Demand Modelling Developments and Competition among Airports

by

Prof. Marc GAUDRY

Département économie et sociologie des transports (DEST)
Institut national de recherche sur les transports et leur sécurité (INRETS), Arcueil, France
and
Agora Jules Dupuit (AJD) www.ajd.umontreal.ca
Université de Montréal, Montréal, Canada

marc.gaudry@inrets.fr ***and*** marc.gaudry@umontreal.ca

Table of contents

The author thanks Professor Jaap de Wit for his comments on a previous version of this paper invited for presentation at the Workshop on *Network strategies of multi-hub airlines and the implications for national aviation policies*, The Netherlands Directorate-General for Civil Aviation and Freight Transport, The Hague, October 28, 2005. This version is prepared for the German Aviation Research Society (GARS) Workshop on Benchmarking of Airports and EU-Liberalization, Hamburg, February 23, 2006.

Version 1, October 24, 2005; revised February 20, 2006.

1. Introduction: models chasing reality

Thought market failure. When Alfred Kahn successfully proposed air deregulation in 1978, thereby ensuring that he would be the last chairman of the United States Civil Aeronautics Board (C.A.B.) abolished in 1984, nobody ─not even Professor Kahn himself─, foresaw the future shape of the air route structure that would progressively ensue, either nationally in the United States or internationally (Kahn, 2003).

This massive thought failure in forecasting performance by academic, policy and business experts occurred despite the existence of the successful *Federal Express* freight hub, initiated in 1973 by Fred Smith in Memphis, Tennessee, and clearly profitable since 1976. Some form of the same failure of vision occurred everywhere: even in perfectly hexagonal hub-shaped France, the beginning of the structuring of the Air France hub at Charles-de-Gaulle airport awaited 1996, a process not completed until perhaps 2003.

Of course, this industry is full of similar non-market forecasting failures : despite the exactly linear growth of Southwest Airlines since 1971 ─it has now been retiring pilots for about 5 years─, who forecasted the rise of the new stable low-cost business model firms challenging incumbent full-service airlines, at least on short high-density markets, and benefitting in the short run from the accumulated excess supply of subsidized airports?

Hub humility : making sense *ex post*. In retrospect, one can make some sense of route restructuring[1], even if the process has yet to stabilize at the international level due to the continuing protective straightjacket of bilateral agreements: we all know the arguments about the high frequencies and leg density economies induced by changes from gridding to hubbing (see Annex 1), and we are all aware that this supply side revolution may have meant, at least for a small minority of U.S. passengers (Morrison and Winston, 1999 or 2002), more circuitous routings fortunately involving almost no interline connections, replaced by on-line connections. And we all wonder how long small European capital cities can keep daily direct non-stop air service to North America in a liberalized environment.

It also makes sense to many of us that, in the next phase of evolution, rising income trends might then slowly increase the value of time and consequently the relative demand for direct passenger flights, if not for freight flights. We shall initiate the discussion of competition among airports by emphasizing their hub character and neglecting (except implicitly in the definitions of land-based access variables) the competitive advantages that may arise from improvements in the complementary[2] links of land modes.

2. A framework: flow, demand and airport system components

Our discussion needs a framework to clarify flow, demand and airport system components.

2.1. Indentifying hub flows of interest within accounting and equation specifications

Key modelling developments and hub stability. If market reality speaks, can an academic researcher say anything to increase the clarity of the message? I will subjectively select a few key new features of models and ask whether they are more likely to imply hub stability or not. In a formal sense, this question raises a problem of air company/air path choice but I will treat it without being very specific about issues such as detailed scheduling and flight coordination (Burghouwt and Wit, 2003), but within a representative schematic demand generating framework.

Two kinds of flows at a hub. We define our interest as that of the flow of air passengers using a hub location (airport) assimilated to a city. They are a number composed of two kinds of flows.

---

[1] Michael Tretheway has also argued many times since 2003 at Hamburg Airport Conferences that this movement was helped by incorrect accounting of multi-leg revenues by airlines, including some double counting of revenues.

[2] For a discussion of complementary and substitution bewteen air and land modes, notably high speed rail, see Gaudry (1998).

For a certain trip purpose *g*, that number equals the sum of the origin-destination (O-D) passengers, for whom the hub *h* is the origin or the destination city of the city pair considered, and of the transfer passengers, passing through *h* on their path between any O-D city pair :

$$[\text{Air flow through } h] = [\text{Direct O-D traffic between } h \text{ and any } j] + [\text{Circuitous traffic through } h], \qquad (1)$$

or

$$T^{g}_{\text{air, h}} = T^{g}_{\text{air, hj}} + T^{g}_{\text{air, ihj}}\ , \qquad i, j \neq h, \qquad (2)$$

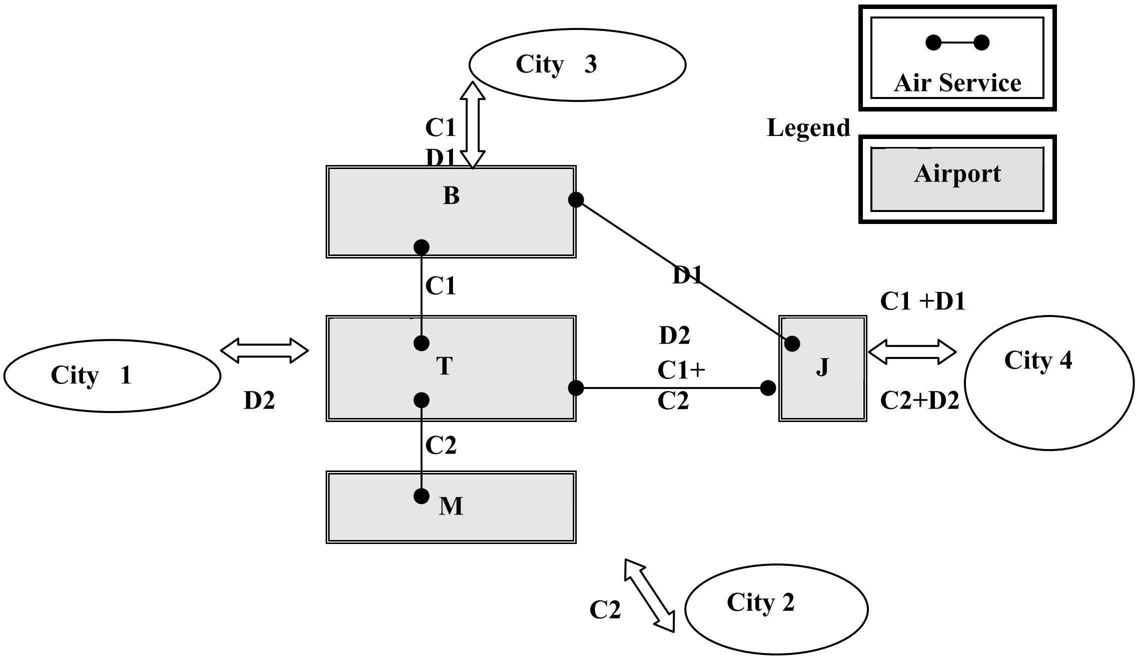
where for simplicity all flows are assumed to be bi-directional between city *h* and any city *j* as well as between any city pair *ij* whose air passengers choose to go through hub city *h.* Also, for similar reasons and without loss of generality, we associate the airport cities to the origin or destination zones. The two components of hub flows defined by Equation (2) amount to much more than transfer flows.

