The Four Approaches to Origin-Destination Matrix Estimation for Consideration by the MYSTIC Research Consortium

Article · January 2000
Source: RePEc

1 author:

Marc J.I. Gaudry
Université de Montréal
162 PUBLICATIONS **1,183** CITATIONS

Some of the authors of this publication are also working on these related projects:

Multiplicative decompositions of aggregate road safety outcomes, notably in DRAG-type models View project

How should the Danish National Land Transport Model activities be housed and governed? View project

The Four Approaches
to Origin-Destination Matrix Estimation
for Consideration by the MYSTIC Research Consortium
by
Marc Gaudry

Agora Jules Dupuit (AJD)
Université de Montréal, Montréal
 www.ajd.umontreal.ca
Bureau d'économie théorique et appliquée (BETA)
Université Louis Pasteur
Strasbourg
Département de sciences économiques
Université de Montréal, Montréal
marc.gaudry@umontreal.ca

Project ST-97-SC.2101

Project Funded by the European Commission Under the Transport
RTD Programme of the 4th Framework Programme ;
(part of Work Package 3. « New Methodologies »)

April 1, 1998; revised July 20, 2002.

Abstract

In this paper, we say a few words about the context, role and classification of Origin-Destination (O-D) matrix determination procedures, with a view to situating existing research streams within a common framework. We distinguish between O-D matrix derivation approaches and proper O-D matrix estimation streams and provide rough outlines of such approaches, but without seeking a complete or exact account of the methodologies of interest. For the first class, we provide simple examples and explain some pitfalls of simplistic derivations based on an assumed proportionality to intermediate economic activity levels generating demand, or to trip ends. For the second class, we focus on the four « pure » approaches to estimation, denoted here as the Sampling, Programming, Quasi-behavioural and Regression streams (whence SPQR), without seriously considering the ways of combining them.

Keywords : origin-destination matrix, input-output, final and total demands, trip ends, network demand, network performance and network supply, O-D matrix derivation, O-D matrix estimation.

Résumé

Le présent article situe brièvement dans leur contexte les diverses méthodes utilisées pour obtenir des matrices Origine-Destination (O-D) dans le but de fournir une typologie des diverses traditions de recherche regroupées en deux grandes tendances. Il est alors souhaitable de distinguer entre les méthodes dites de dérivation des matrices O-D et les méthodes d'estimation au sens strict, sans toutefois rendre compte avec exactitude ou exhaustion des méthodologies pertinentes. S'agissant du premier groupe, nous illustrons les pièges des dérivations fondées sur une certaine proportionalité supposée entre les marges des matrices, ou les activités économiques intermédiaires qui génèrent la demande, et les flux O-D. Il nous est loisible de distinguer dans le second groupe entre l'estimation à partir d'échantillons des flux (S), par des méthodes de programmation (P) ou quasi-comportementales (Q), et par les méthodes de régression (R), sans pour autant envisager de combiner les éléments SPQR.

Mots clés : matrice origine-destination, intrant-extrant, demande finale et totale, marges des matrices, demande dans un réseau, performance du réseau, offre sur le réseau, dérivation des matrices O-D, estimation des matrices O-D.

Zusammenfassung

In diesem Artikel beschreiben wir kurz Inhalt, Bedeutung und Klassifizierung der Bestimmung von Quell-Ziel-Matrizen (O-D), mit dem Ziel die existierenden Forschungsrichtungen in einen allgemeinen Rahmen einzuordnen. Wir unterscheiden zwischen Ableitungs- und Schätzmethoden für O-D Matrizen und geben eine grobe Vorstellung dieser Ansätze, ohne jedoch die verwendeten Verfahren vollständig und exakt zu erläutern. In einer ersten Klasse geben wir einfache Beispiele und erläutern einige Fallen einfacher Ableitungen die auf einem angenommenen Verhältnis zwischen Nachfrage erzeugenden volkswirtschaftlichen Aktivitätsniveaus und deren Randsummen beruhen. In einer zweiten Klasse, konzentrieren wir uns auf die vier « reinen » Schätzansätze, hier beschrieben als Stichprobe, Programmierung, Quasi-Verhalten und Regressionsmethoden. (SPQR).

Stichworte: Quell-Ziel-Matrix, Input-Output, End- und Gesamtnachfrage, Randsummen, Netzwerknachfrage, Netzwerkperformanz, Netzwerkangebot, O-D Matrix Ableitung, O-D Matrix Schätzung.

Table of Contents

List of figures and tables

1. Introduction : the conditional nature of Origin-Destination matrix determination

We first say a few words about the context, role and classification of Origin-Destination (O-D) matrix « estimation » procedures, with a view to situating existing research streams within a common framework. We then distinguish between O-D matrix derivation approaches and proper O-D matrix estimation streams and provide rough outlines of such approaches, but without seeking a complete or exact account of the methodologies of interest.

Context. The establishment of Origin-Destination flows, either by derivation or by estimation methods, is a « top down » problem where the solutions are conditional on the current or presumed level and spatial distribution of economic activities, as illustrated in Table 1.

Table 1. The Top-Down Context of Origin-Destination Matrix Determination

Comments	
The level and distribution of final and intermediate activities is assumed to be given.	**The Economy**
	Aggregate transport expenditures
The O-D pattern is the set of flows consistent with these activities that are often measured in an aggregate way in money or quantity terms.	↓
Top Down link to spatialized networks —>	
The problem is to obtain estimates of the flows in order to be able to modify the network costs or services and calculate new flows, using models.	**O-D flows on Networks**
For each set of flows, it is possible to establish network accounts by mode, jurisdiction (community/national/regional) or function (intercity/municipal). These accounts may contain direct cost (financial, economic or ecological) and revenue values for infrastructure, traffic control and carrier services.	**Modal performance accounts** ---------------- derived firm and government financial recovery
The implicit perspective of O-D estimation is scenario analysis.	

This means that scenario analyses developed from O-D matrix estimates obtained either from derivation approaches or from estimation approaches, and flow models based upon them remain conditional on this level and distribution. We now turn to the first class of ways to obtain O-D matrices: by deriving them from changes in activities.

2. Derivation of matrices from modified activities or trip ends

We shall now outline methods that derive new O-D matrices from spatialized input-output analyses. It is sometimes thought that they can be used as such to determine the impact of changes in GNP or trip ends on transport flows by being adjusted with simple rules of proportionality. To see the complexities involved, and why proportionality is not a straightforward concept, we consider the simplest of cases and assume that there are fixed proportions between transportation and its uses.

We are aware that the world of fixed coefficients of input-output models has been superseded by the more flexible computable general equilibrium model world. However, as the CGE models have yet to be spatialized at the level of transport networks, the « ancestor » input-output models retain their usefulness beyond their pedagogical roles. For instance, ongoing work by the EUNET research consortium (Jin and Williams, 1998) uses input-output analysis to derive O-D matrices.

A. An input-output model without spatial dimensions

Then, if y denotes the (n_S x 1) vector of final demand by in S sectors in a closed economy, z the (n_S x 1) vector of total intermediate demands and x their (n_S x 1) sum, or total output, and assuming a linear technological relationship between sector outputs and intermediate demands given by the ($n_S \bullet n_S$) matrix of input-output coefficients A, such that $z = A \cdot x$, then we have:

(1a) $x = z + y$
(1b) $x = A \cdot x + y$
(1c) $x = (I-A)^{-1}y$

where I is the ($n_S \bullet n_S$) identity matrix and it is clear that changes in final demand y (for transport included) work their way through the economy to determine total resources used x (for transport included), and its intermediate demand component z (for transport included) in ways that are far from proportional. This matters very much for transport because freight demand is primarily of an intermediate nature and passenger demand partakes of both final and intermediate uses. This means that simple transportation models will incorrectly forecast all flows (in particular intermediate ones, and notably freight flows) unless the complex feedbacks involved in the inverse of **(I-A)** are taken into account, as we shall detail in Subsection C below.

B. Deriving new O-D flows from a fixed coefficient spatialized input-output mode

The system (1) is not explicitly spatial, but can easily be made so in order to show the trade (transport) flows among regions of the economy considered, as was recently done for the 20 administrative regions of Italy (Cascetta and Di Gangi, 1996), from whom we borrow the usual accounting system expounded in Moses (1955). Their open-economy specification for R regions is:

(2b) $x = T A \cdot x + T y - im$
(2c) $x = (I-T A)^{-1} (T y - im)$

where all vectors are of dimension ($n_R \bullet n_S$ x 1), all matrices are of dimension ($n_R \bullet n_S$ x $n_R \bullet n_S$), the vector **im** denotes imports from other countries by sector and region, the **A** matrix is now block-diagonal and made up of n_R blocks, each being a regional technical coefficient matrix ($n_S \bullet n_S$), and the new matrix **T** is the trade matrix made up of diagonal matrices, one for each region pair, with elements t (g, i, j) , inter-regional trade coefficients expressing the fraction of sector g production consumed in

region j and acquired in region i. Their examples of $\mathbf{x}$, $\mathbf{A}$ and $\mathbf{T}$ of Equations (2b)-(2c) for an economy of 3 regions ($n_R = 3$) and 2 production sectors ($n_S = 2$) are reproduced in Table 2.

