ECONOMIC ANALYSES OF AIRPORT PRICING AND PRIVATIZATION

•- ·

by

Leonardo Basso

B.E., Universidad de Chile, 1999 M.Sc., Universidad de Chile, 2001

THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

in

The Faculty of Graduate Studies

(Business Administration)

UNIVERSITY OF BRITISH COLUMBIA

.

ň

October 2006

© Leonardo Basso 2006

ABSTRACT

Airport congestion has become a major public policy issue because delays impose important costs on airlines, passengers and shippers (losses have been estimated to be above US\$20 billion annually). The answer economists have given to the airport runway congestion problem has been the use of the price mechanism, under which landing fees are based on a flight's contribution to congestion. However, these schemes have not really been implemented.

Airports have traditionally been owned by governments, but this has been changing: following the UK, many airports around the world have recently been, or are in the process of, being privatized. One of the leading arguments for airport privatization is that privatized airports might well shift towards peak-load or congestion pricing schemes of their runway services, thus reducing delays. Nevertheless, out of the concern that the privatized airports would exert monopoly power, most of the newly privatized airports have been subject to some form of economic regulation. Lately, however, some authors have argued that the regulation mechanisms fell short of being optimal because they would misplace capacity investment incentives. They suggested divestment of regulation or the less-stringent price monitoring.

However, as important as these issues may appear, there have been only a couple of papers that have analytically examined what the outcomes of privatization or divestment of regulation may be. And, although many papers analyze optimal pricing of public airports, most of the papers that deal with privatization and deregulation are fairly descriptive. In this thesis, the effects that competition, privatization and regulation (or absence of it) would have on the performance, pricing structure and capacity investments of airports, and the consequences this will bring to the downstream market (airlines), and final users (passengers) will be examined analytically. What makes this different from previous work on airport privatization is that, here, *vertical structure* models of airport-airlines behavior will be used. That is, it will be recognized that airports provide an essential input, which is used by airlines to produce the output –travel– under oligopolistic conditions. Previous work abstracted from the airline market under the assumption that the airline market was perfectly competitive.

Non-cooperative games will be analyzed, in order to extract lessons and conclusions for public policy, on a number of different issues: (i) whether previous conclusions regarding airport privatization (and airport pricing in general) hold under imperfect airline competition; (ii) whether privatized unregulated airports would use efficient peak-load pricing congestion schemes or not, and under which conditions they would; (iii) whether the arguments that have been put forward in favor of deregulation of private airports hold or not; (iv) whether competition between private airports in multi-airport regions would lead to a more efficient outcome than single ownership.

TABLE OF CONTENTS

Ał	ostrac	et	ii		
Ta	ble o	f Contents	iii		
Li	st of '	Tables	vi		
Li	st of]	Figures	vii		
Ac	knov	vledgements	viii		
De	dicat	ion	ix		
Co)-Aut	horship Statement	X		
1	Intr	Introduction			
	1.1	Research topic and content of the thesis	1		
	1.2	References	4		
2	2 A Survey of Analytical Models of Airport Pricing		5		
	2.1	Introduction	5		
	2.2	The traditional approach to airport pricing	6		
	2.3	The vertical structure approach to airport pricing	16		
	2.4	Relationship between approaches	21		
	2.5	Pricing of airport networks	26		
	2.6	Conclusions and further research	30		
	2.7	References	36		
3	On	Input Market Surplus and its Relation to the Downstream Market Game	39		
	3.1	Introduction	39		
	3.2	The downstream game and the literature	40		
	3.3	Input market surplus and the downstream game	42		
	3.4	Potential function and input market surplus	45		
	3.5	Conclusions	48		
	3.6	References	50		
4	Airj	port Ownership: effects on pricing and capacity	51		
	4.1	Introduction	51		

