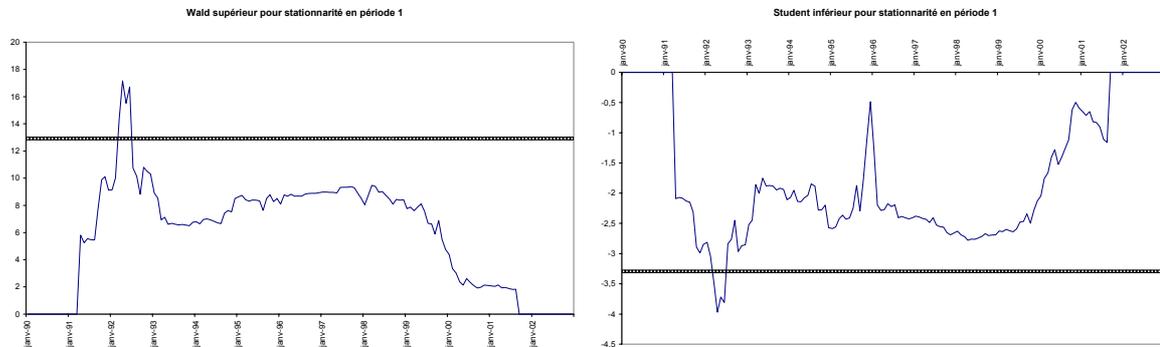
L'application des tests de racine unitaire standards confirme la présence de racine unitaire partout, pour ces partitions. Nos tests concluent à une modification du degré d'intégration précoce.

Cas intégré puis stationnaire



D.4. APPLICATION DES TESTS DE PERRON ET DE NOS TESTS DE MODIFICATION DU DEGRÉ D'INTÉGRATION AUX SÉRIES ÉCONOMIQUES

Cas stationnaire puis intégré



Modélisation auto-régressive de l'indice des ventes en volume dans les grands magasins de la région parisienne Sur période entière, la modélisation auto-régressive de l'indice des ventes en volume dans les grands magasins de la région parisienne s'écrit

$$\begin{aligned} \Delta gdmagrpt &= -0.0484 - 0.8200 \Delta gdmagrpt_{t-1} - 0.5039 \Delta gdmagrpt_{t-2} \\ &\quad (-0.1711) \quad (-11.5698) \quad (-7.0871) \\ \hat{\sigma} &= 3.497325 \end{aligned}$$

Sur la seconde sous-période (mai 1992 - décembre 2002), l'estimation devient

Equation retenue :

$$\begin{aligned} \Delta gdmagrpt &= 0.0360 - 0.7548 \Delta gdmagrpt_{t-1} - 0.4384 \Delta gdmagrpt_{t-2} \\ &\quad (0.1235) \quad (-9.6610) \quad (-5.7872) \\ \hat{\sigma} &= 3.294428 \end{aligned}$$

ANNEXE D. APPLICATION DES TESTS DE PERRON ET DE NOS TESTS DE
MODIFICATION DU DEGRÉ D'INTÉGRATION

Nombre de voyages grandes lignes passant par les gares de Paris

Tests de Perron

$\ln(totgl)$	Rupture sur le niveau (A)	Rupture sur le niveau et la pente (C)
ρ estimé	0.3370	0.0449
Statistique	-8.6218	-22.8604
Date de rupture	juin 1999	novembre 1995
Conclusion	I(0)-I(0)	I(0)-I(0)

Tests standards de racine unitaire sur les sous-périodes éventuelles

Date de rupture	$\ln(totgl)$	juin 1999		novembre 1995	
		Avant	Après	Avant	Après
Test ADF avec constante					
	ρ estimé	0.5447	0.0588	1.0590	0.8274
	t_{ADF}	-4.0753	-4.1539	0.6520	-1.9908
	VC 5%	-2.9152	-2.9370	-2.9370	-2.9152
	Conclusion	I(0)	I(0)	I(1)	I(1)
Test ADF avec constante et tendance					
	ρ estimé	0.3783	0.0472	0.6802	0.3442
	t_{ADF}	-4.7978	-4.0547	-1.4931	-4.1603
	VC 5%	-3.4611	-3.4615	-3.4615	-3.4611
	Conclusion	I(0)	I(0) ^{n.s.}	I(1)	I(0)
KPSS					
	Statistique	0.6876	0.2434	0.7825	0.8907
	VC 5%	0.4572	0.4572	0.4572	0.4572
	Conclusion	I(1)	I(0)	I(1)	I(1)

D.4. APPLICATION DES TESTS DE PERRON ET DE NOS TESTS DE MODIFICATION DU DEGRÉ D'INTÉGRATION AUX SÉRIES ÉCONOMIQUES

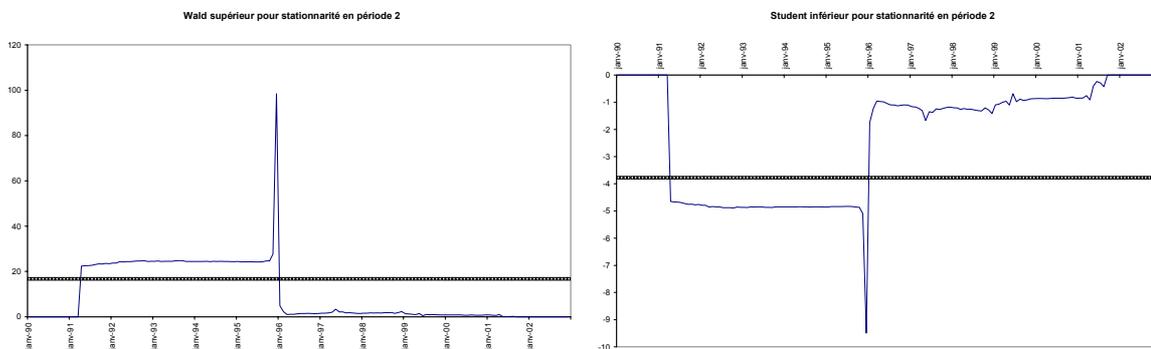
La pertinence de l'application des tests de modification du degré d'intégration à la série du nombre de voyageurs grandes lignes dans les gares de Paris est soumise à caution : les tests de racine unitaire concluent à la stationnarité de cette série.

Il est néanmoins intéressant de voir comment réagissent nos tests.

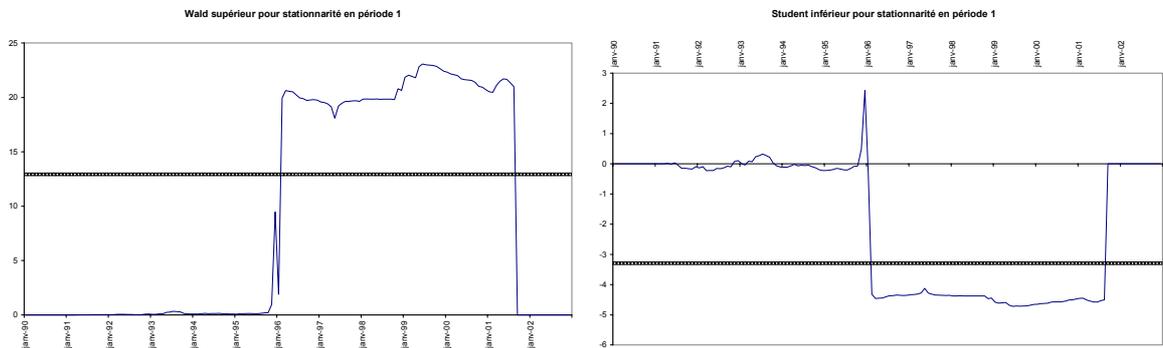
Application des tests de modification du degré d'intégration

$\ln(totgl)$	Statistique	VC 10%	Date	Conclusion
Cas intégré puis stationnaire				
Wald supérieur	98.4738	16.7610	décembre 1995	I(1)-I(0)
Student inférieur	-9.4958	-3.7730	décembre 1995	I(1)-I(0)
Cas stationnaire puis intégré				
Wald supérieur	23.0667	12.9260	décembre 1995	I(0)-I(1)
Student inférieur	-4.7182	-3.2940	juillet 1999	I(0)-I(1)

Cas intégré puis stationnaire



Cas stationnaire puis intégré



Comparaison avec les tests de Perron Les tests de Perron rejettent l’hypothèse de racine unitaire sur l’ensemble de la période, au profit de l’hypothèse de stationnarité autour d’un niveau ayant subi une rupture en juin 1999, ou autour d’un niveau et d’une pente de la tendance soumis à rupture en novembre 1995. Si l’on considère la première date de rupture, les tests de racine unitaire standards indiquent que la série est bien stationnaire sur chacune des sous-périodes. Si l’on prend en compte la deuxième date de rupture, novembre 1995, la série est testée intégrée sur chaque sous-période déagée.

Nos tests amènent à un résultat surprenant : pour une rupture en décembre 1995, la série est à la fois stationnaire puis intégrée et intégrée puis stationnaire. Ce résultat étonnant s’explique par l’importance du choc des grèves de décembre 1995, qui perturbe le test. Notons en outre que le test inférieur de Student pour l’hypothèse alternative ”la série est stationnaire puis intégrée” indique une date de rupture proche elle aussi de la date de rupture issue du test de Perron.

Modélisation auto-régressive du nombre de voyageurs grandes lignes dans les gares de Paris La série étant stationnaire, la modélisation de la série du nombre de voyageurs grandes lignes dans les gares de Paris est effectuée sur la série en niveau, et non la série en différence première. Sur l’ensemble de la période, l’équation suivante est obtenue :

$$\ln(totgl)_t = 8.1489 + 0.4862 \ln(totgl)_{t-1}$$

$$(7.2790) \quad (6.68875)$$

$$\hat{\sigma} = 0.087287$$

D.4. APPLICATION DES TESTS DE PERRON ET DE NOS TESTS DE MODIFICATION
DU DEGRÉ D'INTÉGRATION AUX SÉRIES ÉCONOMIQUES

Pour la sous-période janvier 1996 - décembre 2002, cette équation devient

Equation retenue :

$$\ln(totgl)_t = 11.2628 + 0.1534 \ln(totgl)_{t-1} + 0.1367 \ln(totgl)_{t-2}$$

(12.1769) (2.7962) (2.5685)

$$\hat{\sigma} = 0.055337$$

Taux d'occupation des hôtels à Paris

Tests de Perron

ln(<i>txhot75</i>)	Rupture sur le niveau (A)	Rupture sur le niveau et la pente (C)
ρ estimé	0.5151	0.4748
Statistique	-7.3487	-7.3169
Date de rupture	septembre 1997	août 1996
Conclusion	I(0)-I(0)	I(0)-I(0)

Appliqués à la série du taux d'occupation des hôtels à Paris, les tests de Perron rejettent l'hypothèse de racine unitaire sur l'ensemble de la période. La présence d'une rupture dans le niveau de la série indique une rupture en septembre 1997, tandis qu'une modification éventuelle du niveau et de la pente de la tendance conduit à estimer une date de rupture en août 1996.

Les tests standards de racine unitaire appliqués sur chaque sous-partie semblent confirmer le caractère stationnaire de chacune des sous-périodes dégagées pour la partition en septembre 1997.