In this model, region-to-region trade flows—**the intersectoral O-D flow matrix**—can be expressed as

$$(3) \qquad\qquad \{\, t^{gh}_{ij} \,\} = \mathbf{T\,A} <\mathbf{x}> + \ \mathbf{T} <\mathbf{y}>,$$

where the ($n_R \cdot n_S$ x $n_R \cdot n_S$) matrix of trade flows has elements t^{gh}_{ij} denoting the amount of trade from sector g in region i going into sector h in region j and $<\mathbf{x}>$ and $<\mathbf{y}>$ are ($n_R \cdot n_S$ x $n_R \cdot n_S$) diagonal matrices obtained from vectors $\mathbf{x}$ and $\mathbf{y}$ defined for (2).

It is clear that the intersectorial O-D matrices found in (3) do not respond proportionately to changed final demand $\mathbf{y^*}$, as this new final demand implies by (2c) a new total demand $\mathbf{x^*}$ determined through the complex sectoral and spatial feedbacks involved in the inversion of $\mathbf{(I\text{-}T\,A)}$ even if the matrix $\mathbf{T}$,

Table 2. Example in Cascetta and Di Gangi

		Region 1		Region 2		Region 3	
		Sector 1	Sector 2	Sector 1	Sector 2	Sector 1	Sector 2
Region 1	Sector 1	t(1,1,1)	0	t(1,1,2)	0	t(1,1,3)	0
	Sector 2	0	t(2,1,1)	0	t(2,1,2)	0	t(2,1,3)
Region 2	Sector 1s	t(1,2,1)		t(1,2,2)		t(1,2,3)	
	Sector 2	0	t(2,2,1)	0	t(2,2,2)	0	t(2,2,3)
Region 3	Sector 1	t(1,3,1)	0	t(1,3,2)	0	t(1,3,3)	0
	Sector 2	0	t(2,3,1)	0	t(2,3,2)	0	t(2,3,3)

Trade coefficient matrix $\mathbf{T}$(3*2 x 3*2)

		Region 1		Region 2		Region 3	
		Sector 1	Sector 2	Sector 1	Sector 2	Sector 1	Sector 2
Region 1	Sector 1	a(1,1,1)	a(1,2,1)	0	0	0	0
	Sector 2	a(2,1,1)	a(2,2,1)	0	0	0	0
Region 2	Sector 1s	0	0	a(1,1,2)	a(1,2,2)	0	0
	Sector 2	0	0	a(2,1,2)	a(2,2,2)	0	0
Region 3	Sector 1	0	0	0	0	a(1,1,3)	a(1,2,3)
	Sector 2	0	0	0	0	a(2,1,3)	a(2,2,3)

Technical coefficient matrix $\mathbf{A}$ (3*2 x 3*2)

Region 1	Sector 1	x(1,1)
	Sector 2	x(2,1)
Region 2	Sector 1s	x(1,2)
	Sector 2	x(2,2)
Region 3	Sector 1	x(1,3)
	Sector 2	x(2,3)

Vector of sectorial production $\mathbf{x}$ (3*2 x1)

which expresses the spatial structure of interregional trade, does not change: $\mathbf{y}^* \longrightarrow \mathbf{x}^* \longrightarrow \{ \mathbf{t^{gh}_{ij}} \}^*$. It is appropriate to call this approach O-D matrix **derivation**.

We note that the spatial structure of flows is entirely defined by the matrix **T**, which in effect spatializes the intersectoral and final flows with **ij** indexed coefficients that make it impossible to manipulate changes in trip-ends (**i**-indexed or **j**-indexed) without jointly manipulating the complementary dimension. But sometimes that is too rigid and it is desired to manipulate the changes in origin-demand distinctly from the changes in destination-demand. For this purpose, an adaptation is required that will throw light on the approach proposed by Hyman (1998) to derive O-D matrices from leg (station-to-station) counts for railways, using station choice models. Here the station choices are all-or-nothing and known but the illustration, drawn from Gaudry (1973) who had adopted it from an unavailable handwritten manuscript by Blankmeyer (1971), has the advantage of showing the pitfalls of neglecting « intermediate » network flows, either if new network flows are to be derived from new final demands, or (implicitly) if new demands are to be inferred from new network flows.

C. Deriving new O-D flows with distinct entering and leaving demands

An accounting framework. Consider the city defined by four zones of activity (1,2,3,4) and two transit lines, (BTM, TJ). Each letter stands for a station, and station T is a transfer station. Figure 1 describes the city and its transportation network. If the travel paths of all individuals during a given time period are known, it is possible to know the values of all elements of the flow Matrix 1 where the sources (origins) of passenger flows are read in the left column and the destinations are read across the top row.

Figure 1. A simple transportation system

B = Bourassa
T = Talon
M = Montigny
J = Jardin
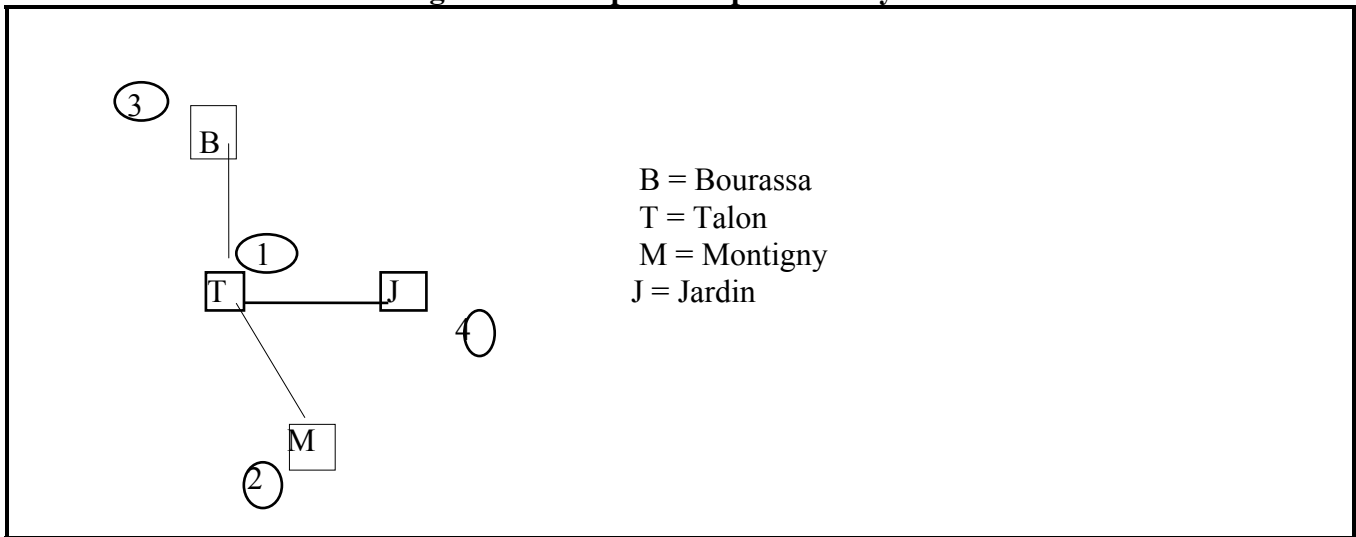
In Matrix 1, passengers are registered :
- when they enter the transportation network : e.g. 200 passengers coming from zone 3 enter the network at Bourassa. The sub-matrix of entering flows may be denoted **E** ;
- when they travel on a link : e.g. 40 passengers go from Montigny to Talon. We shall denote this sub-matrix of network flows by **N** ;
- when they leave the transportation network : e.g. 600 passengers left at station Jardin for zone 4. We shall denote the sub-matrix of leaving flows **L** ;
- the fourth sub-matrix is an origin-destination matrix **O-D** which we shall neglect henceforth.

<div align="center">**Matrix 1**</div>

From	To :	network					outside					
		B	T	M	J	NL	1	2	3	4	FL	TL
network	B		200			200			200		200	400
	T	200		400	600	1200	400				400	1600
	M		400			400		400			400	800
	J		600			600				600	600	1200
	NE	200	1200	400	600	2400					1600	4000
outside	1		400						100	300		
	2			400					100	300		
	3	200					100	100				
	4				600		300	300				
	FE	200	400	400	600	1600						
	TE	400	1600	800	1200	4000						

Matrix 1 may be rewritten in more compact format as Matrix 2 where, denoting the transpose by (′), $N = N'$, $O\text{-}D = O\text{-}D'$ and $E = L'$ because all trips are round-trips.

<div align="center">**Matrix 2**</div>

N	L
E	O-D

Let us examine the following vectors :
- the row vectors

      NE : sub-total of flows entering a station from another station,
      FE : sub-total of flows entering a station from the outside,
      TE : total flow entering a station,

(4)   **TE = NE + FE** ;

- the column vectors

      NL: sub-total of flows leaving a station for another station,
      FL : sub-total of flows leaving a station for the outside,
      TL : total flow leaving a station,

(5)   **TL = NL + FL**.

