Key Features of Data and Model Contributions towards ETIS-PILOT Notably Made in Recent ETIS Workshops

by

Marc Gaudry

Agora Jules Dupuit (AJD)
Université de Montréal
Montréal, Canada
www.ajd.umontreal.ca ; marc.gaudry@umontreal.ca

Département économie et sociologie des transports (DEST)
Institut National de Recherche sur les Transports et leur Sécurité (INRETS), Arcueil, France
www.inrets.fr ; marc.gaudry@inrets.fr

The designations of persons in this text apply to men or women independently from the grammatical gender used
Indépendamment du genre grammatical, les appellations qui s’appliquent à des personnes visent aussi bien les hommes que les femmes

Written between the birthdates of Jules Dupuit (1804-1866) and François Quesnay (1694-1774)

Université de Montréal
Agora Jules Dupuit —Publication AJD-54
18th May to 4th June 2004, July 2005
Abstract

We specify a metamodel to classify contributions made to the construction of the ETIS by combining a 3-level representation of the transport system with the 4-step SPQR representation of data models and models of data defined by the Spotlights Thematic network in 2002.

In effect, we first claim that it is useful to adopt a representation of the transport system that is more complex than simple Demand-Supply representations of other economic markets due to specific characteristics of transport markets. We state that these characteristics can only be properly expounded and classified if a three-level representation of transport system equilibria is adopted and the Performance level of transport systems added to the classical 19th century two-level Supply-Demand structural clarification of previous single-level discussions of economic market outcomes.

We claim also that, once this enriched structure is adopted, it is also useful to combine it with the formal SPQR pedigree code that breaks up into four steps the processes involved in establishing data models (ancestor and derived indicator variables) or in establishing models of data (models proper and derived model results).

The resulting metamodel appears both necessary and sufficient to classify recent and, hopefully, also forthcoming contributions to the establishment of the ETIS.

Table of Contents

1. Introduction : economic flow accounts vs transport flow accounts ................................................................. 5
   1.1. Economic statistics: output and value added ..................................................................................................... 5
   1.2. Economic data: interindustry structure ............................................................................................................. 6
   1.3. Additional transport data specifics .................................................................................................................... 6
2. Transport specific data issues: space and multiple levels ......................................................................................... 7
   2.1. Space: an ETIS tradition already ....................................................................................................................... 7
   2.2. Multiple-layer systems ........................................................................................................................................... 7
3. The SPQR framework for the analysis of ETIS contributions ............................................................................... 10
   3.1. Data: Ancestor Variables ............................................................................................................................... ..11
   3.2. Data: Derived Indicator Variables .................................................................................................................. 12
   3.3. Models............................................................................................................................................................... 12
   3.4. Derived Model Results ..................................................................................................................................... 13
4. Key features of contributions ............................................................................................................................... 14
   4.1. Demand ............................................................................................................................................................ 14
   4.2. Performance ..................................................................................................................................................... 14
   4.3. Supply ............................................................................................................................................................. 15
5. Conclusion: combining three levels and four steps .............................................................................................. 16
6. References ............................................................................................................................................................ 16
1. Introduction: economic flow accounts vs transport flow accounts

To discuss the data of interest for the ETIS, or at least for ETIS-PILOT, it is important to remember the context in which such transport data may not fit in well with the traditional economic data that are the prime responsibilities of member state national or international statistical agencies, such as EUROSTAT or ECE UN, all present at ETIS meetings.

1.1. Economic statistics: output and value added

Indeed these national and international agencies naturally focus on economic data in the framework of the production of National Income and Expenditure Accounts. Even for such purposes, they face considerable difficulties linked to the rarity of public funds, even in advanced OECD countries, despite the relatively recognized theoretical basis for the construction of such accounts of good and service trades among individuals, firms and governments.

Of course, such tasks are not more easily accomplished in countries with recent Marxian accounting procedures and where the establishment of proper income and expenditure accounts may have little or no tradition. It has to be remembered that such developments were not without hurdles in many countries. Although some countries, like Canada, were among the first to develop National Accounts well before the Second World War, others, like Switzerland, have resisted having such accounts until recently and others were just slow: in Germany, for instance, national income measurement came more than 10 years after Canada because theoreticians doubted the existence of a link between such accounting measures and welfare.

