Transport Economic Theory

Transport Economic Theory

Sergio Jara-Díaz
Universidad de Chile
Santiago, Chile

Emerald

United Kingdom – North America – Japan
India – Malaysia – China

Emerald Group Publishing Limited
Howard House, Wagon Lane, Bingley BD16 1WA, UK

Reprints and permission service
Contact: booksandseries@emeraldinsight.com

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

ISBN: 978-0-08-054899-9

To my family
To my students

About the Author

Sergio Jara-Díaz is Professor at the Transport Systems Division of the University of Chile in Santiago, a research unit he helped to found in the early seventies that became a forerunner in establishing transport as a distinctive field of knowledge in that country. He obtained his Ph.D. from the Massachusetts Institute of Technology (MIT) in 1981 with a thesis in the field of Transport Economics. He has published more than seventy research articles in books and journals focusing on users' behaviour and benefits, value of time, scale and scope economies in transport network industries, public transport and pricing. Professor Jara-Díaz has designed and taught courses in Transport Economics both at the University of Chile and MIT, and teaches regularly in many universities in Spain. He has been a member of the Highest Commission of Academic Evaluation at the University of Chile; is a member of the American Economic Association, the Association for European Transport, and a founding member of the Chilean Society of Transport Engineering. He has been running a weekly radio show since 1991 and has published a book of chronicles and short stories.

Prologue

This book is an attempt at settling a debt I have with my students - past and present - and with myself. I have been doing research in transport economics since the seventies and I delivered my first formal basic course in the area in 1982, whose contents have evolved slowly but continuously without changing the structure. From the very beginning I was guided by the conviction that this was not a matter of teaching microeconomics with an application to transport; introducing time and space was unavoidable and that made a significant difference. This posed a challenge that translated into systematic work devoted to construct the basic aspects that I thought constituted the foundations of a body of knowledge that could be properly called Transport Economic Theory.

I remember vividly the day when, twenty eight years ago as a student at MIT, the late Ann Friedlander gave me a pile of papers dealing with the theory of multiple products, asking me to read them and write a short essay on how I would apply this to transport. Matching what I read with my view of transport as a "flows production" activity was instantaneous. In essence, my answer was that this new theory was really the only way to properly understand transport production, as any transport firm produces a vector of origin-destination flows and, therefore, concepts like economies of scope would help the analysis of network growth and shape in transport industries. My dear Professor, whose recent contribution had been the hedonic treatment of transport output, told me that although she could not see exactly what I meant, I had a Ph.D. thesis there. The chapter on transport production is the long run result of a research line that begun that day. I still have that handwritten essay.

Many individuals have contributed to my enjoyment with Transport Economics. Tristán Gálvez introduced me to the world of flows in the early seventies when I was a student at the University of Chile. By the end of the decade I was exposed to the microeconomics of discrete choices in a course taught by Dan McFadden at MIT; I loved the simplicity of the idea that allowed the introduction of quality with only a slight extension of consumer's behavior theory. Later on this was the pillar on which I was able to develop a better understanding of the role of personal income and propose new measures of users' benefits. My work on the microeconomics of demand received the permanent support of my friend and colleague Juan de Dios Ortúzar, from the Catholic University of Chile, whose faith applying my theories has played an important role in this story. Similar appreciation has always been explicitly shown by David Hensher from the University of Sydney in his many books and papers, by Huw Williams from Cardiff University, and by Ken Small from the University of California, Irvine.

Many dear colleagues around the world, perhaps too many to mention, have contributed in different ways with their time and care to keep academic interaction alive and pleasurable. I want to explicitly acknowledge Hani Mahmassani, formerly at U. of Texas at Austin and now at the U. of Maryland, Yossi Sheffi and Nigel Wilson from MIT, Peter Mackie from the U. of Leeds, Pablo Coto-Millán from the U. of Cantabria, Eduardo Martínez-Budría from U. of La Laguna and José Holguín-Veras from RPI. My colleagues at the Transport Division of the University of Chile were always supportive of the work behind the book, and I do thank Francisco Martínez, Marcela Munizaga, Cristián Cortés and Leonardo Basso for their daily life patience.

The final stages of the preparation of this book received the very important editorial help of former student Rodrigo Quijada with the assistance of present student Alejandro Tirachini. Rodrigo offered his help and took his duties as if this book was his; from discussions with him the structure, wording and general presentation of the chapters improved indeed.

