

* The first version of TRIO is the result of collaborative work, much of it long standing, among many people. The **applied theoretical developments** have primarily involved Marcel Dagenais and Richard Laferrière, but would have been quite different without the contributions of Ulrich Blum, Jean-Marie Dufour and Michael Wills, the last of whom had a decisive role, as early as 1976, in the study of Box-Cox and Box-Tukey forms. In addition to these, many others were primarily involved in **applications**: Jean Choquette, Sergio Jara-Diaz, Gertraud Foos, Benedikt Mandel, Juan de Dios Ortuzar and Werner Rothengatter. The core of **scientific programming** is due to Tran Liem, but the **development programming**, as well as the construction and documentation of TRIO proper, primarily involved Louis-Philippe Duclos, Francine Dufort, Jacques Guélat, Pierre Lestage and Pierre Galvan, while Dominique Paskievici and Tran Le made significant contributions. The **design and structure** of TRIO owe a considerable amount to Michael Florian and benefitted from the collaboration of André Babin: in the transportation context, TRIO was conceived as a complement to EMME/2, with which it shares many graphic routines. A number of these 20-some companions over the years appear more formally as authors or collaborators of the TRIO documentation listed on the last page of this summary presentation.

AJD-101
CRT-901

Cur cum TRIO?

de

Marc Gaudry *et al.**

Agora Jules Dupuit, Département de sciences économiques, Université de Montréal, C.P. 6128, Succursale Centre-ville, Montréal, Canada H3C 3J7.

This paper is based on work financed in part by Transport Canada, and which benefitted directly or indirectly from research funded by the National Sciences and Engineering Research Council and by the Social Sciences and Humanities Research Council of Canada, as well as by the *Actions Structurantes* and FCAR programs of the Quebec Government. Benedikt Mandel provided the German translation; Pierre Galvan and Juan de Dios Ortuzar provided the Spanish translation. Carlos Chalhoub was exceptionally innovative in refining the quality of the printed document.

Agora Jules Dupuit — Publication #101

Centre de recherche sur les transports — Publication #901

April 1993 – Revised in October 1993, April 1994, November 1996 and July 2005

Abstract

We describe in summary fashion the professional, scientific and technical features of the TRIO interactive-graphic regression analyst, which is characterized by a full integration of the four basic tasks required for regression work: the management of information, the production of models, the analysis of data or model results, and assistance in report generation.

Key-words: time series and cross-sectional variables, ordinary least squares regression, logit regression, nonlinear transformation, correction of residual errors, graphical analysis, regression result reports.

Résumé

Nous résumons les caractéristiques professionnelles, scientifiques et techniques de l'analyste de régression interactif graphique TRIO qui se distingue par une intégration complète des quatre tâches fondamentales requises pour le travail de régression : la gestion de l'information, la modélisation, l'analyse des données ou des résultats des modèles, et l'aide à la production de rapports.

Mots-clés : variables en séries chronologiques et coupes transversales, régression par moindres carrés ordinaires, régression logit, transformations nonlinéaires, correction des erreurs résiduelles, analyse graphique, rapports de résultats de régression.

Kurzdarstellung

Wir beschreiben zusammenfassend die professionellen, wissenschaftlichen und technischen Merkmale des interaktiv graphischen Regressionsanalyst TRIO, der die vier essentiellen Werkzeuge zur Durchführung von Regressionen vereinigt: Informationsmanagement, Modellerstellung, Analyse der Daten oder Modellergebnisse und Unterstützung bei der Berichterstellung.

Schlüsselworte: Zeitreihen und kreuzsektionale Variablen; Methode der Kleinsten Quadrate und Logit-Regressionsmodelle, nichtlineare Transformationen, Korrektur von Residuenfehlern; graphische Analyse; Regressionsergebnisberichte.

Résumen

Resumimos las características profesionales, científicas y técnicas del analizador de regresión interactivo-gráfico TRIO que se destaca por la integración completa de cuatro tareas fundamentales, necesarias durante el análisis de regresión: la gestión de la información, la modelación, el análisis de los datos y resultados de los modelos y el ayuda a la producción de informes.

Letras-claves: series cronologicas, mínimos cuadrados ordinarios, logit, transformaciones non-lineales, corrección de los errores residuales, analisis gráfico, informes de resultados, interactivo-gráfico.

TABLE OF CONTENTS OF THE 4 PRESENTATIONS

TRIO — The Regression Analyst	1
Information Management	1
Modelling	1
Data and result analysis	3
Report generation assistance	3
Summary Table. TRIO Program Structure: Integrated Tasks and their Main Dimensions	2
TRIO — L'analyste de régression	4
La gestion de l'information	4
La modélisation	4
L'analyse des données et des résultats	6
L'aide à la production de rapports	6
Tableau sommaire. Structure du progiciel TRIO : les tâches intégrées et leurs dimensions principales	5
TRIO — Der Regressionsanalyst	7
Informationsmanagement	7
Modellierung	7
Daten- und Ergebnisanalyse	9
Unterstützung bei der Berichterstellung	9
Tabellarische Übersicht. TRIO Programmstruktur: Komponenten und deren Hauptaspekte	8
TRIO — El analizador de regresión	10
Gestión de la información	10
Modelación	10
Análisis de datos y resultados	12
Ayuda a la producción de informes	12
Tabla resumen. Estructura de TRIO : Tareas integradas y principales dimensiones	11

LIST OF TABLES AND FIGURES COMMON TO THE 4 PRESENTATIONS

Table 1. TRIO database	14
A. The structure of a variable	
B. The structure of a variant	
Table 2. TRIO approach to regression	15
A. Model class and type	
B. Class root models and their nesting extensions	
Illustrative graphs	17
Figure 1. Conditional univariate regression analysis on screen	18
A. Regression line between two variables	
B. Regression line between two variables for given values of other variables	
Example of a TABLEX table of regression results of a model of LEVEL	19
Other examples of TRIO output	20
A. Tree of an ancestor variable and its descendants	
B. Graph obtained using the viewport feature	

APPENDIX AND REFERENCES

Appendix. Definition of model families and model types	21
References	23

OTHER TRIO DOCUMENTATION	See last page
---	---------------

TRIO – The Regression Analyst

TRIO is an open interactive-graphic program that allows users to perform in an integrated fashion most of the tasks involved in the explanation of quantitative and qualitative observations through exploratory analysis of the data and regression modelling. We now comment on these tasks in turn, selecting the significant features of the program indicated in the Summary Table.

1. Information Management

TRIO maintains a data base that contains both **variables** (data) and **variants** (each regression model is a variant), as shown in Table 1. The management of the contents of the data base, keeping trace of everything that matters – including the analyst’s notes –, ensures the complete consistency of the data base and makes it possible to export or import not only variables, but also fully estimated variants.

2. Modelling

Model classes and roots. The regression modelling available in TRIO allows the analyst to start with 3 simple univariate or multivariate models and to build gradually more complex extensions that contain the simpler root models as nested special cases. As shown in Table 2.A, these root models belong to 3 classes, namely LEVEL, SHARE and PROBABILITY, depending on whether the observations on the dependent variable pertain to the *level* of one (or many) variables, to the *market share* of an alternative, or to the categorically discrete *occurrence* of an alternative.

Within each class, the ROOT is the **univariate or multivariate linear starting point** that allows to build, in continuous nesting fashion, progressively more complex extensions that either modify the functional form of the dependent and independent variables of the root model, or take into account various ways of extracting systematic information from the residual errors in order to obtain for them a spherical distribution – that is, purely random and of constant variance. The current ROOT model types are the *Ordinary Least Squares* (OLS) procedure and the *Linear Logit* (LIN-LOGIT) procedure, which are well known and widely used.

In its current state shown in Table 2.B, TRIO allows estimation of such functional form extensions in all three available algorithms (L-1.4, S-1/S-5 and P-2) and further corrections for autocorrelation and heteroskedasticity in L-1.4.

Model type extensions: functional form. The TRIO variants that can be built as extensions of simpler popular forms make extensive use of direct and inverse Box-Cox and Box-Tukey transformations that include as special nested cases the linear and logarithmic cases (for the direct transformations), as well as the exponential case (for the inverse transformations). These transformations are applied to variables, or to complete functions, or to both, depending on the model type or family considered.

SUMMARY TABLE
TRIO Program Structure: Integrated Tasks and their Main Dimensions

INTEGRATED TASKS	MAIN DIMENSIONS		
	<i>Professional</i>	<i>Scientific</i>	<i>Technical</i>
1. Information Management Data base maintains <ul style="list-style-type: none"> • variables (data) • variants (models) Contents include <ul style="list-style-type: none"> • formulas • comments/notes Variables and Variants allow <ul style="list-style-type: none"> • read/write 	<ul style="list-style-type: none"> • complete genealogy • consistency with variables • correction of descendants • trace kept • import/export 	<ul style="list-style-type: none"> • unicity of each specification, including numerical criteria • reproductibility of variables and variants • transferability to others 	<ul style="list-style-type: none"> • IBM or compatible PC and SUN Workstation modular versions • real 32-bit processing • double precision calculations • operating system commands executed within TRIO • ASCII files
2. Modelling Variables <ul style="list-style-type: none"> • creation Variants <ul style="list-style-type: none"> • creation • estimation • parameters • forecasts 	<ul style="list-style-type: none"> • automatic update of descendants • creation and initialization define a version • families of variants can be given names • kept to allow later analysis • kept if desired 	<ul style="list-style-type: none"> • simple or complex variables • LEVEL, SHARE and PROBABILITY model classes • maximum likelihood • simple or complex nested structures can be constructed • by simulation and estimation methods 	<ul style="list-style-type: none"> • interactive creation/editing • interactive creation/editing • indirect batch mode possible • background estimation on SUN • fully documented algorithms, procedures, as well as formulas and statistics
3. Data and Result Analysis Graphs <ul style="list-style-type: none"> • variables • variant results Tables <ul style="list-style-type: none"> • variables • variant results 	<ul style="list-style-type: none"> • explanatory data analysis of variables or expressions • variables transformed by variant parameters • fitted or forecasted values • descriptive statistics • genealogical trees • reanalysis of results without parameter reestimation 	<ul style="list-style-type: none"> • pairwise correlation (y, X_k) and univariate regression on subsets conditional on other variables • one or many y-axes • residuals before and after corrections to obtain spherical distribution • multicollinearity analyses • statistics include elasticities for all variables including dummy variables 	<ul style="list-style-type: none"> • color choice • zoom-window for selected observations • viewport for many graphs on same screen • 4-directional scroll • TABLEX interactive table for easy comparison of variant results
4. Report Generation Assistance Graphs Tables	<ul style="list-style-type: none"> • have titles • constructed and edited interactively • have titles • constructed interactively • comments and formulas can be printed 	<ul style="list-style-type: none"> • types include: horizontal or vertical line, scatter, bar, pie • consistency with data base • choice of 2 definitions of variables 	<ul style="list-style-type: none"> • Tektronics 4014 files in PLOTS directory • printer drivers for dot-matrix, HPGL and PostScript graphics printers • ASCII files in REPORTS directory

The flexibility of this approach – that lets the data determine the appropriate form – can be partially explained with the Illustrative Graphs (not produced by TRIO). Assume that one is trying to explain a dependent variable y_m in terms of K explanatory variables X_k . Questions of interest include :

A. What is the form of the *dependent* variable?

- in models of LEVEL, would a transformation of that variable improve the fit? The case of a simple logarithmic transformation that yields such an improvement is shown in the first two graphs;
- in SHARE and PROBABILITY models, does the underlying probability or share behave in symmetric fashion between limits that are necessarily 0 and 1, or would an asymmetric reaction, perhaps between limiting values different from 0 or 1 (and variable across alternatives), better describe the data? Such shapes are illustrated in the second set of 2 graphs.