Remember also that the measurement of National Income and Product, although started in tolerably recognizable form in 1696 by Gregory King (1648-1712), did not become a thorough practice until much later: 1932 in Great Britain (Clark, 1932) and 1934 in the United States (Kuznets, 1934). If a date is needed, the year of the publication of Hicks’ Social Framework (Hicks, 1936), may be used as a milestone as it became impossible to doubt the correctness of the national income tautologies equating income and output.

In so far as transport is concerned, such accounts, even accepting that they are constructed in value terms and not in quantity terms, pose many problems for our purposes: for instance, private (own account) transport is typically embedded the industrial sectors as they are defined and are not duly identified as transportation: whence the need for Transport Satellite Accounts (TSA) to split each sectoral line into its transport and other component. For a recent exercise of this sort for the USA, consult Fang, B. and X. Han (2004).

---

**1.2. Economic data: interindustry structure**

It might be expected that concern for interindustry flows would ignite interest in transport flows\(^2\). And such concern did arise after the regionalization of input-output matrices (Moses, 1955) itself following closely upon the development of country-wide input-output tables (Leontief, 1941, 1951).

Leontief notes that such tables had first been clearly (if not completely—due to the absence of key agents, investment and trade) conceived, but not filled with actual data, by Quesnay (1759), and thereafter forgotten despite the huge popularity, and multiple editions, of his “normative” *Tableau*. In effect, the interindustry tables clarify the value added process by grouping the intermediate and final outputs in interrelated fashion.

The existence of regional input-output tables prompted transport researchers, despite the sheer size and complexity of the task, to try to link such interregional trade flows to transport flows.

This linkage involved first the transformation from money units into ton or passenger units and then the matching of resulting flows to transport flows at the level of the networks, as effected for instance by Bigras *et al.* (1983) in Canada, by Cascetta and Di Gangi (1996) in Italy and by Yin and Williams (1998) in Great Britain.

Although some links clearly can be established between national accounting systems and transport flows, there is however no sense in which the specifics of transport required for transport models are consistently contained in national accounting systems.

**1.3. Additional transport data specifics**

In effect, transport ministries and firms everywhere have maintained transport-specific databases, some for a century or more, due to the specific needs of transport system management. More than in any other sector of the economy, perhaps, each transport data base is one of a kind and not easily nestable into a consistent set of accounts like national accounts.

Unfortunately, there are no “standard transport accounts”: the heterogeneous spatial and physical nature of transport will indeed pose data problems over and above those found in all economic and industry data. If all data are in reality authored or “signed”, transport data sets will be even more “unique” than general economic data sets may ever be.

Two issues are per force critical: the first is that of space, the second is that of the multiple-level structure of transport systems.

Let us say something about these in turn before we use formal means of isolation of key features of ETIS data contributions. In the absence of “transport accounts”, transport data are more closely related to the models in which they are used than general economic data. Naturally then, data needs will tend to be defined in the context of the use of models rather than in order to satisfy standard reporting requirements.

---

2 In Marxian accounting systems, the “transport branch” was not confused with other branches due to the use of physical quantities instead of monetary values.
2. Transport specific data issues: space and multiple levels

2.1. Space: an ETIS tradition already

It was clear more than 25 years ago, at the start of the very process that led directly to the ETIS, that, in the absence of “transport accounts”, transport data would be more closely related to the models in which they are used than general economic data are.

Indeed, there exists a body of continuous work on the design of the European Transportation policy Information (ETIS) system starting at the end of the 1970’s (COST 33, 1978) with work on intercity passenger demand modelling by seven research institutes\textsuperscript{3}. As a sequel, follow-up work by the same institutes soon stressed the need for regular, comparable and harmonized data for international purposes (COST 252, 1981).

As transport occurs in space among origins and destinations, dimensions formally absent from National Accounting systems and methodologies, a notable step in the gradual process establishing transport data and model needs was the recommendation, on the 12\textsuperscript{th} of May 1987 at the Brussels meeting of the COST 305 Action Management Committee (COST 305, 1988), to decide in favour of the NUTS\textsuperscript{4} zonal system (and equivalent zones in Non-EC Member Countries) hoping that in the long run the existing goods transport zoning system would become compatible with the subdivisions then proposed for passenger transport surveys and already in use for other regional statistics.