ANNEXE D. APPLICATION DES TESTS DE PERRON ET DE NOS TESTS DE
MODIFICATION DU DEGRÉ D'INTÉGRATION

Tests standards de racine unitaire sur les sous-périodes éventuelles

Date de rupture	$\ln(txhot75)$	septembre 1997		août 1996	
		Avant	Après	Avant	Après
Test ADF avec constante					
	ρ estimé	0.6550	0.5342	0.6407	0.6138
	t_{ADF}	-3.4750	-3.3610	-3.3744	-3.2362
	VC 5%	-2.9152	-2.9370	-2.9152	-2.9152
	Conclusion	I(0)	I(0)	I(0)	I(0)
Test ADF avec constante et tendance					
	ρ estimé	0.6444	0.5242	0.5028	0.5825
	t_{ADF}	-3.2105	-3.3895	-3.6772	-3.2770
	VC 5%	-3.4611	-3.4615	-3.4611	-3.4611
	Conclusion	I(1) ^{n.s.}	I(0) ^{n.s.}	I(0) ^{n.s.}	I(1) v
KPSS					
	Statistique	0.3305	0.1580	0.5158	0.3220
	VC 5%	0.4572	0.4572	0.4572	0.4572
	Conclusion	I(0)	I(0)	I(1)	I(0)

Application des tests de modification du degré d'intégration

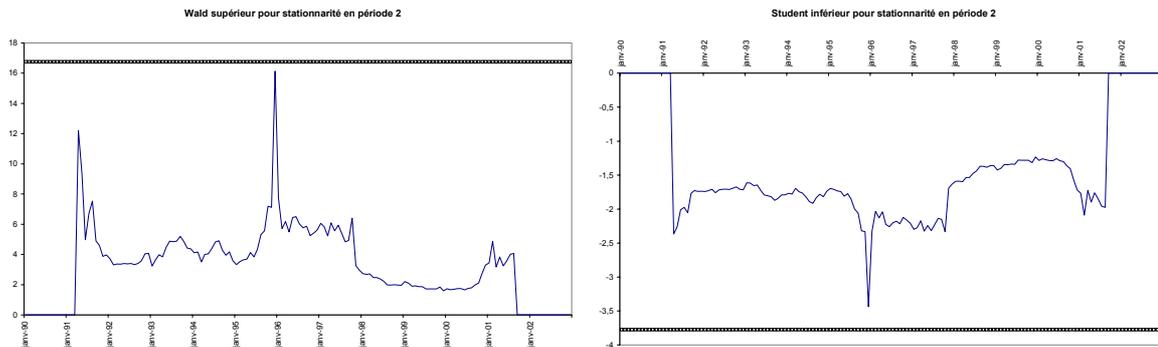
$txhot75$	Statistique	VC 10%	Date	Conclusion
Cas intégré puis stationnaire				
Wald supérieur	16.1343	16.7610	décembre 1995	I(1)-I(0) à 20%
Student inférieur	-3.4384	-3.7730	décembre 1995	I(1)-I(0) à 20%
Cas stationnaire puis intégré				
Wald supérieur	14.0242	12.9260	novembre 1997	I(0)-I(1)
Student inférieur	-3.6312	-3.2940	novembre 1997	I(0)-I(1)

Nos tests indiquent, comme pour la série précédente que la série est à la "stationnaire puis

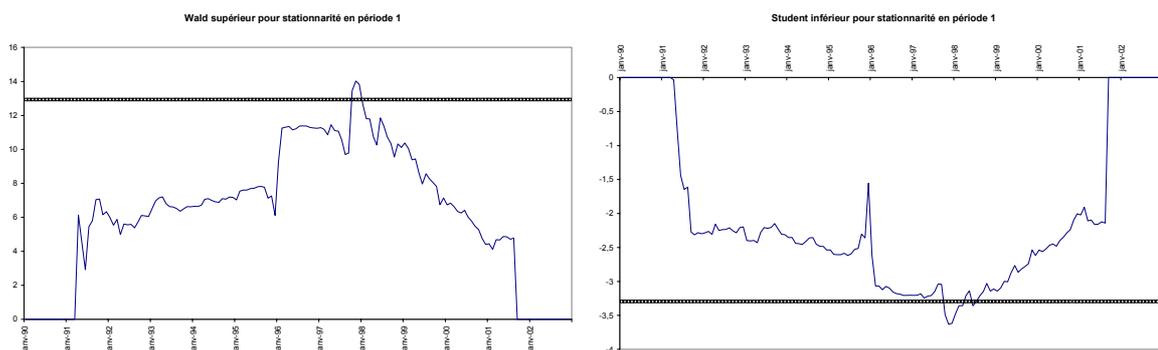
D.4. APPLICATION DES TESTS DE PERRON ET DE NOS TESTS DE MODIFICATION DU DEGRÉ D'INTÉGRATION AUX SÉRIES ÉCONOMIQUES

intégrée” et ”intégrée puis stationnaire”. Les dates de rupture estimées correspondent au mois de décembre 1995 et novembre 1997 (proche du test A de Perron). Ceci incite à penser que la série est stationnaire sur chacune des trois sous-périodes.

Cas intégré puis stationnaire



Cas stationnaire puis intégré



Modélisation auto-régressive du taux d'occupation des hôtels à Paris La modélisation auto-régressive de la série est réalisée sur la variable en niveau, pour la période entière :

$$\begin{aligned}
 txhot75_t = & \quad 0.1427 \quad + \quad 0.5001 \quad txhot75_{t-1} \\
 & \quad (3.2371) \quad \quad (6.2191) \\
 & + \quad 0.1434 \quad txhot75_{t-2} + \quad 0.1508 \quad txhot75_{t-3} \\
 & \quad (1.6025) \quad \quad \quad (1.8793)
 \end{aligned}$$

ANNEXE D. APPLICATION DES TESTS DE PERRON ET DE NOS TESTS DE
MODIFICATION DU DEGRÉ D'INTÉGRATION

Taux d'occupation dans les hôtels d'Ile-de-France

Tests de Perron

<i>t_{xhotidf}</i>	Rupture sur le niveau (A)	Rupture sur le niveau et la pente (C)
ρ estimé	0.6861	0.7063
Statistique	-4.2327	-3.7541
Date de rupture	juillet 1992	juillet 1992
Conclusion	I(1)-I(1)	I(1)-I(1)

Tests standards de racine unitaire sur les sous-périodes éventuelles

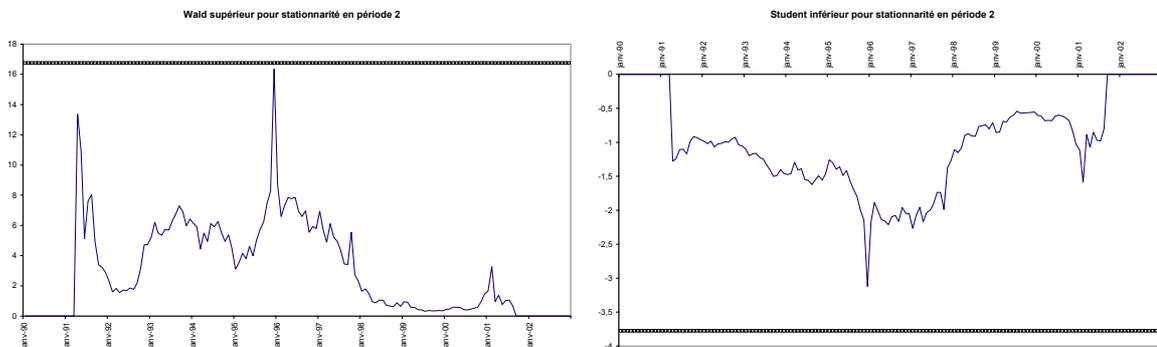
Date de rupture	<i>t_{xhotidf}</i>	juillet 1992	
		Avant	Après
Test ADF avec constante			
	ρ estimé	0.6410	0.9507
	t_{ADF}	-2.1664	-1.0712
	VC 5%	-2.9370	-2.9152
	Conclusion	I(1)	I(1)
Test ADF avec constante et tendance			
	ρ estimé	0.6596	0.8278
	t_{ADF}	-1.8183	-1.6866
	VC 5%	-3.4615	-3.4611
	Conclusion	I(1) ^{n.s.}	I(1) ^{n.s.}
KPSS			
	Statistique	0.2161	1.1217
	VC 5%	0.4572	0.4572
	Conclusion	I(0)	I(1)

D.4. APPLICATION DES TESTS DE PERRON ET DE NOS TESTS DE MODIFICATION DU DEGRÉ D'INTÉGRATION AUX SÉRIES ÉCONOMIQUES

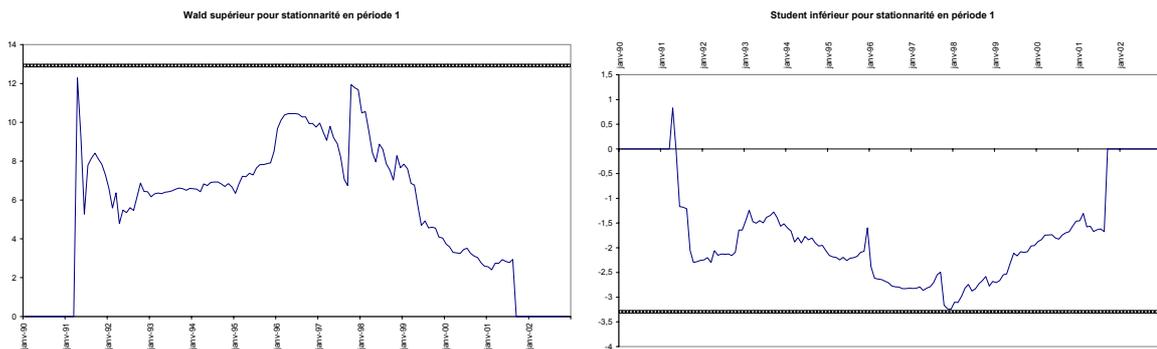
Application des tests de modification du degré d'intégration

<i>txhotidf</i>	Statistique	VC 10%	Date	Conclusion
Cas intégré puis stationnaire				
Wald supérieur	16.3434	16.7610	décembre 1995	$I(1)-I(0)$
Student inférieur	-3.1209	-3.7730	décembre 1995	$I(1)-I(1)$
Cas stationnaire puis intégré				
Wald supérieur	12.2863	12.9260	avril 1991	$I(0)-I(1)$ à 20%
Student inférieur	-3.2429	-3.2940	décembre 1997	$I(0)-I(1)$ à 20%

Cas intégré puis stationnaire



Cas stationnaire puis intégré



Comparaison avec les tests de Perron La série du taux d'occupation des hôtels dans l'ensemble de l'Ile-de-France possède une racine unitaire sur l'ensemble de la période pour les

ANNEXE D. APPLICATION DES TESTS DE PERRON ET DE NOS TESTS DE
MODIFICATION DU DEGRÉ D'INTÉGRATION

tests de Perron, avec rupture en juillet 1992 quel que soit le type de rupture autorisé. Les tests augmentés de Dickey-Fuller appliqués aux sous-périodes confirment cette racine unitaire, tandis que le test KPSS conclut que la série est stationnaire sur la première sous-période.

Notre test supérieur de Wald avec hypothèse alternative "la série est intégrée puis stationnaire" indique une rupture en décembre 1995, non confirmée par le test inférieur de Student avec hypothèse alternative similaire : il y a eu rupture du niveau mais sans doute pas du paramètre auto-régressif.

Notre test supérieur de Wald ayant pour hypothèse alternative "la série est stationnaire puis intégrée" conclut au changement de degré d'intégration en avril 1991. Cette conclusion est confirmée par le test inférieur de Student à un seuil légèrement supérieur à 10%, pour une date de rupture en novembre 1997.

Modélisation auto-régressive du taux d'occupation des hôtels en Ile-de-France Sur l'ensemble la période, l'équation auto-régressive appropriée s'écrit :

$$\begin{aligned} \Delta txhotidf_t = & - 0.0005 - 0.4837 \Delta txhotidf_{t-1} - 0.3975 \Delta txhotidf_{t-2} \\ & (-0.2102) \quad (-5.9993) \quad (-4.5223) \\ & - 0.2190 \Delta txhotidf_{t-3} - 0.2210 \Delta txhotidf_{t-4} \\ & (-2.4918) \quad (-2.7375) \\ \hat{\sigma} = & 0.031077 \end{aligned}$$

Pour la sous-période démarrant en décembre 1997, cette équation devient :

Equation retenue :

$$\begin{aligned} \Delta txhotidf_t = & 0.0019 - 0.4115 \Delta txhotidf_{t-1} - 0.3023 \Delta txhotidf_{t-2} \\ & (0.6646) \quad (-3.4517) \quad (-2.6531) \\ & - 0.2555 \Delta txhotidf_{t-3} - 0.2781 \Delta txhotidf_{t-4} \\ & (-2.3907) \quad (-2.8742) \\ \hat{\sigma} = & 0.021805 \end{aligned}$$

D.4. APPLICATION DES TESTS DE PERRON ET DE NOS TESTS DE MODIFICATION
DU DEGRÉ D'INTÉGRATION AUX SÉRIES ÉCONOMIQUES

Rapport du prix du ticket T et du prix du super

Tests de Perron

<i>compet</i>	Rupture sur le niveau (A)	Rupture sur le niveau et la pente (C)
ρ estimé	0.8537	0.8736
Statistique	-4.3648	-3.6970
Date de rupture	mars 1999	mars 1999
Conclusion	I(1)-I(1)	I(1)-I(1)

Tests standards de racine unitaire sur les sous-périodes éventuelles

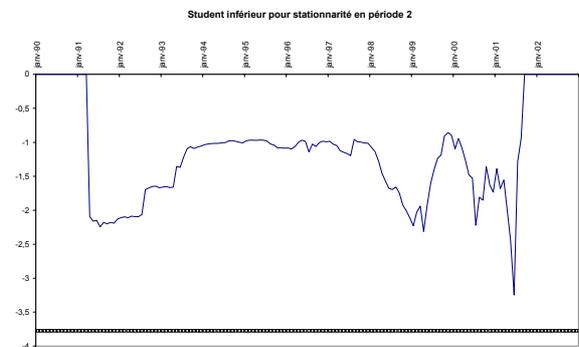
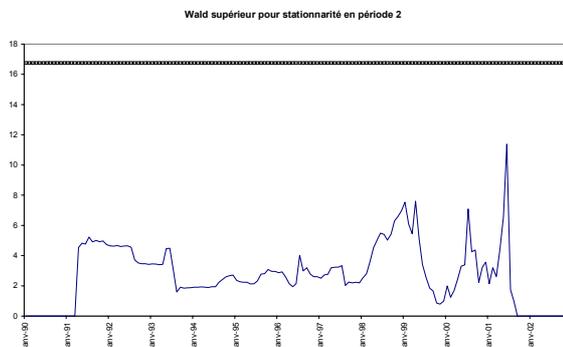
Date de rupture	<i>compet</i>	mars 1999	
		Avant	Après
Test ADF avec constante			
	ρ estimé	0.9747	0.9054
	t_{ADF}	-1.3844	-1.5682
	VC 5%	-2.9152	-2.9370
	Conclusion	I(1)	I(1)
Test ADF avec constante et tendance			
	ρ estimé	0.9222	0.8269
	t_{ADF}	-1.5378	-2.5620
	VC 5%	-3.4611	-3.4615
	Conclusion	I(1) ^{n.s.}	I(1)
KPSS			
	Statistique	1.0801	0.1944
	VC 5%	0.4572	0.4572
	Conclusion	I(1)	I(0)