We have called **FE** and **FL** final flows because the purpose of the transportation system is to serve these flows. The sums of the elements of these vectors are assumed to be equal : they represent the total number of trips made in the network. If e is a column vector of ones, we must have during a given period :

(6) $\mathbf{e'FE'} = \mathbf{e'\ FL} = 1600$,

which simply states that the sum of those who have entered the system from the outside must have left it. This is simply an assumption. It is not an accounting identity. The basic identity of the accounting framework is that for each point (B, T, M, J) the total flows in, out or in transit (in transit flows are

assumed to be zero in our example) must be identical. Thus 1600 passengers entered Talon fom outside the transportation network or from other network points and 1600 passengers must go from Talon to other points in the system or outside. Or, for all network points :

(7) $\mathbf{TE'} = \mathbf{TL}$.

This identity arises because of the accounting procedure. In our example, the sub-totals are also equal because we have made the assumption that the pattern of entering demand is the same as the pattern of leaving demand (all trips are round-trips). So we also have :

(8) $\mathbf{NE'} = \mathbf{NL}$ and

(9) $\mathbf{FE'} = \mathbf{FL}$,

but, like (6), (8) and (9) do not follow from the accounting framework : only (7) does.

Accounting convention (7) will allow us to make consistent forecasts about total passenger flows at given stations and about passenger flows on any link of the network from forecasts of final demand. We assume that estimates of leaving demand $\mathbf{FL}$ are available and make forecasts of the geographic distribution of flows of interest from these estimates. A symmetrical analysis could be made using entering demand $\mathbf{FE}$ (in our example, it would give identical numerical results because $\mathbf{FE'} = \mathbf{FL}$).

Total flows through stations. Consider the sub-matrix of network flows $\mathbf{N}$ and divide each column of this matrix by the appropriate element of $\mathbf{TL'}$. Call the new matrix $\mathbf{A}$.

	B	T	M	J
	0	200/1600	0	0
(10) $\mathbf{A} =$	200/400	0	400/800	600/1200
	0	400/1600	0	0
	0	600/1600	0	0

The typical element of the symmetric matrix $\mathbf{A}$ is a_{ij} , the number of passengers going through i per passenger going through j ; thus $a_{21} = 200/400 = 0{,}5$ and 5 passengers go through Talon for every 10 passengers going through Bourassa from every origin. Of course, $a_{22} = 0$ because we have made the simplifying assumtion that no « in transit » passengers go through Talon. The coefficients a_{ij} may be called the « direct coefficients » : they show the flow through i required to deliver one passenger through j if we ignore the « indirect flows ». We will assume that the elements a_{ij} are constant over time with respect to changes in final demand $\mathbf{FL}$. Were the patterns of activities to change over time in a systematic fashion, the a_{ij} could be made to change a certain amount every time period.

The matrix $\mathbf{A}$ allows us to express network flows in terms of total flows. One may indeed verify that

(11) $\mathbf{NL} = \mathbf{A} \cdot \mathbf{TL}$, for any period t.

To forecast total flows through transportation network points, we recall from (5) that

(12) $\mathbf{TL} = \mathbf{NL} + \mathbf{FL}$

and substitute for the value of **NL** defined in (11) :

(13)   **TL = A•TL + FL**

from which we find :

(14)   **TL - A•TL = FL**
        (I-A) TL = FL
            TL = (I-A)$^{-1}$ FL ,

where **(I-A)$^{-1}$** is the inverse of **(I-A)** and **I** is the identity matrix of order 4.

Expression (14) can be used to forecast total flows **TL*** through network points during period **t** if we are given the **A** matrix and forecasts of final demand **FL*** :

(15) **TL* = (I-A)$^{-1}$ FL***

We note that the elements of **(I-A)$^{-1}$** , denoted **b$_{ij}$** , are called « indirect » coefficients. They mean that we expect an increase in flows through **j** to be acompanied by proportional new flows through **i**. The matrix **(I-A)$^{-1}$** can be written :

(16)   **(I-A)$^{-1}$ =**

$$
\begin{vmatrix}
1-a_{11} & -a_{12} & -a_{13} & -a_{14} \\
-a_{21} & 1-a_{22} & -a_{23} & -a_{24} \\
-a_{31} & -a_{32} & 1-a_{33} & -a_{34} \\
-a_{41} & -a_{42} & -a_{43} & 1-a_{44}
\end{vmatrix}^{-1}
$$

which in our case is equal to

(17)
$$
\begin{vmatrix}
1 & -0,125 & 0 & 0 \\
-0,5 & 1 & -0,5 & -0,5 \\
0 & -0,25 & 1 & 0 \\
0 & -0,375 & 0 & 1
\end{vmatrix}^{-1}
=
\begin{vmatrix}
0,6675 & 0,125 & 0,0625 & 0,0625 \\
0,5 & 1,0 & 0,5 & 0,5 \\
0,125 & 0,25 & 0,25 & 0,125 \\
0,1875 & 0,375 & 0,375 & 0,0125
\end{vmatrix}
$$

Because the **b$_{ij}$** reflect the inter-relationships of the flow pattern, we have for instance that **b$_{34}$** = 0,125 because when 1000 more passengers go through Jardin (from all origins), about 125 more passengers go through Montigny (from all origins). This occurs despite the fact that **a$_{34}$** = 0 , namely that no one goes <u>directly</u> from Montigny to Jardin in our example.

The two-equation demand model forecasts the results of the sum given in (6), namely final trips or the sum of the elements of **FL***. Appropriate assumptions can be made about changes in the distribution of the sum among the elements of **FL***.

Had we started with **FE** instead of **FL**, the matrix **A** would have been defined as the matrix **N** where the elements of each row are divided by **TE′**. Let us call the new Matrix **Â**. In our problem **Â = A′**. We would have had the analog of (11) :

(18) **NE = Â•TE**

and the final result would have been

(19)   **TE* = (I-Â)⁻¹ FE***,

where **FE*** denotes forecasts of final entering demand **FE**.

We have shown different ways in which matrices associated with a system of accounts used in interindustry economics can be derived from new demand conditions and pointed to the importance of intermediate « feedbacks » in this process and, by implication, in the reverse process whereby one would start from new flow matrices **N** and infer new demands that generated these network flows.

3. Estimation of matrices : pure approaches to the problem

By contrast with matrix **derivation**, O-D matrix **estimation** to which we presently turn aims at measuring in terms of money, quantities or vehicles, elements of the matrix { t^{gh}_{ij} } in (3), or their sums by O-D pair, and to relate these flows to actual networks, a point we have to clarify a little before identifying O-D estimation approaches.

A. The context : joint determination of demand and link flows

Except in toy networks, there is no simple relationship between O-D matrices, defined from i to j , and networks, consisting of connected links [a]. As most methods of O-D matrix estimation will make use of link counts, we need a minimum terminology and rough formulation of the relationship between O-D matrices and link counts to proceed. To do this, we shall use a simplified 3-level representation of network equilibrium to state the two-way relationship between O-D matrices and link counts and to present an accounting identity of relevance to our problem.

Some years ago, we introduced (Gaudry, 1976, 1979) a 3-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 or *safety* levels.

A three-level system. Naturally, using the D-C-S system instead of the classical D-S system gave rise to new equilibria, such as the « Demand-Generalized Cost » equilibrium that differs from the « Demand-Supply » equilibrium within the same 3-layer system. We then 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 3 to make it more accessible within the wide transportation subculture.

In this representation of Table 3, the **performance** level determines 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** and given supply actions. 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 links a.

Table 3. Market and Network Analysis : a Three-Level Approach

$$D = Dem\,(\,P,\,C,\,Y,\,A\,)$$

DEMAND PROCEDURE

$$[\,P,\,C\,] = Per\,(\,D,\,[\,S,\,T,\,F\,]\,)$$

PERFORMANCE PROCEDURE

$$[\,S,\,T,\,F\,] = Sup\,(\,SO,\,RE\,,\,[(\,W\,(\,S^*,\,T^*\,)]\,,\,ST)$$
$$ST \equiv (\,P^{**},\,C^{**},\,D^{**}\,)$$

SUPPLY ACTIONS PROCEDURE

where : 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   SO : supplyer objectives
   T : scheduled service levels   RE : regulatory environment
   F : scheduled price, or fare   ST : suppliers' estimate of the state of the system

   [W(.)] : set of minimum cost combinations for the realization of any scheduled (S*, T*)
   D**, P**, C** denote realized values of demand, unit costs and service levels

However, consistency requires that the following identity hold for each link flow:

(20) $v_a = t_{(1,2)} \bullet p_{\cdot(1,2),\,a} + t_{(1,3)} \bullet p_{\cdot(1,3),\,a} + ... + t_{(q)} \bullet p_{\cdot(q),\,a} + ... + t_{(Q)} \bullet p_{\cdot(Q),\,a}$

where $t_{(q)}$ denotes one of Q origin-destination pairs and $p_{(q),\,a}$ designates the proportion of this flow that uses link **a**, contributing to the total flow on this link, v_a. Now if **t** is the O-D matrix represented as a column vector and **v** is a column vector representing traffic volumes on all links of the network, then

(21) $v = \flat\,t$,

where $\flat$ is the proportion matrix with elements $p_{(q),\,a.}$, is the convenient vector notation form of (20). In fact, as it is not possible to observe link counts without errors of observation **e** , the proper specification of (21) is :

(22) $v° = \flat\,t + e$,

where it is clear that the explanation of observed link counts poses the related problems of errors of observation and of the nature of $\flat$. In practice, one estimates $\flat$ with an assignment procedure v = M(t) that produces an « assignment map » (Cascetta and Nguyen, 1988) that differs from the true one but for which (21) holds by construction because assignment procedures do not usually introduce random flows.