B. What is the form of the *explanatory* factors?

- is the form of the effect on the dependent variable monotonically increasing or decreasing? The Box-Cox parameters make it possible to obtain linear, logarithmic or more general power transformations of X_k . Curves 2 and 3 of the J-shaped cloud of observations shown on the third set of graphs show a linear and another monotonic form;
- however, a quadratic shape, such as that shown with line 1 of the fifth graph, would improve the fit. Similarly, the sixth graph contains a sigmoid cloud of observations that should be fitted with a cubic shape. Box-Cox transformations on variables make it easy to explore these different cases.

3. Data and result analysis

Conditional analysis with graphs. TRIO makes it possible to analyse the data or variables generated by variants – to say nothing of their transformations – with on screen computation of correlation and regression intercept and slope for pairs of variables, using if desired only the observations that satisfy certain restrictions on other variables: formally, a *conditional correlation or regression* is calculated. Such a case is illustrated in Figure 1.

Elasticities and other regression results in Tables. The most powerful table generated by TRIO is the interactive TABLEX that permits, for a chosen set of variants, the selection of desired results (e.g. coefficients, elasticities and t-statistics) for each variable: these results appear in user defined order, with full definitions of variables, on tables that easily save 90% of analysis time. An example is included here. In particular, elasticities (and, if desired, cross-elasticities for the SHARE and PROBABILITY variants) are calculated for all variables in a number of ways available on a menu. The Tablex feature for LEVEL variants in particular allows for the calculation of the elasticity of the *Expected value* and of the elasticity of the *Variance* of the dependent variable.

4. Report generation assistance

Another task that is fully integrated in TRIO pertains to the production of analyses that are ready for use in technical reports. A number of interactive edition features of both graphs and tables, and the ability to produce professional quality output files, greatly simplify the incorporation of these graphs, tables or lists, in outside documents. A tree and a graph are shown as examples of output.

TRIO – L’analyste de régression

TRIO est un progiciel interactif-graphique ouvert qui permet à l’usager l’exécution intégrée des tâches généralement nécessaires à l’explication d’observations quantitatives ou qualitatives par l’analyse exploratoire des données et les modèles de régression. En parcourant le Tableau Sommaire, nous noterons les caractéristiques distinctives du progiciel dans l’exécution de ces quatre tâches.

1. La gestion de l’information

TRIO gère une base de données qui regroupe les **variables** (données) et les **variantes** (chaque modèle de régression constitue une variante) comme on peut le comprendre au Tableau 1. La base de données conserve une trace de tout ce qui importe – y compris les notes faites par l’analyste –, garantit la cohérence interne de l’information et permet d’exporter ou d’importer non seulement des variables, mais aussi des variantes complètes estimées.

2. La modélisation

Les classes de modèles et leurs racines. Les méthodes de régression disponibles dans TRIO permettent à l’analyste de formuler au départ trois modèles simples univariés ou multivariés et d’en construire des prolongements progressifs qui comprennent ces “modèles-racines” comme cas particuliers emboîtés. Comme on l’indique au Tableau 2.A, les racines appartiennent à 3 classes, dites de NIVEAUX, de PARTS et de PROBABILITÉS, suivant que la variable dépendante considérée décrit le **niveau** d’une (ou de plusieurs) variable, la **part de marché** d’alternatives (ou options), ou l’**occurrence** discrète d’alternatives catégoriques (ou mutuellement exclusives).

Dans chaque classe, la RACINE est le **point de départ univarié ou multivarié linéaire** qui permet de construire par généralisation continue et successive des prolongements de complexité croissante qui, ou bien modifient la forme fonctionnelle de la variable dépendante et des variables indépendantes du modèle-racine, ou bien permettent d’extraire des erreurs résiduelles l’information systématique comprise dans le but d’en obtenir une distribution sphérique – c’est-à-dire vraiment aléatoire et de variance constante. Les genres de MODÈLES-RACINES disponibles sont le *Moindres Carrés Ordinaire* (M.C.O.) et le *Logit Linéaire* (LIN-LOGIT), qui constituent donc les outils élémentaires requis pour comprendre les prolongements.

Dans son état actuel présenté au Tableau 2.B, TRIO permet les prolongements qui ont trait à l’estimation des formes fonctionnelles dans le cadre des trois algorithmes implantés (L-1.4, S-1/S-5 et P-2) et ceux qui corrigent l’auccorrélation et l’hétéroscédasticité dans le cadre de L-1.4.

Prolongements et genres de modèles : les formes fonctionnelles. Pour construire des variantes qui prolongent les formes communes simples, TRIO fait un usage généreux des transformations Box-Cox et Box-Tukey directes et inverses dont les cas particuliers emboîtés comprennent le linéaire et le logarithmique (pour les transformations directes), ainsi que l’exponentiel (pour les transformations inverses). On applique ces transformations à des variables, à des fonctions, ou aux deux à la fois, suivant la famille ou le genre du modèle considéré.

La souplesse de cette philosophie – qui fait déterminer la meilleure forme par les données – peut être en partie évoquée par les Illustrations Graphiques (voir *Illustrative Graphs*, qui n’est pas produit par TRIO). Supposons en effet qu’il s’agisse d’expliquer une variable dépendante y_m par K variables explicatives X_k . On peut alors se poser les questions suivantes :

TABLEAU SOMMAIRE
Structure du progiciel TRIO : les tâches intégrées et leurs dimensions principales

TÂCHES INTÉGRÉES	DIMENSIONS PRINCIPALES		
	<i>Professionnelle</i>	<i>Scientifique</i>	<i>Technique</i>
1. Gestion de l'information La base de données gère <ul style="list-style-type: none"> • variables (données) • variantes (modèles) Son contenu comprend <ul style="list-style-type: none"> • formules • commentaires/notes Variables et Variantes <ul style="list-style-type: none"> • lues/écrites 	<ul style="list-style-type: none"> • généalogie complète • compatibilité avec variables • correction des descendants • trace conservée • importation/exportation 	<ul style="list-style-type: none"> • unicité des formulations y compris critères numériques • reproductibilité des variables et variantes • transférabilité à autrui 	<ul style="list-style-type: none"> • versions modulaires pour IBM ou compatibles PC et pour postes SUN • précision machine 32 bits • calculs en double précision • exécution dans TRIO d'instructions du système d'exploitation • fichiers ASCII
2. Modélisation Variables <ul style="list-style-type: none"> • création Variants <ul style="list-style-type: none"> • création • estimation • paramètres • prévisions 	<ul style="list-style-type: none"> • mise à jour automatique des descendants • création et initialisation définissent une version • familles de variantes ont des noms • conservés pour analyses subséquentes • conservées s'il y a lieu 	<ul style="list-style-type: none"> • variables simples ou complexes • modèles de NIVEAUX, de PARTS et PROBABILITÉS • maximum de vraisemblance • structures simples ou complexes et emboîtées • par méthodes de simulation et d'estimation 	<ul style="list-style-type: none"> • création/édition interactive • création/édition interactive • exécution en lots possible • estimation en arrière-plan sur postes SUN • algorithmes, formules, procédures et statistiques documentés en détail
3. Analyse des données/résultats Graphiques <ul style="list-style-type: none"> • variables • résultats variantes Tableaux <ul style="list-style-type: none"> • variables • résultats variantes 	<ul style="list-style-type: none"> • analyse exploratoire des données ou expressions • variables transformées par les paramètres de variantes • valeurs calculées ou prévisions • statistiques descriptives • arbres généalogiques • réanalyse sans réestimer les paramètres 	<ul style="list-style-type: none"> • corrélations partielles (y, X_k) et régression univariée sur ensembles conditionnels à d'autres variables • une ou plusieurs ordonnées (axes-y) • erreurs résiduelles avant et après corrections pour obtenir distribution sphérique • analyses de multicollinéarité • élasticités pour toutes les variables, y compris catégoriques ou muettes 	<ul style="list-style-type: none"> • choix de couleurs • zoom-fenêtre pour observations choisies • viewport pour superposition de graphiques sur même écran • déroulement (scroll) à 4 directions • tableau interactif TABLEX facilite comparaison des résultats de variantes
4. Aide à la production de rapports Graphiques Tableaux	<ul style="list-style-type: none"> • peuvent avoir des titres • construits et édités interactivement • peuvent avoir des titres • construits interactivement • commentaires et formules peuvent être imprimés 	<ul style="list-style-type: none"> • de type courbe xy, nuage de points, diagramme en bâtons ou en secteurs • compatibilité avec base de données • choix de 2 définitions de variables 	<ul style="list-style-type: none"> • fichiers Tektronics 4014 dans répertoire PLOTS • utilitaires d'imprimantes à points, HPGL et graphiques PostScript • fichiers ASCII dans répertoire REPORTS

A. Quelle est la forme de la variable *dépendante*?

- dans les modèles de NIVEAUX, une transformation de cette variable améliorerait-elle l'ajustement? Les deux premiers graphiques (1 et 2) montrent qu'une transformation logarithmique a un tel effet;
- dans les modèles de PARTS et de PROBABILITÉS, la part ou la probabilité sous-jacente a-t-elle une forme symétrique entre deux limites nécessairement égales à 0 et 1? Ou obtient-on un meilleur ajustement à l'aide d'une courbe asymétrique située entre des limites différentes de 0 ou de 1 (et variables selon l'alternative considérée)? Des cas de ce genre sont illustrés par le deuxième ensemble de graphiques (3 et 4).