ANNEXE D. APPLICATION DES TESTS DE PERRON ET DE NOS TESTS DE
MODIFICATION DU DEGRÉ D'INTÉGRATION

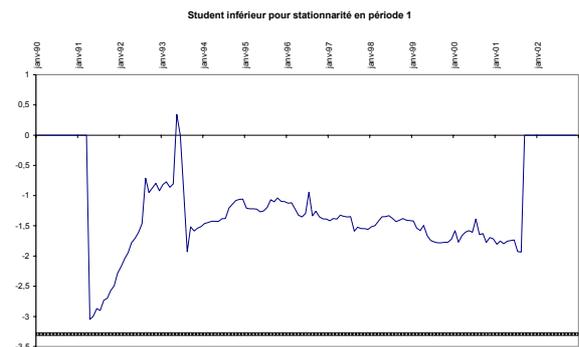
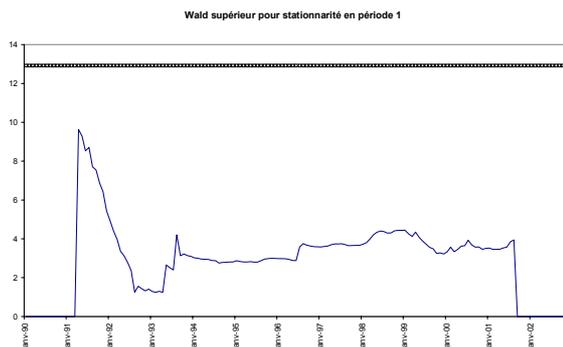
Application des tests de modification du degré d'intégration

<i>compet</i>	Statistique	VC 10%	Date	Conclusion
Cas intégré puis stationnaire				
Wald supérieur	11.3848	16.7610	juin 2001	I(1)-I(1)
Student inférieur	-3.2429	-3.7730	juin 2001	I(1)-I(1)
Cas stationnaire puis intégré				
Wald supérieur	9.6249	12.9260	avril 1991	I(1)-I(1)
Student inférieur	-3.0507	-3.2940	avril 1991	I(1)-I(1)

Cas intégré puis stationnaire



Cas stationnaire puis intégré



Comparaison avec les tests de Perron Pour une rupture du seul niveau de la série, ou pour une rupture du niveau et de la pente de la tendance, les tests de Perron ne permettent

D.4. APPLICATION DES TESTS DE PERRON ET DE NOS TESTS DE MODIFICATION DU DEGRÉ D'INTÉGRATION AUX SÉRIES ÉCONOMIQUES

pas d'infirmer le caractère intégré de la série correspondant au rapport des prix du ticket T et du super. La date de rupture estimée est identique pour les deux tests : mars 1999. Les tests augmentés de Dickey-Fuller confirment la présence de racine unitaire sur les sous-périodes, tandis que le test KPSS laisse penser que la série pourrait être stationnaire sur la seconde sous-période.

Nos tests vont dans le sens des résultats de Perron comme des résultats des tests ADF : la série semble intégrée sur l'ensemble de la période. La date de rupture estimée pour l'alternative "la série est intégrée puis stationnaire" correspond au mois de juin 2001, mais il existe des extrema locaux en janvier ou avril 1999.

Modélisation auto-régressive du rapport du prix du billet T au prix du super Sur la période entière, la modélisation auto-régressive sur la série en différence conduit à l'équation :

Equation retenue :

$$\Delta compet_t = 0.0011 - 0.1320 \Delta compet_{t-1} - 0.0853 \Delta compet_{t-2}$$

(0.6043) (-1.6228) (-1.0486)

Annexe E

Tables de puissance des tests de changement du degré de cointégration

E.1 Processus générateur de données et calculs de puissance

Processus générateur de données 2 500 simulations de Monte Carlo ont été réalisées sur un échantillon de taille 150 en supposant un processus générateur sous l'hypothèse alternative ayant pour relation de cointégration

$$y_t = c + x_t + e_t \text{ (avec constante)}$$

$$y_t = c + t + x_t + e_t \text{ (avec constante et tendance)}$$

avec

$$x_t \sim I(1) \text{ (marche aléatoire)}$$

$$e_t = c_1 + \rho_1 e_{t-1} + u_t \text{ si } t \leq k$$

$$= c_2 + \rho_2 e_{t-1} + u_t \text{ si } t > k$$

Pour chaque type de simulations sont indiquées les valeurs de la date de rupture k , des paramètres auto-régressifs sur la première sous-période (ρ_1) et la deuxième sous-période (ρ_2), des constantes relatives à chaque sous-période (c_1 et c_2).

Le niveau des tests est posé égal à 10%, la date de rupture est estimée en enlevant 10% des observations au début de l'échantillon et 10% des observations en fin d'échantillon.

Les calculs de puissance sont effectués pour un véritable modèle "non cointégré puis cointégré" comme pour un vrai modèle "cointégré puis non cointégré".

Lecture des tableaux de puissance Les tableaux de puissance comportent les renseignements suivants :

- la puissance du test en %, c'est-à-dire le pourcentage de fois où notre test rejette à raison l'hypothèse d'absence de changement du degré de cointégration alors que le processus générateur de données suppose l'existence d'un tel changement du degré de cointégration
- le pourcentage de bonnes dates estimées lorsque le test rejette l'absence de changement du degré de cointégration qui représente la probabilité que la date de rupture estimée \hat{k} soit égale à la véritable date de rupture k
- les intervalles de confiance à 3 et 6 mois (respectivement "IC à ± 3 mois" et "IC à ± 6 mois") autour de la date de rupture k , qui correspondent à la probabilité que la date de rupture estimée \hat{k} soit comprise entre $k - 3$ et $k + 3$ pour l'intervalle de confiance à 3 mois et $k - 6$ et $k + 6$ pour l'intervalle de confiance à 6 mois.

E.2 Puissance des tests suivant la valeur du paramètre auto-régressif des résidus de la relation cointégrée

ANNEXE E. TABLES DE PUISSANCE DES TESTS DE CHANGEMENT DU DEGRÉ DE
COINTÉGRATION

Ha : "Absence de cointégration puis cointégration"

$$\left\{ k = 75, \rho_1 = 1, c_1 = 0, c_2 = 0 \right\}$$

Test supérieur de Wald dans le cas "fallacieux" puis cointégré

Relation de cointégration avec constante

ρ_2	Puissance	Bonne date estimée	IC à ± 3 mois	IC à ± 6 mois
-0.9	98.80	84.696	90.728	94.048
-0.8	98.80	82.955	88.582	92.226
-0.7	98.36	82.432	88.735	91.988
-0.6	98.60	82.759	88.397	92.170
-0.5	98.60	80.446	86.937	91.034
-0.4	98.32	80.146	86.290	89.830
-0.3	97.56	77.368	83.764	87.782
-0.2	97.48	78.170	83.546	88.758
-0.1	97.04	76.381	82.110	87.427
0	96.72	73.904	80.067	84.988
0.1	95.48	73.188	79.136	83.368
0.2	94.00	71.362	77.830	81.362
0.3	93.76	69.113	75.427	79.523
0.4	92.04	66.927	73.750	77.966
0.5	89.88	60.303	68.893	73.522
0.6	83.32	57.801	66.106	70.474
0.7	71.36	52.354	62.668	66.760
0.8	53.00	44.000	57.736	62.718
0.9	28.56	23.529	43.977	50.699

E.2. PUISSANCE DES TESTS SUIVANT LA VALEUR DU PARAMÈTRE
 AUTO-RÉGRESSIF DES RÉSIDUS DE LA RELATION COINTÉGRÉE

Test supérieur de Wald dans le cas "fallacieux" puis cointégré

Relation de cointégration avec constante et tendance

ρ_2	Puissance	Bonne date estimée	IC à ± 3 mois	IC à ± 6 mois
-0.9	98.00	78.776	86.776	90.695
-0.8	97.48	78.006	84.982	88.798
-0.7	97.80	77.832	84.785	88.793
-0.6	97.44	76.847	83.210	87.727
-0.5	96.76	75.734	81.687	86.150
-0.4	96.40	74.647	81.575	86.262
-0.3	95.08	73.244	79.597	83.845
-0.2	94.48	71.423	76.757	81.414
-0.1	92.92	70.857	75.462	81.102
0	92.08	70.156	75.846	81.319
0.1	91.44	68.635	74.148	78.697
0.2	89.16	67.295	72.993	77.524
0.3	87.60	63.014	69.088	73.973
0.4	84.40	61.043	67.014	72.796
0.5	78.60	61.018	67.939	71.348
0.6	74.44	55.562	62.011	65.933
0.7	60.44	48.908	58.107	61.880
0.8	43.20	40.741	54.445 %	57.594
0.9	21.96	19.854 %	36.064 %	42.258 %

ANNEXE E. TABLES DE PUISSANCE DES TESTS DE CHANGEMENT DU DEGRÉ DE
COINTÉGRATION

Test inférieur de Student dans le cas "fallacieux" puis cointégré

Relation de cointégration avec constante

ρ_2	Puissance	Bonne date estimée	IC à ± 3 mois	IC à ± 6 mois
-0.9	98.76	85.622	91.618	94.739
-0.8	98.92	84.068	89.851	93.571
-0.7	98.48	82.981	89.521	92.485
-0.6	98.64	83.496	89.498	93.067
-0.5	98.76	81.045	87.365	91.539
-0.4	98.64	81.144	86.902	90.592
-0.3	97.48	78.416	84.940	88.921
-0.2	97.76	79.173	84.616	89.894
-0.1	97.20	77.449	83.211	88.602
0	96.96	74.381	80.776	85.768
0.1	96.08	74.063	79.933	84.222
0.2	94.64	71.936	78.656	82.459
0.3	94.12	70.038	77.304	81.510
0.4	92.72	67.472	74.633	78.861
0.5	91.04	60.545	70.475	75.087
0.6	84.76	58.377	68.146	72.394
0.7	73.96	52.082	63.925	68.306
0.8	54.84	41.648	57.476	61.998
0.9	25.48	20.565	39.718	46.469

E.2. PUISSANCE DES TESTS SUIVANT LA VALEUR DU PARAMÈTRE
 AUTO-RÉGRESSIF DES RÉSIDUS DE LA RELATION COINTÉGRÉE

Test inférieur de Student dans le cas "fallacieux" puis cointégré

Relation de cointégration avec constante et tendance

ρ_2	Puissance	Bonne date estimée	IC à ± 3 mois	IC à ± 6 mois
-0.9	98.28	79.813	87.669	91.658
-0.8	97.56	79.377	86.470	90.324
-0.7	97.80	79.264	86.379	90.468
-0.6	97.52	78.425	84.618	88.801
-0.5	96.88	77.250	83.279	87.655
-0.4	96.56	75.642	82.809	87.284
-0.3	95.80	74.238	81.003	85.137
-0.2	94.92	72.735	78.170	83.142
-0.1	93.40	71.906	76.831	82.569
0	92.80	70.690	76.595	82.197
0.1	92.36	69.770	75.184	79.559
0.2	90.40	68.274	74.204	78.496
0.3	89.04	63.836	69.990	74.975
0.4	86.76	61.411	68.189	73.490
0.5	80.84	61.851	69.472	73.281
0.6	77.32	55.510	62.908	67.047
0.7	64.16	47.506	56.920	61.285
0.8	45.92	38.415	53.312	56.884
0.9	20.92	17.591	33.462	40.535

ANNEXE E. TABLES DE PUISSANCE DES TESTS DE CHANGEMENT DU DEGRÉ DE
COINTÉGRATION

A l'approche de l'hypothèse nulle

Test supérieur de Wald dans le cas "fallacieux" puis cointégré

Relation de cointégration avec constante (suite)

ρ_2	Puissance	Bonne date estimée	IC à ± 3 mois	IC à ± 6 mois
0.91	25.56	22.066	41.627	48.668
0.92	22.56	21.809	39.894	49.291
0.93	20.32	13.780	33.466	41.536
0.94	17.12	15.888	26.169	32.946
0.95	16.16	11.139	25.744	32.428
0.96	14.84	6.469	14.825	25.067
0.97	14.52	2.479	11.569	20.108
0.98	13.92	2.011	8.045	14.079
0.99	11.28	1.418	5.318	11.347