In view of the interrelated nature of demand and link flows indicated in Table 3, a crucial decision is whether the problem at hand can be treated without congestion, as we first assume, leaving the discussion on congestion for later : we concentrate on the uncongested case, where say $\mathbf{M(t)} = \mathbf{p^f\ t}$, because the proportion matrix $\mathbf{p^f}$ is assumed to be fixed and independent from the trip vector $\mathbf{t}$. In terms of Table 3, this amounts to removing the level of demand D from the determinants of performance.

Preliminary approach zero. But naturally, if the proportions $\mathbf{p}$ are known, or can be easily determined because there is no congestion, and a shortest path or other demand independent assignment is adequate, the question arises whether the vector $\mathbf{t}$ could not be obtained simply as the vector of regression coefficients by applying a least squares algorithm, namely :

Min dist $(\mathbf{v}, \mathbf{v°})$ with respect to $\mathbf{t}$ using Ordinary Least Squares	(Approach 0)

where the distance function to be used is e'e (the residual sum of squares). This tempting approach of the accounting identity (20)-(21) is not possible in general (Robillard and Trahan, 1973; Robillard, 1974, 1975) for the obvious reason that, with $\mathbf{n}$ zones, there are n^2 regression coefficients in vector $\mathbf{t}$ (or n^2 -n if the local flows on the diagonal of the O-D matrix are neglected) a very large number relative to the number of potentially available link counts: as the EU has about 1500 zones at the NUTS 3 level, there would in principle be 2 248 500 regression coefficients (!), requiring for estimation at least as many link counts... to say nothing about the problems with matrix inversion of the matrix of « observations » $\mathbf{p}$ that is full of zeroes. Clearly, **Approach 0** has too many parameters for the data available and is not applicable without further identifying restrictions, as we shall expound presently in our discussion of certain « pure » approaches to the matrix estimation problem.

B. The Sampling approach S

If sampling is used to obtain estimates of the O-D flows, each t_{ij} element of vector $\mathbf{t}$ is a random variable : the independence of drawings made not from a general sample of the population but from « choice-based » samples such as road side interviews **(rsi)** poses non trivial problems as there may well be spatial correlation among drawings due to the network topology (as networks are not ubiquitous) and path choice behaviour. Many complicated questions arise (Hautzinger, 1977). To the extent of our knowledge, none of the available sampling approaches have taken into account the spatial correlation that is expected in *rsi* samples. Still, can one state an objective of the sampling exercise and and refer the reader to the forthcoming summary within the context of the MYSTIC consortium (Clavering and Hautzinger, 1999).

Let us define **imprecision** of an estimate $\mathbf{t_{ij}}\sim$ as

(23) Imprec^{95} ($\mathbf{t_{ij}}\sim$) = max $\{U_{ij} - \mathbf{t_{ij}}\sim , \mathbf{t_{ij}}\sim - L_{ij}\}$,

where $\mathbf{U_{ij}}$ and $\mathbf{L_{ij}}$ are respectively the upper and lower limits of the confidence interval of $\mathbf{t_{ij}}\sim$ with 95 % confidence for an assumed distribution (say, Normal) of the sample estimate.

Let us also define the subset of links that matter as the set of TEN links :

(24) (v^{TEN}) is a subset of (v) .

and the subset of $t_{ij}\sim$ estimates that matter as that using the TEN **:**

(25) $v^{TEN}\sim = \textbf{þ}\, t^{TEN}\sim$.

Then, a possible rule is to :

Minimize the sum of $Imprec^{95}$ $(v^{TEN}\sim)$ subject to a given sample size over all links belonging to the TEN	Un-weighted TEN reference Sub network sampling rule

where *rsi* samples cannot be independent from the assignment matrix **þ,** but that precision statistics should be calculable for each link belonging to the TEN as this matrix is assumed fixed.

In many countries, concern for the accuracy of O-D matrix estimates directly obtained from expanded roadside, terminal, on board or household survey data, is longstanding (in Canada, see Dagenais *et al*., 1979 ; in the United Kingdom, see Department of Transport, 1982), but straightforward discussions of this issue only appeared in scientific journals more recently (e.g. Kuwahara and Sullivan, 1987). The most advanced procedures assessing formally the accuracy of estimates were developed in the United Kingdom, within the MATVAL/ERICA stream of work (DETR, 1998 ; Kirby, 1997 ; Sandman Consultants, 1998), concentrated on *rsi* accuracy estimates.

Trip samples are assumed to be drawn from from the larger population without replacement (according to a hypergeometric or approximate Poisson distribution) or with replacement (according to a binomial or normal distribution), readily allowing for variances to be obtained for each cell value in the matrix. In the case of trips with similar characteristics (e.g. of any length and for any purpose by a mode), it is also possible to merge different samples and obtain variances for the estimates derived from the joint sample, if the zonal system is the same for the different surveys (and sometimes after applying some harmonization procedures to insure identical target populations). Cell estimates are then weighted averages of the original sample values, with weights based on indices of dispersion or variances of the components: analogous weights also serve to obtain variances of the merged values. Procedures in existence before MYSTIC have been presented in Clavering and Kirby (1994) and in DETR (1997).

Matrix merging procedures should not be confused with matrix combining procedures, when the level of spatial detail of two surveys is not the same for all zones, as occurs for instance when national zones are finer that foreign zones in two national surveys and it is desired to combine them, as shown in Figure 2 where each national survey appears to have NUTS-1 national zones and NUTS-0 foreign zones. To obtain NUTS-1 zones for the combined matrix, biproportionate balancing might be used.

Figure 2. Matrix combination problem

*C. The predominantly Programming solving or optimizing approach **P***

We define an optimizing approach as one in which the O-D matrix vector **t** is the result or ***solution*** of a mathematical maximization procedure : more precisely, the vector of $\mathbf{n^2}$ solutions to the problem. There are two classes of approaches in which the vector **t** is obtained as the direct and immediate result of a problem formulation in which a function is maximized ***with respect to* t** *itself* : those obtained as the answers to the formulation of a mathematical program (usually a convex program with linear constraints) and those obtained as regression coefficients of a linear regression problem (ordinary least squares or a generalization of it). In both cases, information that is ***not ij-indexed*** is used in the problem formulation, sometimes in combination with a prior « reference » observed matrix vector *t°*, an approach used frequently and often surveyed (e.g. Abrahamsson, 1998).

It is important to note that this approach uses no behavioural aggregate or disaggregate structural demand model defined in terms of socioeconomic (e.g. Population, Income) or network (e.g. the money or time cost) variables to derive an answer. In fact the only structural variable which is used in its un-weighted linear form in some cases is the travel cost. This is a crucial distinction with the structural approach, where the parameters associated with these variables will be solutions of the problem. As we comment on the two streams of interest here, we will use Table 4 as a sort of summary. We provide here a partial, imprecise and low-brow version of Cascetta and Nguyen (1988), to whom we refer the reader in search of a rigourous treatment of optimization approaches.

Programming models. The idea of all non-sampling appraches to O-D matrix estimation is to find a mechanism that will select or determine a matrix among all possible ones. A first set of approaches establishes the simplest mechanism consistent with the information available, based on the principle of **maximum entropy**. This principle is used in Physics to describe the way in which the elements of a closed system tend towards an arrangement that is the most likely and simultaneously has the greatest « disorder »—or contains the minimum of information. Wilson (1970) had used this approach in transport, notably to derive the fully constrained gravity model. Willumsen (1978) had the idea of maximizing entropy subject to the information contained in the assignment map M(t) = v°**,** where the link counts are **observed**. In parallel, Van Zuylen (1977) used a similar derivation, but using a minimum information formulation that incorporated a « reference » or « prior » matrix *t°* = { t_{ij}~ } in the objective function. These authors had « vehicles » or « trips » in mind as they defined the link counts. These counts constrain the « most likely » or « least informative » arrangement.

A few year later, people trying to obtain the intercity freight flows also drew inspiration from Wilson (1970) who had in particular shown in his book that entropy maximization could be applied to find the most likely inter zonal flows that met the intermediate and macro constraints (1) and (2) of the economy and consequently respected their nationally defined (unspatialized) technical input-output coefficients, a formulation that included as a special case the Leontieff and Strout (1963) gravity form used to move away from constant regional input-output coefficient formulations used in O-D matrix derivations from spatialized input-output accounting systems or models (the trade matrices **T** in (2b) and (2c) above).