Flows on interest in a Trafic Accounting Matrix (TAM). To make our intents clear, Table 1 defines an accounting system to visualize the hub flows of interest in (2) before we use equations to explain them.

Table 1.A. Traffic Accounting Matrix for City 4

Flow Matrix I		**Destination airport**					**Destination city**					**Total**
	O\D	**B**	**T**	**M**	**J**	***NL***	**1**	**2**	**3**	**4**	**FL**	**TL**
Origin airport	**B**								D1+C1			
	T	C1		C2			D2					
	M							C2				
	J	D1	C1+C2 +D2									
	NE											
Origin city	**1**											
	2											
	3											
	4				D1+D2 C1+C2		D2	C2	D1+C1			
	FE											
Total	**TE**											

Table 1.B. Diagram of physical network flows originating in City 4



Part A of Table 1 is a Traffic Accounting Matrix (TAM) reporting on four flows (D1, D2, C1 and C2) that ***entrer Airport J of City 4*** and follow eitheir a direct or a circuitous path to three other cities, as indicated in part B of the same table. In this TAM, known O-D flows correspond to the bottom right hand side quadrant of Part A and passengers are registered when they enter airport J, when they use a flight, and when they leave an airport for the outside. One can apply such a TAM to a complete airport system such as that reported on in Table 2 for the same four cities of Table 1 but without the direct flight from Airport J to Airport B.

Table 2. Flow matrix for the 4-city system without direct connections from Airport J to Airport B

Flow Matrix II		**Destination airport**					**Destination city**					**Total**
	O\D	**B**	**T**	**M**	**J**	***NL***	**1**	**2**	**3**	**4**	**FL**	**TL**
Origin airport	**B**		150 **500** **75**			*725*			25 **200** **25**		250	**975**
	T	25 **200** **25**		**400** **500** 50	**75** **300** 100	*1675*	250 150 100				500	**2175**
	M		**300** **200** 100			*600*		**400** **500** 50			950	**1550**
	J		**25** **400** 250			*675*				**75** **300** 100	475	**1150**
	NE	*250*	*2000*	*950*	*475*	T_i T_J	**500**	**950**	**250**	**475**	2175	
Origin city	**1**		100 25 50			175		50	25	100		
	2			100 **200** **300**		600	100		**200**	**300**		
	3	150 **500** **75**				725	150	**500**		**75**		
	4				**25** **40** 250	675	250	**400**	**25**			
	FE	725	175	600	675	2175						
Total	**TE**	**975**	**2175**	**1550**	**1150**							

Note first that the system O-D matrix is asymmetric. Then, summing up registrations yields **FE**, the vector of flows entering an airport from any outside zone, and **NE**, the vector of flows entering an airport from any other airport. One can also compute **NL** and **FL**, the corresponding vectors of flows leaving an airport for another airport or for outside zones, respectively. Combining these, the total entering flow vector is surely given by **TE** = **FE** + **NE** and the total leaving flow vector by :

$$\mathbf{TL} = \mathbf{NL} + \mathbf{FL}. \qquad (3)$$

As passengers do not accumulate in the airports, the accounting identity of the airport system states that the flow coming through an airport from the outside or from other airports must be equal to the flow leaving that airport for other airports or for the outside:

$$\mathbf{TE'} = \mathbf{TL} \qquad (4)$$

where the (**'**) denotes the transpose operator. We emphasize that the accounting identity of the system is ***not*** that the entering and leaving vectors **FE** and **FL** are the same: in the Table 2 example, the entering and leaving vectors differ due to the asymmetry of the O-D matrix. But, of course, if **e'** denotes the transpose of the unit vector, summing the flows entering all airports from the outside and the flows leaving all airports for the outside must yield the same total number of air trips for the purpose of interest g , as expressed in:

$$\mathbf{e'FE'} = \mathbf{e'FL} = 2175 = T^{g}_{air}\,. \tag{5}$$

In such an accounting system, expounded and developed further in Gaudry (1973), it is possible to pursue the obvious analogies with interindustry input-output tables made up of N and L submatrices (extended with the lower-level ***entering*** and ***O-D*** matrices in Table 3) and show for instance the general lack of proportionality between forecasts of the leaving demand vector **FL*** and forecasts of network leg flows among airports N*. This is done by defining first the familiar matrices of direct (and, later, of indirect) input-output coefficients of interindustry economics now available in two potential formats instead of one, as can be demonstrated with the help of Table 3 where a TAM is written in simplified format:

Table 3. Structure of a Traffic Accounting Matrix

N	L
E	O-D

Thereare two posible matrices of technical coefficients because, after an explicit distinction is made between entering and leaving demand structures E and L, there are two ways to define the operations required to compute direct technological coefficients. Network flows N can be combined either with the leaving flows or with the entering flows; this possibility is absent from input-output tables (Leontief, 1941) that effectively contain only N and L matrices[3] and thereby impose a unique definition of the matrix of technological coefficients A, namely, in the familiar way:

$$A = N \bullet B^{-1}, \text{ or } N = A \bullet B, \tag{6}$$

where B is a diagonal matrix with elements made up of components of the leaving demand vector **TL**:

$$B \equiv \begin{bmatrix} \mathbf{TL_1} & & & \\ & \mathbf{TL_2} & & \\ & & \mathbf{TL_3} & \\ & & & \mathbf{TL_4} \end{bmatrix}$$

so that :

$$\mathbf{NL} = A \bullet \mathbf{TL}\,. \tag{7}$$

and the usual operations can be effected, first by substituting (7) into (3), and so on[4].

