European Transport \ Trasporti Europei

European Transport \ Trasporti Europei
International Journal of Transport Economics, Engineering and Law
Rivista Internazionale di Economia, Ingegneria e Diritto dei Trasporti

ISSN 1825-3997

YEAR XII, NUMBER 36, August 2007

Publisher/Direzione, redazione e amministrazione

I.S.T.I.E.E.
Istituto per lo Studio dei Trasporti nell'Integrazione Economica Europea

Via Lazzaretto Vecchio, 13
34123 TRIESTE
Tel. +39.040.311464
Fax. +39.040.311465
Mail: istiee@units.it

www.istiee.org/te

Print/Stampa
ARTIGRAFICHERIVA SRL TRIESTE

Autorizzazione del Tribunale di Trieste, n. 915 del 31 ottobre 1995

Spedizione in A.O. 70% DC TS
Tassa Pagata/Taxe Perçue

Index

YEAR XII, NUMBER 36, August 2007

Special issue: Pricing, Financing and Investment in Transport

Editors: De Palma, A., Marcucci, E., Niskanen, E., Verhoef, E.

INTRODUCTION

Background

This special issue of European Transport, edited by André de Palma, Edoardo Marcucci, Esko Niskanen and Erik Verhoef, introduces a selection of the papers presented at the 3rd International Kuhmo Conference and Nectar Cluster 2 meeting on "Pricing, Financing, and Investment in Transport", which was held 11-14 July 2006, in Tuusula, Finland. Judging by the quality of papers, and the comments received, the conference was a great success. It attracted over 50 participants, primarily from various European countries and North America. It was preceded by a three day Summer School on "The Economics of Transportation" featuring lectures on a variety of transport economic topics, given by Richard Arnott, Jan K. Brueckner, André de Palma, Amihai Glazer, and Erik Verhoef.

In offering an informal and unique environment for pleasant get-together and fruitful discussions and learning experience, the series of the Kuhmo-Nectar Conference and Summer School have become highly appreciated by all the participants coming from different backgrounds (academia, consultancy, government, independent agencies, international organizations). Encouraged by the positive comments and suggestions for future events, a next edition of this combined Conference and Summer School is scheduled to be held at the University of Urbino, Italy, in July 2007. We can only hope that the initiative will fulfil our expectations in further growing into a vivid and structural yearly highlight on our academic calendars.

From its beginning, this series of conference and summer school have aimed to highlight pricing, financing and investment issues of practical importance in transport, but also more systematically (the summer school) to cover alternative approaches to analyzing these issues and their theoretical backgrounds with the intent of: increasing knowledge and understanding of transport economics, and improve on the quality of research in transport economics; enhancing international cooperation and mutual understanding by attracting participants from different countries, at different levels of economic development, and from international organizations; improving the understanding between economists and other researchers working on transport, land use, regional science, political science, and institutional analysis; facilitating the exchange of views, ideas and experiences between theorists, applied researchers, and practitioners, and help them learn from each other in order to narrow the gap between theory and practice.

This special issue

As said, this special issue presents some of the papers presented at the conference. We will now briefly introduce these papers.

European competitiveness in the global economy is strictly and increasingly linked with an efficient and cost effective transport and port system. The issue of marginal social cost pricing in European seaports, dealt with by Abbes in the first paper, constitutes a key issue not only for European transport policy but also for its future growth opportunities. EU ports are becoming *de facto* multi-product industries offering a range of services and operating under different environments and organizational structures. Many port infrastructures and services are owned/managed by private operators. Notwithstanding the lack of standardization and homogeneity in port characteristics, the European Commission is keen to adopt a common approach to pricing in ports testified by a European Commission White Paper (1998) suggesting the application of marginal social cost pricing, taking into account externalities such as cost of accidents and environmental and congestion costs. Abbes critically discusses the recent changes in the port industry and investigates port pricing in theory, and the possible practical application of marginal social cost pricing in European seaports.

In the second paper, Beria proposes some reflections on transport mega projects in Italy, where he provides a comparative analysis of economic feasibility studies. The author surveys fifteen selected infrastructural transport projects in Italy and provides a critical evaluation of the environmental impact assessments and of the feasibility studies, with the purpose of deriving some general conclusions about the quality, transparency and contents of the assessment, given the Italian normative framework. The evidence from the projects considered suggests that little attention is paid to demand forecast modelling, nor to the use of standardised comparison schemes. Furthermore, some cases of double counting are encountered, as well as some theoretical errors in cost evaluation.

De Palma, Motamedi, Picard, and Waddell contribute the third paper, dealing with the empirical examination of the market for local amenities in the Paris metropolitan region. One important result derived in the paper relates to the considerable inequity in the spatial distribution of these local amenities, which include accessibility, environmental and social indicators. The results are obtained by employing a spatial analysis, using Lorenz curves to study the degree of inequity in the distribution of amenities. The largest divergence is found for noise (due to its concentration near airports), ‘Significant Urban Zones” (areas with high concentrations of social and economic difficulties targeted for government assistance), the presence of water (lakes and rivers) and forests, and the presence of train and subway stations. Recognising that local amenities should be capitalized into the housing market, the authors study the willingness to pay of households for these amenities within the Paris region using alternative specifications of a location choice model. Some important results on the spatial scale of the amenity effects, and how this is captured in a location choice context, are acquired through the estimation of models both at a commune and at a grid cell level. This approach allows the authors to gain new insights into how households in the Paris region trade off amenities against each other, and against housing cost. An important point is that the residential location choice model fits the data moderately better at the smaller scale of the grid cell, compared to the commune.

Dumas discusses, in the fourth paper, the capitalization of land and commercial rents generated by railway infrastructures, as a source of finance for passenger railway companies. He describes the global situation of these activities in the different parts of the world in the period 2003-2004, and subsequently reviews the evolution of these activities during the period 1990-2004. The main conclusion of the study concerns the evidence of a convergent trend in terms of a common diversification strategy and activities. The author finally critically evaluates some consequences of the existence of this convergence, in terms of transport policy implications.

Quality in local public transport and its relevance for transport service contracts is the focus of the fifth paper, by Marcucci and Gatta. Given that service contracts constitute the wide-spread instrument to define bilateral commitments concerning contracts for transportation services and since, in this realm, the frequently divergent interests of public authorities and transport operators have to be reconciled via an appropriate regulatory apparatus, the definition, measurement and evaluation of quality constitutes a salient issue. In fact, in recent years the inclusion of quality requirements in contracts has become common practice, especially when adopting price-cap regulation. The authors suggest a criterion for service quality definition, measurement and integration in contracts for the production of socially valuable transport services. The adoption of choice-based conjoint analysis techniques to analyse customer preferences allows for the estimation of the passengers' evaluation of different service features and the calculation of a robust specification of a service quality index. The authors provide a practical application of the method proposed showing the results achieved in a case study that demonstrates how the procedure could be employed for measuring a service quality index in local public transport in a geographically differentiated market.

Finally, Meléndez-Hidalgo, Rietveld, and Verhoef in the sixth paper consider the more theoretical issue of "Surplus Equivalence", related to measuring benefits from transport infrastructure investments. In fact reductions in transport costs resulting from infrastructure improvements generate benefits measured as surplus changes either at an economy-wide scale (social welfare changes) or, else, as in cost-benefit analysis (CBA), at a transport market level as transport users' surpluses changes. The authors consider an economy with spatially separated markets integrated through a transport network, to study the equivalence between these two benefit measures. The strategy adopted, as also done by Jara-Díaz (1986), considers three different product market competition arrangements, it employs a two-node network and extreme competition assumptions on the production side: perfect competition and monopolistic production with arbitrage. The paper develop the work by Jara-Díaz by additionally considering perfect collusion (monopoly without resale) and Cournot-Nash oligopoly under flow-dependent transport costs (i.e. congestion in transport). The main results produced are based on numerical simulations in a three-node network with and without transshipment nodes.

The overall set of papers thus perfectly illustrates the broadness of the theme of the conference ("Pricing, Financing, and Investment in Transport"), both in terms of the wide variety of concrete issues covered by it, as by the array of methodological approaches that are used to address these.

European Transport \ Trasporti Europei n. 36 (2007): 4-26

Marginal social cost pricing in European seaports

Souhir Abbes [1*]

[1] *University of Nantes, Department of Economics*

Abstract

Europe's competitiveness in the global economy depends increasingly on an efficient and cost effective transport and port system. In the EU, Ports are becoming no different from any other multi-product industry offering a range of services and operating under different environments and organizational structures. Many port infrastructures and services are owned or managed by the private sector. In spite of the lack of standardization and homogeneity in ports, the European Commission is keen to adopt a common approach to pricing in ports. Therefore, a European Commission White Paper (1998) suggests the application of marginal social cost pricing taking into account externalities such as cost of accidents and environmental and congestion costs. The purpose of this paper is to show the recent changes in the port industry and to investigate the issue of port pricing in theory and the possible application of marginal social cost pricing in European seaports.

Keywords: Port pricing; Marginal social cost; Public goods; Privatization.

---

1. Introduction

Seaports are considered as an important link in distribution channels, particularly those involving international trade. The European port sector handles more than 90% of the Union's trade with third countries and around 30% of intra-EU traffic. Moreover, ports are an essential interface between seaborne and land-based modes of transport. In ports, as in many other industries, prices can 'make' or 'break' a port. Efficient prices can lead to prosperity and growth; the wrong ones can cause bankruptcy or the proliferation of subsidies and inefficiency. Europe's competitiveness in the global economy depends increasingly on an efficient and cost effective transport and port system.

Traditionally, ports were considered as a public good. Consequently, pricing and investment in this sector were among the responsibilities of the public port authorities. These responsibilities also covered security and environmental problems. In recent

* Corresponding author: Souhir Abbes (souhir.abbes@univ-nantes.fr)

years, privatization has changed this image. Ports have become no different from any other multi-product industry offering a range of services and operating under different environments and organizational structures. Many port infrastructures or services are now owned or managed by the private sector.

In spite of the lack of standardization and homogeneity in ports (the ownership; organization, and administration of ports as well as their size; pricing and subsidies with the functions and geographical location varying from country to country), the European Commission is keen to adopt a common approach to pricing in ports, assuring the real costs of port services should be borne by the users. Supporting this view, the Green Paper on Seaports and Maritime Infrastructure (1997) discusses common port policies in Europe. An adequate pricing system for port infrastructures and services is one of the main instruments for achieving efficiency in ports. Such a system could improve the efficiency of ports and ensure free and fair competition in the port industry. Hence, the European Commission White Paper (1998) suggests the use of social marginal cost pricing taking into account externalities such as the costs of accidents and environmental and congestion costs.

The purpose of this paper is to investigate the issue of port pricing and to discus social marginal cost pricing in the EU port industry. It is organized in three parts. The first part of this work (section 2) describes the evolution in port management. The following section investigates the theory of optimal pricing in ports. Finally, section four discusses the feasibility of introducing marginal social cost in the European ports.

2. Evolution in port management

2.1. Natural monopoly and public goods in seaports

Various theoretical reasons are given to justify public involvement in both the development and management of ports. Public intervention is mostly justified by either natural monopoly or public goods characteristics of some infrastructure and services provided in seaports.

Port investment poses a problem of indivisibility, high sunk costs and increasing returns to scale[1]. Moreover, many port zones are conditioned by the geographical features of the coast, so the available area is very limited and it is necessary to use it in the most efficient way. Such seaports, which are serving captive hinterlands, have competitive advantages (lower transport costs than a competitive port structure, geographical position, etc.) and face little competition. Therefore, some seaports operate as a local natural monopoly. The "monopoly argument" in seaports is then, only relevant when inter-port competition is imperfect (Goss, 1999). For those ports who serve a captive hinterland, monopolistic power enables ports to discriminate according to the elasticity of demand, so that port users experience high tariffs. The most frequently used regulatory systems to prevent abuse from a dominant monopolist and to prevent market power is the application of price-cap and the limitation of firm profits

[1] More details for these characteristics in seaports will be given in section 3.1.2.

through the rate of return or through the encouraging of intra-port competition[2]. Haralambides (2002) observes that for most ports, such captive hinterlands have diminished and that inter-port competition has become more intensive. Contestable hinterlands, those regions where there is no single port with a clear cost advantage, are more and more numerous and economic rents are smaller for competing ports.

The market is also inefficient for the production of a particular category of goods. These are public goods and services. One can mention for example, streetlights, the police and transport infrastructure. A public good is consumed simultaneously by several people without one person decreasing the quantities available for the others. One understands that the good is only public if there is no congestion. It is the principle of non rivalry in consumption. The second characteristic of public goods is non-exclusion. Indeed, one cannot exclude the consumer who doesn't agree to pay. If a good satisfies this condition, its marginal cost is equal to zero. Non-rivalry and non-exclusion are the intrinsic properties of public goods (Kaul and al. 2001).

However, the model according to which ports are seen as a public good begins to disappear and the present tendency is about the sharing of the functions in the port between the public and the private sectors. Given the importance of the investment in ports, the public nature of port infrastructures has been widely discussed in port economics.

Goss (1990) defines public goods as "those which are unlikely to be provided sufficiently, satisfactorily, or at all by competitive industries". To the general characteristics of a public good, Goss (1990) adds a third one. He alludes to three distinct conditions for the presence of public goods in ports:

- Their joint or non-rivalrous consumption;
- The inability to exclude those who refuse to pay;
- Their non-rejectability of consumption.

Goss (1990) notes that although some infrastructures and port services fulfill these theoretical properties, the practices and the experiences of some European countries have proven that the private sector is also able to exploit the port activity. For example, navigation aids and dredging are non-rivalrous "because their cost will be exactly the same no matter how many people are using them". However, the UK port privatization serves to confirm that the private sector is able to provide these services and that the costs can be recovered through user charges. The supply of a dredged channel, considered as a public good, is really no more than a political decision as there is clearly evidence, particularly from the UK, that the market is able to provide a deep-water channel and make an economic return from it (e.g. Harwich Haven, Southampton etc.). The second condition noted above relates to the inability to exclude those who refuse to pay for the use of the facility. Vessels using a given navigation channel are relatively easily recognizable. Those who refuse to pay would be subject to legal proceedings brought against them. For Goss (1990), the non-rejectability of consumption means that it is impossible for a user to reject the consumption of a good or service. Lighthouses and the police and security services are some examples. However, competition is so intense that some ships refuse these services and divert to other ports especially when

[2] Intra-port competition refers to a situation where two or more different terminal operators within the same port are vying for the same market.

the port serves a contestable hinterland. As Baird (2004) concludes, port users can reject a given port in favor of a competing facility elsewhere in the same region. The public good condition of non rejectability of consumption also therefore lacks substance.

In the port sector, lighthouses are the typical example of a public good. All the ships, no matter their number, can benefit from the light (Kaul and al., 2001). In several countries, notably in GB, lighthouses have been constructed by private investors, by maritime transportation companies or by partnerships[3]. Nevertheless, economists continue to define as public goods those goods that have the same features of lighthouses, specifying that it is not possible to fix a charge against their use given their indivisible costs. In addition to lighthouses, Sloman (1997) classifies pavements, dikes and police as public goods.

To identify public goods in seaports, Notteboom and Winkelmans (2001) classify port activities into two categories. The first one includes services whose costs can be covered by user charges. In this category, one can find pilotage, berthing, handling and storage. The second category includes the services provided without any discrimination between users. This category includes navigation aids, security, and the provision of a dredged channel. These services are defined as public services and should be provided by the public port authorities.

Erol (2000) has a contrary view point concerning the privatization of pilotage. He thinks that it is a very important service that must combine quality and safety. He noted that the privatization of this service in some countries has not succeeded insofar as the competition pushed the private firms to minimize their costs by using unqualified staff and inexpensive equipment. In other countries, firms in a monopolistic position have discriminated between users in service quality and prices. The World Bank (2001, p71) has underlined the risks of privatizing pilotage and several countries have forbidden its privatization. This is the case of the EU countries and the United States.

Some of these suggestions are a reaction to the publication of the European Commission Green Paper (1997). This paper considers that port access, quays and services related to navigation aids are considered as public goods. Consequently, their public financing is justified. The definition of a public good given by the European commission seems to be dictated not by the intrinsic characteristics of the good itself, but by political considerations (Bergantino and Coppejans, 2000). The aim of the following section is to identify to which extent public investment in ports is acceptable, notably in relation to its impact on competition.

2.2. Public financing of ports

The role of ports in economic growth justifies government intervention (Song and al., 2001). The economic influence of a seaport also spreads beyond the industrial and commercial sectors of a nation to include a whole economic region. In practice, the responsibility of investment in ports is divided between the local, regional and federal authorities. In the most widespread model in the EU, the federal authority takes in charge dredging and investments in maritime and land access infrastructures, whereas the regional authority is responsible for the development of the operational infrastructures (construction and maintenance of terminals). It leaves to the private

[3] Sechrest (2003) gives many examples of private provision of public goods from maritime history.

sector the responsibility to provide equipment and probably the other types of superstructure (buildings, etc.). In the autonomous ports in France, the state finances 80% of the investments in maritime access (dredging), 100% of their maintenance, and 60% of the investments in quays. For the remainder, the ports are self-financed, either by borrowing or financed by the local collectivities. However, there are no subsidies in the United Kingdom where seaports (public or private) must be self-financed.

Thus, there are two completely opposed arguments concerning the investment in the maritime ports. On one hand, the public nature of port infrastructures and their value added make it necessary for the state to finance them; on the other hand, subsidies hinder competition, and the non subsidized ports will be relatively penalized. Discussions about subsidies have increased with the increase in competition between Europeans ports.

Public authorities justify subsidies by the high costs of dredging and by the number of jobs that can be created once the port activity is maintained. However, Baird (2004) notes that the creation of employment on a local or national scale through the port activity, in spite of its advantages, cannot justify the subsidies. Indeed, the supply and demand of employment are determined by macroeconomic factors. Subsidies can only displace the supply or the demand from one market to another. Subsidies come with an increase in taxes which makes the port zone less attractive for residents and for industry. We can mention another inconvenience of subsidizing ports: subsidies allow imports to compete with exports which can lower the number of jobs. In some ports like Hamburg, Antwerp, Rotterdam, Le Havre and Bremerhaven, the expenses of dredging maintenance are substantial. Thus, Baird (2004) qualifies these expenses as "public investments" and thinks that the only arguments that can justify subsidies must be the net effects on producers, consumers and on the environment in the whole region concerned. Naturally, every member state is free to finance or not its maritime ports. However, and on a European scale, the investment decisions must take into account the environmental and economic consequences.

Van De Voorde and Winkelmans (2002) think that states should distinguish between public and private infrastructure in order to identify what can be subsidized. They consider that the financing of the basic infrastructures is tolerated. However, they don't explain why these infrastructures must be subsidized, nor to which extent the capital invested constitutes some "public good".

A unique theoretical answer to the identification of public goods and their financing in ports doesn't exist. Indeed, the recent deregulation of port activity has shown that the private sector is also able to exploit infrastructures and to offer port services.

2.3. Port activity deregulation

According to UNCTAD (1992), one can distinguish three port generations. The first remained predominant until the 1950s: the port was limited to providing a refuge, to ensure the transfer of goods and the temporary storage and delivery of goods. The role of the second generation ports expanded to industrial and commercial activities and the port turned into a center of handling and services. This function became more pronounced again with the third generation. Since the beginning of the 1980s, and to this day, the port has become a center of logistics and distribution. Some important

European ports such as Rotterdam and Antwerp have blossomed according to this formula.

The port infrastructure was therefore, for a long time, created, maintained and exploited by the public authorities. This model has begun to disappear with the growth of private capital in the construction of the infrastructure and the supply of port services. The models of port organization differ by the degree of the intervention of the private sector in the supply of infrastructure, superstructure and services. There are mainly four models: the *Landlord* port, the *Tool* port, the *Public Service* port and the *Private Service* port (World Bank 2001).

Table 1: Involvement of the private sector in port activities.

Management model	*Infrastructure*	*Superstructure and equipment*	*Stevedoring*
Public service Port	Public	Public	Public
Tool Port	Public	Public	Private
Landlord port	Public	Private	Private
Private Service Port	Private	Private	Private

Source: Kruk (2005).

The most widespread model in European ports is the Landlord port: the public sector is owner of the infrastructures and is responsible for regulation and control, whereas the private sector supplies services. The port authority plays the mediator's role between the public and private sector by coordinating the process of investment (Alfredo and Sabatino, 2005).

Four types of reform tools have served to increase the involvement of the private sector in ports (World Bank, 2001; UNCTAD, 1995b). Firstly, commercialization consists of splitting the main port activities. It gives more flexibility to the created entities and authority to operate in an autonomous way. In practice, these entities don't enjoy total liberty since the state intervenes in the decision process, in particular pricing and investment policies. Secondly, corporatisation is where the port authority is transformed into a commercial organization. However, the infrastructure remains public property. The objective of the port authority is to make as much profit as private operators. The third tool is liberalization. This method permits the transition from a monopolistic structure to a form where the private sector is more engaged in the operational process and in investment. The last tool is privatization. For several authors, privatization is synonym to reform in the port sector whereas it is only a tool that permits the introduction of the private sector into the port sector (World Bank, 2001, p38). Few ports in the world have been completely privatized (for example in the United Kingdom, Hong Kong and New Zealand). Most countries prefer partial privatization where the terminals are privatized under a management contract or concession agreement or Built Operate Transfer (BOT)[4].

Several motives are at the origin of port reforms: the need for resources to modernize the port, to finance the infrastructure or the facilities, to reduce costs and the public deficit and to increase efficiency in some activities by introducing more flexibility (World Bank, 2001).

[4] See Trujillo and Nombela (1999) for a detailed description of these contracts.

The involvement of the private sector in the port industry seems therefore feasible and desirable. So that, this sector can faces the increasing demand of maritime transportation. The type of the contract and price regulation are determining factors for the success of privatization and in the improvement of port efficiency. According to Song and al. (2001), the empirical studies haven't put into evidence the relation between the management model (public, private, or semi-public) and port efficiency. However, it has been shown that a positive relation exists between the involvement of the private sector in port activity and the increase in output (Estache and al., 2002). These conclusions are coherent with many economic findings that suggest that the separation between transportation activities where there are economies of scale (infrastructure supply) and other activities (the supply of services) reduces the total costs. However, one has to note that the fixing of cap prices can provoke the deterioration of the quality of services and can increase environmental damages.

Public ownership in the seaport industry has usually been justified by the argument that seaports play a key role for national economies, and they have special characteristics that can easily provide the firms running port facilities with market power (expensive specialized assets, sunk costs, indivisibilities and economies of scale) or some public good characteristics. Moreover, in some countries seaports are regarded as focal points for regional development, which justify the subsidies from governments for the building and improvement of port facilities. In the last two decades, a more competitive environment led to a consideration of the role that the public sector must play in the running of seaports.

International experiences have shown that private participation in both these aspects (operations and infrastructure) is beneficial for seaports. These experiences make a case for a revision of the traditional organization of seaports around the world, changes that will prepare ports for a more competitive market and less financial help from governments. The evolution of the port role and the involvement of private and public sector in port activity have induced a need of new models of port pricing. Thus, port pricing focused on the models of pricing that are based on cost but also on the strategic and commercial pricing that adjust more with the new forms of port organization. The forthcoming section is a literature review of the optimal port pricing.

3. Port pricing

3.1. Objectives and constraints

Pricing is the main tool of resource allocation. The major elements that enter into the determination of the pricing system for a port are objectives and constraints.

3.1.1. Port objectives

The politics of pricing depends on the fixed objectives. These objectives depend on the supplier of the service (public or private) and on market conditions. For example, the price fixed to maximize profit (objective of the private supplier, so commercial pricing), is not the one that could maximize the social welfare (objective of the public supplier, so marginal social cost pricing) or the one that permits to maximize the receipts. Sometimes the prices are fixed in order to reach objectives such as security, environment or to preserve a minimal market share. For port infrastructure, Bennathan and Walters (1979) have identified a major distinction between the European and Anglo-Saxon doctrines for setting port prices. They argue that European prices facilitate the economic growth of the port's hinterland, whereas the Anglo-Saxons attempt to ensure that ports cover their costs and, where possible, make a profit irrespective of the effects on the wider local economy.

The different objectives are complex and are often incompatible (see the table below for some examples). Whilst there are many transport pricing objectives, economists often focus on the pursuance of economic efficiency in the transport sector alone. Prices that are socially optimal are seen as the first-best benchmark (marginal cost pricing), which is in most cases politically desirable.

Table 2: Pricing policy objectives and possible conflicts.

Pricing Policy Objectives	*Conflicts*
Economic efficiency vs. profit maximization (or cost coverage)	Efficient pricing of the use of transport capacity may lead to financial losses
Profitability vs. income distribution	Pricing for profitability may lead to higher transport prices with adverse effects on the poor income groups
Economic efficiency vs. macroeconomic policy	Macroeconomic price restraint policies may conflict with the need to increase transport prices during periods of congestion and excess demand

Source: Adapted from United Nations, 2001.

The main constraints for ports are related to the economic character of infrastructure and service production. These constraints lead some ports to choose one pricing system and not another.

3.1.2. Port constraints

Economists recognize that the market remains inefficient in the presence of two major problems that are the economies of scale and external effects. Indeed, when economies of scale exist, pricing at marginal cost does not generate enough revenue for the firm to be financially self-sufficient. Furthermore, the presence of externalities, in particular pollution, and security problems justify the intervention of the government in the production of port infrastructure and services. These constraints influence the chosen prices.

Sunk costs and economies of scale

The magnitude of fixed costs depends on the size of the port and its geographical position. For Rudolf (1995), in general, for container terminals, nearly 80% of the costs are independent of the number of the ships and the volume of the handled merchandise. For bulk merchandise, fixed costs are lower and represent 60% of the invariable costs with the volume. Much port infrastructure and equipment doesn't have any other use if the port activity is interrupted. Sunk costs are irretrievable investments and are sometimes much higher than fixed costs.

In order to analyze economies of scale for ports, Bennathan and Walters (1979) suggest that it is necessary into take in account two parameters which are localization and size effects. The localization effect concerns the optimal localization of the infra and superstructures. As the demand varies, and the technologies of construction and dredging change, the present localization of the infrastructure becomes inefficient. The increase in demand and the construction of new sites entail high fixed costs that cannot be compensated by the decrease in the unit variable cost of the new infrastructures. Analytically, the port moves toward lower cost curves that didn't exist before the move to an optimal localization. The use of more efficient infrastructures and facilities in terms of volume and of deeper accesses (bigger ships can enter the port), permits a reduction of unit costs: this is the size effect. Bennathan and Walters (1979) conclude that, combining localization and size effects, all things being equal and for a given interval of handled quantity, the larger the port, the lower are the costs. Bennathan and Walters (1979) mention that economies of scale in ports are limited by the volume of the output and the size of the market in the hinterland and note, finally, that the structure of the costs in the port depends extensively on the geographical conditions and the indivisibility of the investments: some ports need more or less investment and maintenance according to their geographical location.

There are very few empirical studies on the evaluation of economies of scales for EU ports. The only studies have concerned Spanish ports. Both in infrastructure supply and cargo handling service, the authors have found increasing returns to scale. Four other studies have focused on economies of scale for the ports of Mobile, Melbourne and Ashdod. The results depend on the analyzed activity and the estimated function. The main conclusions are reported in the table below.

Table3: Economies of scale in port activities

Author	*Activity*	*Estimated function*	*Data*	*Scale economies evaluated in the approximation point*
Martínez Budría (1996)	Infrastructure	Monoproductive cost function	Panel data 27 ports in Spain	Increasing (S=3.47)
Jara Díaz and al. (1997)	Infrastructure	Multiproductive cost function	Panel data 27 ports in Spain	Increasing (S=1.43)
Jara Díaz and al. (2002)	Infrastructure	Multiproductive cost function	Panel data 26 ports in Spain	Increasing (S=1.69)
Reker and al. (1990)	Terminal-berth of containers	Production function	Panel data Three terminals in the port of Melbourne	Decreasing
Tongzon (1993)	Terminal-berth of Containers	Production function	Panel data Three terminals in the port of Melbourne	Increasing
Tovar and al. (2005)	Cargo handling	Multiproductive cost function	Panel data Tree port terminals in the Las Palmas port in Spain	Increasing
Chang (1978)	No mention	Production function	Time series for the port of Mobile	Constant
Kim and Shachis (1986)	Infrastructure and Services	Monoproductive cost function	Time series (the port of Ashdod)	Increasing (S=1.3)

Pricing at marginal cost, when there are economies of scale, leads to a financial deficit. Unless the port activity is subsidized, the application of another pricing system is required.

Externalities

In transport infrastructure, external costs are those imposed by the users of the infrastructure on the others. These costs may take the form of congestion, accidents and environmental costs. The efficiency of maritime transportation is heavily dependent on the smooth operation of land transportation. The ease of cargo handling and swift modal transfers are keys to successful intermodal operations. Port congestion poses a serious problem for handling firms and can be sometimes too expensive for them. This cost can come back in terms of more elevated rates of freight, a congestion of the traffic associated with the handling operations, a decrease in the level of security and a loss in terms of competitiveness in the whole region. Port authorities can prevent such problems by changing prices to adapt the supply to the demand, so by imposing congestion charges.

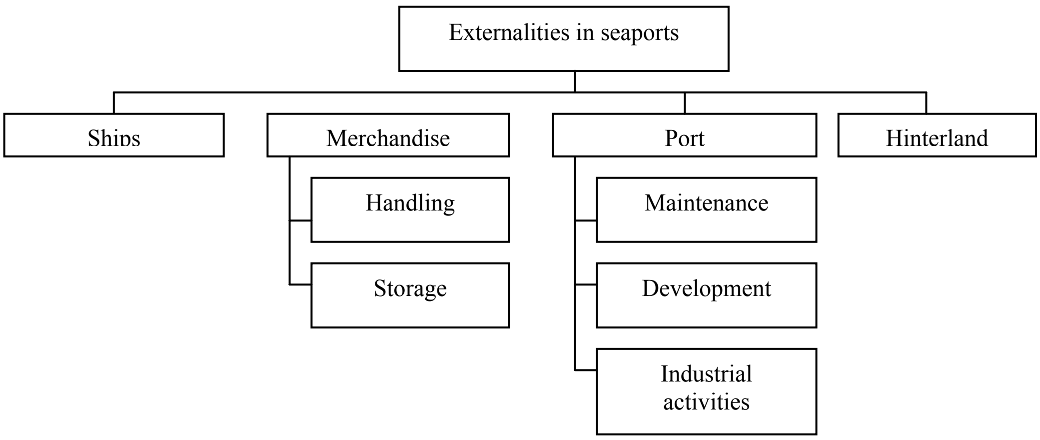


Figure 1: Sources of externalities in seaports.
Source: Adapted from UNCTAD (1993).

Given their position in coastal areas and the great variety of substances handled there, ports (especially those of a certain size) are now considerably complex systems from an environmental point of view. They tend to be associated with water and air pollution, soil contamination, problems related to dust and noise, the generation of waste, dredging operations, movement of ships, lorries and trains, warehouse storage of hazardous substances, etc.

In addition, ports are usually close to urban centers. That means that the impact on the environment, including people, of certain accidents can be very serious. Given the properties of some substances that pass through ports (chemical products, hydrocarbons, fertilizers, etc. (Planas-Cuchi and al., 1997) and the operations that are carried out on them (loading and unloading, storage, transport) the possibilities of there being an accident are not negligible. In fact, there are periodic fires, explosions and toxic releases, with possible consequences of financial losses, etc. In addition to these direct consequences, another quite important aspect of these situations is the negative image they give to the port, with the possible creation of a feeling of rejection among the population (Planas-Cuchi and al., 1998). All this justifies the importance of taking into account external costs when fixing prices.

The economic consequences of the economies of scale and externalities have justified for a long time public intervention into port activity. Considering their geographical position and the network of transportation, several European ports form, for a very big part of their traffic, a natural monopoly, notably when they constitute the necessary passage for a maritime company between two links. Between competition and monopoly, and with the evolution of the role of the port (from regional economic interest to a multiproduct firm), the literature of port pricing has focused on the pricing systems based on costs but also on other systems like strategic and commercial pricing.

3.2. Port pricing: literature review

Optimal pricing, although discussed considerably less in the economic theory for ports than for the other transportation infrastructures, remains the most controversial question concerning port economics. The first discussions on optimal pricing for port infrastructure and services began in the 1970s with Heggie (1974) and Walters (1974) and continue until now with Goss and Stevens (2001) and Haralambides (2002).

The current ideas on port pricing didn't have any theoretical foundations and just expressed points of view and recommendations. For Gardner (1977) for example, port tariffs traditionally based on ships and goods characteristics, should only be based on the nature of goods themselves. Whereas Thomas (1978) notes that port tariffs could form a meaningful proportion of the ocean freight rates. According to him, these charges must take into account several factors such as the nature of the commodities, their volume, the elasticity of demand and the type of the ship.

The end of the 1970s were characterized by the development of these rather simplistic thoughts and pricing systems toward the recommendation of more suitable pricing policies based on costs and efficiency objectives. With the recent participation of the private sector in port activity, other pricing principles, such as commercial or strategic pricing, have emerged. The presence of several actors in the port and the conflicts of interests between them are the main reason for "There is no single solution to the port pricing problem" (Strandenes and Marlows, 2000). The principles of port pricing can be classified in two categories: cost based pricing and alternative pricing methodologies.

3.2.1. Cost based pricing

Port authorities must have at least minimal knowledge of short and long-run costs of infrastructure use. This knowledge is necessary for an efficient running of the port activity, to adopt the best financial and administrative techniques and to take the appropriate investment decisions. As the scheduling of investments and the evaluation of the demand elasticity pass by prices, tariffs must correctly reflect the level of costs. This principle must be applied separately for all the activities in the port (infrastructure and services). The debate on cost based pricing turns mainly around marginal cost pricing (MCP) which is efficient and fair from an economic view and the methods of costs recovery.

Marginal cost pricing and congestion pricing

Marginal cost pricing in ports found its justification in public economy. In fact, most authors consider that ports are public goods, like road infrastructure, and that users should pay for the marginal social cost (MSC). This approach that privileges users who are ready to pay for the totality of the costs including the external costs, has been supported by some economists. However, the only externality that was considered was congestion.

Button (1979) set out to assess the viability of an economic-based pricing system arguing also that the users of a port (when viewed as a public utility) should be charged the full marginal social opportunity cost of the resources that they use. Using a simple model dealing with the economic allocation of car-parking places, he demonstrates that once the port capacity is not optimal, port authorities must levy a congestion charge to eliminate the excess of demand.

Bennathan and Walters (1979) conclude that, under perfect competition hypotheses, optimal port pricing is the one based on the short-run marginal cost if port capacity is optimal. They also note that there are economies of scale both in the supply of port infra and superstructure. So, short-run marginal cost pricing leads to a budget deficit and the port activity must be subsidized. These authors recognize additionally that these perfect conditions don't hold in reality. In general, ports are organized as a monopoly or cartel. The capacity is rarely optimal and the port must always face quay and hangar congestion. Bennathan and Walters (1979) show, supposing that port activity is monopolized, that it is more advantageous for the port authorities to increase the prices once the demand exceeds the supply. First of all, it constitutes an opportunity to appropriate the surplus caused by the growth of the demand. Secondly, these resources can constitute a self-financing for future investments. The third advantage is that congestion taxes encourage a more efficient use of the infrastructure. The main problem of congestion taxes remains administrative difficulties. Moreover, when some port charges are combined, it becomes difficult to fix the optimal tax.

Goss and Stevens (2001) join Bennathan and Walters (1979) in giving some arguments in favor of the short-run marginal cost pricing. They note that, under some conditions, this pricing system maximizes social welfare. The first condition is that all costs must be taken into account including those that don't appear in the accounting. These are externalities like congestion, pollution and noise. The second condition is to use a definition of marginal costs adapted to the accounting system. The last hypothesis is that all prices in the economy must be set to a marginal level. The exclusion of some externalities and the existence of some monopolistic pressures and taxes are barriers to the validity of this theory.

The defendants of MCP distinguish between long-run and short-run marginal cost pricing[5]. At this level, the points of view are different enough. By explaining the consequences of short-run and long-run marginal cost pricing under three different hypotheses (optimal capacity, sub-optimal capacity, and over-capacity in ports),

---

[5] Long-run marginal costs also include the capital costs of increasing capacity to accommodate an increase in output but have proven rather difficult to measure. It is only where capacity is non-optimal that the issue arises. Short-run MCP is seen as offering optimal use of existing capacity, whereas long-run MCP offers appropriate incentives to invest, although it may require regulatory action to ensure that the investment takes place.

Bennathan and Walters (1979) conclude that in all these cases, it is worth fixing prices at short-run marginal cost. Contrary to these two authors, Haralambides (2002) and Bromwich (1978) think that pricing must be based on the long-run costs in order to ensure the economic viability of the port. We can see clearly the difference between the two concepts by taking a look at the following graph, adapted from Haralambides (2002).

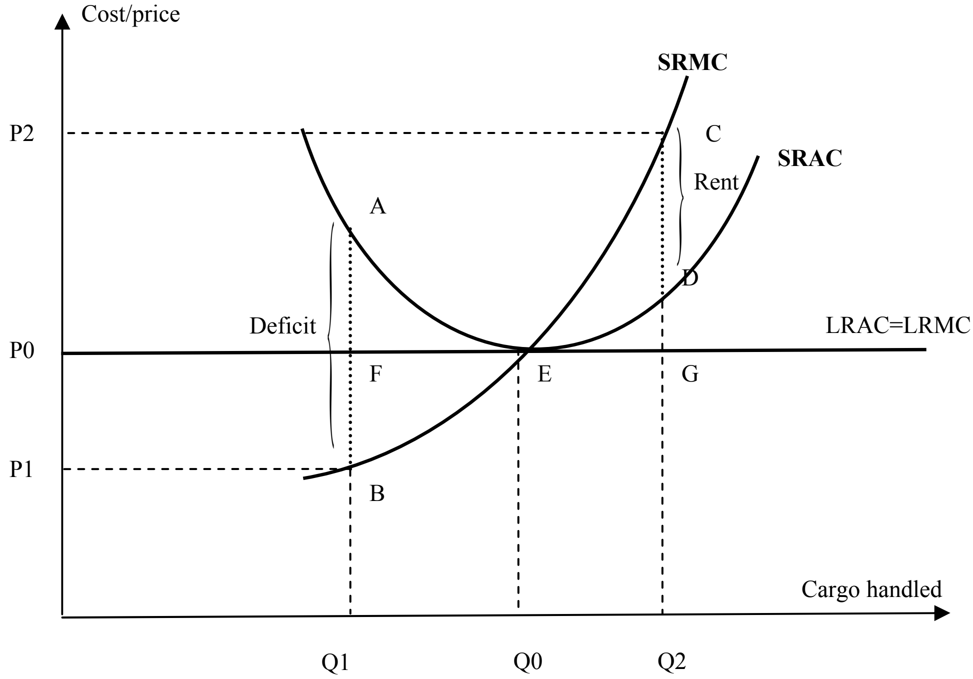


Figure 2: Marginal cost pricing in ports.
Source: adapted from Haralambides (2002, p333).

Assuming that the seaport operates under perfect competition, we suppose that the production structure is adequately described by SRMC (the short-run marginal cost), SRAC (the short-run average cost) and LRAC and LRMC (respectively the long-run average and marginal costs). Suppose first that the port operates under economies of scale and then the level of demand it has to satisfy is Q1, which is smaller than Q0. Pricing at short-run marginal cost (P1) will lead to a deficit of AB. However, pricing at long-run marginal cost (P0) will reduce this deficit to AF. For Haralambides (2002), if the port pursues a cost recovery objective, pricing at long-run marginal cost is more appropriate[6]. Public funding in this case is allowed given that it is temporary and declining. Indeed, given economies of scale, in the long-run (the point F gives the long-run equilibrium), the quantity of output Q1 will be produced by a much smaller port. On the contrary, if the capacity is suboptimal, the port exhibits diseconomies of scale. The demanded quantity of output is given by Q2 and congestion is a chronic problem for the port. At the same time, the port realizes economic rent given by CD. For Haralambides (2002), this situation is not sustainable in the long term. Attracted by the supernormal profits, the port's competitors will invest and expand to capture part of the economic rent, what gives the long-run equilibrium position of the port at the point G. In the case of constant returns to scale (the equilibrium is situated at point E), long-run and short-

[6] However, for Bennathan and Walters (1979), low prices (SRMC< LRMC) reflect the excess of capacity and encourage the use of the infrastructure. Short-run marginal cost is then more appropriate.

run marginal costs are equal. Pricing at marginal cost will attract just enough traffic to cover the port's cost. This was the spirit and philosophy of the European Commission's White Paper on fair payment for infrastructure use which ascertained that "the entire infrastructure complex of the EU as a whole may not exhibit economies of scale". This means that, at least, at an aggregated level, it should be possible to recover total costs through pricing at marginal cost (Haralambides, 2002).

The choice between short-run and long-run marginal costs depends on economies of scale and the efficiency of the investment policy (capacity). The marginal cost determination is complicated, in particular for port services and infrastructures. In fact, several costs are common which make the differentiation between variable and the fixed costs associated to every service difficult enough.

The approach of marginal cost pricing, which is a merely theoretical approach, has been criticized because of the difficulty of the evaluation of marginal costs and the application of this system in the real world. This has led some economists, who recognize all the same that port pricing must be based on costs, to recommend other more convenient pricing systems that permit the recovery of the costs such as the Ramsey pricing and two-part tariffs.

Ramsey pricing and two-part tariffs

In his paper, Button (1979) evokes the drawbacks of marginal cost pricing and suggests three methods to recover the costs. The first consists in subsidizing the port activity, the second consists in applying discriminatory charges between users and the last method intends to apply the two part tariff. In order not to diverge from a marginal cost pricing, Button (1979) suggests that the first tariff component equals the marginal social cost and that the second consists of a fixed charge levied for the right to use the facility.

Walters (1974) also thinks that marginal cost pricing constitutes an appropriate basis for port pricing. But when the port is congested or when there are economies of scale, it is necessary that port authorities set prices higher than average cost (which is not the marginal cost in most cases). Through examples of dredged channels supply, congested ports and cargo handling, Walters (1974) argues that two part tariff is more appropriate for ports than marginal cost pricing. Two part tariffs have the advantages (based on costs) and not the disadvantages (budgetary deficit and subsidies) of a marginal cost pricing.

Heggie (1974) was also one of the first economists to put forth some arguments in favour of cost based pricing. For him, the achievement of the port objectives and fairness are the most important reasons for which pricing must hinge on social costs[7]. Heggie (1974) opposed discriminatory and non transparent subsidies in particular and suggested the revision of the pricing system after every new investment. For Heggie (1974), in the case of congestion, excessive demand requires an intensive use of the infrastructure that should come with an investment. Fairness requires the spent capital to

[7] For Heggie (1974), social costs include environment and accident costs. But these are difficult to estimate.

be collected from the users through pricing rather than local taxpayers[8]. Therefore, it is necessary to apply a suitable pricing system by taking costs into account.

Jansson and Rydén (1979) developed a theoretical model of optimal pricing based on the costs of the inputs. Their proposition is different from the two-part tariff system used in the ports. Indeed, their model divides the port charges into two parts. The first component is a charge per ton of cargo that would be differentiated with respect to the elasticity of demand. And the second component is a charge levied on the carrier to reflect the opportunity cost of using the facility that is, optimal occupancy charges. It is the Value of Service Principle (VSP) that is similar to the Ramsey pricing (Ramsey, 1927) where the costs allocation permits the reflection of the elasticity of demand and recovery of the costs. In this case, the common costs are allocated reflecting differences among different users in elasticity of demand for the specific port services. Monopoly port services tend to be inelastic as long as port costs make up a fairly low share of the price of the cargo though even in monopoly ports there may be alternatives for storage outside the port. Elasticity of demand therefore may be lower for cargo handling services than for the navigational aids offered by the port. If so, this difference should be reflected in the allocation of common costs.

3.2.2. Alternative methodologies

Strategic port pricing

According to this approach, pricing is a strategic issue that must be guided by the port objectives. In his paper, Meyrick (1989) had considered that "insofar as the focus in pricing is on costs at all, it is on the average cost of service provision rather than the marginal cost" and that typically "port accounting systems are incapable of providing a basis for pricing on anything other than an average cost basis". He concluded his paper by suggesting several axioms for port pricing. He argues in particular for full cost recovery and that those costs arising from services or facilities provided for an identifiable user or group of users should be recovered from that user or group of users.

A similar axiomatic approach has been proposed by Talley (1994). He showed that this method can avoid the conflicts which might emerge between marginal cost pricing and full cost recovery in ports. He defines the axiomatic approach of pricing as "a pricing mechanism which determines the prices of the outputs of multi-product firms by allocating the full cost of production to all the outputs". Supposing that the demand for port services is relatively inelastic with respect to port prices and taking the pricing mechanism of Aumann-Shapley (that respects the following five axioms: cost sharing, rescaling, consistency, positivity and additivity), Talley proposes four axioms applicable to port pricing. These are rescaling, attributability, allocating and additivity which he then applies to container terminals. This methodology can help port authorities to determine prices that allow for full cost recovery without having to estimate marginal costs.

Finally, UNCTAD (1995*a*) mentions that pricing is a strategic question. Port authorities have to choose between the economic method based on marginal cost pricing and the financial method based on cost accounting. The chosen method can serve to

[8] Heggie proposes a pricing formula that takes into account operational costs (incliding the costs of capital replacement). See Heggie (1974) for more details.

accomplish the different objectives of the port. It is the CPV approach (Cost, Performance, and Value) that determines an interval for prices. Its pillars are as following:
- A cost based pricing for the port use.
- Pricing based on performance maximizes the output and reduces the congestion.
- Pricing based on the value of the services sufficiently generates income to cover the costs.

In order to improve the competitiveness of the port sector, Strandenes and Marlows (2000) proposed a pricing system based on the quality of the services that is, commercial pricing.

Commercial pricing

In most sectors of the economy, prices depend on the quality of the good or service. For Strandenes and Marlow (2000) port tariffs must be based on the quality of the services since it is quantifiable. The dimension "quality" includes some elements such as the time in port, punctuality, handling with minimum of damages, etc. The authors recommend a two-part tariff system. The first component doesn't depend on the quality of the service. The second one increases with the level of quality (reflected by the duration of port stay, the waiting time etc.). Such a pricing system sensitizes the port authorities and the ship-owners with respect to time (delays of transportation and handling). It also permits the improvement of the competitiveness of maritime transportation in relation to road transportation.

The ratio equilibrium

The ratio equilibrium can be interpreted as a competitive equilibrium of a market economy for which each agent receives a price signal, that consist in paying a share of the total cost personalized in such a way that each agent demands the same quantity of the good (Bergantino, 2002). The solutions obtained fall within the stand alone core: in other words, if each agent has the possibility of constructing the infrastructure that he needs, with existing technology, he would pay more than the share of the cost that he would have been assigned through the ratio equilibrium allocation criteria. Bergantino and Coppejans (2000) apply this concept to allocate maritime infrastructure cost and determine the optimal tariff structure for calculating the usage fees for the access channel to the port of Antwerp.

The general conclusion that one can draw from the review of the literature on port pricing is that "the prices must be based on the costs" (Button, 1979). However, the choice of an optimal "basis" for pricing, in spite of its theoretical interest, hasn't found the necessary attention when it concerned its implementation. The financial results are the privileged objective of the port authorities and the principles of social welfare are little known by decision-makers. The systems of pricing have remained complicated and have rarely reflected the structure of costs (Meyrick, 1991). According to Goss and Stevens (2001), it is in part the economists who excessively complicated the concepts and don't explain the direct relation between the economic theory and port activity. They rarely explain the difference between long-run and short-run costs and the importance to go beyond the costs to include social costs. With regard to social costs, very few authors have noted the importance of internalizing the costs of accidents and the environmental costs (see Heggie, 1974; Button, 1979). The following section intends to investigate the possible implementation of MSCP in European seaports.

4. Implementation of MSC pricing in ports

4.1. The European Commission point of view

On a European level, the question of pricing in transport has often focused on the debate concerning budgetary balance versus marginal social cost. This debate results from the theoretical arguments in favor of the two approaches. The preoccupations of the EC have evolved toward the integration of environmental costs into its pricing policies. In its Green Paper (1995), the EC underlines the importance of setting an efficient and equitable pricing system. The Green Paper indicates that a very clear shift exists between the prices paid by the users and the real costs. In particular, the costs related to accidents, to congestion and scarcity and to pollution are partially covered. The European Commission's 1997 Green Paper advocates a general framework requiring charges to be linked to costs. Different approaches are possible with regard to infrastructure costs: average cost pricing; charging for operating costs only and marginal cost pricing. Outside the port area, the Commission advocates a user-pays policy for all modes of transport, which would make for fair competition and affect the distribution of cargo flows among European ports. In 1998, the European Commission's common transport policy led to a White Paper which established the EU's intentions to apply short-run marginal social cost pricing to all transport modes. The major motivations behind this pricing policy are to improve efficiency, generate revenues and internalize the various externalities (congestion, scarcity, noise, environmental emissions and accidents).

Short-run marginal costs are the additional operating and maintenance costs associated with a marginal increase in output without any increase in physical capacity. If external costs are also included, this is referred to as marginal social cost (Alder and al., 2003). When capacity is optimal, short-run and long-run marginal costs are equivalent. The first one tends to be associated with regimes where government bodies take investment decisions and the second with regulated private infrastructure managers. However, another important issue is the time lags both in the adjustment of demand to price and of capacity to demand. Where capacity is slow to adjust and where demand adjusts reasonably quickly, short-run MCP is likely to be more efficient. This is indeed the approach taken by the EC, although not all economists agree. However, the issue of port pricing and the commission's involvement in it has not only risen from an academic basis but also as a response to the need felt in the port industry itself for a self disciplining mechanism that, if consistently applied, would eventually lead to the recovery of port investments and to future investments that are largely demand driven. This requirement has been the result of the recognition that, in the intensified regional port competition of today and the increasingly tightened fiscal constraints of an integrated Europe, it is no longer acceptable to discriminately and without a formal economic rationale spend taxpayer money on port investment, often aimed at increasing market share at the expense of the other ports, particularly those in neighboring member states (Haralambides and al., 2001). Nevertheless, the question which arises is whether the adoption of any financing or pricing system, or set pricing principles at the European level would be a valid policy option, given the significant constraints that exist and that may prevent or delay the implementation of MSCP in ports. These constraints include economic, organizational and acceptability-related barriers.