	4.2	1.2 The airline market		
		4.2.1	The airline oligopoly model	53
		4.2.2	Equilibrium in the airline market: derived demand for airports and its characteristics	57
	4.3	The ai	rports market	61
		4.3.1	General features	61
		4.3.2	System of private airports	62
		4.3.3	System of public airports	63
		4.3.4	Maximization of joint profits: airlines and airports	68
		4.3.5	On the budget adequacy of public airports	72
		4.3.6	Independent private airports	73
	4.4	Nume	rical simulations	75
	4.5	Final o	comments	79
	4.6	Refere	ences	82
5	Sequential Peak-Load Pricing in a Vertical Setting: the case of airports and airlines			85
	5.1	Introd	uction	85
	5.2	.2 The model		88
	5.3	.3 Analysis of output-market equilibrium		
		Analy	sis of output-market equilibrium	92
	5.4	Analy Airpo	sis of output-market equilibrium rt pricing, traffic, delay and welfare comparisons	92 97
	5.4 5.5	Analy Airpo Airpo	sis of output-market equilibrium rt pricing, traffic, delay and welfare comparisons rt-airline joint profit maximization	92 97 103
	5.4 5.5 5.6	Analy Airpo Airpo Concl	sis of output-market equilibrium rt pricing, traffic, delay and welfare comparisons rt-airline joint profit maximization uding remarks	92 97 103 105
	5.4 5.5 5.6 5.7	Analy Airpo Airpo Concl Refere	sis of output-market equilibrium rt pricing, traffic, delay and welfare comparisons rt-airline joint profit maximization uding remarks	92 97 103 105 106
6	5.4 5.5 5.6 5.7 Con	Analy Airpo Airpo Concl Refere	sis of output-market equilibrium rt pricing, traffic, delay and welfare comparisons rt-airline joint profit maximization uding remarks ences e Facility Rivalry in Vertical Structures	 92 97 103 105 106 110
6	 5.4 5.5 5.6 5.7 Con 6.1 	Analy Airpo Airpo Concl Refere gestibl Introd	sis of output-market equilibrium rt pricing, traffic, delay and welfare comparisons rt-airline joint profit maximization uding remarks ences e Facility Rivalry in Vertical Structures uction	 92 97 103 105 106 110 110
6	 5.4 5.5 5.6 5.7 Con 6.1 6.2 	Analy Airpo Airpo Concl Refere gestibl Introd	sis of output-market equilibrium rt pricing, traffic, delay and welfare comparisons rt-airline joint profit maximization uding remarks ences	 92 97 103 105 106 110 110 113
6	 5.4 5.5 5.6 5.7 Con 6.1 6.2 	Analy Airpo Airpo Concl Refere gestibl Introd The m 6.2.1	sis of output-market equilibrium rt pricing, traffic, delay and welfare comparisons rt-airline joint profit maximization uding remarks ences e Facility Rivalry in Vertical Structures uction	 92 97 103 105 106 110 110 113 114
6	 5.4 5.5 5.6 5.7 Con 6.1 6.2 	Analy Airpo Airpo Concl Refere gestibl Introd The m 6.2.1 6.2.1	sis of output-market equilibrium	 92 97 103 105 106 110 110 113 114 117

			6.3.1 Closed-loop duopoly	121
			6.3.2 Open-loop duopoly	124
			6.3.3 Comparisons: closed-loop vs. open loop duopoly	125
		6.4	Monopoly and the social optimum	126
			6.4.1 Monopoly	126
			6.4.2 The social optimum	129
		6.5	Vertical integration	133
		6.6	Concluding remarks	136
		6.7	References	138
	7	Con	nclusions	141
		7.1	Summary of the thesis	141
		7.2	Future work	143
	Aj	pendi	ices1	145
·		App	pendix A	145
		App	pendix B	149
		App	pendix C	159

LIST OF TABLES

Table 2.1	Summary of papers using the traditional approach	34
Table 2.2	Summary of papers using the vertical structure approach	35
Table 4.1	Parameter values for the numerical simulation	75
Table 4.2	Results of numerical simulation	77
Table c.1	Parameter values for the numerical simulation	164

LIST OF FIGURES

.