Independently from its extended input-output interpretation, the TAM found in Table 2 can be used to focus visually on the different objects explained by particular transport models. For instance, the Trip Generation step of the four-step transport planning format involves explaining the vectors T_i and T_j, noted there in dark.

In this analysis, we first match cities to airports, which imposes a particular structure of the network entering and leaving vectors **NE** and **NL** because flows are not allowed to enter or leave an airport in relation to more than a single zone. We then isolate the hub values of interest by underlining them or putting them in bold in

---

[3] Quesnay (1759) had a single square (4x4) matrix where column and row totals « balanced » after some hard work, but did not distinguish between intermediate and final flows, as did Leontief by specifying N and L parts, keeping the N matrix square for inversion purposes. After Blankmeyer (1971), we simply interpret L as reporting on flows « towards the outside », which naturally calls for the « from the outside » counterpart matrix E and for their O-D combination in the remaining quadrant of Table 3. This spatial extension differs from the standard spatialization of Leontief systems expounded by Moses, (1955) and in use in transport (e.g. Cascetta and Di Gangi, 1996), where a diagonal spatial trade matrix T is used to multiply a redefined diagonalized matrix A and suitably redefined output vectors.

[4] From **TL** = A•**TL** + **FL**, one collects terms and writes **TL** = $(I-A)^{-1}$ **FL** thereby deriving C = $(I-A)^{-1}$, the famous matrix of indirect coefficients. Forecasting N* involves the reverse path from **FL*** to **TL*** and then to N* from N* = A•B* in (6).

Table 2. The underlined entries circumscribe the O-D passengers flows and isolate them from the **transfer passengers** flows. These are the flows to be understood, that is explained.

Flows of interest generated within an equation framework. We explain them by associating analytical expressions to the two kinds of flows using hub *h* and consequently explicate (2) as follows:

$$T^{g}_{air,h} = \left\{ Pop_h^{*} \bullet \left\{ \Sigma_j \, [Pop_j^{*}]^{\alpha} \bullet [U_{ij}]^{\beta} \bullet [U_{air,ij} / \Sigma_{m,ij} \, U_{m,ij}] \right\} \right\}, \qquad j=1,\ldots,Z, \qquad (8)$$

$$+ \Sigma_i \Sigma_j \left\{ [Pop_i^{*} \bullet Pop_j^{*}]^{\alpha} \bullet [U_{ij}]^{\beta} \bullet [U_{air,ij} / \Sigma_{m,ij} \, U_{m,ij}] \right\} \bullet \left[U_{air,ihj} / \Sigma_{air,ikj} \, U_{air,ikj} \right], \qquad i,j \neq h, \qquad (9)$$

where (8) refers to O-D passengers and (9) to transfer passengers, α and β are parameters, there are modes m= 1,..., M, and zones i, h, j = 1,..., Z, and :

Pop_j^{*} : income and activity weighted measures of population at location *j*; (10)

$[U_{ij}]$: the aggregate utility of the modes from *i* to *j*; (11)

$[U_{air,ij} / \Sigma_{m,ij} \, U_{m,ij}]$: the *air* modal market share from *i* to *j*; (12)

$\left[U_{air,ihj} / \Sigma_{air,ikj} \, U_{air,ikj} \right]$: the share of air trips from *i* to *j* going through hub *h*. (13)

Note that a possible definition of a successful hub is a location where the ratio of (9) to (8) is high.

2.2. Airport flows within a system: demand, performance and supply components

These demand equations contain many variables that are determined by the « supply side » of the air transport system. This notion requires some development because we believe that it is insufficient without the further introduction of the notion of performance. Why and how?

The 19th Century: from one to two levels. Until the end of the 18th Century, explanations tended to relate variables of interest, say imports or exports flows, 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 and expected signs.

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. Fur us, 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 » (D-C-S) to distinguish it from « Demand-Supply » (D-S) 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), within the system definition indicated in Table 3 to which we now turn.

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

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

SO : supplyer objectives
RE : regulatory environment
ST : supplyers’ estimate of the state of the system

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

Most recent version: Gaudry (1999).

More conceptual details on a three-layer system. Naturally, using a 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 3 where, without loss of generality, the Supply dimension side was then illustrated by being related simultaneously to road infrastructure and public transport road vehicle services in a potentially congested network.

In this representation of Table 3, 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***.

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 concluded 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 eeonomies", (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 is important for our airport system problem to our classification to distinguish among the three layers: Demand, Performance Supply, jointly determine airport passenger and freight transfer and delay times that themselves become objects of analysis along with "demand" and "supply" quantities.

3. New modelling developments and the sensitivity of hub flows

3.1. Trends in the analysis an extended gravity-type structure

Gravity and the IIA axiom. The gravity structure of both components (8) and (9) is arguably the most successful modelling structure in all of economics, compatible with many theories and still formidably robust empirically in both mathematical form and broad specification.

First used in transportation by Lill (1891) on Vienna-Brünn-Prague passenger rail flow data, it is written here in so-called « Generation-Distribution » form, without the double constraints that garantee that all of what comes out of *i* or comes into *j* is assigned. We exclude such constraints because they bias the estimation of parameters by imposing a very special restriction on error terms and thereby artificially induce a form of substitution among all O-D pairs that is absent from the core model structure.

Indeed, this basic structure of the Gravity form, written in this simplified form is :

$$(\text{Flow indicator}_{\mathbf{ij}}) \leftarrow (\text{Activities}_{\mathbf{i\ and\ j}}\ ;\ \text{Socioeconomic}_{\mathbf{i\ and\ j,}}\ ;\ \text{Ease of interaction}_{\mathbf{ij}}), \qquad (14)$$

where the ease or (dis-)utility of interaction is an impedance metric that can be of many types (generalized cost, geographical, preference or taste, cultural, legal or regulatory, *etc*.). This specification (14) makes abundantly clear the fact that, because the flow from *i* to *j* depends only on variables that have *i* and *j* indices, the characteristics of other paths or locations have no influence on the flow from *i* to *j*. Consequently, the ratio of the demand for two O-D pairs does not depend on the characteristics of other links or locations than those involved in the ratio, a property consistent with Luce's Independence from irrelevant Alternatives (IIA) axiom of choice theory.