Neither research along these years nor this book have stolen time from my family, as nothing could have deprived me from the pleasures of making my children sleep, of playing, singing, studying and talking with them along their lives. This, of course, was induced by my own education with my parents Sergio and Elena. My sons Pedro (Santiago, 1977) and Francisco (Boston, 1980), and these many years with the woman I love, Momy, are the meaning of life to me. They have always been first priority.

Sergio Jara-Díaz
Santiago, January 2007.

Presentation

The study of Transport is the study of movement, of displacements of individuals and things in both space and time. But unlike the displacement of, for example, water particles in a piping system, the distinctive feature of Transport as an area of knowledge is the involvement of human *will* in the process. On one hand, our "particles" have will; each one needs to go from some point to some other point in space-time, and this makes them no-interchangeable. On the other hand, some elements of our "piping system" have will as well; they decide how many particles can go together, how frequently they can travel, or how fast they move. Those that are responsible for the particles' will, be it passengers or freight, are the **users** of the transportation system, and those responsible for the moving elements of the piping system are the **operators**.

Will means tastes, preferences, decision-making, objective pursuing, perceptions, rationality; all together. It means *behavior*. So in principle, if we are interested in analyzing the behavior of either transport users or operators, we can rely upon fairly well established theories of economic behavior. On the users' side, individual demand for trips could be studied with the concepts of consumer theory, and freight demand could be looked at as part of firms' decisions, either to bring inputs to the plant or to distribute output to markets, such that transport is an input to the complete production process. On the other side, the behavior of transport operators can be in most cases understood using the tools of the theory of the firm, although in this case the product is the vector of displacements itself. But despite being useful as forms to approach the problem, basic production and consumption theories do not explicitly account for the key dimensions in transportation processes: **time and space**. In fact, introducing them is nearly what it is all about to formulate a transport economic theory. And this is exactly the intention behind this book.

Let us briefly point out some of the key issues that **have** to be included in a framework to understand and analyze transportation activities. First, for the operator to be able to produce movements, what is being moved has to be physically present; this is common to nearly all services and makes an obvious difference with respect to the production of goods in general, which does not require the consumer' presence. This has been sometimes referred to as the "non-storability" property of transportation processes. Second, this same fact can be looked at from the users' viewpoint, which means that own time is needed to actually realize the product; again, this is common to all kinds of consumption, but here it is usually the single most important dimension of the "product". Thirdly, as each complete displacement in space-time is a different product, transportation firms usually generate a vector of products as opposed to a single (scalar) output; thus, the supply side of transportation services should be looked at as multioutput processes. Although it is true that most productive activities involve more than one product, here this characteristic is unavoidable when going into minimally relevant economic analysis. Movements in different directions are different products, and simple descriptions like ton-kilometers only hide what is part of the essence of transportatio n analysis: origins, destinations, networks.

Therefore, many characteristics make the economic analysis of transport systems operation very different from a straight application of the concepts contained in microeconomic textbooks. As explained, most (if not all) of these characteristics are not exclusive to the analysis of movements, but their simultaneous presence (plus others not mentioned here) make transport economics a somewhat specialized field as opposed to an area of application. And this is so accounting only for technical-analytical considerations, not yet including the social and/or political dimensions usually involved.

The book is organized around four central topics: the transport firm itself; individual's decision-making regarding travel; accounting users' benefits from changes in the transport market; and optimal pricing, encompassing all the previous elements; a chapter has been devoted to each one of them. In Chapter 1, Transport Production, the basic theory of the firm is extended to account for time and space explicitly, showing the multiproduct nature of the transport firm. It is discussed how operator's decisions about route structure, frequency, points served and so on, plus their relation with network characteristics, determine input requirements to produce a desired output; simple cyclical systems help illustrating the concepts. Next, a basic supply-oriented microeconomic tool is introduced: the cost function. Its evident multioutput nature is used to understand rigorously the meaning of scale, scope and complementarity in transportation operations. Building on this, traditional analysis like determining marginal costs or degree of scale economies –which are typically used to study the convenience of expanding production- are looked at here paying attention to the spatial dimension of transport production, unveiling along the way some mistakes that are still common in the analysis of transport cost structures and industry organization.

In Chapter 2, Travel Demand and Value of Time, individuals' behavior is explored starting from the traditional utility-maximization approach, but considering discrete-choice formulations that are able to handle the type of decisions people have to make when facing alternative modes of transport. Needless to say, the amount of time needed to be spent by using an alternative appears as a fundamental variable, just as important –or even more- than price is in traditional consumer modeling. All services require consumers to put in their time in order to provide the service in question, but in the case of transport time is undoubtedly at the center of the issue. Not only that, unlike time spent in a restaurant or at the movies, time spent in mandatory transport is normally undesired and even unpleasant, so people would gladly pay for reducing it as much as possible. As a consequence, this willingness-to-pay-to-reduce-travel-time, or more broadly, individuals' valuation of their own time, is a key element in the transport field. However and despite its obvious importance, it is surprising to realize how slow the progress has been regarding our understanding of the way people manage their time in general. The role and evolution of time as a variable in consumer theory is discussed in this Chapter, which finishes with the presentation of a more complete model encompassing leisure, work, travel, and their values.