Figure 2.1	Network Structure and airline competition (Brueckner, 2005)	28
Figure 2.2	Possible route structures	29
Figure 5.1	Schematic representation of the results in Proposition 5.5	102
Figure 6.1	Consumer distribution and facilities' catchment areas	115
Figure b.1	Free-entry long-run equilibrium	151

ACKNOWLEDGMENTS

I am profoundly indebted to my Professors Anming Zhang and Tom Ross for innumerable comments, notes and suggestions on the drafts of my papers, which materially improved this thesis. But I would specially like to thank them for their friendship, encouragement and advise throughout the process.

I would also like to thank Sergio Jara-Diaz for sharing his lucid vision of science and academia with me. It has definitely helped shaping my love for this profession.

I owe immeasurable gratitude to my parents, Myriam and Patricio, and to Vero Hirsch. This would have been impossible without their love, support, and encouragement throughout so many years.

Many others have also helped me with important comments and suggestions about my work: I am grateful to Ralph Winter, David Gillen, Tae Oum, Bill Waters, Yossi Berechman, Hugh Neary, Markus von Wartburg, Heng Ju, Andrew Yuen, and all the faculty at the OPLOG and SBE divisions.

To my fellow students Anita Parkinson, Jeff Colpitts, Neil Walsh, Alex Boonekamp, David Patient, Xiaowen Fu, Benny Mantin, Fredrik Odegaard, Jonathan Patrick, Ali Ali, Kim Trottier, Kelly Folley, Chris Bidner, Ben Sand and Ken Jackson: thank you for your friendship and support. Special gratitude goes to my good friend Miranda Selinger.

Finally, I want to thank Elaine Cho from the Sauder School of Business PhD office, for being so helpful on every single problem or trouble I had. You really made things easier!

Financial support for my studies was possible thanks to the Canadian National Railway Fellowship in Transportation Studies and the Izaak Walton Killam Memorial Pre-Doctoral Fellowship. I am grateful to these institutions and UBC for their support. Further financial support from the Social Science and Humanities Research Council of Canada (SSHRC) through my advisor's grants and the Phelps Centre for the Study of Government and Business is gratefully acknowledged.

DEDICATION

To my parents, Myriam and Patricio

"La Universidad de Chile ha sido, y seguirá siendo, la conciencia de Chile. Quienes creen que podrán someterla por la fuerza no entienden, no podrán entender nunca que para un universitario la libertad es como la vida misma y que, por ello, no renunciaremos nunca a ser libres".

Patricio Basso, Santiago, Chile, 1986 (desde la cárcel, durante la dictadura)

"The University of Chile has been, and will continue to be, the consciousness of Chile. Those who think it can be subjugated by force do not understand, will never understand that, for an academic, freedom is like life itself and, therefore, we will never renounce to be free".

Patricio Basso, Santiago, Chile, 1986 (from jail, during the dictatorship)

CO-AUTHORSHIP STATEMENT

The first drafts of the papers contained in Chapters 5 and 6 of this thesis were written by me, but have subsequently benefited from major revisions by my Advisor and co-author, Professor Anming Zhang.

1 INTRODUCTION

1.1 Research Topic and Content of the Thesis

Airport pricing has attracted the attention of economists and policy-makers for at least thirty years. Economists have been interested in finding efficient pricing practices, which would need to take into account the increasing levels of congestion at many airports. Early work includes Levine (1969) –who was critical on weight-based runway charges and the fact that prices were 'too low' to reflect congestion levels–, and Carlin and Park (1970) – who advocated social marginal cost pricing, or congestion tolls, as a way to fight airport congestion in the short run. But, given the wave of privatizations of public enterprises throughout the world –which started in the UK in the 1980s–, economists were also urged to discuss about what would be the efficiency gains that privatization of airports would bring, about whether regulation would be necessary or not and, if it was, about how it should be implemented. All these issues, new and old, gave rise to a literature on airport pricing, ownership and regulation which has been, if not massive, considerable. The present thesis attempts to make a contribution to the economic analyses of airport pricing and privatization, through the presentation of five papers, which are assembled in Chapters 2 to 6.