4.2. Barriers to the implementation of MSC pricing

4.2.1. Economic constraints

The analysis of present pricing policies in European ports, conducted as part of the ATENCO study, demonstrated the substantial diversity prevailing among EU ports with regard to their financing, accounting and charging practices. This diversity is deeply rooted in various judicial and cultural traditions, as well as in the divergent port management styles, related responsibilities and degree of autonomy. Current port pricing practices have been mainly based on empirical intuition and past trends. These tariff levels do not reflect the actual costs levied by the port operations and they do not recover all the costs, thus creating severe inefficiencies such as congestion as well as sources of significant financial loss. Meersman and al. (2002) argue that outlining a typology of current port pricing schemes is not possible due to the complex, untransparent and archaic systems in existence at present.

European ports work under widely different conditions. Some ports suffer significant overcapacity, while others are lacking in infrastructure. Indeed, it is difficult to say that short-run MCP is an efficient pricing basis for all ports. Another argument, is the one that "in attempting to apply MCP to ports, a practical problem arises – the inability of ports to determine the marginal costs of their services" (Talley, 1994) and the difficulty to estimate external costs. Mayers and Proost (2003) showed that efficiency, equity and acceptability of any reform in pricing, depend not only on the change in transport prices, but also on the way the extra revenues are used. Here, one can wonder if implementing MSCP in privatized or semi-privatized ports is appropriate. Appling the polluter pays principle in seaports must be highly regulated to ensure that private operators use efficiently tax revenues to reduce port externalities. Furthermore, efficient pricing in the port sector will not be able to bring about the expected welfare effects if the rest of the related infrastructure is not priced accordingly. The issue thus appears to be reaching a standstill, particularly in view of the fact that in most countries, ports are considered as part of the country's infrastructure and thus State investment in ports is considered as 'public investment' outside the reach and mandate of the European Commission. In this particular context, Haralambides and al. (2001) note that, a voice that is often loudly raised, by both the commission (recently) and the port industry, argues that MC pricing applied to ports only, will make port services 'unilaterally' more expensive thus penalizing the union's efforts to check road traffic and promote short sea shipping.

Some authors note that pricing matters (at least in a liberal economic environment such as that in the EU) ought to be, ideally, left to the producers (ports) themselves (Haralambides and al., 2001). They argue also that greater private sector interest in the port industry, as well as in the rest of the Europe's infrastructure, necessitates some form of cost based pricing that would allow the recovery of port investments. This could, however, disturb the existing "equilibrium" among ports that has been established over the years as a result of each port's particular characteristics such as geographical location, proximity to markets, navigational constraints, subsidies and types of financing. Among competing container ports, like those of Western Europe, such a "disturbance" can have marked impacts on ports' market shares, as a result of the ease with which carriers can nowadays switch between ports. Swahn (2002) argues that port infrastructure pricing, given certain boundary conditions of accounting and transparency, could be left up to ports as market actors without causing any significant distortions.

Finally, Matthews and al. (2002) state that "MCP is clearly much easier to implement where the infrastructure manager is a public body, funded largely from general taxation, as in Sweden". On the other hand, Jansson and Shneerson (1982), Verhoeff (1981) and Dowd and Flemming (1994) think that it is more difficult to allocate costs in *Public Ports* in a way that is largely arbitrary compared to costs in *Landlord Ports* where more clear lines of responsibility and accountability exist.

4.2.2. Organizational constraints

If the aim of a global European policy is the establishment of a "level playing field" among competing European ports, it should be recognized that any assessment of potential improvements cannot be solely undertaken in terms of purely market based considerations. In contrast to many other sectors where liberalization and market based rules have been widely credited as instrumental to the creation of better and best practices, the diversity of the port sector requires uniform methods of accounting and cost recovery using a step-by-step approach. Undoubtedly, such an approach should consider, at least in the short-run, national perception on the appropriate role of public investment, still the prerogative of member states. However, considerable progress could be made through efforts aimed at harmonizing definitions and classifications of port infrastructure. Current classifications (eg. investment inside or outside the port area) often lack an economic rationale and are instead based on technical or geographical considerations aimed at determining whether investment costs should be allocated directly to users or to society at large (Haralambides and al. 2001).

Pricing external costs will need to establish some organizational structure to ensure that taxes are effectively used to reduce congestion and environmental damage. Given the different port models, this seems to be relatively complicated. The lack of data and the lack of harmonization of pricing principles across EU countries is also not considered. Both parameters imply that, to make some headway in formulating institutional change, substantial attention should be devoted to historical trajectories and path dependencies associated with specific (sub)national port financing and pricing routines. If not, the danger exists that substantial unintended policy effects might arise.

Alder and al. (2003) note also that, for ports, there is no agreed pricing formula defined at the EU level or EC Directive, as there is currently for rail. "For inland shipping and waterways, the major barrier to the implementation of any pricing scheme on the river Rhine is the Mannheim convention (1868), as one of its main principles is exemption from navigation duties. However, this does not prevent the introduction of port related charges, but rather will cause difficulties in introducing en-route charges to account for emission externalities" (Alder and al., 2003).

4.2.3. Acceptability

The increasing transformation of ports (at least for competing regional container terminals) from public to private enterprises raises the issue of the desirability and fairness of pricing methods focusing on the "user" rather than the "general taxpayer" (Green Paper, 1997). Cost recovery implies that port revenues will be generated from the user of a facility, who will have to somehow pass these costs on to the final consumer. This consumer will in all likelihood have to pay higher prices for the goods he consumes but, at least in efficient markets, he is compensated by correspondingly paying fewer taxes (for infrastructure investments). Obviously, the problem that arises is the acceptability of such a switch. By changing prices, port authorities and private companies fear that users will move to other ports.

The power of oligopolistic shipping lines and hub ports are likely to pose serious political barriers to MCP implementation. Indeed, in the seaport sector, several actors may be involved in the vertical port activities chain and the horizontal port activity cluster. Because of the potential of unintended policy impact, different financing disciplines externally imposed on ports may disturb the effective horizontal and vertical linkages among the various port actors.

5. Conclusion

We have witnessed in recent years a significant change in the organization of the port industry in the EU. The British port sector is a very good example in this respect. Private sector intervention has affected all the port services and activities, even those that where considered as public goods and those that must be controlled by the public authority for security and environmental reasons (aids to navigation, pilotage, dredging, etc.). While such sweeping change has not occurred in the case of continental ports, there has been an unmistakable trend towards greater autonomy and a more substantial private stake in goods handling. We can find different management models in EU ports where the involvement of the private sector varies widely (*Public Service port, Tool Port, Landlord Port* and *Private Service Port*). In spite of these changes, the European legislation seems to lack precision when it concerns identifying and financing public goods in ports.

Neither is the situation clear concerning pricing. The Green Paper (1997) suggests several pricing methods (marginal cost pricing, average cost pricing, etc.) to ensure cost recovery and to apply the user-pays principle. To establish a common European transport policy, the White paper (1998) recommends the application of marginal social cost pricing (MSCP) to all transport modes. This pricing approach will improve efficiency and ensure free and fair competition in the transport sector. Nevertheless, the EC has not specified which costs should be taken into account in the port sector. Although MSC constitutes a first best in economics, its application seems to be controversial in the port sector. We can find economic, organizational and acceptability related barriers to its implementation.

On the one hand, if there are economies of scale, MSCP leads to a financial deficit. Therefore, port activity needs to be subsidized. On the other hand, subsidies hinder competition, especially in some economic regions like the EU. So, free and fair competition and MSCP in European seaports are two incompatible goals. Even if we suppose that the port authority may let several private operators do a bid to acquire the port operation where the private operators also mention the prices they will charge to the users and the subsidy they wish to receive from the port authority, several organizational constraints will form a barrier to the application of MSCP. Moreover, this will certainly create inefficiencies and acceptability problems in the port sector. The lack of European harmonization in what concerns accounting, financing, the physical infrastructure limits and the costs to take into account are also barriers to the implementation of MSCP. Other reasons are related to the difficulty of estimating marginal costs, both internal and external. Finally, there are acceptability constraints. For the users, port prices and dues will be higher if they have to pay for external costs too. Ports authorities and operators fear that they will move to other ports. For the consumer, the price of goods will be higher. Here, the problem arises whether they will agree to pay higher prices for goods and less tax for infrastructure investment.

Therefore, although some aspects of MSCP are important and appreciable (attention to environment, necessity of a cost based pricing and for harmonization to improve efficiency, etc.), this pricing principle does not seem appropriate to meet the objectives of the European commission. We think that MSCP permits the application of users/pays principle only if marginal costs are calculated for each activity when analyzed individually. Given the economic characteristics of seaport costs and the diversity of port activities, this seems to be very difficult. Moreover, European Commission objectives concern the development of a level playing field between ports and the reduction of public financing in port infrastructure. MSCP is not, by definition, consistent with all these objectives.

References

Alder, N., Nash, C. and Niskanen, E. (2003) "Barriers to efficient cost-based pricing of rail, air and water transport infrastructure in Europe", paper prepared for the fourth seminar of the *IMPRINT-EUROPE Thematic Network*, "Implementing Pricing Policies in Transport: Phasing and Packaging", Brussels.

Alfredo, J. and Sabatino, P. (2005) *Legal framework in the context of port commercialization*, http://www.sabatinop.com/art/art1.htm.

ATENCO (1999) "Analysis of the cost structure of the main TEN ports", European Commission, *Fourth Framework Research Programme*.

Baird, A. J. (2004) "Public goods and the public financing of major European seaports", *Maritime Policy and Management* 31 (4), pp. 375-391.

Bennathan, K. J and Walters A. A. (1979) *Port pricing and investment policy for developing countries*, Oxford University Press: Oxford.

Bergantino, A. S. and Coppejans, L. (2000) "Shipowner preferences and user charges: Allocating port infrastructure costs", *Transportation Research* Part E36, pp. 97-113.

Bergantino, A. S. (2002) "The European Commission approach to port policy: some open issues", *International Journal of Transport Economics* Vol XXIX (3), pp. 337-379.

Bromwich, M. (1978) "Port costs: an alternative approach", *International Journal of Transport Economics* 5 (3), pp. 227-243.

Button, K. J. (1979) "The economics of port pricing", *Journal of Maritime Policy and Management* 6: 201-207.

Chang, S. (1978) "Production function and capacity utilization of the port of Mobile", *Maritime Policy and Management* 5, pp. 297-305.

Dowd, T. J. and Fleming, D. (1994) "Port pricing", *Maritime Policy and Management* 21 (1), pp. 29-35.

Erol, C. A. (2000) "Pilotage services and competition", *Turkish Shipping World Magazine*, October 2000 issue.

Estache, A., Gonzalez, M. and Trujillo, L. (2002) "Efficiency gains from port reform and the potential for yardstick competition: lessons from Mexico", *World Development*, 30 (4), pp. 545-560.

European Commission (1995) *Green Paper towards fair and efficient pricing in transport policy: Options for internalizing the external cost of transport in the European Union*, COM (95) 691, Brussels.

European Commission (1997) *Green Paper on Seaports and Maritime Infrastructure*, COM (97) 678, Brussels.

European Commission (1998) *White Paper on Fair Payment for Infrastructure Use: A phased approach to a common transport infrastructure charging framework in the EU*, COM (98), 466 Final, Brussels.

Gardner, B. M. (1977) "Port pricing, an alternative approach", In *Transport of steel exports - an investigation into the scope for rationalization*, Department of Maritime Studies. University of Wales, pp. 347-365.

Goss, R. (1990) "Economic Policies and seaports, part III: are port authorities necessary?" *Maritime Policy and Management* 17 (4), pp. 235-249.

Goss, R. (1999) "On the distribution of economic rent in Seaports", *International Journal of Maritime Economics* 1(1): 1-9.

Goss, R. and Stevens, H. (2001) "Marginal Cost Pricing in Seaports", *International Journal of Maritime Economics* 3 (2), pp. 128-138.

Haralambides, H. E., Verbeke, A, and Musso, E. (2001) "Port financing and pricing in the European Union: Theory, Politics and Reality", *International Journal of Maritime Economics* 3, pp. 368-386.

Haralambides, H. E. (2002) "Competition, Excess Capacity, and the Pricing of Port Infrastructure", *International Journal of Maritime Economics* (4), pp. 323-347.

Heggie, I. G. (1974), "Charging for port facilities", *Journal of Transport Economics and Policy* 8, pp. 3-25.

Jansson, J. O. and Rydén I. (1979) "Swedish seaports: Economics and Policy", *Working Paper*, The Economic Research Institute, Stockholm School of Economics, Stockholm.

Jansson, J. O. and Shneerson, D. (1982) *Port Economics*, MIT Press, Cambridge.

Jara Díaz, S., Martínez-Budría, E., Cortes, C. and Vargas, A. (1997) "Marginal costs and scale economies in Spanish ports", *25th European Transport Forum*, Proceedings Seminar L, PTRC, London, pp. 137-147.

Jara Díaz, S., Martinez-Budría, E., Cortes, C. and Basso, L. (2002) "A multioutput cost function for the services of Spanish ports' infrastructure", *Transportation* 29 (4), pp. 419-437.

Kaul, I., Grundberg, I and Stern, M. A. (2001) "Global public goods", *International Cooperation in the 21st Century,* London: Oxford University Press.

Kim, M. and Sachish, A. (1986) "The structure of production, technical change and productivity in a port", *International Journal of Industrial Economics* 35 (2), pp.209-223.

Kruk, C, B. (2005) "Port Reform process Review", *Transport forum*, March 2005.

Matthews B., Nash C., Nilsson J-E, and Farkas G. (2002) "Institutional and Technological Barriers to Implementation - Rail Transport", *MC-ICAM* Task 5.2 part 1.

Martínez-Budría, E. (1996) "Un estudio econométrico de los costes del sistema portuario español", *Revista Asturiana de Economía* 6, pp. 135-149.

Mayers, I. and Proost S. (2003) "Reforming transport pricing: an economic perspective on equity, efficiency and acceptability", In Schade J. and Schalg B. *Acceptability of transport pricing Strategies*, Chapter 6, Elsevier Science.

Meersman, H., Van de Voorde, E, and Vanelslander, T. (2002) "Port pricing issues: Considerations on economic principles, competition and wishful thinking", Paper presented at 2nd *IMPRINT-EUROPE*, Brussels.

Meyrick, S. (1989) "Port Pricing: some observations on principles and practice", *ASRRF Working Paper* 5, Centre for Transport Policy Analysis, University of Wollongong.

Meyrick, S. (1991) "How far does economic theory really take us?", Proceedings from *New Thinking in Port Pricing*, University of Wollongong.

Notteboom, T. E. and Winkelmans, W. (2001) "Reassessing public sector involvement in European ports", *International journal of Maritime Economics* 3 (2), pp.242-259.

Planas-Cuchi, E., Montiel, H. and Casal, J. (1997) "A survey of the origin, type and consequences of fire accidents in process plants and in the transportation of hazardous material", *Trans. Ichem E* 75, pp.3–8.

Planas-Cuchi, E., Vilchez, J. A., Pérez-Alavedra, F. X. and Casal, J.(1998) "Effects of fire on a container storage system, a case study", *Journal of Loss Prevention in Process Industries* 11 (5), pp.323–331.

Ramsey, F. (1927) "A contribution to the theory of taxation", *Economic Journal* (37), pp.47-61.

Rekers, R. A., Connell, D. and Ross, D. I. (1990) "The development of a production function for a container terminal in the port of Melbourne", *Papers of the Australasian Transport Research Forum* 15, pp.209-218.

Rudolf, C. D. (1995) "A cost Comparison of Modern Container Cranes" Proceedings of the conference*: Ports'95, American Society of Civil Engineer*s.

Sechrest, L. J. (2003) "Private provision of public goods, theoretical issues and some examples from maritime history", *Working Paper*, Sul Ross State University, Texas.

Sloman, J. (1997) *Economics*, 3rd Edition. Hemel Hempstead, UK: Prentice Hall Europe.

Song, D. W., Cullinane, K. and Roe. M. (2001) "The productive efficiency of container terminals: an application to Korea and the UK", *Working Paper*, Aldershot, UK: Ashgate.

Strandenes, S. P. and Marlows, P. B (2000) "Port pricing and competitiveness in short sea shipping", *International Journal of Transport Economics* 23 (3).

Swahn H. (2002) "Marginal cost pricing in the maritime sector: Cost calculation, acceptance and Swedish infrastructure charging practice", Paper presented at *3rd IMPRINT-EUROPE Thematic Network Seminar*, Brussels.

Talley, W. K. (1994) "Port pricing: a cost axiomatic approach", *Journal of Maritime Policy and Management* 21, pp.61-76.

Thomas, B. J. (1978) "Port Charging Practices", *Journal of Maritime Policy and Management* 5, pp.117-132.

Tongzon, J. L. (1993) "The Port of Melbourne Authority's pricing policy: its efficiency and distribution implications", *Maritime Policy and Management* 20 (3), pp.197-203.

Tovar, B., Jara-Díaz, S. and Trujillo, L. (2005) "Multioutput analysis of cargo handling firms: An application to a Spanish port", *Transportation* 32 (3), pp.275 – 291

Trujillo L., and Nombela G. (1999) "Privatization and Regulation of the Seaport Industry", the *World Bank, Policy Research Working Paper*.

UNCTAD (1992) "The Development and Improvement of Ports", *United Nations Conference on Trade and Development*, TD/B/C.4/13 Geneva.

UNCTAD (1993) "Sustainable Development for ports", *United Nations Conference on Trade and Development*, 27/8, Geneva, Switzerland.

UNCTAD (1995*a*) "Strategic Port Pricing", *United Nations Conference on Trade and Development* /SDD/PORT/2, Geneva.

UNCTAD (1995*b*) "Comparative Analysis of Deregulation, Commercialisation and Privatisation of Ports", *United Nations Conference on Trade and Development*, Geneva.

United Nations (2001) *Sustainable transport Pricing and Charges: Principles and Issues*, Bangkok, New Delhi.

Van De Voorde, E. and Winkelmans, W. (2002) "A General Introduction to Port Competition and Management", In *Port competitiveness: An Economic and Legal Analysis of the Factors Determining the Competitiveness of Seaport*, edited by M. Huybrechts, H.

Verhoeff, J. M. (1981) "Seaport Competition: Some fundamental and political aspects", *Maritime Policy and Management* 8, pp.49-60.

Walters, A. A. (1974) "Marginal cost pricing in ports", *The Logistics and Transportation Review* 11: 299-308.

World Bank, (2001) World *Bank Port Reform Tool Kit, Alternative port management structures and ownership models*, Washington DC, United States.

European Transport \ Trasporti Europei n. 36 (2007): 27-46

Transport megaprojects in Italy. A comparative analysis of economic feasibility studies into EIAs.

Paolo Beria [1*]

[1] *Politecnico di Milano, Milan, Italy*

Abstract

The paper presents a survey conducted on the environmental impact assessments and the feasibility studies for fifteen selected infrastructural transport projects all over the Italian territory and approved under the so called "Objective Law". The survey collects data concerning the demand forecasts, the definition of alternatives and the cost benefit analysis, with their level of detail.

The aim of such work is to derive some general considerations about the quality, the transparency and the contents of the assessment, in the context of the Italian normative framework.

For the analysed projects, it is possible to recognise a lack of contemplation to the definition and in the construction of the demand forecast model, the use of extremely standardised schemes of comparison, the presence of double counting and some theoretical errors concerning costs evaluation. Moreover, the parallel analysis of different projects on the same study-area shows, in some cases, the absence of any co-ordination among projects, as well as different outputs from the transport demand forecast model; this creates problems in validating and comparing the results of each analysis.

After an introduction to the normative framework, the adopted methodology is presented. After that, the paper – reflecting study objectives – includes first, an analysis of alternatives definition and tools used for the decision support system; second, the discussion and comparison of the demand forecasts quality; third, the approach used for the quantitative economic assessment. Then, as example, the standardised railways operator procedure has been analysed and commented. Finally some conclusions will be drawn.

Keywords: Evaluation; CBA cost benefit analysis; Transport infrastructure; EIA environmental impact assessment.

Normative context and paper aims

Due to a real or apparent deficit in national infrastructure, transport planning in Italy has, of late, been intensive. The 2nd General Plan of Transport and Logistics ("PGTL")

* Corresponding author: Paolo Beria (paolo.beria@polimi.it)

was published in the year 2000, followed in 2001 by the "Objective Law" (L. 443/2001, "*Legge Obiettivo"*). This law consists of delegating infrastructure and national development to the Government. The Government can produce a list of public and private investments, following some in-house criteria and guidelines (but not declared on a plan nor in law), like the filling of the infrastructural gap between the North and the South of the country or the increase of national competitiveness. The law declares its own consistency with the General Plan of Transport and Logistics of 2000, but in fact the Government can freely express a list of infrastructure, as it did. Consequently the PGTL, if still in effect, has been made redundant as a planning document by the Objective Law[1].

A key point to be explored in such planning context, is the relation between the economic feasibility study (Cost Benefit Analysis in the Italian normative) and the Environmental Impact Assessment (EIA)[2]. The law 443/01 (and the following decree DLGS. 190/2002) partially reform the procedure of the EIA. Among other measures, the reform makes both the approval and start of the chosen investments easier and faster.

Concerning the verification of such evaluations, the L. 308/2004 establishes that (comma 9, letter f) the EIA procedures must consider the cost/benefit ratio of the project from the environmental, economic and social point of view. APAT, the former National Agency of Environmental Protection, published a document containing the guidelines for the evaluation of EIA studies (APAT, 2001). Among the criteria of evaluation and possible rejection of a project by the Ministry of Environment, there are the "criteria of preventive verification":

> *... they take their origin from different evaluation contexts (economic cost/benefit, technological feasibility). The EIA* ***must simply verify*** *that these evaluations took place in order to avoid the occurrence of useless or technically wrong investments that produce […] unnecessary impacts on the environment. ...* [3].

In addition, one must consider that the feasibility studies and the subsequent EIA (even if not indicated by the EU normative as a tool for economic impact verification) are the only aspects of the process in which a compulsory and public socio-economic evaluation of projects takes place.

Alternatively, it would be useful defining the main and unavoidable elements forming a correct evaluation (alternatives selection, demand forecast methods, economic feasibility algorithm). Only a correct assessment can aim to justify the choice of one project from a shared and clear planning context, as in the Italian situation. The use of "shopping lists" in addition to inconsistent evaluations can potentially justify projects without any social or economic justification. Of course, the definition of a correct evaluation is not entirely obvious, but the international practice and literature can help in doing so.

The aim of this study is thus to verify in a sample of EIAs if some key aspects of the evaluation and decision making that concern infrastructures take place in a correct, shareable and transparent way. The presented material itself can provide some

---

[1] Brambilla et al (2003) reports 80 interventions in the initial phase, becoming over 200.
[2] defined by Italian law DPCM 27/12/1988.
[3] APAT, 2001, page 19. The bold and the translation are by the author.

significant judgements about the quality of projects and allows us to draw some conclusions about the inadequacy of the Italian normative concerning project selection.

Methodology

The approach used for the analysis of the quality of assessment is based on a comparison grid, including the most relevant aspects concerning the evaluation of investments[4]. The focus is on the aspects of defining alternatives, demand forecast and economical analysis; all of which concern the decision making process. The analysis is done by answering whether some aspects of project assessment, that theory and international practice suggest, are present and correctly treated. The filling of the grid is, of course, a partially arbitrary operation because it sometimes implies some judgements by the author, but with acknowledgement to literature and laws.

Table 1: projects analysed.

	Proponent	Short name	Project description
H1	CTM	Region Lazio	Highway link Roma - Formia: "Corridoio Tirrenico Meridionale".
H2	A4 Milan-Turin	ASTM	Upgrade of Turin – Milan highway.
H3	Pedemontana	Region Lombardia	Highway link Dalmine-Como-Varese, Gaggiolo pass and related works.
R1	Paullese	Milan Province	Upgrade of provincial road between Peschiera Borromeo and Spino d'Adda
R2	Valtellina	ANAS	Accessibility to Valtellina valley
R3	SS77	ANAS	SS77 Val di Chienti; Umbria - Marche road system ("Quadrilatero Umbria - Marche")
R4	Modena – Sassuolo	ANAS	Highway link Modena - Sassuolo - Campogalliano
L1	Brenner	BBT	Upgrade of rail link Munich – Verona and Brenner Base Tunnel.
L2	Verona-Brenner	RFI	Railway line Verona – Fortezza, 4th track.
L3	Foligno-Fabriano	RFI	Railway line Orte – Falconara. Foligno – Fabriano section doubling.
L4	Falconara	RFI	Falconara (Ancona) rail node and link between Orte – Falconara line and eastern coastal line.
L5	Terni-Spoleto	RFI	Railway line Orte – Falconara, doubling of Spoleto – Terni section.
L6	Pontremolese	RFI	Railway line "Pontremolese"
L7	Rho-Gallarate	RFI	Upgrade of Rho – Arona rail line, section Rho – Gallarate.
L8	Verona-Padua	RFI	High speed/High capacity railway line Verona – Padova.

Elaboration of the author.

The grid, which will be fully detailed in the following sections, aims to include and systemise the most relevant aspects forming a functional, transparent and impartial evaluation, pointing out potential biases and errors present in the EIAs. Good practice is suggested by the theory cited in literature concerning the implementation of EIA procedures (Baccaro et al., 2005 for the Italian normative; European Commission, 1997)

[4] The quality of environmental analysis, even if extremely relevant for the decision and the design of the infrastructure, has been excluded from our comparison.

or by experience and common sense (for example when is stated that it is meaningless to compare only one alternative). Specifications are provided in the following sections.

The sample analysed is made of fifteen homogeneous[5] projects, listed in the Table 1. Seven of them are road or highway projects, eight are rail lines or nodes. For all of them the official EIAs and CBAs have been analysed[6].

The grid has primarily been filled in with articulate descriptive answers. However, to be easily represented here, it has been simplified using a yes/no format. In the following sections, the most relevant aspects are illustrated in table and graph format. The green bars of graphs represent the number of projects for which a positive answer has been given, in red the negative. For some projects data has been incomplete. In this case the bar is shorter than the total number of projects (15 in total or 7 road + 8 rail). More details about methodology and assumptions will be given in the next sections.

Definition of alternatives and decision tools

The issue of defining alternatives is very relevant and delicate. According to the common theory of assessment (see next paragraph), the definition of alternatives should be done during the planning phase in order to highlight the most compatible and feasible solution. In this respect, the EIA should not consider alternatives, but highlight environmental concerns. In this sense, according to directive 97/11/CE (art 5, comma 2), the evaluation activity in the EIA can be limited to a function of support.

Nevertheless some path alternatives are presented in EIAs where their purpose is not only an environmental impact minimisation. These are usually analysed, and chosen, also in terms of functionality, technical feasibility, investment cost minimisation (even if considering investment only, instead of life cycle total cost), consistency with planning tools and so on.

The scheme used here for comparison is articulated into four main areas: the number of alternatives considered, their extent, the level of description, the use of such scenarios. The following table includes all the results for each project and for each criterion considered[7]. The following paragraphs will comment the results.

[5] All projects refer to the same laws and have been assessed during the same government. However, the authors of the projects and EIAs are various.

[6] In some cases the Ministry of Environment required additional parts concerning the aspects analysed. When additions can be considered more satisfying, this study considered them, instead of the original documents.

[7] In the table, “yes” means that the element is present and/or correct, “no” that is absent and such absence constitutes a conceptual error. The blank means that it was not possible to answer or that the answer has no sense. For example, “no alternatives” is applicable only if the other criteria are “no”.

Table 2: results of the analysis of alternatives.

	Project:	H1	H2	H3	R1	R2	R3	R4	L1	L2	L3	L4	L5	L6	L7	L8
Alternatives and scenarios considered	Do nothing			yes	yes	yes		no	yes	yes	yes	no	no	no	no	no
	Base case	no	no	yes	no	yes		yes	yes	no	no	no	no	no	no	no
	One alternative only	yes	yes		yes	yes	no			yes						
	>2 projects	no	no	yes	no	no	yes	yes	yes	no	yes	yes	yes	yes	yes	yes
Extent of alternatives	modal alternatives	no	no	no	no	no	no	no	yes	no	no	no	no	no	no	no
	corridor alternatives	no	no	no	no	no	no	no	yes	no	no	no	yes	no	no	no
	path alternatives	no	no	yes	no	yes	yes	yes	yes	no	yes	yes	yes	yes	no	yes
	minor differences	no	no	yes	yes	no	yes	yes	no	yes	no	no	yes	no	yes	no
	no alternatives	yes	yes													
Level of description	description only	no	no		yes		yes	yes						yes		yes
	partial design	no	no	yes	no		no	no	yes	yes	yes	yes		no		no
	full design	no	no	no	no	yes	no	no	no	no	no	no	yes	no	yes	no
Use of the alternatives	simulation model	no	no	no		yes	yes	no	yes	no	no	no	no	no	no	no
	evaluation model	no	no	no	yes	yes	no	no	no	no	no	no	no	no	no	no

Source: elaboration of the author.

Alternatives and scenarios considered

A consistent evaluation requires the definition of a base scenario and some alternative projects. The literature is full of prescriptions for the building of a correct base scenario. Among others, EU guidelines (Florio, 2003) give the definition of a "do minimum" scenario (the less costly solution to the problem) together with a "do nothing" (to keep the present situation). For Eijgenraam (2000) it should be a combination of the best, alternative application of the available investment resources and the best possible alternative solution for the problem that we want to solve with the project. In general, national guidelines apply similar definitions for the do-minimum as reference scenario (for example: HM Treasury, 1997).

The first step of this survey then aims to verify the presence of a well defined base case or, at least, a do nothing scenario (it would be *wrong* to compare future alternatives with a reference solution that will never exist). On second thought, one must verify the existence of more than the chosen alternative only, in order to make significant its choice. The feasibility of a project, in fact, doesn't demonstrate *a priori* that it was the best in absolute terms and that some better unexplored alternatives didn't exists[8]. In certain cases the best solution could be the reference solution.

[8] The discard of socially viable alternatives can be accepted only in the case of technical limitations. This, however, is the only instance and the presence of more than one alternative nevertheless makes the selection more significant.

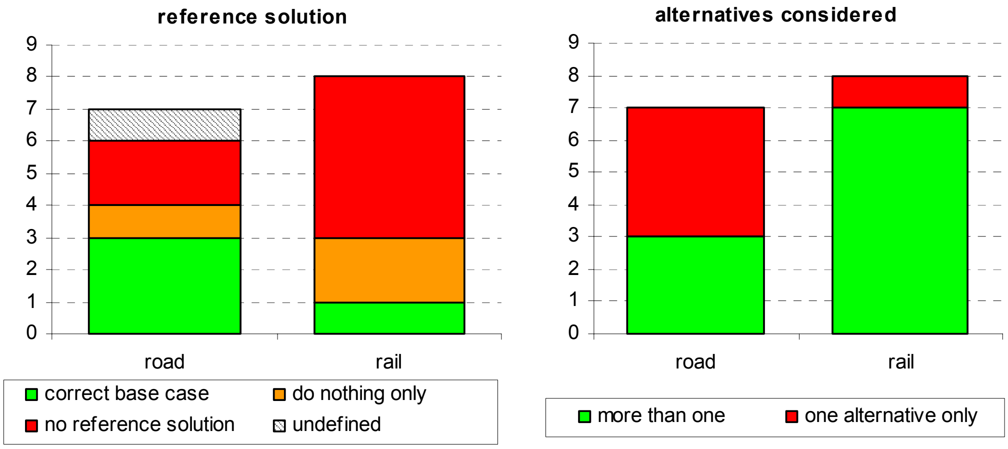


Figure 1: results, reference solution and number of considered alternatives.
Source: elaboration of the author.

The results of this analysis doesn't seem generally positive. Only 10 projects out of 15 include more than one alternative: the chosen project is not compared, but simply verified. Even more problematic is the absence of a base scenario in the majority of projects (present only for 4 out of 15). Rail projects tend to show more than one alternative, and are consequently better in this respect, but show shortcomings for the base case. Road projects show the opposite effect.

Extent of alternatives

Apart from the actual presence of alternatives, their extent is even more important. One can consider modal alternatives, corridor alternatives, path alternatives, minor detail differences or no alternatives. The majority of projects (10 out of 15) include only path alternatives or (7 out of 15) minor differences. Real corridor or modal alternatives are present only in 2 projects (Brenner Basis Tunnel and Terni-Spoleto line). Two projects consider no alternatives at all (A4 Milan-Turin and Roma-Formia). Rail projects have been better in these terms with respect to road ones.

Use of alternatives

A key point is the use of defined (and sometimes fully designed) alternatives. If such definition is irrelevant for demand forecasts and evaluation phase, the alternatives can be considered fictitious.

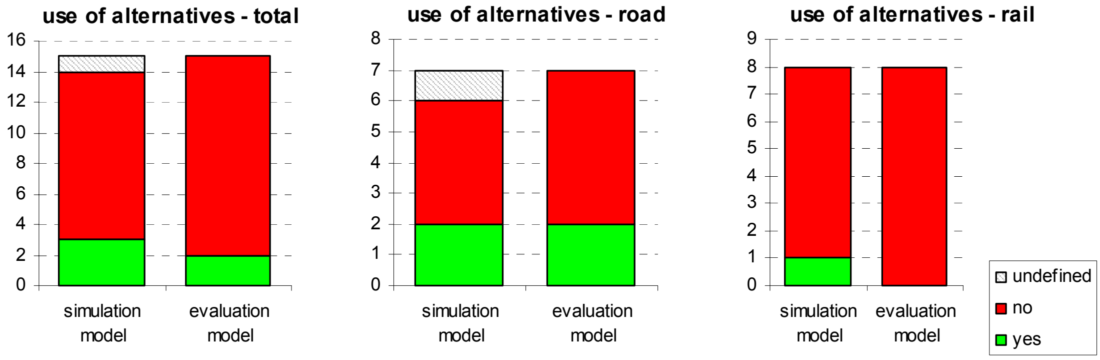


Figure 2: use done with defined alternatives.
Source: elaboration of the author.

The results are poor; the projects in which alternatives enter in the demand model or the evaluation phase are three and two, respectively, out of 15. One can affirm that the alternatives are usually formally presented in order to fulfil the EIA's requests, but these are irrelevant for the choice. Sometimes the choice of the path is done *a priori* using technical criteria only. The most complete study is the *Valtellina* road access, with alternatives used both in simulation and CBA. The only rail project is the Brenner Basis Tunnel.

Demand forecasts

The demand forecasts are key elements in any transport study. Any infrastructural improvement should be the consequence of demand, real or presumed, and therefore the nature of this demand must be justified. A demand study is therefore essential in demonstrating the necessity of the project, before any socio-economic feasibility consideration, showing its impacts on ecosystems and human beings.

The Italian law requires explicitly to define and report[9] the degree and level of demand fulfilment with and without the project, plus the quantitative forecast of the demand-supply ratio. The building of infrastructure with relatively small amounts of traffic is not in accordance with the Ministry of Environment guidelines (ANPA, 2001), since it causes unnecessary environmental damages.

The scheme used for analysis and comparison is formed by the following areas; the nature of the forecasts, the methods used for the forecasts and for flow assignment, the transparency and reproducibility of the procedure, the boundaries of the analysis, and some technical aspects of the models used. The full results are in following table[10], details come further.

[9] DPCM 27/12/1988, Art. 4, c. 2, lett. b) and c).

[10] In the table, "yes" means that the element is present or is correct, "no" that it is absent and such absence constitutes a conceptual error, "?" that the answer wasn't possible due to lack of information. The blank means that the issue was not present. For example, in the section "method used" it has been marked "yes" only for the method used.

Table 3: results of the analysis of demand forecasts.

	Project:	H1	H2	H3	R1	R2	R3	R4	L1	L2	L3	L4	L5	L6	L7	L8
Quality of the forecast	weak considerations	no	yes								yes	yes	yes	yes		yes
	qualitative evaluation	no	no								no	no	no	no		no
	quantitative evaluation	yes	?	yes	yes	yes	yes	yes	yes	yes	no	no	no	no	yes	no
	traffic surveys & data	yes	yes	no	yes	yes	yes	yes	yes	yes	no	no	no	no	no	no
Method of demand forecast	qualitative / "supply side"									yes	yes	yes	yes	yes	yes	yes
	use of trends	yes	yes	yes	yes	yes	yes	yes	yes	yes			yes		yes	
	link to GNP					yes	yes									
	use of elasticity															
	macroeconomic scenarios		yes	yes					yes							
Model used for flows estimation	assignment model	yes	no	yes	?	?	?	yes	yes	no	no	no	no	no	no	no
	integrated land use model	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no
Clarity and transparency	clear data and hypotheses	yes	yes	no	no	yes		yes	yes	no	no	no	no	no	no	no
	possible to reproduce	yes	yes	no	no	yes	no	yes	yes	yes	yes	yes	no	yes	no	no
Boundaries	time horizon correct	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes		yes	yes	no	yes
	space boundaries correct	yes	no	yes	no	yes	yes	yes	yes	yes	no		no	no	yes	no
Model	presence of sim model	yes	no		yes	yes	yes	yes	yes	yes	no	no	no	no	no	no
	only declared			yes	yes											
	calibrated	yes		yes	?	yes	yes	yes	yes	yes						
	validated	no		?	no	no	no	?	?	?						
The model is common with other projects?		no	no	no	no	?	no	no	yes	yes	no	no	no	no	no	no

Source: elaboration of the author.

Quality and nature of the forecast

Forecasts do not always follow a quantitative rational procedure, as the problem suggests, but sometimes are ruled by weaker qualitative considerations. Only for 9 projects the forecasts are treated in quantitative terms.

Another element that should be explicitly analysed and documented is the presence of traffic surveys and the proof of simulations. Traffic surveys are necessary to size the infrastructure, to be the basis of forecasts, but also, in theory, to calibrate and validate the models.

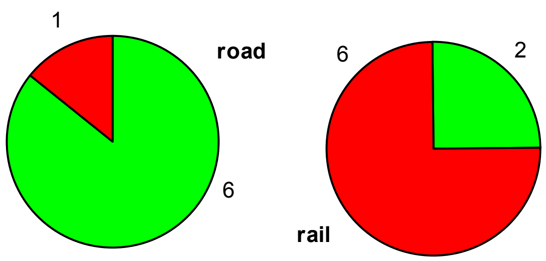


Figure 3: presence of traffic surveys in public documentation.
Source: elaboration of the author.

Traffic surveys are very common for road projects, but not for rail (only the two Brenner projects give indications about the actual and presumed traffic). At first sight this is counterintuitive, since the rail EIAs are designed by a sector of the rail operator itself: the operator should know, at least, the actual number of passengers/freight, but no data is provided apart from the number of trains.

Method used for forecasts

Many ways currently exist to produce forecasts. For simplicity's sake, only six are considered. These are not "methods", but the ways in which the designer approached the problem. More than one can be used at the same time:

- Qualitative. The future demand is quantified by qualitative assumptions only, for example making an hypothesis for the number of future trains and their load factors.
- Use of trends. The application of trends to actual traffic or demand is the simplest way to make forecasts. The trends can derive from various sources, of course to be always declared.
- Link to GNP. A common hypothesis is to link the increase of demand to GNP rate.
- Use of elasticity. That is to say the use of generalised cost elasticity in order to link the demand with the supply, where necessary according to some scenarios. Multi-modal models, if used, include an implicit elasticity.
- Macroeconomic scenarios. The definition of scenarios, more than one, can simplify the forecast allowing to make undemonstrated hypotheses. On the other side it is necessary the clear explanation of all the assumptions done and the use of the same scenarios until the end of the process.
- Interviews and stated preferences. Another way of forecasting future demand, frequently used in sectors such as Local Transport, consists in the use of interviews together with the technique of stated preferences.

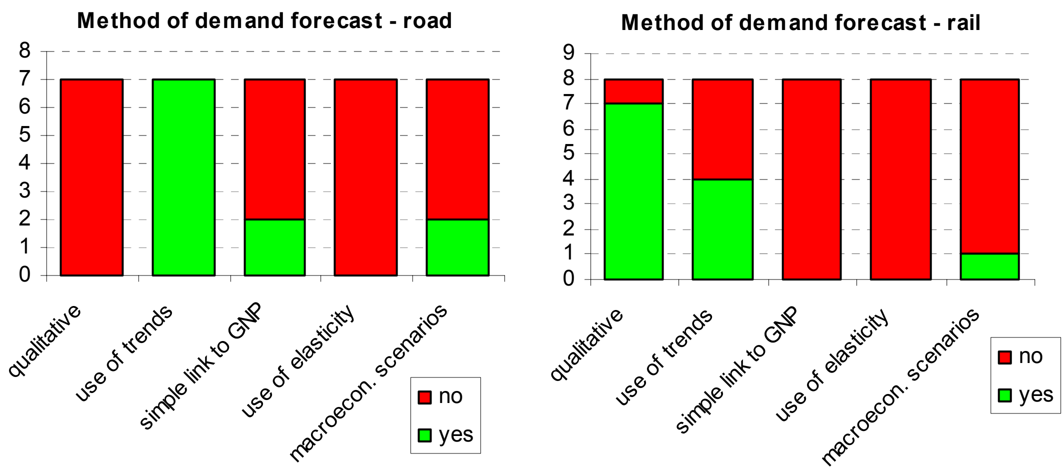


Figure 4: methods used for the forecast of the demand.
Source: elaboration of the author.

The survey shows clearly a predilection for trends in the road project, sometimes according to different scenarios (A4 Highway and "Pedemontana" Highway). Rail sector projects are done, apart from Brenner, by the same designer (RFI) and use the same approach to future demand estimation, described later in the example, and defined "supply side approach". The "method" can be classified as qualitative, and the quoted "commercial analysis"[11] is not provided. No one project used elasticity or other economically founded method, nor stated preferences.

Another issue is the presence of an assignment model, to distribute the predicted flows on the present and the future network. This approach, quite common in Local Transport or Road sectors, is rarely applied in the analysed projects, despite its importance (it is available only for the Roma-Formia, "Pedemontana" and Modena Sassuolo projects).

Other aspects

Concerning the other aspects, it is worth reporting the lack of clarity in data used – only for 5 projects out of 14 all data sets are clearly specified – with the possibility to reproduce the forecasts. This is more common, for example in road projects, but simply because of the large use of simple trends.

Only five road projects and two rail projects include a flow simulation model, at least declared, even if results are not always reported. For all the projects using models, some traffic data are available, so one can suppose that such data has been used for a calibration phase (even if clearly declared only in few cases). Finally, no proof of model validation is evident, since none of the designers who produced such a large model considered the importance of using a third party checking. Only two projects use a generic model, for example built in advance or prepared for general uses (Brenner Tunnel and Verona-Brenner line has a common model) rather than a project specific model.

Of this part some comments can be drawn. The study of demand is extremely complex and costly, especially for large projects such as those considered. The impacts are region-wide or nation-wide and the models to be used cannot usually be produced in a short time and for single projects. But, conversely, the analysed projects are extremely important and their total cost is very relevant The costs to produce an acceptable traffic model, given its importance, would be more than justified. Moreover, these projects often came from large agencies, like RFI – National Railways and ANAS – National Roads. These agencies, since one of their functions is to design infrastructural investments, should have among their operative tools one national scale model common for all projects and adequately set, calibrated and validated, as exercised by other agencies (the public transport companies of big cities, for example).

Approach to economic assessment

The analysed projects can be classified as "large", sometimes as "megaprojects", not only due to the relevant amount of money required (Warrack, 1993; Haynes, 2002), but also for the complexity of the decisional process. Nevertheless the "Objective Law" is characterised by the presence, at the same time, of mega, large and medium projects, despite the important differences among these categories. The socio-economic feasibility is clearly required by Italian laws concerning EIA (DPCM 27/12/1988),

[11] see the example section of this paper for details.

while the European original version does not (directive 85/337/EEC and amended by 97/11/EC).

Table 4: results of economic analysis survey.

	Project:	H1	H2	H3	R1	R2	R3	R4	L1	L2	L345	L6	L7	L8
Method used	CBA economic	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes		yes
	CBA financial													
	multicriteria analysis			yes							yes			yes
	impact analysis													
	no economic assessment								yes				yes	yes
Consistency	intertemporal discount	yes		yes	yes	yes	yes	yes	yes	yes	yes	yes		yes
with theory	shadow pricing	yes		yes	yes	yes	yes	yes	yes	yes	yes	yes		yes
	elasticity explicit*	no			no				no	no	no	no		no
	correct indicators (NPV, IRR)	yes		yes	yes	yes	yes	?	yes	yes	yes	yes		yes
	sensitivity analysis for SDR	yes	yes	yes	no	yes	?	yes	yes	?	no	no		yes
	full sensitivity analysis	?		yes	no	yes	?	yes	yes	yes	yes	yes		yes
	time horizon plausible	yes	?	yes	yes	yes	yes	yes	yes	yes	yes	yes		yes
	external costs	yes	no	yes	no	yes	yes	no	yes	yes	yes	yes		yes
	continuous in time	?		yes	yes	yes	yes	no	yes	yes	yes	yes		yes
Improvements	Risk Analysis	no	no	no	no	no	no	no	no	no	no	no		no
	Option Value	no	no	no	no	no	no	no	no	no	no	no		no
	MOCPF or similar	no	no	no	no	no	no	no	no	no	no	no		no
	model integrated CBA	no	no	no	no	no	no	no	no	no	no	no		no
Evident errors*	no double counting	yes		yes	yes	no	yes	yes	yes	yes	yes	yes		yes
	no computational errors	?		yes	?	?	yes	yes	yes	?	no	?		?
	no errors in gen. traffic B	no		yes	no	no	no	yes	no	no	no	no		no
	no omissions of C or B	yes	no	yes	no	yes	yes	no	yes	no	no	no		no
	consistent B calc (base-proj)	no		yes	no	yes	yes	yes	no	no	no	no		no
	consistent C calc (base-proj)	yes		yes	yes	yes	yes	yes	no	yes	yes	yes		yes
	real alternatives	no	no	no	no	no	no	no	no	no	no	no		no
Data used	no relevant errors in input	?		yes	yes	yes	no	yes	?	**	**	**		no
	macroeconomic input ok	?		yes	no	yes	yes	yes	yes	yes	yes	yes		yes
	correct shadow pricing	?		yes	yes	yes	yes	yes	yes	yes	yes	yes		yes
	correct external costs	?		yes	no	yes	yes	no	yes	no	no	no		?
	general coher. with literat.	?		yes	yes	yes	no	no	yes	yes	yes	yes		yes
Results	Positive VAN	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes		yes
	Pos. VAN for sensitivity too	?	no	?		?	?	yes	no	yes	no	?		?

*: excluding data errors and models
**: the values used are EU average instead of available Italian values.
Source: elaboration of the author.

In this section the number of assessments reported is 12, plus one with no analysis. This is due to the fact that the Orte – Falconara rail corridor (Foligno – Fabriano section doubling, Falconara rail node and Spoleto – Terni section) has three EIAs, one for each section, but only one comprehensive CBA. The most expensive is the Brenner Basis Tunnel (4.500 M€) followed by Verona-Padova HST (more than 2.500 M€) and

Fologno-Fabriano rail (less than 2.000 M€). Other projects cost 500 M€ or less; for some of them the figures are not available.

The approach used for the survey and comparison among considered projects is structured into seven main topics: which method has been used, the consistency with the theory of CBA, the presence of less conventional and codified "improvements", the most evident theoretical errors and the recognised data errors. At the end, a check has been done on the transparency and the results of the assessments. The previous table contain the results for all cases[12].

Method used

The majority of projects are assessed using a CBA. Only some of them include also a multicriteria analysis to choose some minor issues (Pedemontana, Orte-Falconara, Verona-Padova). For three projects the EIA does not include any socio-economic method for alternatives choice, apart from technical or environmental aspects.

Excluding the multicriteria analyses, from this point on the focus is on the CBAs.

Consistency with theory of CBA

The procedure of CBA is well established in literature, at least regarding its fundamental aspects. Many guidelines are available (Florio, 2003 for EC structural funds; CERTU, 2004 for France; NUVV, 2001 for Italy; COBA manual for UK, Eijgenraam, 2000 in The Netherlands, etc.). Among the possible schemes, this study refers mainly to the one proposed by the European Commission for the access to structural funds (Florio et al., 2003), but these elements are widely accepted. The survey considers the presence of the following issues:

- intertemporal discount. The use of a social discount rate to discount future costs and benefits (Florio, 2003, section 2.5.4).
- shadow pricing. The use of coefficients to correct distortions in prices (Florio, 2003, section 2.5.1).
- elasticity explicit or demand curve. The elasticity or the demand curve should be clearly defined and used to calculate the surplus of new customers (Florio, 2003, box 3).
- correct indicators (NPV, IRR) (Florio, 2003, section 2.5.5).
- sensitivity analysis for SDR. The presence of the sensitivity for the social discount rate.
- full sensitivity analysis. The presence of the sensitivity analysis for other aspects (Florio, 2003, section 3.3.7).
- time horizon plausible. The temporal extent of the analysis, depending on the characteristics of the projects (Florio, 2003, section 3.3).
- external costs. The presence of environmental and non-environmental externalities (Florio, 2003, section 2.5.2).

[12] In the table, "yes" means that the element is present or is correct, "no" that it is absent and such absence constitutes a conceptual error, "?" that the answer wasn't possible due to lack of information. The blank means that the that the issue was not present. For example, in the section "method used" it has been marked "yes" only for the method used.

· continuous in time. If the transitory periods (building phases, etc.) are, correctly, considered.