Test supérieur de Wald dans le cas "fallacieux" puis cointégré

Relation de cointégration avec constante et tendance (suite)

ρ_2	Puissance	Bonne date estimée	IC à ± 3 mois	IC à ± 6 mois
0.91	19.76	16.599	33.603	40.080
0.92	17.60	12.500	29.547	35.683
0.93	14.64	17.213	30.873	36.064
0.94	13.96	8.883	21.204	28.656
0.95	13.48	8.605	20.771	29.969
0.96	12.32	5.844	15.584	24.025
0.97	11.80	1.356	6.102	12.204
0.98	9.52	1.261	7.145	10.086
0.99	10.00	0.800	2.800	9.200

E.2. PUISSANCE DES TESTS SUIVANT LA VALEUR DU PARAMÈTRE
 AUTO-RÉGRESSIF DES RÉSIDUS DE LA RELATION COINTÉGRÉE

Test inférieur de Student dans le cas "fallacieux" puis cointégré

Relation de cointégration avec constante (Suite)

ρ_2	Puissance	Bonne date estimée	IC à ± 3 mois	IC à ± 6 mois
0.91	23.04	19.097	36.631	43.923
0.92	20.48	16.211	29.296	38.280
0.93	17.48	10.526	25.171	33.637
0.94	15.24	6.299	14.960	20.733
0.95	13.60	5.000	13.530	18.529
0.96	11.56	3.114	7.958	16.262
0.97	11.88	1.010	7.744	13.804
0.98	11.28	1.064	3.901	8.865
0.99	9.12	0.877	3.509	7.018
1	9.68	0.413	7.023	10.328

Test inférieur de Student dans le cas "fallacieux" puis cointégré

Relation de cointégration avec constante et tendance (Suite)

ρ_2	Puissance	Bonne date estimée	IC à ± 3 mois	IC à ± 6 mois
0.91	18.04	13.082	27.938	35.477
0.92	16.32	7.843	22.303	28.675
0.93	13.16	11.854	23.709	30.397
0.94	12.28	4.235	15.310	22.475
0.95	11.60	4.483	13.448	20.690
0.96	10.60	2.264	10.942	18.112
0.97	11.08	1.083	5.415	10.108
0.98	8.12	0.000	5.418	8.374
0.99	8.84	0.905	2.714	8.596
1				

ANNEXE E. TABLES DE PUISSANCE DES TESTS DE CHANGEMENT DU DEGRÉ DE
COINTÉGRATION

Ha : "Cointégration puis absence de cointégration"

$$\left\{ k = 75, \quad \rho_2 = 1, \quad c_1 = 0, \quad c_2 = 0 \right\}$$

Test supérieur de Wald dans le cas cointégré puis "fallacieux"

Relation de cointégration avec constante

ρ_1	Puissance	Bonne date estimée	IC à ± 3 mois	IC à ± 6 mois
-0.9	95.24	12.852	62.915	74.297
-0.8	94.24	17.020	56.280	69.523
-0.7	92.40	16.580	57.317	71.429
-0.6	91.76	15.301	50.828	64.908
-0.5	90.68	14.954	50.287	64.049
-0.4	90.28	14.311	46.833	60.392
-0.3	88.28	13.140	46.533	59.309
-0.2	86.92	13.668	44.640	59.320
-0.1	84.72	13.409	43.155	57.414
0	82.36	12.190	40.408	54.978
0.1	80.44	12.233	40.478	53.656
0.2	77.76	11.831	38.373	52.724
0.3	74.88	10.630	34.135	47.918
0.4	71.08	9.454	30.163	43.331
0.5	66.28	10.682	29.933	41.400
0.6	59.36	7.682	24.730	35.511
0.7	50.44	5.630	17.922	26.725
0.8	32.44	4.316	12.208	18.621
0.9	18.04	3.548	7.318	10.201

E.2. PUISSANCE DES TESTS SUIVANT LA VALEUR DU PARAMÈTRE
 AUTO-RÉGRESSIF DES RÉSIDUS DE LA RELATION COINTÉGRÉE

Test supérieur de Wald dans le cas cointégré puis "fallacieux"

Relation de cointégration avec constante et tendance

ρ_1	Puissance	Bonne date estimée	IC à ± 3 mois	IC à ± 6 mois
-0.9	94.56	9.856	53.681	65.864
-0.8	92.04	11.951	46.805	60.059
-0.7	91.28	11.350	44.216	58.589
-0.6	87.56	10.598	40.703	53.950
-0.5	87.96	10.277	39.199	53.570
-0.4	86.80	10.553	37.327	51.198
-0.3	84.76	10.901	38.415	51.252
-0.2	81.52	9.323	35.034	48.528
-0.1	79.76	9.880	34.203	48.145
0	75.80	9.393	32.559	45.171
0.1	72.56	8.655	32.138	45.975
0.2	69.00	8.348	29.449	42.667
0.3	64.40	6.894	27.453	39.874
0.4	61.12	7.199	25.589	36.191
0.5	55.28	6.223	20.912	31.911
0.6	48.68	4.519	14.873	22.598
0.7	37.52	2.985	11.727	18.124
0.8	26.16	2.905	7.798	11.927
0.9	15.76	1.015	5.837	7.868

ANNEXE E. TABLES DE PUISSANCE DES TESTS DE CHANGEMENT DU DEGRÉ DE
COINTÉGRATION

Test inférieur de Student dans le cas cointégré puis "fallacieux"

Relation de cointégration avec constante

ρ_1	Puissance	Bonne date estimée	IC à ± 3 mois	IC à ± 6 mois
-0.9	94.92	41.949	65.765	69.691
-0.8	94.00	13.064	63.043	74.379
-0.7	92.72	17.362	56.938	70.087
-0.6	92.24	16.609	56.687	70.751
-0.5	90.92	15.872	51.388	64.658
-0.4	90.40	14.562	49.803	64.455
-0.3	88.88	14.381	46.814	60.485
-0.2	87.40	13.456	45.724	59.314
-0.1	85.16	13.822	44.440	59.176
0	82.76	11.890	41.036	55.488
0.1	81.24	13.621	43.682	57.961
0.2	78.64	12.055	38.606	53.152
0.3	75.88	11.965	40.326	53.767
0.4	73.00	10.754	34.685	48.180
0.5	68.04	10.935	30.629	42.798
0.6	61.56	9.425	30.356	43.452
0.7	52.00	7.862	25.472	36.779
0.8	33.88	3.778	13.459	20.542
0.9	18.08	4.646	9.955	15.043

E.2. PUISSANCE DES TESTS SUIVANT LA VALEUR DU PARAMÈTRE
 AUTO-RÉGRESSIF DES RÉSIDUS DE LA RELATION COINTÉGRÉE

Test inférieur de Student dans le cas cointégré puis "fallacieux"

Relation de cointégration avec constante et tendance

ρ_1	Puissance	Bonne date estimée	IC à ± 3 mois	IC à ± 6 mois
-0.9	93.92	10.349	53.748	65.844
-0.8	91.72	11.819	47.232	60.054
-0.7	90.80	11.278	44.228	58.633
-0.6	87.76	10.665	40.884	53.693
-0.5	88.12	10.213	39.491	53.927
-0.4	87.04	10.340	38.051	51.930
-0.3	84.8	11.274	39.010	51.888
-0.2	81.96	8.980	35.480	48.951
-0.1	80.24	9.721	34.346	48.005
0	76.56	9.195	33.230	45.613
0.1	73.64	8.365	31.994	45.953
0.2	70.80	8.192	30.000	43.333
0.3	66.20	7.130	28.580	41.148
0.4	64.08	7.303	25.093	36.642
0.5	57.76	6.233	21.191	32.549
0.6	51.40	4.591	16.109	24.512
0.7	40.20	4.478	13.037	20.699
0.8	27.28	3.519	9.384	14.223
0.9	17.12	1.402	6.076	8.412

ANNEXE E. TABLES DE PUISSANCE DES TESTS DE CHANGEMENT DU DEGRÉ DE
COINTÉGRATION

A l'approche de l'hypothèse nulle

Test supérieur de Wald dans le cas cointégré puis "fallacieux"

Relation de cointégration avec constante (suite)

ρ_1	Puissance	Bonne date estimée	IC à ± 3 mois	IC à ± 6 mois
0.91	16.80	1.905	6.428	10.951
0.92	15.48	2.584	6.976	10.593
0.93	14.64	1.913	5.464	7.649
0.94	13.96	1.719	6.305	9.745
0.95	13.08	1.835	6.116	9.480
0.96	10.48	7.252	11.070	13.742
0.97	11.12	5.036	11.152	14.390
0.98	9.00	3.111	12.443	15.109
0.99	9.76	4.508	10.655	12.705
1	8.76	8.676	14.612	17.809

Test supérieur de Wald dans le cas cointégré puis "fallacieux"

Relation de cointégration avec constante et tendance (suite)

ρ_1	Puissance	Bonne date estimée	IC à ± 3 mois	IC à ± 6 mois
0.91	14.76	0.813	3.794	5.691
0.92	12.88	0.621	3.417	5.592
0.93	13.80	2.029	5.508	8.118
0.94	11.80	1.695	4.068	4.407
0.95	11.16	0.717	5.376	7.526
0.96	11.32	1.413	6.006	8.126
0.97	11.24	0.712	3.204	5.695
0.98	8.52	3.286	7.979	9.856
0.99	7.56	4.233	7.407	10.052
1	7.08	3.390	12.43	13.560

E.2. PUISSANCE DES TESTS SUIVANT LA VALEUR DU PARAMÈTRE
 AUTO-RÉGRESSIF DES RÉSIDUS DE LA RELATION COINTÉGRÉE

Test inférieur de Student dans le cas cointégré puis "fallacieux"

Relation de cointégration avec constante (Suite)

ρ_1	Puissance	Bonne date estimée	IC à ± 3 mois	IC à ± 6 mois
0.91	17.28	2.083	8.101	12.498
0.92	15.76	3.553	8.631	13.200
0.93	14.96	2.139	6.149	9.090
0.94	14.92	2.681	9.115	14.476
0.95	14.20	2.535	7.323	10.704
0.96	11.00	7.636	13.090	16.000
0.97	11.64	7.560	17.526	19.589
0.98	9.28	3.879	15.516	18.533
0.99	9.64	5.394	11.619	14.939
1	9.44	10.169	21.61	25.847

Test inférieur de Student dans le cas cointégré puis "fallacieux"

Relation de cointégration avec constante et tendance (Suite)

ρ_1	Puissance	Bonne date estimée	IC à ± 3 mois	IC à ± 6 mois
0.91	15.00	1.867	4.800	7.200
0.92	13.72	0.875	5.540	8.456
0.93	14.68	2.997	7.083	10.079
0.94	11.92	2.349	5.370	6.042
0.95	11.84	0.338	6.082	7.434
0.96	12.16	1.645	5.922	7.896
0.97	11.52	0.694	4.165	7.636
0.98	8.72	3.211	8.257	10.551
0.99	7.80	5.128	9.232	12.310
1	7.36	6.522	17.391	21.739

**E.3 Puissance des tests suivant la valeur de la constante dans
la modélisation des résidus pour la sous-période cointégrée**

Ha : "Absence de cointégration puis cointégration"

$$\left\{ k = 75, \rho_1 = 1, \rho_2 = 0.8, c_1 = 0 \right\}$$

Test supérieur de Wald dans le cas "fallacieux" puis cointégré

Relation de cointégration avec constante

c_2	Puissance	Bonne date estimée	IC à ± 3 mois	IC à ± 6 mois
-2	46.80	61.453	73.419	76.153
-1	50.44	47.978	62.411	67.011
-0.5	51.76	45.131	59.042	64.529
0	53.52	45.740	60.015	64.052
0.5	52.16	45.706	59.509	63.651
1	49.16	47.030	61.187	65.418
2	46.84	61.998	73.100	76.771

Relation de cointégration avec constante et tendance

c_2	Puissance	Bonne date estimée	IC à ± 3 mois	IC à ± 6 mois
-2	34.72	63.249	68.204	70.161
-1	40.56	44.280	55.423	59.073
-0.5	41.76	38.027	52.203	57.184
0	46.76	38.152	52.010	57.314
0.5	42.56	37.218	51.034	55.734
1	39.08	43.193	56.191	60.285
2	33.56	60.548	69.130	71.394

E.3. PUISSANCE DES TESTS SUIVANT LA VALEUR DE LA CONSTANTE DANS LA MODÉLISATION DES RÉSIDUS POUR LA SOUS-PÉRIODE COINTÉGRÉE

Test inférieur de Student dans le cas "fallacieux" puis cointégré

Relation de cointégration avec constante

c_2	Puissance	Bonne date estimée	IC à ± 3 mois	IC à ± 6 mois
-2	48.56	60.544	72.488	75.453
-1	51.56	47.091	62.064	66.796
-0.5	53.72	43.559	58.897	64.557
0	54.80	43.869	59.344	64.162
0.5	54.36	43.046	57.175	61.810
1	50.68	45.856	60.457	64.720
2	48.48	60.231	72.196	75.744