This thesis has been written following The University of British Columbia's *manuscript* style. In this case, the thesis has an *Introduction* and a *Conclusions* chapters, while the rest of the chapters are published, in-press, accepted, submitted or draft manuscripts written in a consistent format. This has the advantage that each chapter is a stand-alone unit which can be read it on its own, but it is obtained at the expense of some *linearity* of the thesis. In this sense, this introduction attempts to provide a more unified vision of the papers in the thesis, by presenting them as sequentially and linked as possible. However, some circularity, that is, cross references between papers, was necessarily introduced.

Chapter 2, entitled "A Survey of Analytical Models of Airport Pricing" is a literature review, and is the reason why I do not present one here in the Introduction. In this paper, analytical airport pricing papers are reviewed and their main insights are highlighted. We give special consideration to how particular features of the model may drive the results, and we look for similarities and differences across papers.¹ Among other things, we show that the literature may be grouped into two broad *approaches*, which we have called the *traditional approach* and the *vertical structure approach*. The traditional approach uses a classical one-market partial equilibrium model where the demand for airports depends on airport charges and on congestion costs of both passengers and airlines; the airline market at the airport is not formally modeled. The vertical structure approach instead recognizes that airports provide an input for the airline market –which is modeled as a rather simple oligopoly– and that it is the equilibrium of this downstream market which determines the airports' demand for airports is therefore a derived demand from final

¹ Given the time at which the literature review was written and the state of advance of the other papers, Chapter 2 does include Chapter 3 –and partially Chapter 4– as part of the papers reviewed, but not Chapters 5 and 6.

passengers' consumption of airline services. We show with an example that airports' demand does not carry enough information to derive welfare-maximization results unless the airline market is perfectly competitive. Hence, vertical structure models, which formally consider the airline market, are more appropriate when there is market power at the airline level.

That last point has in fact to do with the more general question of how "consumer surplus" measures coming from areas under input demand curves -such as the airport's demandcapture the effects of direct purchasers -the airlines- and downstream final consumers -the passengers. In other words, on how to obtain a proper measure of social welfare when analyzing an input market. The most general results available in the literature regarding the relation between input and output markets surplus, hinge on a number of strong simplifying assumptions, which in fact make them inapplicable for the airport case. Chapter 3, entitled "On Input Market Surplus and Its Relation to the Downstream Market Game" relax those assumptions, providing a systematic way to assess how much information about the downstream market is captured in the derived demand for inputs. And, while this is not directly a paper about airport pricing, its application to airport markets is important given the derived nature of the airport's demand. The main proposition of this paper is obtained by linking the input markets surplus question to results from another -seemingly unrelatedstream of literature, which characterizes a function that firms in oligopoly collectively, yet unintentionally, maximize. I show that the input markets surplus change measure (obtained by integration under the input demands derived from the equilibrium of a downstream oligopoly game) is equal to the change in a function for which critical points coincide with the equilibria of the downstream game. In particular, if the downstream game is *potential*, the input market surplus is shown to be equal to the change in the exact potential function. This proposition synthesizes and generalizes the established results on the relation between input and final market surplus measures, providing guidance to policy analysts who seek to infer the total welfare effects of input market price changes from information on the input market demands only.

Chapter 4 is entitled "Airport Ownership: effects on pricing and capacity". It has been argued in the literature that privatized airports would charge more efficient congestion prices and would be more responsive to market incentives for capacity expansions. Furthermore, the privatized airports would not need to be regulated since price elasticities are low, so allocative inefficiencies would be small, and collaboration between airlines and airports, or airlines countervailing power, would solve the problem of airports' market power. Given the results of Chapters 2 and 3, this paper uses a model of vertical relations between airports and airlines in order to adequately set-up the central planner benchmark case (the *public* airport). With this model I examine, both analytically and numerically, how ownership affects airports' prices and capacities. The results show a rather unattractive picture for privatization when compared to first- and second-best benchmark cases. I find that: (i) private airports would be too small in terms of both traffic and capacity and, despite the fact that they may be less congested, they would induce important deadweight losses; (ii) the arguments that airlines' countervailing power or increased cooperation between airlines and airports may make regulation unnecessary seem to be overstated; and (iii) things may deteriorate further if privatization is done on an airport by airport basis rather than in a system. Also, I show that two features of air travel demand that have not been incorporated previously in the literature –demand differentiation and schedule delay cost–play important roles on airports' preferences regarding the number of airlines using the airport.