Recent attempts have been made by many authors to bring some measure of competition or complementarity from other O-D pairs into more or less complex variants of (14). They met with little success despite extensive work, for instance to demonstrate the presence of such forces with spatial autocorrelation techniques simultaneously with tests of the multiplicative form of (14) on Canadian and German multimodal intercity passenger data (Gaudry *et al.*, 2006).

Gravity and trip purpose. The form (8)-(9) is better at handling trip purposes other than tourism, the fastest growing of the passenger air markets and in many air markets the dominant trip purpose, but recent sophisticated attempts to explain tourism flows have been successful (Last, 1998). However, the structure used by that author is still IIA consistent.

Further attempts to improve it by introducing competition among O-D pairs demonstrated, with an excellent NUTS-3 Europe-wide data set based on French and British data, that such competition existed, but only within very limited circles (about 300 km in radius) around origin and destination zones (Gaudry *et al.*, 1998). No other attempts have yet been made to model for instance the competition among different tourism zones, such as the Mediterranean, the Atlantic coast of Europe and the Baltic sea.

Influence on competitive hub stability. It is not clear whether the demonstration of the existence of competition among O-D pairs would favour passenger transfers at hubs : we conjecture that the demonstration of greater competition among O-D pairs may reduce the density economies associated with hubbing by spreading demand. A demonstration of the presence of complementarity among large subsets of cities would have the opposite effect.

We have to remember that the hubs have two markets : the O-D direct market and the circuitous transfer market. Most of the successful hubs have a large « home base » of O-D demand (the first term in (8)). Effects on economies of density happen with both types of airport clients.

3.2. Trends in the analysis of individual terms of the extended gravity form

Activity terms. It is extremely difficult to develop international Generation-Distribution models beyond variables other than population, perhaps corrected by income and some measure of employment structure. We do not see any new modelling trends in this area, except the innovation introduced by Last (1998) who successfully weighted the trips by the propensity to travel by age class. But, as a form of aggregation of market segments, this has no definable effect on the structure of the hubs markets.

Total Utility term U. All models have to find a way of aggregating the mode and path opportunities offerred between any two points : much progress has occurred in this area due to the use of the denominator of the mode choice model (typically logit) as indicator, according to the U_{ij} format used in (8) and (9).

What is less known is that a simple mesure of distance will work as well, and yield almost identical elasticities (in absolute value), as the denominator of a sophisticated model : Last (1998) has demonstrated so much with three trip purposes (business, private and vacation). The reason is of course that the spatial structure of modal origin-destination generalized costs and distance are extremely correlated. Consequently, increased sophistication in the description of alternatives will have no impact on the comparative advantages of differrent hubs.

The classical Linear Logit mode choice model. The work horse of mode choice modelling is the classical Linear Logit model that explains the probability of choice (or the market share) of a mode *m* as:

$$p\left(m\right)=\frac{\exp(V_m)}{\sum\limits_{j\in C}\exp(V_j)} \tag{15}$$

with all utility functions defined as

$$V_i = \beta_{i0} + \sum_n \beta_{in}^i X_n^i + \sum_s \beta_{is} X_s + u_i \tag{16}$$

where we have changed slightly the usual notation to identify the X_n^i (the upper index denotes the mode and the lower ones the utility function index and the network variable), the network characteristics that belong to a particular mode and therefore vary across alternatives, and it is clear that in (16) the X_s denote socio-economic characteristics of consumers that are common across alternatives.

The year 1977 was a turning point for this model as three new streams of work simultaneously extended it: (i) hierarchies we developed (e.g. Williams, 1977); (ii) what became known later as the « mixed » Logit arose through the treatment of the coefficients in (16) as random (Johnson, 1977, 1978); (iii) the Standard Box-Cox Logit was defined by applying Box-Cox transformations to the explanatory variables (Gaudry and Wills, 1978).

As generally applied, the hierarchies of nested linear logits do not change the characteristics of the model for our purposes, essentially because they are still IIA consistent across main branches. Similarly, randomization poses considerable problems of credibility because the distributions of coefficients of the model are unknown and because its information matrix[5] does not have a closed form and hence its efficiency bound is not defined (Cirillo, 2005). It is of course a formal way of dealing with a type of market segmention, long practiced in combinatorial fashion by energetic analysts.

More importantly, Orro *et al.* (2005) have shown in effect that the recent popularity of mixed logit may well be due to the fact that the true relationships are not linear and should have their curvature estimated rather than postulated. Their tests of curvature involved the Box-Cox (1964) transformation, defined as :

---

[5] This point was recently brought to our attention by Lasse FridstrØm.

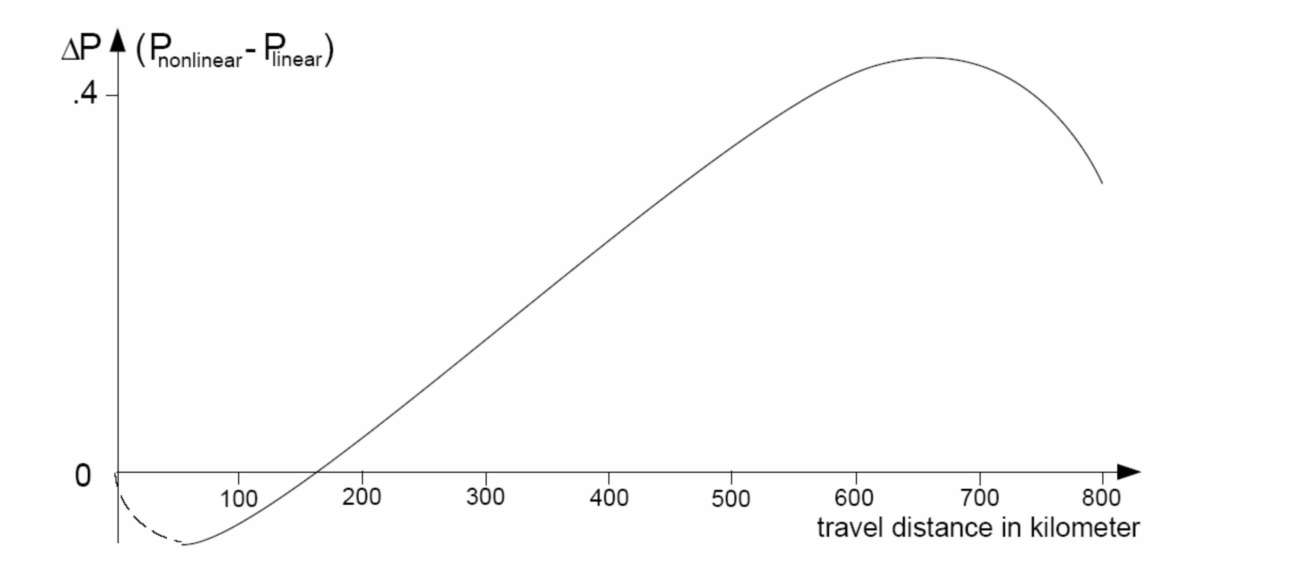
$$X_k^{(\lambda_x)} = \begin{cases} \dfrac{X_k^{\lambda_k} - 1}{\lambda_k} & si \quad \lambda_k \neq 0, \\ \ln X_k & si \quad \lambda_k \rightarrow 0. \end{cases} \tag{17}$$