In the current state of affairs, the development of a first pilot ETIS site, this intertwining of data and models has not changed, even if our emphasis, in this analysis of recent contributions to ETIS workshops during the period 2003-2004, is on data.

2.2. Multiple-layer systems

The 19\textsuperscript{th} Century: from one to two levels. Until the end of the 18\textsuperscript{th} Century, explanations tended to relate variables of interest, say imports or exports, to hypothesized causes, whatever these may have been: money supply and the price level, etc.

Implicitly, single-layer (i.e. single equation) systems were effectively in use: not surprisingly, the expected directions of the various effects of “causes” were often confusedly expressed, as compared to what will be the case when observations on quantities or flows will be thought of as resulting from the interaction of supply and demand: it will then be clear that many “causes” jointly affect supply and demand and have effects of very uncertain directions on equilibria, effects often consisting mixtures of underlying structural effects of various strengths.

\textsuperscript{3} The seven institutes were: Centro Studi sui Sistemi di Transporto (Italy); Deutschen Forschungs-und Versuchsanstalt Fur Luft-und Raumfahrt (DFVLR, Germany); Ecole Polytechnique Fédérale de Lausanne (EPFL, Switzerland); Institut de Recherche sur les Transports (IRT, France); Netherlands Transport Research Institute (NVI, Netherlands); Royal Institute of Technology (KTH, Sweden) and Transport and Road Research Laboratory (TRRL, United Kingdom).

\textsuperscript{4} NUTS for \textit{Nomenclature des Unités Territoriales Statistiques}. 
Groenewegen (1987) reminds us that the phrase “supply and demand” was initiated in the context of price determination by Steuart-Denham (1767) and used infrequently until Ricardo (1817) used it in a chapter heading. Until the 1830’s, the terms were rarely used in the modern sense, i.e. as a function of price: Cournot (1838) was the first to give such a systematic and symmetrical exposition.

As one fleshes out these structural relationships, two structural equations, one for Supply and one for Demand (and an equilibrium condition), are deemed essential to the explanation of market data on quantities and prices in most markets.

The clarity of this structural mechanism progressively made it self-evident to all who read Cournot, or 4 to 5 years later Dupuit (1844), even if the empirical conditions under which one could “identify” each equation were not obvious until Working (1927) pointed out that such unscrambling (“identification”) from the data was possible for each equation as long the other moved independently: if you imagine that the supply schedule is fixed and that the demand schedule moves by itself, clearly observed points will draw the supply curve, and conversely…

The 20th century: from two to three levels in transportation. However, the simplest way of thinking of transportation, and perhaps also of other systems, is not to adopt this two-level Demand-Supply formulation: it is to add a third level, the determination of Performance that depends on both Demand and Supply, as researchers effectively do in structural transport analysis.

Some years ago, we introduced (Gaudry, 1976, 1979) this three-level structure to capture the fact that realized transportation service levels often differ from supplied service levels through a third and explicit level between the classical supply and demand levels. We first called the resulting structure « Demand-Cost-Supply » to distinguish it from « Demand-Supply » structures of classical Economics. In that new structure, costs denote realized money, time, crowding comfort or safety levels. We also estimated a complete three-layer bi-modal urban model system on these lines (Gaudry, 1980).

More conceptual details on a three-layer system. Naturally, using the D-C-S system instead of the classical D-S system gave rise to new unheard-of equilibria, such as the « Demand-Generalized Cost » equilibrium that differs from the « Demand-Supply » equilibrium within the same 3-layer system.

To make the enriched formulation more accessible within the wide transportation planning subculture, we then subsequently relabeled the D-C-S system as a D-P-S (Demand-Performance-Supply) system and changed the notation (Florian and Gaudry, 1980, 1983) to that used in Table 1 where, without loss of generality, the Supply dimension side relates simultaneously to road infrastructure and public transport road vehicle services in a potentially congested network.