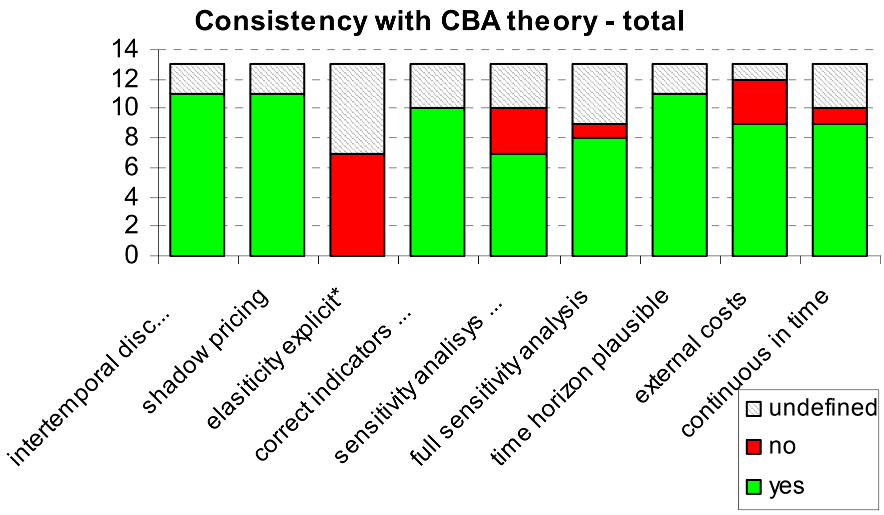


Figure 5: presence and correctness of theoretical aspects of CBA.
Source: elaboration of the author.

The results are positive. The theoretical foundations are generally respected, even if no project reports the elasticity used for generated traffic and sometimes external costs are neglected (always the case for road projects).

Improvements

Some improvements have been proposed by theory in order to correct some uncovered or failing aspects of standard CBA. These are still not very used in the common practice. The main focuses are the Risk Analysis, the Option Value, the Marginal Opportunity Cost of Public Funds, the use of model integrated CBAs. None of the analysed projects consider these aspects.

Theoretical errors

Some errors have been identified. These are quite general and represent simply the categories of errors found in the survey.

· Double counting. The presence of benefits (or even costs) calculated twice into two different forms. For example: the gains in real estate values due to a transport improvement together with the time gains for the inhabitants.
· Computational errors. In general, all the calculus errors revised.
· Errors in generated traffic benefits. The benefits of generated traffic must be calculated as surplus gain using a quantity(cost) diagram and hypothesising a demand curve. The lack of these elements suggest a wrong benefit, usually calculated as the difference between initial and final costs and multiplied by all the final users. In this case the benefit is overestimated.

- Omissions of some costs or benefits. The lack of some standard and commonly introduced costs or benefits. For example: the lack in the analysis of external environmental costs.
- Consistent benefit calculus (base-project). This refers to the correctness of the base and project scenarios to calculate benefits. For example: it is wrong to calculate benefits as the difference between the project scenario and the year-zero situation, instead of the base case projection to the analysed year.
- Consistent Cost calculus (base-project). See before.
- Use of real alternatives. This is the most common error: the comparison is made of one alternative only or of irrelevant alternatives.

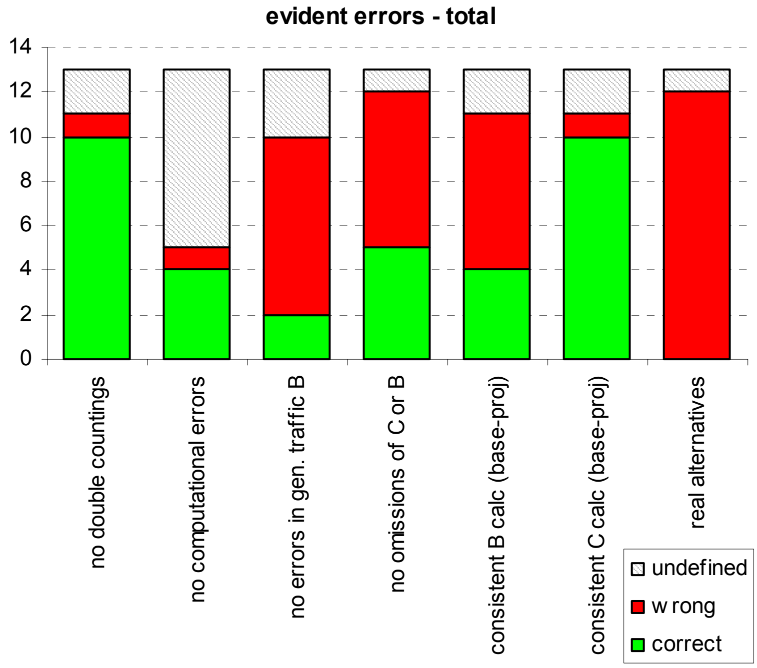


Figure 6: review of the most evident errors revised in CBAs.
Source: elaboration of the author.

The evaluation of these errors is extremely difficult: the documents provided were often very short and incomplete and the reproduction of the analyses to check errors is sometimes impossible. Although only for a small number of projects the verification is well grounded, some double counting has been found in one project only (*Valtellina* accessibility), such as some minor computational errors (Orte-Falconara rail). The calculation of the costs is generally correct, too. More problems raise in the generated traffic determination, the frequent omission of relevant costs or benefits, and the wrong determination of benefits. For all 12 projects the presence of real alternatives (also where correctly defined in the design part) is, in the author's opinion, absent.

Data errors

Evaluating this area can be as arduous as those detailed in the previous section. For the sake of simplicity, the answer has been set as positive by default, apart from evident and demonstrable biases. All rail projects used a public source for external costs, but reported general values even if the specific Italian values were available (see next sections for details). The introduction of errors in data can obviously be accidental, yet despite this, all data used should be justified or quoted. Nevertheless this sometimes does not happen.

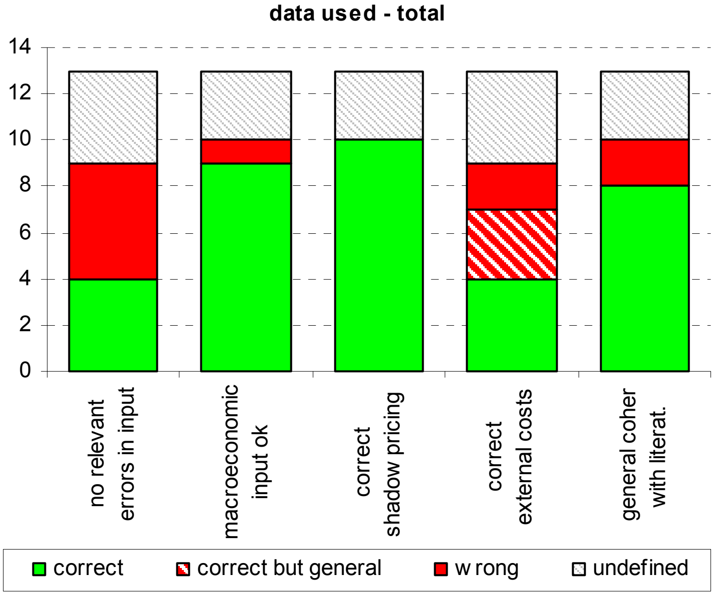


Figure 7: presence of errors in the data used.
Source: elaboration of the author.

Clarity and transparency

The data sets used are generally lacking and unclear (only 4 out of 12 projects were clearly explained). The reproducibility of the analyses is generally possible, but this is due to the simplicity of the procedures.

Results

All the analysed projects are positive according to the produced CBA. For 3 projects the sensitivity analysis points out negative results for some aspects. However, 2 projects are always positive (Modena-Sassuolo link and Verona-Brenner doubling), whatever are the proposed variation of parameters.

Some useful concluding remarks can be made. The CBAs show a good consistency with theory, including almost all the parts of the shared algorithm. The problems appear in the operative parts and in the transparency. Some unacceptable and even evident errors have been found, usually in favour of the project. The alternatives are always irrelevant or absent. Data errors or unclear sources have been revised. The fact that all the projects are positive is, in one aspect, obvious (the alternative is always one), but in another also worrying.

An example of Cost Benefit Analysis: the RFI standard procedure.

A good example of standardised procedure to Cost Benefit Analysis for infrastructural projects is the one used by RFI[13]. All the rail projects analysed can be classified as "large" projects, involving an average investment of 1.670 M€ (210 M€ of investment for the cheapest project, the Falconara node, and 4.500 M€ for the most expensive, the Brenner Basis Tunnel). For all these large, costly and impacting projects, the only and official document describing CBA is usually about 30 pages long, including the information concerning the demand forecast. An article published in the technical journal of RFI, describes briefly both the theory and an example of such an approach (Cicini et al., 2005). The case study used in the article and described in the following pages is one of the projects previously analysed, the line Orte – Falconara in central Italy. The official documents analysed for this study are exactly the same as the article, apart for some data, sometimes different in the EIA[14].

The procedure used for all the projects is very simple and procedural. The structure is summarised in the Table 5.

Table 5: structure of CBA in Cicini et al. (2005).

costs	- benefits
▪ incremental costs of investments	▪ reduction of road passengers transport costs
▪ incremental costs of infrastructure exercise	▪ time savings for actual demand
▪ incremental costs of train exercise	▪ lower external costs

Source: Cicini et al. (2005).

Despite the simplicity, which may seem to be excessive for projects of such importance, some aspects of the procedure raise relevant doubts:

- the study states that the new competitiveness of the rail mode comes from the "removal of capacity constraints", even if these constraints are never demonstrated nor analysed.
- the demand for the infrastructure is never calculated using a model, but "comes from a commercial analysis of the transport operator about the slots it would be interested to buy" (Cicini et al., 2005, page 10). Such analysis is never included in the report nor publicly available, but simply quoted. Moreover, the transport operator is run by the same owner of RFI and this statement is simply a declaration. An independent simulation would be more convincing.

  the quantification of passengers and freights traffic [...] has been done using a "supply side" approach (sic)*; commercial analyses has been carried out analysing the actual transport demand and defining a new commercially sustainable supply, based on the exercise regime hypothesised after the realisation of projects.*[15]

  In practical terms, the method consists in inducting the CBA with a future arbitrary supply and using the same load factors.

[13] RFI is the society, part of the national railways *Ferrovie dello Stato S.p.A.*, owner of the infrastructures and responsible for maintenance and circulation.

[14] details are given later in this section.

[15] Cost Benefit Analysis of Orte-Falconara rail corridor. The same sentences with minimal variations can be found in every RFI CBA. The same CBA can be read as case study on Cicini et al. (2005).

- the general consistency with other projects is doubtful. Projects are analysed singularly, no comprehensive model for all the national projects exists[16].
- the opportunity cost of public funds is quoted, but not included in the analysis (*cit*, page 12).
- the necessity of realising the scheme completely in order to have the benefits is given as a hypothesis, even if not demonstrated nor realistic. Partial improvements (like selective doublings) can generally give disproportionately good benefits at lower costs (*cit*, page 18).
- part of the surplus generated by new traffic is not considered (the time and money savings due to modal change or the surplus generated by a new displacement).

Moreover, some conceptual errors can be raised:

- benefits are calculated as the difference between the "do-nothing" scenario at year 0 and the one with it at year *n*. A correct approach would consider the difference between the demand at year *n* without the project ("base scenario") and with it (Florio, 2003).
- no relevant alternatives (i.e. use or improvement of parallel lines) are considered.
- the demand is calculated as purely dependent from supply, using average (and unspecified) load factors. The "commercially sustainable supply" of X trains per day, is multiplied by an average (national?) load factor. This implies an unspecified hypothesis of completely frequency-elastic demand, which is clearly false, especially for rail mode.
- the amount of passenger·km, later used for the determination of all the marginal benefits and costs, is calculated as follows; the average load factor is multiplied by all the train·km produced, as if the train were completely full for the whole journey. The (positive) effect is illustrated in Table 6: the "real" pkm of the example should be 30.000, while the document would report 40.000 pkm. The effect is that the longer the trains' journeys, the better the CBA result, even if in reality the trains could be even empty.

Table 6: example of pkm calculation in RFI procedure.

stations	*A*		*B*		*C*
Distance		100 km		100 km	
real load factor		100 pax		200 pax	
real pkm		10.000 pkm		20.000 pkm	
assumed load factor[17]		200 pax		200 pax	
assumed pkm		20.000 pkm		20.000 pkm	

Source: elaboration of the author.

Finally, a couple of problems and doubts can be voiced about the data used:

- part of the external costs used, taken from international literature (INFRAS, 2000), are biased. The source reports both the European average and the country specific data. The CBA of article (Cicini et al, 2005) uses the European data, introducing a bias that is in favour of the train system. The data used in the official

[16] Some projects can be considered as alternative to each other: for example the access to alpine passes (Simplon, Frejus and Gotthard) are competing for freights.
[17] The load factors used are extremely high, seeming too similar to peak load factor. See further.

CBA for freight cannot be found in the quoted source. Curiously, with such varying data, the final results of the article and official CBA are identical.

Table 7: External costs used by RFI in the two documents, compared with the quoted source.

	article	Passengers *official CBA*	INFRAS (Italy)	article	Freight *official CBA*	INFRAS (Italy)
Road	0.087	0.087	0.078	0.072	0.088	0.072
Rail	0.020	0.020	0.020	0.019	0.004	0.026

Source: Cicini et al. (2005), EIA official documentation, INFRAS (2000).

- the load factors, although extremely important for the estimation of costs and benefits, are never declared. In the article, the load factor for regional trains is 486 passengers/train and for long distance is 375 passengers/train. The values are not explicitly given among the inputs, but they can be simply calculated from the data. It is possible to compare them with the official data published on the operator website (www.trenitalia.it, visited 07/02/2007). It provides the average number of trains and passengers per day for the regional services of Umbria and Marche. The average regional values (including lines both more and less important) calculated are 189 pax/train and 134 pax/train respectively[18]. The bias introduced with respect to the average regional values is considerable.

Conclusion

In the paper the framework of the comparative analysis has been defined. A survey of some projects has also been done. Some numerical information has been derived, demonstrating quite clearly the poor quality of produced analyses, given the inadequacy of the EIA tool for the socio-economic evaluation.

The survey pointed out some general patterns concerning the analysed aspects.

- The alternatives, if present, are definitely fictitious, since they aren't incorporated into models and/or in the economic evaluation. The alternatives are generally limited to path variants, no modal or corridor alternatives are considered.
- The base scenarios are seldom correct, with respect to literature and common sense.
- Forecast methodologies are extremely simple, demand and supply relationships appear unexplored, trivial extrapolation of historic data sets prevails. Rail project forecast methodology may be seen as inconsistent.
- Only the road projects provide sufficient traffic data and surveys and analyse it by assignment models. Rail projects ignore actual traffic.
- Lack of use of generic nation-wide models inside the larger agencies.
- Economic Cost Benefit Analysis is the method generally adopted for social feasibility.

[18] *Umbria*: 21000 trips per day using 111 trains. *Marche*: 22256 trips per day using 166 trains. Of course these values are indicative, but gives an indication quite precise of plausible load factors.

- The basic theoretical assumptions of the model are respected, apart from the definition of a demand curve and the consequent welfare gain calculation.
- Many conceptual errors have been found. These errors, generally speaking, introduce a bias in favour of the selected project.
- Some data errors or unjustified values have been found.

The EIA in its original form does not aim to select the most welfare enhancing projects, since the criteria used are different than the ones suggested in this paper. Nevertheless, since no public economic ex-ante evaluation of alternatives is available and the EIA is structured *as if* it were an ex-ante assessment, it would be essential that it is carried out according to a unanimously accepted, transparent and well founded method. An alternative would be to limit the extent of the EIA to only environmental impacts and moving the CBA to a former phase, where the alternatives are still present and different choices remain possible.

References

ANPA (2001) *Linee guida V.I.A. - Parte Generale*, ANPA, Rome (Italy).

Baccaro, G., Belvisi, M. and Sini, V. (2005) *Dispositivi legislativi internazionali, comunitari e nazionali in materia di VIA*, APAT, Roma (Italy).

Berechman, J. and Paaswell, R. E. (2005) "Evaluation, prioritization and selection of transportation investment projects in New York City", *Transportation*, Vol. 32, No. 3, pp. 223-249.

Borgnolo, C. (2004),*Institutional analysis. Italy*, in ECORYS et al. (2004), *Monitoring the development of national transport policies and their impact on Community transport objectives*, Rotterdam (NL).

Brambilla, M. (2003), *Analisi costi – benefici del progetto del Ponte sullo Stretto di Messina*, proceedings of the conference "Legge Obiettivo e Valutazione dei Progetti", 16th Semptember 2003, Università Cattolica, Milano (Italy).

Brambilla, M., Erba, S. and Ponti, M. (2003) "La pianificazione e la valutazione delle infrastrutture in Italia: dal piano nazionale dei trasporti alla Legge Obiettivo", proceedings from *VI SIET conference*, Palermo (Italy).

Cicini, P., Macchia, G. and Giuseppetti, A. (2005) "Progetti ferroviari: dall'analisi finanziaria all'analisi economica", *Argomenti*, Year 3, 5: 6-37, RFI, Rome (Italy).

CERTU (2004) *Evaluation des projets d'exploitation sur les axes principaux d'agglomeration*, Certu, Lyon (France).

EIB (2005), *RAILPAG guidelines*.

Eijgenraam, C. J. J. et al. (2000) *Evaluation of infrastructural projects; guide for cost-*benefit *analysis*, CPB Netherlands Bureau for Economic Policy Analysis and Netherlands Economic Institute, Rotterdam (The Netherlands).

European Commission (1997) *Update of 5-years EIA Report*, European Commission. http://ec.europa.eu/environment/eia/eia-support.htm.

Florio, M. (2001) *La valutazione degli investimenti pubblici. I progetti di sviluppo nell'Unione Europea e nell'esperienza internazionale – volume primo*, Franco Angeli, Milano (Italy).

Florio, M. et al. (2003) *Guide to cost- benefit analysis of investment projects*, European Commission.

Flyvbjerg, B. et al. (2002) "Underestimating costs in public projects. Error or lie?", *APA Journal*, Vol. 68, No. 3, Chicago (Illinois - USA).

Flyvbjerg, B. et al. (2003) *Megaprojects and risk. An anatomy of ambition*, Cambridge U.P., Cambridge (Great Britain).

Flyvbjerg, B. et al. (2004) "What Causes Cost Overrun in Transport Infrastructure Projects?", *Transport Reviews*, Vol. 24, No. 1, pp. 3–18.

Haynes, W. (2002) "Transportation at the millennium: In search of a megaproject lens", *The Review of Policy Research*, Vol. 19, No. 2.

HM Treasury (1997) *The Green Book. Appraisal and evaluation in central government*, HM Treasury, London (UK).

Maibach, M. et al. (2000) *External costs of transport*, INFRAS/IWW, Zürich/Karlsruhe.

Næss, P. (2006) "Cost-benefit analyses of transportation investments: neither critical nor realistic", *Journal of Critical Realism*, Vol. 5, No. 1, London (Great Britain).

NUVV (2001) *Studi di fattibilità delle opere pubbliche. Guida per la certificazione da parte dei Nuclei regionali di valutazione e verifica degli investimenti pubblici (NUVV), Conferenza dei presidenti delle Regioni e delle Province Autonome*, Roma (Italy).

Pearce, D., Atkinson, G. and Mourato, S. (2006) *Cost-Benefit Analysis and the Environment. Recent developments*, OECD.

Cittadini, M. R. (2005a) *La linea AV Milano-Genova e il problema del Terzo Valico dei Giovi*. Taken from www.ricerchetrasporti.com.

Cittadini, M. R. (2005b) *Il nuovo collegamento ferroviario Torino-Lione*. Taken from www.ricerchetrasporti.com.

Warrack, A. (1993) "Megaproject decision making. Lessons and strategies", *Western Centre for Economic Research Information Bulletin*, No. 16.

Zambrini, M. (2003) *Osservazioni sullo studio di impatto ambientale del Ponte di Messina*, Istituto di Ricerche Ambiente Italia.

Zambrini, M. (2004) *La costruzione della rete AV/AC. Dalla finanza di progetto alla finanza creativa*, Istituto di Ricerche Ambiente Italia.

Acknowledgements

I wish to thank Stefano Erba and Marco Brambilla for the preliminary work they carried on in analysing with me the *mountains* of documents, papers, files, cds, used for this survey! Many thanks also to Jonathan Ongley, for the complete revision of this paper.

European Transport \ Trasporti Europei n. 36 (2007): 47-74

Accessibility and environmental quality: inequality in the Paris housing market

André de Palma [1*], Kiarash Motamedi [2], Nathalie Picard [2], Paul Waddell [3]

[1] *University of Cergy-Pontoise and Ecole Nationale de Ponts et Chaussées, France.*
[2] *University of Cergy-Pontoise, France.*
[3] *University of Washington, USA.*

Abstract

In this paper we examine empirically the market for local amenities in the Paris metropolitan region. We find first that there is considerable inequality in the spatial distribution of these local amenities, including accessibility, environmental and social indicators. We use a spatial representation and Lorenz curves to examine the degree of inequality in these amenities, and this provides evidence that some amenities (or disamenities) are much more inequitably distributed than others. The most extremely unequally distributed amenities are noise (due to its concentration near airports), "Redevelopment Areas", presence of water (lakes and rivers) and forests, and presence of train and subway stations. Some indicators, such as the "Poulit accessibility" measure, were by contrast remarkably constant over the region. We recognize that local amenities should be capitalized into the housing market, and explore the willingness to pay of households for these amenities within the Paris region using alternative specifications of a location choice model. One of the core questions we examine is the spatial scale of the amenity effects and how this is captured in a location choice context. By estimating models at both a commune and at a grid cell level, we obtain new insights into how households in the Paris region trade off amenities against each other and against housing cost. We find that the residential location choice model fits the data moderately better at the smaller scale of the grid cell compared to the commune.

Keywords: Inequality; Efficiency; Local public goods; Residential location; Integrated model; Transportation modelling; Paris area.

1. Introduction

Integrated land use and transportation models have received increased attention in research and practice over the past decade, principally based on their ability to examine the combined effects of land use and transportation policies on the endogenous system of urban development and patterns of travel, and their ability to represent the long-term induced travel effects of expansion in transportation capacity (Waddell, Ulfarsson,

---

* Corresponding author: André de Palma (andre.depalma@u-cergy.fr).

Franklin and Lobb, 2007). The utility of integrated models goes well beyond the examination of induced travel demand effects, however. There is considerable potential to use these tools to examine questions such as efficiency and equity in markets for local public goods (and local public bads). In addition, there has been growing interest in using integrated models to examine the spatial patterns and distributional effects of environmental externalities such as vehicle-related emissions (Wegener, 2004). In this paper, we undertake an analysis of *inequality* in the spatial distribution of accessibility and environmental quality in the Paris metropolitan region, using an integrated land use and transport model system, and examine how these (dis)amenities are capitalized into the housing market.

Our approach integrates recent work in urban land use modelling (see, e.g. Waddell, Borning, Noth, Freier, Becke and Ulfarsson, 2003 and Waddell, Ulfarsson, Franklin and Lobb, 2007) with recent advances in dynamic transportation modelling (see, e.g. de Palma, Lindsey and Kilani, 2005) to provide an integrated platform for land use and transportation policy analysis. UrbanSim is a land use model system that simulates the location choices of households and firms, the real estate development choices of developers and real estate prices resulting from the interactions of these agents in urban real estate markets (see Waddell, Ulfarsson, Franklin and Lobb, 2007 for details). A range of urban amenities and disamenities, including accessibility, land use, and social composition are incorporated in the model system through the utilities of location choice and as determinants of dwelling prices. UrbanSim is a microsimulation model system that simulates individual households and firms (or jobs), and models the changes in the composition of the population of households and firms, and their locations, on an annual basis. The aggregate targets for population and employment are defined for the region as a set of constraints, or control totals. Models representing relocation, location choice, real estate development and prices are connected via their operations on a shared database, which is updated from one simulation year to the next. Accessibilities are measured by interfacing UrbanSim with a travel model, and combining results of traffic assignment with the spatial patterns of activity location predicted by UrbanSim.

METROPOLIS simulates the dynamic traffic patterns over the course of a day, based on the spatial distribution of households and jobs and the configuration of the transportation network (see de Palma, Marchal and Nesterov, 1997 for details). Transportation demand is represented at the microscopic level allowing representation of population heterogeneity and providing detailed information about travellers' costs. Transportation supply is represented by a macroscopic model of road traffic that computes travel times on links based in this paper on a simple congestion function. The information transferred from METROPOLIS to UrbanSim is a matrix that gives the individual travellers' surplus in each Origin–Destination pair. In UrbanSim, the accessibility measures from METROPOLIS influence residential and business location choices, real estate development and real estate prices. UrbanSim then simulates these location choices, which are in turn used to update an O-D matrix that is then used in the next iteration of METROPOLIS. This cycle is repeated at time steps that could be as small as one year (the time step for UrbanSim), to reflect the joint evolution of the transportation system conditions and urban development patterns. More details are provided in de Palma, Motamedi, Picard and Waddell, 2005.

By linking these two model systems, it is possible to examine in a more comprehensive way than has been possible previously the implication of a range of urban externalities on travel, land use, and environmental outcomes. Our intent is to apply the integrated model system to an analysis that examines inequality and efficiency in the markets for local public goods. We concentrate on two types of externalities: accessibility and environmental quality (or to be more precise, its degradation by pollution). Note that in order for a public good to have the potential to influence intra-metropolitan location choice, prices and urban development, it should meet two necessary conditions: (1) it is unevenly distributed over the region; (2) it is valued by households.

In the next section, we present a brief review of economics literature concerning urban externalities and inequality in urban areas. The third section describes the region under study and available data, and discusses spatial disparities and population inequalities in the region. We propose in this paper a simple measure of inequality based on Lorenz curves. In the fourth section, we examine through a discrete choice model of residential location how households in the region trade-off these local public goods against housing cost. This provides direct evidence on their willingness to pay for these local public goods. Estimation results are reported and discussed in the fifth section. The sixth section concludes this paper.

2. Externalities and urban form

In many situations, the cost associated with the action of an agent is not equal to the cost born by the society (social cost). If the cost born by an individual is smaller than the social cost, externalities are negative (pollution, noise and accident provide the easiest examples). On the other hand, if the social cost is smaller than the individual cost (or if the social benefit is larger than the individual benefit), externalities are positive (green spaces provide by some individuals provide an example of such positive externalities). Externalities are endemic in urban life, and motivate much of urban public policy and sometimes, pricing or quantity regulations. Externalities represent a market failure, in the sense that the market does not produce an efficient outcome in the presence of externalities. They justify, according to Pigou (1920), government intervention. Pigouvian taxes or subsidies introduce prices that internalise the externalities (for example, road pricing can address congestion externalities)[1]. Such taxes can be positive (as in private transportation, to take into account congestion and pollution) or negative (in this case, it is a subsidy, as in the education sector).

Coase (1960) later argued that the market could internalize such externalities by means of private contracts between affected parties. Coase's Theorem can be summarized as follows: if transaction costs among private persons are nil and there are no income effects, the externalities may lead to a mutually beneficial agreement between the parties without any government intervention. However, as Coase knew, transaction costs are seldom zero, and other solutions are required. In particular, the assumption of zero transaction cost is not acceptable in the context of urban and regional development.

We are concerned in this paper with the examination of externalities in the urban context. Pollution caused by industrial plants in the early 20th century precipitated the

[1] See also the early contribution of Knight, 1924. The Pigou-Knight controversy is described in detail in Pahaut and Sikow, 2006.

development of zoning policy in the United States, and led to the widespread use of land use regulation to separate land uses that were considered incompatible. Traffic congestion and air pollution are notorious externalities arising from private travel choices. Massive public investments attempt to reduce congestion, generally with limited success due to the underlying incentives to over-consume transportation arising from its under-pricing (the exclusion of external costs from the calculations of the individual traveller). Of course, the situation in the US and in Europe is very different with respect to transportation costs: the price of gasoline is probably too low is the US and too high in Europe (see the details on the computation of the optimal price of gasoline in Small and Parry, 2005).

Although urban externalities can be positive, as in the case of knowledge spillovers, or technological externalities, and other interactions that lead to agglomeration economies and the rise of employment centers (see for example, Krugman, 1991), most literature addressing urban externalities focuses on the negative effects of certain land uses such as polluting industries, of low-density urban form generally referred to as sprawl, and of travel patterns dominated by extensive use of single-occupancy vehicles. A variety of approaches to determine the nature and magnitude of these externalities has been developed, but the most common approach is the use of hedonic regression to estimate the effect of the externalities on housing prices. Since locational amenities and disamenities are likely to be capitalized into the price of housing due to the locational fixity of the housing stock and its durability, house prices provide a convenient measure of the degree to which housing consumers value the presence of a particular externality. Many studies, numbering at least in hundreds, have attempted to measure in different localities the housing price impacts of urban externalities, especially with respect to environmental degradation. Note that this approach is represented as capturing both point-source and broader non point source effects from water, air and noise pollution. Early research along this line was reviewed by Brown and Li, 1980. For a recent review of studies of the housing price effects of air quality, water quality, undesirable land uses, neighbourhood effects, and multiple environmental effects, see Boyle and Kiel, 2001. To our knowledge, the housing price effects of traffic congestion have not yet been systematically explored, though there is considerable research on the capitalization of various measures of accessibility in housing prices.

Much has also been written about the externalities associated with urban sprawl (see, for example, Verhoef and Nijkamp, 2003, Koland, 2006, and Chan, 2004), a term used to describe low-density, auto-oriented development. A systematic attempt to estimate the costs of sprawl was undertaken by Burchell et al., 1998, though the problems associated with measuring sprawl have not been easy to overcome. Brueckner (2000) argued that urban sprawl represents three related forms of market failure: “These are the failure to account for the benefits of open space, excessive commuting because of a failure to account for the social costs of congestion, and failure to make new development pay for the infrastructure costs it generates.” This assessment led Brueckner to advocate for development impact fees and for congestion pricing as means to internalise the externalities related to urban sprawl.

Aside from the efforts to measure the effects of externalities on housing prices noted above, there appears to be a dearth of research examining broader influences of externalities in urban areas on outcomes such as residential location or business location or on real estate development. Nor is there any significant body of literature that provides an analytic framework for addressing transportation and land use interdependencies in the context of urban externalities. We suggest that this is an important gap in the literature, and that an integrated land use and transportation analytic framework is useful to better inform public policies intended to address these urban externalities.

Beyond the political implications of inequality, we anticipate that inequality in the spatial distribution of amenities or disamenities produced by local externalities impact households differentially, and influence their location preferences. Social externalities such as those produced by the social composition of neighbourhoods interact with the social preferences of households making location choices to influence patterns of social clustering and segregation. We explore these and other interactions between household preferences and local externalities in Section 5.

3. Descriptive analysis of the study area

The Ile de France is the capital region of France. It is a large metropolitan region of over 11 million inhabitants, with the city of Paris at its core (See Appendix Fig. 8). There are two spatially-nested levels of administrative units we will refer to in this paper: (1) 1300 "*Communes*" corresponding to municipalities outside Paris, and to "Arrondissements" or large administrative neighbourhoods with local governance within Paris; and (2) 8 *districts* that are administrative units overlaying communes. In order to provide a context for our analysis of the inequality of externalities in this region, we present below indicators of accessibility and environmental quality for the communes. An additional unit of analysis we use in this paper is a grid of 500 meter resolution, resulting in approximately 50,000 cells in the region (see appendix figure 8 where dots represent cells[2]). We compare below the variability of the different local attributes at the three geographical levels, and comment on the consequences of these inequalities.

Table 1 presents accessibility and environmental variables and indicates their average values in each district. The cells' attributes are weighted by the number of households in the cell in order to compute these averages. Table 1 also indicates average values of other variables, which are defined only at the commune level. These averages are also weighted by the number of households in the commune.

Table 2 shows the variance decomposition at the three relevant geographical levels: district, commune and grid cell (and therefore only for which are available at grid cell level). For each variable X, the model is:

$$X_{dcg} = \mu + \eta_d + \xi_c + \varepsilon_g, \tag{1}$$

where, according to the analysis of variance model, X_{dcg} denotes the value of X in grid cell *g* located in commune *c* and in district *d*, μ is the average level of X in Ile-de-France, η_d, ξ_c and ε_g are independent random Gaussian variables with zero mean and variances $\sigma_d{}^2, \sigma_c{}^2$ and $\sigma_g{}^2$, respectively. The total variance of X is denoted by

[2] See also appendix figures 10 and 11 where square represents a cell with its 8 neighbouring cells.

$\sigma^2 \equiv \sigma_d{}^2 + \sigma_c{}^2 + \sigma_g{}^2$, and the fraction of variance (reported in Table 2) at level *i*, *i=d, c, g*, is: $\alpha_i = 100 \bullet \sigma_i{}^2 / \sigma^2$, with $\alpha_d + \alpha_c + \alpha_g = 100$.

Table 1: Average accessibility and environmental variables, by District (over grid cells and population-weighted).

Variable \ District	*75*	*92*	*94*	*93*	*78*	*91*	*95*	*77*
Number of Grid Cells	420	704	980	952	9400	7399	5187	24194
Number of Communes	20	36	47	40	262	196	185	514
Distance and accessibility								
Subway and tramway stations	8.17	0.80	0.65	0.72	0.00	0.00	0.00	0.00
Subway and tramway stations around	71.93	7.14	5.77	6.49	0.00	0.00	0.00	0.00
Train stations	0.98	0.75	0.44	0.53	0.53	0.47	0.68	0.30
Train stations around	9.35	6.89	3.94	4.76	4.29	4.07	5.60	2.35
Accessibility to employment (Public Transit)	50.93	49.29	48.02	48.25	45.20	44.30	45.97	41.86
Accessibility to employment (Private Car)	54.45	53.88	53.42	53.48	51.18	51.15	51.82	48.60
*Accessibility to employment (M)**	*-72.9*	*-72.3*	*-73.9*	*-74.0*	*-73.8*	*-93.3*	*-68.8*	*-89.9*
Average travel time (Public Transit)	*28.08*	*31.54*	*38.45*	*38.54*	*47.84*	*52.04*	*46.24*	*48.86*
Average travel time (Private Car)	*16.19*	*16.08*	*16.31*	*16.90*	*24.13*	*36.58*	*20.31*	*35.59*
Accessibility to shops (Public Transit)	33.24	32.25	31.98	32.06	30.28	30.25	30.94	28.66
Accessibility to shops (Private Car)	35.58	35.30	35.30	35.31	34.06	34.27	34.48	32.96
Private car travel time variability	*2.72*	*4.11*	*3.07*	*3.46*	*3.55*	*6.65*	*3.71*	*3.56*
Distance to arterial	2.03	1.28	0.74	0.67	3.27	1.62	2.71	2.15
Distance to highway	1.48	1.19	1.27	1.28	2.78	3.12	2.34	4.68
Distance to Châtelet	3.57	8.95	11.42	10.88	27.64	26.36	20.27	41.50
Environment								
% surface in Noisy areas (severe noise: >96dB)	0.00	0.00	4.44	0.03	0.00	2.32	1.36	0.76
% surface of Parks and Gardens	7.59	9.13	9.54	8.65	13.09	12.66	10.60	12.74
% surface of Water	0.90	1.37	1.80	0.65	1.42	0.98	0.71	1.43
% surface of forests	0.27	1.54	1.20	1.01	7.15	7.28	3.97	8.02
% surface of Sporting facilities	1.43	3.05	2.48	3.03	2.42	2.36	2.89	1.77
% surface of Public facilities	9.32	13.72	13.20	12.69	22.67	22.31	17.46	22.52
% surface of Open Space	8.95	12.13	12.60	12.54	16.18	15.65	14.31	14.99
% surface in Redevelopment area**	5.01	8.32	5.23	8.14	2.18	6.67	6.57	1.63

Source: Authors' computations from IAURIF GIS database and Metropolis computations.
* Travellers' surplus computed by METROPOLIS, see relevant section. ** Redevelopment Areas, translated from French: 'Zone Urbaine Sensible'[3]. The variables in italics are computed at the commune level and not represented in Table 2.

We comment hereafter on Tables 1 and 2 together. Note that the variance decomposition and district average values give only limited information concerning the inequalities in distance, accessibility and environmental variables. A more complete overview is given by Lorenz curves, which are used to represent the distribution of a characteristic Y among a population[4]. We now consider specific amenities reported in Tables 1 and 2.

[3] There is no precise translation for the term 'Zones Urbaines Sensibles', but these are areas with high concentrations of social and economic difficulties targeted for government assistance or redevelopment investments. We will use the term Redevelopment Areas for this.

[4] The Lorenz curve was developed by Max O. Lorenz in 1905 for representing income distribution. It shows for the bottom x% of households (plotted on the x-axis), what percentage y% of the total Y corresponds to them (plotted on the y-axis). Perfect Equality (each household has the same quantity of Y)

Table 2: Variance decomposition of accessibility and environmental variables over different scales.

Variable	*District (σ_d)*	*Commune (σ_c)*	*Cell (σ_g)*
Distance and accessibility			
Number of subway and tramway stations	60.5%	22.1%	17.4%
Number of subway and tramway stations around	65.2%	22.3%	12.4%
Number of train stations	10.0%	23.1%	66.9%
Number of train stations around	17.0%	33.4%	49.6%
Accessibility to employment (Public Transit)	35.1%	64.9%	0.0%
Accessibility to employment (Private Car)	39.1%	60.9%	0.0%
Accessibility to shops (Public Transit)	34.9%	65.1%	0.0%
Accessibility to shops (Private Car)	35.8%	64.2%	0.0%
Distance to nearest arterial	7.8%	84.1%	8.1%
Distance to nearest highway	7.8%	83.0%	9.2%
Distance to Châtelet (Paris center)	42.7%	56.9%	0.4%
Environment			
% surface in Noisy area (severe noise: >96dB)	1.8%	62.4%	35.9%
% surface of Parks and Gardens	5.2%	17.2%	77.6%
% surface of Water	7.4%	19.3%	73.4%
% surface of forests	2.5%	36.2%	61.3%
% surface in 'Redevelopment Areas'	3.9%	17.3%	78.8%
% surface of Public Facilities	0.5%	19.3%	80.3%
% surface of Open Space	1.9%	37.6%	60.5%
% surface of Sporting Facilities	2.0%	11.3%	86.7%

Source: Authors' computations from IAURIF GIS database.

Subway, tramway and train stations

The number of (subway, tramway and train) stations around a cell corresponds to the number of stations in a cell and its 8 neighbouring cells. As a consequence, the average number of stations around a cell is approximately 9 times the average number of stations in the cell. The small differences are explained by the fact that some adjacent cells are located in a district different from the one of the cell considered. The fraction of the variance at the grid cell level is significantly lower for the stations around the cell compared to the stations in the cell, which goes along intuition. Train stations are distributed rather equally between the 8 districts (see Table 1), and between the communes (most of the variance is at the cell level), although they are less concentrated in the outer ring, especially in the East.

The inequalities are more prominent in the distribution of subway or tramway stations: more than 60% (resp. 70%) of households live in cells with no train (resp. subway) stations either in the cell or in the 8 surrounding cells (Appendix Fig. 1). These percentages are even larger when restricting to the number of stations in a single cell. The large inequalities in the presence of subway stations are consistent with the fact that they are concentrated in Paris and in the inner ring surrounding Paris. Consequently, the

corresponds to the diagonal, while Perfect Inequality (a single household has all Y) corresponds to the x-axis together with a vertical line at x=100%. A Lorenz curve is always increasing and convex, and located between the lines of Perfect Equality and of Perfect Inequality. When the Lorenz curve is farther from the Perfect Equality line, this reflects more inequality. In this paper we apply this idea to measure spatial inequalities in an urban area.

largest fraction of the variance for the numbers of subway and tramway stations (both in the cell and around it) is at the district (α_d) level (see Appendix Fig. 10 and 11).

Distance to main roads or center of region

The distances are given in kilometers. As seen in Table 2, most of the variability for the distance to arterial or to highway is at the commune level, which suggests that arterials and highways are rather equally distributed between districts, but not between communes. Table 1, on the other hand, suggests a different interpretation, since it shows important differences between district averages. However, the apparently large differences between district averages in Table 1 hides even larger differences between communes in the same district. This apparent inconsistency is partly due to the fact that District 77 is very large (nearly half the total number of cells in Ile-de-France, corresponding to the entire East part of the outer ring) and the distance to arterial or to highway is particularly unevenly distributed in this large district (see Appendix Fig. 12 and 13). Tables 1 and 2 (and intuition) agree that a large fraction (43%) of the variability in (Euclidian) distance to Châtelet is at the district level (Châtelet is the main regional trains station located in the centre of Paris, see Appendix Fig. 9), and an even larger fraction (57%) of the variance is at the commune level.

The distances to Châtelet, highway or arterial are more unequally distributed than travel times (see Appendix Fig. 2 and 3), with the 20% of the population farthest from Châtelet (resp. from the closest arterial or highway) sharing 45% (resp. 50% and 55%) of the distance to Châtelet (resp. to closest arterial or highway).

Travel times

The travel times are given in minutes. The average travel time by Private Car is slightly more unequally distributed than the average travel time by Public Transit (see Appendix Fig. 3, 14 and 15). Note that the average travel time by Private Car roughly increases with the distance to Paris centre, whereas the average travel time by Public Transit is low in some regions of the outer ring. This is probably because the averages are computed using O - D matrix and people living in the outer ring tend to travel by car for longer trips and by Public Transit for shorter trips[5]. The accessibility to employment by Private Car consistently increases with the distance to the Paris Centre, whereas the accessibility to employment by Public Transit varies less regularly with the distance to Paris Centre, depending on the geographical distribution of train stations.

Accessibility measures

We now consider a set of accessibility variables, computed using the method proposed by Jean Poulit in 1974, based on the logarithm of product supply at each destination. It takes into account the opportunity Q_j (number of employments or shops) at destination *j*. The utility S_{ij} of travelling from origin *i* to destination *j*, subtracting travel cost (C_{ij}), is:

$$S_{ij} = \lambda \log(Q_j) - C_{ij}, \tag{2}$$

[5] See Wenglenski (2002) for the distribution of travelled distance for home-to-work trips for different socio-professional categories of workers over the region.

where the weighting factor is $\lambda = \frac{\alpha}{a}$, α is the value of time (VOT), and a is an empirical coefficient used in the gravity trip distribution model. These parameters have been estimated using travel survey data (see IAURIF/THEMA, 2005).

Accessibility from origin i is the log-sum of the accessibilities over all the destinations, j, given by:

$$S_i = \lambda \log\left(\sum_j Q_j \exp\left[-\frac{C_{ij}}{\lambda}\right]\right). \tag{3}$$

It could be presented in monetary units as traveller's surplus. Four Poulit accessibility measures are calculated for two trip purposes (professional and shopping) with two alternative modes (private cars and public transit).

The variability of Poulit accessibility (to employment or to shops) at the grid cell level is clearly null. Cells have the same access attributes within the same travel model zone (which are nearly synonymous with communes). The between-districts differences represent more than 1/3 of the variance and mainly reflect differences between the city of Paris, inner ring and outer ring. Note that the coefficient of variation of the accessibility variable is very low, so one point difference represents a large fraction of the variance. Indeed, Lorenz curves (see Appendix Fig. 4 and 16 to 19) show that the four Poulit accessibility measures are very equally distributed among the population, suggesting that these accessibility measures do not really vary within the region. Therefore, it seems that the Poulit accessibility measures may significantly underestimate inequalities compared to other approaches to measuring accessibility, such as travel time.

We therefore also consider another accessibility measure, unfortunately restricted to access to employment using private car. The "accessibility to employment (M)" variable corresponds to the average travellers' surplus (integrated over the morning peak) for home to work travels computed by METROPOLIS (see de Palma, Motamedi, Picard and Waddell, 2005 for more details). In addition to the travel time cost, this measure takes into account schedule delay cost omitted in the Poulit accessibility measure (but does not consider the quality of the opportunities offered at the final destination). This measure is more unequally distributed than Poulit accessibility but much less than travel times (see Fig. 3 in the Appendix which illustrates this with a Lorenz curve).

Environment

The variability of the environment variables is mostly at the grid cell level, which means that, if households are sensitive to environmental variables, the relevant location choice model should be estimated at the grid cell level rather than at the commune level. The between-district differences in environmental variables are very limited, which suggests that the average quality of environment is highly localized, and averages out across cells within districts. However, Table 1 reveals important differences in forests between Paris, inner ring and outer ring districts, as well as important differences in noise exposure. Note that the severe noise limit (96 dB), as Appendix Fig. 20 shows, corresponds to the zones located around Roissy airport in the North and Orly airport in the south. The apparent paradox between Table 1 and 2 concerning forests and noise is similar to the one discussed above for the distance to arterial or to highway. Note also that the noise measured this way in District 94 is very large because a large airport is located in this small district, close to very dense cells (recall that Table 1 presents population-weighted average values). Commune differences are important concerning the surface of forests, and even more concerning the exposure to severe noise. A decrease of 0.4 to 1.1 percent of dwelling prices for each decibel after 55 has been reported (CGP, 2001). It indicates a 40% depreciation of housing prices for dwellings in a cell completely located in noisy areas.

Recreational areas are distributed relatively unequally over the region. The blue curve in Appendix Fig. 5 shows the fraction of the cumulated parks and gardens surface as a function of the fraction of the cumulated population (ordered by increasing parks and gardens surface in the cell). It is similar to Lorenz curves found for income distributions in developed countries. It shows that the 20% of households who benefit the least from parks and gardens surfaces live in cells which together contain only 5% of parks and gardens surfaces, whereas the 20% of households who benefit the most from parks and gardens surfaces live in cells which together contain 50% of parks and gardens surfaces. The distribution of sports spaces is far more unequal, since more than 40% of the population lives in cells with no sports spaces, whereas the 20% of households who benefit the most from sports spaces live in cells which together contain 80% of sports spaces, as shown in Appendix Fig. 5 and 25. The distribution of forests and water surfaces are even more unequal, since nearly 80% and nearly 90%, of the population live in cells with no forests or no water, respectively. We notice that here, Water concerns land covered permanently by water as lakes and rivers and is considered as an amenity. The rather equal distribution of parks and gardens is confirmed by Appendix Fig. 5, 22 and 23, which show that both population and parks and gardens are approximately concentrated in the same areas, mainly in Paris and inner ring. On the other hand, Appendix Fig. 22 shows that forests are unequally distributed among the population because forests are concentrated in the least dense regions. The reason for the large degree of inequality in water distribution is totally different: water is concentrated in a very small number of grid cells, mainly along the rivers, as shown in Appendix Fig. 24.

The disamenities are more unequally distributed. Indeed, less than 15% of the population lives in cells classified as “Redevelopment Areas”. Appendix Fig. 21 shows that these zones are mainly located in the inner ring. The inequalities are even more striking concerning severe noise, which affects only 1.6% of the population, in the sense that 1.6% of the households live in a grid cell affected by severe noise.

4. Specification of the residential location choice model

To the extent that households have preferences for local amenities and disamenities (we will refer to the collection of those simply described in the preceding section), spatial variations in these amenities should be capitalized into housing prices, which are higher in Paris and the Western part of the inner ring, and lower in outer ring, especially in the Eastern part. We turn now to the specification of a discrete choice model of residential location to examine willingness-to-pay for local amenities, and the trade-offs households make among them.

We wish so to investigate what is the most appropriate spatial scale for studying preferences for locational amenities. This raises the question of the most relevant level for estimating a household location choice model. We are asking whether people have preferences for communes or for smaller more homogeneous geographical units. In order to answer this question, we estimate two location choice models: one at the commune level, and the other one at the grid cell level. This is somewhat different from some prior work that has estimated the effects of right hand side variables at different levels of aggregation (see, for example, Guo and Bhat, 2005), since here we also vary the level of the unit of location choice.

Assuming that all the dwellings i located in the grid cell k, k=1,…,K, which is located in commune j, $j=1,\dots,J$, have the same observable attributes (since we do not have information on structural attributes of the housing units), household h, $h=1,\dots,N$, have the same expected utility $V_i^h = V_k^h$ for them. We assume that expected utility is a linear combination of grid cell attributes X_k and commune attributes Z_j, in which the marginal utilities of grid cell and commune attributes can be household-specific. Expected utility is therefore of the form:

$$V_i^h = X_k \beta_h + Z_j \gamma_h, \tag{4}$$

where β_h and γ_h denote the household-specific marginal utilities of grid cell and commune attributes, respectively.

If households choose among communes rather than grid cells, then expected utility of a dwelling located in commune j only depends on Z_j and is of the form:

$$V_j^h = Z_j \delta_h = Z_j \gamma_h + E\left(X_k \beta_h \mid Z_j\right), \tag{5}$$

where δ_h mixes the marginal utilities of commune attributes and the marginal utilities of grid cells attributes, weighted by their distribution in the commune. The exact formula for $E\left(X_k \beta_h \mid Z_j\right)$ corresponds to the log-sum in a nested logit model (see Anderson, de Palma and Thisse, 1992 for details).

The total number of dwellings in Ile-de-France is denoted by I; the number of dwellings in commune j and in grid cell k are denoted by C_j and G_k, respectively. The probability for household h to choose a dwelling i is given by the Multinomial Logit formula:

$$P_i^h = \frac{\exp\left(V_i^h\right)}{\sum_{i'=1}^{I} V_{i'}^h}, \tag{6}$$

Since all the dwellings located in k have the same expected utility, and the same probability of being selected, Equation (4) implies that the probability that household h selects grid cell k is:

$$P_k^h = G_k P_i^h = \frac{G_k \exp\left(V_k^h\right)}{\sum_{k'=1}^{K} \sum_{i' \text{ in } k'} V_{i'}^h} = \frac{\exp\left(V_k^h + \log\left(G_k\right)\right)}{\sum_{k'=1}^{K} \exp\left(V_{k'}^h + \log\left(G_{k'}\right)\right)}, \tag{7}$$

Similarly, the probability that household h selects commune j is:

$$P_j^h = C_j P_i^h = \frac{C_j \exp\left(V_j^h\right)}{\sum_{j'=1}^{J} \sum_{i' \text{ in j'}} V_{i'}^h} = \frac{\exp\left(V_j^h + \log\left(C_j\right)\right)}{\sum_{k'=1}^{K} \exp\left(V_{j'}^h + \log\left(C_{j'}\right)\right)}, \tag{8}$$

We report below the estimation results for the household location choice models at the level of 500 x 500 meter cells within the study area, and at the level of communes. These analyses extend the previous results reported in de Palma, Motamedi, Picard and Waddell, 2005, which were based strictly on a model specified at the level of communes.