Valuation of Users' Benefits in Transport Systems constitutes Chapter 3, and there, as the name suggests, different measures to estimate users' benefits arising from changes in the transport system –like those a new transport project would make- are presented and discussed. The simple and most widely used measures are shown and their limitations exposed, followed by less popular but rigorous measures derived directly, as opposed to the former, from the general microeconomic theory. The chapter also covers two other related topics. First, an analysis is made about the link between benefits measures directly in the transport market and those one could measure in the markets that generated the transport demand in the first place. This means examining, for example, benefits from commercial activities produced by people who traveled for shopping. Second, the issue of aggregating benefits from different users for matters of determining benefits at a society level is discussed. This means, stated in more practical terms, analyzing the difference between transport projects that are financed by users directly and those undertaken with tax money.

In the final chapter, Optimal Transport Pricing, a look is taken at different pricing strategies and their consequences. It starts presenting the desired case where price reflects economic efficiency, which helps illustrating that neither private transport nor public transport markets would produce such case if left alone, due to the presence of externalities. The necessary correction -from a regulatory point of view- that needs to be implemented for each case in order to reach efficiency is also presented. Next, alternative forms of pricing aiming at different goals other than economic efficiency are analyzed.

Transport industries are certainly among the most interesting as a subject of study. The inherent spatial characteristic of this activity and its repercussions for firm cost structure analysis, the presence of externalities of significant magnitude, the fact that users do not want this product by itself but only as a mean to solve some other need, the unavoidable restriction faced by all individuals of having to organize their activities, travel among them, within a 24-hour time frame, and the rather public nature of this market, which forces national and regional authorities to design, evaluate and finance transport projects with social welfare in mind, make this field of knowledge a very special and challenging one. This book aims at providing a theoretical ground for its understanding.

1. Transport production and cost structure

1.1. Introduction

As transport activities mean movement of individuals and goods in both time and space, the analysis of transport production involves the assignment of resources to generate trips between several different points in space during various periods. As a consequence, the microeconomic analysis of transport production is far from a simple extension of the traditional theory of the firm. In this chapter we present the underpinnings of a microeconomic theory of the transport firm, with particular emphasis on the nature of the technical relations between inputs and outputs (production or transformation function) and the use of the cost function as a tool to obtain valuable information for the design of transport policies as pricing and regulation.

The chapter begins with the notion of transport production, including the definition of transport output, the role of space, the idea of operating rules, and the concept of scale, all of which are illustrated using simple cyclical systems. Then the cost function and its properties regarding the calculation of marginal costs, economies of scale and economies of scope, are presented and explained within the context of transport systems analysis. A synthesis of the empirical work using transport cost functions is then offered, with special emphasis on the adequate treatment of output in its specification, and on the difficulties with the prevailing approach to analyze industry structure. Improved procedures to calculate scale and spatial scope economies correctly when output aggregates are used are included, plus a discussion about the analysis of the industry structure considering such measures.

1.2. Transport production

1.2.1. Product and technology

Basically, the production of goods and services can be synthetically described using the concepts of inputs, outputs and technology. Inputs have to be acquired by the firm in order to be combined - within the boundaries of process-specific rules – so as to produce outputs. For a given level of outputs, the firm has to choose type and amount of inputs, as well as a subset of combination rules. Technology defines all feasible input combinations. Formally, let X be the inputs vector (quantity/time unit) and Y the outputs vector. Then,

Definition 1.1: Technology
The technology T is defined by all $(X, Y) \in T \Leftrightarrow Y$ can be produced from X.

A transport process is the immediate effect of the action of transporting, i.e., moving some physical entity from a certain origin in space-time to a certain destination in space-time. We can associate this concept with that of "product" in an economic sense, with some reservations. To describe a product we refer to its qualitative characteristics, assigning a name for simplicity (e.g., oranges, shoes, etc.). To measure a product we need a physical unit of reference, and a quantity in terms of these physical units (e.g., 5 tons of oranges, or

1000 pairs of shoes). When we talk about a production process we need flow units, as opposed to stock units (e. g., 1000 pairs of shoes per week).