In this representation of Table 1, the achieved Performance \[ P, C \] contains actual queues, the level of congestion and risk, as well as other forms of modal performance (effective capacity, occupancy or load factors and crowding, etc.) conditional on both actual Demand D and given Supply actions \[ S, T, F \].
For instance, in a network equilibrium there is a set of values of \( P \), \( C \) and \( D \) that simultaneously satisfy the demand functions and the conditions required by the performance procedures. For our purposes here, money and time performance by origin-destination pair on the network have to be consistent with the demands generated with these transportation conditions, a non-trivial problem as the dimensions of the demand functions (from \( i \) to \( j \)) are not the same as the transportation conditions on individual network links \( a \).

**Table 1. Market and Network Analysis: a Three-Level Approach**

<table>
<thead>
<tr>
<th>Demand Procedure</th>
<th>Performance Procedure</th>
<th>Supply Actions Procedure</th>
</tr>
</thead>
<tbody>
<tr>
<td>( D = \text{Dem}(P, C, Y, A) )</td>
<td>([P, C] = \text{Per}(D, [S, T, F]))</td>
<td>([S, T, F] = \text{Sup}(\text{SO}, \text{RE}, [(W(S^<em>, T^</em>))], \text{ST}))</td>
</tr>
<tr>
<td>( \text{ST} = (P^{<strong>}, C^{</strong>}, D^{**}) )</td>
<td>( W(.) ): set of minimum cost combinations for the realization of any scheduled ((S^<em>, T^</em>)) and where ( D^{<strong>}, P^{</strong>}, C^{**} ) denote realized values of demand, unit costs and service levels.</td>
<td></td>
</tr>
</tbody>
</table>

with:  
- \( D \): market demand  
- \( P \): out-of-pocket unit expenditures  
- \( C \): levels of service  
- \( Y \): consumer socio-economic characteristics and their budget  
- \( A \): economic activity  
- \( S \): quantity supplied  
- \( T \): scheduled service levels  
- \( F \): scheduled price, or fare  
- \( \text{SO} \): supplier objectives  
- \( \text{RE} \): regulatory environment  
- \( \text{ST} \): suppliers’ estimate of the state of the system

Other three-layer structural systems. This three-layer specification applies to many regulated markets, notably those of communist economies (Gaudry and Kowalski, 1990) or to sectors of market economies (such as health) where the prices and freely determined wait times are not allowed to clear the markets but are “centrally planned” and regulated.

Note in passing that modeling the three levels in this way to explain observations generated in the absence of D-S equilibrium is much simpler than using disequilibrium econometrics—a difficult combinatorial game—or some subtle forms of hysteresis. As an example of the former, Portes et al. (1987) stunningly conclude that the Polish economy was in «excess supply half of the time during the 1960’s and during the years 1976-1978»! It would have seemed more appropriate to build a model in which the length of the queues for housing, cars, etc. was explained by a Performance level, as we do in transport, and to raise in this context issues of identification without which there is a real “supply of revelations on centrally planned economies”, (Podkaminer, 1989).
Similarly, the explicit modeling of **Performance** (queues, etc.) should take precedence over the search for hysteresis in labor markets (Blanchard and Summers, 1986). In health studies, do not conclude that reducing the number of doctors reduces public health expenditures: look at the length of queues, at the black market and at the market for side-privileges and model these explicitly.

But of course, modeling **Performance**, as in transport, is hard work. It does not suffice to model the stationarity of socialist economy shortages (Kornai, 1982): “insatiable demand” does not exist, but queues do (Kornai and Weibull, 1977) and they reestablish equilibrium.

In consequence, it will be important to our classification to distinguish among the three layers: Demand, Performance Supply, but this is very important to transport and may become useful elsewhere as the “free” determination wait-time becomes an object of analysis along supply quantity\(^5\).

### 3. The SPQR framework for the analysis of ETIS contributions

**A metamodel, not an ontology.** In order to isolate the key features of contributions to ETIS workshops, we need a framework or metamodel. As we do not need to allow for the simultaneous application of different perspectives (of, for example, traveler, modeler, policy maker, etc.) to one and the same reality (that of ETIS data), we only need one metamodel to effect our selection of key features: we do not need an ontology like those in construction for medical data where all perspectives (doctor, patient, pharmacologist, geneticist, lawyer, priest, insurer, etc.) simultaneously matter for the database (Smith, 2002).