5. Empirical results

We now turn to the results from model estimation. We begin with a comparison of two models, one estimated at the grid cell level and the second at the commune level. We then proceed to a grid cell level model that combines grid cell and commune variables. Table 3 reports the estimation results for the household location choice model at the grid cell level, using only grid cell variables and at the commune level, using only commune variables. The overall explanatory power at the grid cell level is slightly larger than the explanatory power at the commune level (McFadden's LRI=.2414 and 0.2272, respectively). Table 4 provides results from a grid cell level specification that used both grid cell and commune level variables (LRI=0.2440). These new results are largely consistent with those presented in Table 3, which we focus on in the following discussion of these results.

We note that the different results obtained for location choice at two different levels of geography reflect several influences. One is a measurement effect related to aggregation bias. The larger the geographic scope, the more a measure reflects an averaging of variation that occurs within it, producing some level of aggregation bias due to loss of information in the aggregation process. See, for example, Zellner, 1962, for an early test for aggregation bias, and an assessment of aggregation bias in the context of choice models by Allenby and Rossi, 1991. The second influence is behavioural and perceptual. Households and individuals perceive neighbourhoods and the effects of various kinds of amenities differently at different spatial scales. A third influence is a reflection of what geographers have referred to as the Modifiable Aerial Unit Problem (MAUP), which reflects the general finding that models estimated on different levels of geography tend to produce different coefficients, indicating that a model is not independent of the scale at which it is applied (Openshaw, 1984). Note that it is easy to confound these effects. Guo and Bhat, 2004, have explored the question of separating behavioural from MAUP effects in the context of residential location choice in the San Francisco Bay area, and proposed a multi-scale logit model (MSL). They used parcels represented by household travel survey respondents as the universal choice set, and sampled alternatives from within this sample. It is not clear how such a small sample of the housing inventory (the survey respondents' parcels represent a very small fraction of the housing stock) would be able to adequately address problems of MAUP since the sample of housing alternatives would be very sparsely distributed. Our approach differs in that we use exhaustive data of all movers within the region from 1998 to 1999 as the sample of agents, and represent the full available housing stock in the universal choice set, representing these by grid cells in one model, and communes in the second. Parcel-level data for the Paris region was unfortunately not available, and we recognize that the results will depend in part on the aggregate units of geography we were able to use.

An unusual aspect of our study is that we were able to obtain information on the prior residential location, which we find to be extremely important in influencing residential location choice. The most significant variable in our results is (by far) the dummy variable indicating that the cell or commune is located in the same district as the one in which the household lived before it moved. This indicates that households have a strong preference for relocating not too far from their previous residence, which can be partly explained by the proximity to their current employment (indeed, 53% of active persons work and live in the same district) or to other attachment to their neighbourhood and the limit of their search extent. Unfortunately, the previous commune of residence was not indicated in our data, so we could not measure the preference to stay in the same commune.

We found that price effects significantly varied as a function of household head age, and household income and size. We consider both the prices of flats and of houses, and their interactions with some household's characteristics. We were able to use both variables because we observe separately prices for flats and for houses, and their correlation is not too large (79% in real terms, 82% in logs). The results commented below show that larger and older households are more sensitive to the prices of houses than to the prices of flats. This is consistent with the fact that such households have a higher probability to live in a house than in a flat, compared to singles or young families. In future research, we will explore this issue in more detail and model the choice between a house and a flat simultaneously with the location choice. The insignificant coefficient of Log(Price of Flat) in Table 3 means that we can not determine for a reference individual, aged 40, living alone and with a yearly income equal to exp(9.97)=21,400 €[6] a significant sensitivity to the prices of flats, possibly due to omitted variables. The positive coefficient of the income/price interaction variable Log(Price of Flat)* (Log(Income)- $\overline{\text{Log(Income)}}$) means that willingness to pay increases with income. According to commune level estimates, older and larger families are slightly more sensitive to the prices of flats[7]. Results not reported here show that the coefficient of flats price is not sensitive to household head age or family size at the grid cell level. Our reference individual seems to prefer communes with higher house prices. One can suspect that omitted variable bias (prices are higher in places with better unobserved amenities, and their positive effect is not modelled explicitly in the location regression) dominates the true negative price effect. However, with the same income and family size, houses price effect becomes negative when household head is aged over 49 (commune level estimates) or 62 (grid cell level estimates)[8]. Similarly, at the reference income and age, the house price coefficient is negative for families with more than 2 members.

---

[6] This corresponds to the average (in log terms) per capita household income in Ile-de-France. Household per capita income corresponds to household income divided by the square root of the number of household members, which is an equivalence scale commonly used.

[7] This is implied by the negative signs of the coefficients of Log(Price of Flat)* (Age-40)/10 and of Log(Price of Flat)* (Number of hh members – 1).

[8] This figure corresponds to age at which the marginal price effect is zero for the reference income and family size, namely 0.08735*Log(Price of House)-0.10359*Log(Price of House)* (Age-40)/10=0 at the commune level.

Table 3: Estimation results at commune and grid cell level.

Variable	*Estim. at commune level*		*Estim. at cell level*	
	Coefficient	**t-statistic**	**Coefficient**	**t-statistic**
Same district as before move	2.53879	275.91	2.53494	272.47
Paris	-0.25412	-10.10	-0.16871	-10.29
Log(Price of Flat)	0.00644	0.20	0.02178	0.94
Log(Price of Flat)* (Age-40)/10	-0.02976	-1.74		
Log(Price of Flat)* (Log(Income)- $\overline{\text{Log(Income)}}$)	0.40919	8.20	0.24849	5.84
Log(Price of Flat)* (Number of hh members – 1)	-0.04236	-2.40		
Log(Price of House)	0.08735	3.70	0.13934	7.41
Log(Price of House)* (Age-40)/10	-0.10359	-8.78	-0.11371	-20.90
Log(Price of House)* (Log(Income)- $\overline{\text{Log(Income)}}$)	0.18693	5.78	0.14874	5.17
Log(Price of House)* (Number of hh members – 1)	-0.10932	-8.38	-0.13038	-20.43
Number Subway stations around			-0.00412	-2.37
Number Subway stations in the commune / cell	0.00518	4.59	0.00146	0.21
Number Railway stations around			0.01326	3.85
Number Railway stations in the commune / cell	-0.00940	-4.86	0.00616	0.60
Average travel time from j, commuting (TC) [hr]	0.02483	0.80	0.04007	1.45
TC*(Dummy female) [hr]	-0.37377	-8.28	-0.29400	-6.56
Distance to highway [km]	-0.00594	-3.10	-0.00146	-0.74
Distance to arterial [km]	-0.00798	-2.98	-0.01500	-5.87
Distance to Châtelet (Paris centre) [km]	0.00167	2.94	-0.00054	-1.06
% households with 1 member * 1 member in h	1.87670	20.62	2.27023	30.73
% households with 2 members* 2 members in h	1.33649	4.20	1.77322	10.03
% households with 3+ members* 3+ member in h	2.22967	22.16	2.07516	27.98
% hh with no working member * no working member in h	7.79690	33.91	5.38190	33.49
% hh with 1 working member * 1 working member in h	-0.93134	-5.84	0.50248	4.58
% hh with 2+ working member * 2+ working member in h	1.66425	14.83	0.40430	5.01
% hh with a young head	0.45006	3.01	1.02140	10.86
% hh with a young head * young head in h	4.77740	28.34	3.24295	29.59
% hh with a middle age head * middle aged head in h	-0.69337	-4.16	-0.16208	-1.51
% Rich hh * Rich h	3.28038	28.36	2.87636	36.58
% Medium Income hh * medium income h	1.62120	10.11	1.92296	17.00
% Poor hh * poor h	0.45889	3.53	1.06997	12.11
% households with a foreign head * foreign head in h	5.04570	26.74	4.57810	37.26
% households with a foreign head * French head in h	-1.99493	-17.44	-1.30553	-16.71
% of surface in Redevelopment Area * Rich h	0.67811	4.97	-0.22815	-4.28
% of surface in Redevelopment Area * Med. Inc. h	-0.08340	-0.78	-0.19330	-5.16
% of surface in Redevelopment Area * Poor h	0.43164	4.45	0.16457	4.99
Log of the number of residential units	0.06682	9.26	0.02295	3.27
% of Flats in total dwellings * Foreign head in h	1.36775	15.42	0.67595	11.37
% of Flats in total dwellings * French head in h	0.48469	7.28	0.16077	3.76
% of Flats in total dwellings * (N. of members - 1)	-0.08459	-4.29	-0.05561	-4.29
% of Flats in total dwellings * young head in h	-0.13120	-2.43	-0.03680	-1.01
% of Flats in total dwellings * old head in h	-0.76601	-8.40	-0.18806	-2.82
% of surface in Noisy area	-0.08929	-1.56	-0.07667	-1.52
% of surface covered by Forest	-0.11708	-3.09	-0.29739	-5.05
% of surf. covered by Forest * N. of Children	0.42292	6.87	0.46115	4.99
% of surface covered by Water	0.20360	1.62	0.33284	4.24
% of surface covered by Parks and Gardens	-0.42727	-4.22	-0.15159	-2.93
% of surf. covered by Parks * N. of children	-0.12942	-0.71	0.61447	6.67
% of surface covered by Sporting facilities	0.44171	2.50	-0.37138	-4.06
% of surf. covered by Sporting facilities * N of children	0.26755	0.83	0.78619	4.93

Notes: “%” sign represents the proportion. “N. of children” counts family members aged 11 or less.
Source: Authors’ estimations

Table 4: Estimation results at grid cell level with grid cell and commune variables.

Variable	*Communes' variables*		*Cells' variables*	
	Coefficient	**t-statistic**	**Coefficient**	**t-statistic**
Same district as before move	2.53671	271.53		
Paris	-0.22675	-8.43		
Log(Price of Flat)	-0.01300	-0.51		
Log(Price of Flat)* (Age-40)/10	0.28139	6.29		
Log(Price of Flat)* (Log(Income)- $\overline{\text{Log(Income)}}$)	0.14508	7.11		
Log(Price of Flat)* (Number of hh members – 1)	-0.13010	-22.93		
Log(Price of House)	0.25108	8.44		
Log(Price of House)* (Age-40)/10	-0.13981	-19.70		
Log(Price of House)* (Log(Income)- $\overline{\text{Log(Income)}}$)			-0.00322	-1.71
Log(Price of House)* (Number of hh members – 1)	0.00605	5.16	0.00097	0.14
Number Subway stations around			-0.00322	-1.71
Number Subway stations in the commune / cell	0.00605	5.16	0.00097	0.14
Number Railway stations around			0.02115	5.77
Number Railway stations in the commune / cell	-0.01494	-7.10	0.00677	0.66
Average travel time from j, commuting (TC) [hr]	0.01065	0.33		
TC*(Dummy female) [hr]	-0.35600	-7.73		
Distance to highway [km]			-0.00067	-0.33
Distance to arterial [km]			-0.01314	-4.96
Distance to Châtelet (Paris centre) [km]			-0.00028	-0.47
% households with 1 member * 1 member in h	-0.16085	-1.15	2.30278	20.02
% households with 2 members* 2 members in h	-0.28286	-0.75	2.04371	9.65
% households with 3+ members* 3+ member in h	-0.08790	-0.59	2.11213	19.15
% hh with no working member * no working member in h	5.09950	16.50	3.30625	15.56
% hh with 1 working member * 1 working member in h	-2.63547	-12.42	1.65813	11.18
% hh with 2+ working member * 2+ working member in h	1.53570	9.18	-0.04502	-0.36
% hh with a young head	-0.40537	-2.16	1.11631	9.36
% hh with a young head * young head in h	1.93974	9.14	2.64059	19.00
% hh with a middle age head * middle aged head in h	-0.86585	-4.05	0.00740	0.05
% Rich hh * Rich h	0.60752	3.91	2.56647	23.79
% Medium Income hh * medium income h	-0.37902	-1.73	1.97329	12.54
% Poor hh * poor h	-0.80207	-4.71	1.35812	11.62
% households with a foreign head * foreign head in h	1.21668	4.63	3.99283	23.27
% households with a foreign head * French head in h	-0.10054	-0.67	-1.30069	-12.58
% of surface in Redevelopment Area * Rich h	0.85358	5.58	-0.35104	-6.05
% of surface in Redevelopment Area * Med. Inc. h	0.21783	1.82	-0.22336	-5.43
% of surface in Redevelopment Area * Poor h	0.24909	2.28	0.12894	3.56
Log of the number of residential units	0.01889	2.46	0.00828	1.11
% of Flats in total dwellings * Foreign head in h	1.31465	10.50	0.06254	0.73
% of Flats in total dwellings * French head in h	0.33011	3.75	0.07201	1.23
% of Flats in total dwellings * (N. of members - 1)	-0.08926	-3.23	0.00257	0.14
% of Flats in total dwellings * young head in h	0.01850	0.23	-0.08934	-1.59
% of Flats in total dwellings * old head in h	-1.30230	-8.72	0.45902	4.19
% of surface in Noisy area	-0.14111	-1.94	0.01074	0.17
% of surface covered by Forest	-0.05748	-1.46	-0.25683	-4.21
% of surf. covered by Forest * N. of Children	0.33436	5.07	0.30723	3.15
% of surface covered by Water	0.35641	2.70	0.24600	2.99
% of surface covered by Parks and Gardens	-0.52131	-4.90	-0.05872	-1.08
% of surf. covered by Parks * N. of children	-0.26030	-1.36	0.56111	5.73
% of surface covered by Sporting facilities	0.25816	1.41	-0.36015	-3.89
% of surf. covered by Sporting facilities * N of children	0.48338	1.46	0.75183	4.63

Notes: "%" sign represents the proportion. "N. of children" counts family members aged 11 or less.
Source: Authors' estimations

Interestingly, larger or older families are more sensitive to the price of houses than flats,[9] which is consistent with the fact that, when they become older and/or have children, families have a stronger tendency to live in a house rather than in a flat. On the other hand, the price coefficient is more sensitive to income for flats than for houses. For the richest households, the net price effect is positive even in very large households with an old head. This reflects the fact that richer families are more sensitive to unobserved amenities and less sensitive to prices (for example they may be willing to pay a premium to live in neighbourhoods with exclusively wealthy neighbours).

Households prefer living in communes with many subway or tramway stations, but dislike the close proximity of such stations. Note that a station located in the commune but not in the cell or surrounding cells is clearly beneficial since it improves accessibility. On the other hand, a very close station may be detrimental because of the noise and crowding, and the clustering of business activity around stations, so the sign of the overall effect is not clear a priori.

Table 4 shows that households are mainly sensitive to stations in the commune or in the surrounding cells rather than in the cell they live, with positive and significant effects for the commune and negative for nearby stations. Railway stations around a cell were positive and significant in their influence on location choices, while it was negative in commune and not significant at the 5 percent level for cell-level concentration. This suggests that individuals value proximity to railway stations, and are willing to walk or drive (or take the bus) relatively long distances (more than 500 meters, to go to adjacent cells.

All the distance variables have a negative sign meaning that, ceteris paribus, households prefer to locate close to arterials or highways, or close to the centre of Paris (Châtelet). The negative coefficient for Paris dummy reflects the fact that, ceteris paribus, individuals prefer the suburbs to Paris. However, it seems more plausible that this negative coefficient rather reflects the non-linear effect of other variables, which are particularly high in Paris, and interactions with the distance variables. For example, the number of subway stations is particularly high in some cells within Paris. If the marginal utility of a subway station is decreasing with the number of subway stations, then imposing a linear specification for the effect of the number of subway stations leads to overestimate this effect in the cells or communes with a very large number of stations, which happen to be all located inside Paris. The (negative) Paris coefficients then corrects for this overestimation. This interpretation is consistent with the fact that Paris coefficient is significantly larger (in absolute terms) in the commune estimates than at the grid cell level (very large numbers of stations are more common in communes than in grid cells).

We tested (results not reported here) a non-monotonous specification for the distance to arterial or to highway. The idea was that the close proximity to a highway is a nuisance because of pollution and noise, whereas individuals also dislike a too large distance to highway, for accessibility considerations. However, we failed to estimate the intuitive inverse U-shaped effect for those distance variables, with a large number of specifications for the covariates. This suggests that accessibility and travel time considerations dominate pollution and noise nuisance considerations with respect to location choice, at least at the scale we are measuring these effects.

[9] The coefficient of Log(Price of House)* (Number of hh members – 1) is larger than the coefficient of Log(Price of flat)* (Number of hh members – 1) in absolute terms.

Female-headed households dislike communes with larger average commuting times by transit. This is consistent with the tendency towards shorter commutes by women reported elsewhere. We failed to estimate significant effects for average travel time using Private car or accessibility variables, probably because of the large correlation between the variables measuring distances, average travel times and accessibility.

Households tend to prefer locating close to households with a rather young head. As far as age, household size or income is concerned, households tend to locate close to similar households. According to Table 4, these segregation effects seem to be more discernible at a small neighbourhood level, since the measures of social structure of population computed for grid cells are more significant than those computed for communes. This is consistent with intuition, since as the scale is increased one would expect to aggregate across more heterogeneous neighbourhoods. Note that the counter-intuitive negative coefficients of some commune level variables in Table 4 might, similarly to Paris dummy, reflect some non-linearities of the effect of the corresponding composition variables.

Foreign households tend to locate in grid cells and, to a lesser extent, communes with more foreign households, whereas French households prefer to locate in places with less foreign households. Poor households tend to locate close to other poor households, and the same is true for rich households. According to grid cell estimates, rich or medium income households tend to locate far away from ‘Redevelopment Areas’, whereas poor people have a stronger tendency to locate in such neighbourhoods, because they can not afford better locations. The positive effect of for rich or medium income households is harder to explain.

As expected, households attempt to avoid noisy areas. According to Table 4, this effect is more important at the commune level than at the grid cell level, suggesting that the negative effect of airports are not limited to the grid cells suffering from severe noise, but influence a broader area. Households with children like living close to forests, whereas households without children seem to avoid forests. However, this probably reflects a displacement effect (more forests means less room for dwellings) rather than a preference effect. This hypothesis is supported by the more negative or less positive signs at the grid cell level (where the displacement effect is a priori more important) than at the commune level. Households seem to appreciate the proximity of water, but this is not very significant, and there seems to be some displacement effect at the grid cell level.

Households tend to locate very close to parks and gardens, especially when they have children. This effect is important at the grid cell level, but negative and hardly significant at the commune level. The difference between commune and grid cell estimates might result from the fact that parents are only interested in parks and gardens available within walking distance. In addition, parks and gardens are rather uniformly distributed over communes. Households seem to appreciate the availability of sports spaces within the commune, even when they do not have children. The negative coefficient at the grid cell level may result from the displacement effect, since sports spaces are often large and occupy a significant fraction of a grid cell.

We have studied the effect of density, which is positive and larger at the commune level than at the grid cell level (results not reported here). The positive coefficient for the logarithm of the number of households in the cell reflects the fact that new households tend to locate in more populated areas. This suggests that dense areas are more dynamic, attract more people and rotation is more important (more households move in and out). The difference in the coefficients for cell and commune density results from a displacement effect (denser parts have less room for new households to come in). This displacement is likely to be more important at a small geographical level, since a large density in a cell probably means a large density everywhere in the cell, whereas a large density is a commune may corresponds to a very large density in some parts and more room for new households in other parts of the commune. These results on interactions of household characteristics with social characteristics of neighbourhoods are consistent with a substantial body of literature on residential segregation tendencies. Note that the analysis used importance sampling, meaning that the alternative cells were included in the choice set with a probability proportional to the number of dwellings, so the variable "Number households" was never zero. The coefficients of the number of collective dwellings (by itself and crossed with household size) means that only singles are attracted by collective dwellings, whereas 2-member households are not sensitive to the number of collective dwellings and larger families prefer locations with less collective dwellings. The opposite effects prevail for the number of individual dwellings, except that only families with more than 6 members are attracted by places with more individual dwellings. These two coefficients reflect both a displacement effect (more individual dwellings imply less room for new inhabitants) and a preference effect (larger families prefer a neighbourhood of individual dwellings). The displacement effect dominates for most families (6 members or less).

Compared to the results reported in Table 3, the main difference in Table 4, from the grid cell estimation using grid cell and commune variables is that we can see more complex interactions among the scales at which different kinds of amenities are more significant. No clear dominance of one scale appears to emerge from these results for different types of effects. In some cases, the commune and grid cell effects are both significant and in agreement, while in others they have opposite signs, and in others only one or the other is significant. We do not attempt to generalize from this any clear lessons about the most natural scales for measuring differing locational amenity effects.

6. Conclusions

In this paper we have examined empirically local amenities in the Paris metropolitan region. We find first that there is considerable inequality in the spatial distribution of these local amenities, including accessibility, environmental and socio-economic indicators such as the distribution of households' income, or the fraction of foreign-born population. We use spatial representation and Lorenz curves to examine the degree of inequality in these amenities over the population, and this provides evidence that some amenities (or disamenities) are much more unequally distributed than others. The most extremely unequally distributed amenities are noise (due to its concentration near airports), Redevelopment Areas, presence of water and forests, and presence of train and subway stations. Some measures, such as the Poulit accessibility measure, were remarkably insensitive, by contrast, appearing to be more ubiquitous.

We have recognized that local amenities are generally capitalized into the housing market, and explored the willingness to pay of households for these amenities within the Paris region using alternative specifications of a location choice model. One of the core questions we examined was the spatial scale of the amenity effects and how this is captured in a location choice context. By estimating models at both a commune and at a grid cell level, we have generated new insights into how households in the Paris region trade off amenities against each other and against housing cost.

We find that the residential location choice model fits the data moderately better at the smaller scale of the grid cell level compared to the commune level. Some have previously argued that models are likely to fit better at more aggregate levels, but this is not what we find. This could be due to some combination of better measurement (less aggregation error), and a more accurate representation of households preferences of households. We have not completely avoided the MAUP problem, however, since both units of analysis we have examined are aggregate in nature. Nevertheless, our results are largely consistent at the two levels of geography, with some notable exceptions. Some amenities appear to be more localized in their effect while others are broader in their effect.

This study has provided considerable insight into the way local amenities are distributed in the Paris region, and how these influence household location choices. Further elaboration of this research may explore ways of generating simulation-based sensitivity analysis to explore the trade-offs among amenities and cost for different classes of households. We also wish to further examine the potential endogeneity of prices, since amenities should be capitalized into housing prices.

References

Allenby, G. M. and Rossi, P. E. (1991) "There Is No Aggregation Bias: Why Macro Logit Models Work", *Journal of Business & Economic Statistics*, 9, 1, pp. 1-14.

Anderson, S., de Palma, A. and Thisse, J. (1992) *Discrete choice theory of product differentiation*, MIT Press, Cambridge, Ma.

Boyle, M. A. and Kiel, K. A. (2001) "A survey of house price hedonic studies of the impact of environmental externalities", *Journal of Real Estate Literature,* 92, pp. 117-144.

Brown, H. J. and Li, M. M. (1980) "Micro-Neighborhood Externalities and Hedonic Housing Prices", *Land Economics*, 56, 2, pp. 125-141.

Brueckner, J. K. (2000) "Urban Sprawl: Diagnosis and Remedies." *International Regional Science Review,* 23, 2, pp. 160-171.

Burchell, R. W. et al. (1998) *The Costs of Sprawl Revisited,* Transit cooperative research program, Report No. 39, Transportation Research Board, National Academy of Science Press, Washington, D.C, USA.

Chan, S. (2004) "Drawing the Line: The Effect of Urban Growth Boundaries on Housing Prices in the San Francisco Bay Area", *PhD. thesis*, Public Policy Department, Stanford University.

Coase, R. H. (1960) "The problem of social cost", *Journal of law and economics,* 3, pp. 1-44.

CGP, Commissariat Général du Plan, (Boiteux M.) (2001) Transports : choix des investissements et coût des nuisances (rapport Boiteux 2), Commissariat Général du Plan, Paris, France.

de Palma, A., Marchal, F. and Nesterov, Y. (1997) "METROPOLIS: A Modular System for Dynamic Traffic Simulation", *Transportation Research Record,* 1607, pp. 178-184.

de Palma, A., Lindsey, R. and Kilani, M. (2005) "Congestion pricing on an urban road network: A study using the dynamic traffic simulator METROPOLIS", *Transportation Research, Part A*, 39, pp. 588-611.

de Palma A., Motamedi, K., Picard, N. and Waddell, P. (2005) "A model of residential location choice with endogenous housing prices and traffic for Paris region", *European Transport*, 31, pp. 67-82.

de Palma, A., Picard, N. and Waddell, P. (2007) "Discrete choice models with capacity constraints: an empirical analysis of the housing market of the greater Paris region", *Journal of Urban Economics*, in press.

de Palma, A., Lindsey, R. and Proost, S. (2007) *Investment and the Use of Tax and Toll Revenues in the Transport Sector*, in press.

Guo, J. Y., and Bhat, C. R. (2004) "Modifiable Areal Units: Problem or Perception in Modeling of Residential Location Choice?", *Transportation Research Record*, 1898, pp. 138-147.
IAURIF and THEMA/TT&R. (2005) *SIMAURIF, Modèle dynamique de SIMulation de l'interAction Urbanisation-transports en Région Ile-de-France*. Report for DRAST/ PREDIT, N°05MT6008, Institute for urban planning and development of the Paris Ile-de-France region (IAURIF).
Jones, P. (2003). "Acceptability of Road User Charging: Meeting the Challenge". In J. Schade and B. Schlag (Eds.), *Acceptability of Transport Pricing Strategies* (pp. 27-62), Elsevier, Amsterdam.
Knight, F. (1924) "Some fallacies in the Interpretation of Social Cost", *Quarterly Journal of Economics*, 38, 4, pp. 582-606.
Koland O. (2006) "Urban sprawl, car-related pollution and settlement structure: Insights from a two-region CGE model", presented in European summer school on Computable General Equilibrium Modelling in Environmental and Resource Economics, Venice.
Krugman, P. (1991). *Geography and Trade*, MIT Press, Cambridge, MA.
Lorenz, M. O. (1905). *Methods of measuring the concentration of wealth,* Publications of the American Statistical Association, 9, pp. 209-219.
Lewis, G. H. and Michel, R. C. (Eds) (1990). *Microsimulation Techniques for Tax and Transfer Analysis,* The Urban Institute Press.
Openshaw, S. (1984) "The Modifiable Areal Unit Problem." *Concepts and Techniques in Modern Geography*, 38, GeoBooks, Norwich.
Pahaut, S. and Sikow, C. (2006) "History of thought and prospects for road pricing", *Transport Policy, Modelling of Urban Road Pricing and its Implementation*, 13, 2, pp. 173-176.
Pigou, A. C. (1920) *The Economics of Welfare,* Macmillan, London, UK.
Small, K. A. and Parry, I. W. H. (2005) "Does Britain or The United States Have the Right Gasoline Tax?", *American Economic Review*, 95, pp. 1276-1289.
Verhoef, E. T. and Nijkamp, P. (2003) *Externalities in the Urban Economy,* Tinbergen Institute Discussion Paper No. 78/3.
Waddell, P., Borning, A., Noth, M., Freier, N., Becke, M. and Ulfarsson, G. (2003) "UrbanSim: A Simulation System for Land Use and Transportation", *Networks and Spatial Economics,* 3, pp. 43-67.
Waddell, P., Sevcikova, H., Socha, D., Miller, E. and Nagel, K. (2005) "Opus: An Open Platform for Urban Simulation", paper presented at the Computers in Urban Planning and Urban Management Conference, London, U.K.
Waddell, P., Ulfarsson, G. F., Franklin, J. and Lobb, J. (2007) Incorporating Land Use in Metropolitan Transportation Planning, *Transportation Research Part A: Policy and Practice*, 41, pp. 382-410.
Wegener M. and Lautso K. (2004), "A new generation of tools to facilitate decision- making (PROPOLIS project)", presented in Sustainable Urban Management and Land Use, From European research to application at local level Conference, Prague, Czech Republic.
Wenglenski, S. (2002) "Parcours effectif à l'emploi versus accès potentiel à l'emploi: une mesure des contraintes des actifs dans la métropole parisienne", presented in 38th conference of French speaking regional science association (ASRDLF), August 2002, Trois-Rivières, Canada.
Zellner, A. (1962) "An Efficient Method of Estimating Seemingly Unrelated Regressions and Tests for Aggregation Bias", *Journal of the American Statistical Association*, 57, pp. 348-368.

Acknowledgements

This study has been performed as part of the SIMAURIF project, in collaboration with the Institute for urban planning and development of the Paris Ile-de-France region (IAURIF), and financed by the French Ministry of transportation (PREDIT) and the Transportation department of Ile-de-France region (DREIF). The first and the third authors would like to acknowledge the French Ministry of Research grant 03396. The last author also acknowledges the National Science Foundation Grants EIA-0121326 and IIS-0534094. We wish to acknowledge the collaboration of Dany Nguyen-Luong (IAURIF) and José Moyano (adpC). Finally, this paper benefited from the comments of participants in The 53rd Annual North American Meetings of the Regional Science

Association International in Toronto, Nov. 2006 and in particular Jan Rouwendal, for very helpful comments. Finally, we thank the anonymous referees and the editors for their fruitful comments.

Appendix

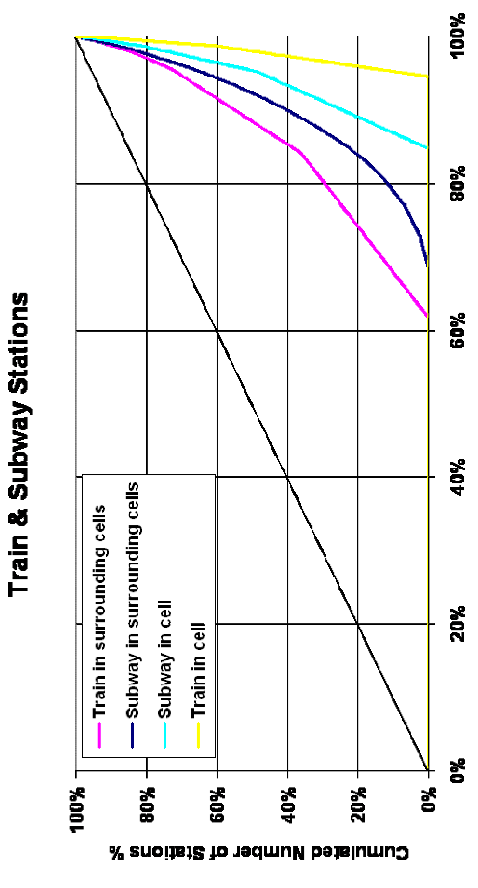


Figure 1: Lorenz curve for Train and Subway stations in the cell or surrounding.

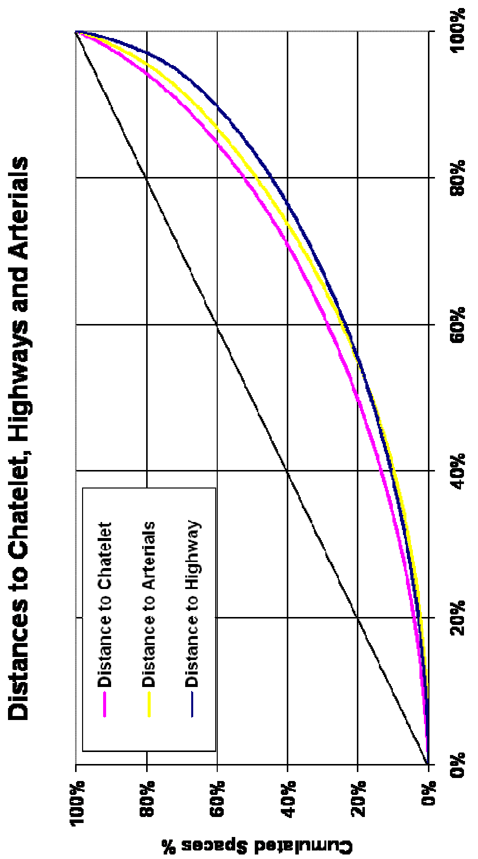


Figure 2: Lorenz curve for Distances to Paris Centre, highway and arterial.

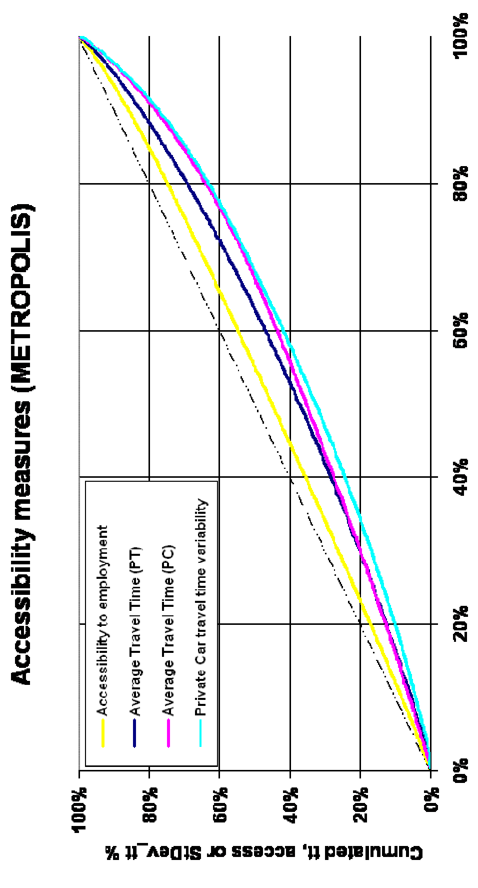


Figure 3: Lorenz curve for different measures of Accessibility (by METROPOLIS).

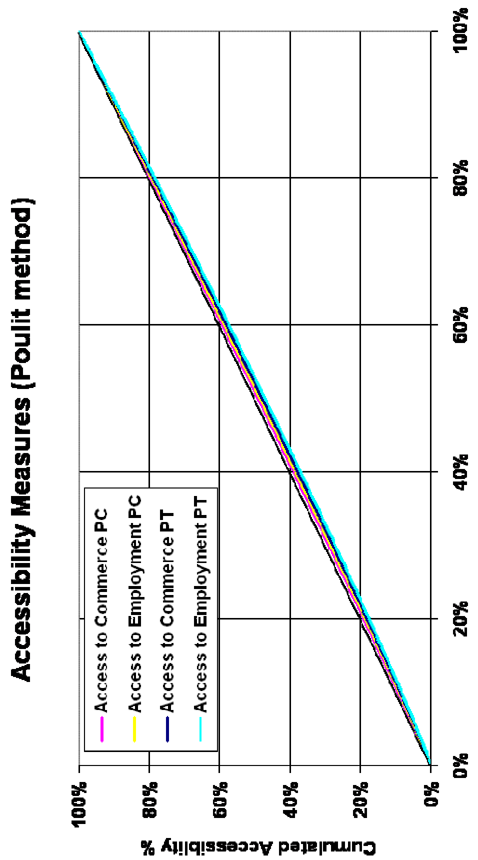


Figure 4: Lorenz curve for different accessibility measures (by Poulit method).

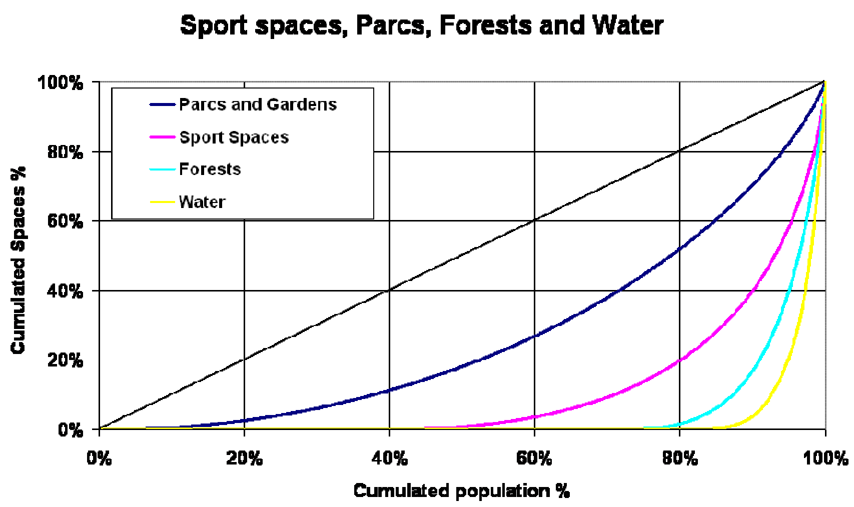


Figure 5: Lorenz curve for Sporting facilities, Parks, Forests and Water.

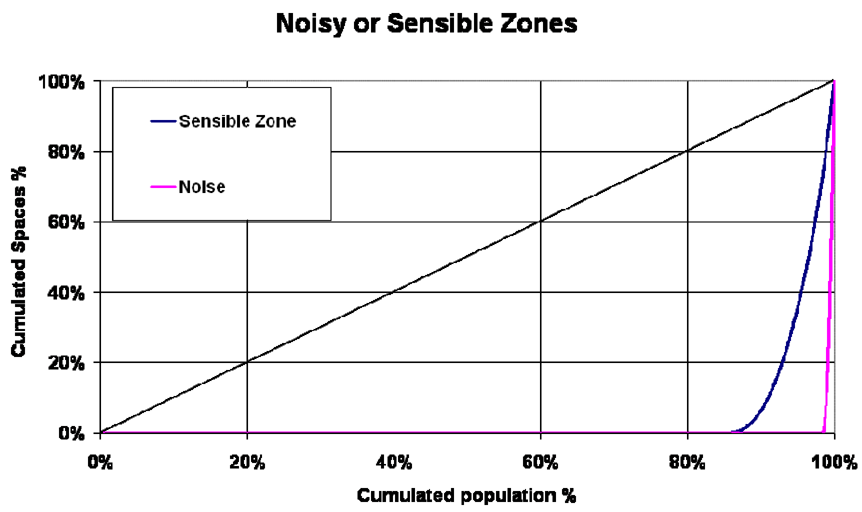


Figure 6: Lorenz curve for Noisy and Redevelopment Areas.

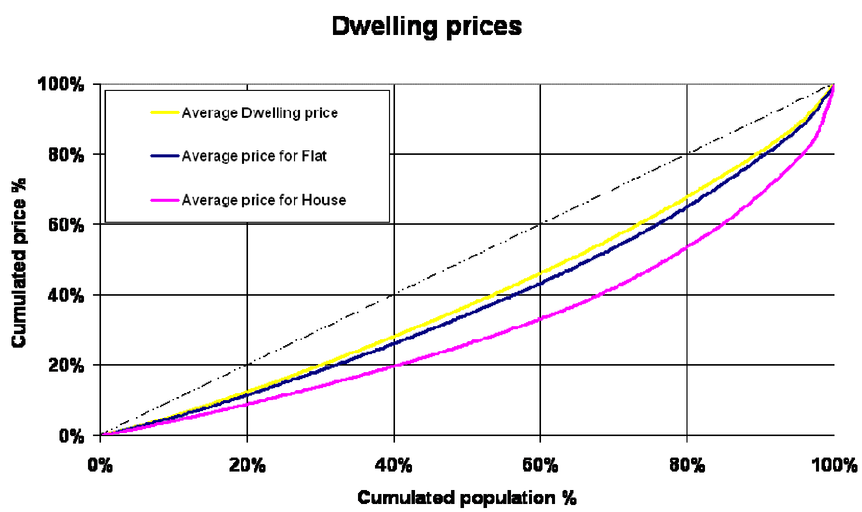


Figure 7: Lorenz curve for Dwelling Prices.

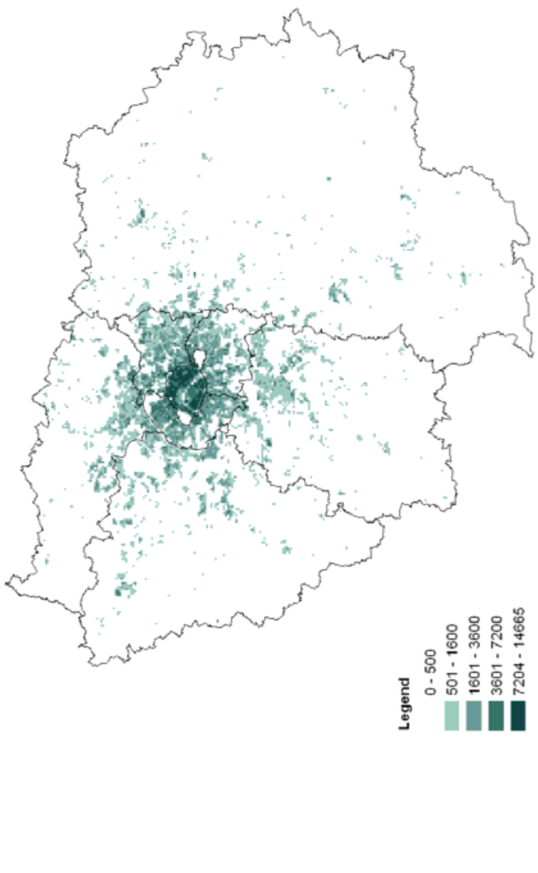


Figure 8: Distribution of population over the region.

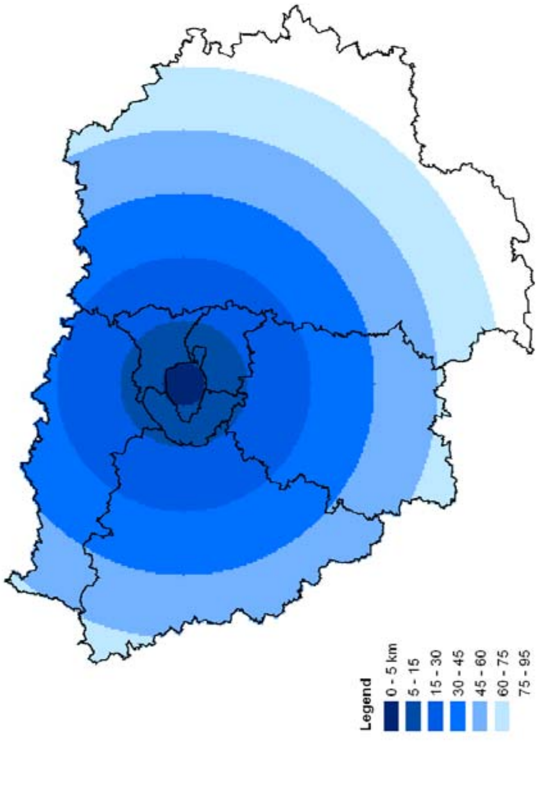


Figure 9: Distance to Paris Center.

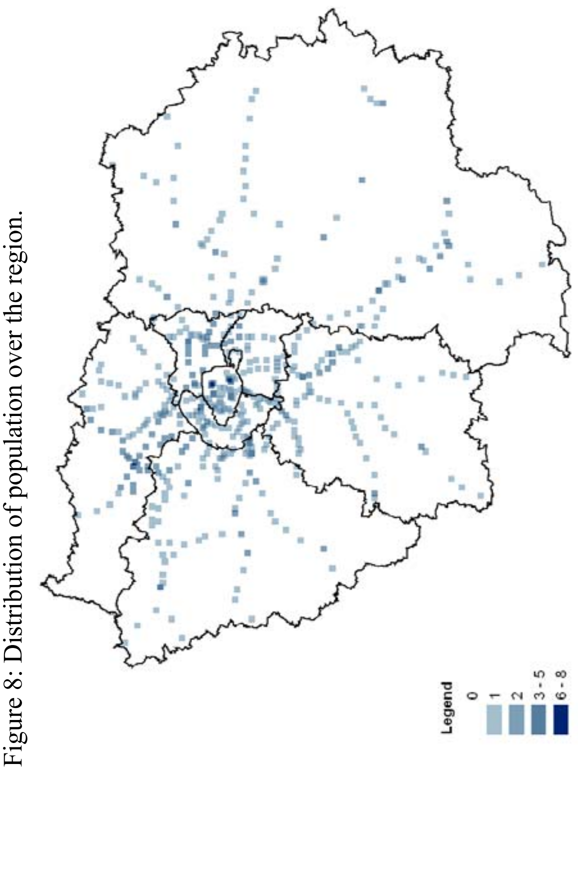


Figure 10: Number of train stations in surrounding cells.

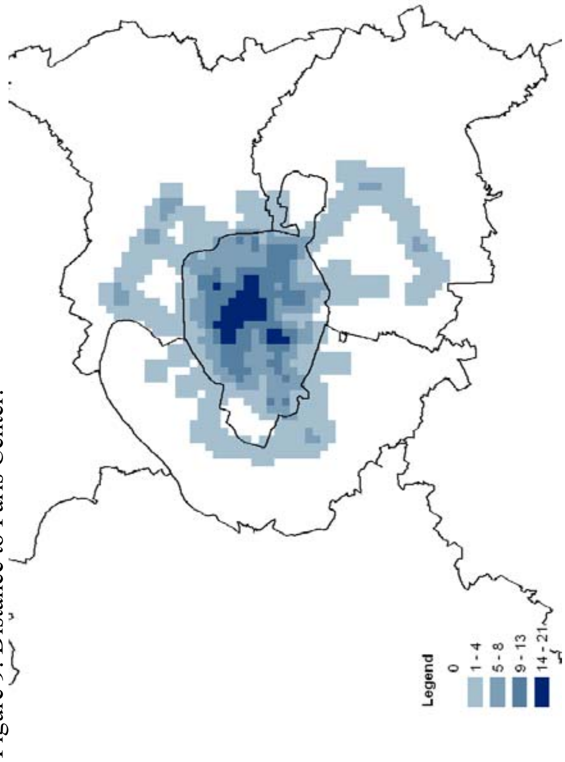


Figure 11: Number of Subway stations in surrounding cells.

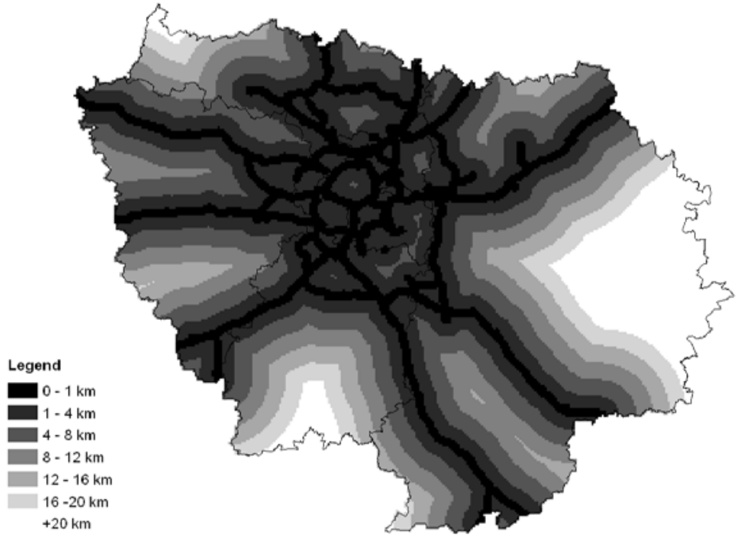


Figure 12: Distance to the nearest highway.

Figure 13: Distance to the nearest arterial.

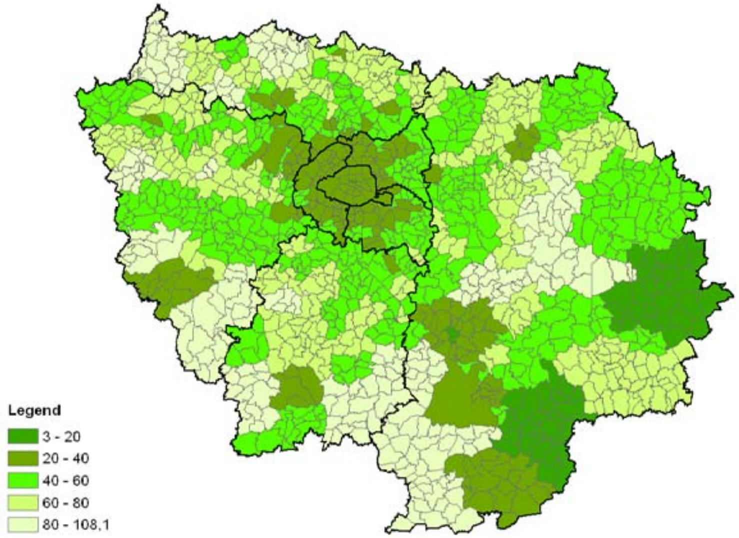


Figure 14: Average travel time by Public Transit in minutes.

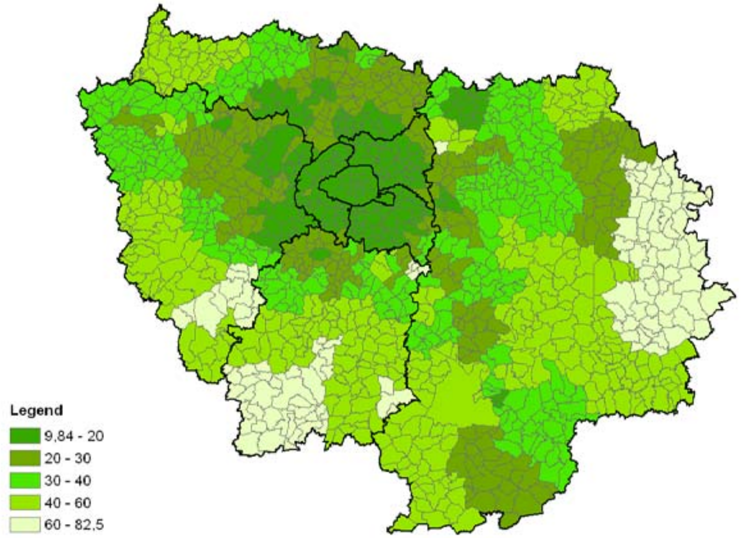


Figure 15: Average travel time by Private Car in minutes.

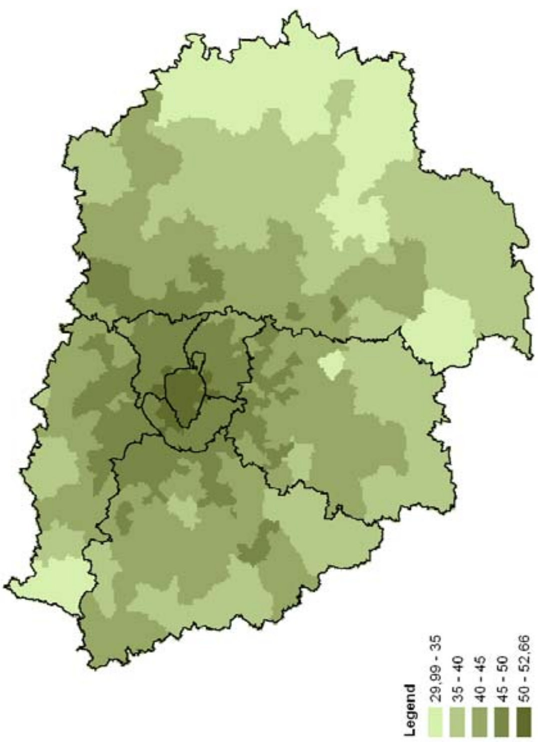


Figure 16: Accessibility to Employment by Public Transit (Poulit method).