In Part I of a freight study, a formulation using an entropy/information type maximand was used by Bigras *et al.* (1983) to derive the most likely interprovincial flows among Canadian provinces that met the national and provincial input-output constraints of type (1)-(2) for 64 commodity groups covering the complete set of freight flows in Canada. In contrast with the Italian case, where the 20 regional input-output matrices for 1991 (FORMEZ, 1995) identify the regional origin of each the 11 commodity groups , the Canadian ones did not, thereby prompting the formulation used.

Table 4. Predominantly programming optimization approaches

Max f(t) with respect to	Form of f(t) — Sum over (all i *and* j)	Subject to all $\mathbf{t}_{ij}$ nonnegative and (all i *or* j		M (t) = v° observed link flows	Subject to — M (t) = v calculated link flows	Authors Year	#
$\mathbf{t}$ each element t_{ij} of the vector	-($\mathbf{t}_{ij}$ Log$_e$ $\mathbf{t}_{ij}$) entropy	$\mathbf{c}_{ij}$ $\mathbf{t}_{ij}$ = C total cost constraint type	$\mathbf{O}_i$ and $\mathbf{D}_j$ trip end constraints	$\mathbf{M}$ (t) = v° observed link flows	$\mathbf{M}$ (t) = v calculated link flows		n^2
"	"	**Programming Models**		"		Willum-sen1978	n^2
"	-($\mathbf{t}_{ij}$ Log$_e$ [$\mathbf{t}_{ij/}$ $\mathbf{t}_{ij}\sim$]) information			"		Van Zuylen 1977	n^2
$^\mathbf{g}\mathbf{t}$ each element $^g t_{ij}$ of the vector ($)	and sum over all goods g -($^\mathbf{g}\mathbf{t}_{ij}$ Log$_e$ $^\mathbf{g}\mathbf{t}_{ij}$)	total value of all goods	by region r			Bigras *et al.* 1983	g n^2
all $^\mathbf{g}\mathbf{t}$ flows are in money or are in tons	-($^\mathbf{g}\mathbf{t}_{ij}$ Log$_e$ [$^\mathbf{g}\mathbf{t}_{ij/}$ $^\mathbf{g}\mathbf{t}_{ij}\sim$]) information ; $^\mathbf{g}\mathbf{t}_{ij}\sim$ from a model for each good g	total $ value or total tons	input-output			Picard *et al.* 1985	g n^2
trucking flows $\mathbf{t}$ for a given good g are in tons	-($\mathbf{t}_{ij}$ Log$_e$ [$\mathbf{t}_{ij/}$ $\mathbf{t}_{ij}\sim$]) information ; $\mathbf{t}_{ij}\sim$ from a sample for each good g		inter-regional flows also known			Picard and Gaudry 1993	n^2
$\mathbf{t}$	$\mathbf{p}_{ij}$ $\mathbf{t}_{ij}$ - $\mathbf{t}_{ij}\sim$ [Log$_e$ $\mathbf{t}_{ij}$] where $\mathbf{t}_{ij}$ = $\mathbf{t}_{ij}\sim$ / $\mathbf{p}_{ij}$ and $\mathbf{p}_{ij}$ is sampling rate of Poisson process of mean $\mathbf{p}_{ij}$ $\mathbf{t}_{ij}\sim$	possibly		"		Spiess 1987	n^2
$\mathbf{t}$	Squared regression errors $\mathbf{e}$	**Regression model**			Solution values satisfy this constraint	Cascetta 1984	n^2

But, as NUTS-1 interregional flows are not good enough for transport if the networks of interest are interurban in nature, it can be useful to obtain the most likely interurban flows subject to regional or provincial constraints. So Picard *et al.* (1985) defined in Part II of the study a more disaggregated version of the model to simultaneously obtain, for the same 64 commodity groups for 67 zones—in which case the product g n^2 equals 287 296—defined at NUTS-3 type (interurban) level, ***total flows*** by value or by tonnage, using a « reference » initial value defined by a transport-sensitive model. In effect, this multiplicative-type regression model for each commodity yielded the kinds of results obtained by sophisticated intercity freight models and the maximization procedure « corrected » these flows in order to make them compatible with the input-output structure of the economy. This made it possible to show that, when transport conditions were modified, standard trip distribution models tended to overstate the impact on transport demand because they failed to take the constraints into account. A good summary of this stream of work can be found in Picard *et al.* (1988).

In a the Third step of the study, flows by mode had to be obtained. For instance, as the totals shipped or received by zone by ship and rail were well known, and the flows by truck were known on an interprovincial basis, the truck « trip ends » were obtained as a residual and the following problem was formulated (Picard and Gaudry, 1993) , which we show in full to demonstrate how the top-down constraints of the economy described in Section 2 above can be used to derive O-D matrices by mode and commodity (as the problem was solved separately or each of the 64 commodities—not jointly as in the model of totals specified for Part I and Part II) :

$$(26) \qquad MIN \sum_i \sum_j T_{ij} \ ln \ \frac{T_{ij}}{\widetilde{T}_{ij}}$$

subject to

$$(27) \qquad \sum_j T_{ij} = 0_i \qquad\qquad i = 1,...,67;$$

$$(28) \qquad \sum_i T_{ij} = D_j \qquad\qquad j = 1,...,67;$$

$$(29 \qquad \sum_{i \in I} \sum_{j \in J} T_{ij} = T_{IJ} \qquad\qquad I,J = 1,...,8;$$

$$(30) \qquad T_{ij} \geq 0 \qquad\qquad i,j = 1,...,67;$$

where:

T_{ij} : freight flow forwarded by truck from zone i to j ;

$\widetilde{T}_{ij}$: sample (CIGGT) freight flow forwarded by truck from zone i to j ;

T_{IJ} : freight flow forwarded by truck from province I to J ;

O_i : difference between total production and freight flows shipped by rail and boat from zone i ;

D_j : difference between total consumption of zone j and freight flows arriving in j by rail and boat.

It is possible to add to this problem specification the requirement that the total cost of transport sum up to the total known expenditure on transportation or, as Nguyen (1983) has done, to replace this total cost constraint by constraints derived from the information on utilized paths and traffic counts (22). Also, Spiess (1987) has provided an interesting interpretation of the « reference matrix » vector t_{ij}~ as a Poisson distributed variable : in this case the reference matrix is not just helpful but a central part of the resulting Maximum likelihood maximand shown in Table 4. With given marginal totals (trip ends) and a uniform sampling rate, this objective function collapses to t_{ij}~ [Log$_e$ t_{ij}], an entropic homomorph.

Regression models. We noted above (Approach 0) that Equation (22) cannot be exploited directly because the number of observations on links counts is insufficient. Cascetta (1984) augments the number by using estimates of the O-D matrix $t_{ij}\sim$ and defining a new vector of observations $[t_{ij}\sim, v^o]'$ on the dependent variable, combined with new regressors $[I, p^f]'$, with I the identity matrix of size n^2 by n^2. This extension of (22) makes it possible not only to obtain t as regression coefficients but also to start dealing formally with problems on heteroskedasticity or non-independence among the residuals by generalized least squares. Bell (1991) has recently proposed restrictions to insure non negativity of the fitted values. This is a minority sub-stream, without significant large or known applications.

D. The Quasi-behavioural production approach Q

In all above cases, the desired flow matrix is the variable with respect to which the maximization occurs. The rule used to formulate the maximand—the « demand » model—is extremely simple and has no proper behavioural component. We now want to turn to a set of approaches that are slightly older that the above and are « **Quasi-behavioural** » in the sense that the trip rates estimated can be interpreted as technological/behavioural coefficients in the spirit of trip generation and trip attraction coefficients : they are a half-way house between the above (that have n^2 parameters with the help of which the O-D matrix is estimated) and the still older behavioural models (that have k parameters—one per structural variable) to be described in the next section. The overparametrization problem in (20)-(22) can be solved by reducing the number of parameters to be estimated. For instance, if the unknown t_{ij} are set equal to the product of generation and attraction factors for each zone,

(31) $t_{(ij)} = G_i \cdot A_j$

the number of unknown parameters falls from ($n^2 - n$) to $2n$ and the problem becomes a manageable nonlinear minimization problem (of e'e with respect to the A_i and B_j) if the matrix p^f is assumed to be known ; the products of estimates of the A_i and B_j yield estimates of t_{ij} flows from link counts. This approach defines a « conceptually transitional » trip rate approach :

(32) $v_a = A_1 B_2 \cdot p_{\cdot(1,2),a} + A_1 B_3 \cdot p_{\cdot(1,3),a} + ... + A_i Bj \cdot p_{\cdot(q),a} + ... + A_Q B_Q \cdot p_{\cdot(Q),a} + e_a$.