which implies so called Standard Box-Cox Logit utility functions written as:

$$V_i = \beta_{i0} + \sum_n \beta_{in}^i X_n^{i(\lambda_{in}^i)} + \sum_s \beta_{is} X_s^{(\lambda_{is})} \tag{18}$$

Impact of non linearity on forecasts. These functions have considerable implications for the behaviour of market shares because the marginal impact of price and service characteristics are not constant any more. For instance, Mandel *et al.* (1997) show with a discrete choice data set that forecasts of the impact of high speed rail ICE service on the market share of the train in Germany differ considerably between the linear assumption (incorrect) and the (optimal) non linear one (the power parameter equal to 0,25 is applied to the price and time dimensions of (18)), as the following Figure 1 drawn from that paper makes clear:

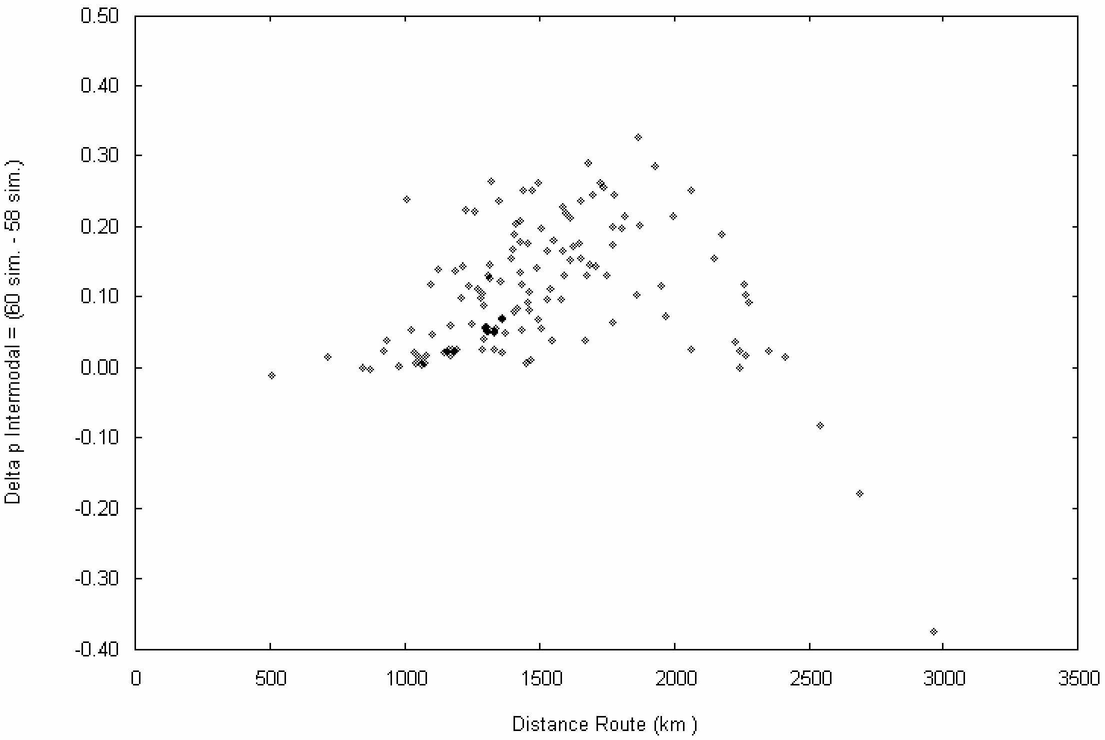
Figure 1. Comparing Linear and Box-Cox Logit forecasts of an ICE train scenario in Germany



A number of key European models now use Box-Cox transformations with functions of type (18). It is also true in freight that the non linear forms yield higher market shares after improving an alternative than the linear form: assuming a 10% decrese in the price of intermodal transport across the Pyrénées (other land modes are the classical rail and truck), Figure 2 shows the difference, documented in Gaudry *et al.* (2005), between the Box-Cox Logit forecast (model 60) and the linear forecast (model 58). We note the same thing as in the passenger model above: intermediate flows increase more with non-linear models than with linear models.

If that holds generally, use of Linear Logit models, for instance in Europe-wide passenger models, as with the VACLAV model (Schoch, 2000) used in TEN-STAC (2004), or with a current freight demand models in Germany (Seltz, 2004), could make a difference.

Figure 2. Comparing Linear and Box-Cox Logit forecasts of an intermodal train gains



A practical form of the Universal Logit. But non linearity holds further promises, because the transformation also makes it possible to extend considerably the specification found in (18) by allowing for the inclusion in the utility functions of the charateristics of other alternatives, in the so called Generalized Box-Cox Logit form:

$$V_i = \beta_{i0} + \sum_n \beta^i_{in} X_n^{i(\lambda^i_{in})} + \sum_n \beta^j_{in} X_n^{j(\lambda^j_{in})} + \sum_s \beta_{is} X_s^{(\lambda_{is})} \tag{19}$$

which provides perhaps the first workable form of McFadden's (1975) Universal Logit. To the extent that these ***other*** characteristics are successfully introduced, the Logit becomes a complete system of demand equations. By allowing for a complex pattern of substitution and complementarity, it probably makes the market shares more stable, although this has yet to be tested on a case by case basis.