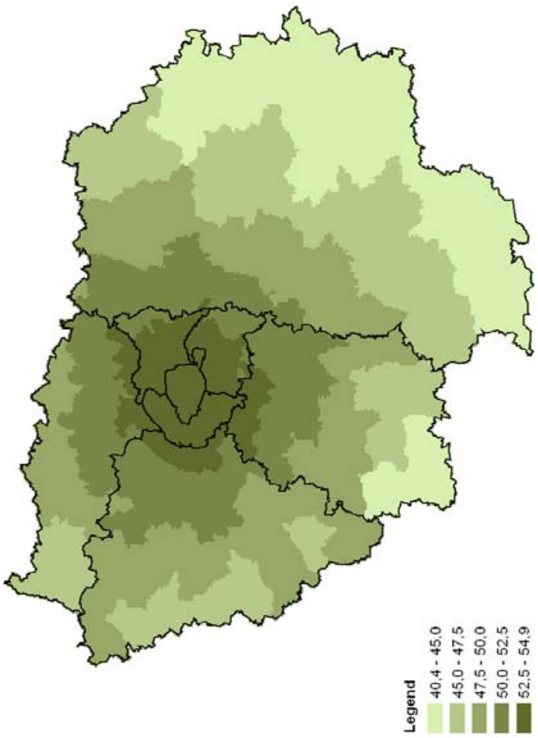


Figure 17: Accessibility to Employment by Private Car (Poulit method).

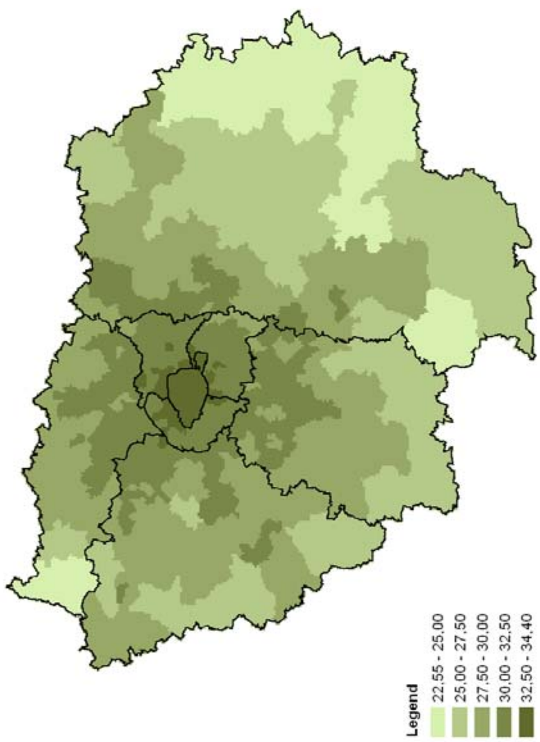


Figure 18: Accessibility to Commerce by Public Transit (Poulit method).

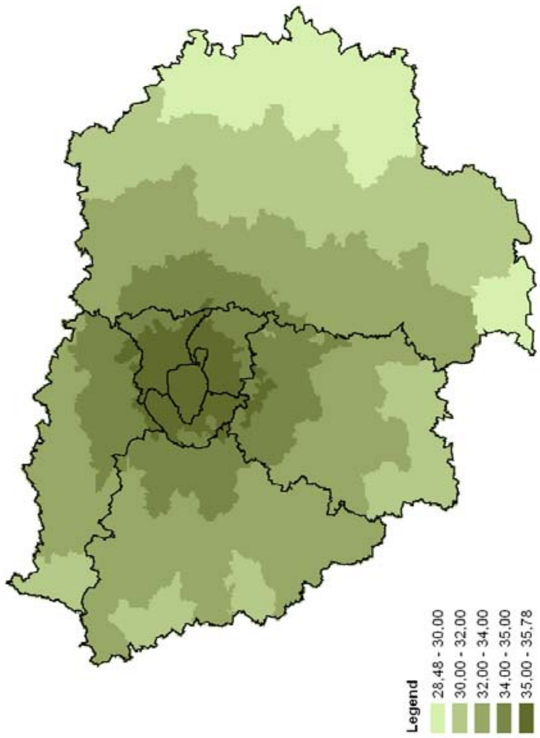


Figure 19: Accessibility to Commerce by Private Car (Poulit method).

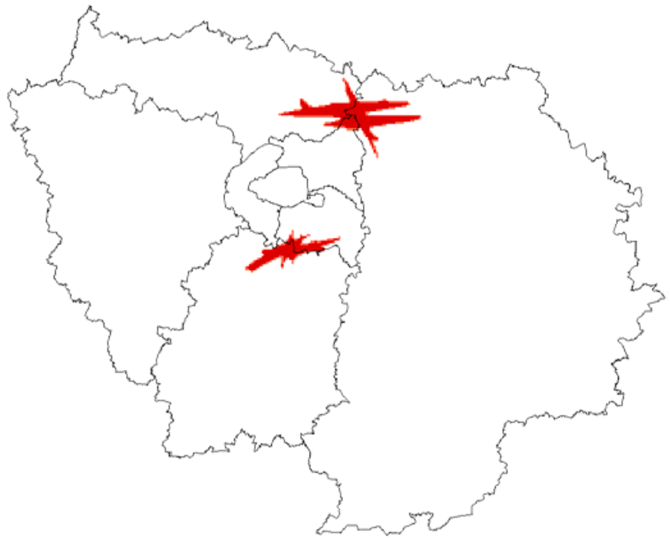


Figure 20: Noisy areas near airports.

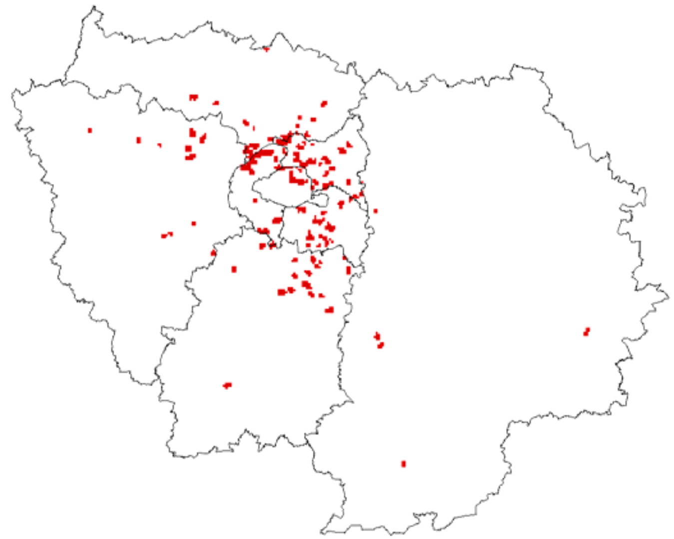


Figure 21: Redevelopment Areas.

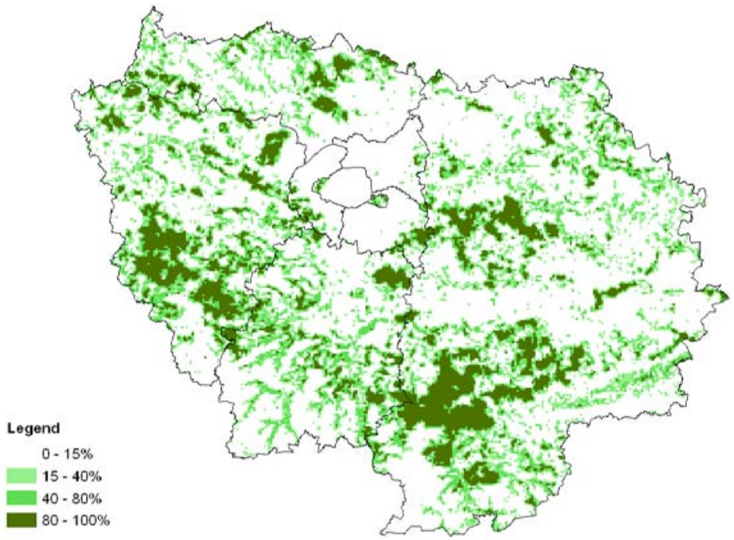


Figure 22: Proportion of cells' surface covered by Forest.

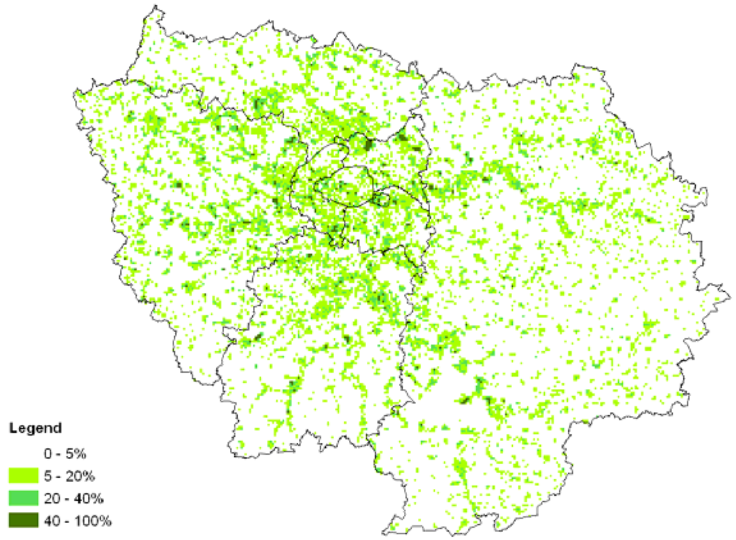


Figure 23: Proportion of cells' surface in Parks and Gardens.

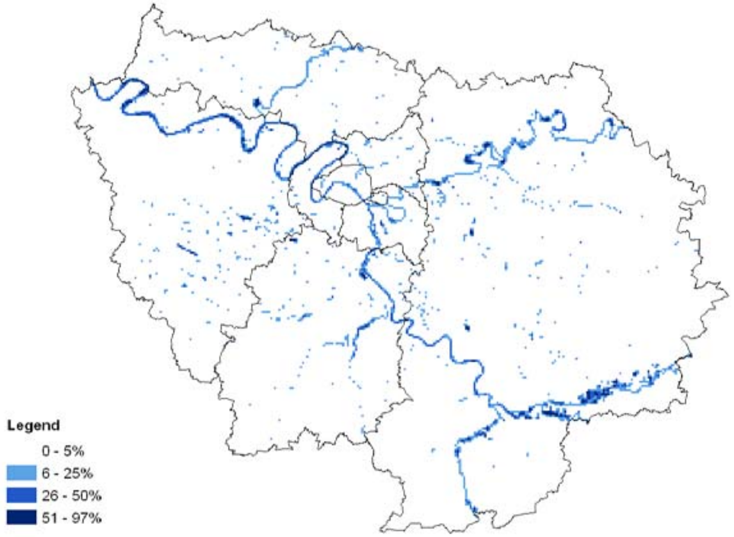


Figure 24: Fraction of cells' surface covered by Water.

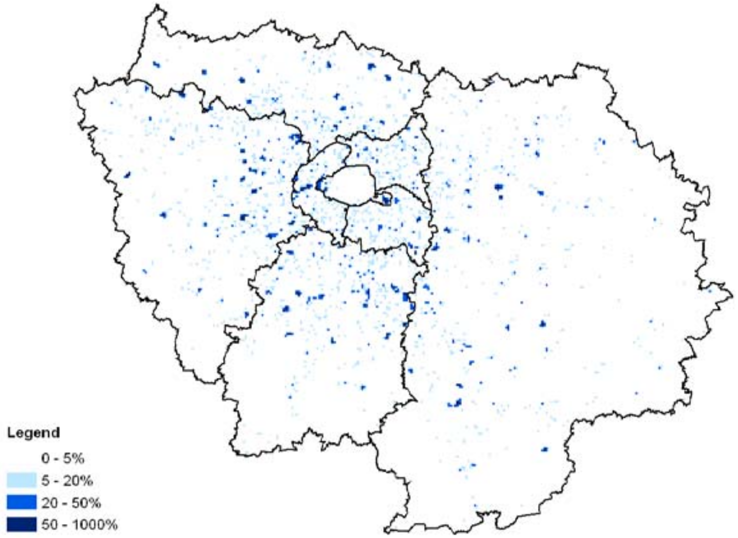


Figure 25: Percent of cells' surface covered by Sporting facilities.

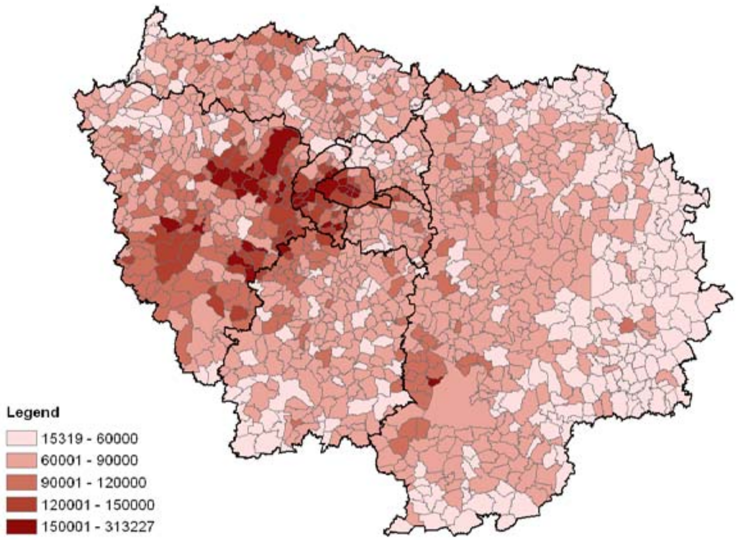


Figure 26: Average price for Flats.

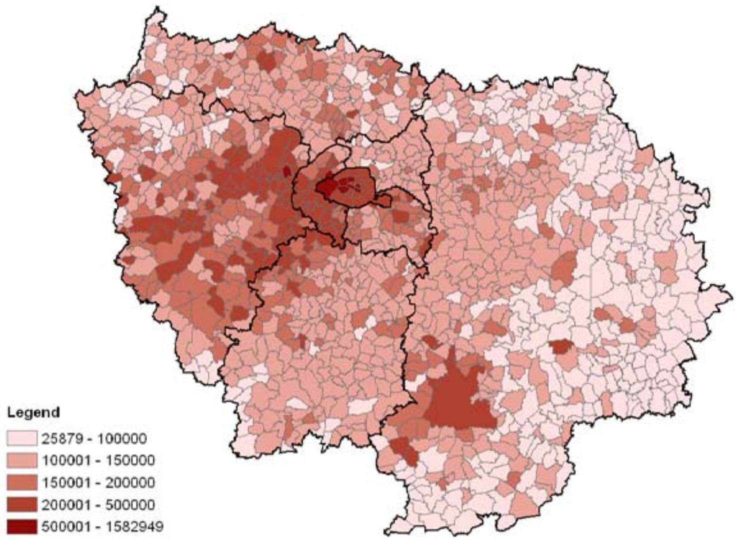


Figure 27: Average price for Houses.

European Transport \ Trasporti Europei n. 36 (2007): 75-91

Diversification activities of passenger railway companies

Emmanuel Doumas [1*]

[1] *PSE-ENPC, Ecole Nationale des Ponts et Chaussées*

Abstract

This paper is about diversification activities, and, more accurately, activities of capitalization of the land and commercial rent generated by railway infrastructures, of passenger railway companies. The global situation of these activities in the different parts of the world in 2003-2004 is first described. Then, the evolution of these activities during the period 1990-2004 is studied. A convergence towards a common diversification strategy and towards activities in stations is observed. Some consequences of the existence of this convergence in terms of transport policies are discussed.

Keywords: Diversification; Passenger railway companies; Station; Rent; Land rent; Railway infrastructure.

1. Introduction, definition, categorization of diversification activities of passenger railway companies and data

a) Introduction

«In Japan, from the early 60s, and the emergence of the car-oriented society, it was understood that, in order to enhance the competitiveness of railways, there were only two ways: high speed trains, and the diversification activities.»

These are words of Mr.Yamanouchi[1], former president of the JR EAST company, the railway company that carries every day the largest number of passengers, in Japan, and in the world. These words show well the fundamental difference that has existed at least from the last 30 years between the passenger railway companies in the different parts of the world regarding their business strategy: While some countries in Europe have developed high speed trains, it seems that it is only in Japan that this second way of supporting railways, that is, the development of the strategy of diversification activities

* Corresponding author: Emmanuel Doumas (emmanuel.doumas@ponts.org).

[1] Interview, Japan, April 2005.

of passenger railway companies, has appeared. The analysis of this "second way" is precisely the subject of the present paper.

The aim of this study is to be an *empirical basis* of numerous economic and financial (linked to the land and commercial rent generated by transport infrastructure, to the financial strategy of the firm, or to the organizational reforms of the different transportation modes, and especially to the railway reforms), urban and geographical (the links between cities and transportation, transportation and sustainable development issues, for example) or even architectural studies (the role of the station in the city, the design of stations and so on...). This empirical basis is aimed to be qualitative (description of diversification activities of passenger railway companies, and its evolution) and quantitative (analysis of the financial results of passenger railway companies by segment of activity). The results will be shown from the early 1990s until 2004, for a wide range of countries and passenger railway companies located all over the world. By doing this, this paper is aimed to cope with a lack in academic research, to the best of our knowledge: until now, there had been no global empirical description, in time and space, of the situation of diversification activities of passenger railway companies. Indeed, a few studies have already raised the question of the diversification activities of passenger railway companies. But they only studied some specific segments of the railway market (one country, one group of passenger railway companies, one railway company, or even one station), and often, in a short interval of time. For example, Aveline (2003), Aoki (1999) have studied the Japanese private passenger railway companies called "*Major*, Killeen and Shoji (1997), the Japanese private passenger railway companies called "*Minor*", Bussolo and Doumas (2005), the European and Japanese stations, Aveline (2003), and Doumas (2003), the Japanese railway companies called "JR", Komatsubara (2004) the diversification activities in stations.

The present study is aimed to be a *global* empirical study because it is wishing to allow the reader to have an overall view of diversification activities of passenger railway companies. By saying this, it is not meant that all passenger railway companies of any part of the world will be covered, nor that the level of accuracy of the description will be the same for all passenger railway companies (data that are currently available at an acceptable cost are not sufficient for such a work[2]). Instead, this study is willing to answer, as far as currently available data can tell us about, to the following questions: Are there some companies or some parts of the world specifically involved in these diversification activities? Concerning their contents and their recent developments, is it possible to observe a move at a global scale?

Consequently, the *reasons* why the present situation of diversification activities of passenger railway companies is such as what is observed will not be the central subject of this study. In particular, because it certainly deserves another study per se, the approach to diversification activities of passenger railway companies in terms of theory of the multiproduct firm, economies of scope etc. will be only briefly discussed.

This study may be of some interest for searchers involved in transportation economics, urbanism, geography, or architecture, but also for representatives of public bodies in charge of transportation policies (and, especially, to those in charge of railway policies), and finally for the representatives of companies involved in transport issues, and of course, more especially for the representatives of railway companies.

[2] Readers who would like to know more about the specific situation of some countries or some companies may also refer to DOUMAS & OKI 2006.

b) Data

The reasons why a global study dealing with the situation and evolution of diversification activities of passenger railway companies has not been achieved yet are certainly the difficult access to data, and the natural barriers of languages in which these data are shown. In this study, the following data have been used:

- Concerning the European companies: the International Union of Railway statistics (UIC), the data of the Yokohama 2004 and Rome 2005 international congresses[3], annual reports of passenger railway companies. Some other data were directly asked to railway companies.

- Concerning the American, African and Asian –except Japan- companies: the UIC statistics, the data of the Yokohama 2004 and Rome 2005 international congresses, and the annual reports of passenger railway companies.

- Concerning the Japanese companies: the UIC statistics, the data of the Yokohama 2004 and Rome 2005 international congresses, the annual reports and *stock book*s of passenger railway companies[44] Some other data were directly asked to railway companies.

c) Definition and categorization of diversification activities of passenger railway companies

This study focuses on the diversification activities of passenger railway companies which are activities of capitalization of the land and commercial rent generated by railway infrastructures (that is to say, stations and tracks). We call *activities of capitalization of land rent* (generated by railway infrastructures) activities the profit of which comes from the renting (lease), or the sale of some land, the price of which depends on the presence of some railway infrastructure. We call *activities of capitalization of commercial rent* (generated by railway infrastructures) activities the profit of which profit comes from the sale of some goods, other than land, at a place where the price of land depends on the presence of some railway infrastructure.

Thus, a real estate business located next to a station belongs to what we call *activities of capitalization of land rent*, but a supermarket business located next to a station belongs to what we call *activities of capitalization of commercial rent*. Let us note that the categorization of a business may depend on the way it is managed by the agent who owns the land where it is located: from the point of view of the railway company that directly operates its shops in the station, a shopping activity is an activity of capitalization of commercial rent, but from the point of view of the railway company that earns a rental lease from an outside company that operates the shop in the station, this same shopping activity is an activity of capitalization of land rent. Let us also note

[3] These congresses were patronized by the UIC, gathering a great number of railway companies. Their subjects were the diversification activities of railway companies and their diversification activities in stations.

[4] *Stock books* (*Yuuka shouken houkokusho souran*) are official documents made by companies listed on the Japanese stock exchange and strictly regulated by laws.

that this definition does not include all cases: the fact that a good is sold here or there may not be connected to the presence of some railway infrastructure. Yet, we think this definition is accurate enough for the purpose of the present study. This implies that the diversification activities that are mainly dealt with in this study are:

- Retail activities in, and next to stations. Retail includes shopkeeping in stations, supermarkets, department store and malls (but also restaurants and cafes) in, and next to stations.
- Real estate activities next to stations and tracks, that may include housing, or office rental lease and sale businesses.
- Leisure activities in, and next to stations. These are hotels, amusement parks, tourism, movie theatres and theatres, private schools and so on.

Concerning retail activities, distinguishing *activities of capitalization of land rent* and *activities of capitalization of commercial rent* allows us to emphasize an interesting aspect of diversification activities of passenger railway companies, which is their two different business models: direct management, or lease from outsiders[5]. In other respects, it will be seen in the next sections that "leisure" activities are not often profitable, and one could thus wonder why some passenger railway companies develop such activities. This is partly due to the fact[6] that these activities are often to be consumed on the place they were bought, and that their consumption takes time (cinema, restaurants and so on). Thus, consumers remain in the station, and may more naturally be encouraged to consume other kinds of goods, such as retail (department store located near restaurants) or train (amusement parks located near stations in the outskirts of a city, thus attracting additional demand for rail trips: this is in fact the general phenomenon of complementary products).

What are called *activities of capitalization of land rent* and *activities of capitalization of commercial rent* in this study cover most diversification activities of passenger railway companies, namely, nearly all diversification activities which are directly or not connected to some passenger railway infrastructure. Thus, diversification activities of railway companies which are not included in this study are diversification activities of freight railway companies (logistics etc.) and activities which have no link with railway infrastructure (for instance, insurance activities).

2. Diversification activities of passenger railway companies: situation in 2003-2004

This section aims to describe as accurately as possible the situation of diversification activities of passenger railway companies in 2003-2004 in the different parts of the world. The analysis, in addition to the description of the different activities, uses as an indicator the percentage of each activity or segment of activity of the passenger railway company in its global operating revenues and operating income. The passenger railway

[5] The distinction between *activities of capitalization of land rent* and *activities of capitalization of commercial rent* may seem fuzzy at a fist glance. It is indeed not usual to the theory of the firm. Nevertheless, we think this distinction is relevant to the present study because, as will be seen in the next sections, it shows particularly well the contrast that exists between Japan and the rest of the world.

[6] Please also see DOUMAS and OKI 2006.

companies are distinguished on the basis of the geographical area they belong to, their size and their statutes. The five geographical areas that are studied here are: Japan, Asia except Japan, America, Africa and Europe. In the case of Japan, three kinds of passenger railway companies are distinguished: “JR”, “*Major*” and “*Minor*” companies[7]. The JR companies are the companies that were born after the privatization of the national railway company JNR in 1987. A characteristic of this privatization process was vertical integration and horizontal separation, and this is why geographically separated companies were created. There are 6 JR passenger companies: JR EAST (Tokyo area), JR CENTRAL (Nagoya area), JR WEST (Osaka area), JR HOKKAIDO (Hokkaido island area), JR KYUSHU (Kyushu island area) and JR SHIKOKU (Shikoku island area). The passenger railway companies others than JR are private companies that were never public companies. Among these companies, 15 are sorted out, these are the so-called “*Major*” companies, the biggest Japanese passenger railway companies except the JR (in terms of operating revenues, number of passengers carried or number of passenger-kms). The private companies that were never public and which are not the “*Majors*” are called the “*Minor*” companies. The “*Minor*” companies that are studied are a group of 31 companies which are the companies listed on the stock exchange we could gather data about (there are more than 100 “*Minor*” companies in Japan).

This categorization of passenger railway companies is partly justified by the number of passengers carried by each group of railway company (for instance, 8 590 M for JR, 6 430 M for *Majors*, and 5 660 M for Western Europe companies). It has to be noted that, since this study deals with activities of capitalization of land and commercial rent, it is indeed the number of passengers, and not the number of passenger-km, which is the key indicator of the size of the passenger railway company. In other respects, this categorization mainly follows the one used by the UIC in its 2002 statistics.

a) Asia except Japan / America / Africa /Eastern Europe

Data concerning passenger railway companies of these areas are scarce. The only available statistics are the UIC statistics. These are 2002 statistics, that do not cover all the countries of these areas, and the only available result by segments is based on the categorization: “Passengers”, “Freight”, “Infrastructure” and “Others”. Activities of capitalization of land and commercial rent are mostly classified in “Others” segment (which, thus, mostly gives an overvaluation of it), but it can also be classified in “Passengers” segment. Thus, in term of categorization of activities, the results, strictly speaking, do not exist (since a real accurate description of diversification activities do not exist, except in the case of AMTRAK –USA). Nevertheless, it can be argued that, looking at the few official sources of information (the web) the only existing diversification activities are shopkeeping and restaurants in stations, and a few real estate businesses. In terms of operating revenues by segments, the results of the Africa / Asia except Japan / Eastern Europe areas are shown here:

[7] JR are inter-cities carriers, local and urban railways. Major are urban railways. Minor are urban and local railways.

Table 1: Operating revenues (OR)[8] by segment of activity, 2002. Asia except Japan /Africa / Eastern Europe (ME: Millions of Euros; sources: UIC, annual reports of companies)

OR, ME 2002	***Passengers***	***Others***
Total Eastern Europe	1 696	148
% total	92	8
Total Africa	303	5
% total	98.4	1.6
Total Asia Except Japan	11 256	243
% total	97.9	2.1

In the case of AMTRAK, activities of capitalization of land and commercial rent ("Food and beverage" + "lease rental") accounts for nearly 8% of total operating revenues ["Passengers" + "Capitalization of land and commercial rent" + "Others"][9]. It can therefore be said that, apart from AMTRAK, diversification activities, and activities of capitalization of land and commercial rent of passenger railway companies in Asia except Japan, America, Africa and, to a less extent, Eastern Europe, are under developed. This is not a surprise, but the present study allows us to give an order of the magnitude of the weakness of this development.

b) Japan

In the case of Japan, numerous and accurate data are available. This can be explained by the fact that Japan is the country where passenger railway companies have the most developed the strategy of diversification of activities. Data are mainly come from the "*stock books*" (*Yuuka shouken houkokusho souran*) of the railway companies.

i) JR Companies

There are two groups of JR Companies: on the one hand, the JR companies that operate their railway network on the island of Honshu, the main island of Japan (there are three of them : JR EAST, JR CENTRAL and JR WEST), which are profitable and listed on the stock exchange, and on the other hand, the three other JR companies, operating networks in the other three main islands of Japan (Hokkaido -JR HOKKAIDO, Kyushu -JR KYUSHU and Shikoku -JR SHIKOKU), which are smaller, often in the red, and not listed on the stock exchange. The segments of activity that are distinguished in the *stock books* are "Transport" (railway, bus, ferry), "Retail" (shops, restaurants, shopping centre, department stores in or near stations, directly operated by child companies), "Real Estate" (office, residential and commercial lease in or near stations), and "Others" (hotel, advertisement, credit cards, tourism, other leisure activities). In terms of operating revenues and operating income by segment, the results are the followings:

[8] OR = sales for the period + change in trade receivables over the period.
[9] AMTRAK Annual Report 2004.

Table 2: Operating revenues (OR) and operating income (OI)[10] by segment. Japan. JR CENTRAL, JR EAST and JR WEST (MY: Millions of Yens; source: *stock books* of JR CENTRAL, JR EAST, JR WEST).

MY (2003)	***Transport***	***Retail***	***Real Estate***	***Others***	*total*
OR outside group	3 728 815	727 690	263 824	373 941	5 094 275
% total OR outside group	73.2	14.3	5.2	7.3	100
OI	672 344	37 658	70 179	30 710	810 895
% total OI	82.9	4.6	8.7	3.8	100

Activities of capitalization of land and commercial rent (segments “Retail”, “Real Estate”, and one part of the segment “Others”) account for 27% of total operating revenues (30% for JR EAST and JR WEST), and 17% of total operating income (26% for JR WEST and JR EAST) of these companies. These figures greatly differ with the ones presented in the previous subsection and show well the special nature of Japanese passenger railway companies.

The data which are available in the case of JR HOKKAIDO, JR KYUSHU and JR SHIKOKU are less accurate. There are official accounting documents, data from the UIC and, sometimes, annual reports, but, since they are not listed on the stock exchange, there is no *stock book*. Only one segment includes all diversification activities (the segment “Others”: retail, real estate, tourism, hotel, and so on). The results are the followings:

Table 3: Operating revenues (OR) and operating income (OI) by segment. Japan. JR HOKKAIDO, JR KYUSHU and JR SHIKOKU (MY: Millions of Yens; sources: annual reports of the companies).

MY, 2003	***Rail***	***Other transports***	***Others***	*total*
OR outside group[11]	247 466	5 027	23 762	276 255
% total OR outside group	89.6	1.8	8.6	100
OI	-44 341	-1 228	7 261	-38 308

It is to be noted that in the case of these companies, diversification activities account for the only positive part of total operating income.

ii) “Major” companies

“*Major*” companies are the 15 most important Japanese private passenger railway companies except JR Companies. These are private companies that were born nearly one hundred years ago.

This study focuses on 14 of them[12]. All of them operate networks located in Honshu, the main island of Japan. These networks are mostly small (between 30 and 300 kilometres), urban (suburbs of Tokyo, Osaka and Nagoya) and carry a great number of

[10] OI = earnings generated by the investment cycle and operating cycle for a given period (in most cases, statements of accounts state OI = OR – operating costs.

[11] “Outside group” means that does not include transactions from one subsidiary company to another subsidiary company of the group, or from one subsidiary company to the parent company.

[12] A financial scandal linked to the SEIBU company occurred a few months ago, showing that some of the figures shown in its stock books were questionable.

passengers every year. The data that are used in this section come from the "*stock books*" of these companies. The segments of activity that are distinguished are "Transport" (railway, bus, taxi, airline, ferry and so on), "Retail" (shops, restaurants, shopping centre, malls, department stores in or near stations, directly operated, by subsidiaries or not, by the railway group), "Real Estate" (office, residential and commercial lease in or near stations), "Leisure" (hotel, advertisement, tourism, amusement parks, cinema, theatre, other leisure activities) and "Others" (construction works, industry, rolling stock industry, information systems and so on). In terms of operating revenues and operating income by segments, the results are the followings:

Table 4: Operating revenues (OR) and operating income (OI) by segment, 2003. Japan. *Major* Companies (MY: Millions of Yens; source: *stock books* of the companies).

MY (2003)	***Transport***	***Real Estate***	***Retail***	***Leisure***	***Others***	*total*
OR outside group	2 326 682	848 901	2 168 881	1 014 323	871 688	7 230 498
% total OR outside group	32.2	11.7	29.9	14.1	12.1	100
OI	266 065	138 528	28 265	1 915	32 240	467 038
% total OI	56.9	29.7	6.1	0.4	6.9	100

Activities of capitalization of land and commercial rent (segments "Retail", "Real Estate" and "Leisure") account for 56% of total operating revenues and 36% of total operating income. This is clearly a different business model, in which the major part of revenues do not come from train ticket sales, but from activities of capitalization of land and commercial rent. Let us note that "Leisure" segment, though it accounts for 14% of total operating revenues, only accounts for 0.4% of total operating income, showing financial results which are, as mentioned before, often in the red.

iii) "Minor" companies

"*Minor*" companies are Japanese private passenger railway companies others than *Major* and JR companies. There are more than 100 of these companies in Japan. Here, only 31 are studied, which are the companies listed in the stock exchange and whose "*stock book*" show operating results by segments. These companies are usually much smaller than JR or *Major* companies (in terms of number of passengers carried or total operating revenues), but the biggest ones carry every year as many passengers as many national companies in the world (for instance, SHIN-KEISEI carried 100 Millions passengers in 2001). The data that are used in this section come from the "*stock books*" of these companies. The segments of activity that are distinguished are the same as in the case of *Major* companies, but the content slightly differs: "Retail" segment do not include many Department stores, and "Leisure" segment includes many "Sports" activities, especially the management of golf-courts. In terms of operating revenues and operating income by segment, the results are the followings:

Activities of capitalization of land and commercial rent ("Retail", "Real Estate" and "Leisure" segments) account for 61% of total operating revenues and 66% of total operating income. "Leisure" segment seems more profitable in the case of *Minor* companies than in the case of *Major* companies, since its stands for 7% of total operating income (10.9% of total OR).

Table 5: Operating revenues (OR) and operating income (OI) by segment, 2003. Japan. *Minor* Companies (MY: Millions of Yens; source: *stock books* of the companies).

MY, 2003	***Transport***	***Real Estate***	***Retail***	***Leisure***	***Others***	*total*
OR outside group	285 059	62 426	389 054	98 742	64 302	899 583
% total OR outside group	31.7	6.9	43.2	10.9	7.1	100
OI	7 070	13 966	3 845	2 253	3 106	30 240
% total OI	23.4	46.2	12.7	7.5	10.3	100

iv) Japan

It is now possible to dress a global picture of diversification activities of Japanese passenger railway companies. In order to build comparable figures, it is necessary to be less accurate, but in spite of this, some strong results can be shown, which are summarized in table 6.

Table 6: Operating revenues and operating income by segment, 2003. Japan (MY: Millions of Yens).

MY, 2003	***Transport***	***Activities of Capitalization of Land and Commercial Rent***	***Others***	*total*
OR outside group	6 593 049	5 597 603	1 309 931	13 500 606
% total OR outside group	48.8	41.5	9.7	100
OI	899 910	303 870	66 056	1 269 861
% total OI	70.9	23.9	5.2	100

Activities of capitalization of land and commercial rent (segments “Retail”, “Real Estate” and “Leisure” of the tables previously seen; “Leisure” activities of JR Companies are not included in these segments) account for 41% of total operating revenues and 24% of total operating income of the Japanese passenger railway companies that were studied in the present study.

It is to be noted that nearly all Japanese passenger railway companies operate activities of capitalization of land rent as well as activities of capitalization commercial rent: some of their subsidiaries directly operate shops in stations, hotels etc[13]. This contrasts with the situation of European passenger railway companies, as will be shown in the next sections.

c) Northern and Western Europe

In the case of Northern and Western Europe, different sources of data are available (2002 UIC Statistics, annual reports of different passenger railway companies, data from 2004/2005 UIC congresses, data from interviews with representatives of passenger railway companies), but, to let figures be comparable, the only 2002 UIC Statistics and different existing annual reports are used in this section. The main problem one faces when trying to collect data about diversification activities of European passenger railway companies is the lack of homogeneity of the different national statements of accounts, especially in the choice of segments. There is no common regulation aiming

[13] For more information about this specific point, please refer to DOUMAS & OKI 2006.

to make these different choices of segmentation be comparable. Thus, each company produces its own segments of activity, making the international comparison very difficult. Some countries show segments such as “Hub development and Operations” (NS, Netherlands) that include retailing in stations, “Property Rental Income” (NR, UK) or even “Real Estate” (CFF, Switzerland), but most statements of account of European passenger railway companies show segments of activity including, with some activities of capitalization of land and commercial rent, a large number of other activities, making the whole analysis quite difficult.

Thus, while, thanks to the different annual reports and statements of account, one can describe accurately the different diversification activities of European passenger railway companies, it is unfortunately very difficult to compare their segments of activity and financial results. Activities of capitalization of land and commercial rent are, in the case of European passenger railway companies, mostly shopping activities in stations. These are not department stores or shopping centers, but different kinds of shops, mostly passenger flows dedicated shops. European passenger railway companies do not directly operate these shops, but they earn a rental lease which is function of the operating results of the shop. Some railway companies, like CFF (Switzerland), also show segments dedicated to Real Estate activities (mostly located around stations). In terms of operating revenues by segment, one can try to make the different figures taken from the different accounting frameworks be comparable, and the results are the following, which are, because of the different uncertainties that exist, to be analysed with caution:

Table 7: Operating revenues and operating income by segment, 2003. Northern and Western Europe (ME: Millions of Euros; source: annual reports of the companies).

OR, 2003, ME	***Transport (passenger)***	***Activities of capitalization of land and commercial rent***	***Others***	*total*
Total “Northern and Western Europe”[14]	23 841	1591	2094	27 526
%	86.6	5.8	7.6	100

d) Global results

Having in mind the different uncertainties of the different results shown before, one can now show global results on the situation of activities of capitalization of land and commercial rent of passenger railway companies in the different parts of the world in 2002/2003/2004. First, most passenger railway companies are involved in only a few sub-segments of diversification activities (renting space for passenger flow dedicated shops in stations). Then, Japan is a very special case: its different kinds of passenger railway companies (JR, *Major* or *Minor*; regardless of their size) operate all kind of diversification activities, from real estate to amusement parks, which account for a large part their total operating results. In terms of operating revenues by segment, global results are the followings:

[14] DB, CFF, FS, RENFE, NS, CIE, DSB, VR, NSB

Table 8: Global operating revenues (OR) by segment, 2003 (sources: please see previous tables).

% OR, 2003	***Transport***	***Activities of capitalization of land and commercial rent***	***Others***
Japan	49	41	10
Asia Except Japan	98	<2 ?	<2?
North America	>85	<7	<8
South America	?	?	?
Africa	98	<2	<2
Western / Northern Europe	87	<6	<7
Eastern Europe	92	<8 ?	<8 ?

The case of Japan could be explained by different geographical or historical reasons: density of population, available land, level of development of road networks, time spent in transport -commuting time- and so on[15]. Yet, many of these reasons may in fact be seen as being themselves linked to the development of diversification strategies by passenger railway companies.

3. Evolution of diversification activities of passenger railway companies from 1990 to 2004

As seen in section 2, activities of capitalization of land and commercial rent of passenger railway companies mostly exist in two parts of the world: Japan, and to a less extent, Europe. Thus, section 3 now studies the evolution of diversification activities of passenger railway companies in Japan and Europe, by showing the evolution of their operating results by segments.

a) Japan

This section studies the case of JR and *Major* companies. This is because, unfortunately, data concerning the evolution of results of *Minor* companies were not available.

i) JR

This subsection analyses the case of Honshu JR companies, which are, as seen before, the JR companies that developed the most diversification activities. Data are from the "*stock books*" of these companies from 1992 (the first JR company –JR EAST- to publish its results by segments did it in 1992) to 2004. The evolution of segments of activities show, from the very beginning, a clear separation between "Transport" and "Retail" segment. On the other hand, "Real Estate" segment appeared late (in 1998 for JR EAST, in 2000 for JR CENTRAL, in 1996 for JR WEST).

[15] A more detailed analysis of these issues can be found in Doumas and Oki 2006.

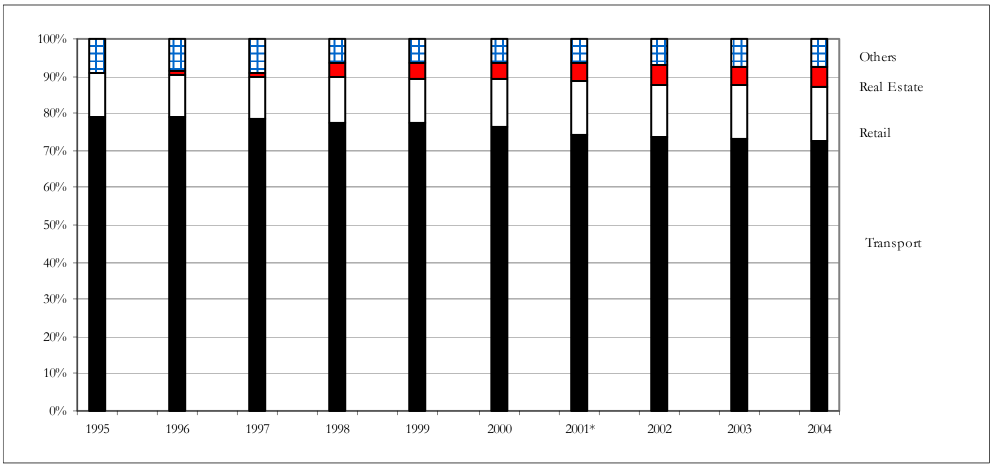


Figure 1: Evolution of operating revenues by segment (% of total operating revenues), 1995-2004, Japan, JR CENTRAL, JR EAST and JR WEST (source: *stock books* of the companies).

These results show that, between 1995 and 2004, the percentage of activities of capitalization of land and commercial rent in total operating revenues (income) has raised up by 7% (14%). It is also to be noted that the "Retail" segment has become profitable with delay, and that the percentage of the "Real Estate" segment in total operating income has raised up from 1% in 1996 up to 9% in 2004. How to explain this evolution? JR companies, when they have been privatized in 1987, have wished to develop a strategy of diversification of activities taking as a model the one *Major* companies had developed for many years. This strategy has led to the development of the segments "Retail" and "Real Estate", but this development took time and it is only from 1992 to 1995 that these segments appeared in the statements of accounts.

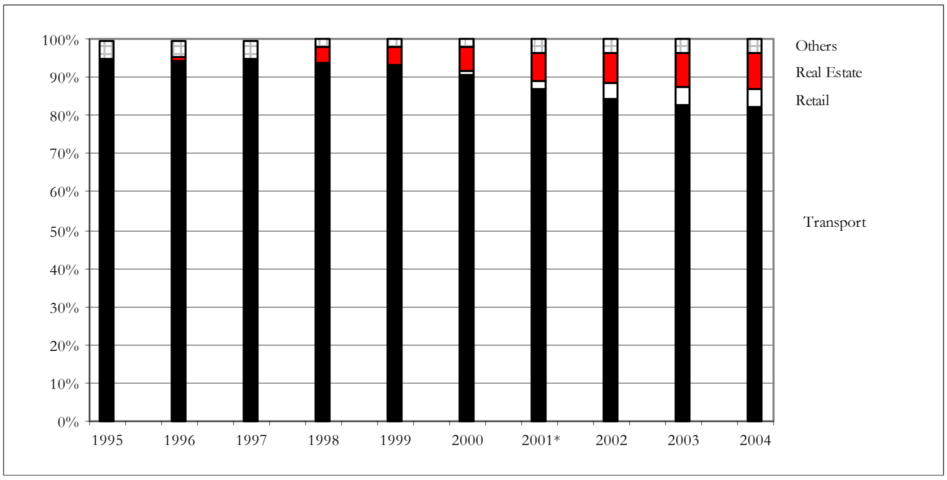


Figure 2: Evolution of operating income by segment (% of total operating income), 1995-2004, Japan, JR CENTRAL, JR EAST and JR WEST (source: *stock books* of the companies).

ii) "Major" companies

Data that were used in this section come from the *stock books* of *Major* companies from 1991 (the *Major* companies began to publish their results by segments in 1991) to 2004. The evolution of the segments of activity of *Major* companies shows that these companies were already operating diversified activities in 1991, and that the segments "Transportation", "Retail", "Real Estate" and "Leisure" already existed in this same year. Nevertheless, it is to be noted that, between 1991 and 2004, 7 out of the 14 companies added to the existing segments, "Retail" (HANKYU, MEITETSU, TOBU) or "Leisure" (HANSHIN, KEISEI, NANKAI, SOTETSU) segments. It means that there has been among these companies a convergence on the period 1991-2004 towards a model of a diversified *Major* company operating three kinds of activities of capitalization of land and commercial rent: "Retail", "Real Estate" and "Leisure" (13 out of 14 in 2004). In terms of evolution of operating revenues by segment, the results are the followings:

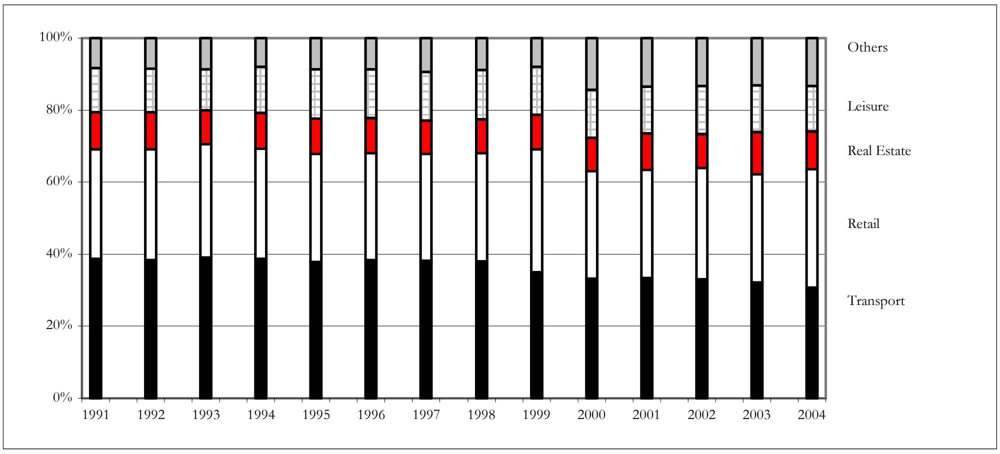


Figure 3: Evolution of operating revenues by segment (% of total operating revenues), 1991-2004, Japan. *Major* companies (source: *stock books* of the companies).

The results[16] confirm that diversification activities had already been developed in 1991 (activities of capitalization of land and commercial rent accounted for 53% of total operating revenues in 1991, and 56% in 2004). Indeed, *Major* companies began to operate such activities one hundred years ago. The evolution of the share of total operating revenues of each of the segments shows a global stability. But, regarding operating income, the percentage of "Real Estate" segment has fallen (from 47% in 1991 to 30% in 2004), while the percentage of "Transportation" segment has raised up (from 34% in 1991 to 52% in 2004). Moreover, "Leisure" segment is in the red 10 years out 14. How to explain this evolution? The fall of the percentage of "Real Estate" segment, and the associated rise of the percentage of "Transportation" segment in total operating income might be explained by the general fall of the prices of land on the period of time studied here (mechanism of speculative bubble). The fall of the percentage of "Transportation" segment in total operating revenues may be explained by the evolution of the Japanese demography in the same period of time, a consequence of which was the fall of the train ridership. Finally, the convergence towards a business model including three segments of capitalization of land and commercial rent ("Retail", "Real Estate" and "Leisure") may be seen as a change from a diversification strategy that was initially, especially in the beginning of the 1990s, strongly oriented towards Real Estate activities (the percentage of the segment "Real Estate" in total operating income reaches 54% in 1995), towards a more diversified diversification strategy,

[16] The evolution of operating income by segment can be found in DOUMAS & OKI 2006

including leisure and retail activities (sports, tourism, shopping center and so on) certainly partly aimed to better lower business cycle related risks. In this more diversified strategy, the role of the station is emphasized (shops, sports activities, movie theatres and theatres and other leisure activities take place in the station).

iii) Japan: global results

In the case of Japan, from 1991 to 2004, global results show a convergence of the diversification strategies of the different passenger railway companies. It is a convergence towards a diversified business model including retail, real estate and, most of the time, leisure activities. It is also a convergence on the relative weights of the different segments of the diversification strategy: More accurately, the diversification strategy has tended to let each of its segments take a more equal part in the global strategy. A consequence of this is that activities taking place in or around stations, which had been previously neglected compared to Real Estate activities, have been taking more and more weight in the global strategy of passenger railway companies.

b) Europe

It was shown in the last section that it is very difficult to make the different statements of accounts of the different European passenger railway companies be homogeneous. It can be therefore easily understood that the description of the evolution of their results by segments is even more difficult. In this section, a panel of European passenger railway companies is studied (SNCF, DB, FS, CFF and DSB): these are the companies for which some time series data could be gathered. Data are from annual reports and other official documents, from the UIC 2004/2005 congresses' presentations, and from interviews with the representatives of some of these companies[17].

In the case of **SNCF**, it has to be noted that the *Direction du Développement des Gares* (stations' department), one of its aims is to develop commercial activities in stations, was created only in 1997, while it was in year 2000 when a subsidiary company called A2C was specially created to enhance the commercial results (retail and services) of stations. Some big stations (Gare du Nord, Gare de Lyon, Gare Saint-Lazare and so on) have started to be refurbished since 2000. The evolution of the operating revenues of different activities in stations from 2000 to 2004 shows a progressive rise of these activities (approximately by 10% for operating revenues).

In the case of **DB**, operating revenues by segment are available from 1999. Their evolution shows a progressive rise of the weight of "passenger stations" activities in

[17] It goes without saying that a more complete study should more accurately describe the present state and the evolution of diversification activities of European passenger railway companies. Readers who are more specifically interested in these companies may refer to DOUMAS&OKI 2006, or to the *Actes du Premier congrès mondial sur l'aménagement des gares « NEXT STATION »*. In particular, the situation of the UK would deserve a more accurate presentation. Nevertheless, the present situation of stations in this country (a change of management responsibility occurred when NETWORK RAIL replaced RAILTRACK ; NETWORK RAIL, the present infrastructure manager, now manages directly only the 17 biggest stations in the UK, the others being managed by local Train Operator Companies) makes the global data collection particularly difficult.

global results. Besides, as in the case of SNCF, lots of stations have been refurbished from 1997 to 2004 (Köln, Hanover, Nuremberg, East Berlin and so on).