B. Quelle est la forme des facteurs **explicatifs**?

- l'effet sur la variable dépendante est-il monotone croissant ou décroissant? Les paramètres Box-Cox rendent possibles des transformations linéaire, logarithmique ou par une puissance quelconque de X_k . Les courbes 2 et 3 du nuage d'observations en J du graphique 5 montrent une forme linéaire et une forme monotone;
- toutefois, il est clair que la forme quadratique de la courbe 1 du même graphique 5 améliorerait l'ajustement. De la même façon, le graphique 6 montre un nuage d'observations de forme sigmoïde qui exigerait une forme cubique. Il est facile d'explorer ces diverses hypothèses avec les transformations Box-Cox.

3. L'analyse des données et des résultats

Analyse conditionnelle par des graphiques. Il est possible d'analyser avec TRIO les données ou les variables produites par les variantes – ou des transformations de ces variables – à l'aide de calculs de corrélation et de régression (ordonnée à l'origine et pente), pour des paires de variables, en n'utilisant si on le désire que les observations qui obéissent à certaines conditions définies sur d'autres variables : il s'agit alors de corrélation ou régressions conditionnelles. La Figure 1 illustre cette possibilité.

Elasticités et autres résultats de régression par tableaux. Le tableau le plus puissant de TRIO est le TABLEX interactif qui permet de choisir, pour un groupe de variantes considéré, des résultats désirés (par exemple le coefficient, l'élasticité et la statistique du t de Student) associés à chaque variable. Nous présentons ici un exemple de ces tableaux qui font visualiser dans un ordre choisi par l'utilisateur des résultats dans un format qui utilise notamment les définitions complètes des variables et fait généralement épargner plus de 90% du temps nécessaire à la comparaison de résultats. En particulier, les élasticités (et, si nécessaire, les élasticités croisées dans le cas des variantes de PARTS et de PROBABILITÉS) sont calculées pour toutes les variables explicatives selon plusieurs formulations disponibles au menu. Dans le cas des variantes de NIVEAUX, le TABLEX autorise entre autres le calcul de l'élasticité de l'**espérance mathématique** et de l'élasticité de la **variance** de la variable dépendante.

4. L'aide à la production de rapports

La dernière des tâches complètement intégrées dans TRIO concerne la production d'analyses utilisables dans des rapports. Plusieurs fonctions d'édition de graphiques et de tableaux, ainsi que la possibilité de produire des fichiers de sortie de qualité professionnelle, simplifient grandement l'incorporation de ces graphiques, tableaux ou listes à des documents externes. À titre d'exemples d'aides-rapports, nous montrons un arbre et un graphique.

TRIO — Der Regressionsanalyst

TRIO ist ein offenes, interaktiv graphisches Programm, mit dem der Benutzer auf umfassende Weise nahezu alle Untersuchungen zur Erklärung quantitativer und qualitativer Beobachtungen mittels Informationsanalyse der Daten und Regressionsmodellierung durchführen kann. Nachfolgend beschreiben wir einzelnen diese Aufgaben anhand der signifikanten Merkmale des Programms, die in der tabellarischen Übersicht aufgeführt sind.

1. Informationsmanagement

TRIO verwaltet eine Datenbasis, die sowohl **Variablen** (Daten) als auch **Varianten** (jedes Regressionsmodell gilt als Variante) beinhaltet, wie Tabelle 1 zeigt. Das Datenbasis-Management stellt die Konsistenz des Datenmaterials — inklusive der Anmerkungen des Benutzers — sicher und ermöglicht nicht nur den Im- und Export von Variablen, sondern auch vollständig geschätzter Varianten.

2. Modellierung

Modellklassen und Basismodelle. Mit Hilfe der Regressionsmodellierung von TRIO kann der Anwender mit drei einfachen univariaten oder multivariaten Modellen beginnen und nach und nach komplexere Erweiterungen schaffen, die die Basismodelle als verknüpfte (nested) Spezialfälle beinhalten. Wie Tabelle 2.A zeigt, lassen sich die Basismodelle drei Klassen zuordnen — LEVEL, SHARE und PROBABILITY —, maßgeblich hierfür ist, ob die Beobachtungen der abhängigen Variablen das **Niveau** (Level) einer (oder mehrerer) Variablen, den **Anteil** (Share) einer Alternative am Gesamtvolumen (z.B.: Marktanteil) oder die diskrete **Entscheidung** für eine Alternative (Probability) darstellen.

Innerhalb jeder Klasse ist das Basismodell der univariate oder multivariate lineare Ausgangspunkt, von dem aus sukzessiv komplexere Erweiterungen möglich sind. Diese modifizieren die Funktionsform der abhängigen und unabhängigen Variablen des Basismodells oder extrahieren auf unterschiedliche Weise systematische Informationen aus den Residuenfehlern, um eine sphärische Verteilung — Zufallsverteilung mit konstanter Varianz — zu erreichen. Als Basismodelltypen stehen die **Methode der Kleinsten Quadrate** (Ordinary Least Squares / OLS) und die **Lineare Logit Methode** (LIN-LOGIT) zur Verfügung, die weitverbreitet anerkennende Anwendung finden.

Die aktuelle Programmversion von TRIO ermöglicht es, die oben beschriebenen erweiterten Funktionsformen — siehe Tabelle 2.B — mit den drei zur Verfügung stehenden Algorithmen (L-1.4, S-1/S-5 und P-2) zu schätzen und mit L-1.4 weitere Korrekturen zur Autokorrelation und Heteroskedastizität durchzuführen.

Erweiterungen der Modelltypen: Funktionsform. Die mit TRIO generierten Varianten, die Erweiterungen einfacher weitverbreiteter Funktionen sind, verwenden in umfassender Weise direkte und inverse Box-Cox- und Box-Tukey-Transformationen, die auch die linearen und logarithmischen Fälle (direkte Transformation) sowie den Exponentialfall (inverse Transformation) einschließen. Diese Transformationen werden auf Variablen und / oder komplette Funktionen, in Abhängigkeit des betrachteten Modelltyps oder der Modellfamilie, angewandt.

Die Flexibilität dieses Ansatzes — wodurch die Daten die geeignete Funktionsform bestimmen — kann anhand von Graphiken (Illustrative Graphs) verdeutlicht werden. Bei dem Versuch eine abhängige Variable y_m im Sinne von K erläuternden Variablen X_k zu erklären, stellen sich nachfolgende interessante Fragen:

TABELLARISCHE ÜBERSICHT
TRIO Programmstruktur: Komponenten und deren Hauptaspekte

<i>KOMPONENTEN</i>	<i>LEISTUNGSMERKMALE</i>		
	<i>Anwendung</i>	<i>Wissenschaft</i>	<i>Technik</i>
1. Informationsmanagement Datenbank verwaltet <ul style="list-style-type: none"> • Variablen (Daten) • Varianten (Modelle) sowie zugehörige <ul style="list-style-type: none"> • Formeln • Kommentare/Notizen Variablen/Varianten <ul style="list-style-type: none"> • Ein-/Auslesen 	<ul style="list-style-type: none"> • Vollständige Genealogie • Konsistenz mit Variablen • Korrektur aufbauender/nachfolgender Variablen/Varianten • Nachvollziehbarkeit • Im- und Export 	<ul style="list-style-type: none"> • Eindeutigkeit jeder Spezifikation einschliesslich numerischer Kennzeichnung • Reproduzierbarkeit von Variablen und Varianten • Übertragbarkeit 	<ul style="list-style-type: none"> • Versionen für IBM-kompatible PC's und SUN Workstations (modulhaft) • 32-Bit-Verarbeitung • Rechengenauigkeit "double precision" • Ausführung von Betriebssystembefehlen aus dem Programm • ASCII-Dateien
2. Modellierung Variablen • Erstellung Varianten • Generierung <ul style="list-style-type: none"> • Schätzung • Parameter • Prognosen 	<ul style="list-style-type: none"> • Automatisches "update" generierter Variablen • Generierung und Initialisierung definieren eine Version • Namensgebung der Variantenfamilie • Speicherung zur späteren Auswertung • optionale Verwendung 	<ul style="list-style-type: none"> • Einfache oder komplexe Variablen • LEVEL-, SHARE- und PROBABILITY-Modellklassen • Maximum Likelihood • Konstruktion einfacher oder komplexer verknüpfter Strukturen • durch Simulations- und Schätzmethoden 	<ul style="list-style-type: none"> • Interaktive Erstellung/Editierung • Interaktive Erstellung/Editierung • Indirekte Stapelverarbeitung möglich • Hintergrundverarbeitung auf SUN Workstations • Vollständige Dokumentation der Algorithmen und Prozeduren sowie der Formeln und Statistik
3. Daten- und Ergebnisanalyse Graphen • Variablen • Varianten-Ergebnisse Tabellen • Variablen • Varianten-Ergebnisse	<ul style="list-style-type: none"> • Erklärende Datenanalyse von Variablen oder -konstellationen • Transformierte Variablen durch Variantenparameter • Angepasste oder prognostizierte Werte • Deskriptive Statistik • Stammbäume • Neuauswertung von Ergebnissen ohne erneute Parameterschätzung 	<ul style="list-style-type: none"> • paarweise Korrelation (y_i, X_{jk}) und univariate Regression auf bedingten Teilmengen • Eine oder mehrere y-Achsen • Residuen vor und nach den Korrekturen, um eine sphärische Verteilung zu erreichen • Multikollinearitätsanalysen • Statistik umfasst auch die Elastizität aller Variablen, einschliesslich der der Dummy-Variablen 	<ul style="list-style-type: none"> • Farbwahl für Graphenelemente • "zoom-window" für ausgewählte Beobachtungen • Darstellung mehrerer Graphen auf einem Bildschirm • Bildschirmrollen in 4 Richtungen • TABLEX, interaktive Tabelle zum leichten Vergleich von Variantenergebnissen
4. Berichterstellungshilfe Graphen Tabellen	<ul style="list-style-type: none"> • Überschriften einfügbar • Interaktive Konstruktion und Editierung • Überschriften einfügbar • Interaktive Konstruktion • Optionale Ausgabe von Kommentaren und Formeln 	<ul style="list-style-type: none"> • Möglichkeiten: Horizontale oder vertikale Polygonzüge, Streu-, Balken-, Tortendiagramme • Konsistenz mit der Datenbasis • Wahl zweier Variablendefinitionen 	<ul style="list-style-type: none"> • Tektronics-4014-Dateien im Verzeichnis PLOTS • Druckertreiber für Nadeldrucker, HPGL- und PostScript-Drucker • ASCII-Dateien im Verzeichnis REPORTS

- A. Welche funktionale Form besitzt die **abhängige** Variable?
- LEVEL-Modelle: Verbessert eine Transformation dieser Variablen die Güte der Abbildung? Der Fall einer einfachen logarithmischen Transformation, die eine solche Verbesserung erbringt, wird in den ersten beiden Graphiken (1,2) gezeigt.
 - SHARE- und PROBABILITY-Modelle: Verhält sich die zugrunde liegende Wahrscheinlichkeit oder der Anteil symmetrisch zwischen den Grenzwerten, die notwendigerweise bei 0 und 1 liegen, oder beschreibt eine asymmetrische Reaktionskurve, vielleicht zwischen Grenzwerten ungleich 0 und 1 (und variabel über die Alternativen), die Daten besser? Dies wird im zweiten Graphikpaar (3,4) verdeutlicht.
- B. Welche funktionale Form weisen die **erklärenden** Faktoren auf?
- Ist die Wirkung auf die abhängige Variable monoton steigend oder fallend? Die Box-Cox-Parameter ermöglichen lineare, logarithmische oder allgemeinere Potenztransformationen von X_k . Die Kurven 2 und 3 auf der J-förmigen Beobachtungswolke im dritten Graphikset (5,6) zeigen eine lineare und eine weitere monotone Form.
 - Eine quadratische Form, wie sie mit Funktion 1 in dieser Graphik aufgezeigt wird, verbessert die Abbildungsgüte der Beobachtungswolke entscheidend. Entsprechend beinhaltet die sechste Graphik eine sigmoide Beobachtungswolke, die mit einer kubischen Form angepaßt werden sollte. Box-Cox-Transformationen der Variablen erleichtern die Untersuchung dieser unterschiedlichen Fälle.