Figures 3 compares, for the trans-pyrenean model referred to above, forecasts made with a Generalized Box-Cox Logit form that includes a set of cross prices[6] in addition to own price (model 61), as in (19), to forecasts made with the linear form (model 58) of the same model. The pattern of the same as that found in the previous graph.

Non linearity may therefore make a difference to our problem both because it implies different market share in middle-range markets and because it allows explcitly for both substitution and complementarity in the Logit model: we are back in classical demand systems where all prices matter and utility is not presumed to be separable and additive.

Path and company choice. Generally, path choice models are of the Linear Logit kind, with the same properties as in mode choice applications except one: that of the constants, to be discussed shortly. Until as recently as 2004, Lufthansa was using a Linear Logit procedure for their path choice problems. As the number of paths used by air passengers tend to increase with distance (from say 3 within Germany up to 16 Germany and the United States), the modelling procedure is of some import.

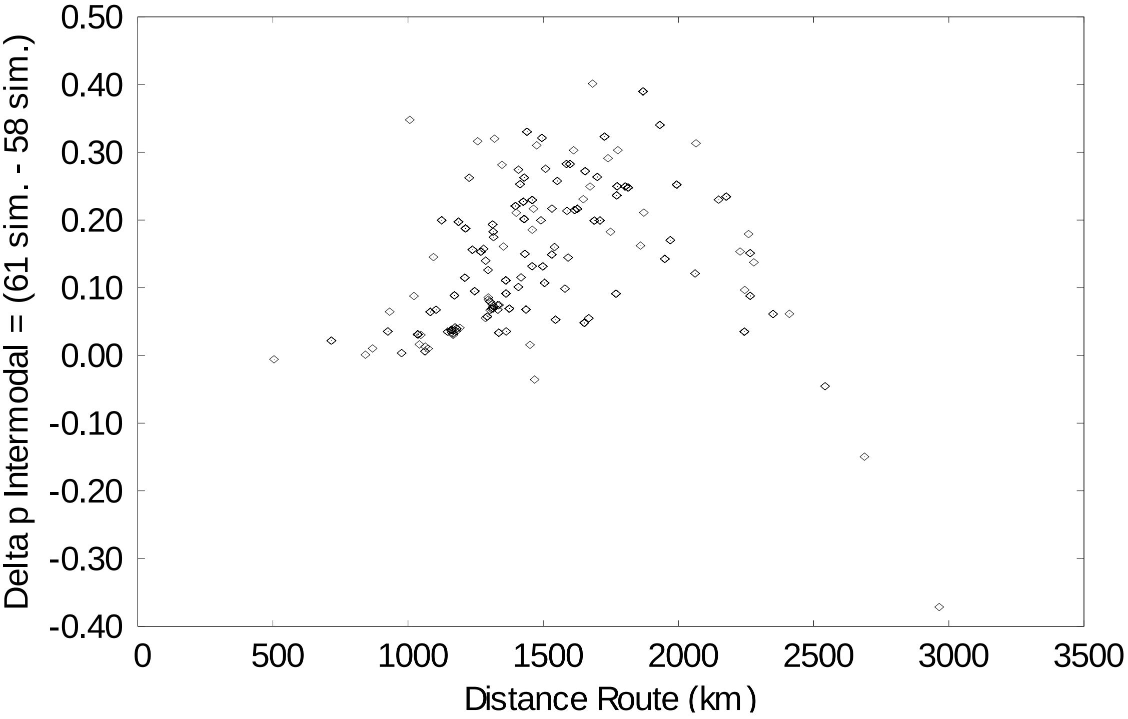
---

[6] This is the model in use since 2005 by the French ministry of transport to forecast freight land mode market shares across the Pyrénées.

The emergence of the non linear Box-Cox Logit model implies that some paths will incorrectly dominate other paths under a linear form, but these effects have yet to be studied in comparative fashion as they are expected to mimick those found for the mode choice models noted above.

Use of non linearity of the Box-Cox type should therefore have a strong impact on ***marketing studies*** and Stated Preference surveys : it implies that experimental data should contain sets of variables that are orthogonal not in linear but in non linear space : the correlation that matters for the reproduction of reality and the establishment of statistical causality is defined among the optimal nonlinear forms of the variables of interest, a point particularly well made in Cirillo (2005).

Figure 3. Comparing Generalized Box-Cox and Linear Logit forecasts of intermodal train gains



In the analysis of air carrier and path choice, there are few exceptions to the use of a Linear Logit and such exceptions as exist, such as the PODS procedure derived from an earlier Boeing (Boeing Co, 1996) procedure, works very much like a Linear Logit (Carrier, 2003), no doubt partly because its utility functions are linear. More importantly, it attempts to take into account capacity constraints on links, which generates some complementarity among them, in the spirit of MAPUM (Soumis *et al.*, 1979). This should favour high frequency hubs and stabilise path assignments by complementary block groupings (such as Toronto-Paris, Montreal-Paris and Montreal-Toronto-Paris) despite the simplistic linearity assumptions.

In this context, a problem with the Logit is that its mode-specific constants are underidentified so that one ends up estimating differences in their coefficients with respect to an arbitrarily chosen reference one, as in:

$$\left.\begin{array}{l} V_i = (\beta_{i0} - \beta_{r0}) + \sum\limits_n \beta^i_{in} X^i_n + \sum\limits_s (\beta_{is} - \beta_{rs}) X_s \\ V_i = \qquad \beta^{\nabla}_{i0} \qquad + \sum\limits_n \beta^i_{in} X^i_n + \sum\limits_s \qquad \beta^{\nabla}_{is} \qquad X_s \end{array}\right\} \tag{20}$$

where the problem is illustrated for these constants and for the socio-economic variables common to all functions. This problem is minor in mode choice problems if no new mode is to be introduced, but in company/path choice procedures, it matters whether the path constants are equal because the game is precisely to introduce new paths all of the time.

The reason why this is a hard problem in air networks is that there is no natural labelling of paths serving an O-D pair, as there is in mode choice analysis, because it hardly makes sense to define a reference path as one might define a reference mode. One way out of this quandary is to use an Inverse Power Transformation (IPT) envelope (Gaudry, 1981) applied to a Logit (IPT-L) path choice model, as was done by Laferrière (1987) to identify all his path-specific air constants for a Canada-wide model built with a 100% sample (16 million individual trips) of domestic air trips made on Air Canada and Canadian Pacific Airlines in 1983.