In the case of **FS**, two subsidiaries, part of the capital of which belongs to private exterior investors, have been created in order to enhance the development, and especially the commercial development, of stations: GRANDI STAZIONI (development of the 13 largest Italian stations) in 2001, and CENTO STAZIONI (development of the 100 "middle-sized" Italian stations) in 2002. Moreover, another subsidiary, FERROVIE REAL ESTATE, in charge of the management of the real estate assets of the group, has appeared in 2003. The evolution of the percentage of operating revenues and operating income of these subsidiaries in global operating results of FS group shows a rising weight of these activities.

In the case of **CFF**, it has to be noted that there is a "Real Estate" segment of activity in the statements of account, which is in fact also a business segment of the company, and which includes commercial activities in stations as well as real estate activities of the company. It was created in 2003 and was accounting for 14% of total operating revenues and 43% of total operating income ("total" = ["Real estate" + "Passenger"] of the company in this same year.

In the case of **DSB**, in 2002, a "Restaurants and Shops" segment was included in the results by segment of the company, and it was accounting for 4% (and 10% in 2003) of total ([«Restaurants and Shops» + «Passenger and freight» + « Traffic contracts»]) operating revenues in 2002.

In the other European countries, data concerning the evolution of activities of capitalization of land and commercial rent is scarcer. Nevertheless, from the mass of documents and unordered pieces of information that could be gathered, it can be said that a model of commercialization of European stations was born in the beginning of the 2000's. This is only the beginning of the story, and all European passenger railway companies are not yet involved in this move. Nevertheless, it has clearly begun and it can be seen for example in the systematic creation of bodies especially devoted to the management of commercial activities in stations in many European passenger railway companies.

c) Towards a convergence?

Once the evolutions of diversification activities of the different passenger railway companies studied in this paper have been presented, one may wonder if they share common regularities, regardless of the nationality or the category of the passenger railway company. The answer that our study gives to this question seems to be positive. We think indeed that the data that were presented show a convergence of activities of capitalization of land and commercial rent of passenger railway companies towards activities located in stations. In Europe, these are the only diversification activities that have been developed for some years. In Japan, it is more than 10 years since JR Companies focus their strategy of diversification of activities on stations. What's more, *Major* companies, and even if these companies have been diversifying their activities for decades, also tend to focus their attention especially on stations.

The reasons that make these different groups of passenger railway companies focus their diversification strategies towards the station obviously depend on the situation they face. In the case of European passenger railway companies, commercial activities in stations are mostly the only activities of capitalization of land and commercial rent that can anyway be developed (some are public companies; for most of them stations are the major assets they own, except tracks and rolling stocks –whereas some Japanese companies also own much land around stations and tracks[18]). In the case of the Japanese JR Companies, it is the lack of land, in addition to the bursting of the land speculative bubble in Japan at the very moment when these companies were about to undertake their diversification strategy that made them focus their attention on stations. Finally, in the case of *Major* companies, in addition to the reasons previously mentioned, it may also be the wish to minimize business cycle related risks (concerning real estate and shopping activities for example) in the diversification strategy that made it converge towards activities in stations. Thus, the reasons that made each of these groups of passenger railway companies focus their attention on the station are different, but the global result seems to be robust: passenger railway companies have tended to diversify their activities for at least 15 years, and part of the direction of this diversification process seems to be the station.

There are several consequences to this fact. First, it should give incentives to passenger railway companies or countries that have not yet undertaken the systematic development of commercial activities in stations to observe and study the foreign experiences in this field and its results[19] . Then, it should give incentives to the representatives of public bodies, and to searchers of the different fields involved in studies linked to stations (transport, economics, urbanism, geography, architecture) to focus their attention on the consequences of this new empirical fact -which seems to be robust-, that is, the systematic development of commercial activities in stations and, more generally, the systematic development of activities of capitalization of the land and commercial rent generated by railway infrastructures.

Conclusion

In this study, the present situation, and the evolution of diversification activities of passenger railway companies in the different parts of the world have been presented. The particular case of Japan, and its different groups of passenger railway companies (JR, *Major* and *Minor*), has been emphasized. Besides, a convergence of diversification activities towards activities in stations in Japan and Europe has been shown and discussed.

Some important questions remain: first, what is specifically due to economic mechanisms (and especially agglomeration mechanisms), in the move observed towards commercial activities in stations? What are the link between phenomena of capitalization of land and commercial rent generated by railway infrastructures, by passenger railway companies, and the privatization process of these companies, or, more generally speaking, the change of ownership of these companies? The

[18] Thus, the move that is observed may in some cases be a partly regulation induced one (if, for example, a public company is ordered to get rid of some of its assets, including land, when privatized).

[19] It is to be noted that, in 2004 and 2005, the first two international congresses of the history of rail dealing with diversification activities -and especially with commercial activities in stations- of railway companies, have taken place (UIC Yokohama 2004 and Rome 2005 congresses).

consequences, in terms of public economics, of the answers given to these questions may be of the highest importance.

Finally, a generalization of the method and contents developed in this study is obviously possible for other means of transportation. While transport infrastructures differ, most of them generate land and commercial rent, which will be, or not, capitalized by transport operators. The cases of subway stations and airports seem to be the closest to the railway infrastructure one. But gas stations, ports, or shops that can be found in ships, may also be interesting subjects of research. In fact, until now, the only barrier to this general study of the land and commercial rent generated by transport infrastructures is the access to data.

Acknowledgments

Mr Makoto AOKI. Associate Professor. Faculty of Business Administration, Tokyo Keizai University; Mr. Kiyotaka YUGUCHI. Assistant Professor. Sagami Women's University. Tokyo; Mr. David MEUNIER. PSE-ENPC, Ecole Nationale des Ponts et Chaussées; Mr. André DE PALMA. Professor. Université de Cergy Pontoise, ENPC; Mr. Philippe ROUMEGUÈRE. Former chairman. UIC; Mr. Pascal LUPO. Head of Direction des Gares et de l'escale, SNCF; Ms Kae DEBAYLES-OKI Parvis. Comments received by two anonymous referees were also very helpful.

References:

Aoki, M. (1999) *Ootemintetsuto chihoumintetsuni miru keieitakakukano bunsekihikaku – kantouno otemintetsuto hokurikuno chihoumintetsuwo reinishite (A comparative analysis of Business diversification on private railway companies in Tokyo Metropolitan and Hokuriku region),* nihonkai chiiki kenkyu center

Aoki, E., Imashiro, M., Katô, S. and Wakuda, Y. (2000) *A History of Japanese Railways: 1872-1999*. East Japan Railway Culture Foundation

Aveline, N. (2003) *La ville et le rail au Japon -l'expansion des groupes ferroviaires privés à Tôkyô et Ôsaka*. Paris, CNRS éditions, collection Asie Orientale

Bussolo, M. and Doumas, E. (2006) "Evolution of the role of the railway station within the organization of railway companies", *Actes du Premier congrès mondial sur l'aménagement des gares «NEXT STATION»*

Doumas, E. and Oki, K. (2006) *Activités de diversification des entreprises de chemin de fer : Une convergence vers la gare ?,* Editions Techniques Ferroviaires, 133p

Doumas, E. and Oki, K. (2006) *Les grandes gares dans les grandes villes. Situation actuelle de leurs configurations spatiales et offres commerciales a Osaka, Londres et Pari*s, Editions Techniques Ferroviaires, 49p, Avril 2006

Killeen, B. J. and Shoji, K. (2001) "Diversification strategy of the minor private railway companies in Japan", *working paper*

Komatsubara, M. (2004) "Successful Business Models for Railway Non-fare Box Businesses", *Joint Business Opportunities UIC Seminar Yokohama 2004*

Saito, T. (1997) "Japanese Private railway companies and their business diversification", *Japan Railway&Transport Review*

European Transport \ Trasporti Europei n. 36 (2007): 92-106

Quality and public transport service contracts

Valerio Gatta [1*], Edoardo Marcucci [2]

[1] *Dipartimento di Statistica, Probabilità e Statistiche Applicate (DSPSA)*
Facoltà di Scienze Statistiche, Sapienza Università di Roma
[2] *Istituto di Scienze economiche, matematiche e statistiche (ISEMS)*
Facoltà di Economia, Università degli studi di Urbino

Abstract

Service contracts are the natural method to set bilateral commitments. In transport service context, public authorities and transport operators have different goals, therefore regulation plays an important role especially failing competition. After a brief description of the most important regulatory procedures, we focus our attention on the quality framework in service contracts. In recent years the inclusion of quality requirements in contracts is becoming common practice, especially when adopting price cap regulation. This paper suggests a criterion for service quality definition, measurement and integration in contracts for the production of socially valuable transport services. Using choice-based conjoint analysis to analyse customer preferences we estimate the passengers' evaluation of different service features and calculate a robust specification of a service quality index from the customers' point of view. A case study demonstrates the procedure to follow for measuring service quality in local public transport differentiated by geographical service segments.

Keywords: Service quality; Stated preferences; Service contracts.

Introduction

Public authorities and transport operators are both involved in the provision of public transport services. There is a contrast between the social goals of the former and the private ones of the latter. Private firms maximise profits without considering social welfare. Regulation plays an important role especially failing competition. Service contracts are the natural method to set bilateral commitments. A contract between the authority and the operator constitutes the instrument to induce firms in naturally non-competitive markets to act in line with social targets. Only in a few countries in Europe, the relation between authorities and transport operators are not regulated by a service contract.

The question of regulatory procedures has generated an extensive literature. The traditional rate-of-return (ROR) regulation has been examined by many authors (Averch

* Corresponding author: Valerio Gatta (valerio.gatta@uniromal.it)

and Johnson, 1962; Baumol and Klevorick, 1970; Bailey, 1973; Das, 1980 and others) who agreed that ROR induces firms to produce inefficiently causing damages to consumers. Other regulatory procedures have been developed. Sappington and Sibley (1988) proposed the "incremental surplus subsidy scheme" which induces a subsidized firm in a natural monopoly to price at marginal cost and use the cost-minimizing input mix. Various authors recognized the importance of billing algorithms as a potentially strategic key for increasing social welfare. Boiteux (1960), Williamson (1966) and others identified the optimal time-of-use prices. Willig (1978) and Panzar (1977) formalized a regulatory procedure including multipart and self-selecting tariffs. Another famous regulatory procedure is the price cap system where price is set by the regulator and is adjusted over time (Acton and Vogelsang, 1989) leading to more specific investment in cost-effective innovation (See Train, 1991 for a wide literature review).

Public transport has long been dominated by a production-oriented approach, but it is now progressively moving towards a more customer-oriented one and in recent years the inclusion of quality requirements in contracts is becoming common practice. In this paper we focus our attention on the quality framework in service contracts and following Hensher *et al.* (2003) we suggest a criterion for service quality definition, measurement and integration in contracts for the production of socially valuable transport services. This contractual context is becoming more relevant since in recent contributions (Bergantino *et al.*, 2006) there is a clear and specific reference to service quality factors in regulatory schemes based on price cap. Bergantino *et al.* (2006) specifically refer to a price-quality cap system.

This paper is structured as follows. Section 2 illustrates the various approaches developed to tackle the problem of quality definition and measurement, stressing the advantages connected to the approach adopted in this paper. Section 3 describes how the method proposed could be used in the context of contractual definition of quality when preparing a public transport service contract. Section 4 shows a case study that demonstrates the procedure to follow for calculating a service quality index (SQI). Finally section 5 proposes some concluding remarks.

Measuring Service Quality

The issue of quality is contentious. Although it is recognised as a key management tool, it still remains a fairly subjective concept. Quality is often related to the notion of standards, but in many cases the existing standards are linked to performance determinants which are not very important for the customer. We reject the resulting assumption of different kinds of quality, such as "expected" and "perceived" quality or "targeted" and "delivered" quality and believe that there is only one sort of quality and it must be strongly user-oriented, that is, based on customer preferences.

A second problem concerns the measurement method. Difficulties arise from the specific and subjective nature of services. The distinctive characteristics of intangibility, heterogeneity, inseparability and perishability make services unique and different from goods and thus rendering service quality evaluation more complicated than manufacturing quality control. The most popular tools are basically customer satisfaction surveys in which respondents are asked to evaluate quality factors one at a time. Data are generally analysed by multivariate statistical techniques like factor analysis, principal components, regression or structural equation models. SERVQUAL, proposed by Parasuraman, Zeithaml and Berry (1988), is the method that has attracted the greatest attention. It is a multiple-item scale for rating both the expectations and the perceptions of the service performance on a seven-point Likert scale. They measure service quality by means of the disconfirmation model, calculating the degree and direction of discrepancy between consumers' perceptions and expectations about different dimensions of the service. Other famous methods are, for example, SERVPERF (Cronin and Taylor, 1992), Normed Quality (Teas, 1993) and Zone Of Tolerance (Zeithaml *et al.*, 1993).

The intent to overcome some critical factors pertaining to the above methods like conceptual basis, psychometric problems or troubles with the usage of Likert scales such as the well-documented tendency for respondents to choose central response options rather than extreme ones, the impact of the number of scale points used, the influence of the format and the verbal labelling of the points and the transformation from ordinal data to cardinal data, induced us to search for a new approach for measuring service quality. Following Hensher *et al.* (2003) we adopt an alternative approach with the same level of general appeal (Gatta, 2006).

First of all, quality is linked with the concept of utility gained by the consumers. Every service implies a certain level of utility depending on its characteristics. The higher is the level of quality delivered, the greater is the corresponding utility. Another crucial point is the assumption that individuals' preferences are captured by utility functions. The higher is the utility level of a service, the greater is the probability that a consumer chooses that service.

In order to represent service quality as determined by consumers, we suggest to employ a stated preference (SP) survey in which individuals are asked to choose, according to their preferences, among a set of options. The basic idea is that users buy a package of service characteristics (attributes) when deciding to travel on a bus. In particular, we recommend choice-based conjoint analysis[1] (CBCA), a decompositional method that estimates the structure of consumers' preferences given their choices between alternative service options (Mc Fadden, 1974; Louviere and Woodworth, 1983). Such method was originally developed in marketing research field with the objective to identify the structure of customers' preferences for available or not yet available products on the market. Respondents typically observe profile descriptions of two or more products and pick the most preferred from the set. The flexibility and the rich information that can be gathered by using CBCA, have allowed its application also in transport, environment and medicine.

CBCA asks the agent to explicitly choose among the profiles, thus mimicking actual market choice, rather than rating or ranking alternatives. This is the characteristic distinguishing CBCA from other types of conjoint analysis. CBCA provides less information compared to the other two methodologies, but it is also easier for the agents to understand and respond to the choices proposed, since it reproduces a context similar to that they are, in reality, accustomed to. In fact they are asked to compare a set of alternatives and select the one providing the highest utility. Furthermore, this method does not require any assumptions to be made about order or cardinality measurement (Louviere, 1988).

[1] The seminal paper is Mc Fadden (1974).

The main drawbacks can be related to the increasing burden of information that respondents have to process before choosing, when the number of attributes rises[2] as well as to the reliance on people's stated intentions, as it occurs whenever making use of questionnaires. However, in our case, we are not precisely interested in estimating demand curves, rather to identify the relative weights of the various attributes in determining quality levels.

In CBCA, the options in each choice set are constructed in terms of levels of different service attributes and designed by the researcher. The package of service attributes with the highest utility is chosen. Therefore, through the users' conjoint evaluations of the attributes, and thus through their choices, we are able to estimate utility functions and identify the relative importance of the relevant quality attributes. Besides, by means of this method, we are able to determine the global satisfaction (or utility) that a passenger obtains from the actual service and how this might change under alternative service level delivered, as well as the contribution of each elemental attribute to the overall service quality level (Hensher and Prioni, 2002). This method is more reliable than those in which attributes are evaluated one at a time (e.g. SERVQUAL, SERVPERF, Normed Quality, Zone Of Tolerance) because the data gathered from the latter lack the information about trade-offs between attributes.

The major theoretical aspects are now briefly recalled. According to random utility theory (RUT) proposed by Thurstone (1927), utility is modelled as a random variable in order to reflect the assumption that the decision-maker has a perfect discriminative capability, while the analyst has incomplete information (Ben Akiva and Lerman, 1985) deriving from unobserved alternative attributes, unobserved individual characteristics or measurement errors (Manski, 1977). The utility that individual q associates with alternative i is given by

$$U_{iq} = V_{iq} + \varepsilon_{iq}, \tag{0.1}$$

where V_{iq} is the deterministic part of the utility and ε_{iq} is the random term. The deterministic term is a linear in the parameters function of the attributes of the alternatives

$$V_{iq} = \overline{\beta}\overline{X}_{iq}, \tag{0.2}$$

where $\overline{X}_{iq}$ is the vector of attributes as perceived by individual q for alternative i, and $\overline{\beta}$ is the vector of related parameters[3].

Mc Fadden (1974) supposed that an individual facing a finite choice set selects the alternative that maximizes utility. He proposed a probabilistic approach where the probability that individual q chooses alternative i from choice set C (J alternatives) is

[2] However, according to Hensher (2004), cognitive burden doesn't come from the increase of information that respondents have to process due to the product of the number of attributes and number of alternatives associated with each choice set, on the contrary limited information may in itself be especially burdensome where it is an incomplete representation of the attribute space that matters to an individual. He found that choice complexity is linked with the relevancy issue.

[3] The bar in $\overline{\beta}$ and $\overline{X}$ represents a vector, although such a bar usually indicates a mean value.

$$\begin{aligned} P_q(i \mid C) &= P[(\varepsilon_{jq} - \varepsilon_{iq}) < (V_{iq} - V_{jq})], \quad \forall j \neq i \\ &= \int_{\varepsilon} I(\varepsilon_{jq} - \varepsilon_{iq} < V_{iq} - V_{jq}; \forall j \neq i) f(\varepsilon_q) d\varepsilon_q \end{aligned} \tag{0.3}$$

where $f(\varepsilon_q)$ is the joint density of the random vector $\varepsilon_q = (\varepsilon_{1q}, \ldots, \varepsilon_{Jq})$, $I(\bullet)$ is the indicator function equalling 1 when the expression in parenthesis is true and 0 otherwise. That probability is a multidimensional integral over the density of the unobserved portion of utility. Equation (0.3) is referred to as a random utility model (RUM) explanation of observed choices. Different assumptions about the distribution of the random term imply different discrete choice models that can be used to analyze the gathered choice data with the purpose of estimating the β-parameters and calculating a SQI.

Service quality, service contracts and incentives

A key element involves the inclusion of quality in contracting procedures. We assume that the service contract between the authority awarding it and the operator producing the service should explicitly foresee a minimum level of service quality measured in terms of a given level of SQI. The intention is to achieve the best possible service from the user's point of view according to operator's capabilities to fulfil the requirements specified by the authority.

The question of how to establish a minimum SQI level is not trivial. The task is complicated by the significant asymmetry in the available information. Usually, regulators dispose of less information about the firms' cost and demand function than do the firms themselves and hence do not know whether the firm is able to provide the targeted quality level or the exact cost it has to bear. However, regulators know that firms act in order to maximize their own profit (Train, 1991). When contracting there is an important distinction between *ex-ante* quality, as stipulated in the contract, and *ex-post* quality, after the contract is let. After fixing a SQI target, one has to start a monitoring system so that the quality of the service supplied can be kept under control. In order to get good results it is essential that the SQI monitoring system is assigned to an independent party.

Strong incentives result from the appropriate definition and measurement of a SQI. In fact, a better service quality produces higher user satisfaction and a more attractive service that implies new customers and consequently higher revenues for the operators who may invest them to improve service quality (QUATTRO Research Consortium, 1998). SQI targets should be modified over time according to individual preferences variations.

Empirical measurement of a SQI

Case study description

In this paragraph we illustrate a case study conducted in five geographical areas of the *Marche* a region in central Italy. We describe a procedure for measuring and integrating service quality in local public transport contracts. The project was carried out in collaboration with the local transport operator. The interview we prepared was composed of two sections: in the first one respondents were asked to provide information about their current trip, referred to as revealed preference (RP) data, and, in addition, about their socioeconomic characteristics; in the second one the interviewee had to make repeated choices between three alternatives, one representing her current trip (status quo) and two hypothetical trips (different bundles of trip attribute level), referred to as SP data. The literature identifies a set of attributes (see, for example, Hensher *et al.*, 2003; Friman *et al.*, 2001; Tranportation Research Board, 1999) and, in particular, we initially considered a group of 18 and we asked people to assign them a degree of importance[4]. Based on focus groups with customers and local operators we ended up with the following five attributes as the most appropriate dimensions to characterize service quality from a user's perspective[5]: bus fare (Cost); amount of delay at bus stop (Delay); bus travel time (Trip Length); bus frequency - number of buses per hour (Frequency); amount of time between service inception and service closure (Availability).
The attribute levels were selected as percentage changes from the status quo (figure 1). Even though this choice may introduce an endogeneity bias (De Palma and Picard, 2005) and implies complexity and cumulative cognitive burden for respondents, it anchors attributes to current experience. Respondents choose within a more realistic choice set and are not forced to choose one of the hypothetical alternatives they would never choose had they the opportunity to do so. Anchoring the SP exercise to the RP choice both avoids poor quality and inappropriate responses and provides an escape to the "no-reply" problem (Stopher, 1998; Hensher *et al.*, 2003; Bradley, 1993; Carson *et al.*, 1994; Louviere *et al.*, 2000)[6]. Through a formal experimental design, that is a full profile random design[7], the attribute levels were combined into bus options and we constructed 8 choice sets per interview. One of these choice sets had a control function, in fact it was formed by three fixed-design alternatives: the best possible one, the worst one and the current one. To allow for a rich variation in the combination of attributes

---

[4] The 18 attributes were: bus travel time, bus fare, bus delay, bus frequency, bus availability, information at bus stop, general cleanliness on board, seats at bus stop, cover bus stop, driver kindness, capability to cover the territory, availability of tickets, safety on board, seat availability on bus, opportunity to purchase different kinds of tickets, time to reach bus stop by car, access to bus facilities, time walking to bus stop.
[5] We considered only the first five attributes in terms of degree of importance in order to avoid cognitive burden and fatigue for the interviewee. According to Aaker and Day (1990) in SP experiments four or five attributes are usually selected.
[6] In order to avoid forcing respondents, alternatively, one should include a "no choice" or a "don't know" option in the choice set. However, it may lead to a serious loss of information.
[7] We use the "Shortcut" method of design generation included in Sawtooth Software's CBC product, in which profiles for each respondent are constructed using the least often previously used attribute levels for that respondent, subject to minimal overlap. Each one-way level frequency within attributes is balanced (Sawtooth Software, 1999).

levels we prepared 25 different versions of the survey form[8]. Overall for the five geographical service segments, we administered 264 interviews either on board or at the bus stops associated with the main routes. We used a random sampling strategy to select the sample[9]. The minimum acceptable sample size *n* is determined by the desired level of accuracy of the estimated probabilities about the proportion of decision makers that choose an alternative. In our case, we have $k = 3$ alternatives, therefore it requires for a simultaneous confidence statement of specified precision about the parameters of a multinomial population. According to Tortora (1978), the minimum sample size is given by the following formula:

$$n = \frac{Bp_i(1-p_i)}{\vartheta^2}, \tag{0.4}$$

where B is the upper $(\alpha / k)x100th$ percentile of the χ^2 distribution with 1 degree of freedom for a specified confidence level $\beta = 1-\alpha$; p_i is the true proportion closest to 0,5; ϑ is the percentage level of allowable deviation between the estimated and the true proportion (that is the precision level). If there is no prior knowledge about the values of p_i's, the most precautionary situation in terms of sample size can be made assuming some $p_i = 0,5$.

For $\beta = 95\%$ and $\vartheta = 6\%$, the minimum sample size suggested would be 398. However, each respondents contributes a number of observations, that is, each individual does $r = 8$ choice scenarios. In practice, we may obtain the required n choices from n/r respondents, but this holds only if decision makers treat each choice occasion as being an independent decision task. In the latter case, we would end up with about 50 individuals[10].

For econometric analysis, we ignored all the interviews in which the agents failed to answer correctly the control choice exercise (14 over 264, net used 250).

Delay	Cost	Trip length	Frequency	Availability
+100%	+50%	+50%	+50%	+20%
+50%	+25%	Same as now	Same as now	+10%
Same as now	Same as now	-50%	-50%	Same as now
-50%	-25%			-10%
-100%	-50%			-20%

Figure 1: attributes and their levels selected for the choice experiments.

[8] In this study we employed paper-and-pencil interviews. We are presently constructing computer-based interviews in which attribute levels are still selected as percentage changes from the status quo but they are expressed in the choice profiles as cardinal numbers. Basically, we are producing a software for customized interviews.

[9] We do not have information about the demand levels for each area because the service company itself does not know how many people are using the bus.

[10] The sample size for each area is equal or greater than 45, the exact amount requested in case of equidistribution of the proportions (when $p_i = 0,\overline{3} \quad \forall i = 1,2,3$).

Revealed Preference data

Table 1 provides information about the socioeconomic variables we included in the survey for each geographical segment.

Table 1: socioeconomic data by area.

SOCIOECONOMIC VARIABLES		***AREA*** 1	2	3	4	5	*All*
GENDER	Female	60,0%	53,3%	55,3%	45,0%	55,3%	53,8%
	Male	40,0%	46,7%	44,7%	55,0%	44,7%	46,2%
AGE	<30	95,4%	55,6%	53,2%	71,7%	61,7%	69,7%
	30-50	3,1%	31,1%	29,8%	21,7%	27,7%	21,2%
	>50	1,5%	13,3%	17,0%	6,7%	10,6%	9,1%
NATIONALITY	Foreign	6,2%	-	12,8%	6,7%	14,9%	8,0%
	Italian	93,8%	100,0%	87,2%	93,3%	85,1%	92,0%
INCOME (€ annual)	<2500	72,2%	45,8%	37,0%	41,7%	42,3%	45,8%
	2500-5000	22,2%	12,5%	22,2%	27,8%	19,2%	21,4%
	5000-10000	-	12,5%	14,8%	8,3%	23,1%	12,2%
	10000-15000	5,6%	12,5%	11,1%	5,6%	3,8%	7,6%
	15000-20000	-	12,5%	3,7%	8,3%	7,7%	6,9%
	>20000	-	4,2%	11,1%	8,3%	3,8%	6,1%
MAIN OCCUPATION	Employed full time	-	33,3%	19,1%	16,7%	21,3%	16,7%
	Self-employed worker	3,1%	4,4%	6,4%	3,3%	4,3%	4,2%
	Student	75,4%	33,3%	36,2%	60,0%	38,3%	51,1%
	Student-worker	20,0%	8,9%	8,5%	13,3%	8,5%	12,5%
	Retired or pensioner	-	6,7%	12,8%	3,3%	4,3%	4,9%
	Unemployed	-	4,4%	4,3%	-	10,6%	3,4%
	Housewife	1,5%	8,9%	4,3%	3,3%	6,4%	4,5%
	Other	-	-	8,5%	-	6,4%	2,7%
FREQUENCY OF BUS USAGE	Almost every day	46,2%	37,8%	42,6%	8,3%	42,6%	34,8%
	1-3 times a week	29,2%	22,2%	27,7%	13,3%	25,5%	23,5%
	Once every two weeks	9,2%	13,3%	8,5%	58,3%	19,1%	22,7%
	Rarely	15,4%	26,7%	21,3%	20,0%	12,8%	18,9%
TRIP PURPOSE	Study	87,7%	28,9%	34,0%	38,3%	27,7%	46,2%
	Work	4,6%	26,7%	27,7%	15,0%	38,3%	20,8%
	Leisure	4,6%	28,9%	12,8%	33,3%	23,4%	20,1%
	Other	3,1%	15,6%	25,5%	13,3%	10,6%	12,9%
AVAILABILITY OF ALTERNATIVE MEANS OF TRANSPORT	None	63,1%	26,7%	31,9%	60,0%	44,7%	47,3%
	Car	18,5%	33,3%	44,7%	31,7%	25,5%	29,9%
	Motorcycle	1,5%	6,7%	2,1%	-	8,5%	3,4%
	Scooter	10,8%	13,3%	10,6%	3,3%	17,0%	10,6%
	Other	-	-	4,3%	-	-	0,8%
	More than one	6,2%	20,0%	6,4%	5,0%	4,3%	8,0%
TOTAL RESPONSES (N)		65	45	47	60	47	264

The sample is composed of 142 females and 122 males. In all but area 4 males are less than females. Although the average age is 30, the 69,7% of the sample is under 30 (95,4% in area 1). About half of the sample is composed by students and the annual income is under 2.500€ (about 75% in area 1). Almost all of the sample is made up of Italians. As regards the frequency of bus usage, the most representative class (34,8%) is the first one, that uses the bus "almost every day", except for area 4 where "once every two weeks" has a share of 58,3%. The data just described suggest that the sampled people have knowledge of the service delivered and the conclusions drawn from their responses should be considered reliable. Study is the main trip purpose (46,2%) while in half of the cases there is no availability of alternative means of transport so the bus is a forced choice.

Table 2 reports the average attribute levels associated with the current trip and the relative average cutoffs by segment.

Table 2: RP data on current trip: average attribute levels and relative average cutoffs.

	AREA					*All*
	1	2	3	4	5	
ATTRIBUTES IN CURRENT TRIP						
COST (€)	0,65	0,73	0,71	2,05	0,99	1,05
St. dev.	0,23	0,12	0,15	0,49	0,33	0,63
[min; max]	[0,13; 0,80]	[0,38; 0,80]	[0,19; 0,83]	[0,45; 3,20]	[0,40; 2,75]	[0,13; 3,20]
DELAY (minute)	1,80	2,79	1,74	1,84	1,49	1,92
St. dev.	1,67	2,22	1,78	3,26	2,20	2,34
[min; max]	[0; 6]	[0;10]	[0; 5]	[0; 15]	[0; 10]	[0; 15]
TRIP LENGTH (minute)	9,14	10,79	14,37	48,71	21,21	21,33
St. dev.	4,42	4,56	6,56	13,16	3,79	16,88
[min; max]	[2; 20]	[3; 20]	[5; 30]	[8; 60]	[10; 30]	[2; 60]
FREQUENCY (n°buses/h)	2,9	2,0	1,7	1,1	1,9	1,9
St. dev.	0,99	0,50	0,49	0,41	0,25	0,90
[min; max]	[0,67; 4]	[1; 4]	[1; 2]	[1; 4]	[1; 2]	[0,67; 4]
AVAILABILITY (minute)	796	816	772	853	857	818
St. dev.	59,30	70,14	40,38	95,57	49,83	89,44
[min; max]	[720; 1055]	[720 1055]	[615; 840]	[720 1080]	[690; 960]	[615; 1080]
ATTRIBUTES CUTOFF						
COST_Cutoff	0,98	1,10	1,11	2,86	1,41	1,52
	(+50%)	(+51%)	(+57%)	(+40%)	(+42%)	(+45%)
DELAY _Cutoff	10,87	9,35	11,13	12,86	10,71	11,07
	(+504%)	(+235%)	(+539%)	(+597%)	(+617%)	(+477%)
TRIP LENGTH _Cutoff	16,80	17,53	21,92	61,96	31,40	30,50
	(+84%)	(+63%)	(+53%)	(+27%)	(+48%)	(+43%)
FREQUENCY_Cutoff	1,5	1,1	0,9	0,8	1,2	1,1
	(-47%)	(-42%)	(-47%)	(-21%)	(-37%)	(-41%)
AVAILABILITY_Cutoff	628	677	551	685	664	641
	(-21%)	(-17%)	(-29%)	(-20%)	(-23%)	(-22%)

The fractions of the variables expressed by minutes are rounded as general numeric type variables.
In brackets the percentage variations with respect to the actual levels.

Before commenting the results in Table 2, we briefly clarify the meaning of cutoffs. A cutoff is a self-imposed constraint by the decision maker. According to Swait (2001) we have extended the traditional compensatory utility maximization framework incorporating attribute cutoffs into the decision problem formulation. We asked people to state their upper bounds for variables which have a negative impact on utility and lower bounds for those with positive impact on utility. Then allowing respondents to violate the self-imposed constraints at a potential cost leads to the formulation of a penalized utility function and implies non-linearities in the preference function.

As we can see in Table 2, the average bus fare is approximately 1€ while on average the maximum level of bus fare that respondents are willing to pay is about 1,50€. The high fare experienced in area 4 is due to the fact that user in that area don't often buy monthly tickets[11]. The average delay is about 2 minutes while the cutoff is about 11 minutes. If we look at the percentage increase of delay and compare it with the other percentage variations, we find a surprising result: while agents are willing to accept an increase of 477% of current delay, on the other hand, the increase (or decrease) in other attributes is never greater than 50%. This fact reflects an underestimation of the actual delay because of the large number of interviews that were administered at the terminus. Bus travel time is around 21 minutes while the cutoff is about 30 minutes. The number of buses per hour is about 2 while the cutoff is not much greater than 1. Finally, the time interval between the first and the last bus is 818 minutes and users are willing to accept a decrease till 614 minutes.

Econometric results

Now we turn our attention to the issue of parameter estimation. We may obtain information about the relative importance of the attributes using discrete choice models. Multinomial logit (ML), is derived from the assumption that the error terms of the utility functions are independent and identically Gumbel distributed. The choice probabilities of ML are expressed as follows

$$P_q(i \mid C) = \frac{e^{\lambda \bar{\beta} \bar{X}_{iq}}}{\sum_{j=1}^{J} e^{\lambda \bar{\beta} \bar{X}_{jq}}} \,, \qquad (0.5)$$

where λ is the scale parameter inversely related to the variance of the error term[12].

We have data from 5 different areas so we cannot simply estimate 5 different MLs. In order to make meaningful comparisons between attributes parameters associated with different geographical segments we follow Hensher *et al.* (2003); we pool the data and use the nested logit (NL) structure as a trick to reveal differences in scale. Normalising the scale parameter for one segment (area 1) allows variation for the other four segments.

---

[11] In fact, when users bought a monthly ticket, the bus fare was calculated dividing the price for that type of ticket by the average number of trips made in a month.

[12] In most cases, the arbitrary decision about λ does not matter and can be safely ignored, but since the scale factor affects the values of the estimated taste parameters (the larger the scale, the bigger the coefficients), one should never directly compare the coefficients from different choice models (Adamowicz et al., 1998).

The econometric results with the scaled coefficients are summarised in Table 3. Besides the five attributes we included the cutoff violations (e.g. COST_VC), that is the positive amount by which the lower and upper cutoffs for each attribute are violated, as explanatory variables. Therefore the parameters associated to these quantities should be negative, representing marginal disutilities. In the final model we included all the variables which have significant parameter[13].

Table 3: final model with the scaled coefficients.

VARIABLE	*AREA*				
	1	2	3	4	5
	$\tilde{\beta}$ -COEFFICIENT				
COST	-2,8287	-0,8358	-1,6680	-0,7472	-1,4181
	(0,3048)	(0,2642)	(0,2299)	(0,1201)	(0,2055)
DELAY	-0,2414	-0,1094	-0,0492	-0,1212	-0,0886
	(0,0439)	(0,0245)	(0,0327)	(0,0300)	(0,0345)
TRIP LENGTH	-0,0332	-0,0261	-0,0303	-0,0168	-0,0292
	(0,0167)	(0,0133)	(0,0119)	(0,0038)	(0,0072)
FREQUENCY	0,2251	0,5747	0,3757	0,6132	0,1588
	(0,0564)	(0,0809)	(0,0905)	(0,2133)	(0,0872)
AVAILABILITY	0,0034	0,0040	0,0027	0,0033	0,0029
	(0,0007)	(0,0006)	(0,0006)	(0,0007)	(0,0006)
COST_VC	-	-6,0148	-	-0,8768	-1,6492
		(1,3018)		(0,3818)	(0,7550)
TRIP LENGTH _VC	-	-	-0,1153	-	-
			(0,0572)		
FREQUENCY _VC	-	-	-	-2,1301	-0,7756
				(0,5669)	(0,2956)
SCALE PARAMETER	1,0000	0,6378	0,7653	0,9529	0,6802
	(fixed)	(0,0584)	(0,0760)	(0,0822)	(0,0638)

Log likelihood function	-1598.892
Restricted log likelihood	-5394.436
Chi squared	7591.089
Degrees of freedom	35
Prob[ChiSqd > value] =	.0000000
R-sqrd = 1-LogL/LogL* =	.70360

In brackets the standard errors for the parameter estimates.

[13] We investigated a number of interactions between attributes and socio-economic variables, but they did not add significantly to the overall goodness-of-fit. It was possible to estimate interactions coefficients only for pooled data since interactions within nests could not be estimated due to limited availability of data. Even if this problem could potentially be overcome by increasing the number of observations, the results obtained with pooled data indicate limited explanatory power achievable through this method. However given a sufficiently high number of observations one could avoid using a nested logit specification since the scale parameters may not prove statistically different from 1 when individual characteristics are introduced in the specification. Moreover, we implicitly assume that the residuals corresponding to different questions for the same individual are independent, which may not hold. A possible extension could be the use of a random coefficient model, with random terms specific to the individual and identical across questions. These specific issues remain to be tested in future research endeavours.

The overall explanatory power of this non-linear model is very high, a pseudo-R^2 of 0,7 is equivalent to approximately 0,9 for a linear model (Domencich and Mc Fadden, 1975) and this is in line with a similar study conducted in a different environment (Hensher *et al.*, 2003). The interpretation is very interesting since the weights of the attributes vary between areas. Area 1 is the most price sensitive while area 4 is the least price sensitive. Area 2 is characterised by a very high cutoff of bus fare. Area 1 shows the highest coefficients for delay and trip length. People in area 4 are the most sensitive to bus frequency and the associated cutoff has a significant impact. People in area 2 are the most responsive to service availability. Finally, the scale parameters are all statistically significant and different from one (except for area 4) indicating that the data cannot be pooled.

Calculating a Service Quality Index

In drafting contracts, it is crucial to take into account the local conditions and the distinctive characteristics of the public transport system considered. Hence, setting the minimum SQI level should be context-specific.

In order to calculate a SQI for each area we first calculate the SQI measure for each user through the formula

$$SQI_q = \sum_{k=1}^{K} \tilde{\beta}_k X_{kq} \ . \tag{0.6}$$

The SQI for user q is obtained by multiplying the RP attribute levels, as perceived by user q, by the appropriate scaled $\tilde{\beta}$-parameter in Table 3 and summing across the k attributes (in this case five). Then for each geographical segment s the overall SQI is measured by taking the individual SQI average for the sampled users in each area:

$$SQI_s = \frac{\sum_{q=1}^{n_s} SQI_q}{n_s} \ . \tag{0.7}$$

Table 4 shows the overall SQIs and the contributions of each attribute by area.

The various SQIs assume values between 0,69 (area 5) and 3,18 (area 2) and the mean is 1,24. As expected, bus fare, delay and travel time are sources of negative utility while service frequency and service availability offer positive contributions. In particular service availability is the most important attribute in explaining user satisfaction in each segment. Increasing the amount of time between service inception and service closure has the greatest effect in improving the SQI.

Moreover, SQI measures for different scenarios of public transport service can be calculated, since different mixes of attribute levels produce different SQI indexes.

Table 4: SQI and attributes contributions per area.

ATTRIBUTES	AREA					
	1	2	3	4	5	All
SQI_COST*	-1,84	-0,61	-1,19	-1,53	-1,40	-1,36
SQI_DELAY*	-0,43	-0,31	-0,09	-0,22	-0,13	-0,25
SQI_TRIP LENGTH*	-0,30	-0,28	-0,44	-0,82	-0,62	-0,49
SQI_FREQUENCY*	0,66	1,12	0,62	0,66	0,31	0,67
SQI_AVAILABILITY*	2,70	3,25	2,11	2,80	2,53	2,68
SQI	0,78	3,18	1,02	0,88	0,69	1,24

*Contributions account for cutoffs' influence

Concluding remarks

This paper has analysed service quality measurement and its integration in service contracts so to provide correct regulatory incentives via the introduction of the proposed quality specification and measurement. A case study illustrates the mechanism. In order to obtain reliable results, in the future a carefully structured sampling plan is needed.

Using SP methods and CBCA we estimate passengers' evaluation of different bus service features which users perceive to be the sources of utility and via discrete choice models we calculate a SQI.

Future research will pursue three different goals, one more strictly related to the methodological issues and the remaining two both related to the practical impact that SQI measurement might have.

As it is for further methodological investigation, as a referee mentioned, one should also test other possible explanations of the results obtained. In particular, one could consider the effects of random sampling variation (were this strong enough, it could explain the variability of the coefficients that are now imputed to differences across regions), individual characteristics' heterogeneity and observed attributes' endogeneity.

Whereas for the practical impact that SQI measurement might have, we would like to explore the role SQI might have within a service quality contract based on a price-quality cap as recent contributions underline (Bergantino *et al.*, 2006; Billette de Villemeur *et al.*, 2003; Cremer *et al.*, 1997) as well as to study the potential applications a SQI might have in defining a marketing strategy aimed at increasing profits. In fact, from the supplier's point of view, there is a need to establish the optimum trade-off between the service quality and its supply cost. The proposed method may also provide a useful performance assessment tool, in fact the operators may well understand where to focus their investment in order to reach a high level of service quality and increase their competitive advantage without wasting financial resources in relatively less important attributes amelioration.

The focus on quality should be a shared goal by the authorities and operators involved in the provision of transport services and the adoption of the suggested framework could prove a first step in this direction.

References

Aaker, D. and Day, G. S. (1990) *Marketing research*, John Wiley, New York.

Acton, J. and Vogelsang, I. (1989) "Introduction", Symposium on Price Cap Regulation, *Rand Journal of Economics*, 20(3), pp. 369-372.

Adamowicz, W., Louviere, J. and Swait, J.(1998) "Introduction to Attribute-Based Stated Choice Methods", Final report to Resource Valuation Branch, Damage Assessment Center, NOAA, U.S. Department of Commerce, January.

Averch, H. and Johnson, L. (1962) "Behavior of the Firm Under Regulatory Constraint", *American Economic Review*, 52(5), pp. 1053-1069.

Bailey, E. (1973) *Economic Theory of Regulatory Constraint*, Lexington Books, Lexington, MA.

Baumol, W. and Klevorick, A. (1970) "Input Choices and Rate-of-Return Regulation: An Overview of the Discussion", *Bell Journal of Economics and Management Science*, 1(1), pp. 162-190.

Ben-Akiva, M. and Lerman, S. R. (1985) *Discrete Choice Analysis: Theory and Application to Travel Demand*, MIT Press, Cambridge.

Bergantino, A. S., Billette de Villemeur, E. and Vinella, A. (2006) "A Model of Partial Regulation in the Maritime Ferry Industry", VIII SIET conference, 29th of June, Trieste, Italy.

Billette de Villemeur, E., Cremer H., Roy, B. and Toledano, J. (2003) "Optimal Pricing and Price-Cap Regulation in the Postal Sector", *Journal of Regulatory Economics*, 24(1), pp. 49-62.

Boiteux, M. (1960) "Peak-Load Pricing", *Journal of Business*, 33(2), pp. 157-179.

Bradley, M. A., (1993) "Behavioral models of airport choice and air route choice", Sixth International Conference On Travel Behaviour Research, Santiago, Chile, June 1993.

Carson R., Louviere J., Anderson D., Arabie P., Bunch D., Hensher D., Johnson R., Kuhfeld W., Steinberg D., Swait J., Timmermans H., Wiley B., (1994), "Experimental Analysis of Choice", Marketing Letters, 5(4), pp. 351-368.

Cronin, J. J. and Taylor, S. A. (1992) "Measuring service quality: a reexamination and extension", *Journal of Marketing*, 56(3), pp. 55-68.

Cremer, H., De Rycke, M. and Grimaud, A. (1997) "Service Quality, Competition, and Regulatory Policies in the Postal Sector", *Journal of Regulatory Economics*, 11(1), pp. 5-19.

Das, S. (1980) "On the Effect of Rate of Return Regulation under Uncertainty", *American Economic Review*, 70(3), pp. 456-460.

De Palma, A. and Picard, N. (2005) "Route choice decision under travel time uncertainty", *Transportation Research A: Policy and Practice*, 39(4), pp. 295-324.

Domencich, T. and Mc Fadden, D. (1975) *Urban travel demand, a behavioural analysis*, North-Holland Publishing, Oxford.

Friman, M., Edvardsson, B. and Garling, T. (2001) "Frequency of negative critical incidents and satisfaction with public transport services, I", *Journal of Retailing and Consumer Service*, 8(2), pp. 95-104.

Gatta, V. (2006) "Evaluating Service Quality: a Stated Preferences approach", unpublished Ph.D. Thesis, Department of CNAPS, Faculty of Statistics, *La Sapienza* University of Rome.

Hensher, D. A. and Prioni, P. (2002) "A service quality index for area-wide contract performance assessment regime", *Journal of Transport Economics and Policy*, 36(1), pp. 93-113.

Hensher, D. A., Stopher, P. and Bullock, P. (2003) "Service quality - developing a service quality index in the provision of commercial bus contracts", *Transportation Research A*, 37(6), pp. 499-517.

Hensher, D. A. (2004) "How do Respondents Handle Stated Choice Experiments? – Information processing strategies under varying information load", Working Paper ITS-WP-04-14.

Louviere, J. and Woodworth, G. G. (1983) "Design and Analysis of Simulated Consumer Choice or Allocation Experiments: An Approach Based on Aggregate Data", *Journal of Marketing Research*, 20(4), pp. 350-367.

Louviere, J. (1988) "Conjoint analysis modelling of stated preferences. A review of theory, methods, recent developments and external validity", *Journal of Transport Economics ad Policy, 22*(1), pp. 93-119.

Louviere J., Hensher D., Swait J., (2000), "Stated Choice Methods: Analysis and Applications", Cambridge University Press, UK.

Manski, C. (1977) "The structure of random utility models", *Theory and Decision*, 8, pp. 229–254.

Mc Fadden, D. (1974) "Conditional logit analysis of qualitative choice behaviour", in Zarembka P., ed., *Frontiers in Econometrics*, Academic Press, New York, pp. 105-142.

Panzar, J. (1977) "The Pareto Dominance of Usage Insensitive Pricing", in Dorick H., ed., Proceedings of the Sixth Annual Telecommunications Policy Research Conference, Lexington Books, Lexington, MA.

Parasuraman, A., Zeithaml, V. A. and Berry, L. L. (1988) "SERVQUAL: a multiple item scale for measuring consumer perceptions of service quality", *Journal of Retailing*, 64(1), pp. 12-37.

QUATTRO Research Consortium, (1998) "Quality approach in tendering urban public transport operations", Transport Research - Fourth Framework Programme Urban Transport – Final Report, Office for Official Publications of the European Communities, Luxembourg.

Ramsey, F. (1927) "A Contribution to the Theory of Taxation", *Economic Journal*, 37(1), pp. 47-61.

Sappington, D. and Sibley, D. (1988) "Regulating without Cost Information: The Incremental Surplus Subsidy Scheme", *International Economic Review*, 29(2), pp. 297-306.

Sawtooth Software Inc., (1999) "CBC User Manual", Sequim: Sawtooth Software.

Swait, J. (2001) "A Non-compensatory Choice Model Incorporating Attribute Cutoffs", *Transportation Research B*, 35(10), pp. 903-928.

Stopher, P. R. (1998) "A review of separate and joint strategies for the use of data on revealed and stated choices", *Transportation*, 25(2), pp. 187-205.

Teas, R. K. (1993) "Expectations, performance evaluation, and consumers' perceptions of quality", *Journal of Marketing*, 57(4), pp. 18-34.

Thurstone, L. (1927) "A law of comparative judgment", *Psychological Review*, 34(4), pp. 273-286.

Tortora, R. D. (1978) "A Note on Sample Size Estimation for Multinomial Populations", *The American Statistician*, 32(3), pp. 100-102.

Train, K. E. (1991) *Optimal Regulation: The Economic Theory of Natural Monopoly*, The MIT Press, Cambridge.

Transportation Research Board (1999) "A handbook for measuring customer satisfaction and service quality", Transit Cooperative Research Program, Report 47, National Academy Press, Washington, D.C.

Williamson, O. (1966) "Peak-Load Pricing and Optimal Capacity under Indivisibility Constraints", *American Economic Review*, 56(4), pp. 810-827.

Willig, R. (1978) "Pareto Superior Non-Linear Outlay Schedules", *Bell Journal of Economics*, 9, pp. 56-69.

Zeithaml, V. A., Berry, L. L. and Parasuraman, A. (1993) "The Nature and Determinants of Customer Expectations of Service", *Journal of the Academy of Marketing Science*, 21(1), pp. 1-12.

Acknowledgements

Financial help form CIPE 17/2003 contract act N. 1804, 23/12/2003 titled "Interdisciplinary analysis for the definition of urban mobility policies in the Marche Region aimed at containing individual transport and increasing local public transport" is gratefully acknowledged. We would also like to thank the participants to the SKIM Software & Sawtooth Software conference on "Design & Innovations" held in Munich (Germany), the 3rd International Kuhmo Conference and Nectar Cluster 2 meeting on "Pricing, Financing, and Investment in Transport" held in Tuusula (Finland), the SIEP seminar on "Public services: the new trends in regulation, production and financing" held in Pavia (Italy), as well as the valuable suggestions and comments of Nathalie Picard and the anonymous referees for their insightful and helpful comments.

European Transport \ Trasporti Europei n. 36 (2007): 107-140

On the change in surpluses equivalence: measuring benefits from transport infrastructure investments

José Meléndez-Hidalgo [1*], Piet Rietveld [1], Erik Verhoef [1,2]

[1] *Department of Spatial Economics, Free University of Amsterdam*
[2] *Tinbergen Institute*

Abstract

Reductions in transport costs resulting from infrastructure improvements generate benefits that can be measured as surplus changes either at an economy-wide scale (social welfare changes) or, as is common practice in cost-benefit analysis (CBA), at a transport market level as transport users' surplus changes. In this paper we look at an economy with spatially separated markets embedded in a transport network (a spatial price equilibrium model) to study the equivalence between these two benefit measures. Three different product market competition arrangements are considered. A similar question and strategy is presented in Jara-Díaz (1986) employing a two-node network and extreme competition assumptions on the production side: perfect competition and monopolistic production with arbitrage. We extend his work by additionally considering perfect collusion (monopoly without resale) and Cournot-Nash oligopoly under flow-dependent transport costs (i.e. congestion in transport). Numerical simulations in a three-node network with and without transshipment nodes, illustrate our main results.