**Previous work towards the ETIS.** Fortunately, such a metamodel has been prepared with ETIS in mind by the SpotlightsTN project (2000-2002) which elaborated a documentation structure, or pedigree code, that could apply to data and models recognized by the ETIS. We will use this structure here and, to explain it, quote at length from the paper elaborated and approved by the project scientific committee (Gaudry *et al.*, 2002) that outlined the SPQR structure, “code” or metamodel of data and models.

**Four objects to which the four-dimensional SPQR code is applied.** The metamodel is applied to four types of data and modelling objects: two kinds of data (**variables** and **derived indicators**); two kinds of models (**models** proper and the use of **derived model result**s).

The SPQR treatment consist in qualifying for any object the 4 dimensions indicated in Table 2. It is difficult to state fewer dimensions than this, or to define a simpler vocabulary, as both intuitive and very formal claims somehow involve those steps or dimensions, even summarily.

The idea of having 4 elements, instead of the simpler mapping structure with 3 (input, relation, output) is to give some flesh to the « relation »: it is to distinguish between the presumed existence of relationships and the quantification of the relationships.

---

\(^5\) For instance, many firms determine their logistic structure by first deciding on the appropriate wait time for their product in each market: this wait time decides the number of depots necessary and their location.
Table 2. The SPQR quartet of dimensions applied to information objects

<table>
<thead>
<tr>
<th>Quartet</th>
<th>Step description</th>
<th>Idea of the step or dimension</th>
</tr>
</thead>
<tbody>
<tr>
<td>S</td>
<td>1. Input data sample</td>
<td>A datum is used; a regression sample; (“saisie” de données)</td>
</tr>
<tr>
<td>P</td>
<td>2. Propose hypothesis</td>
<td>A vision (theoria) is formulated</td>
</tr>
<tr>
<td>Q</td>
<td>3. Quantify relationships</td>
<td>A functional relationship is calculated</td>
</tr>
<tr>
<td>R</td>
<td>4. Report output results</td>
<td>A set of raw results is obtained</td>
</tr>
</tbody>
</table>

This SPQR foursome is indeed sufficient to handle the construction of elementary variables, of derived indicators, of models proper and of derived uses of model results, as we see for each in turn.

3.1. Data: Ancestor Variables

How are simple variables generated? Many data are collected by individuals from casual observation, but most data—especially economic data—are collected by firms and governments in the course of ordinary activities, particularly production and taxation activities. Most economic data, however, are actually produced by models—some of which are quite complicated. This is summarized in Table 3A.

Sampling. Sampling for ancestor variables is often informal and “unscientific” even if it involves complicated, more or less intuitive, procedures aimed at drawing a measurement. But often a “scientific” procedure is used with explicit (and often false) assumptions used to derive precision estimates for the sampled values. Whether the sample comes from « revealed» (actual) preference data, stated preference data or experimental [preference] data, the sampling rules always involve some theory and some calculation to obtain the target population result from the sample estimate. Samples can be casual or much more formal.

Postulated theory. Statistical sampling theory is known and appreciated. However many other theories are often involved in the construction of economic data. The construction of economic accounts relies on economic theory, for instance.

Quantification. Sophisticated computational procedures distinct from the theories proper are used to calculate target population estimates. They draw from many toolboxes: statistical distribution theory for some independent variables, matrix operations for some related variables (e.g. national income accounts may require use of multidimensional balancing and scaling), etc. Transport variables, such as time or cost among origin-destination pairs, result from abstract constructs.

Resulting variables. Clearly, even the simplest of variables are constructs. There is almost no produced ancestor variable that has not had judgmental interventions. All variables are in effect « authored », even the most reliant on experimental conditions. Measures of GNP exclude components linked to illegal activities and include judgmental constructs like “owner-occupied housing costs”. In transportation, there is no unique way to define even shortest path in a network because this depends on the representation of access arcs (dummy links) and the representation of the network (even in road networks where the presence of transfers among lines and common lines is not a real issue). General costs are even more complex constructs, as model-derived or assumed weights are used for different elements.
Simple economic activities like employment require non-trivial appreciation of part-time employment and hours worked.