3. Daten- und Ergebnisanalyse

Bedingte Korrelations- oder univariate Regressionsanalyse am Bildschirm. TRIO ermöglicht die Analyse der durch Varianten erzeugten Variablen oder Daten — ungeachtet ihrer Transformationen — mit der direkten Bildschirmausgabe der Berechnung von Korrelation und Regressionsschritten sowie der Steigung für Variablenpaare, in der — falls gewünscht — nur die Beobachtungen, die gewisse Bedingungen für andere Variablen erfüllen, verwendet werden: formal wird eine **bedingte Korrelation oder Regression** berechnet. In Bild 1 wird solch ein Fall dargestellt.

Elastizitäten und andere Regressionsergebnisse in Tabellenform. Die mächtigste Tabelle, die von TRIO erstellt wird, ist das interaktive TABLEX, das dem Benutzer ermöglicht, für einen gewählten Satz von Varianten gewünschte Ergebnisse (z.B.: Koeffizienten, Elastizitäten oder t-Statistiken) für jede Variable auszuwählen. Diese Ergebnisse erscheinen in der vom Benutzer gewählten Reihenfolge in Tabellenform, mit vollständigen Definitionen der Variablen, was gut 90% der Analysezeit einspart. Ein Beispiel hierfür ist beigefügt. Im besonderen werden Elastizitäten (und falls erwünscht, auch Kreuzelastizitäten für die SHARE- und PROBABILITY-Varianten) für alle Variablen berechnet, entsprechend der im Menü angebotenen Möglichkeiten. Das Tablex für die LEVEL-Varianten erlaubt speziell die Berechnung der Elastizität des **Erwartungswertes** und die Elastizität der **Varianz** der abhängigen Variablen.

4. Unterstützung bei der Berichterstellung

Die von TRIO erstellten Analysen können direkt in Fachberichte oder Veröffentlichungen integriert werden. Eine Reihe von interaktiven Editiermöglichkeiten sowohl für Graphiken als auch für Tabellen und die Fähigkeit druckfertige Ausgabedateien zu produzieren, vereinfachen die Einbindung dieser Graphiken, Tabellen oder Listen in externe Dokumente erheblich. Ein Baum und eine Graphik werden als Ausgabebeispiele gezeigt.

TRIO — El analizador de regresión

TRIO es un programa interactivo-gráfico que permite ejecutar en forma integrada la mayoría de las etapas requeridas para explicar observaciones de naturaleza cuantitativa y cualitativa en análisis exploratorios de los datos y en modelos de regresión. Las características principales del programa, incluidas en la Tabla Resumen, serán presentadas en detalle en las secciones siguientes:

1. Gestión de la información

TRIO administra una base de datos compuesta de variables (datos) y variantes (cada modelo de regresión constituye un variante) como se muestra en la Tabla 1. La base de datos que mantiene un registro de toda modificación importante — incluyendo notas personales del analista —, garantiza la coherencia interna de la información y permite la exportación e importación no solo de variables sino que de variantes completamente estimados.

2. Modelación

Clases de modelos y raíces. Los métodos de regresión disponibles en TRIO permiten formular al analista, en un principio, tres modelos simples (univariados o multivariados) y luego construir extensiones más complejas que contienen a estos “modelos raíces” como casos particulares anidados. Tal como se muestra en la Tabla 2.A, los “modelos raíces” pueden pertenecer a tres clases: NIVEL, PROPORCION y PROBABILIDAD, según si la variable dependiente considerada describe el **nivel** de una (o varias) variable(s), la **proporción de mercado** de una alternativa, o la **ocurrencia** discreta de una alternativa categórica (o mutuamente excluyente).

Dentro de cada clase, la RAÍZ es el **modelo de referencia lineal univariado o multivariado** que permite construir, en forma continuamente anidada extensiones progresivamente más complejas que, o bien modifican la forma funcional de las variables dependiente e independientes del modelo raíz, o permiten extraer información sistemática de los errores residuales a fin de obtener para ellos una distribución esférica — esto es, puramente aleatoria y de variación constante. Los tipos de MODELOS RAÍCES disponibles son: **Mínimos Cuadrados Ordinarios** (OLS) y **Logit Lineal** (LIN-LOGIT), cuáles constituyen procedimientos ampliamente conocidos y utilizados en la práctica.

En su estado actual, que se presenta en la Tabla 2.B, TRIO permite estimar extensiones a estas formas funcionales en todos los algoritmos implementados (L-1.4, S-1/S y P-2). Además, en el caso de L-1.4 también permite corregir problemas de autocorrelación y heteroscedasticidad.

Tipos de extensiones: forma funcional. Las variantes de TRIO que pueden ser construidas como extensiones de las populares formas simples anteriores, utilizan las transformaciones directas e inversas del tipo Box-Cox y Box-Tuckey. Estas incluyen como casos particulares anidados las formas lineal y logarítmica (para la transformación directa) y la forma exponencial (para la transformación inversa). Estas transformaciones se pueden aplicar tanto a variables como a funciones completas o ambas a la vez, dependiendo del tipo o familia de modelo considerado. La flexibilidad de esta filosofía — que deja que los datos determinen la forma funcional más apropiada — puede ser parcialmente explicada con los Gráficos Ilustrativos “*Illustrative Graphs*” (no producidos por TRIO) que presentamos a continuación. En efecto, supongamos que se debe explicar una variable dependiente y_m en términos de K variables explicativas X_k . Las siguientes preguntas pueden ser de interés:

TABLA RESUMEN
Estructura de TRIO : Tareas integradas y principales dimensiones

TAREAS INTEGRADAS	PRINCIPALES DIMENSIONES		
	<i>Profesional</i>	<i>Científica</i>	<i>Técnica</i>
1. Gestión de la información El banco de datos administra <ul style="list-style-type: none"> • variables (datos) • variantes (modelos) Su contenido incluye <ul style="list-style-type: none"> • fórmulas • comentarios/notas Variables y Variantes <ul style="list-style-type: none"> • leídas/escritas 	<ul style="list-style-type: none"> • genealogía completa • coherencia con variables • corrección de los descendientes • trazas conservadas • importación/exportación 	<ul style="list-style-type: none"> • unicidad de la formulación y de los criterios numéricos • reproductividad de las variables y variantes • transferabilidad 	<ul style="list-style-type: none"> • versiones modulares para IBM o PC compatible y para estaciones de trabajo SUN • precisión máquina de 32 bits • cálculos en doble precisión • ejecución de instrucciones del sistema operativo en TRIO • archivos ASCII
2. Modelos Variables • creación Variantes • creación <ul style="list-style-type: none"> • estimación • parámetros • predicciones 	<ul style="list-style-type: none"> • ajuste automático de los descendientes • creación e inicialización definen una versión • familias de variantes tienen nombres • conservados para análisis subsecuentes • conservadas si es necesario 	<ul style="list-style-type: none"> • variables simples o complejas • modelos de NIVEL, PROPORCION y PROBABILIDAD • máxima de verosimilitud • estructuras simples o complejas y anidadas • utilizando métodos de simulación y estimación 	<ul style="list-style-type: none"> • creación/edición interactiva • creación/edición interactiva • ejecución en lotes posible • estimación de fondo en estaciones SUN • algoritmos, fórmulas, procedimientos y estadígrafos documentados
3. Analisis datos y resultados Gráficos • variables <ul style="list-style-type: none"> • resultados de variantes Tablas • variables <ul style="list-style-type: none"> • resultados de variantes 	<ul style="list-style-type: none"> • análisis exploratorio de los datos o expresiones • variables transformadas por los parámetros estimados • valores calculados o predichos • estadísticas descriptivas • árboles genealógicos • re-análisis sin volver a estimar parámetros 	<ul style="list-style-type: none"> • correlación parcial (y, X_k) y regresión univariada condicional a otras variables • una o varias ordenadas (eje-y) • error residual antes y después corrección para obtener distribución esférica • análisis de multicolinealidad • elasticidades calculadas para todas las variables, incluyendo las mudas 	<ul style="list-style-type: none"> • selección de colores • zoom en pantalla para observaciones escogidas • ventana para sobreponer varios gráficos en una misma pantalla • desplazamiento (scroll) en 4 direcciones • TABLEX interactiva facilita comparación de los resultados de variantes
4. Ayuda producción de informes Gráficos Tablas de Resultados	<ul style="list-style-type: none"> • tienen títulos • construidos y editados interactivamente • tienen títulos • construidas interactivamente • comentarios y fórmulas pueden ser impresos 	<ul style="list-style-type: none"> • de tipo curva xy, nube de puntos, diagrama de barras o de tortas • coherencia con la base de datos • elección entre 2 definiciones por variable 	<ul style="list-style-type: none"> • archivos Tektronics 4014 en el directorio PLOTS • utilitarios de impresoras matriciales, HPGL y gráficas PostScript • archivos ASCII en el directorio REPORTS

- A. ¿Cuál es la forma de la variable *dependiente*?
- ¿Mejorarían los resultados si se aplicase una transformación (por ejemplo Box-Cox) a esta variable en modelos de NIVEL? Los primeros dos gráficos (1 y 2) muestran la mejora producida por una simple transformación logarítmica;
 - ¿Se comportan en forma simétrica entre sus límites naturales [0 y 1] las probabilidades o proporciones de mercado subyacentes en modelos del tipo PROPORCION o PROBABILIDAD? ¿Sería posible obtener un mejor ajuste si se utilizara una curva asimétrica situada entre límites de 0 y 1 (y variables con las alternativas consideradas)? El segundo conjunto de gráficos (3 y 4) presenta dos ejemplos de este tipo.
- B. ¿Cuál es la forma de los factores **explicativos**?
- ¿Es la forma del efecto sobre la variable dependiente monótonicamente creciente o decreciente? Los parámetros de la transformación Box-Cox permiten obtener transformaciones de X_k lineales, logarítmicas o de cualquier potencia. Las curvas 2 y 3 de la nube de observaciones con forma de J, que se presenta en el tercer conjunto de gráficos, muestran una forma lineal y una monótonica respectivamente;
 - Sin embargo, es claro que la forma cuadrática como se muestra en la curva 1 del quinto gráfico, mejora el ajuste. Similarmente, el sexto gráfico contiene una nube de observaciones sigmoïdal que sería mejor representada por una forma cúbica. Las transformaciones de Box-Cox facilitan la exploración de estas diversas hipótesis.