Relation de cointégration avec constante et tendance

c_2	Puissance	Bonne date estimée	IC à ± 3 mois	IC à ± 6 mois
-2	34.72	63.249	68.204	70.161
-1	40.56	44.280	55.423	59.073
-0.5	41.76	38.027	52.203	57.184
0	46.76	38.152	52.010	57.314
0.5	42.56	37.218	51.034	55.734
1	39.08	43.193	56.191	60.285
2	33.56	60.548	69.130	71.394

ANNEXE E. TABLES DE PUISSANCE DES TESTS DE CHANGEMENT DU DEGRÉ DE
COINTÉGRATION

Ha : "Cointégration puis absence de cointégration"

$$\left\{ k = 75, \rho_1 = 0.8, \rho_2 = 1, c_2 = 0 \right\}$$

Test supérieur de Wald dans le cas cointégré puis "fallacieux"

Relation de cointégration avec constante

c_1	Puissance	Bonne date estimée	IC à ± 3 mois	IC à ± 6 mois
-2	25.56	40.532	61.659	66.978
-1	31.08	21.493	37.453	42.472
-0.5	32.52	6.642	16.974	22.632
0	34.44	3.833	10.686	16.726
0.5	32.24	8.313	18.238	25.558
1	30.52	19.528	37.222	41.807
2	26.08	41.411	62.730	66.103

Test supérieur de Wald dans le cas cointégré puis "fallacieux"

Relation de cointégration avec constante et tendance

c_1	Puissance	Bonne date estimée	IC à ± 3 mois	IC à ± 6 mois
-2	26.36	42.185	57.815	62.063
-1	26.28	17.047	28.158	30.136
-0.5	29.04	5.510	13.500	17.079
0	27.04	2.663	8.285	11.984
0.5	26.12	5.360	11.485	15.620
1	25.52	15.831	27.116	30.877
2	27.36	41.667	57.456	60.672

E.3. PUISSANCE DES TESTS SUIVANT LA VALEUR DE LA CONSTANTE DANS LA MODÉLISATION DES RÉSIDUS POUR LA SOUS-PÉRIODE COINTÉGRÉE

Test inférieur de Student dans le cas cointégré puis "fallacieux"

Relation de cointégration avec constante

c_1	Puissance	Bonne date estimée	IC à ± 3 mois	IC à ± 6 mois
-2	28.20	43.404	66.384	71.206
-1	34.36	17.346	38.183	44.353
-0.5	32.44	19.359	37.608	43.035
0	36.60	4.809	12.896	19.343
0.5	34.24	9.229	21.612	28.154
1	34.16	20.609	39.579	44.966
2	29.56	20.072	40.262	45.280

Test inférieur de Student dans le cas cointégré puis "fallacieux"

Relation de cointégration avec constante et tendance

c_1	Puissance	Bonne date estimée	IC à ± 3 mois	IC à ± 6 mois
-2	29.48	42.605	60.651	64.993
-1	29.04	17.906	30.991	33.745
-0.5	30.64	6.136	15.798	20.106
0	29.16	2.881	9.740	14.542
0.5	27.28	6.305	13.784	18.184
1	28.92	16.736	30.015	34.719
2	31.44	41.603	60.051	64.249

E.4 Puissance des tests suivant la véritable date de rupture

Ha : "Absence de cointégration puis cointégration"

$$\left\{ \rho_1 = 1, \rho_2 = 0.8, c_1 = 0, c_2 = 0 \right\}$$

Test supérieur de Wald dans le cas "fallacieux" puis cointégré

Relation de cointégration avec constante

<i>k</i>	Puissance	Bonne date estimée	IC à ± 3 mois	IC à ± 6 mois
25	68.40	27.602	50.175	66.023
50	63.56	33.543	50.159	56.704
75	52.96	42.900	58.233	62.991
100	45.84	57.068	70.506	73.823
125	40.88	67.906	81.311	83.953

Test supérieur de Wald dans le cas "fallacieux" puis cointégré

Relation de cointégration avec constante et tendance

<i>k</i>	Puissance	Bonne date estimée	IC à ± 3 mois	IC à ± 6 mois
25	61.76	23.446	46.180	61.141
50	51.88	31.457	44.950	50.502
75	43.20	37.407	50.186	53.890
100	35.08	50.399	64.196	67.616
125	33.88	66.824	80.164	82.289

Test inférieur de Student dans le cas "fallacieux" puis cointégré

Relation de cointégration avec constante

k	Puissance	Bonne date estimée	IC à ± 3 mois	IC à ± 6 mois
25	70.80	25.932	47.515	62.317
50	65.20	32.086	49.203	56.320
75	54.56	41.276	57.992	63.490
100	45.84	53.490	67.102	71.204
125	33.56	62.574	80.094	83.552

Test inférieur de Student dans le cas "fallacieux" puis cointégré

Relation de cointégration avec constante et tendance

k	Puissance	Bonne date estimée	IC à ± 3 mois	IC à ± 6 mois
25	65.24	20.478	40.528	53.404
50	54.88	29.883	43.440	49.563
75	45.32	33.892	48.102	52.956
100	36.16	46.350	60.620	65.598
125	27.16	62.887	77.909	81.444

ANNEXE E. TABLES DE PUISSANCE DES TESTS DE CHANGEMENT DU DEGRÉ DE
COINTÉGRATION

Ha : "Cointégration puis absence de cointégration"

$$\left\{ \rho_1 = 0.8, \rho_2 = 1, c_1 = 0, c_2 = 0 \right\}$$

Test supérieur de Wald dans le cas cointégré puis "fallacieux"

Relation de cointégration avec constante

<i>k</i>	Puissance	Bonne date estimée	IC à ± 3 mois	IC à ± 6 mois
25	13.48	3.858	12.760	19.585
50	22.60	4.248	9.204	14.160
75	33.44	4.426	13.038	19.976
100	49.60	4.194	12.903	20.483
125	70.12	6.617	22.876	38.391

Test supérieur de Wald dans le cas cointégré puis "fallacieux"

Relation de cointégration avec constante et tendance

<i>k</i>	Puissance	Bonne date estimée	IC à ± 3 mois	IC à ± 6 mois
25	12.56	5.096	9.872	14.011
50	18.36	3.050	6.755	9.371
75	27.00	2.519	7.260	12.001
100	39.80	3.819	9.348	13.972
125	59.28	4.723	16.463	29.013

Test inférieur de Student dans le cas cointégré puis "fallacieux"

Relation de cointégration avec constante

k	Puissance	Bonne date estimée	IC à ± 3 mois	IC à ± 6 mois
25	15.12	5.291	15.609	26.190
50	24.28	5.272	10.873	16.968
75	36.24	4.636	14.901	22.516
100	51.20	4.688	14.611	22.580
125	69.76	6.537	22.133	37.328

Test inférieur de Student dans le cas cointégré puis "fallacieux"

Relation de cointégration avec constante et tendance

k	Puissance	Bonne date estimée	IC à ± 3 mois	IC à ± 6 mois
25	13.28	5.723	11.445	16.565
50	19.68	3.252	7.724	10.774
75	28.40	2.676	8.732	14.647
100	41.12	3.696	10.700	16.829
125	59.72	5.291	18.620	31.413

Annexe F

Causalité au sens de Granger

Les contributions les plus connues de Granger (1969 [62] et 1980 [63]) à l'économétrie sont sans nul doute celles liées à l'analyse de la causalité. Ses travaux ont en effet permis de définir et de formaliser de façon simple des notions relatives à la causalité, au travers de quatre définitions : la causalité (au sens de Granger), la répercussion dans les deux sens ("feedback"), la causalité instantanée et la causalité décalée. Ces définitions sont rappelées dans le premier paragraphe. Leur application est évoquée dans le deuxième paragraphe.

Enfin, en 1988 [64], Granger montre que si deux séries intégrées à l'ordre 1 sont reliées par une relation de cointégration, alors il doit exister une causalité au sens de Granger dans au moins une direction. Cette théorie est développée dans le dernier paragraphe de cette section.

F.1 Définitions relatives à la causalité

L'ensemble des définitions suivantes s'applique au cas de séries stationnaires.

La causalité au sens de Granger est définie par le fait que " x_t cause y_t " si la prévision de la série y_t est améliorée quand toute l'information disponible est prise en compte (i.e. les valeurs présentes et passées de x_t et y_t) et non la seule information relative au présent et au passé de y_t . Autrement dit, " x_t cause y_t " si la connaissance du passé de la série de x_t permet de réduire la variance de l'erreur de prévision de y . Cette définition, présente dans l'article de 1969 [62] a été axiomatisée dans l'article de 1980 [63], par l'utilisation de probabilités conditionnelles.

Definition 3 : Causalité au sens de Granger. Soit $\Omega_{t-1}(y)$ l'ensemble des valeurs passées de y_t , avec $\Omega_{t-1}(y) = \{y_s, s \leq t-1\}$. Désignons par le sigle "MSE" l'erreur quadratique moyenne, c'est-à-dire la moyenne des carrés des résidus. " x_t cause y_t " si

$$MSE [y_t | \Omega_{t-1}(y)] > MSE [y_t | \Omega_{t-1}(y), \Omega_{t-1}(x)]$$

La définition suivante de non causalité au sens de Granger peut aussi être considérée. " x_t ne cause pas y_t " si prendre en compte les valeurs passées de x_t ne modifie pas la prévision de la série y_t .

Definition 4 : Non causalité au sens de Granger. " x_t ne cause pas y_t " au sens de Granger si

$$MSE [y_t | \Omega_{t-1}(y)] = MSE [y_t | \Omega_{t-1}(y), \Omega_{t-1}(x)]$$

L'idée de "feedback" correspond à l'existence d'une causalité dans les deux sens, où " x_t cause y_t " et " y_t cause x_t ". Chaque variable a des répercussions sur l'autre variable.

Definition 5 : "Feedback". Si

$$MSE [y_t | \Omega_{t-1}(y)] > MSE [y_t | \Omega_{t-1}(y), \Omega_{t-1}(x)]$$

et

$$MSE [x_t | \Omega_{t-1}(x)] > MSE [x_t | \Omega_{t-1}(x), \Omega_{t-1}(y)]$$

" x_t cause y_t " et " y_t cause x_t ", il y a causalité dans les deux sens.

Autre concept important, il est possible d'avoir une "causalité instantanée", alors que le fait de dire " x_t cause y_t " (au sens de Granger) évoque implicitement l'antériorité de la formation de la série x_t par rapport à celle de y_t : la cause doit précéder l'effet. Cette causalité instantanée provient généralement de la trop faible fréquence d'échantillonnage des données (l'intervalle de temps entre deux données collectées est trop grand pour capter de façon satisfaisante les relations

causales) ou de l'aggrégation de données temporelles. Une autre raison à la causalité instantanée provient de l'incomplétude de l'ensemble de l'information disponible : on a testé si x_t cause y_t , mais il existe une variable w_{t-1} qui cause à la fois y_t et x_t et n'appartient pas à l'ensemble de l'information disponible. Enfin, plus rarement, les variables réagissent véritablement dans un laps de temps non mesurable.

Definition 6 : *Causalité instantanée.* Si

$$MSE [y_t | \Omega_{t-1}(y), \Omega_{t-1}(x)] > MSE [y_t | \Omega_{t-1}(y), \Omega_t(x)]$$

alors " x_t cause y_t instantanément" : la prévision de y_t est meilleure si la valeur présente de x_t est prise en compte au lieu d'être omise.

Si " x_t cause y_t ", il est possible que cette causalité n'apparaisse qu'avec un certain décalage, i.e. un nombre de retard m . Par conséquent, les valeurs de $x_{t-1}, \dots, x_{t-m+1}$ n'amélioreront pas la prévision de y_t .

Definition 7 : *Causalité décalée.* Si " x_t cause y_t ", il existe un nombre entier m correspondant au décalage entre y_t et sa cause au sens de Granger x_t : m est la dernière valeur de k telle que

$$MSE [y_t | \Omega_{t-1}(y), \Omega_{t-(k+1)}(x)] > MSE [y_t | \Omega_{t-1}(y), \Omega_{t-k}(x)]$$

F.2 Application des tests de causalité à un modèle bivarié

Soit y_t et x_t deux séries stationnaires de moyenne nulle. Pour tester la causalité dans chacun des sens, construisons un modèle de la forme :

$$\begin{aligned} y_t &= \sum_{j=1}^l a_j y_{t-j} + \sum_{j=1}^l b_j x_{t-j} + e_t \\ x_t &= \sum_{j=1}^l c_j x_{t-j} + \sum_{j=1}^l d_j y_{t-j} + u_t \end{aligned}$$

où e_t et u_t sont des bruits blancs non corrélés et l est le nombre de retards à prendre en compte

En suivant les définitions précédentes, " x_t cause y_t " si l'ensemble des paramètres b_j est non nul ; de façon inverse " y_t cause x_t " si l'ensemble des paramètres d_j est différent de 0. Si l'ensemble des b_j et des d_j est différent de 0, il existe une causalité dans les deux sens, un "feedback". Pour tester la causalité instantanée, le modèle suivant, plus général, sera utilisé :

$$\begin{aligned}
 y_t + b_0 x_t &= \sum_{j=1}^l a_j y_{t-j} + \sum_{j=1}^l b_j x_{t-j} + e_t \\
 x_t + d_0 x_t &= \sum_{j=1}^l c_j x_{t-j} + \sum_{j=1}^l d_j y_{t-j} + u_t
 \end{aligned}$$

L'article de 1980 évoque aussi l'importance de l'utilisation de données ultérieures à l'échantillon pour confirmer les relations de causalité mises en évidence. Cette prévision de données ultérieures, et l'évaluation des performances, sont mises en oeuvre dans l'article de Ashley, Granger et Schmalensee (1980 [10]).