Further, if one accepts an ordering of air paths, some other gains can be expected from the IPT-L form, such as the identification of captivity to certain paths (or airlines and airports) and the detection of left over non linearity or asymmetry in the response to such ordered path utilities.

This new orientation is therefore also promising for the proper modelling of air company/path competition.

4. Conclusion

We have collected the main points in Table 4. Overall, the identification of path-specific constants in air networks and the introduction of nonlinearity seem promising to learn more about what current air flow patterns must be saying. If the basic structure of models is IIA consistent, how can can competition among tourism destinations be modelled ?. Competition among airports has to be more than a matter of improvements in access modes to the different airports.

Table 4. Impacts of modelling developments on hub competition and stability

Likely impact on transfer passenger hub demand :		***Favourable***	***Unfavourable***
General Gravity Structure Features			
i)	Substitution among O-D pairs		***probably***
ii)	Complementarity among O-D pairs	***probably***	
The Role of Components			
iii)	Activity terms	no influence	no influence
iv)	Aggregate utility term U	no influence	no influence
v)	Mode choice		
	non linearity	**probably important**	
	segmentation and linear mixed logit	no influence	no influence
vi)	Path choice		
	PODS-type assignment	yes	
	the identification of constants in a Logit	**important**	

5. References

Blanchard, O., J. and L. H. Summers, "Hysteresis and the European Unemployment Problem", in S. Fischer (ed), NBER Macroeconomics Annual, Vol 1, 409 p, M.I.T. Press, 1986.

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

Boeing Co., The Decision Window Path Preference Model : Summary Discussion, Marketing and Business Strategy, Boeing Commercial Airplanes, The Boeing Company, 1996.

Box, G.P. and D.R. Cox, « An Analysis of Transformations », Journal of the Royal Statistical Society, Series B, 26, 211-243, 1964.

Burghouwt , G. and J. de Wit, "The Temporal Configuration of European Airline Networks", 33 p., Publication AJD-74, Agora Jules Dupuit, Université de Montréal, Juillet 2003. http://planning.geog.uu.nl/medewerkers/g_burghouwt.htm; www.aaeconomics.com.

Carrier, E., Modeling Airline Passenger Choice : Passenger Preference for Schedule in the Passenger Origin-Destination Simulator (PODS), M.A, Thesis, Civil and Environmental Engineering, M.I.T., 136 p., June 2003.

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.

Cirillo, C. «Optimal Design for Mixed Logit Models », Groupe de recherche sur les transports, Facultés Universitaires Notre-Dame de la Paix, Namur, 11 p., Association for European Transport and contributors, 2005.

Cournot, A.-A., Recherches sur les Principes Mathématiques de la Théorie des Richesses, 1838.

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", Transportation Research B 17, 2, 147-154, 1983.

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, 188 p., 1973.

Gaudry, M., "A Note on the Economic Interpretation of Delay Functions in Assignment Problems", Lecture Notes in Economics and Mathematical Systems, 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, Behavioural Travel Modelling, Croom Helm, Ch. 6, 138-163, 1979.

Gaudry, M., "A Study of Aggregate Bi-Modal Urban Travel Supply, Demand, and Network Behavior Using Simultaneous Equations with Autoregressive Residuals", Transportation Research B 14, 1-2, 29-58, 1980.

Gaudry, M. "The Inverse Power Transformation Logit and Dogit Mode Choice Models", Transportation Research B 15, 2, 97-103, 1981.

Gaudry, M, « Key Substitution-Complementarity Features of Travel Demand Models, with Reference to Studies of High Speed Rail Interactions with Air Services », in Calzada, M. et M. Houée, éd. Actes du Colloque Déplacements à Longue Distance. Mesures et Analyses, Ministère de l'Equipement, du Logement et des Transports, et ESA Consultants, Paris, 412-457 1998 ; and in COST 318, Interactions between High-Speed Rail and Air Passenger Transport, Final Report of the Action, European Commission, Directorate General Transport, Brussels, 204-238, www.cordis.lu/cost-transport/src/cost-318.htm, 1998.

Gaudry, M., «The Four Approaches to Origin-Destination Matrix Estimation : Some Considerations for the MYSTIC Research Consortium », Working Paper N° 2000-10, Bureau d'économie théorique et appliquée (BETA), Université Louis Pasteur, 23 p., Septembre 1999.

Gaudry, M., Briand, A, Paulmyer, I. et C.-L. Tran, "Choix modal transpyrénéen ferroviaire, intermodal et routier: un modèle Logit Universel de forme Box-Cox Généralisée », 22 pages, Avril 2005.

Gaudry, M., Heinitz, F., Last, J. and B. Mandel, « Methodological Developments within the Quasi-Direct Format Demand Structure : the Multicountry Application for Passengers MAP-1 », Publication CRT-98-46, Centre de recherche sur les transports, Université de Montréal, Working Paper N° 9815, Bureau d'économie théorique et appliquée (BETA), Université Louis Pasteur, 67 p., octobre 1998.

Gaudry, M., and J.S. Kowalski, "Demand-Performance-Supply Equilibria in Centrally Planned Economies". Communist Economies 2, 2, 205-221, 1990.

Gaudry, M., Mandel, B. and W. Rothengatter, « Introducing Spatial Competition through an Autoregressive Contiguous Distributed (AR-C-D) Process in Intercity Generation-Distribution Models for Canada and Germany", Ch. 10 in Gaudry, M. and G. Tegnér, eds, Transport Demand and Mode Choice: The Impact of Functional Form, Elsevier Science, Oxford, forthcoming 2006.

Gaudry, M. and M. Wills, "Testing the Dogit Model with Aggregate Time Series and Cross Sectional Travel Data", Transportation Research B 13, 2, 155-166, 1978.

Groenewegen, P., “Supply and Demand”, The New Palgrave, A Dictionary of Economics 4, 553-556, 1987.

Johnson, L. W., Stochastic parameter estimation: an annotated bibliography. International Statistical Review 45, 257-272, 1977.

Johnson, L.W., Regression with Random Coefficients. Omega 6, 71-81, 1978.

Kahn, A., Lessons from Deregulation, AEI-Brookings Joint Center for Regulatory Studies, 88 p., 2003.

Kornai, J., Growth, Shortage and Efficiency, University of California Press, Berkeley and Los Angeles, 1982.

Kornai, J. and J. W. Weibull, “The Normal State of the Market in a Shortage Economy: a Queue Model”, Discussion paper, The Institute for International Economic Studies, Stockholm University, 1977.