3. Análisis de datos y resultados

Análisis condicional mediante gráficos. TRIO posibilita el análisis de los datos o variables generadas con los resultados provenientes de variantes estimados — o de transformaciones de estas variables — mediante cálculos en pantalla de correlaciones o regresiones, (intercepto y pendiente) para pares de variables, utilizando sólo las observaciones que satisfagan ciertas restricciones en otras variables si se desea; esto es, formalmente se calcula una **correlación o regresión condicional**. Este caso se ilustra en la Figura 1.

Tablas de elasticidades y otros resultados de regresión. La tabla más poderosa que genera TRIO es el TABLEX interactivo que permite, para un conjunto de variantes previamente estimados y seleccionados, escoger los resultados deseados (por ejemplo, coeficiente, elasticidad y estadígrafos t de Student) para cada variable incluida en los variantes seleccionados. Estos resultados aparecerán en el orden especificado por el usuario, incluyendo una definición completa de las variables, en tablas como las del ejemplo a continuación, presentación que permite ahorrar fácilmente 90% del tiempo dedicado al análisis de resultados. En particular, se pueden calcular las elasticidades (y si se desea, las elasticidades cruzadas en el caso de variantes del tipo PROPORCION y PROBABILIDAD) para todas las variables, según varias formulaciones disponibles. Además, para modelos del tipo NIVEL, el TABLEX permite calcular elasticidades del **valor Esperado** y de la **Variánza**, de la variable dependiente.

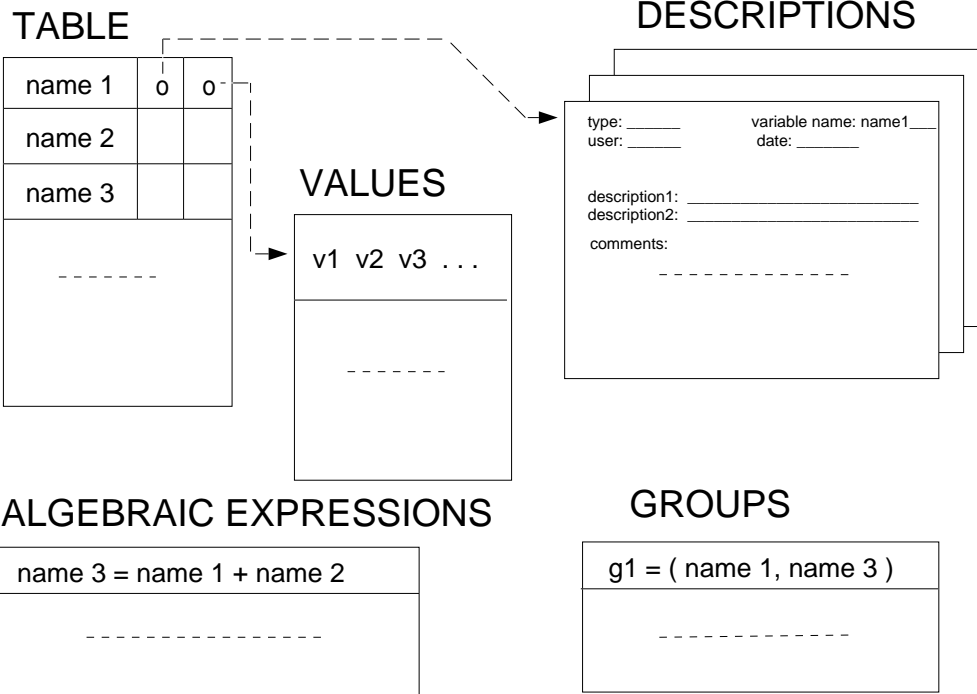
4. Ayuda a la producción de informes

La última tarea, completamente integrada en TRIO, es la capacidad de producir resultados listos para ser incluidos en informes técnicos. Una variedad de funciones de edición interactiva, tanto de gráficos como de tablas, así como la habilidad para producir archivos de salida de calidad profesional, simplifica la incorporación de estos gráficos, tablas o listados en documentos externos. Como ejemplo de ayuda a la producción de informes, mostramos a continuación un árbol genealógico y un gráfico.

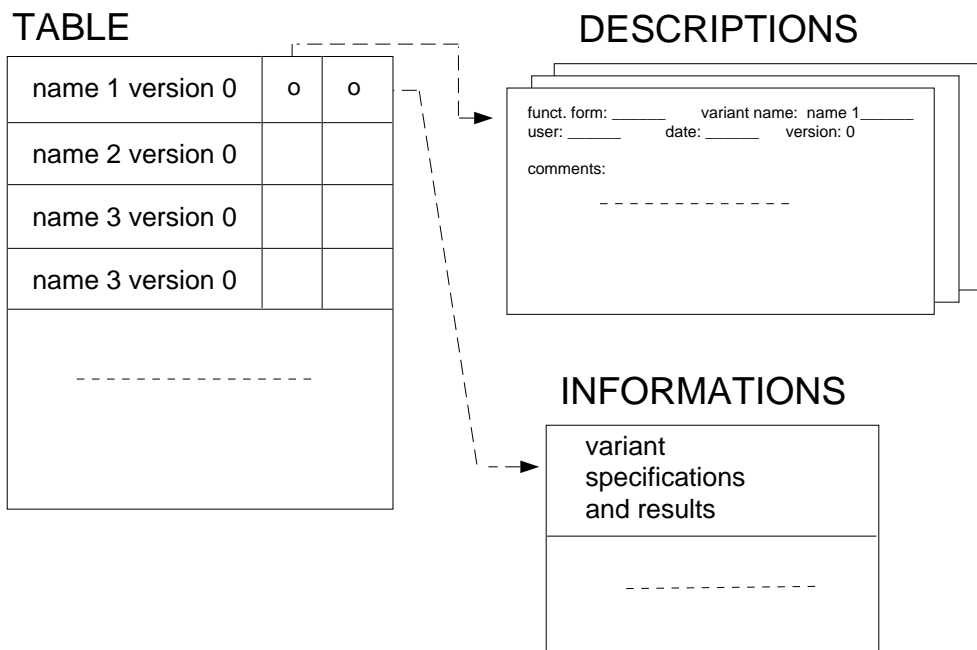
COMMON FIGURES AND TABLES

TABLE 1. TRIO DATABASE

A. The Structure of a Variable



B. The Structure of a Variant



**Table 2. TRIO APPROACH TO REGRESSION
A. Model Class and Type**

Class	Values of the dependent variable y_m				Form of the $f_m(\cdot)$		Avail. in TRIO v. 2.0
	in the regression		in the sample		in systematic and stochastic parts		
	$\sum_m y_m = 1$	range	nature	range	#	name*	
LEVEL	NO	$-\infty < y_m < \infty$	cont.	$-\infty < y_m < \infty$	L-1.4	OLS (Root)	yes
		$0 < y_m < \infty$		$0 < y_m < \infty$		BC-GAUHESEQ	yes
		$-\infty < y_m < \infty$		$-\infty < y_m < \infty$	L-2	GAUSEQ	
		$-\infty < y_m < \infty$		$-\infty < y_m < \infty$	L-3	GAUMEQ	
		$-\infty < y_m < \infty$		$L_1 \leq y_m \leq L_2$	L-4	AUSEQ-T	
SHARE	YES	$0 \leq y_m \leq 1$	cont.	$0 < y_m < 1$	S-1	LIN-LOGIT (Root)	yes
		$0 \leq y_m \leq 1$		$0 < y_m < 1$		BC-LOGIT	yes
		$0 \leq y_m \leq 1$		$0 < y_m < 1$	S-2	S-DOGIT	yes
		$0 \leq y_m \leq 1$		$0 < y_m < 1$	S-3	G-DOGIT	yes
		$0 \leq y_m \leq 1$		$0 < y_m < 1$	S-4	LIN-IPT-LOGIT	yes
		$0 \leq y_m \leq 1$		$0 < y_m < 1$	S-5	BT-IPT-LOGIT	yes
PROB.	YES	$0 \leq y_m \leq 1$	discr.	$y_m = 0$ or 1	P-2	LIN-LOGIT (Root)	yes
		$0 \leq y_m \leq 1$		$y_m = 0$ or 1		BC-LOGIT	yes
Defines CLASSES				Defines MODEL FAMILIES and TYPES			

B. Class Root Models and their Nesting Extensions

CLASS	TYPE		EXTENSIONS				
	Root	Family or Type (name*)	Functional form estimation			Corrections of residuals	
			on y_m		on X_k	Generalized Autocorrelation	Generalized Heteroskedasticity
			Asymmetry	Non-0 tails			
LEVEL	OLS	OLS				yes	yes
		BC-GAUHESEQ	yes		yes	yes	yes
SHARE	LIN-LOGIT	LIN-LOGIT					
		BC-LOGIT	yes		yes		
		S-DOGIT	yes	yes	yes		
		G-DOGIT	yes	yes	yes		
		LIN-IPT-LOGIT	yes	yes	yes		
		BT-IPT-LOGIT	yes	yes	yes		
PROBABILITY	LIN-LOGIT	LIN-LOGIT	yes				
		BC-LOGIT	yes		yes		

* Names of model families or model types are defined in the Appendix.