Moreover, imagine now that survey estimates $t_{ij}\sim$ and cordon estimates $A_i\bar{}$ and $B_j\bar{}$ of trip ends are both also available. One may then simply write, using gross-up factors s_{ij} related to the (perhaps uniform) survey sampling rates yielding the t_{ij}

(33) t_{ij} $= s_{ij} [t_{ij}\sim \bullet *A_i\bar{} \bullet *B_j\bar{}]$

where the starred values may denote trip end totals or shares...Of course, mixing **rsi** or other **ij-indexed** information, such as that obtained from household surveys, with **i-indexed** or **j-indexed** information derived form cordon counts could be seen more formally as sampling on both the « trip ends » t_i or t_j (the total number or trips originating from or destined for a zone) and the the interzonal fows t_{ij} in order to obtain an estimate of the O-D matrix vector t. Such an estimate could, but need not be further adjusted by formal maximization on the A_i and B_j. In his seminal paper, Robillard (1975) noted that Equation (32) reduced the numbers of parameters to be estimated to $2n$, thereby formulating a manageable non-linear regression problem in t_i and t_j vectors whose products yield the desired O-D flows. But Debaille (1977, 1979), working with data from Roanne (France), applied the idea in her NEMROD model to update—or « reconstitute », as she said—the matrix from a previous estimate $t_{ij}\sim$, in effect transforming the problem (32)-(33) into (34) :

(34) $v_a = a_1 b_2 \cdot t_{12}\sim\bullet p_{\cdot(1,2),a} + a_1 b_3 \cdot t_{13}\sim\bullet p_{\cdot(1,3),a} + ... + a_i bj \cdot t_q\sim\bullet p_{\cdot(q),a} + ... + a_Q b_Q \cdot t_Q\sim\bullet p_{\cdot(Q),a} + e_a,$

that is, into a formal minimization with respect to these $2n$ « quasi-behavioural » parameters.

E. The structural behavioural Regression approach **R**

O-D models by mode. As with all good ideas, the oldest stream arose independently in different places, some of which were not generally available or known : for instance Montreal planners developed a full set of urban models on the basis of aggregate behavioural models calibrated on link counts only (Arbour *et al.*, 1969, 1971a, 1971b), as many planners may well have done elsewhere when denied the benefits of large and expensive travel surveys, and Wills (1971) developed an intercity model for the Province of British Columbia in an unpublished M.A. thesis.

But the principal motive prompting the first generally available analyses was the desire to simplify urban transport models. As Low (1972), working with data from West Virginia, said, he wanted to: « effectively combine into one process what is usually handled in a series of three or four sub models, each with its own set of errors ». The same motive is clear in work by Overgaard (1972), working with data from Silkeborg (Denmark), by the Danish Road Directorate in rural and urban areas, including Jentsen and Nielsen (1973), described in Bendtsen (1974), in OECD (1974), or in Holm *et al*. (1976), and by Högberg (1976) in Sweden.

To understand the core idea, consider again Equation (22), i.e. (35), and transform it into (36) :

(35) $v^\circ = \text{þ } t + e$

(36) $\quad v_a = [f(X_{(1,2)\,1},..,X_{(1,2)k}]\, p_{\cdot(1,2),\,a} + [f(X_{(1,3)\,1},..,X_{(1,3)k}]\, p_{\cdot(1,3),\,a} + ..+ [f(X_{(i,j)\,1},..,X_{(i,j)k}]\, p^{\cdot}_{(q),\,a} + ..+ e_a$

where it is clear that the X_k variables are simply the variables one normally finds in structural demand models, such as the travel cost from i to j, activities at i or j, etc., and the functional form questions are the same as they are in regression models. This problem has only k parameters (plus the parameters used to extract information from the regression residuals). It is also obvious that one can formulate the likelihood of observing the link count values as one would any dependent variable in structural models.

Between 1971 and 1976, authors used predetermined functional forms, linear or multiplicative, with simple terms like Population and Employment. Soon after, Gaudry and Lamarre (1978) compared fixed functional forms on an intercity road traffic model for the Province of Quebec and introduced heteroskedasticity corrections. But simultaneously, in an extraordinary piece of work that is still unsurpassed, Wills (1978) produced in his Ph.D. thesis maximum likelihood estimates of O-D matrices from link counts for the car and bus modes, for both British Columbia and for Canada as a whole (plus neighbouring U.S. cities).

For both modes, Wills used fixed-form (multiplicative) models and discussed thoroughly the problem of the assignment of regression constants and spatial correlation among residuals that arise in the context of multivariate structural demand models with flows assigned to the network using a variety of assignment techniques. Moreover, for the bus mode he also estimated a fixed-form intervening opportunities model. For the car mode, he also used a flexible-form model with distinct Box-Cox tranformations on the dependent and 2 independent variables and showed that the resulting estimates were clearly superior to those obtained with the usual multiplicative fixed form. In particular, extensive use was made of Box-Cox and Box-Tukey transformations to determine the optimal forms of the models .

The Box-Cox transformation is defined as :

$$
(37) \qquad \qquad X^{(\lambda)} \;=\; \begin{cases} \dfrac{X^{\lambda}\,-\,1}{\lambda}, & \lambda \neq 0, \\[3em] \ln X, & \lambda = 0. \end{cases}
$$

(38)

and it gives great flexibility as it includes, when used on both dependent and independent variables, many forms used in practice. This breath and depth of modelling was not approached by others in later similar attempts, e.g. Willis and May (1981) and Han *et al.* (1981).

Wills' approach was recently updated for intercity car and truck flows among 208 cities (148 Canadian and 60 bordering American cities) by Transport Canada. In that work, Leore (1996a) first tested the approach with the following simple fixed-form model :

$$
(39) \quad T_{ij} \;=\; \alpha \; (P_i P_j)^{\beta} \; D_{ij}^{\;\delta} \; L_{ij}^{\;\gamma} \; I_{ij}^{\;\theta} \; C_{ij}^{\;\mu},
$$

where $P_i P_j$ is the population product, D_{ij} the time-distance separating the nodes, L_{ij} is a linguistic pairing index defined as the difference in the percentages of each city's population with English as a mother tongue, I_{ij} is the weighted average per capita personal income of the city-pair, and C_{ij} is a network centrality variable expressing the relative location of the nodes in the network. $\alpha, \beta, \delta, \gamma, \theta, \mu$ are parameters to be estimated. The inclusion of the centrality variable takes account of the different propensities to travel for nodes located near the geographic centre of the network (defined roughly by highway 400 through central Ontario) compared with those on the periphery (i.e. a heartland-hinterland distinction).

Then, Leore (1996b) tested a more general notion of travel impedance, that of intervening opportunities, the idea that persons will not travel as much to a city if a larger one comes between them and the destination, using a particular case of Wills' (1986) specification. This involves multiplying (39) by a complex proportionality factor », PGO, that is a function of the cumulative opportunities between an origin and a destination:

$$
(40) \qquad PGO_{i,j(i)} = Z_{i,j(i)} - Z_{i,j-1(i)}
$$
where,

$Z_{i,j(i)}$ = a function of cumulative opportunities from i to $j(i)$;
$Z_{i,j-1(i)}$ = a function of cumulative opportunities from i to $j-1(i)$, the centroid immediately preceding $j(i)$,

and the first term of cumulative opportunities function is:

$$
(41) \qquad Z_{i,j(i)} \;=\; \left(\sum_{k=i+1}^{j} U_{i,k} \;+\; e^{\upsilon} \right)^{(\lambda)}
$$

where,

λ represents a Box-Cox transformation measuring the impact of intervening opportunities on O-D trips, υ represents a Box-Tukey shift parameter, and $U_{i,k}$ represents the opportunity function, summed over all nodes intervening between a given origin and destination, in this case :

$$(42) \quad U_{i,k} = e^{\gamma * \ln(P_k) + \delta * \ln(D_{i,k})}$$

or equivalently:

$$(43) \quad U_{i,k} = P_k^{\gamma} D_{i,k}^{\delta}$$

where P represents the population of the kth node intervening between a given origin and destination node; D represents the distance separating the origin node and the kth node intervening between the origin and destination; and γ, δ are parameters.

Using O-D estimates by mode in total demand and mode choice models. In the second part of his thesis, published as a part of Gaudry and Wills (1977, 1978), Wills specified flexible form models of total intercity passenger flows and mode choice (4 modes). These demand models, produced with the estimated O-D matrices for two of the modes, included in particular Box-Cox logit models that demonstrated decisively both the theoretical and practical superiority of Box-Cox forms over the linear forms. Although the behavioural (structural) models used were aggregate, the same issues arise with disaggregate data. In the case of the logit model, for instance, nonlinearity means that modifications of transportation conditions (say price or time) do not have constant effects on the choice probabilities : in consequence, the values of time derived will depend on trip duration and on how much time is gained, as can be deduced from the Table 5 below where derivatives with respect to the representative utilities clearly make the point.

Table 5. Non constant returns in a Box-Cox logit model

	$\dfrac{\partial U_m}{\partial X_{mk}} = \beta_{mk} X_{mk}^{(\lambda_{mk}-1)}$	$\dfrac{\partial U_m^2}{\partial X_{mk}^2} = \beta_{mk}(\lambda_{mk}-1) X_{mk}^{(\lambda_{mk}-2)}$	Returns
$\lambda = -1$	β_{mk} / X_{mk}^2	$-2\beta_{mk} / X_{mk}^3$	Decreasing
$\lambda = 0$	β_{mk} / X_{mk}	β_{mk} / X_{mk}^2	Decreasing
$\lambda = 1$	β_{mk}	0	Constant
$\lambda = 2$	$\beta_{mk} X_{mk}$	β_{mk}	Increasing

During the period 1977-1982, Transport Canada produced all of its intercity passenger forecasts with these models, that is with O-D matrices for two modes obtained by sophisticated models and total demand and mode choice models of no smaller refinement. It is clearly well within our reach to estimate O-D matrices for certain modes with models of comparable complexity to that of model (39)-(43), partly described above, and to estimate total and mode choice demand models using flexible forms as Wills has done more than 20 years ago. To the extent of our knowledge, no-one since Arbour *et al*. (1969) has estimated a <u>multimodal</u> model on the link counts of many modes simultaneously. This could be done with either aggregate or disaggregate models, or with a combination of these.