F.3 Liens entre cointégration et causalité

Selon Granger [64], l'existence d'une relation de cointégration entre deux variables implique l'existence d'une causalité au sens de Granger dans au moins une direction. En effet, si les séries intégrées à l'ordre 1 x_t et y_t sont cointégrées (c'est-à-dire telles que $y_t - Ax_t = z_t$, où z_t est stationnaire), elles peuvent être générées par un modèle à correction d'erreur de la forme :

$$\begin{aligned}
 \Delta y_t &= \gamma_1 z_{t-1} + \sum_{i=1}^l f_i \Delta x_{t-i} + \sum_{i=1}^m g_i \Delta y_{t-i} + e_{1t} \\
 \Delta x_t &= \gamma_2 z_{t-1} + \sum_{i=1}^n h_i \Delta y_{t-i} + \sum_{i=1}^o k_i \Delta x_{t-i} + e_{2t}
 \end{aligned}$$

avec $\gamma_1 * \gamma_2 \neq 0$, $e_{1t} \sim MA(q_1)$ et $e_{2t} \sim MA(q_2)$

D'après ces équations, les variations des variables x_t et y_t à court terme dépendent en partie de la valeur passée de z_t , la déviation par rapport à la situation d'équilibre de long terme défini par l'équation $y_t - Ax_t = 0$. En conséquence, Δy_t et/ou Δx_t sont causées au sens de Granger par z_{t-1} , lui même fonction de y_{t-1} et x_{t-1} , ce qui implique que x_t cause y_t et/ou que y_t cause

x_t , au sens de Granger. S'il peut apparaître surprenant que l'existence d'une dynamique de long terme ait pour conséquence une causalité (associée à la prédictivité de la série à court terme), la causalité est nécessaire à l'existence de cette dynamique de court terme.

Dans le modèle à correction d'erreur, deux sources peuvent amener à conclure que " x_t peut causer y_t " au sens de Granger :

- si $\gamma_1 \neq 0$, " x_t peut causer y_t " par le biais de la variable z_{t-1} ,
- si les termes Δx_t sont présents dans l'équation de court terme (i.e. que les paramètres f_i ne sont pas tous nuls).

Examinons la forme de la causalité introduite par la variable z_{t-1} dans le modèle simple

$$\begin{aligned} y_t &= Aq_t + u_{1t} \\ x_t &= q_t + u_{2t} \end{aligned}$$

où q_t est une série intégrée à l'ordre 1, plus précisément une marche aléatoire telle que $\Delta q_t = a_t$. a_t , u_{1t} et u_{2t} sont des bruits blancs de moyenne nulle. On a

$$\begin{aligned} z_t &= y_t - Ax_t = Aq_t + u_{1t} - A(q_t + u_{2t}) = u_{1t} - Au_{2t} \\ \text{et } \Delta y_t &= A\Delta q_t + \Delta u_{1t} = Aa_t + \Delta u_{1t} = -u_{1t-1} + Aa_t + u_{1t} \end{aligned}$$

Les deux derniers termes de cette équation correspondent à l'erreur de prévision à un pas, quand la prévision a été réalisée à la date $t - 1$ en prenant en compte les valeurs de x_t et y_t jusqu'à cette date. Le premier terme $-u_{1t-1}$ est égal à la partie de Δy_t que l'on peut prévoir. Non observable directement, elle est corrélée avec z_{t-1} sauf si ce processus est de variance nulle. Si la variable z_t n'est pas utilisée dans la modélisation, alors u_{1t-1} dépend uniquement d'une somme des variations passées de y_t , cette somme incluant la somme des valeurs passées de a_t , conduisant à un terme de nuisance ayant une variance parfois importante. u_{1t-1} sera alors peu corrélé avec la somme des variables Δy_t passées. L'oubli de z_t conduit à une mauvaise spécification du modèle et l'oubli potentiel de valeurs passées de x_t dans la prévision de y_t .

Annexe G

Historique des faits marquants ayant une incidence sur le trafic de la RATP

1989

- mai - juin : Seconde étape de l'interconnexion de la ligne A du RER avec la ligne SNCF de Cergy - Pontoise : la ligne A comporte désormais une cinquième branche qui permet aux habitants du secteur concerné de traverser Paris et d'atteindre la banlieue Est de la capitale sans avoir à emprunter de correspondance. L'intervalle des trains sur cette ligne a été ramené de 2 minutes et 30 secondes à 2 minutes pendant la période de surpointe, dans le sens principal de la charge, grâce à la mise en service du SACEM (Système d'Aide à la Conduite, l'Exploitation et la Maintenance).
- juin : Création du titre "Paris-Visite", en remplacement de l'ancien "Paris-Sésame". Ciblé sur les touristes, valable 3 ou 5 jours sur les réseaux de la RATP et de la SNCF, il comporte un coupon avec accès éventuel aux aéroports d'Orly et de Roissy-Charles de Gaulle sur les dessertes exploitées par la RATP et la SNCF.

Grèves RATP.

- juillet : Grèves RATP.

– septembre : Augmentation des tarifs de 4.05% pour la RATP.

– novembre, décembre : Grèves RATP.

1990

– janvier : Achèvement de l'interconnexion Ouest de la ligne A du RER.

Grèves RATP.

– février, avril : Grèves RATP.

– mai : Prolongement de la ligne D du RER (SNCF) jusqu'à Orry la Ville.

Grèves RATP.

– juin : Autrement Bus 94.

Grèves RATP.

– août : Augmentation des tarifs de la RATP de 4.50%.

– septembre : Interruption de service pour renouvellement de l'ensemble des équipements.

– octobre : Achèvement de l'interconnexion de la ligne D du RER : tous les trains desservent la gare de Châtelet-les-Halles.

Menaces de guerre dans le Golfe Persique.

Grèves RATP.

– novembre, décembre : Grèves RATP.

1991

– janvier : Guerre du Golfe et plan Vigipirate.

Grèves RATP.

– avril : Grèves RATP.

– mai : Autrement bus 93.

Grèves RATP.

– juin : Extension du nombre de zones de validité des titres de la gamme Orange de 5 à 8 zones. La carte Orange couvre désormais la totalité de l'Ile-de-France, et la nouvelle zone de compétence du Syndicat des Transports Parisiens ("STP") correspond à la totalité

ANNEXE G. HISTORIQUE DES FAITS MARQUANTS AYANT UNE INCIDENCE SUR
LE TRAFIC DE LA RATP

de l'Ile-de-France. Chaque département de la Grande Couronne est représenté au conseil d'administration du STP.

Grèves RATP.

- août : Augmentation des tarifs de 5.80% pour la RATP.

Suppression de la première classe sur le réseau Métro.

- septembre : Ouverture de la gare "Porte de Clichy" sur la ligne C du RER SNCF, gare en correspondance avec la ligne 13 du métro.

- octobre : Mise en service de la liaison "Orlyval" entre Anthony et l'aéroport d'Orly.

Mise en service du nouveau funiculaire.

Grèves RATP.

- novembre, décembre : Grèves RATP.

1992

- Mise en service de SIEL (Système d'Information sur l'Exploitation de la Ligne) sur la ligne A du RER : les voyageurs disposent ainsi d'une information en temps réel sur les cinq prochains trains et l'état du trafic en cas de perturbation.

- Sortie du premier baromètre image - satisfaction, qui mesure l'appréciation du client et s'inscrit dans le développement d'une image de qualité de service.

- 6, 30 et 31 mars : Grèves RATP.

- avril : Autrement Bus 92 Nord 1^{ère} phase.

La ligne 1 du métro est prolongée vers l'ouest avec l'ouverture de deux stations "Esplanade de la Défense" et "Grande Arche de la Défense" (zone 3); elle bénéficie ainsi d'une nouvelle correspondance avec la ligne A du RER ainsi qu'avec la ligne SNCF vers Versailles ou Saint-Nom la Bretèche.

Prolongement vers l'Est de la ligne A du RER jusqu'à "Chessy-Marne la Vallée" (zone 5), permettant ainsi la desserte du parc d'attractions Eurodisney.

- juillet : Autrement Bus 93 Nord 1^{ère} phase.

1^{ère} étape de la mise en service de la ligne 1 du tramway (T1) en site propre entre "Bobigny Pablo Picasso" et "La Courneuve - 8 mai 1945".

Grèves RATP.

- août : Grèves RATP.
- septembre : Augmentation des tarifs de 5.80% pour la RATP.

Grèves RATP.

- octobre : Autrement Bus 92 Nord 2^{ème} phase. Grèves RATP.
- 17 - 20 et 23 novembre : Grèves RATP.
- décembre : Autrement Bus 93 Nord 2^{ème} phase.

Ouverture de la gare Bussy-Saint-Georges sur la ligne A du RER.

Le 7 et 16, Grèves RATP.

2^{ème} étape de la mise en service de la ligne 1 du tramway (T1), entre "La Courneuve - 8 mai 1945" et "Saint Denis - Gare RER". La ligne T1 dessert 21 stations, dont quatre possèdent des correspondances avec des infrastructures radiales : la SNCF (gare de Saint-Denis), le métro ("Saint-Denis Basilique" ligne 13, "La Courneuve - 8 mai 1945" ligne 7 et "Bobigny Préfecture ligne 5).

Création de la ligne Roissy-Bus.

1993

- février : Reprise de l'exploitation de Orlyval par la RATP.

Nouvelles clés de partage des recettes Carte Orange.

- 10, 11, 12 et 16 mars : Grèves RATP sur le réseau ferré.
- mai : Augmentation des tarifs de 5.8% pour la RATP.

Grèves RATP.

- juin : Autrement Bus Sud 1^{ère} phase.

Grèves RATP.

- 20 - 24 septembre : Grèves RATP.
- octobre : Autrement Bus Sud 2^{ème} phase.

Mise en service de la première ligne d'autobus en site propre continu dénommée "Trans Val de Marne" (TVM). Cette nouvelle ligne relie entre elles 5 communes du Val de Marne

ANNEXE G. HISTORIQUE DES FAITS MARQUANTS AYANT UNE INCIDENCE SUR
LE TRAFIC DE LA RATP

(Chevilly-Larue, Thiais, Choisy-le-Roi, Créteil et Saint-Maur) entre le Marché International de Rungis et la gare RER de Saint-Maur-Créteil. Elle assure la correspondance avec plusieurs lignes de métro, RER, SNCF ou de bus. Des écrans vidéo informent en temps réel les voyageurs sur les temps d'attente.

Le 12, grèves RATP.

- 10, 18 et 26 novembre, 1, 2 et 9 décembre : Grèves RATP.

1994

- Mise en service de SIEL (Système d'Information sur l'Exploitation de la Ligne) sur la ligne B du RER.
- Sur le réseau bus, première mise en place et expérimentation d'un nouveau système à l'exploitation des autobus appelé Altair. Ce système de radiolocalisation, en transmettant toutes les 20 secondes la position des autobus, permet aux voyageurs d'être informés au point d'arrêt en temps réel sur les temps d'attente et les perturbations éventuelles et au régulateur de faire, en temps réel, des manoeuvres de régulations. En outre est mis en place un partenariat actif avec la Préfecture de Police et la Mairie de Paris consistant à surveiller les itinéraires des lignes d'autobus afin d'améliorer la régularité et la vitesse commerciale.
- Publication du nouveau schéma directeur de la région d'Ile-de-France ("SDRIF").
Lancement et financement par le Syndicat des Transports Parisiens d'une politique de renforcement de la présence humaine sur les réseaux desservant les quartiers difficiles.
- mars, avril : Grèves RATP.
- mai : Prolongement de la ligne A du RER SNCF avec ouverture de 2 nouvelles gares : "Cergy le Haut" et "Neuville-Université".
- juin : Grèves RATP.
- septembre : Augmentation des tarifs de 5.20% pour la RATP.
- octobre : Grèves RATP.
- novembre : Prolongement de la ligne B du RER à "Aéroport Roissy Charles de Gaulle 2 - TGV".
Mise en service du TGV Eurostar.