Table 3. Pedigree elements for ancestor and derived variable indicators

<table>
<thead>
<tr>
<th>A. Ancestor variable pedigree</th>
<th>B. Derived indicator variable pedigree</th>
</tr>
</thead>
<tbody>
<tr>
<td>S Sample</td>
<td>S Ancestor Variable 1</td>
</tr>
<tr>
<td>P Theory</td>
<td>Ancestor Variable 2</td>
</tr>
<tr>
<td>Q Quantifier</td>
<td>R Population Value Var.1</td>
</tr>
<tr>
<td>—›</td>
<td>—›</td>
</tr>
<tr>
<td>P Theory</td>
<td>Derived Indicator 1</td>
</tr>
<tr>
<td>Q Quantifier</td>
<td></td>
</tr>
</tbody>
</table>

3.2. Data: Derived Indicator Variables

The same also holds for indicators obtained by combining «primitive» variables, which we call ancestor variables, as indicated in Table 3B. In both cases the quartet applies. In terms of sampling, the problems of obtaining derived indicators typically compound the problems associated with the individual ancestor variables, but this is not easily recognized as one performs simple algebraic or arithmetic operations on the ancestors to obtain them. This is also true of complex variables obtained by combining heterogeneous ancestor variables.

3.3. Models

It may be more obvious that the four dimensions easily apply to models. So we will be brief. For an overview, refer to Table 4A and Table 4B.

Sample. Model pedigrees take the input data, or variables, as given. So, «data models» belong to data pedigrees, not to model pedigrees: the latter deal with the modelling of data. There should be a pedigree for all input data (variables) used for a model, or at least for the input database of the model.

Postulated theory. Different kinds of theories are used for different models and the theory of each model should be justified in the P part of the pedigree.

Quantification. There are two principal aspects to quantification. The first has to do with statistical criteria used to obtain model parameters, and the second relates to the numerical nature associated with carrying out computations, since quantification methods or algorithms are in practice embodied in computer programs for specific hardware allowing for a certain level of precision.
Results. Raw parameters are obtained as primary models outputs. They normally have a value and dimensions that depend on units of measurement. More often than not, they cannot be understood or used without considerable difficulty. Whence the need to derive understandable statistics from them and study their properties.

3.4. Derived Model Results

In effect, using model results is a form of modelling that itself requires a pedigree. Consider for instance a system of non-linear simultaneous equations. Presume that the estimates are correct, because they result from correctly applied model Quantification procedures. Consider Table B.

Table 4. Pedigree elements for models and derived model results

Sample. The individual outputs of the model (specification, estimated parameters) constitute the raw information needed to derive information from the modelling exercise and constitute the primary input data for the derivation of model results; sometimes other elements, such as forecasts of variables outside of the domain of estimation, are also required as input to be “seized” for the derivation of results, for instance forecasts.

Postulated theory. Many questions asked about any given model cannot be answered without some modelling to obtain derived results. One question might be about the stability and stationary behaviour of the estimated system. Other questions might be about model fit, for each equation. Questions of marginal effects (partial derivatives) or rates of substitution, such as the value of time (the ratio of two partial derivatives) could be important. Or it will often be useful to forecast with the whole system because elasticities evaluated at a particular point may not be the desired result from the model.

Quantification. Answering the questions asked will involve a complex calculation quite distinct from, but derived from, the raw model proper and require particular ways of applying the postulated theory of derived results.
Results. Usable model results almost always require some such complexity, and perforce imply a pedigree. One must also bear in mind the policy questions being asked and how would the decision maker interpret, in turn, processed model results. The history of model use is a summary way of reporting on results: forecasting, for instance, is part of the results of a model.

4. Key features of contributions

Of course, there are many levels of documentation in SPQR, as remaining unquoted parts of the source paper makes clear: one might consider that, without a regulation to that effect, ETIS data (perhaps supplied by EUROSTAT), will not be satisfactorily documented.