4. Congestion complications

It is clear however that congestion is a problem for all approaches, because the path choice probabilities p in Equation (6) then depend on flows, as the flows depend on link costs. The equilibration required for an equilibrium is not trivial because, in Table 3, the demands are defined from i to j and the performance occurs on links a, from which unique ij measures are required to establish a Demand-Network equilibrium.

Often heuristics are used, as pointed out recentlly (Zhu and Hensher, 1995), for instance in work done by the Danish Road Directorate, where convergence conditions are not known. However, Nguyen (1977) has defined equilibrium conditions following Wardrop's formulation (Wardrop, 1952) in which each user minimizes his cost (the assignment is at average (not marginal) cost on each link).

As pointed discussed in Cascetta and Nguyen (1988), the solution to the problem that the asignment map cannot be reduced to a closed form function of the trip vector (i.e. there is no explicit form for the user equilibrium operator M(t)) is to generate an approximate linear map. But unicity of the solution and convergence of the algorithm always pose a problem, as there is no known convergent algorithm.

To see why the problem matters, consider the 6-node (4 of which are simultaneously origin and destination zones), 6-link network of Figure 3. Both matrices A and B are consistent with the observed flow pattern. However, if travel time on link 2→4 is modified from 2 to 3 minutes, the resulting flow pattern will not change if Matrix A is corrrect, but will be profoundly modified if Matrix B is correct—with strong implications for the profitability of links.
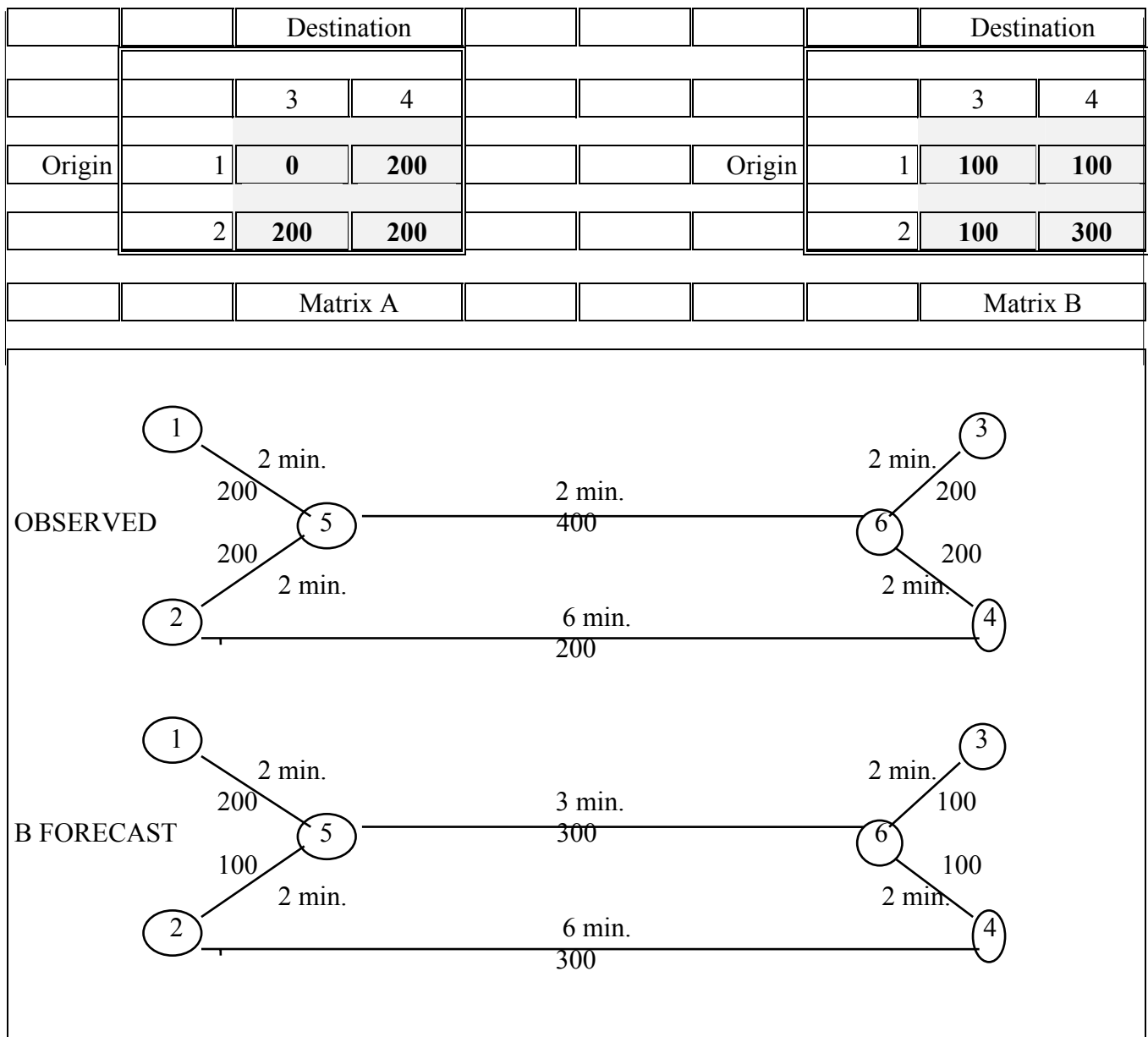
		Destination							Destination	
		3	4						3	4
Origin	1	0	200			Origin	1		100	100
	2	200	200				2		100	300
		Matrix A							Matrix B	

Figure 3. Implications of knowing true matrix if there is congestion

5. Conclusion

We have distinguished between matrix derivation approaches inspired by the accounting systems of inter industry economics and matrix estimation approaches. We have identified « pure » approaches to the problem of O-D matrix estimation, in increasing age order—if sampling is excepted. Clearly some approaches are mixed. For instance, MVESTM (DETR, 1998) appears to combine all 3 approaches : Willumsen and Debaille for the form of the demand model and a regression problem (with 2n + a parameters ?) to compare reference and calculated values not only for «generated» trips and link counts, but also for trip ends ! Clearly, there are many ways to obtain O-D matrices and it is very hard to see how, to obtain cheap, easily updated matrices for even one country—to say nothing of the problem of generating such matrices for the EU—combinations of approaches within a single maximand (not formulated here) would not carry the day at least for the basic flows needed to develop TEN network accounts.

6. References

Abrahamsson, T., « Estimation of Origin-Destination Matrices using Traffic counts—A Literature Survey », IR-98-201, International Institute for Applied Systems Analysis, Laxenburg, May 1998.

Arbour, D., Bégin, P., Gohier, J.-P., Latour, G. et Thivierge, J., « Le modèle de transport », Textes polycopiés présentés au Comité AMUR-ETEN. Service d'Urbanisme, Ville de Montréal, 1969.

Arbour, D., Bégin, P., Gohier, J.-P., Latour, G. et Thivierge, J., « Le modèle de transport : schéma », Service d'Urbanisme, Ville de Montréal, 1971a.

Arbour, D., Bégin, P., Gohier, J.-P., Latour, G. et Thivierge, J., « Le modèle de transport : simulations métro 1981», Service d'Urbanisme, Ville de Montréal, 1971b.

Bell, M.G., « The Estimation of Origin-Destination Matrices by Constrained Generalized Least Squares », Transportation Research B, 25, 1, 13-22, 1991.

Bendtsen, P.H., « Simplification of Trafic Models », Trafic Engineering and Control, 15, 817-819, 1974.

Bigras, Y., Hamelin, P. et Nguyen, S., « Un modèle des flux interrégionaux de marchandises au Canada basé sur le concept d'information », Publication # 232, Centre de recherche sur les transports, Université de Montréal, 1983.

Blankmeyer, E.C., « A Metro Model », Unpublished manucript, Economics Department, Princeton University, 1971.

Cascetta, E., « Estimation of Trip Matrices from Trafic Counts and Survey Data : a Generalized Least-Squares Estimator », Transportation Research B, 18, 4/5, 289-300, 1984.

Cascetta, E. and M. Di Gangi, « A Multi-Regional Input-Output Model with Elastic Trade Coefficients for the Simulation of Freight Transport Demand in Italy », Transport Planning Methods, PTRC Education and Research Services Ltd, Volume P 404-2, 13 p., 1996.

Cascetta, E. and Nguyen, S., « A Unified Framework for Estimating or Updating Origin/Destination Matrices from Traffic Counts », Transportation Research B, 22, 437-455, 1988.