Keywords: Cost-benefit analysis; Indirect benefits.

1. Introduction

In recent years a renewed interest has emerged on the effects arising outside the transport market when transport infrastructures are improved (SACTRA, 1999). The interest on this issue in economics can be traced back to Tinbergen (1956) where, for a three-region economy under monopolistic production, a comparison is made between transport users' cost savings and economy-wide effects (e.g. aggregate production changes) after a reduction in the cost of transportation for a pair of regions. In a related vein of research Jara-Díaz and Friesz (1982) employed a spatial price equilibrium (SPE) model to study the equivalence between the economy-wide changes in trade surplus and the transport users' benefits for a two-region economy under perfect competition in production. Their paper initiated the use of SPE models to investigate the properties of

[*] Corresponding author: José Meléndez-Hidalgo (jmelendez@feweb.vu.nl)

an implicitly *derived transport demand.*[1] Jara-Díaz (1986) returned to this topic considering perfect competition and monopolistic production, the latter with the possibility of resale (i.e. arbitrage), within a two-node one-link transport network.

In this paper we further develop the analysis in Jara-Díaz (1986) and Jara-Díaz and Friesz (1982), (hereafter referred as JD and JD-F, respectively) considering different competition arrangements in production and more complex transport networks. In JD-F the equivalence between users' benefits measured on a derived transport demand and those benefits measured as economic (trade) surplus is analyzed under perfect competition. Their network equilibrium model comprises a two-region economy within a two-links (transport modes or routes) network, given that their main interest is on mode interaction when a uni-modal improvement is in place. This comparison is dubbed the *surplus equivalence problem* in JD, where more analytical detail is provided using an SPE model with two regions embedded in a single link network, under both perfect competition and monopolistic production with resale.[2] We will be employing similar SPE models with elastic derived transport demands and in addition to the perfect competition case will consider imperfect competition in the form of monopoly over all production sites without arbitrage (i.e. perfect collusion) and Cournot-Nash oligopoly.

Our work differs from JD-F in that we allow only for one link per pair of nodes, but we are still able to discuss route choice for the three node case network. Additionally, in contrast with JD, we study the case of more than two nodes relying on numerical simulations. A network increasing in complexity is considered aiming at disentangling the effects of alternative assumptions on both the type of competition in the product market and the shape of the transport network. Finally, our approach for addressing economy-wide welfare differs from both JD-F and JD, as we look at the entire magnitude of social welfare and not only at the surplus from trade.

The network equilibrium models employed in this paper correspond to a partial equilibrium concept and are especially suited for situations were price and output changes in the transportation sector do not significantly affect capital and labor markets. Income effects, intermediate goods and inventory issues are also ruled out. [3] Moreover, the focus here is on steady-state network equilibrium models featuring interregional one-commodity trade, as opposed to network models emphasizing passenger flows (Beckman et al 1956) or those investigating freight transportation in a dynamic context (Friesz et al. 2001).[4] Additionally, we allow for more than two nodes and the existence of transshipment nodes, in the simulation exercise, in order to be able to discuss the interaction of congestion effects, Wardrop equilibria and the possible presence of Braess's paradox type of phenomena.

The rest of the paper consists of three main sections. Firstly, in the next section a review of what has been discussed in the literature up to now is presented and put it in the perspective of our approach. The discussion in this section is conducted mainly through graphical analysis. Additionally to the usual spatial perfect competition case,

[1] See also Galvez and Jara-Díaz (1975). This related literature is summarized in Freisz (1985).

[2] The conditions characterizing equilibrium for the monopolistic production case in JD are quite different from those used here, but also from the literature on SPE under imperfect competition (see Harker, 1986). As will be discussed later in the paper, Jara-Díaz (1986) assumes a spatial monopoly with arbitrage, as defined in Takayama and Judge (1971).

[3] Waquil and Cox (1995) consider intermediate production within SPE models.

[4] The generalization to multi-commodity economies is made possible through a multi-copy network formulation which in the case of non-separable functions implies no loss of generality when using a single-commodity formulation (Freisz et al. 1998).

both the spatial monopoly without resale and Cournot-Nash oligopoly cases are discussed in detail. In section three we use numerical simulations to evaluate the surplus equivalence in a three region economy embedded in two different transport networks. Finally, the last section concludes and discusses future work.

2. Surplus change equivalence in a two-region economy

In this section a graphical exposition is employed to review what has been already pointed out in JD-F with respect to benefits measures comparison using a SPE model under perfect competition. More important, based on this review we extend their analysis to the cases of perfect collusion without resale across regions and to Cournot-Nash oligopoly. A two-region economy is assumed and transport costs are taken initially as a constant (i.e. independent of the flows). [5] Before going to the main issue in this section, more detail on the *surplus change equivalence problem* will be given. In particular, we discuss how users' benefits in the transport market itself are usually measured, both for the case with and without congestion effects. [6] Additionally, we discuss the composition of total welfare in a SPE model.

Surplus change equivalence

The main concern in this paper is the comparison between users' benefits arising in the transport market itself as a change in transport consumers' surplus (TCS), with those benefits arising at an aggregate level as changes in total social welfare. This research question can be said to start formally with Tinbergen (1957). After this work, as far as we know, no further contributions were seemed to be cast in a similar or related manner until Kanemoto and Mera (1982), which explicitly recognized his work in the context of a first-best general equilibrium economy subject to non-marginal transport cost reductions. JD-F and JD adopted a related approach and the latter called the comparison the *surplus equivalence problem*. As analyzed in JD, the comparison is between the "trade" surplus at an economy-wide scale and the TCS, from an equilibrium without trade to one with positive trade.[7] This is in contrast with the previous and more recent literature on this issue where the comparison is usually reported as the ratio of the economy-wide welfare change over the transport market measure when both the initial and final equilibriums are characterized by positive trade flows. Under the assumptions taken in JD-F (i.e. perfect competition in production) there is no difference between both approaches but when the consequences of intermediate competition arrangements are to be analyzed the correct comparison corresponds to the last interpretation. In this respect, our work departs from JD but employs the second interpretation; consequently we refer to the *surplus change equivalence,* instead of the (trade) surplus equivalence, and compute the economy-wide welfare considering both the traded and non-traded

[5] At the end of the section the extension to transport cost depending on the flows is discussed briefly.
[6] This discussion is referred later in the paper when the numerical simulations are presented.
[7] As discussed in Jara-Díaz and Freisz (1982) the equivalence holds as well for a comparison between two equilibriums with positive trade. In this case the change in surpluses also holds, but in their paper the economy-wide surplus still refers only to the trade surplus.

quantities. The common practice for this second approach is to report results as a ratio of economy-wide benefits to the change in TCS. The benchmark for this ratio is a level of one when distortions exist neither in the transport market nor in the rest of the economy.

The main components of the surplus change equivalence are two users' benefit measures. One is for the entire economy and can take as many possibilities as there are available according to welfare economics, but in general will be restricted by the background economic equilibrium trade model employed. For instance, in Meléndez-Hidalgo et. al (2003) both the equivalent and compensated variations are used to compute total welfare changes and in Kanemoto and Mera (1982) more sophisticated total welfare measures are proposed. In both of these papers a general equilibrium approach is employed to model trade across space complicating the issue of measuring total welfare. In a partial equilibrium framework, as is the case in this paper, total welfare or *the social welfare* (SW) *function* is simply composed of net surplus to consumers plus producers' profits (including transportation costs). Only if there are available instruments for the government to influence the allocation of resources, the government surplus would be added to this two. [8]

The second component of the surplus change equivalence problem is a measure of direct benefits based on a derived demand for these services and captured in the change in TCS. As for the previous welfare measure, the TCS change can be computed based in as many ways available within welfare economics with respect to consumers' benefits. These options are again restricted by the chosen economic model. For the case of SPE models, due to the unknown functional form of the transport services demand and the absence of income effects, there are only two possible ways of measuring the TCS. One may compute it as the integral over an implicit expression for the derived demand or as an approximation using the *rule of a half*. [9]

In Figures 1 and 2 we show as shaded areas the users' benefits arising in the transport market as a result of a decrease in the equilibrium transport cost. These benefits are measured on a (inverse) derived transport demand, $D_t(\cdot)$, which in the case of a two-region SPE model corresponds, as it will be shown later, to the amount of production (consumption) in the exporting (importing) region that is traded due to a positive gap in the willingness to pay between the importing and exporting regions -at a given transport cost level.

[8] In Cremer et al (2003), a related paper, the social planner welfare function or SW is decomposed into the sum of consumers' (net) surplus, the transport network operator income (or merchandizing surplus) and the producer surplus (excluding transport costs). This is equivalent to the approach followed here but since we are not considering social planner decisions we prefer to exclude the distinction they made.

[9] If it is the case that the "derived transport demand" curve moves from an initial equilibrium to the final one, after an infrastructure improvement, then the consumer's surplus might be path dependent if certain conditions do not hold (Jara-Díaz and Friesz (1982)). In this case we employ a linear path from the initial to the final equilibrium, which is the same as assuming that the final effect is the sum of marginal consecutive effects from changes in the direction assumed.

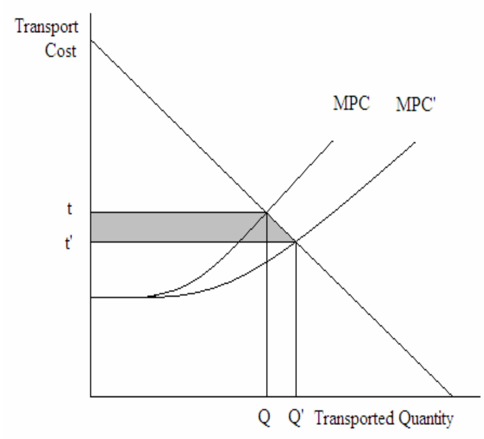


Figure 1: Users' Benefits in the Transport Market without congestion charging.

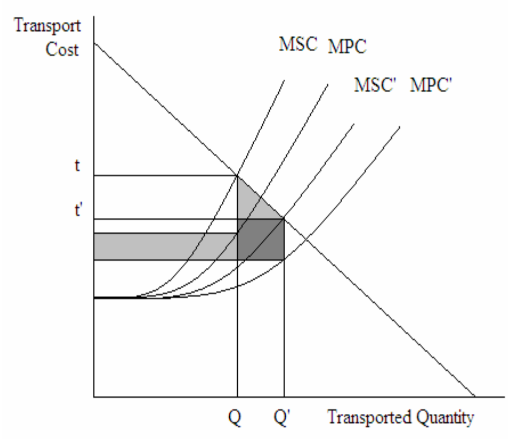


Figure 2: Users' Benefits in the Transport Market with congestion charging.

A reduction in the transport cost from t to t' is shown in both Figures. The traditional Marshallian consumers' surplus variation gives the users' benefits as the change in TCS:

$$\Delta TCS = -\int_{t}^{t'} Dt(s)ds$$

These benefits result both from cost savings in transporting the previous level of trade but also from increased trade induced by the decrease in the transport cost. This latter component has been interpreted previously in the literature as the average of the maximum and minimum possible benefits for the induced trade when the rule of a half is applied to approximate the users' benefits (Neuberger (1971)).[10]

The case of no congestion charging is shown in Fig.1. As discussed in Mohring (1976), among others, the benefits are measured with respect to the marginal private costs (MPC) of transportation when this is a function of the flow.[11] In this case the transport benefits measured as a change in TCS will include a transport market welfare loss associated with the congestion externality.[12] In Fig.2 the case of congestion charging is shown. To the traditional rule of a half components (the triangle and the horizontal rectangle in light-colour) has to be added the vertical rectangle representing the net increase in congestion charging revenue resulting from the exogenous decrease in transport costs.

[10] A first difference arises here with respect to JD-F as they consider modal choice leading to a linear (Hoteling type) integral of the demand function. To go around the path dependency problem under linear integration they assume that Green's condition holds.

[11] The MPC is equivalent to the average cost of transportation, used in other contexts.

[12] This is not shown in the Figure but it would be clear if, as in Fig.2, both the marginal private and social costs are drawn. In that case the equilibrium at the intersection of the marginal private costs with the derived demand will correspond (at the equilibrium level of flow) with a level of marginal social cost above the private level. The area under the marginal social cost and above the transport demand from the point of intersection of this curve with the transport demand, to the equilibrium defined as in Fig. 1, will correspond to the deadweight loss due to the congestion externality.

In the rest of this section we will discuss the equivalence between users' benefits measures under three different assumptions for the industrial organization in production. The first case analyzed is the one originally addressed in JD-F for perfect competition without congestion effects. It would be argued later in the paper that the case with congestion is equivalent to the one without congestion with the only qualification that the users' benefits in the transport market must be computed as shown in Fig.2. The second case analyzed is the perfect collusion assumption which has not been discussed before in the literature. JD considers monopolistic competition in production but employs different conditions to characterize the equilibrium (i.e. resale is possible). Again the case without congestion is the one discussed in detail but the same discussion applies to the case with congestion effects applying the same qualification as in the perfect competition case. Finally, we also contribute to the literature on the surplus equivalence problem by analysing the Cournot-Nash competition case. To simplify the exposition this latter case is presented for a situation were only one of the firms in exporting to the other region meaning that both firms compete in the importing market but the exporting firm behaves as a monopolist in its own market. In what follows the exposition will be mainly based on graphs. In section three we formalize the models before discussing numerical simulations for a three regions economy embedded in two different networks.

Perfect competition

Our final aim is to compare changes in SW with the change in TCS -measured as the area to the left of the derived transportation demand- both arising after an initial equilibrium with trade is disturbed by an improvement in the capacity of a transport link between a pair of regions. In Fig. 3 and 4 this exercise is presented as in JD-F for a two-node-one-link network and the resulting (inverse) derived transport demand, $D_t(\cdot)$, is shown in Fig.5. In contrast with their paper we assume the homogeneous good is transported from region 1 to region 2, with direct and inverse demand and supply functions per region respectively given by:

$$D_i = D(p_i); \qquad p_i = \theta(D_i)$$
$$S_i = S(p_i); \qquad p_i = \psi(S_i)$$

with i={1,2} and excess demand and supply functions given by:

$$ED_2(p_2) = D_2(p_2) - S_2(p_2)$$

$$ES_1(p_1) = D_1(p_1) - S_1(p_1)$$

In Fig.3 the equilibrium prices arising with positive trade, p_1 and p_2, are shown. A different pair of prices is associated with both an initial and final level of the transport cost, t and t', respectively. In Fig. 4 the corresponding equilibrium excess supply and excess demand quantities are derived based on the difference in quantity demanded and supplied in a particular region for each transport cost level.

In contrast to JD-F when computing the economy-wide surpluses in each region we use a quantity formulation of the SPE instead of a price formulation[13]. This is reflected in the figures when measuring the economy-wide welfare (shaded areas), as we consider not only the net surplus from *trade* (the triangles shown in JD-F and JD) but the entire surplus in consumption and production for both regions. As discussed before, this distinction is central when considering competition arrangements others than perfect competition in production, as we do when analyzing Cournot-Nash competition and Perfect Collusion.

Figure 3 shows that the result given in JD-F concerning the trade surplus equivalence when perfect competition is assumed in both the transport-using and transport sectors, has a parallel in the quantity formulation.[14] From a quantity formulation perspective we can re-state their finding as follows: the change in total social welfare from a situation with trade across regions, as compared to one without trade, will correspond to the entire TCS arising for the final transported quantity, T. The change in social welfare is precisely the net trade surplus that JD-F emphasize. Their focus is on the benefits of trade and consequently they use the excess supply and demand to characterize these benefits and to link them with the derived transport demand[15]. In contrast, in Fig.3 and Fig 4 our emphasis is different. In those figures it is shown that in a quantity formulation the equivalence that holds is that of the users' benefits (change in TCS) with the change in SW when comparing two decentralized equilibriums. This approach is more general as those equilibriums can include one without trade as in JD and JD-F (and as discussed above), but also the more general case we are interested in where we compare two equilibriums with positive trade (i.e. positive and not prohibited trade costs). The *change* in TCS is also captured by the excess supply and demand curves net surpluses as shown in Fig. 4, but again, this result applies only to perfect competition conditions meaning that is of no use in our more general approach.

In Fig. 3 we show two decentralized equilibriums arising from two different levels of the transport costs. In both equilibriums region 1 is exporting the homogeneous good to region 2. In the original equilibrium, region 1 produces S_1 and consumes D_1 (less than S_1) at a price p_1 while region 2 produces at S_2 and consumes D_2 units of the good ($> S_2$) at a price p_2. At those prices there would be an interest in trade from region 1 to region 2 and assuming that the transport cost is initially a constant this trade will be conducted until the point where the difference in prices (p_1-p_2) equals the transport cost level, t.[16] The total social welfare for this (initial) equilibrium will correspond to the sum of both light-shaded regions in Fig. 3, which is the sum of consumers' (net) surplus in both

[13] The domain for integration when computing the SW refers to quantities in the quantity formulation and to prices, as in JD-F, for the price formulation.

[14] This is not a surprising result in the light of the isomorphism of both formulations already discussed in Takayama and Woodland (1970).

[15] JD-F also argues that the change in TCS when comparing two equilibria with positive transport cost is also equivalent to the change in SW. This is basically the same approach to look at the equivalency in surpluses as the one taken in our paper.

[16] The transport cost level represents the willingness-to-pay to move the units from region 1 to region 2 and under perfect equilibrium is equal to the difference in prices between the importing and exporting region. Jara-Díaz (1986) shows that this latter condition together with the conservation flow condition stating the equality between excess demand and excess supply, can be used to construct an inverse derived transportation demand in terms of the inverse excess demand and supply functions. Based on this latter result, the TCS can be shown to be equal to the change in SW from autarky to the equilibrium with trade. See Jara-Díaz (1986), pag. 383.

regions and the firms' profits including transport costs. [17] After this equilibrium is disturbed by a decrease in the cost of transportation from t to t', the new equilibrium will correspond to more trade from region 1 to 2 and new prices at p_1' and p_2'. The gains arising from the infrastructure improvement will be shared by producers of the exporting region and consumers of the importing one according to the elasticities of demand and supply in both regions. [18] Total welfare in the new equilibrium will include not only the light-shaded areas but also the dark-shaded ones, which correspond to the increase in welfare due to increased trade. The change in social welfare is composed of *saved costs* of transporting the previous traded quantity –transferred in part both to producers and consumers- and extra benefits arising from induced trade after the reduced transport cost opens a gap between initial equilibrium prices across regions that is subsequently reduced to zero with more trade.

At this point it should be easier to see the parallel with JD-F. Since in their analysis distortions in the economy are ruled-out, their focus is on the benefits from trade at an economy-wide level. Ignoring the surplus arising in autarky is innocuous under their assumptions. As a consequence the sum of surpluses arising on an excess supply and excess demand will completely capture this extra surplus as shown in Fig.4. While assuming perfect competition there is no serious implication from this approach since there are no distortions in production and the whole extra benefit arising from a reduction in transport costs comes entirely from the new trade flows. If we want to look at other types of industrial organization in production the price formulation is no longer illustrative of what is happening at the level of total welfare, then our emphasis on the entire surplus (e.g. total social welfare). [19]

The surplus (change) equivalence under perfect competition is illustrated in Fig. 3 and Fig. 5. As a result of a decrease in transport costs from t (= p_1-p_2) to t' (= p_1'-p_2'), the change in total welfare as shown in Fig.3 is completely captured by the users' benefits measured in the transport market itself, as a change in the transport consumers' surplus (TCS). This is shown in Fig. 5 where the derived transport demand, D_t, arising from this simple model is depicted. The TCS change corresponds to the dark-shaded area in Fig.5 which exactly matches the sum of dark-shaded areas in Fig.3 (and Fig.4).[20] The new transport cost will match the difference in prices at the new equilibrium and is associated with a level of trade given by T′ (= $D_1' - S_1' = S_2' - D_2'$), which is greater than the level of trade associated with the previous equilibrium, T (= $D_1 - S_1 = S_2 - D_2$). The sharing of benefits across regions depends on their elasticities of the excess demand and supply, as discussed before.

[17] Besides we are in a competitive economy, there are non-zero profits (e.g. normal profits) for the non-marginal firms. These are measured as the difference between the total sales minus the integral of the supply curve up to the equilibrium quantity.

[18] See Jenkins and Kuo (2006) for more on the sharing of benefits and for an application of this framework to the actual evaluation of infrastructure improvements across countries.

[19] The price formulation is still indicative of what is happening with the benefits of trading even in the case of imperfect competition, but drawing a parallel to the discussion in perfect competition we can carried out all the analysis based on the quantity formulation.

[20] This is the main result in relation to the surplus equivalence issue shown in Jara-Díaz and Friesz (1982) and Jara-Díaz (1986) for the SPE model in price formulation.

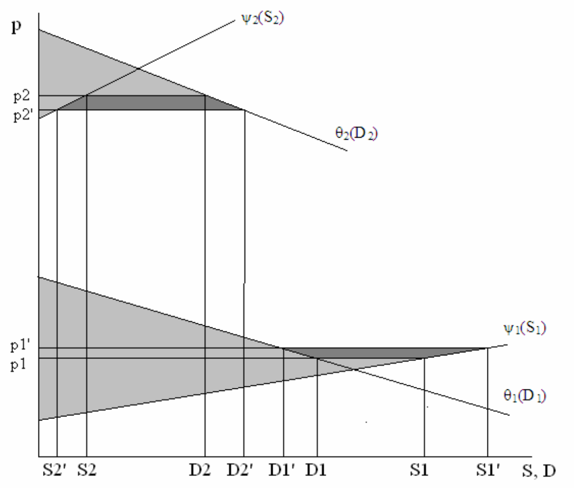


Figure 3.

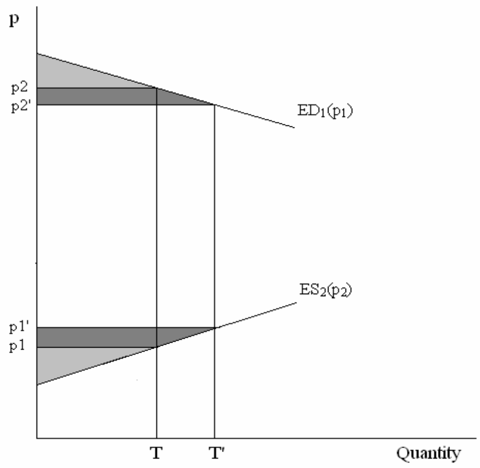


Figure 4.

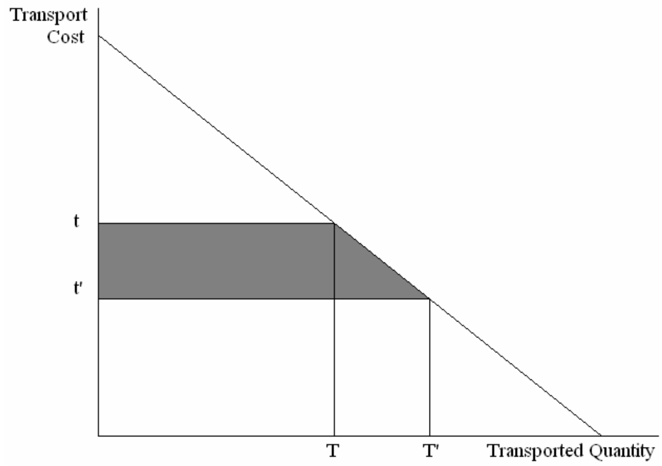


Figure 5.

Perfect collusion (spatial monopoly without resale)

The competitive SPE model makes a strong assumption concerning the market structure. It supposes that markets are perfectly competitive within a region; that is, price will always equal marginal cost in every region since there are infinite number of both consumers and producers per region. Analyzing spatial markets with a perfect competition assumption has been criticized in the literature (Scotchmer and Thisse, 1992; Sheppard and Curry, 1982). The existence of spatial monopolies is well recognized and the very existence of a spatial structure put to question the use of the perfect competition assumption if one recognizes that a firm closer to a market will have an advantage over a firm farther away when transportation is not costless (Harker, 1986). In this sub-section we relax the assumption of perfect competition in production, initially considering a spatial monopoly in production or as we prefer to interpret it here, perfect collusion among regional producers. This competition arrangement is different to the monopolistic competition considered in JD, as he allows for resale among regions meaning that prices between a pair of regions cannot differ more than the magnitude of the transport cost between them.[21]

A first attempt to consider imperfect competition within SPE was taken by Takayama and Judge (1971). In their setting production is undertaken by a monopolist operating in all regions. If resale among regions is not possible, the monopolist discriminates prices across regions. They showed that when there is trade from a region *i* to another region *j*, the marginal revenue of the firm in market *j* equals the marginal costs of the firm in region *i* plus the transportation cost from *i* to *j*. In the monopoly case prices are higher and production levels lower as compared to the competitive case. Takayama and Judge (1971) also explored monopoly pricing when resale among regions is possible, that is, arbitrage is allowed across regions. This impose additional constraints for the monopoly, namely prices among regions can not differ more than the transportation cost between them. The latter approach is the one followed in JD when he refers to monopolistic competition. The former interpretation, allowing for price discrimination, is the one we use in this paper.

It was shown in the previous sub-section that under perfect competition the change in SW corresponds exactly with the change in TCS, when a reduction in transport costs is in place. For the perfect collusion case this result does not hold anymore. The main difference concerns the measuring of TCS changes. For the perfect competition case and at a given transport cost level, the demand together with the supply curves (sum of marginal costs across active firms in a region) define in every region an equilibrium price level such that the traded quantity is a positive gap between the amount demanded and supplied at the importing region particular price, which in turn matches a positive gap between the supplied and demanded quantity in the exporting region at a lower price. This latter price differs from the price of the importing region by the level of the transport cost. J-D shows (pgs 380-82) that for a two region economy a derived transportation demand can be constructed from the equilibrium relation between prices across regions and the excess demand and supply. It follows from this that the willingness to pay for the movement of the traded units (e.g. derived transport demand) is in fact constructed based on the final demand and supply functions.

The equilibrium condition in the perfect collusion case corresponds to an equalization of marginal costs and marginal revenues in each region and not in prices as the perfect competition case. This equalization define the equilibrium demanded and supplied quantities which in turn fulfil a conservation of flow condition implying an equalization of "desired" trade flows across regions. [22] Consequently it is the marginal revenue instead of the final demand which defines, together with the marginal cost functions, the willingness to pay for trading goods in the perfect collusion case. The equilibrium condition across regions in the perfect collusion case says that the exporting region marginal cost plus the transportation cost must be equal to the importing region marginal revenue. Then the benefits from trade are captured in a derived transportation demand constructed from these curves. This is in parallel to the procedure employed in J-D for the perfect competition case but with the marginal revenue curve replacing the total demand and the total supply being replaced by the marginal cost function in each

[21] See Takayama and Judge, (1971).

[22] More formally, a conservation flow condition is behind the equalization of "desired" and "realized" trade flows.

region.[23] Since the SW is measured in the same way irrespective of the competition arrangement in production, it follows that the users' benefits will in general differ from the change in SW, after an infrastructure improvement.

In Figures 6 and 7 the equilibrium before and after the transport cost reduction is shown for the perfect collusion case. In Fig. 6 the equilibrium is, as before, one where region 2 imports from region 1 before and after the improvement in infrastructure. The SW will correspond to the sum of all grey areas. In contrast to the case of perfect competition, the transported quantity is defined as a gap between the marginal cost and marginal revenue curves at the equilibrium across regions. This equilibrium results from the equalization of marginal revenue and costs in each region and the equalization of the marginal revenue for the importing region with the sum of the marginal and transport cost from the exporting region. The traded quantity is not anymore defined based on the market demand as in perfect competition. The benefits from trade (e.g trade surplus) are shown as in JD as the triangular grey areas forming from the crossing of the marginal revenue and marginal cost curves. The two-firm "cartel" decides on how to allocate production as to maximize total profits and in equilibrium discriminate prices across regions.

Figure 7 shows the consequences of a reduction in the transport cost level. After the reduction had defined a new equilibrium, characterized by a lower price in region 2 but a higher one in region 1 (not shown in the price scale to keep the figure clear enough), the change in SW is no longer the sum of the horizontal dark-grey areas as in the previous case (Fig.3). In addition to the trade surplus (horizontal dark areas) there is a net effect related to mark-up trade-off across regions which is exercised by the "cartel" in order to maximize total profits. Two vertical dark-grey areas reflect this trade-off. In the exporting region (region 1) there is a loss in profits because fewer units are consumed locally and instead are now exported. There is also a net loss in consumer surplus given by the reduction in demand (= D_1'- D_1) and reflected by the upper triangle in the horizontal dark-shaded bar in the lower set of curves in Fig.7. Additionally, with the increase in trade and demand in region 2 (=D_2'- D_2) there is a net increase in consumer surplus and also an increase in profits dictated by the mark-up exercised (e.g. profits) by the cartel in the new units imported from region 1. This is reflected by the horizontal bar in the upper set of curves in that figure. The net effect defines the extra component in the change in SW, as compared to the perfect competition case. The difference in users' benefits measured in the transport market itself and those reflected as aggregate welfare changes will correspond in this case precisely to the net sum of the integrals between the total demand and marginal revenue curves, from the original quantity demanded (D_1 and D_2) to the new quantities (D_1' and D_2') in both the importing and exporting market. Being that in general one of these areas is negative (for the exporter) and one is positive (for the importer), the change in SW will differ from the change in TCS. The change in profits is expected to be always positive since the monopolist maximizes across all regions. Besides that, the net effect can still be negative if the net changes in consumers' surplus across regions end up being negative.

[23] The focus changes here from a market view to a firm view of benefits and costs since we are assuming that there is a "cartel" of regional firms behaving as a monopolist over the entire set of regions in the economy. The willingness to pay for transporting the good in this case refers to in a direct way to the firm, and indirectly to the final consumer.

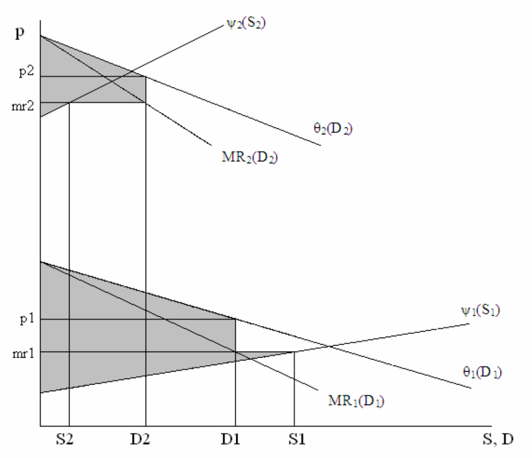


Figure 6.

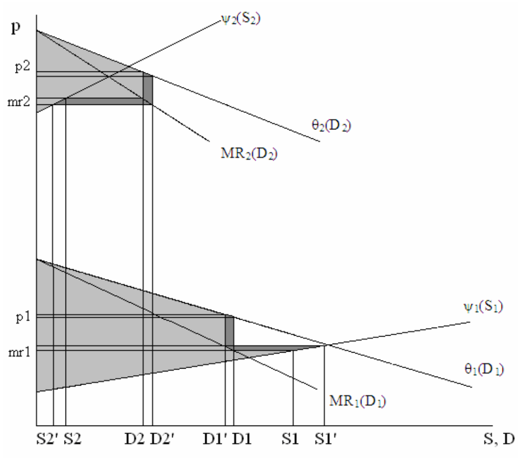


Figure 7.

The magnitude of the difference will depend on the elasticity of demand in each market. The possibility for the "cartel" of exercising market power and discriminating in prices across regions creates a gap between the benefits at an aggregate level and those arising in the transport market itself. After a reduction in the transport cost the cartel trade-offs profits margins across regions in order to maximize profits, meaning that the change in total profits will always be positive. In contrast, the net change in consumer's surplus included in the horizontal bar can be positive or negative and could even compensate the net change in profits.

Cournot-Nash competition

Hashimoto (1984) has extended the perfect competition SPE model to a spatial Nash non-cooperative equilibrium model. He considers one firm per-region competing *à la* Cournot-Nash with firms from other regions. A comparison of the Nash equilibrium model with the perfect competition and monopoly cases shows that the differentials in interregional prices in the Nash equilibrium are greater than in the perfect equilibrium case and lower than in the monopoly case. Harker (1986) and Nagurney (1993) also present a spatial oligopoly model where the profit function of each firm considers the transportation costs from the production plants to the demand markets. We follow closely the formulation by Harker (1986).

In order to simplify the discussion of the Cournot-Nash competition case, we make the extreme assumption that there is only one firm controlling production in each region-market. In a more general case we could take the situation where more than one firm operates in a particular region. As pointed out by Harker (1986), our assumption is not restrictive in the sense that any particular region, in which two or more firms operate, can always be decomposed into sub regions in which only one firm operates. Then we consider two firms, firm 1 and 2, supplying both its internal market and the other's region market. Each firm distribute its production between the two markets in a way that maximizes its total profits and takes transport costs as part of the total costs of supplying the markets. [24] In particular, firm 1 and 2 supply the "internal" market up to a point where the marginal costs of production equals the marginal revenue based on the

[24] This is similar to the segmented market assumption in Brander and Krugman (1983).

residual perceived demand in that market. This residual demand results from subtracting the amount supplied by other firms from the total demand, in this case, the amount supplied by firm 2 and firm 1, respectively. Both firms supply consumers in the "competing" region, and they do that until a point where the marginal revenue of supplying that region, based on the residual demand in that market, equal the sum of the marginal cost of its production (in the region of origin) plus the transport cost incurred in the last unit shipped. The residual demand in this case corresponds to the subtraction from the total demand in the "foreign" market of the amount supplied by other firms, in our framework the amount supplied by the local producer. An important difference arises at this point in comparison with previous industrial organizations in production because both firms might be supplying both markets and consequently we could be in a situation of *two-way trade*. [25] The implications from this to the surplus equivalence are significant since even though an infrastructure improvement might only affect a one-way link, the final equilibrium will be reflected also in the trade going on an opposite direction (as in the two-link model of JD-F).

The initial equilibrium in this environment can be seen more clearly in Figures 8 to 10. To further simplify the analysis we are making the extra assumption that in equilibrium there is only trade from region 1 to region 2, that is, region 1 producer exercise full market power over its own region demand and additionally competes in region 2 market with the local producer. Fig.8 is basically equivalent to the bottom part of Fig.6 where region 1 producer is the exporter and also behaves monopolistically over its own region total demand. The benefits from trade in that region correspond to the dark-grey area, that is, the total benefits arising from the amount of goods sent to region 2, T_{12} ($= S_1 - D_1$). In Fig.9 and Fig. 10 the equilibrium is shown for regional market 2. In this market the situation is different since there is strategic interaction among firms. The total market demand, $D_2(P_2)$, is assumed to take a linear form given by $\theta_2(D_2) = \rho_2 - \eta_2 D_2$; then the perceived inverse residual demand for the local firm, $\theta_{22}(D_2)$, results from subtracting horizontally D_{21}, the demand in region 2 served by region 1 firm. In terms of the inverse demand this is equivalent to subtract a constant, $\eta_2 D_{21}$, from $\theta_2(D_2)$, which is graphically shown as vertically subtracting this magnitude from the total inverse demand. The perceived inverse residual demand is then given by:

$$\theta_{22}(D_{22}) = \rho_2 - \eta_2 (D_{22} + D_{21}) = (\rho_2 - \eta_2 D_{22}) - \eta_2 D_{21}$$

with ($D_{21}=T_{12}$) in equilibrium. The associated residual marginal revenue faced by firm 2 is also shown in Fig.9, and the equilibrium production corresponds to a point where both marginal revenue and marginal costs for this firm coincide ($S_2 = D_{22}$). The equilibrium for firm 1 in region 2 is shown in Fig.10. There, the residual demand faced by firm 1 is constructed as the result from vertically subtracting a quantity $\eta_2 D_{22}$ from the inverse demand, which is related to the part of total demand in that market which is served by the local firm. The equalization of the associated residual marginal revenue with the sum of the equilibrium marginal revenue in region 1 plus the transport cost incurred in the last unit sent from region 1 to 2, define the amount of firm 1's total production directed to region 2 ($D_{21}=T_{12}$). More concretely, firm 1 allocate its total production between the local and foreign markets based on an equalization of marginal costs to marginal revenue which in the case of the foreign market, and due to the

[25] Cross-hauling is due to strategic interactions across firms, as discussed in Brander (1981).

assumption of full market power in the local market, corresponds to the sum of the marginal cost in the region of origin plus the transport cost per unit of good sent. This last sum is shown in Fig.9 and Fig.10 as a horizontal line. This line also helps to show how the SW level associated with this equilibrium is computed.

The final demand in region 2, D_2, will be composed of the sum of the quantity served by each firm ($D_2=D_{22} + D_{12}$), and this sum plugged into the total demand defines an equilibrium price, p_2. Consequently, SW being the sum of welfare in each region will corresponds to all shaded areas in Figures 8 and 9. The area in Fig.8 is exactly the same as the one discussed for the exporting region in the perfect collusion case, as we are assuming that region 2 does not export to region 1. The area in Fig. 9 is quite different from the perfect collusion case. It includes the net consumers' surplus plus the net local producer surplus, including a mark-up over the average cost of production incurred in all units sold, (p_2-mc_2), plus a smaller mark-up area defined over all the units sent from region 1 and sold in region 2, (p_2-mr_2-t_{12}). The total amount demanded in equilibrium in region 2 is in general bigger than the one obtained under perfect collusion.

An important difference also arises when computing the users' benefits based on the derived transport demand. In principle, one might then to think that we should be able to apply a similar approach as the one employed to derive the users' benefits in the perfect collusion case, that is, instead of using the final demand as in the perfect competition case we can refer to the marginal revenue curve to define a derived transportation demand. In the particular case that we are dealing with here, we would expect that the reference will be to the residual marginal revenue distinguish between the sales in the "local" and "foreign" markets. In fact, this approach is no longer valid here due to the strategic interaction that characterizes firm's behaviour. In Fig.10, the benefits from trade are valued as the area below the residual marginal revenue curve faced by firm 1 in region 2 and above the sum of the marginal revenue in the region of origin plus the transport cost per unit sent. This is the derived transportation demand in this case, defining the willingness to move goods from one region to the other. The total benefits from transporting T_{12} units of the good correspond to the triangular dark-shaded area in that figure. In Fig. 8, the benefits from trade are shown, as in the perfect collusion case, as the triangular dark-shaded area. For the particular equilibrium shown there the TCS, for a given transport cost level, will be the sum of these two areas.

The main complication with the computation of users' benefits (i.e. ΔTCS) under these circumstances is that due to the Cournot's expectations at play in the oligopoly case, the residual marginal demand and revenue curves do not have a fixed position from equilibrium to equilibrium. More specifically, when an original equilibrium is disturbed by an improvement in the infrastructure, the distribution of production across markets will vary and consequently also the location of the residual demands and marginal revenues. This complicates the analysis as we know it until now since we will have to incorporate these re-locations in residuals demands when computing the benefits in the transport market itself and subsequently when trying to compare them with the total social welfare change. Following Mohring (1956), among others, a linear path of integration is employed.

In Fig. 13 it is shown what is at stake when an initial equilibrium is disturbed by a decrease in the transport costs. The reduction in this cost increases the profitability for region 1 of allocating an extra unit of its production to region 2, as compared to its local market. This increase firm 1 exports and in turn weakens the position of region 2 local firm; a contraction in the residual demand for firm 2 results from this. This contraction in firm 2 residual demand generate a compensating expansion in region 1 residual demand which in turn is reflected in region 1 as an extra increase in production that leads to a higher marginal costs, compensating the increase in the residual marginal revenue in region 2. The process continues until a new equilibrium is achieved. The reduction in transport cost generates a re-composition in market share in region 2 that favours the exporting region firm.

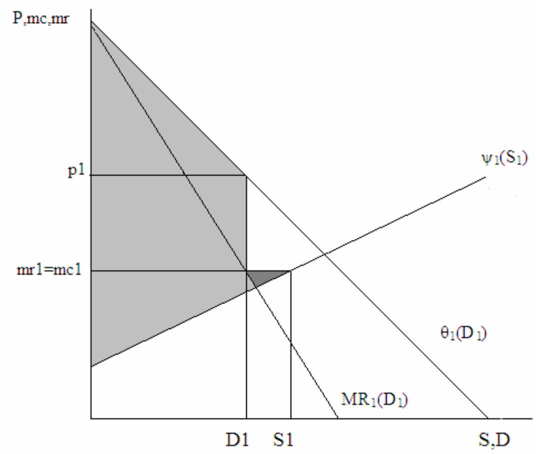


Figure 8.

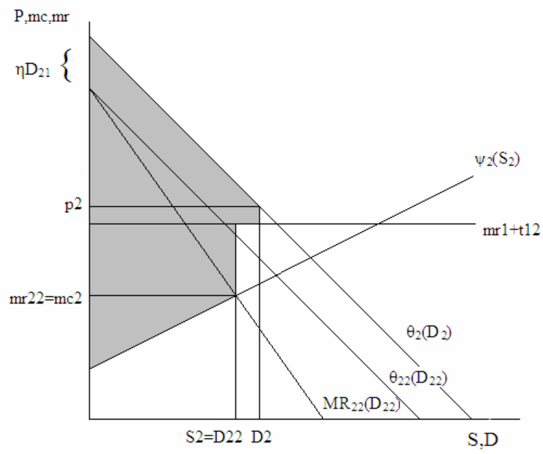


Figure 9.

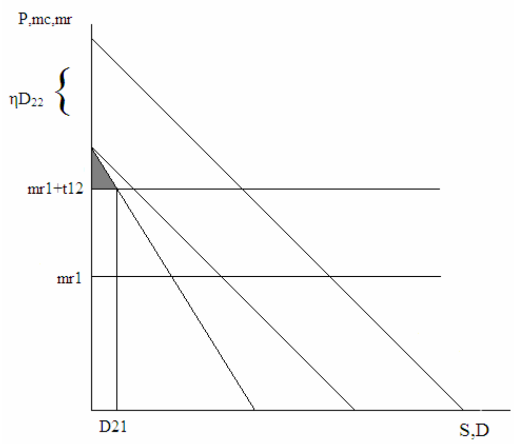


Figure 10.

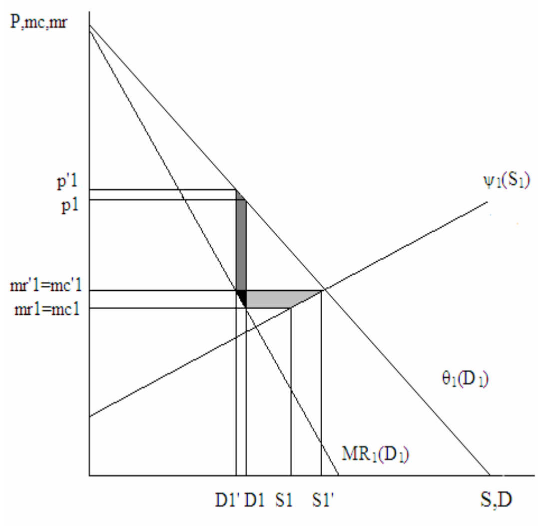


Figure 11.

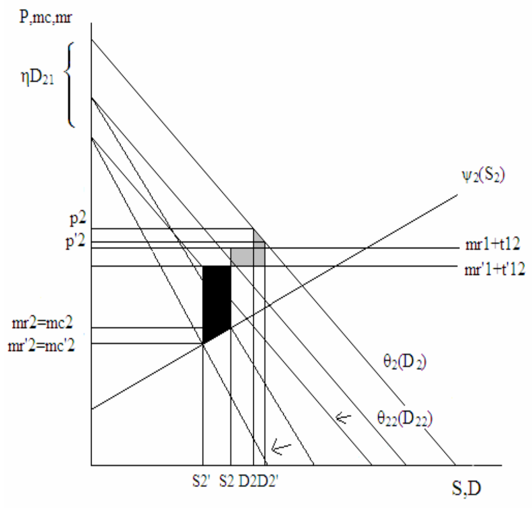


Figure 12.

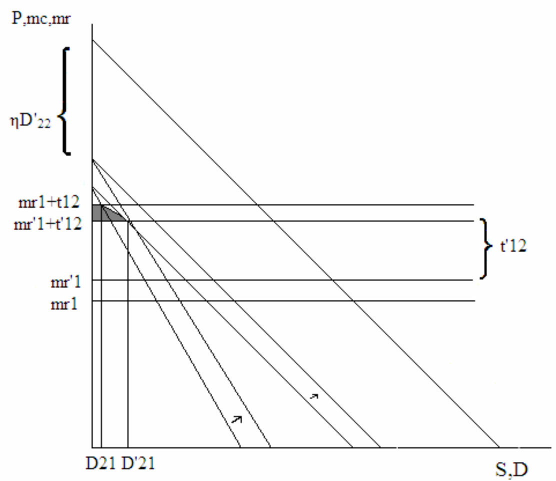


Figure 13.

The users' benefits arising from the transport infrastructure improvement will correspond to the sum of extra benefits from trade in region 1 and region 2. As before, the net benefits from trade in region 1 correspond to the horizontal (light) grey area in Fig. 11. A distinct case occurs in region 2 since the curve used to value the benefits from trade moves from the initial to the final equilibrium. To compute the benefits in this case we use a linear path of integration from the original to the final equilibrium assuming that the total benefits correspond to the sum of consecutive marginal movements in the direction of the final equilibrium. The corresponding benefits are the light-grey area in Fig.13.

Now we are ready to compare the users' benefits with the change in SW under Cournot-Nash competition. In general the change in SW will correspond to the change in the shaded areas in Fig. 8 and 9. In Fig. 11 and Fig. 12 we show, respectively, the change in SW of regions 1 and 2 after the infrastructure improvement defines a new equilibrium. As before, the case of region 1 is parallel to the case of perfect collusion, being the change in welfare the sum of a positive term defining the benefits from extra trade and a negative area defining the losses from consumers and producers of a reallocation of production from the local to the foreign region. There is again a small area in which the benefits and dis-benefits associated with the new equilibrium are cancelling out. For region 2, where there is competition between firms in supplying the market, the situation is significantly different. Increased competition faced by firm 2 results in a market share loss, S_2-S_2', for this firm. Consequently its profits are reduced in the mark-up that was before applied to these units, and this is reflected as a negative vertical area in black. Firm 1, in contrast, increases its mark-up over the previous units and also increases its profits applying the new equilibrium mark-up over the extra units sell in region 2. Part of the extra profits correspond to a transfer from firm 2 and does not appear as welfare increase. The other part appears as a horizontal positive area in light grey. A final positive triangular area in light grey at the top corresponds to the benefits going to the consumers from extra consumption at a lower price that additionally is not translated as extra profits to the producers.

It is clear from the previous discussion that the change in SW under Cournot-Nash competition will in general not coincide with the users' benefits captured in the transport market. Even more, the former could be higher or lower than the latter, with the final magnitude defined among other things by the elasticity of demand in each market, the number of regions-firms and the shape of the transportation network. There is also not possible to derive a clear link between the two measures of benefits as was the case for the perfect competition and perfect collusion case. The presence of strategic interactions among firms complicates the comparison between benefits measures.

In the next section we discuss in detail the general case of n regions and formalize the equilibrium condition discussed conceptually thus far. Additionally, congestion in transport and alternative network configurations are discussed. This is also in preparation for the numerical exercises that illustrate and complement the previous sections arguments.

3. Surplus equivalence in complex networks with congestion effects

In this section we discuss in more detail the main building blocks of SPE models. The focus of this paper is on inter-regional (e.g. between cities) freight flows. There are two main classes of network models which have been used to analyze inter-region freight movements: *spatial price equilibrium* models and *freight network equilibrium models* (Harker, 1986). The first class focuses on the producer-consumer-shippers interactions in the economy without explicitly determining the microeconomics of these activities. The transportation sector is represented by a directed graph with nodes and links. The transportation costs are not derived from a model of carrier behavior, but are stated as fixed values or as functions of the flows on a discrete network. Additionally producers' and consumers' behavior is incorporated by defining supply and demand functions for each region. In the general case -independently of the producer's industrial organization- the shippers are assumed to behave according to the following two equilibrium principles[26]:

[26] An important distinction arises in a Cournot-Nash type of competition as compared to perfect competition and perfect collusion. In the former the marginal revenues in the origin and destination both refers to the same firm whereas in the other two types of industrial organization they refer to the region.

- If there is a flow of commodity i from any pair of regions (k,l), then the marginal cost of commodity i in k plus the transportation costs from k to l will equal the marginal revenue (e.g. equal price under perfect competition) of the commodity in l.
- If the marginal costs of commodity i in k plus the transportation costs from k to l is greater than the marginal revenue of commodity i in l, then there will be no flow from k to l.

The SPE were originally described qualitatively in Enke (1951), and later formalized by Samuelson (1952) and Takayama and Judge (1964, 1970). In our work we use formulations and extensions from Florian and Los (1982), Friesz et al (1983) and Dafermos and Nagurney (1985), among others.

The second major class of predictive intercity freight models is the *freight network equilibrium models*, in which the focus is on the shipper-carrier interaction. The generation of trips from each region is imposed and assumed to be known in this type of models, in contrast with the SPE that solves for the trip (e.g. trade flows) generation via the interaction between supply and demand functions embedded in a network. An attempt to integrate both types of models is presented in Harker and Friesz (1986a, 1986b) as a *generalized spatial price equilibrium*, and can be seen as a possible extension of the model used here. In the rest of the paper, the focus will be on SPE.