In this context, however, our concern is simply with using the SPQR structuring to give significance to particular contributions, not to impose the SPQR pedigree car on recognised data. Consider Table 5 and the contributions in turn.

4.1. Demand

M: The contribution by Mandel (2004) indicates that a very complex set of models (including VAČLAV and VIA) is to be used to generate a passenger flow matrix on the basis of existing and available data enriched by the year 2000 Europe-wide DATELINE Survey. In Table 5, this contribution is understood as using Models of data to generate data.

It is a good, if very complicated, example of the use of models to merge various sources to generate what is in effect a “most likely” synthetic matrix resulting from the joint use of a full set of passenger models.

C: The paper by Chen (2004) also shows how data and models are combined to generate most likely Origin-Destination freight matrices.

Clearly, it would hardly be possible to make any statement about this very complicated procedure without the SPQR framework.

4.2. Performance

Three contributions are made to the estimation of performance.

Sw: Swimba and Schnell (2004) effectively present different data model for the construction of level-of-service variables for each of the passenger modes. For instance, the air service levels are sampled through the use of internet questions put to airline web sites for specifically defined Origin-Destination pairs; the information is combined with weights (choice probabilities) obtained from models proper to define indices of service.

Such procedures, based on internet sites can be automatized, as the program by Marchal (2004) has recently formalised.
**R&**: using in particular socio-economic data such as GNP, population and the growth rate of the local economy, Reynaud et al. (2004) show how these data can be combined to define designation criteria for the TEN-T and indicate how such indicators, based for instance on long-distance and transborder flow estimates, could conflict with local conditions.

**J**: Jantunen (2004) emphasizes how methodology used to assess external effects, such as accidents, emissions and noise, creates a demand for the necessary input data to compute such measures.

**Table 5. Key features of contributions to ETIS Workshops (2003-2004)**

<table>
<thead>
<tr>
<th>TRANSPORT SYSTEM</th>
<th>DEMAND</th>
<th>PERFORMANCE</th>
<th>SUPPLY</th>
</tr>
</thead>
<tbody>
<tr>
<td>Passenger (P) or Freight (F)</td>
<td>P</td>
<td>F</td>
<td>P</td>
</tr>
<tr>
<td><strong>DATA MODELS</strong></td>
<td>S</td>
<td>P</td>
<td>Q</td>
</tr>
<tr>
<td><strong>MODELS OF DATA</strong></td>
<td>S</td>
<td>P</td>
<td>Q</td>
</tr>
<tr>
<td>Authors</td>
<td>M</td>
<td>C</td>
<td>Sw</td>
</tr>
</tbody>
</table>

**4.3. Supply**

**S&**: Drawing from the currently ongoing TEN-STAC project, Schmedding and Schoch (2004) present both data models of models of data used to generate network supply variables for the different modes. In particular, synthetic rules are used in the passenger model to add local flows to long-distance flows.

**N**: Newton (2004) presents procedures for the establishment of freight network services and costs that are “classically” of a Data model sort.
5. Conclusion: combining three levels and four steps

To classify contributions made to the construction of the ETIS, a metamodel is required. We have argued that such a metamodel may be usefully specified by combining a 3-level representation of the transport system with the 4-step SPQR representation of data models and models of data.

In effect, we have first claimed that it is useful to adopt a representation of the transport system that is more complex than simple Demand-Supply representations of other markets due to specific characteristics of transport markets. We have stated that these characteristics can only be properly expounded and classified if a three-level representation of transport system equilibria is adopted and the Performance level of transport systems is added to the classical 19th century two-level Supply-Demand structural clarification.

We have also claimed that, once this enriched structure is made clear, it is then also useful to be formal about the data and model contributions made to the ETIS by applying to them the SPQR pedigree code recommended by the Spotlights Thematic Network. This code breaks up into four steps the processes involved in establishing data models (ancestor and derived indicator variables) or in establishing models of data (models proper and derived model results).

Our metamodel appears sufficient to classify recent and, hopefully, forthcoming contributions to the establishment of the ETIS.

6. References


Quesnay, F., Tableau Économique, 3rd edition, Versailles, 1759.


Steuart-Denham, J., 1767.