1995

- Dans le cadre de l'amélioration de la qualité de service offerte aux voyageurs, des standards de service et des indicateurs de mesure sont définis pour le métro, dans une logique de service client (régularité, disponibilité des escaliers mécaniques et des distributeurs automatiques de billets, accueil aux guichets, netteté des stations). Ces indicateurs sont mesurés tous les trimestres à l'exception de l'accueil (tous les semestres).
- Amélioration de l'offre de transport au travers de renforcement des flancs de pointe sur les lignes 2, 3, 6, 7 et 8 du métro.
- mars : Autrement Bus Boucles de la Marne (restructuration qui a concerné 10 communes et 26 lignes d'autobus de la banlieue Est).
Restructuration de 3 lignes puissantes de la banlieue Nord.
Grèves RATP.
- avril : Grèves RATP.
- mai : Ouverture de la liaison SNCF "la Verrière-Saint Quentin" - "la Défense".
Grèves RATP.
- juin : Fête de la Musique.
Grèves RATP.
- 25 juillet : Attentat à la station Saint-Michel (ligne B du RER) : 8 morts et 117 blessés.
- août : Augmentation des tarifs de 4.4% pour la RATP.
Réforme tarifaire en vue de simplifier la structure tarifaire : un seul ticket permet au voyageur d'avoir accès au réseau banlieue, au métro et au RER dans Paris intra-muros.
Attentats avec report des transports en commun souterrains vers les transports de surface.
- septembre : Interconnexion de la ligne D avec ouverture du tronçon central Châtelet - Gare de Lyon.
Attentats et mise en place du plan Vigipirate.
- octobre : Montée en puissance du plan Vigipirate.
17 octobre : Explosion d'une bombe dans le RER C entre les gares Musée-d'Orsay et Saint-Michel (29 blessés).

ANNEXE G. HISTORIQUE DES FAITS MARQUANTS AYANT UNE INCIDENCE SUR
LE TRAFIC DE LA RATP

Grèves RATP.

- 24 novembre - 16 décembre : Arrêt complet des bus, métros, RER et trains. Mise en place d'un plan minimum de dépannage pour permettre aux Franciliens non motorisés de se rendre de banlieue à Paris.

1996

- Mise en place du système contractuel d'alarme sociale. Elaboré avec les organisations syndicales, cet accord indique explicitement que le recours à la grève est un recours ultime, et qu'il faut privilégier toute forme de prévention.
- janvier : Allègement du plan Vigipirate.
- avril, mai, juin : Grèves RATP.
- juillet : Augmentation des tarifs de 5.20% pour la RATP.
- septembre, octobre : Grèves RATP.
- 3 décembre : Explosion d'une bombe à la gare de Port-Royal, dans une voiture du RER B (3 morts, 2 blessés très graves, 6 graves et près de 80 blessés sérieux ou légers). Réactivation totale du plan Vigipirate.

1997

- Inauguration du Tram Val-de Seine (ligne T2 du tramway).
- avril : Aménagement du couloir bus en site propre entre Concorde et Châtelet.
- mai : Lancement du Plan de Déplacement Urbain ("PDU").
- août : Journées Mondiales de la Jeunesse qui ont donné lieu à une fréquentation supplémentaire sur le métro et le RER
- septembre : Refonte du réseau Noctambus avec 3 créations de lignes sur Paris et extensions en banlieue.
- 1^{er} octobre : Octroi de la gratuité des transports sur tous les réseaux d'Ile-de-France lors du pic de pollution.

1998

-
- Elaboration et mise en place par le Syndicat des Transports Parisiens du chèque mobilité réservé à certains demandeurs d'emplois en Ile-de-France.
 - Lancement d'une tarification en faveur des jeunes scolaires et étudiants : la carte Imagine R.
 - janvier : Ouverture de la gare "La Plaine Stade de France" sur la ligne B du RER SNCF. Prolongement de la ligne 57 de bus jusqu'à la Porte de Bagnolet.
 - juin : Coupe du Monde de football. Celle-ci n'a pas eu d'influence sur les déplacements en transport en commun, compte-tenu d'une perte de clientèle touristique.
Restructuration du réseau bus nord parisien, avec amélioration de la desserte du Stade de France et de l'université Paris XIII.
 - octobre : Ouverture de la ligne 14 du métro, appelée "METEOR", avec mise en service de 7 stations ("Bibliothèque", "Cour Saint Emilion", "Bercy", "Gare de Lyon", "Châtelet-Halles", "Pyramide" et "Madeleine").
Mouvement social sur les bus le 6.

1999

- 2 et 3 juin : Grèves RATP. 7 juin : Cessation d'activité à la RATP suite au décès d'un agent de la RATP.
- juillet : Mise en service de la ligne E du RER SNCF entre "Haussman Saint-Lazare" et "Chelles-Gournay".
- août : Mise en service de la nouvelle branche de la ligne E du RER SNCF entre "Haussman Saint-Lazare" et "Villiers-sur-Marne"
- septembre : Journée sans voiture le 22.
Suppression de la première classe dans le RER et les trains de banlieue, tarification avec un seul billet sur les lignes d'autobus de la RATP et sur le tramway.
Prolongement de la ligne 88 de "Quai André Citroën" à "Hôpital Georges Pompidou".
Grève sur la ligne B du RER.
- octobre : La ligne d'autobus PC est exploitée désormais en 3 tronçons indépendants afin d'en améliorer la vitesse et la régularité, et en site protégé sur le tronçon sud.

ANNEXE G. HISTORIQUE DES FAITS MARQUANTS AYANT UNE INCIDENCE SUR
LE TRAFIC DE LA RATP

- décembre : Interruption du trafic sur les réseaux due à la tempête du 26. Gratuité des transports en commun le 31.

2000

- Lancement de deux formules d'abonnement "illimité" donnant "libre accès" aux salles de cinéma. Offertes dans plus des deux tiers des écrans parisiens, ces formules présentent près de la moitié des établissements et regroupent 70% des fauteuils.
- février : Grève RATP.
- avril : Restructuration du réseau d'autobus dans le sud des Hauts-de-Seine et dans le nord de l'Essonne.
- mai : Modification du stationnement payant : la durée du stationnement résidentiel a été portée à 7 jours consécutifs au lieu de 24 heures consécutives.
- juillet : Dézonage de la carte Imagine R pendant les petites vacances scolaires.
- août : Prolongement de la ligne C du RER SNCF jusqu'à Pontoise.
- septembre : Inauguration des aménagements réalisés sur les quais Rive Gauche et sur le quai François Mauriac.
Du 26 au 28, perturbations du trafic ferroviaire provoquées par les arrêts de travail sur le réseau de la SNCF Ile-de-France et sur les différents réseaux de la RATP
- octobre : Mondial de l'automobile jusqu'au 15 : métro et autobus sont renforcés les week-ends et un combiné "ticket de métro + salon " est vendu dans toutes les stations de métro.
- novembre : Inauguration de la ligne d'autobus PC2 en site protégé entre "Porte d'Italie" et "Porte de la Villette".
- décembre : Mise en service de la ligne 14 du métro et de la ligne C du RER SNCF à la station "Bibliothèque François Mitterrand".
Mise en service de deux lignes SNCF d'autobus de nuit en correspondance à Châtelet avec le réseau Noctambus.
Gratuité des transports le 31 à partir de 17 heures.

Le Syndicat des Transports Parisiens ("STP") devient le Syndicat des Transports d'Ile-de-France ("STIF").

2001

- Renforcement de l'offre de transport durant les flancs de pointe descendants d'après midi, en soirée ainsi que les samedis et dimanches sur la ligne 11 du métro. Création d'une ligne d'autobus de banlieue et de sept lignes sur les services communaux de banlieue.
- 1 et 8 février : Grèves RATP.
- mars : Lancement de "Transition", nouvel abonnement forfaitaire long, dézonné le week-end, les jours fériés et les petites vacances scolaires. Du 29 au 31, fortes grèves SNCF.
- avril : Fortes crues de la Seine qui ont entraîné la fermeture des voies sur berges pour une durée cumulée allant jusqu'à 42 jours pour le tronçon Bercy - Austerlitz.
Poursuite des grèves SNCF jusqu'au 13.
Mise en service de la gare "Val d'Europe" sur la ligne A du RER.
Le 25, incendie à la gare de Lyon. Le 28, grève RATP.
- juin : Prolongement de la ligne 302 depuis "Porte de la Chapelle" jusqu'à la Gare du Nord, afin de renforcer les liaisons entre la banlieue Nord et le nord de Paris.
Mise en service de la nouvelle gare routière à la Défense.
Fermeture le 25 de la ligne T1.
- juillet : Fermeture des voies sur berges.
Mise en service de 7 kilomètres de "couloirs protégés" pour autobus, taxis et cyclistes.
- septembre : Remise en service de la ligne T1.
Prolongement de la ligne 139 à la gare de Saint-Ouen (ligne C du RER SNCF).
Attentats aux Etats -Unis et remise en place de Vigipirate.
- octobre : Pas de Salon de l'auto.
Le 16, grève sur tous les réseaux.
- novembre : En raison de l'indisponibilité des Terminaux Points de Vente dans environ 200 stations, de nombreuses lignes de péage sont mises hors tension le 19, les données ont été corrigées.

ANNEXE G. HISTORIQUE DES FAITS MARQUANTS AYANT UNE INCIDENCE SUR
LE TRAFIC DE LA RATP

Inauguration du pôle multimodal de la gare du Nord.

Le 29, l'exploitation est perturbée par un accident de voyageur sur la ligne A du RER et la rupture d'un rail à Châtelet-les-Halles.

- décembre : Inauguration du "Coeur Transport", pôle multimodal de la Défense.

Gratuité des transports publics d'Ile-de-France à partir de 17 heures le 31 décembre.

2002

- Renforcement de l'offre de service du métro durant les flancs de pointe descendants d'après-midi, en soirée, ainsi que les samedis et dimanches sur les lignes 4 du métro (mi-janvier), 13 (fin avril) et 6 (début mai).

Renforcement de l'offre de transport sur le RER durant l'heure de pointe du matin et les flancs de pointe sur la ligne A (septembre), pour les flancs de pointe sur la ligne B (mi-décembre).

Renforcement du service pour 46 lignes de banlieue et sur toutes les lignes du Réseau Régional de Surface ("RRS").

- janvier : Modification des tarifs du stationnement payant dans le but de rééquilibrer la circulation dans Paris : revalorisation du coût du stationnement rotatif, forte baisse du stationnement résidentiel.
- mars : Lancement par le Syndicat des Transports d'Ile-de-France de la carte Solidarité Transport permettant aux bénéficiaires de la Couverture Maladie Universelle d'acheter des billets demi-tarif en carnet ou à l'unité sur les réseaux de transport d'Ile-de-France.
- juillet : Hausse des tarifs des transports publics en Ile-de-France.

Du 15 juillet au 24 août : Fermeture de la ligne C du RER entre "Invalides" et "Austerlitz" pour travaux de rénovation.

Du 21 juillet au 18 août : Opération Paris-Plage et réouverture le 22 août des voies sur berges aux automobilistes.

- août : Attribution du passe Navigo aux porteurs d'Imagine R étudiants ou de Carte Intégrale.

-
- octobre : Service renforcé sur le réseau ferré et mise en place d'un service de navettes par autobus pour relier une quinzaine de sites ouverts au public lors de l'opération Nuit Blanche organisée par la ville de Paris dans la nuit du samedi 5 au dimanche 6.

Le 21, grève sur la ligne 8 du métro.

- du 26 décembre au 1er janvier 2003 : Rencontres œcuméniques de Taizé.

31 décembre : Mise en place du service "Grande Nuit" avec ouverture toute la nuit des lignes 1, 2, 6, 9 et 14 du métro, de toutes les lignes du RER, des lignes SNCF de banlieue et du Transilien et renforcement du service assuré par le Noctambus. Le STIF a décidé de pérenniser la gratuité des transports collectifs en région parisienne 17 heures le 31 décembre et 12 heures le 1^{er} janvier.

2003

- janvier : Création par le STIF du ticket T, qui remplace le billet U, et est valide également sur le réseau OPTILE.

Le 20, grève sur la ligne 2 du métro ; le 21, grève sur la partie SNCF du RER B.

- février : le 1, mouvement de grève sur tous les réseaux de la RATP et fortes chutes de neige ; le 12, interruption du trafic sur la ligne 14 du métro entre 8h45 et 12h30 en raison d'une panne du système informatique.

- mars : du 10 au 12, mouvements sociaux sur le réseau SNCF Paris Est.

Le 12, 17 et 21 : mouvements sociaux dans la fonction publique.

- avril : Grèves dans la fonction publique entraînant des perturbations sur les différents réseaux de la RATP.

- 13 -16 mai : Mouvements sociaux pour tout le secteur public avec fortes perturbations sur le réseau SNCF Ile-de-France et sur les différents réseaux de la RATP.

- 2 - 12 juin : Grèves dans les transports en commun et dans la fonction publique.

Le 19, grève dans les transports en commun et dans la fonction publique. Panne générale des Terminaux Points de Vente dans les stations de métro.