Clavering, J. and H. Hautzinger, H., « « Matrix Merging and Combination—Theory and Methods », Version 3.1, 25th August 1999, Forthcoming in the Final Report of the MYSTIC Consortium, September 1999.

Clavering, J. and H. Kirby, « Data Accuracy as an Aid to Model Development : towards a Manual of Guidance », Proceedings of the European Transport Forum, PTRC Education and Research Services, London, 11 p., 1994.

Dagenais, M., Fisk, C., Gaudry, M. et P. Hamelin, "Analysis of Microtape Data from the 1977 Passenger Travel Survey", Publication CRT-217, Centre de recherche sur les transports, Université de Montréal, 220 p., 1979.

Debaille, S., « Un modèle de reconstitution de matrices origine-destination en milieu urbain », Rapport I.R.T. No 27, Institut de Recherche des Transports, Décembre 1977.

Debaille, S., « Reconstitution de matrices origine-destination en milieu urbain : le modèle NEMROD », Bulletin RECHERCHE-TRANSPORTS No 26, 13-23, Institut de Recherche des Transports, Avril 1979.

DETR. U.K. Department of Transport, Environment and the Regions, « Removing Double-Counting from RSI Databases ». Final report of work carried out by Sandman Consultants Ltd and University of Leeds Institute for Transport Studies, 1997.

DETR. U.K. Department of Transport, Environment and the Regions, « Appendix D. The Statistics Used within ERICA », 13 p., 1998 (?).

DOT. U.K. Department of Transport, « Trip Matrix Building and Validation Program Suite, User Documentation », Highway Engineering Computer Branch, HECB/R/27, London, 1982.

Florian, M., and M. Gaudry, "A Conceptual Framework for the Supply Side in Transportation Systems", Transportation Research B 14, 1-2, 1-8, 1980.

Florian, M., and M. Gaudry, "Transportation Systems Analysis: Illustrations and Extensions of a Conceptual Framework", <u>Transportation Research B</u> 17, 2, 147-154, 1983.

FORMEZ, « Tavole e modelli intersectoriali biregionale e multiregionali », <u>Effeti Economici della spesa pubblica. Strumenti Quantitativi a Supporto delle Decisioni</u>, Progetto per la formazione di analisti finanza pubblica per lo sviluppo, 1995.

Gaudry, M., "The Demand for Public Transit in Montreal and its Implications for Transportation Planning and Cost-Benefit Analysis", Ph.D. Thesis, Economics Department, Princeton University, and University Microfilms, Ann Harbor, Michigan, 188 p., 1973.

Gaudry, M., "A Note on the Economic Interpretation of Delay Functions in Assignment Problems", <u>Lecture Notes in Economics and Mathematical Systems</u>, 118, 368-381, 1976.

Gaudry, M., "Six Notions of Equilibrium and their Implications for Travel Modelling Examined in an Aggregate Direct Demand Framework", in Hensher, D. and P. Stopher, eds, <u>Behavioural Travel Modelling</u>, Croom Helm, Ch. 6, 138-163, 1979.

Gaudry, M. and L. Lamarre, "Estimating Origin-Destination Matrices from Traffic Counts: A Simple Linear Intercity Model for Quebec", <u>The Logistics and Transportation Review</u> 15, 5, 631-642, 1979.

Gaudry, M.J.I. and Wills, M.J. , « Estimating the Functional Form of Travel Demand Models », Publication #63, Centre de Recherche sur les Transports, Université de Montréal, 1977, and <u>Transportation Research</u> 12, 4, 257-289, 1978.

Han, A.F., Dowling, R. G., Sullivan, E. C., and May, A. D., « Deriving Origin-destination Information from Routinely Collected Trafic Counts », Volume II, Trip Table Synthesis for Multipath Networks, Research Report UCB-ITS-RR-81-9, Institute of Transport Studies, University of California, Berkeley, 231 p., 1981.

Hautzinger, H., « Statistical Analysis of Travel Behavior : Some Methodological Considerations », <u>Proceedings of the Seventh International Symposium on Transportation and Traffic Theory</u>, Kyoto, 1977.

Högberg, P., « Estimation of Parameters in Models for Trafic Prediction : a Non-Linear Regression approach », <u>Transportation Research</u> 10, 263-265, 1976.

Holm, J., Jensen, T., Nielsen, S.K., Christensen, A., Johnsen, B., and Ronby, G., « Calibrating Trafic Models on Trafic Census Results Only » », <u>Trafic Engineering and Control</u>, 17, 137-140, 1976.

Hyman, G., « The Conversion of Rail Station-to-Station Matrices using Station Choice Models », DETR. U.K. Department of Transport, Environment and the Regions, 4 pages, March 1998.

Jin, Y. and I. Williams, « The Use of a Generalised Input-Output Structure in Transport Demand Forecasting and Assessment », Paper presented at the 8th World Conference on Transport Research, Antwerp, Belgium, July 1998.

Kirby, H.R., « Combining and Modelling Trip Data », Manuscript, Transport Research Institute, Napier University, Edinburg, 7 p., 1997 (?).

Kurihara, M. and E.C. Sullivan, « Estimating Origin-Destination Matrices from Roadside Interview Surveys », <u>Transportation Research B</u>, 21, 233-248, 1987.

Leore, R., "Estimating O-D Travel from Link Traffic Counts: Can We Really Get There From Here?", <u>Proceedings of the Canadian Transportation Research Forum's 31st Annual Conference</u>, May 26-29, pp. 806-819, 1996a.

Leore, R. « « An Economic Model of Inter-Urban Trafic on the Canadian Highway Network », Report TP 12800E, Economic Analysis, Transport Canada, 35 p.,August 23, 1996b.

Low, D., « A New Approach to Transportation System Modelling ». <u>Traffic Quarterly</u> 26, 391-404, 1972.

Moses, L., « The Stability of Interregional Trading Patterns and Input-Output Analysis », <u>American Economic Review</u>, 1955.

MYSTIC, A technical Annex, Peter Davidson Consultancy, 1997.

Nguyen, S., « Estimating an O-D Matrix from Network Data : A Network Equilibrium Approach », Publication # 39, Centre de recherche sur les transports, Université de Montréal, 1977.

Nguyen, S., Modèles de distribution spatiale tenant compte des itinéraires », INFOR 21, 4, 270-292, 1983.

OCDE Road Research Group, Urban Trafic Models : Possibilities for Simplification, OCDE, Paris, 1974.

Picard, G. and M. Gaudry, "A Box-Cox Logit Model of Intercity Freight Mode Choice", Publication CRT-898, Centre de recherche sur les transports, Université de Montréal, 17p., 1993. Forthcoming in Transportation Research E, 1997.

Picard, G., Nguyen, S. et M. Gaudry, "Fret: un modèle de simulation des flux de marchandises au Canada", Les Cahiers Scientifiques du Transport 17-18, 183-200, 1988.

Picard, G., Paskievici, D. et Nguyen, S. « Une version désagrégée du modèle de simulation de flux de marchandises TOMM », Publication # 377, Centre de recherche sur les transports, Université de Montréal, 63 p., 1985.

Robillard, P., « Estimating the O-D Matrix from Observed Link Volumes », Publication # 1, Centre de recherche sur les transports, Université de Montréal, 1974, and Transportation Research 9, 123-128, 1975.

Robillard, P. and Trahan, M. « Estimating the O-D Matrix and Network Characteristics from Observed Link Volumes ». Publication # 128, Département d'informatique, Université de Montréal. Proceedings of the International Conference on Transportaion Research. First Conference, 736-740, 1973.

Sandman Consultants, « MYSTIC-Merging Origin-Destination Data », Discussion Note, 6 p., 10 February 1998.

Spiess, H., « A Maximum Likelihood Model for Estimating Origin-Destination Matrices », Transportation research B, 21, 5 , 395-412, 1987.

Wardrop, J.G. « Some Theoretical Aspects of Road Traffic Research », Proceedings of the Institute of Civil Engineers, Part II, 1, 325-378, London, 1952.

Willis, A. E. and A.D. May, « Deriving Origin-destination Information from Routinely Collected Trafic Counts », Volume I, Trip Table Synthesis for Single Path Networks, Research Report UCB-ITS-RR-81-8, Institute of Transport Studies, University of California, Berkeley, 248 p., 1981.

Wills, M.J., « Development of the Highway Network, Trafic Flow and the Growth of Settlements in Interior British Columbia », M.A. Thesis, Department of Geography, The University of British Columbia, Vancouver, 1971.

Wills, M.J., « Linear and Nonlinear Estiomators of the O_D Matrix », Ph. D. Thesis, Department of Geography, The University of British Columbia, Vancouver, 342 p., 1978.

Wills, M.J. , « A Flexible Gravity-Opportunities Model for Trip Distribution », Transportation Research ,Vol. 20B, No. 2, pp. 89-111, 1986.

Wilson, A.G., Entropy in Urban and Regional Planning, Pion, London, 1970.

Zhu, W. and Hensher, D.A., « Estimating an Origin-Destination Trip Matrix from Link Trafic Counts for Large Networks », Working Paper ITS-WP-95-3, Institute of Transport Studies, The Univerity of Sydney, 22 p., February 1995.