ILLUSTRATIVE GRAPHS

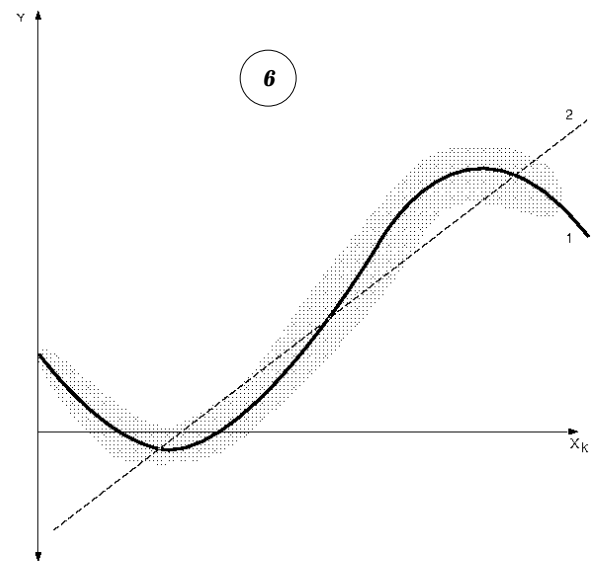
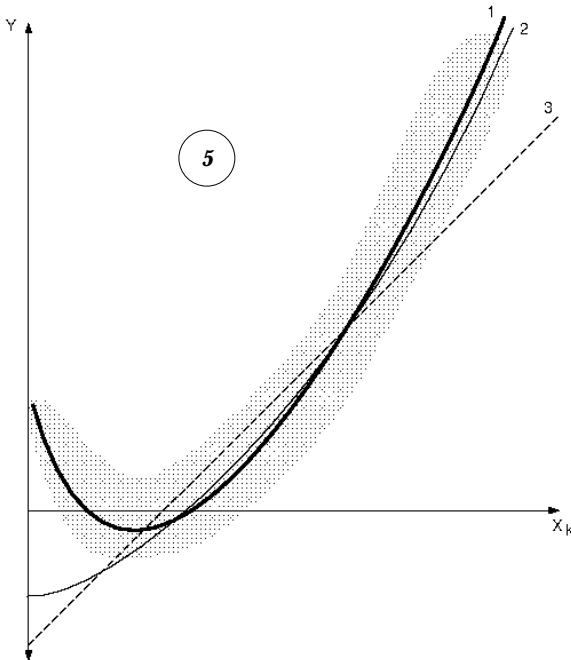
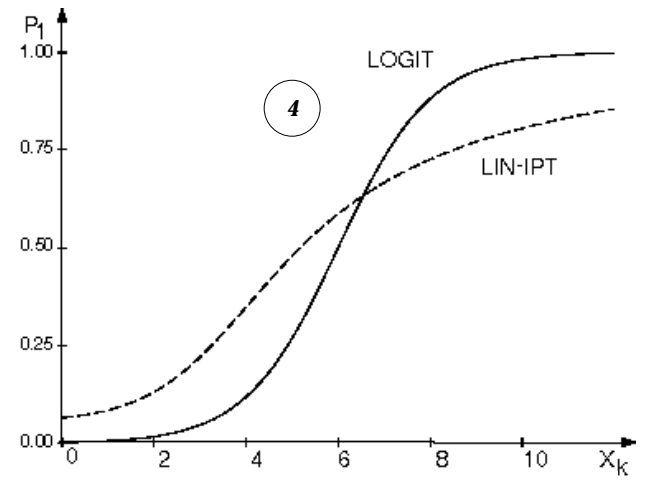
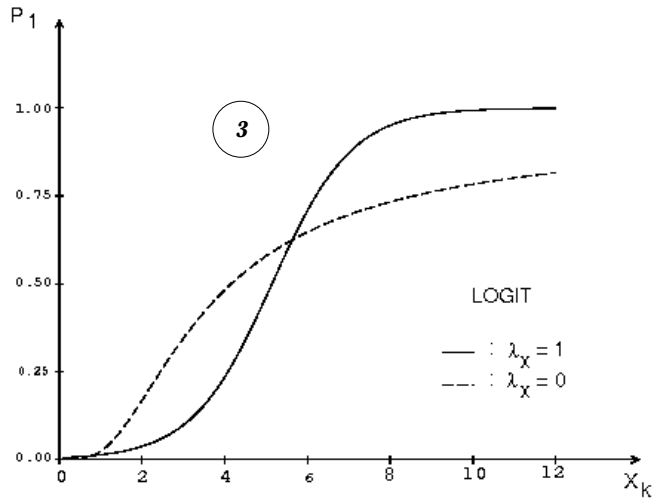
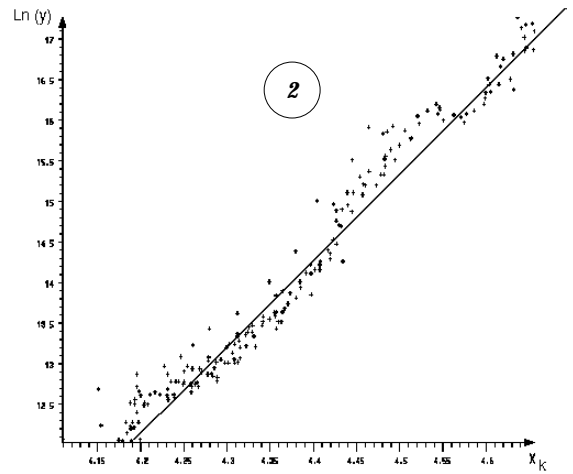
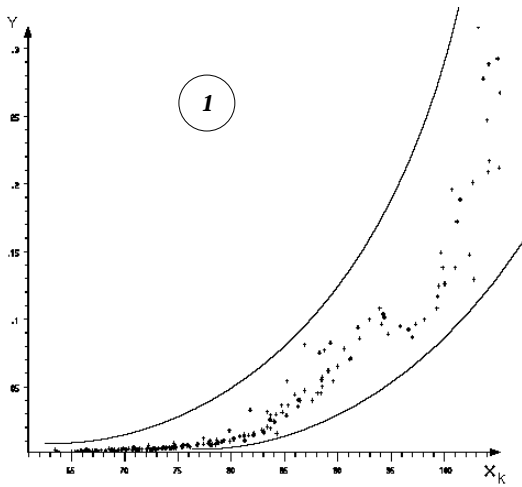
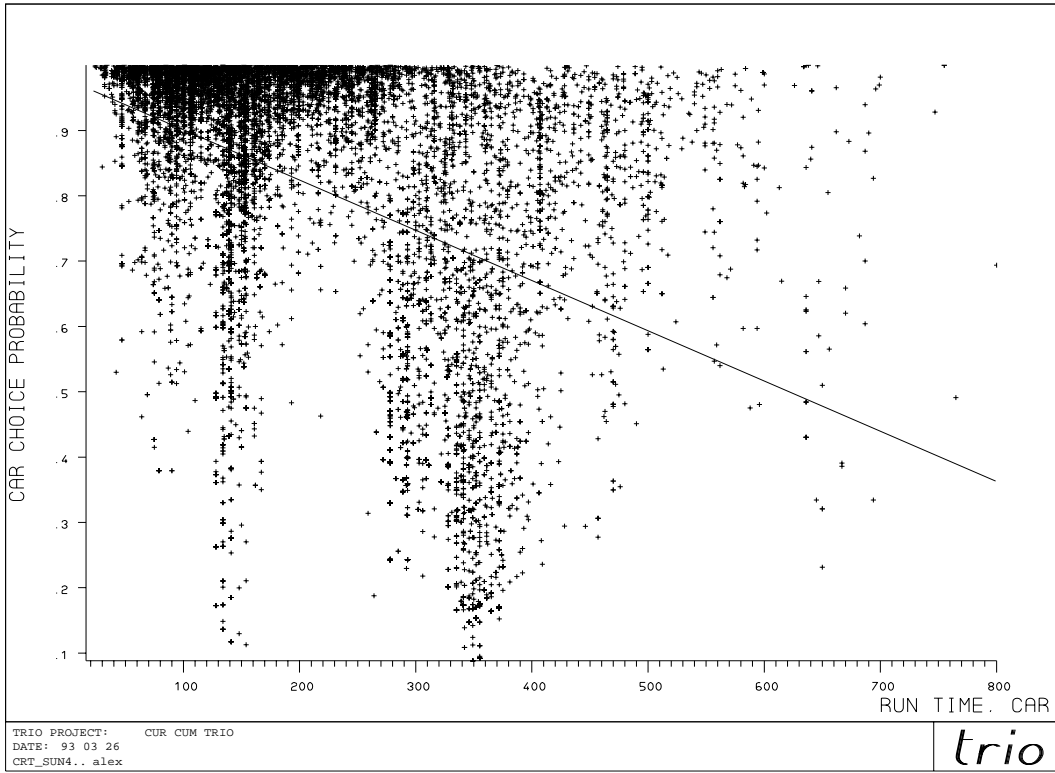
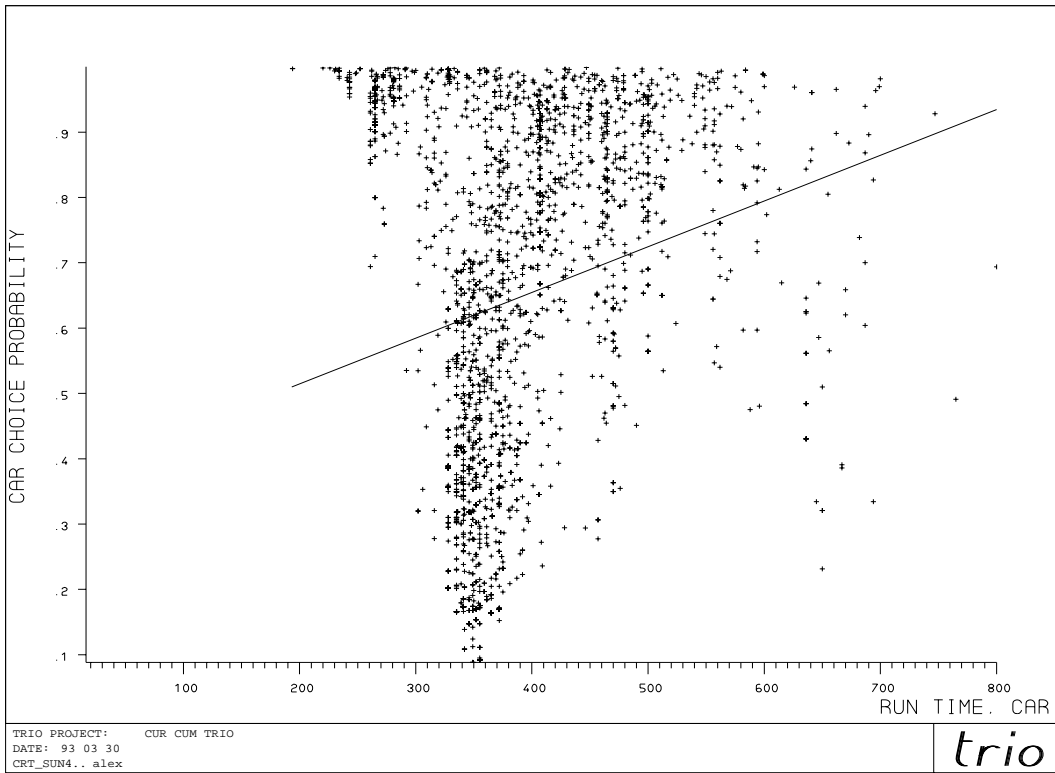