But to measure a transport process we would need: a qualitative description of what is being transported, a physical unit of reference, quantity (flow) in terms of these units, and origin and destination in space-time. **The need to explicitly establish origin and destination in space-time is the characteristic that distinguishes more clearly a transport product from the traditional concept**. The transport firm has to use vehicles, terminals, rights-of-way, energy, labor, and so on, to produce movements - freight and/or passengers - from several origins to several destinations during different periods. Thus, **the output of a transport process is a *vector***,

$$Y = \left\{ y_{ij}^{kt} \right\} \in R^{KxNxT} \tag{1.1}$$

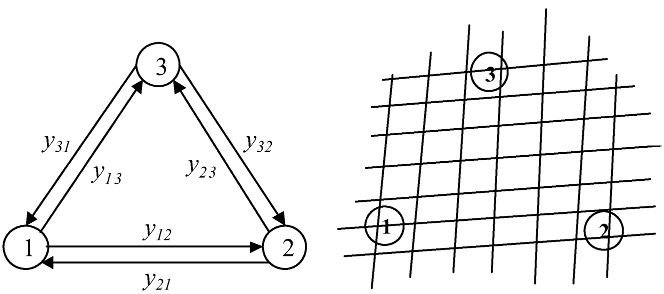
where each component y_{ij}^{kt} represents the flow of type k moved from origin i to destination j (O-D pair ij) within period t, for example passengers from Paris to Frankfurt during a specific weekend (K, N and T are the number of flow types, the number of O-D pairs, and the number of time periods, respectively).

Expression (1.1) is fairly general, including the possibility of a transport firm dealing with several flow types (persons, goods of different types), but it is quite important to stress that **a transport firm produces multiple products mainly because of the presence of time and *space* (periods and origin-destination pairs), not by the handling of multiple flow types**. The word "product" is used in this chapter to indicate a given flow type in a given period between a given O-D pair. Even if a firm only offers passenger service, it still will be offering multiple products. Moving passengers from New York to Buenos Aires for Christmas will usually involve different inputs combinations than doing it from Tokyo to Moscow in June; they are indeed two different products. Unlike the classical theory of the firm, a transport company participates in several markets simultaneously, each with its own demand curve and its own marginal costs, although the latter are usually interrelated. And note the spatial dimension –much more importantly than time- is the key aspect distinguishing the transport industry from other economic activities.

Now, for a given set of flows in Y, the firm has to make several choices: number and capacity of vehicles (fleet size), design of the rights-of-way (location, flow capacity), design of terminals (location, loading-unloading capacity), vehicle frequencies, and so on. Some of these decisions involve choosing the characteristics of inputs, and some are related with their use, i.e. with the form in which inputs are combined to accommodate the flow vector. We will call these latter types of choices **operating rules**. Because transport production takes place on a network, a transport firm has to decide, as well, **a service structure** –i.e. the generic way in which vehicles will visit the nodes to produce the flows– and a **link sequence**. These two endogenous decisions define a **route structure**, which has to be chosen using exogenous spatial information, namely the **O-D structure** of demand (defined by the vector Y), the **location of the nodes** and the **physical network**. Note that the need to make a decision on a route structure is, in the end, a consequence of the spatial dimension of product.

For a given type of transport firm (for example interurban buses) some of the decisions related with the acquisition of inputs are constrained, because of the existence of common infrastructure (for example the road system) or the rigidity of input markets (for example fleet size). On the other hand, decisions on operating rules are generally made within the boundaries of existing inputs.

Example 1.1
Consider an O-D system with three nodes, a single period and a single flow type, as in Figure 1.1.a, located on a physical network, as in Figure 1.1.b. For a given set of flows $\{y_{ij}\}$, the appropriate combination of inputs and operating rules would depend on many factors. Three possible service structures are shown in Figure 1.2 (Jara-Díaz, 2000). Structure (a) corresponds to a general cyclical system (Gálvez, 1978), structure (b) corresponds to three simple cyclical systems (direct service) and structure (c), where a distribution node is created, is known as *hub-and-spoke* and is very common in air transport (note that *hub* H may or may not coincide with an origin or destination node). Regarding vehicle assignment to fleets, which is part of the service structure, there is no choice but one fleet (one frequency) in case (a), three fleets in case (b) and one, two (with three alternatives) or three fleets in case (c). If a cyclical counter-clockwise system like the one in Figure 1.2.a was chosen, a possible route structure could be the one shown in Figure 1.3.



(a) **(b)**

Figure 1.1. ***O-D structure and physical network***