Spatial price equilibrium is also one of the two most studied steady-state concepts of network equilibrium (Friesz, 1985). The other type is the *user equilibrium* which is mostly employed in urban passenger networks but as pointed out in Florian and Los (1981) and Harker (1985), can also be fully consistent with SPE featuring multi-paths for a given origin-destination (O-D) pair. This last point will be emphasized in the numerical simulations section when considering trans-shipment nodes and average cost pricing in the transportation sector. [27]

Network and industry configuration

In this section we first specify the network topology for a homogeneous one-commodity economy with single paths between each O-D pairs. [28] The discrete shipper network is represented by a finite-directed graph, $G[L,W]$, with L and W denoting the full set of nodes and arcs, respectively. The indices i, j, k, and l refer to nodes of the network. Define $W = \{w = (ij);\ i \in L, j \in L\}$, to be the set of all origin-destination pairs connecting pairs of regional centroids of trade represented by pairs of nodes in L. All

---

Additionally, it should be notice that the conditions are given in very general terms, for instance, in the case of perfect competition the marginal revenue is equal to the equilibrium price.

[27] As will be discussed later in the paper when more than two paths are possible a region can play the role of trans-shipment even though it has positive production and consumption itself; a users' equilibrium arise in trade flows going through alternative paths for the same origin-destination pair. For further discussion on users' equilibrium see Fernandez and Friesz (1983). Additionally Dafermos and Nagurney (1985) present an alternative view of the incorporation of user equilibrium in spatial price equilibrium models.

[28] The case for path choice will be introduced when discussing transshipment nodes in the simulations section. Harker (1985a) introduces the concept of path variables.

interactions within agents in the same centroid are conducted through the price system so an implicit assumption of zero transport cost for trade within each centroid applies. [29]

For each region $l \in L$, S_l is the supplied quantity in this region and D_l is the demanded quantity. The flow between O-D pair $(ij) \in W$ is denoted by $T_{(ij)}$. Conservation of flow in every region implies:

$$S_l - D_l + \sum_{i \in L} \sum_{(il) \in W} T_{(il)} - \sum_{j \in L} \sum_{(lj) \in W} T_{(lj)} = 0 \qquad (1)$$

for all $l \in L$. The market clearing condition requires that total demand equal total supply, which is obtained summing up over all $l \in L$ in (1):

$$\sum_{l \in L} D_l - \sum_{l \in L} S_l = 0 \qquad (2)$$

For each O-D pair, a function $c_{(ij)}(T_{(ij)}, K_{(ij)})$ is defined as the average (marginal private) cost incurred in transporting an unit of good between i and j. This is the cost that has to be paid when shipping between O-D pair $(ij) \in W$, and depends negatively in infrastructure capacity in the relevant link, $K_{(ij)}$, and positively in the transport flow, $T_{(ij)}$, in that link. [30] Furthermore $c_{(ij)}(\cdot,\cdot)$ is continuous, monotone and strictly convex in $T_{(ij)}$, implying that we are focusing on situations where congestion effects occur in each link. There is a cost per-unit of infrastructure provided that will represent the cost side of a CBA in our context, but we assumed it null at this point in order to focus on user's benefits. Allowing this cost to be different from zero will made possible to analyze the interaction of first and second best policies in infrastructure provision within the surplus equivalence question (Meléndez-Hidalgo et. al. 2006).

Let us define, for each region $l \in L$, $\psi_l(S_l)$ as the inverse supply function with associated long-run supply function $S_l(\psi_l)$, and $\theta_l(D_l)$ as the inverse demand function with associated long-run demand function $D_l(\theta_l)$. These functions are assumed, without loss of generality, separable in prices and quantities for simplicity in computation of equilibrium and simulations (Friesz et.al. 1983).

Perfect competition

The competitive SPE conditions compatible with atomistic behavior in production and consumption (Samuelson, 1952) can be written as:

$$\text{if } T_{(ij)} > 0, \ then\ \psi_i(S_l) + c_{(ij)}(T_{(ij)}) = \theta_j(D_j) \qquad for\ all\ (ij) \in W.$$
$$\text{if } \psi_i(S_i) + c_{(ij)}(T_{(ij)}) > \theta_j(D_j), \ then\ T_{(ij)} = 0 \qquad for\ all\ (ij) \in W.$$

These conditions state that a shipper or trader will purchase the good in a perfectly competitive market at node i for the price $\psi_i(S_i)$, and will incur an economic price of

[29] Thore (1991) considers transport cost for shipping within a given region.
[30] Following Harker (1985) and Florian and Los (1981), $c_{(ij)}$ is the economic price (freight rate plus value of level of service) if a transportation firm offers service between $(ij) \in W$.

transportation, $c_{(ij)}(T_{(ij)})$, for the move of goods from i to j. The shipper takes this economic price of transportation as given. If the price he could receive in the perfectly competitive market at region j, $\theta_j(D_j)$, exceeds the sum of the purchase price and the transportation charges (= *delivered price*), then he will continue to ship the good until the sale price and the delivered price are equal. If the delivered price is greater than the sale price, the shipper will lose money in shipping from i to j and hence he will not ship anything between this O-D pair (condition (ii)).

Conditions (i) and (ii) and (1) can be formulated as a complementarity problem, a variational inequality problem or as a simple mathematical programming problem depending on the functional forms and network conditions assumed. In our context the mathematical program describing the decentralized solution can be written as:

$$\text{Max} \qquad \sum_{l \in L} \int_0^{D_l} \theta_l(s)ds - \sum_{l \in L} \int_0^{S_l} \psi_l(s)ds - \sum_{(ij) \in W} \int_0^{T_{(ij)}} c_{(ij)}(s, K^0_{(ij)})ds \tag{3}$$

$$\text{subject to} \quad S_l - D_l + \sum_{i \in L} \sum_{(ij) \in W} T_{(il)} - \sum_{j \in L} \sum_{(ij) \in W} T_{(lj)} = 0 \;\; for \;\; all \;\; l \in L$$

$$D_l, S_l \geq 0 \;\; for \;\; l \in L$$
$$T_{(ij)} \geq 0 \;\; for \;\; all \;\; (ij) \in W$$
$$M_l^S \leq S_l \leq U_l^S \;\; for\ all\ l \in L;$$
$$M_l^D \leq D_l \leq U_l^D \;\; for\ all\ l \in L;$$
$$M_l^T \leq T_{(ij)} \leq U_l^T \;\; for\ all\ (ij) \in W.$$

where M_l^Z, U_l^Z, with $Z = \{D,S\}$, corresponding to upper and lower bounds on the supplies, demands and O-D flows as discussed in Harker(1986). The formulation in (3) reflects a decentralized initial equilibrium before the increase in capacity $K_{(ij)}$, then the 0 superscript in the capacity index.

A distinction should be made here between first and second best policies in the transport sector depending on if congestion charging is in place or not (Williamson et.al., 2001). This will affect the objective function in (3) as well as in subsequent formulations with alternative product market industrial organizations. Is not our aim in this paper to consider optimal policies per-se. In what follows we will ignore any other possible instruments at hand to the social planner for optimizing total welfare, except from congestion taxes. The social planner will either choose between implementing congestion charging in every link of the network, at a rate $\tau_{(ij)}$, and a policy of no-intervention at all. [31] In the latter case, as shown in the last element of the objective function in (3), a Wardropian user optimal equilibrium concept for network flows and costs is at stake, i.e. each shipper seeks to minimize its transportation cost.

[31] In a latter section when we consider more than one path between origin-destination pairs, these two policy options in transportation markets will clearly correspond to Wardropian System and User Equilibrium, respectively; see Harker (1985) and Fernandez and Friesz (1983). At this point the first-best level of the congestion charge is calculated as is usual in the literature; more explicitly in our context we follow van den Bergh and Verhoef (1996).

In equilibrium used paths between any O-D pair have the same cost and unused paths have costs at least as high (Fisk, 1987). The existence of alternative paths for a given O-D pair is not explicitly considered in the formulation that we are using here for the SPE, and this will be amended later in the path formulation section and incorporated in the numerical simulations section of the paper when discussing transshipments nodes for better representing transportation infrastructure. Even though, it is important to notice that the conservation flow condition in (1) implicitly allows an exporting region to use a "third-region" for conducting its shipments to a final importing region. The third region in that case will be playing the role of a transshipment node (in this case with its own market) and this might lead to more than one shipping paths for a particular O-D pair. At this stage of the paper we just need to notice that this possibility exists and consequently we introduce the Wadropian equilibrium that applies depending on if there is congestion charging in place or not (i.e. user or system equilibrium).

The average cost function in (3) is replaced by the marginal cost function (Beckmann, 1967; Yang and Huang, 1998) when first-best congestion charging is applied, leading to a Wadropian social optimal equilibrium for network flows and costs. The marginal cost function is defined as:

$$c^{mg}_{(ij)}(T_{(ij)},K^0_{(ij)}) = c_{(ij)}(T_{(ij)},K^0_{(ij)}) + T_{(ij)}\frac{dc_{(ij)}(T_{(ij),}\cdot)}{dT_{(ij)}} \quad \text{for all } (ij) \in W$$

Establishing tolls, $\tau_{(ij)}$, given by:

$$\tau_{(ij)} = T_{(ij)}\frac{dc_{(ij)}(T_{(ij)},\cdot)}{dT_{(ij)}} \qquad \text{for all } (ij) \in W$$

Consequently the only change in the objective function is to consider an additional (negative) term representing a congestion charge, which is assumed to be totally passed on to the consumers. [32] The mathematical program describing the decentralized solution when congestion charging is practiced will be:

$$\text{Max} \qquad \sum_{l\in L}\int_0^{D_l}\theta_l(s)ds - \sum_{l\in L}\int_0^{S_l}\psi_l(s)ds - \sum_{(ij)\in W}c_{(ij)}(T_{(ij)},K^0_{(ij)})T_{(ij)} \tag{3b}$$

subject to the same equalities and inequalities. Observe that whereas the objective function in (3) has no economic interpretation, the one in (3b) corresponds to the total social welfare for the economy under analysis.

JD shows that, total welfare changes from a situation with no trade between regions to one with positive trade across regions will correspond to the transport consumers' surplus resulting in the decentralized solution with positive trade, measuring this surplus based on a derived transport demand. This holds for non-flow dependent transport costs and perfect competition both in production and consumption. When transport cost does depend on the flow, it is straightforward to show that this relationship will continue to

[32] For a discussion on congestion charges incidence under imperfect competition in production see Van Dender (2004).

hold if the perfect competition assumption is kept for production and consumption and a complete CBA is implemented (e.g. including congestion taxes revenue changes). We will go into more detail on this relation later in the paper. Before that we discuss how the total social welfare is computed in both the before and after transport-cost-change equilibriums.

Social welfare corresponds to the sum of consumer's surplus and producer's surplus in every region (including transport costs) and in perfect competition (SW_c) is computed as follows:

$$SW_c = \sum_{l\in L}\int_0^{D_l}\theta_l(s)ds - \sum_{l\in L^I}\theta_l(D_l)S_l - \sum_{l\in L^I}\theta_l(D_l)[D_l - S_l] - \sum_{l\in L^E}\theta_l(D_l)D_l$$

$$+\sum_{l\in L^I}\theta_l(D_l)S_l + \sum_{l\in L^E}\theta_l(D_l)D_l + \sum_{l\in L^I}\theta_l(D_l)[D_l - S_l] - \sum_{l\in L}\int_0^{S_l}\psi_l(s)ds$$

$$-\sum_{(ij)\in W}c_{(ij)}(T_{(ij)},K^z_{(ij)})T_{(ij)}$$

with L^I and L^E referring to the set of import and export nodes, respectively. The SW_c reduces to:

$$SW_c = \sum_{l\in L}\int_0^{D_l}\theta_l(s)ds - \sum_{l\in L}\int_0^{S_l}\psi_l(s)ds - \sum_{(ij)\in W}c_{(ij)}(T_{(ij)},K^Z_{(ij)})T_{(ij)} \qquad (4)$$

which is gross benefits to consumers minus total costs including the cost incurred in transportation.

When there is congestion charging the government surplus will become explicit (non-zero) in the social welfare function. [33] In perfect competition the difference between prices in each node will match the sum of average costs plus the congestion charge, $\tau_{(ij)}$, then the social surplus for the decentralized solution with congestion charging will be:

$$SW_c = \sum_{l\in L}\int_0^{D_l}\theta_l(s)ds - \sum_{l\in L}\int_0^{S_l}\psi_l(s)ds - \sum_{(ij)\in W}c_{(ij)}(T_{(ij)},K^0_{(ij)})T_{(ij)} - \sum_{(ij)\in W}\tau_{ij}T_{(ij)} + \sum_{(ij)\in W}\tau_{ij}T_{(ij)} \qquad (5)$$

Expression (5) is equivalent to (4).

Perfect collusion (spatial monopoly without resale)

Takayama and Judge (1971) formulated their model in a simple network where the monopolist takes as given the transport costs, e.g. transport rate. Harker (1986) shows how a different interpretation of the transportation costs leads to two distinct models of spatial monopolies. In the first case, it is recognized that in most situations the producing firm will buy the transportation service from a motor carrier, railroad, or other carrier. Under these conditions producers will take the economic price of transportation as given when deciding how much to produce and send to each market. [34] The following mathematical program results:

$$\text{Max} \qquad \sum_{l\in L}\theta_l(D_l)D_l - \sum_{l\in L}\int_0^{S_l}\psi_l(s)ds - \sum_{(ij)\in W}\int_0^{T_{(ij)}} c_{(ij)}(s, K_{ij}^0)ds \tag{6}$$

subject to (1), and D_l, $S_l \geq 0$ for all $l \in L$, and $T_{(ij)} \geq 0$ for all $(ij) \in W$.

From the Kuhn-Tucker conditions of this problem we can derive the transport market equilibrium and consequently compute the transportation consumption surplus. These conditions are necessary and sufficient and letting π_l denote the dual variable of constraint (1) we have:

$$\begin{array}{ll}
(\theta_l + D_l\theta_l^{'} - \pi_l)D_l = 0 & \text{for all } l \in L \\
\theta_l + D_l\theta_l^{'} - \pi_l \leq 0 & D_l \geq 0 \\
(-\psi_l + \pi_l)S_l = 0 & \text{for all } l \in L \\
-\psi_l + \pi_l \leq 0 & S_l \geq 0 \\
(-c_{(ij)} + \pi_j - \pi_i)T_{(ij)} = 0 & \text{for all } (i,j) \in W \\
-c_{(ij)} + \pi_j - \pi_i \leq 0 & T_{(ij)} \geq 0
\end{array}$$

what these conditions imply in terms of trade is that if there is flow between region i and j, then the private marginal cost of transporting one unit ($c_{(ij)}$) plus the marginal production cost in i (ψ_i) equals marginal revenue ($\theta_j + D_j\theta_j^{'}$). If these conditions do not hold for a particular O-D pair, then there is not trade in that particular pair. If there is congestion charging in practice, the firm's profit (or cartel) maximization problem will be written as:

$$\text{Max} \qquad \sum_{l\in L}\theta_l(D_l)D_l - \sum_{l\in L}\int_0^{S_l}\psi_l(s)ds - \sum_{(ij)\in W} c_{(ij)}(T_{(ij)}, K_{(ij)}^0)T_{(ij)} \tag{7}$$

subject to (1) and non-negativity constraints. In both (8) and (9) we are considering $c_{(ij)}$ as the average cost price of transport. The Kuhn-Tucker conditions in this case are:

$$\begin{array}{ll}
(\theta_l + D_l\theta_l^{'} - \pi_l)D_l = 0 & \text{for all } l \in L \\
\theta_l + D_l\theta_l^{'} - \pi_l \leq 0 & D_l \geq 0 \\
(-\psi_l + \pi_l)S_l = 0 & \text{for all } l \in L \\
-\psi_l + \pi_l \leq 0 & S_l \geq 0 \\
(-c_{(ij)} - T_{(ij)}c_{(ij)}^{'} + \pi_j - \pi_i)T_{(ij)} = 0 & \text{for all } (i,j) \in W \\
-c_{(ij)} - T_{(ij)}c_{(ij)}^{'} + \pi_j - \pi_i \leq 0 & T_{(ij)} \geq 0
\end{array}$$

[34] Florian and Los (1982) discuss in much detail the different interpretations possible for the industrial organization of the transport sector.

The only difference here is that instead of the marginal private cost of transportation is the marginal “social” cost the one used in the previous conditions.

Concerning the social welfare, in the first formulation it will correspond to the sum of consumer’s surplus and firm’s profits minus transportation costs for each regional market, while in the case of congestion charging we will explicitly add the government surplus although as before, it will cancel out against toll payments by transporters. The social welfare function under monopoly will have the same reduced form as in perfect competition (SW_m):

$$SW_m = \sum_{l \in L} \int_0^{D_l^e} \theta_l(s)ds - \sum_{l \in L} \int_0^{S_l^e} \psi_l(s)ds - \sum_{(ij) \in W} c_{(ij)}(T_{(ij)}^e, K_{(ij)}^0) T_{(ij)}^e \tag{8}$$

which is equivalent to (5). For the case where congestion charging is in place a parallel case also be drawn with respect to the competitive situation. The addition of the government surplus will result in an expression of welfare equal to (6).

Cournot-Nash oligopoly

The model we implement here follows closely Harker (1986). If Q denotes the set of firms operating in the market and each firm $q \in Q$*'s* control only one production site ($Q = L$). The total amount demanded in region $l \in L$ is given by D_l. Defining D_{lq} as the amount shipped by firm q to region l (or the amount demanded by the consumers in region $l \in L$ from firm q), and $\tilde{D}_{lq}$ as the amount supplied by all other firms to region $l \in L$:

$$\tilde{D}_{lq} = \sum_{j \in Q, j \neq q} D_{lj}$$

and

$$D_l = \sum_{j \in Q} D_{lj} \tag{9}$$

we can defined a spatial Cournot-Nash equilibrium as the solution, for every firm, of the following programming program:

$$\text{Max} \quad \sum_{l \in L} \theta_l(D_{lq} + \tilde{D}_{lq}) D_{lq} - \int_0^{S_q} \psi_q(s)ds - \sum_{(qj) \in W} \int_0^{T_{(qj)}} c_{(qj)}(s, K_{(qj)}^0)ds \tag{10}$$

subject to an optimal strategy vector x_q, followed by all other firms (including supplies, demands and transport flows). The amount supplied by all other firms to region $l \in L$,

$\tilde{D}_{lq}$, is taken as given in the perspective of a particular firm. [35] At this point, we also assume $c_{(ij)}(T_{(ij)})$ is given, even in the presence of congestion charging. This last assumption could be seen controversial when a more realistic view of the interaction shipper-carrier is considered in a Cournot-Nash environment since a partial market power should be recognized for either size in this context. If for example, transport services are produced under oligopoly conditions then in the presence of congestion a partial internalization of the externality created when transporting using a fixed capacity should be recognized. This is precisely the point raised recently for airline transportation in Brueckner (2004) and Verhoef and Pels (2004).

For the mathematical program formulation in (13) there are associated Kuhn-Tucker conditions similar to the ones obtained for the monopoly case. Important insight can be obtained from these conditions:

$$(\theta_l(\tilde{D}_{lq} + D_{lq}) + D_{lq}\partial\theta_l\,(\tilde{D}_{lq} + D_{lq})/\partial D_{lq} - \pi_l)D_{lq} = 0 \qquad \text{for all } l \in L$$

$$(\theta_l(\tilde{D}_{lq} + D_{lq}) + D_{lq}\partial\theta_l\,(\tilde{D}_{lq} + D_{lq})/\partial D_{lq} - \pi_l) \le 0 \qquad D_{lq} \ge 0$$

$$(\theta_q(\tilde{D}_{lq} + D_{qq}) + D_{qq}\partial\theta_q\,(\tilde{D}_{lq} + D_{qq})/\partial D_{qq} - \pi_q)D_{qq} = 0 \qquad \text{for q}$$

$$(\theta_q(\tilde{D}_{lq} + D_{qq}) + D_{qq}\partial\theta_q\,(\tilde{D}_{lq} + D_{qq})/\partial D_{qq} - \pi_q) \le 0 \qquad D_{qq} \ge 0$$

$$(-\psi_q + \pi_q)S_q = 0$$

$$-\psi_q + \pi_q \le 0 \qquad S_q \ge 0$$

$$(-c_{(qj)} + \pi_j - \pi_q)T_{(qj)} = 0 \qquad \text{for all } (q, j) \in W$$

$$-c_{(qj)} + \pi_j - \pi_q \le 0 \qquad T_{(qj)} \ge 0$$

The first two conditions imply that if there is demand in a particular region for the product of firm *q*, π_l will equal the marginal revenue for firm *q* in that region. Similarly conditions three and four imply that if there is demand in the region that the firm is located, then π_q will equal the marginal revenue for firm *q* in its location. The last four conditions concern costs of production and transportation. Conditions six and seven imply that if there is production, π_q will equal the marginal cost of production whereas the last two conditions state that if there is flow between region *j* and the production site of firm *q* then the cost of transportation plus the marginal cost of production will equal the marginal revenue for firm *q* in region *j*. An important distinction with respect to the perfect collusion case is that the marginal revenue functions here are based on a residual demand, instead of the total demand in a particular market. This residual demand is the part of total demand that firm *q* considers it is facing, after subtracting the expectation of supply for that market from other firms.

[35] Equilibration algorithms for oligopoly models usually employ diagonalization/relaxation algorithms. These algorithms applied to oligopoly problems have an interesting interpretation as a Cournot expectation dynamic process. This process assumes that each player (firm) establish its strategy (production level) by trying to maximize its individual profit, considering that the other players will set their own strategies as the same as the last round. After all the firms make their choice, the game moves to another state and the firms play again. This process continues until no firm can increase its profits. The Cournot expectation process can be interpreted as a Jacoby method.

Concerning total welfare computations a similar procedure as the one followed for the perfect collusion case applies here. Total social welfare will depend again on the congestion pricing policy but in general will consist of the sum of consumer's surplus and firm's profits (including transport costs) over all regions.

$$SW_o = \sum_{l \in L} \int_0^{D_l} \theta_l(s)ds - \sum_{l \in L} \theta_l(D_l)D_l + \sum_{q \in Q} \sum_{l \in L} \theta_l(D_l)D_{lq} - \sum_{q \in Q} \int_0^{S_{lq}} \psi_{lq}(s)ds$$

$$- \sum_{q \in Q} \sum_{(qj) \in W} c_{(qj)}(T_{(qj)} K^0_{(qj)})T_{(qj)} \tag{11}$$

which using (11) reduces to:

$$= \sum_{l \in L} \int_0^{D_l} \theta_l(s)ds - \sum_{q \in Q} \int_0^{S_{lq}} \psi_{lq}(s)ds - \sum_{(ij) \in W} c_{(ij)}(T_{(ij)}, K^0_{(ij)})T_{(ij)} \tag{12}$$

Following a similar path of argumentation as in the cases of perfect competition and perfect collusion, a similar expression can be shown to apply for the case where congestion pricing is in practice. A final important distinction to notice here is that the level of total supply will lie between these two extremes and it can be shown (Haurie and Marcotte, 1996) that in the limit, when the number of firms goes to infinity, it will coincide with the perfect competition production level.

Path formulation: trans-shipment nodes

In this section we discuss how the previous model formulation changes when trans-shipment nodes are explicitly considered. A variable path formulation explicitly considers trans-shipment nodes which better represent certain types of transport infrastructure. The network is again a finite-directed graph, *G[N,A]*, with *N* and *A* denoting the full set of nodes and arcs respectively. The set of centroids L is a subset of *N*. Thus, the network contains nodes that are neither origins nor destinations, but represent transportation facilities such as rail yards or ports (Harker, 1986). The indices *a* and *p* refer respectively to an arc or link and a path in the network. The indices *i, j, k,* and *l* refer to nodes of the network and P_{ij} denote the set of all paths connecting node *i* to node *j*, consequently $p \in P_{ij}$ is a particular path between regions *i* and *j*. For $\delta_{ap} = 1$ link *a* is contained in path *p* while $\delta_{ap} = 0$ otherwise. The flow on link *a* is represented by f_a while h_p is the flow on path p; consequently $f_a = \Sigma_{p \in Pij}\, \delta_{ap}\, h_p$. The average (unit) cost function of transportation on arc *a* will be denoted by $c_a(f)$, while the average cost on path p is $C_p(h) = \Sigma_a\, \delta_{ap}\, c_a$.

The equilibrium conditions change according to the product competition assumed. For instance, for the perfect competition case π_l denote the commodity price at node (market) *l*, $D_l(\pi)$ is the commodity demand at region *l*, $S_l(\pi)$ is the commodity supply at region *l* and *f, D, S, c* and π are the full row vectors of link flows, demands, supplies, unit link costs and prices. In a particular network with *m* arcs and *n* nodes (regions) a flow pattern *(f , π)* satisfying the following conditions is a competitive spatial price equilibrium:

$$h_p[C_p(h) + \pi_i - \pi_j] = 0 \ \ \forall \ \ (i, j, p \in P_{ij}) \qquad \text{(a)}$$

$$C_p(h)_p + \pi_i - \pi_j \geq 0 \ \ \forall \ \ (i, j, p \in P_{ij}) \qquad \text{(b)}$$

$$D_l(\pi) - S_l(\pi) + \sum_k \sum_{p \in P_{lk}} h_p - \sum_k \sum_{p \in P_{kl}} h_p = 0 \quad \forall l \qquad \text{(c)}$$

$$f_a - \sum_P \delta_{ap} h_p = 0 \ \ \forall a \qquad \text{(d)}$$

$$h \geq 0, \ \ \pi \geq 0 \qquad \text{(e)}$$

Conditions (a) and (b) consist of a generalization of the spatial price equilibrium conditions discussed in the previous formulation based on Samuelson (1952). Condition (a) implies that for utilized paths between a given origin-destination pair delivered commodity prices equals local ones. Condition (b) states that paths for which delivered price exceeds local price will not be utilized. The flow conservation condition is stated in (c); (d) is the definition of flow already given before while (e) corresponds to non-negativity conditions. Conditions (a) to (e) can also be used to represent multiple-mode and multiple-commodity equilibrium problems using the device of a multi-copy network originated in Aashtiani (1979). Similar modifications can be performed to cast the two other product competition arrangements in a path-formulation.

This closes the formalization of the SPE previously discussed in a graphical manner. In the next section the formalization of the models is employed to conduct two simulation exercises which illustrate and complement the previous discussion. The idea there is to construct two simple examples where the same conditions (i.e. parameters) holds for both cases except for the transportation network employed. Similar trade flow paths and magnitudes are obtained but it is shown that the different between infrastructure improvements benefits can be quite significant despite the similitude in demand and supply primitives. The shape of the transport network drives the results.

4. Simulation results

In this section we use numerical simulations to illustrate and complement the main conceptual results already provided in the paper. A three region economy embedded in two different transport networks is the workhorse employed to study the different impacts arising from an infrastructure investment under alternative competition arrangements. The two networks differ on the existence of a transhipment node. In both network cases the supply-demand nodes are characterized by the same set of parameters and the cost link parameters are such that similar patterns of trade are obtained. This is in order to capture in our results the distinguish effects arising from the differences in transport network shape. For both types of networks the following functional forms are used:

$\Psi_l(S_l) = \alpha_l + \beta_l S_l$ (Cost function)

$\Theta_l(D_l) = \rho_l - \eta_l D_l$ (Demand function)

$c_a(f_a) = \kappa_a + \nu_a f_a^2$ (Transport Cost function per link)

where $\alpha_l, \beta_l . \rho_l, \eta_l, \kappa_a$ *and* ν_a are constants. The transport cost function is interpreted as average private costs (taking into account level of service) and then expressions for the total costs and social marginal costs can be obtained from it.

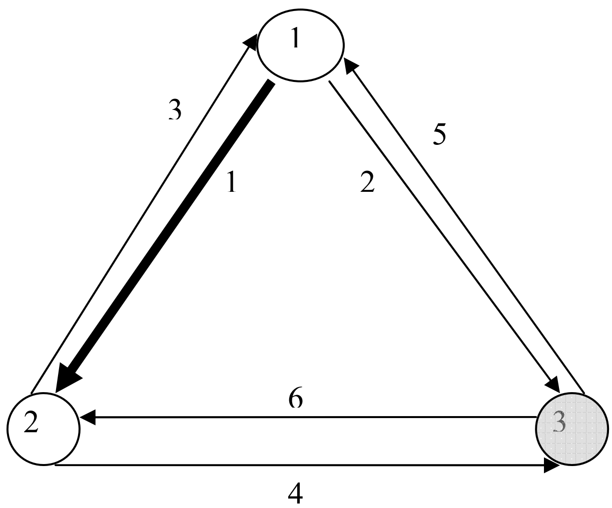
Table 1a shows the calibration of parameters for the demand and supply in the two network cases considered. These values together with the transport network parameters generate a certain pattern of trade across regions.

Table 1: Parameters for Production and Consumption Functions per Region.

Region l ∈ L	α	β	ρ	η
1	0.1	0.05	1.9	0.02
2	0.2	0.04	2.7	0.001
3	0.15	0.03	3.0	0.03

The first type of network extends the two region economy previously analyzed with a third production-consumption node and two related transport links. Again we are here considering a one-good economy. This case has not been analyzed in the literature and as an extension of the previous case, allows us to study network issues in the simplest setting possible. Table 2 shows the parameters for the transport cost function.

First network



An important issue raised by this case is that whenever there is a change in the cost of transporting through a particular link, the demand on other links will change resembling a general equilibrium effect in the transport markets. This interaction has been addressed recently in a different framework by Kidokoro (2003) and is included in the results shown later in Table 3 when computing the change in TCS in all links. For the cases of monopoly and oligopoly "indirect benefits" are present and captured by a level of the usual multiplier different that one. Under perfect competition the ratio of social welfare changes to user's benefits (i.e. multiplier) takes always a value of one, both for the case of no tolling and optimal tolling. For the latter, the change in congestion charges revenue adds up with the usual transport benefits.

Table 2: Parameters for Transportation Functions per Link.

Arc $a \in$ A	κ	ν
1	0.1	*0.04*
2	0.2	0.04
3	0.1	0.02
4	0.3	0.03
5	0.1	0.01
6	0.4	0.08

A reduction in 10% on the capacity related parameter (ν) in link 1 is the infrastructure improvement scenario considered; starting from the initial level in Table 2. Trade flows differ for the different competition arrangements. Under perfect competition trade flows are positive in link 1, 2 and 6 for all equilibriums. Region 1 exports to both regions 2 and 3; region 3 exports to region 2 and imports from region 1. Region 2 imports from the other two regions. This structure of trade includes route choice for region 1 exports to region 2. Is already established in the literature that a model with this features will show uniqueness of equilibrium only at the level of link flow, but multiple equilibria will arise at the level of path flow (Fernandez and Friesz, 1982). Taking this into account in the computation of changes in TCS implies calculating benefits at the level of link flow. For the monopoly case trade goes from regions 1 and 3 to region 2. Finally, when producers interact in a Cournot-Nash fashion the trade flows resemble the pattern of the perfect competition case, but in addition there is trade from region 3 to region 1. This extra flow is implying that there is two-way trade between these two regions. Table 1 show the results for the three alternative competition assumptions in production.

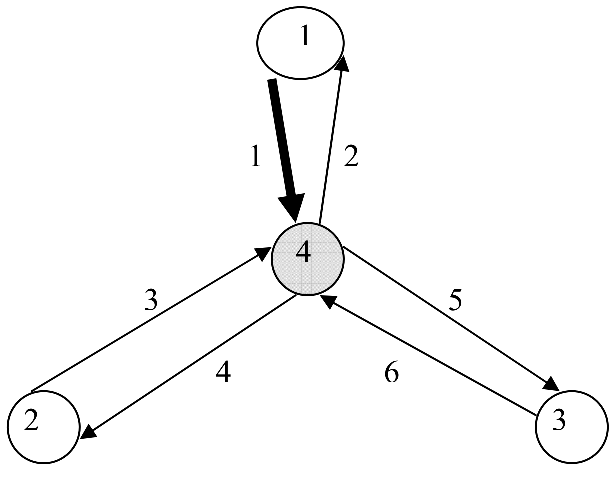
Since we assume either optimal or zero toll levels in all links, transport benefits for the perfect competition case (CSPE) will match economy-wide welfare changes in both cases. For the perfect collusion case the situation is quite different. Without tolls transport benefits overestimate the economy-wide effects with the gap explained both by profits made on the extra units (both imported and exported) and the net loss in consumer surplus for these extra units transferred from the exporting the to importing region. As shown in Table.3, as the link 1 capacity expands the overestimation persists when optimal tolling is implemented but turns into underestimation when no tolling is practised. This is so because even though the presence of congestion externalities affects welfare negatively, the fact that there is not toll allows the monopolist to reach output levels which bring profit and net consumer surpluses effects that compensate that overall compensate this externality. On the other hand, when congestion is optimal charged, the externality is eliminated but the level of output in which the monopolist is manoeuvring is such that the net profit effects is compensated by a negative net consumers surplus effect, then the multiplier is always below one. The Cournot-Nash case also shows overestimation of economy-wide benefits and under the parameter values we chose it is more severe than the perfect collusion case. The interesting point here is that under the same functional form for supply and demand, the strategic interaction among firms can create significant negative indirect effects even when congestion is charged. As capacity expands to reach a very small value for ν_1, eventually the gap between both multipliers shrink but there are still negative indirect benefits (i.e. disbenefits).

Table 3: Results from a three node network without transshipment nodes.

Transport Parameters		Multipliers					
Fixed	*Variable*	***CSPE*** No Tolling	***CSPE*** Optimal Tolling	***MONOP*** No Tolling	***MONOP*** Optimal Tolling	***OLIGOP*** No Tolling	***OLIGOP*** Optimal Tolling
0.1	0.04						
0.1	0.036	1.000	1.000	0.983	0.946	0.456	0.723
0.1	0.032	1.000	1.000	1.001	0.949	0.490	0.728
0.1	0.029	1.000	1.000	1.017	0.952	0.523	0.734
0.1	0.026	1.000	1.000	1.032	0.955	0.553	0.739
0.1	0.023	1.000	1.000	1.048	0.959	0.587	0.746
0.1	0.021	1.000	1.000	1.061	0.963	0.617	0.751
0.1	0.019	1.000	1.000	1.073	0.966	0.643	0.757
0.1	0.017	1.000	1.000	1.086	0.970	0.671	0.763
0.1	0.015	1.000	1.000	1.099	0.975	0.702	0.770
0.1	0.0135	1.000	1.000	1.112	0.980	0.731	0.776
0.1	0.005						
0.1	0.0045	1.000	1.000	1.202	1.044	0.776	0.853

The previous network was the simplest possible extension from J-D and J-D and F when considering congestion. Even in this network there is route choice in the perfect competition case when trade goes from region 1 to region 2 directly but also through region 3. Route choice is not possible in the case of monopoly or Cournot-Nash oligopoly. The next network structure model explicitly a trans-shipment (without route choice) in order to compare the magnitude of the multiplier arising simply from a different network structure.

Second network



The second network corresponds to a case where all trade flows have to arrive first to a consolidation point (trans-shipment node), where network resources are shared between shippers. Table 4 shows the parameters for the transport cost function in each link. The infrastructure improvement is again in link 1. The pattern of trade is similar to the one discussed for the previous network. An distinctive situation occurs for the case of monopoly and Cournot-Nash because trade goes from regions 3 and 1 to region 2, implying that the last link to reach region 2 is shared by these two flows. Once, the infrastructure is improved in link 1 this generate negative spillover effects to the flow going from 3 to 4. In an extreme situation this can lead to negative social welfare changes resembling a Braess paradox type of phenomena at the level of the entire economy. [36] The sign of the indirect benefits change for the monopoly case and the magnitude of the extra benefits (above the usual transport benefits) is considerable. Again, this benefits are higher without tolling because due to the demand and supply schedules parameters, the monopolist without congestion charging can reach output level where the net consumer surpluses effects turns out to be positive and adds up with the profits effect.

Table 4: Parameters for Transportation Functions per Link.

Arc $a \in$ A	κ	ν
1	0.05	*0.02*
2	0.05	0.01
3	0.05	0.01
4	0.07	0.05
5	0.03	0.01
6	0.02	0.01

Table 5: Results from a three node network with one transhipment.

		Multipliers					
Transport Parameters		***CSPE***	***CSPE***	***MONOP***	***MONOP***	***OLIGOP***	***OLIGOP***
Fixed	*Variable*	No Tolling	Optimal Tolling	No Tolling	Optimal Tolling	No Tolling	Optimal Tolling
0.05	0.02						
0.05	0.018	1.000	1.000	4.223	2.090	0.406	0.513
0.05	0.016	1.000	1.000	4.010	2.130	0.442	0.518
0.05	0.014	1.000	1.000	3.805	2.178	0.475	0.523
0.05	0.0126	1.000	1.000	3.637	2.229	0.507	0.528
0.05	0.011	1.000	1.000	3.491	2.220	0.535	0.533
0.05	0.01	1.000	1.000	3.368	2.207	0.560	0.538
0.05	0.009	1.000	1.000	3.273	2.247	0.579	0.542
0.05	0.008	1.000	1.000	3.178	2.294	0.735	0.546

In the oligopoly case, considerable indirect effects arise. Again, as congestion is realised through capacity improvements, the multipliers increases. Under the parameter chosen the strategic interaction across firms is such that the re-composition of market shares and the net consumer welfare effects compensate the usual transport benefits.

[36] Venables (1999) discuss this possibility for the case of transport market with a similar transport network shape.

5. Conclusions

Our aim in this paper has been to study in more detail the difference in benefits measures arising from an infrastructure improvement employing a spatial price equilibrium framework, originating in Samuelson (1952). Previous contributions had looked at the topic in a simplified fashion assuming extreme product competition arrangements and network environments. We contribute with this literature by extending these previous contributions in several ways. First, apart from the usual perfect competition assumption we consider monopolistic behaviour across regions without the possibility of resale and Cournot-Nash interaction across regions-firms. Additionally we employ flow depend transport cost functions and allow for the existence of optimal tolling or no tolling at all. Finally, we explain in detail the formalization of the models to more than two regions and complex networks and illustrate our main points with numerical simulations in a three region network with and without transhipment nodes.
Our main contribution is to disentangle the relationship between economy-wide changes and transport users' benefits under spatial monopoly without resale and spatial Cournot-Nash firm interaction. For the former case, the relationship is clear and driven by the net effect of net consumer surpluses changes and net profit changes. On the contrary, for the latter there is no clear connection that we can derive between transport users' benefits and welfare changes because the strategic interaction between firms makes difficult the measurement of the transport benefits, as compared with the perfect competition and collusion cases. In the last section, some illustrative numerical simulations are shown which lead us to conclude that the sign and magnitude of the indirect benefits are significantly influenced by the shape of the transportation network.

References

Aashtiani, H. Z. and Magnanti, T. L. (1981) "Equilibria on a congested transportation network" *SIAM Journal on Algebraic and Discrete Methods* 2, pp. 213–226.

Brander, J. A. (1981) "Intra industry trade in identical commodities" *Journal of International Economics* 11, pp. 1–14.

Brander, J. and Krugman, P. (1983) "A 'Reciprocal Dumping' Model of International Trade" *Journal of International Economics*, 15, pp. 313-321.

Beckmann, M., McGuire, C. B. and Winsten, C. B. (1956) *Studies in the economics of transportation*. Yale University Press, New Haven, Connecticut; also published as Rand-RM-1488-PR, Rand Corporation, Santa Monica, CA, May 12, 1955.

Brueckner, J. (2002) "Airport congestion when carriers have market power", *American Economic Review*, 92, 5, pp. 1357-1375.

Cremer, H., Farid, G. and Laffont, J. J. (2003) "Access to Pipelines in Competitive Gas Markets", *Journal of Regulatory Economics*, Vol. 24, No. 1, pp. 5-33.

Dafermos, S. and Nagurney, A. (1985) "Isomorphism between spatial price and traffic equilibrium models". *Lefschetz Center for Dynamical Systems* Report No. 85–17, Division of Applied Mathematics, Brown University, Providence, Rhode Island.

Enke, S. (1951) "Equilibrium among spatially separated markets: Solution by electronic analogue", *Econometrica* 10, pp. 40–47.

Fernández, J. L. and Friesz, T. L. "Equilibrium Predictions in Transportation Markets: The State of the art". *Transportation Research*, Vol. 17B, No. 2, pp. 155-172, 1983.

Friesz, T. L. (1985) "Transportation network equilibrium, design and aggregation: Key developments and research opportunities." *Transportation Research., Part A*, 19A(5–6), pp. 413–427.

Friesz, T. L., Bernstein, D., Smith, T. E., Tobin, R. L. and Wie, B.W. (1993) “A variational inequality formulation of the dynamic network user equilibrium problem”, *Operations Research* 41, pp. 179–191.

Friesz, T. L., Bernstein, D., Mehta, J., Tobin, R. L. and Ganjalizadeh, S. (1994) “Day to day dynamic network disequilibrium and idealized traveler information systems”. *Operations Research* 42, pp. 1120–1136.

Friesz, T. L., Harker, P. T. and Tobin, R. L. (1984) “Alternative algorithms for the general network spatial price equilibrium problem”. *Journal of Regional Science* 24, pp. 475–507.

Friesz, T. L., Tobin, R. L. Smith, T. E. and Harker, P. T. (1983) “Anonlinear complementarity formulation and solution procedure for the general derived demand network equilibrium problem.” *Journal of Regional Science* 23, pp. 337–359.

Florian, M. and Los, M. (1982) “A new look at static spatial price equilibrium models”. *Regional Science and Urban Economics* 12, pp. 579–597.

Friesz, T. L., Tobin, R. L., Smith, T. E. and Harker, P. T. (1983) “A nonlinear complementarity formulation and solution procedure for the general derived demand network equilibrium problem.” *Journal of Regional Science* 23, 3, pp. 337–59.

Harker, P. (1986) “Alternative models of spatial competition.” *Operations Research*, *34*, pp. 410-425.

Hashimoto, H. (1995) “A Spatial Nash Equilibrium Model”, in: P. Harker, ed., *Spatial Price Equilibrium: Advances in Theory, Computation and Application* Springer-Verlag, 1995.

Haurie, A. and Marcotte, P. (1985) “On the relationship between Nash and Wardrop equilibrium” *Networks* 15, pp. 295–308.

Jara-Díaz, S. R. (1986) “On the relation between users’ benefits and the economic effects of transportation activities”, *Journal of Regional Science*, 26(2), pp. 379-391.

Jara-Díaz, S. R. and Friesz, T. L. (1982) “Measuring the benefits derived from a transportation investment,” *Transportation Research B*, 16B, pp. 57-77.

Kanemoto, Y. and Mera, K., (1985) “General Equilibrium Analysis of the Benefits of Large Transportation Improvements,” *Regional Science and Urban Economics 15*, pp. 343-363.

Kidokoro, Y. (2003) “Cost-Benefit Analysis for Transport Networks -Theory and Applications”, *Journal of Transport Economics and Policy*. Journal of Transport Economics and Policy, Volume 38, Number 2, pp. 275-307(33).

Melendez-Hidalgo, J., Rietveld, P. and Verhoef, E. (2003) “Transport Infrastructure, Spatial General Equilibnium and Welfare” working paper VU.

Melendez-Hidalgo, J., Rietveld, P. and Verhoef, E. (2006) “On the Benefits from Transport Infrastructure Investment in a Spatial Oligopoly”, working paper VU.

Mohring, H. (1976) *Transportation Economics*, Ballinger Publishing Co.

Nagurney, A. (1993) *Network economics: a variational inequality approach*. Kluwer Academic Publishers, Boston, Massachusetts.

Neuburger, H. (1971) “User benefit in the evaluation of transport and land use plans,” *Journal of Transport Economics and Policy*, 5, pp. 52-75.

Newbery (1998) *Efficiency Benefits of transport cost reductions*, report to SACTRA.

Pels, E. and Verhoef, E. (2004) “The economics of airport congestion pricing”, *Journal of Urban Economics*, 55, 2, pp. 257-277.

SACTRA (1999) Standing Advisory Committee on Trunk Road Assessement: T*ransport and the Economy*.

Samuelson, P. A. (1952) “Spatial price equilibrium and linear programming” *American Economic Review*, Vol.42, pp. 283-303.

Samuelson, P. A. (1948) “International trade and the equalization of factor prices.” *The Economic Journal*, Vol. 58 (June), pp. 163-184.

Samuelson, P. A. (1962) “The gains from international trade once again.” *Economic Journal.* Vol.72 (December), pp. 820-29.

Schotchmer, S. and Thisse, J. F. (1992) “Space and competition: a puzzle”, *Annals of Regional Science* 26, pp. 269-286.

Sheppard, E. and Curry, L. (1982) “Spatial price equilibria” *Geographical Analysis* 14, pp. 279–304

Takayama, T. and Judge, G. G. (1964) “Equilibrium among spatial separated markets: A reformulation”. *Econometrica*, *32*, pp. 510-524.

Takayama, T. and Judge, G. G. (1971) *Spatial and temporal price and allocation models*. Amsterdam: North-Holland.

Takayama, T. and Woodland, A. (1970) “Equivalence of price and quantity formulations of spatial equilibrium: purified duality in quadratic and concave programming”. *Econometrica* 38, pp. 889-906.

Thore, S. (1991) *Economic logistics: The optimization of spatial and sectoral resource, production, and distribution systems*. New York: Quorum Books.

Tinbergen, J. (1957) "The appraisal of road construction: two calculation schemes" *Review of Economics and Statistics* 39, pp. 241-249.

Tobin, R. L. and Friesz, T. L. (1983) "Formulating and solving the spatial price equilibrium problem with transshipment in terms of arc variables". *Journal of Regional Science* 23, 2, pp. 187–98.

van den Bergh, J C. J. M., Nijkamp, P. and Rietveld, P. ed. (199) *Recent advances in spatial equilibrium modelling: Methodology and applications*. New York: Springer.

van Dender, K. (2005) "Duopoly prices under congested access", *Journal of Regional Science*, 45, 2, pp. 343-362.

Venables, A. and Gasiorek, M. (1998) The Welfare Implications of Transport improvements in the Presence of Market Failure, report to the Standing Advisory Committee on Trunk Road Assessment. UK.

Waquil, P. D. and Cox, T. L. (1995) *Spatial equilibrium with intermediate products: Implementation and validation in MERCOSUR*. Staff Paper Series, Department of Agricultural Economics, University of Wisconsin – Madison, USA.

Yang, H. and Huang, H. J. (1998) "Principle of marginal-cost pricing: how does it work in a general road network?" *Transportation Research*32A, pp. 45–54.

Instructions to Authors

Papers should be written in English and submitted electronically to trasportieuropei@istiee.org and, by regular mail, in printed form in duplicate to:

prof. Romeo Danielis
Dipartimento di Scienze Economiche e Statistiche, Facoltà di Economia
Università degli Studi di Trieste. P.le Europa, 1, 34100 Trieste, Italy
Phone: +39-0405587076 - Fax: +39-040567543

Submission of a manuscript is considered to be a representation that it has been neither copyrighted (or if copyrighted is clearly marked so that the appropriate permission can be obtained) nor published, that it is not being submitted for publication elsewhere, and that, if the work results from a military contract, it has been released for open publication. As a condition of final acceptance of a paper for publication in European Transport\Trasporti Europei, the author(s) must indicate if their paper is posted on a working paper website, other than their own. They are responsible for assuring that, if any part of the paper has been copyrighted for prepublication as a working paper, the copyright can and will be transferred to ISTIEE when the paper has been accepted. This includes both print and electronic forms of the paper. On acceptance, the text, or any link to full text, must be removed from the working paper websites, other than the author's own website. Other material such as book reviews and announcements should also be sent to the Editor.

Manuscripts should contain only endnotes. Figures are required in a form suitable for photographic reproduction. Any one of a number of forms will be acceptable, e.g., laser printer drawing, original black ink drawings, or high-quality glossy prints. Lettering should be uniform in size and style and sufficiently large to be legible after reduction. Figures should be designated by arabic numbers and referred to in the text by number. Figure legends should be collectively provided on a separate sheet rather than placed on the figures themselves. Tables may be typed on sheets separated from the text. Each table should have a caption that makes the table entries clearly independent of the text; complicated column headings should be avoided. All tables should be numbered and referred to in the text by number.

In mathematical expressions, authors are requested in general to minimize unusual or expensive typographical requirements; for example: authors are requested to use the solidus wherever possible in preference to built-up fractions, to write complicated exponentials in the form exp() and to avoid subscripts and superscripts on subscripts or superscripts. Subscripts and superscripts should be shown large and clear, Greek letters and unusual symbols should be labeled on first occurrence, as should subscript "zero", to distinguish it from the letter "oh". Whether each letter is capital or lower case should be unambiguous. Equation numbers must be at the right.

Each paper must be accompanied by an abstract of about 100 words. The abstract should be adequate as an index and should summarize the principal results and conclusions. The first section of the article should not be numbered. References to related previous work should be reasonably complete, and grouped at the end of the paper. References in the text should be cited by the author's surname and the year of publication, e.g.: (Jansson, 1980), (Marguier and Ceder, 1984).The following format should be used for references:

Article in a journal:
Swait, J. (2001) "A non-compensatory choice model incorporating attribute cut-offs" *Transportation Research, Part B: Methodological*, 35, 10, pp. 903-28.

Chapter in a book:
Regan, A. and Garrido, R. (2001) "Modelling Freight Demand and Shipper Behaviour: State of the Art, Future Directions", In: Hensher, D. (eds) *Travel Behaviour Research: The Leading Edge*, Pergamon, Amsterdam.

Working paper:
Gavish, B. and Graves, S.C. (1981) "Scheduling and Routing in Transportation and Distribution Systems: Formulations and New Relaxations", *Working Paper 8202*, Graduate School of Management, University of Rochester, Rocherster, NY.

Book:
Urban, D. (1993) *Logit – Analyse. Statistische Verfahren zur Analyse von Modellen mit qualitativen Response-Variablen*, Gustav Fischer, Stuttgart.

Dissertation:
Jaillet, P. (1985) Probabilistic Traveling Salesman Problems, *Ph.D. thesis*, Massachusetts Institute of Technology, Cambridge, MA

Presentation at a conference:
Maggi, R. and Bolis, S. (1999) "Adaptive Stated Preference Analysis of Shippers' Transport and Logistics Choice", *World Transport Research* -Proceedings from the 8th World Conference on Transport Research, (H. Meersman, E. Van de Voorde, W. Winkelmans eds.), Pergamon, Amsterdam.

Authors are responsible for revising their proofs, and should limit alterations to the strict minimum. The editorial management of ISTIEE reserves the right to accept only those changes that affect the accuracy of the text.