FIGURE 1. CONDITIONAL UNIVARIATE REGRESSION ANALYSIS ON SCREEN

A. Regression Line Between Two Variables



B. Regression Line Between Two Variables for Given Values of Other Variables



EXAMPLE OF A TABLEX TABLE OF REGRESSION RESULTS FOR A MODEL OF LEVEL

TRIO module: 4.15 Date: 92/11/17 User: CRT_SUN4
 Project: CUR CUM TRIO

 Comparison of 6 Models
 =====

B E T A S		TYPE =	LEVEL-1	LEVEL-1	LEVEL-1	LEVEL-1	LEVEL-1	LEVEL-1
ELASTICITIES S (EP)		VARIANT =	sch_lin	sch_bc1	sch_bc1a	sch_bc1a	ad_bc1au	ad_bc2au
(UNCOND. T-STATISTICS)		VERSION =	0	0	0	0	0	0
(COND. T-STATISTICS)		DEP.VAR. =	xcsnfafa	xcsnfafa	xcsnfafa	xcsnfafa	xcanfa	xcanfa
Real MUCTC	fscd		-.19E+08	-.25E+04	-.16E+06	-.23E+08		
student transit fare			-.376	-.292	-.419	-.410		
			(-2.31)	(-.28)	(-.25)	(-.26)		
			(-2.31)	(-1.10)	(-2.39)	(-2.74)		
				LAM 1	LAM 1	LAM 1		
Cars in MUCTC area	nc		-.91E+01	-.23E+01	-.60E+01	-.10E+02	-.18E+02	-.17E+01
			-.923	-.710	-1.138	-1.036	-.393	-.329
			(-1.38)	(-.69)	(-1.31)	(-1.27)	(-1.63)	(-.47)
			(-1.38)	(-.80)	(-1.66)	(-1.69)	(-3.42)	(-2.47)
				LAM 1	LAM 1	LAM 1	LAM 1	LAM 1
Consumer price index, Montreal	p		-.24E+05	-.26E+03	.80E+03	-.38E+04	.59E+04	.36E+03
			-1.081	-1.596	.577	-.166	.111	.060
			(-.71)	(-.44)	(.27)	(-.11)	(.36)	(.22)
			(-.71)	(-.74)	(.30)	(-.12)	(.47)	(.26)
				LAM 1	LAM 1	LAM 1	LAM 1	LAM 1
Real average weekly earnings, Montreal	rawem		.36E+05	.47E+02	.57E+03	.34E+05	.52E+05	.55E+04
			.915	.213	.270	.823	.576	.538
			(1.49)	(.21)	(.24)	(.37)	(.61)	(.49)
			(1.49)	(.28)	(.47)	(1.03)	(3.47)	(3.27)
				LAM 1	LAM 1	LAM 1	LAM 1	LAM 1

[...]

=====

GENERAL STATISTICS		TYPE =	LEVEL-1	LEVEL-1	LEVEL-1	LEVEL-1	LEVEL-1	LEVEL-1
		VARIANT =	sch_lin	sch_bc1	sch_bc1a	sch_bc1a	ad_bc1au	ad_bc2au
		VERSION =	0	0	0	0	0	0
		DEP.VAR. =	xcsnfafa	xcsnfafa	xcsnfafa	xcsnfafa	xcanfa	xcanfa
LOG-LIKELIHOOD			-2439.82	-2483.26	-2456.45	-2297.34	-2414.20	-2407.54
PSEUDO-R2 :	(E)		.907	.909	.934	.932	.971	.973
	(L)		.687	.477	.619	.942	.971	.973
	(E) ADJUSTED FOR D.F.		.890	.891	.919	.915	.964	.965
	(L) ADJUSTED FOR D.F.		.630	.377	.533	.928	.964	.966
AVERAGE PROBABILITY (Y=LIMIT OBSERV.)			.000	.000	.000	.000	.000	.000
SAMPLE :								
	- NUMBER OF OBSERVATIONS		169	169	169	169	169	169
	- FIRST OBSERVATION		13	13	13	13	13	13
	- LAST OBSERVATION		181	181	181	181	181	181
NUMBER OF ESTIMATED PARAMETERS :								
- FIXED PART :								
	. BETAS		27	27	27	27	28	30
	. BOX-COX		0	1	1	1	1	2
	. ASSOCIATED DUMMIES		0	0	0	0	0	2
- AUTOCORRELATION			0	0	4	4	2	2
- HETEROSKEDASTICITY :								
	. DELTAS		0	0	0	1	1	1
	. BOX-COX		0	0	0	1	1	1
	. ASSOCIATED DUMMIES		0	0	0	0	0	0

=====

OTHER EXAMPLES OF TRIO OUTPUT

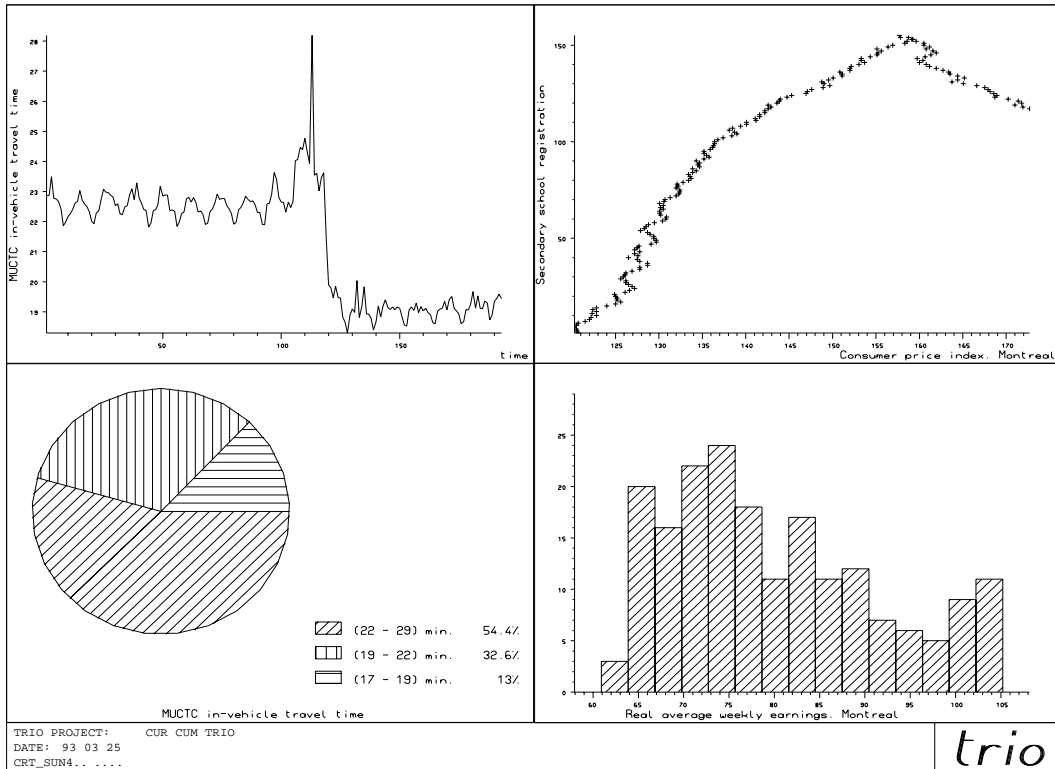
A. Tree of an Ancestor and its Descendants

```

TRIO module: 5.10                      Date: 93/03/26                      User: CRT_SUN4...
Project: CUR CUM TRIO
-----
EXTENDED TREE OF VARIABLE DESCENDANTS

!-AUTO-----!-AUPPOP
!-AUTOEQ56----!-AEC56PAD
!              !-AEC56POP
!-AUTOL1-----!-REBUT-----!-COAUT-----!-PGRPCKM
!              !              !-KMPARR1-----!-PRMTRAR1
!              !              !              !-PRMMAGR1
!              !              !-KMPAR-----!-PRMTRA
!              !              !-PRMMAG
!              !-AUTOEQ56----!-AEC56PAD
!              !-AEC56POP
!-AUTOPAD
!-AUTOPE
!-AVUG-----!-AVUGPPOP
!-AVUGPAD
!-AVUGPE
!-AVUGPE
!-COAUT-----!-PGRPCKM
!              !-KMPARR1-----!-PRMTRAR1
!              !              !-PRMMAGR1
!              !-KMPAR-----!-PRMTRA
!              !              !-PRMMAG
!              !-AUTOEQ56----!-AEC56PAD
!              !-AEC56POP
!-PERPA
!-PERPAVUG
!-POPPAU
!-REBUT-----!-COAUT-----!-PGRPCKM
!              !-KMPARR1-----!-PRMTRAR1
!              !              !-PRMMAGR1
!              !-KMPAR-----!-PRMTRA
!              !              !-PRMMAG
!              !-AUTOEQ56----!-AEC56POP
!              !-AEC56PAD
!-VUGPA
    
```

B. Graph Obtained Using the Viewport Feature



Appendix. Definition of model families and model types

LEVEL CLASS

The algorithms designed for the explanation of data on levels include:

L-1.4 : This procedure (Liem, Dagenais and Gaudry, 1993) naturally allows the user to specify the simplest root procedure of the LEVEL class, or more complex ones, namely:

OLS : Ordinary Least Squares;

BC-GAUHESEQ : Box-Cox Generalized AUtoregressive HEteroskedastic Single EQUation;

More precisely, this procedure for **single equation** model estimation consists of a Box-Cox regression with multiple-order autocorrelation and a very general form of heteroskedasticity which includes as nested special cases most forms used in practice. Distinct Box-Cox transformations can be used on the dependent variable, on groups of explanatory variables and in the model of heteroskedastic residuals. The procedure assumes that y_m is *a priori* a doubly censored variable but that the actual data used for model estimation include no limit (mass point) observations; it verifies *a posteriori* the validity of this assumption by calculating the probability that each observation be at the limits.