Samedi 21 : Fête de la Musique. Le service "Grande Nuit" est assuré dans la nuit de samedi à dimanche : fonctionnement toute la nuit des lignes 1, 2, 4, 6, 9 et 14 du métro, des lignes

ANNEXE G. HISTORIQUE DES FAITS MARQUANTS AYANT UNE INCIDENCE SUR
LE TRAFIC DE LA RATP

A et B du RER, renforcement du service offert sur les Noctambus et sur les réseaux de la SNCF. Vente d'un ticket spécial valable du samedi 17 heures au dimanche 7 heures.

- du 12 juillet au 14 août : Fermeture de la ligne C du RER entre Invalides et Austerlitz.
- août : Augmentation des tarifs des transports publics.

Le 13, interruption de service sur la ligne A entre Auber et Nation suite à une chute de caténaire, le service reste perturbé le 14.

Fin de l'opération Paris - Plage qui a attiré 3 millions de voyageurs.

Du 23 au 31, Championnats du Monde d'Athlétisme au Stade de France. Mise en place d'un plan de transport RATP permettant d'acheminer les spectateurs, les athlètes et accrédités. Canicule.

- du 16 au 22 septembre : Semaine européenne de la mobilité et du transport public, se terminant le 22 par la journée "En ville sans ma voiture". Augmentation sensible de l'offre de transport et du trafic sur les différents réseaux de la RATP.
- du samedi 4 au dimanche 5 octobre : Service renforcé sur le métro et les lignes de bus Noctambus lors de la Nuit Blanche. La ligne 14 est restée ouverte et gratuite de 1h00 à 5h30 et 5 lignes spécifiques de bus ont assuré gratuitement la desserte des lieux d'animation.
- novembre : Tenue du Forum Social Européen à Paris, Saint-Denis, Bobigny et Ivry pour lequel la RATP a mis en place un renforcement de son offre de transport.

Renforcement du service sur 8 lignes du réseau de bus de Paris (38, 43, 47, 65, 67, 76, 86 et 96) de façon à prendre en compte l'accroissement de mobilité des Franciliens en début de soirée et le soir, mais aussi le week-end, d'améliorer l'offre de service sur les axes bénéficiant de nouveaux couloirs de bus sécurisés (entre Châtelet et Gare de l'Est, rue de Rivoli, boulevard Saint-Germain, entre Saint-Augustin et Trinité...) et la prise en compte des futures lignes Mobilien, et d'améliorer les relations entre Paris et les communes limitrophes. Ce renforcement représente 84 000 kilomètres supplémentaires par an et devrait induire plus de 3 millions de voyages supplémentaires sur les bus parisiens en 2004.

- décembre : Perturbations sur de nombreuses lignes de bus suite à d'importantes manifestations sur la voie publique.

Prolongement du RER SNCF "Eole" jusqu'à Tournan-en-Brie avec desserte de six nouvelles gares dans l'Est de l'Ile-de-France réparties entre les départements de Seine-Saint-Denis et de Seine-et-Marne.

Prolongement de la ligne T1 de Bobigny à la gare de Noisy-le-Sec, sur 2.9 kilomètres, avec desserte de 5 nouvelles stations. 6 lignes de bus du secteur vont adapter leur offre à ce prolongement.

Prolongement de la ligne 14 de Madeleine à Saint-Lazare, assurant ainsi la correspondance avec les lignes 3, 9, 12, 13 du métro, le RER E et la gare SNCF.

Mise en place du service "Grande Nuit" lors de la Saint-Sylvestre.

2004

- janvier : Fin de la gratuité sur les réseaux à 12h.

Suspension de l'interconnexion de la ligne A du RER à Nanterre-Préfecture les samedi 10 et dimanche 12.

Le 21, journée de grève à la SNCF entraînant la suppression des interconnexions à Nanterre-Préfecture et Gare du Nord et une perte de trafic sur les réseaux de la RATP.

A partir du 22 : Perturbations sur les lignes de bus dues à de nombreuses manifestations sur la voie publique (de 11 à 16 lignes selon les jours) et aux mesures restrictives de circulation mises en place à l'occasion de la visite d'Etat du président chinois Hu Jintao (du 26 au 29)

Fermeture pour travaux de rénovation de 8 stations réparties sur 3 lignes de 21h à fin de service et de 1 station sur toute la durée de service.

- février : Perturbations sur les lignes de bus dues à des manifestations sur la voie publique les 5, 7 et 26, et aux déplacements dans Paris du chef d'Etat d'Israël.

Fermeture pour travaux de rénovation de 13 stations réparties sur 4 lignes de 21h à fin de service et de 3 stations réparties sur 3 lignes sur toute la durée de service.

- mars : Fermeture pour travaux de rénovation de 15 stations réparties sur 6 lignes de 21h à fin de service et de 4 stations réparties sur 3 lignes sur toute la durée de service.

Perturbations sur les lignes de bus dues à de nombreuses manifestations sur la voie publique (jusqu'à 35 lignes le 25).

ANNEXE G. HISTORIQUE DES FAITS MARQUANTS AYANT UNE INCIDENCE SUR LE TRAFIC DE LA RATP

Les dimanche 7 et 14 : Interruption de service dans les deux sens du RER B entre les gares d'Aulnay-sous-Bois et Aéroport de PARIS 2 TGV suite à d'importants travaux en gare SNCF d'Aulnay-sous-Bois du 6 mars au 2 mai.

Le 11 : Attentats meurtriers dans les transports en commun de Madrid. Mise en place du Plan Vigipirate Rouge dans les transports collectifs franciliens et provinciaux. Forte augmentation du nombre d'interruptions de service pour cause de paquets suspects à partir de cette date.

Le 26 : Grèves RATP.

- avril : Fermeture pour travaux de rénovation de 15 stations réparties sur 4 lignes de 21h à fin de service et de 5 stations réparties sur 5 lignes sur toute la durée de service.

Perturbations sur les lignes de bus dues à des manifestations sur la voie publique en début de mois (jusqu'à 34 lignes le 3 avril), et à la visite d'Etat de la Reine d'Angleterre les 5 et 6.

Le 8 : Interruption de service sur la ligne A du RER avec évacuation des gares de la Défense à Nation de 20h05 à 21h45 suite à une alerte à la bombe.

Le 27 : Appareils de vente et de péage hors service entre la Croix de Berny et Saint-Rémy-Chevreuse sur la ligne B du RER en raison d'un incendie sur les câbles alimentant le poste haute tension de la Croix de Berny.

- mai : Fermeture pour travaux de rénovation de 10 stations réparties sur 3 lignes de 21h à fin de service et de 4 stations réparties sur 4 lignes sur toute la durée de service.

Perturbations sur les lignes de bus dues à des manifestations sur la voie publique les 14, 15, 16, 18, 25, 27, 29 mai (13 à 35 lignes selon les jours).

Mise en place de Libobus (ligne 578), nouveau service pour les salariés de l'Ile-Saint-Denis, desservie du nord au sud par des minibus le matin et le soir.

Le 4 : Panne du système informatique d'Orlyval et mise en place d'un service de remplacement par autobus entre Antony-RER et Orly-Ouest de 6h20 à 20h30.

Le 13 : Grève nationale SNCF ayant perturbé le trafic en Ile-de-France (40% à 50% des trains ont circulé).

– juin : Fermeture pour travaux de rénovation de 18 stations réparties sur 4 lignes de 21h à fin de service et de 9 stations réparties sur 5 lignes sur toute la durée de service.

Mise en service d'une nouvelle ligne de bus (394) entre Bourg-la-reine et Issy-les-Moulineaux.

Perturbations sur les lignes de bus dues à des manifestations sur la voie publique les 3, 5, 10, 12, 15, 19, 26 et 29 (de 15 à 37 lignes selon les jours) et de la visite d'Etat du président américain G.W. Bush.

Le 7 : Fortes perturbations sur les trafics RATP et SNCF de 3h30 à la mi-journée suite à des coupures de courant liées à un mouvement de grève des personnels EDF.

Le 8 : Alerte pollution avec incitation à prendre les transports en commun et limitation de la vitesse automobile de -20km/h.

Le 21 : Fête de la Musique avec le Service "Grande Nuit" et la vente d'un ticket spécial.

Le 28 : Baisse d'alimentation électrique occasionnant une baisse de service de -25% sur le RER A et sur de nombreuses lignes de métro.

Annexe H

L'amélioration des méthodes d'analyse et de prévision à la RATP

L'ensemble de ce travail répond de façon précise à la demande de la RATP, qui désirait améliorer ses méthodes d'analyse et de prévision des trafics sur les sous-réseaux.

Il a permis de **formaliser des méthodes** employées couramment dans l'unité Tarification Ventes Résultats du département Commercial de la RATP, comme le calcul du trafic du jour ouvrable moyen qui élimine l'impact du calendrier ou des grèves ponctuelles à la RATP lors des analyses mensuelles.

De même, la **reconstitution d'un historique** des faits marquants ayant une incidence sur les trafics de la RATP, la définition précise des séries constituant la base de données économiques décrivant la conjoncture économique en Ile-de-France ou encore l'explication de leurs variations passées, constituent un socle de travail sur lequel l'unité pourra s'appuyer lors de ses analyses sur longue période.

La **classification des stations de métro** ainsi que des gares de RER en fonction des motifs à l'origine des déplacements et des modes de rabattement¹ a permis d'enrichir le rapport annuel

¹Cette classification a été réalisée grâce à une analyse factorielle des correspondances et une classification hiérarchique ascendante.

édité par l'entité² : chaque classe possède une dynamique différente.

En 2003, sur le métro et le RER, la chute de trafic a été plus marquée pour les stations caractérisées par des motifs non obligés (achats et loisirs), du fait de la morosité économique et de la mauvaise saison touristique.

Sur le métro, le poids de chacun des motifs est demeuré stable. Sur le RER, la proportion des déplacements ayant pour motif le travail ou les affaires professionnelles diminue au profit des déplacements pour les achats et loisirs.

Le caractère opérationnel de nos modélisations a été accru par la **construction, sous Excel, de fichiers permettant le calcul des rétro-prévisions et des prévisions** :

- Le fichier de rétro-prévision permet d'expliquer les variations de trafic observées par les variations de la situation économique en Ile-de-France.
- Le fichier de prévision permet de réaliser des prévisions en fonction du seul passé des variables de trafic (modélisation auto-régressive) ou en fonction de variables économiques (modélisations de long terme et de court terme dans un cadre cointégré). Pour les modèles issus de l'analyse de la cointégration, les prévisions des séries décrivant l'environnement économique en Ile-de-France sont calculées à partir de leur modélisation auto-régressive. Elles supposent que l'évolution future de ces variables sera similaire à celle observée par le passé. L'utilisateur a toutefois la possibilité de modifier ce scénario implicite en imposant un taux de croissance pour les mois à venir.

Dans ces fichiers, l'ensemble des calculs se met à jour automatiquement : la seule tâche de l'utilisateur consiste à alimenter la base de données.

Avec l'aide d'un stagiaire³, des macros ont été programmées en Visual Basic pour permettre la mise à jour des paramètres de chacun des trafics étudiés si les comportements des clients se modifiaient, ou étudier d'autres types de trafic (par exemple, pour modéliser les trafics d'une classe de stations).

²Ce rapport annuel s'intitule "Analyse des résultats 2003" et décrit l'environnement des transports, le service offert, le trafic mesuré ainsi que les ventes de titres à la RATP.

³Yannick TOTEU FOTSING

ANNEXE H. L'AMÉLIORATION DES MÉTHODES D'ANALYSE ET DE PRÉVISION À
LA RATP

Le fichier de rétro-prévision a d'ores et déjà été utilisé pour expliquer l'évolution des trafics de la RATP au premier semestre 2004 sur les différents sous-réseaux. Il a permis de mettre en évidence certains problèmes de comptage ou des pertes de trafic dues aux manifestations dans la fonction publique au cours du premier semestre 2003.

Variations de trafic du premier semestre 2004 par rapport au premier semestre 2003				
	Métro	RER	Bus Paris	Bus banlieue
Données brutes	+4.42%	+8.39%	-0.42%	+3.71%
<i>Impact calendrier et effet ponctuel des grèves</i>	+2.10%	+2.56%	+2.79%	+1.35%
Données JO	+2.32%	+5.82%	-3.21%	+2.36%
<i>Impact évolution économique (d'après les élasticités identifiées)</i>	+1.12%	+1.22%	+1.22%	+2.47%
Données JO corrigées de l'impact de la conjoncture économique = part non expliquée de la variation de trafic	+1.19%	+4.60%	-4.43%	-0.11%
Explications possibles : 2003	<ul style="list-style-type: none"> Pertes de trafic dues aux manifestations dans la fonction publique 	<ul style="list-style-type: none"> Pertes de trafic dues aux manifestations dans la fonction publique Problème de comptage des Cartes Intégrale et Imagine R 		
2004	<ul style="list-style-type: none"> Amélioration de la qualité de service et de l'offre plus forte que par le passé 	<ul style="list-style-type: none"> Amélioration de la qualité de service et de l'offre plus forte que par le passé 	<ul style="list-style-type: none"> Impact de la lutte contre la fraude (ou problème de comptage ?) 	<ul style="list-style-type: none"> Impact de la lutte contre la fraude