Laferrière, R., Une Aggrégation Nouvelle des Itinéraires du Transport Aérien (ANITA), 119 p., Publication AJD-87, Agora Jules Dupuit, Université de Montréal, décembre 1987.

Last, J., “Reisezweckspezifische Modellierung von Verkehrsverflechtungen”, Diplomarbeit, 132 p., Universität Karlsruhe, Februar 1998.

Leontief, W., The Structure of the American Economy, 1919-1929, Cambridge, Mass., 1941.

Lill, E., Das Reisegesetz und zeine Anwendung auf den Eisenbahnverkehr, Vienna, 1891.

Mandel, B., Gaudry, M. and W. Rothengatter, A Disaggregate Box-Cox Logit Mode Choice Model of Intercity Passenger Travel in Germany and its Implications for High Speed Rail Demand Forecasts. The Annals of Regional Science 31, 2, 99-120, 1997.

McFadden, D., “On independence, structure, and simultaneity in transportation demand analysis”, Working paper 7511, Institute of Transportation and Traffic Engineering, University of California, Berkeley, CA, 1975.

Morrison, S.A. and C. Winston, A Profile of the Airline Industry. Ch. 1, p. 1-26, in Gaudry, M. and R. R. Mayes, eds, Taking Stock of Air Liberalization, Kluwer Academic Press, 304 p., 1999. www.wkap.nl/prod/b/0-7923-8387-7.

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

Morrison, S.A. et C. Winston, Le profil du secteur du transport aérien. Ch. 1, p. 3-40, in Gaudry, M. et R. R. Mayes, dir., La libéralisation du transport aérien : Bilan et perspectives, Les Presses de l’Institut du Transport Aérien, Paris, 319 p., 2002. http://www.ita-paris.com/.

Orro, A., Novales, M. and F.G. Benitez, «Nonlinearity and Taste Heterogeneity Influence on Discrete Choice Model Forecasts», 18 p., Association for European Transport and contributors, 2005.

Parker, R. and J, Walker, « Estimating the Utility of Time-of-day Demand for Airline Schedules Using a Mixed Logit Model, 19 p. , July 2005

Podkaminer, L., “Macroeconomic Disequilibria in Centrally Planned Economies: Identifiability of Econometric Models Based on Theory of Households’ Behavior Under Quantity Constraints”, Zaklad Badań Statystyczno-Ekonomicznych Zeszyt 30, Warsaw 1989.

Portes, R., Quandt, R., Winter, D. and S. Yeo, « Macroeconomic Planning and Disequelibrium Estimates for Poland, 1955-1980» , Econometrica, 55, 1, 19-41, 1987.

Prossaloglou, K. E. and F.S. Koppelman, The Choice of Carrier, Flight and Fare Class, Journal of Air Transport Management 5, 4, 1999.

Quesnay, F., Tableau Économique, Versailles, 3ème édition, 1759.

Ricardo, D., The Principles of Political Economy and Taxation, London, 1817.

Schoch, M., « VACLAV-VIA, The IWW-MKmetric Passenger Model », Institut für Wirtschaftspolitik und Wirtschaftsforschung, Universität Karlsruhe (TH), 20 p., September 2000.

Seltz, T., « The BVU Nested Box-Cox Logit Freight Mode Choice Model Developed for the German Federal Transport Plan in 2002 », Presentation at the MODEV-M Workshop, Fondation Reille, Paris, 22nd June 2004.

Soumis, F., Ferland, J. A. and J. M. Rousseau, « MAPUM, A Model for Assigning Passengers to a Flight Schedule », Publication CRT-142, Centre de recherche sur les transports, Université de Montréal, août 1979.

Steuart-Denham, J., An Inquiry into the Priniples of Political Economy, 1767.0

Stigler, G. J. and C. Friedland, Chronicles of Economics, The University of Chicago Press, 1989.

TEN-STAC: **T**rans-**E**uropean **N**etwork — **S**cenarios, **T**rafic Forecasts and **A**nalysis of **C**orridors, European Commission Contract N° : ETU/B5-7000A-SI2.346797, www.nea.nl/ten-stac , 2004.

Williams, H.C.W.L., « On the Formation of Travel Demand Models and Economic Evaluation Measures of User Benefit ». Environment and Planning, 9a(3), 285-344, 1977.

Working, E. J., What do statistical “demand curves” show ? Quarterly Journal of Economics, 41, 212-35, 1927.

6. Annex 1. The simple cost and frequency analytics of hubbing

What happens to route length and flight frequency when one changes from bilateral to hub structures ? Imagine a country A with 5 cities, located at its center and corners, and a second country B of hexagonal form with 7 cities similarly located, as in Figure AB.

Figure AB. Two spatial configurations

In each case denote by **c** the length of the link from a corner city to the central city. Now compare the length of a network of direct non-stop bilateral links among all cities to a hub-shaped network centered on the central city. For country A, the hub-based network is 3,41 times shorter than the other and, for country B, it is 4,73 times shorter[7].

Service frequency will increase correspondingly when, using the same planes, the grid structure is replaced by the hub structure. To see this, first note that, in each case the number of bilateral links among **n** cities is $\mathbf{n(n-1)/2}$. In each case the length of the hub-shaped network is $\mathbf{(n-1)c}$ and the corresponding length of the grid-haped network is equal to $\mathbf{(n-1)c((n-1)/2 + ((n-1)/2)^{1/2})}$.

The obviously positive difference between the two configurations is $\mathbf{(n-1)c\ ((n-1)/2 + ((n-1)/2)^{1/2}\ -1)}$ for each city. To further compare actual gains between cities, one would have to further specify that length **c** is equal for both cities, but this has not been assumed in this case.

---

[7] Le nombre de liens bilatéraux entre **n** villes est $\mathbf{n(n-1)/2}$. Dans les deux cas présentés, la longueur du réseau en étoile est, pour chaque pays, égal à $\mathbf{(n-1)c}$ et le réseau de liens bilatéraux est de longueur $\mathbf{(n-1)c((n-1)/2 + ((n-1)/2)^{1/2})}$. Alors la différence entre les deux configurations est $\mathbf{(n-1)c\ ((n-1)/2 + ((n-1)/2)^{1/2}\ -1)}$. On pourrait s’arranger pour que la longeur **c** soit égale dans les deux cas, mais cela n’est pas vrai dans les figures présentées.