L-2 : this procedure (Gaudry *et al.*, 1981) is for a **single linear equation** model which contains endogeneous explanatory variables and multiple-order autocorrelation of the residuals. It basically combines ordinary least-squares with Box-Jenkins techniques to specify the appropriate multiple-order autocorrelation scheme and uses the Cochrane-Orcutt iterative solution technique to obtain the parameters. In two-stage least-squares terminology, the SSGAU option produces second-stage estimates and contains the least-squares option LSGAU as a special case; both can be used with or without the generalized autoregressive model of the residuals.

L-3 : this procedure (Gaudry and Liem, 1981) pertains to **multiple linear equations** in which each equation has its specific generalized autoregressive model of the residuals. When equations are *grouped* in order to take into account the contemporaneous correlation among the residuals of different equations, two estimation techniques can be used:

ISUGAU, an iterated version of Parks' (1987) seemingly unrelated procedure generalized to multiple-order autocorrelation cases.

TSGAU, an iterated third stage of a three-stage least-squares generalized autoregressive estimator applied after second stage regressors have been obtained for each equation separately and independently.

If, by contrast, equations are considered *simultaneously*, two techniques can be used:

IFIVER, an iterated version of Fair's (1972) full-information instrumental variables efficient estimator.

MIFIDA, a modified iterated version of Dhrymes' (1971) full-information dynamic autoregressive estimator.

The user can obtain non-iterated versions of the above four estimators by requesting only one iteration.

L-4 : this procedure (Dagenais, 1982), designed for first-order AUtorelated Single EQUation linear TOBIT models, allows for the explicit recognition of a lower limit L_1 (such as 0) and of an upper limit L_2 (if necessary) in the data. The explicit recognition of such mass points means that L-4 is a censored model type, as Maddala (1983), for instance, uses the notion of censoring.

SHARE CLASS

The algorithms (Liem *et al.*, 1993) designed for the explanation of data on shares include:

S-1 : this algorithm allows the user to specify the simplest root procedure of the SHARE class, or more complex ones:

LIN-LOGIT : **LIN**ear exponent **LOGIT** model;

BC-LOGIT : **Box-Cox LOGIT** model.

The first, LIN-LOGIT, is the classical LOGIT form with **linear-in-parameters** “representative utility” or exponent functions. The supplementary dimensions of the second procedure arise through the fact that the BOX-COX-LOGIT (Gaudry and Wills, 1978) is defined by adding Box-Cox transformations on all variables of a LOGIT. This makes possible the use of identical sets of explanatory variables in all exponent functions, an option that can be called the EXTENDED, or GENERALIZED, formulation (Gaudry, 1978) and provides a substitute to the hierarchical logit by addressing directly a cause of the correlation among residuals of alternatives in the multinomial logit. The STANDARD option is defined by using the same sets of explanatory variables as those one would use in the LIN-LOGIT.

S-2 : the STANDARD LOGIT model (Gaudry and Dagenais, 1979) adds to the LOGIT formulation parameters that make it possible to represent compulsive consumption, or captivity to alternatives. TRIO allows the user to simultaneously use Box-Cox transformations on the explanatory variables, as in S-1.

S-3 : the GENERALIZED LOGIT model (Gaudry and Dagenais, 1978) transforms the single-alternative “captivity” parameters of the STANDARD LOGIT model into parameters pertaining to each combination of alternatives. The TRIO user must decide of the functional form of the explanatory variables before using this option.

S-4 : The INVERSE POWER TRANSFORMATION (IPT) (Gaudry, 1981) makes it possible to define a first family, the LINEAR-IPT-LOGIT, where additional parameters applied to the exponential functions of the logit model itself represent captivity to alternatives and can make the shape of the response curve asymmetric. Moreover, they also allow the user to specify an EXTENDED formulation by the use of identical sets of explanatory variables in all exponent functions; a STANDARD option is defined by using the same sets of explanatory variables as those one would use in the LIN-LOGIT. The TRIO user is allowed to simultaneously use direct Box-Cox transformations on the explanatory variables.

S-5 : The second family, the BOX-TUKEY IPT LOGIT, offers the same possibilities as S-4, but through transformations applied to the representative utility functions of the logit model itself. The TRIO user is allowed to simultaneously use Box-Cox transformations on the explanatory variables.

PROBABILITY CLASS

The algorithms designed for the explanation of the occurrence of mutually exclusive states include:

P-2 : this procedure (Liem and Gaudry, 1987) allows the user to specify the simplest root procedure of the PROBABILITY class, or more complex ones.

LIN-LOGIT : **LIN**ear exponent **LOGIT** model;

BC-LOGIT : **Box-Cox LOGIT** model.

These procedures are analogous to the S-1 procedure and allow for STANDARD and EXTENDED options.

References

- Dagenais, M. G. (1982), “The Tobit Model with Serial Correlation”, *Economics Letters* 10, 3/4, 263–267.
- Dhrymes, P. J. (1971), “Full-information Estimation of Dynamic Simultaneous Equation Models with Autoregressive Errors”, Discussion paper No. 203, The Wharton School of Finance and Commerce, University of Pennsylvania.
- Fair, R. C. (1972), “Efficient Estimation of Simultaneous Equations with Autoregressive Errors by Instrumental Variables”, *Review of Economics and Statistics* 54, 4, 444–449.
- *Gaudry, M. (1978), “A Generalized Logit Mode Choice Model”, Publication #98, Centre de recherche sur les transports, Université de Montréal.
- Gaudry, M. (1981), “The Inverse Power Transformation Logit and Dogit Mode Choice Models”, *Transportation Research B* 15, 2, 97–103.
- *Gaudry, M. and Dagenais, M. (1978), “The Generalized Dogit Mode Choice Model”, Publication #107, Centre de recherche sur les transports, Université de Montréal.
- Gaudry, M. and M. J. Wills (1978), “Estimating the Functional Form of Travel Demand Models”, *Transportation Research* 12, 4, 257–289.
- Gaudry, M. and Dagenais, M. (1979), “The Dogit Model”, *Transportation Research B* 13B, 2, 105–111.
- Gaudry, M., Liem, T. C. and Sabourin, Y. (1981), “GAUSEQ: A Generalized AUtoregressive Single EQuation Estimation Procedure”, Publication #214, Centre de recherche sur les transports, Université de Montréal.
- Gaudry, M. and Liem, T. C. (1981), “GAUMEQ : A Generalized AUtoregressive Multiple EQuation Estimation Procedure”, Publication #215, Centre de recherche sur les transports, Université de Montréal.
- Liem, T. C., Dagenais, M. and M. Gaudry (1993), “LEVEL: The L-1.4 program for BC-GAUHESEQ regression — Box-Cox Generalized AUtoregressive HEteroskedastic Single EQuation models”, Publication #510, Centre de recherche sur les transports, et Cahier #8714, Département de sciences économiques, Université de Montréal. The first version was refereed in *The American Statistician*, June 1980.
- Liem, T. C. and M. Gaudry (1993), “PROBABILITY: The P-2 program for the Standard and Generalized BOX-COX LOGIT models with disaggregate data”, Publication #527, Centre de recherche sur les transports, Université de Montréal.
- Liem, T. C., Gaudry, M. and R. Laferrière (1993), “SHARE: The S-1/S-5 programs for the Standard and Generalized BOX-COX LOGIT and DOGIT and for the Linear and Box-Tukey INVERSE POWER TRANSFORMATION-LOGIT models with aggregate data”, Publication #899, Centre de recherche sur les transports, Université de Montréal.
- Maddala, G. S. (1983), *Limited-Dependent and Qualitative Variables in Econometrics*, Cambridge University Press, New-York.
- Parks, R. W. (1967), “Efficient Estimation of a System of Regression Equations when Disturbances are both Serially and Contemporaneously Correlated”, *Journal of the American Statistical Association* 62, 500–509.

* Stars denote papers that are integrated into “TRIO Model Types, Version 1.1”, Publication #904, Centre de recherche sur les transports, Université de Montréal, 1994.

OTHER TRIO DOCUMENTATION

- [CRT-441] • Babin, A., Dagenais, Florian, M., Gaudry, M., Guélat, J. , Lestage, P. and T. C. Liem "TRIO, An Open Interactive Graphic System for Demand Model Estimation", *Proceedings of the World Conference Transport Research 1986*, 1690-1704, The Centre for Transport Studies, University of British Columbia, or Publication #441, Centre de recherche sur les transports, Université de Montréal, January 1986 and March 1987.
- Describes the general design concepts and the open computer based system, shows examples of graphic output. Shows how TRIO was conceived using the same graphical user interface as EMME/2.*
- [CRT-901]
[AJD-101] • Gaudry, M. *et al.* "Cur Cum TRIO?" Publication #901, Centre de recherche sur les transports, Université de Montréal, April 1994.
- Summarizes in English, French, German and Spanish the professional, scientific and technical features of TRIO, notably the full integration of the four basic tasks required for regression work: the management of information, the production of models, the analysis of data or model results, and assistance in report generation. Shows examples of text, model results analysed with TABLEX tables and graphic output.*
- [CRT-902]
[AJD-102] • Gaudry, M., Lestage, P., Guélat, J. and P. Galvan, "TRIO Tutorial Version 2.0", Publication #902, Centre de recherche sur les transports, Université de Montréal, February 1994.
- Conducts the user through a first session with TRIO. Principal steps covered include: creation of a database, model estimation with LEVEL, SHARE and PROBABILITY algorithms, use of the TABLEX table to examine regression results and production of graphs.*
- [CRT-903]
[AJD-103] • Gaudry, M., Duclos, L.-P., Dufort, F. and T. C. Liem, "TRIO Reference Manual, Version 2.0", Publication #903, Centre de recherche sur les transports, Université de Montréal, November 1994.
- Provides full description of all functions available in TRIO, as they pertain for instance to the system environment, the database, the variable and variant editors for all model classes, the analysis of data and model results, through tabular and graphical methods.*
- [CRT-904]
[AJD-104] • Gaudry, M., Dagenais, M., Laferrière, R. and T. C. Liem, "TRIO Model Types, Version 2.0", Publication #904, Centre de recherche sur les transports, Université de Montréal, January 1995.
- Contains a summary of the regression theory and methods used in TRIO. Demonstrates the properties and usefulness of the methods with selected applications. Describes in detail all available model types or specifications, as well as the estimation techniques, which are programmed in TRIO.*
- [CRT-996] • Gaudry, M., Duclos, L.-P., Dufort, F. and P. Galvan, "TRIO Student Tutorial, Version 2.0", Publication #996, Centre de recherche sur les transports, Université de Montréal, August 1994.
- This student manual conducts the TRIO user through a first session with TRIO using a subset of the full TRIO TUTORIAL database and especially tailored examples. Principal steps covered include: creation of a database, model estimation with LEVEL, SHARE and PROBABILITY algorithms, use of TABLEX interactive table to examine regression results, and production of graphs.*