On the closing sections of Chapters 2 and 4, I state that two lines of work that should be explored are: (i) periodic demand, which may induce peak-load pricing practices sequentially, at both the airport and the airlines levels; and (ii) the case of geographic competition between airports, that is, what happens in multiple airports regions (such as New York or San Francisco), because in this case results may be hindered. Chapter 5, "Sequential Peak-Load Pricing in a Vertical Setting: the case of airports and airlines", examines the first line of work. We investigate airport peak-load pricing (PLP) and analyze both the price level and price structure with vertically differentiated peak and off-peak travel. Using a vertical structure of airport and airlines in which both players may use peakload pricing, we carry out an analysis for a private, unregulated airport and for a public airport that maximizes social welfare. We find that compared to the public airport, a private, profit-maximizing airport would charge both, higher peak and off-peak runway prices, as well as a higher peak/off-peak price differential. As a consequence, airport privatization would lead to both fewer total air passengers and fewer passengers using the premium peak hours for their travel, both of which reduce social welfare. Although those passengers who still use the peak period benefit from less congestion delays, overall it is not economically efficient to have such a lower level of peak congestion. The analysis also shows that whilst private airports will always use peak-load pricing a public airport may, somewhat surprisingly, actually charge a peak price that is lower than the off-peak price. Here the public airport, on the surface, is not practicing the peak-load pricing, but such pricing structure is nevertheless socially optimal. Again, the case where a private airport strategically collaborates with the airlines is examined.

Chapter 6, "Congestible Facility Rivalry in Vertical Structures" examines the second line of work identified in Chapters 2 and 4. We investigate rivalry between congestible facilities and its effects on facility charges, capacities and congestion delays. The analysis is conducted under a vertical facility-carrier-consumer structure, with imperfectly competitive output (carriers) markets. We find that the duopolists' equilibrium prices increase with both the consumers' value of time and the carriers' cost sensitivity to congestion delays; entrance of a new carrier to any of the facilities depresses the prices charged by both facilities; and lower marginal cost of the carriers at one facility will induce a higher facility price at that facility but a lower facility price at the other facility. In terms of service level, we find that the duopoly facilities provide longer congestion delays than a monopolist only if capacity decisions are made prior to the facility pricing decisions. When the capacity and pricing decisions are made simultaneously, the duopolists would provide the same level of service quality (delays) as the monopolist. Furthermore, monopoly pricing and capacity choices result in a higher level of service quality (shorter delays) than the social optimum. Our analysis shows that when the monopolist vertically integrates with the carriers at the facilities, it would provide the same congestion level as the central planner. Nevertheless, this monopoly service level is not socially optimal in a second-best sense. In effect, in the fully ex-ante symmetric case, it is too low with respect to the second best.

1.2 References

- Carlin, A. and Park, R.E. (1970) Marginal Cost pricing of Airport Runway Capacity. American Economic Review, 60, 310-319
- Levine, M.E. (1969) Landing fees and the airport congestion problem. Journal of Law and Economics, 12, 79-108.

2 A SURVEY OF ANALYTICAL MODELS OF AIRPORT PRICING¹

2.1 Introduction

Airport pricing has attracted the attention of economists for some time now, starting with Levine (1969). Most of the attention has been devoted to the efficiency of pricing practices by airport authorities and the need to take into account congestion which, even in the early 70s, was afflicting passengers and airlines. The alleged inefficiencies of actual pricing practices plus the factual wave of privatizations and/or partnerships that started in the late eighties throughout the world (following the example of the UK) induced, in addition, a focus on the effects of privatized airports will pursue maximization of profits, but it has usually been accepted that airports enjoy a local monopoly position because they have a captive market. Besides, there would be sizeable economies of scale on airport infrastructure provision and airport operations (Doganis, 1992). Out of the concern that private airports would exert market power in user charges, many private airports are under some type of economic regulation such as rate-of-return or price caps.

The work on airport pricing has been, if not massive, considerable. Some old questions, such as, how should we use the price mechanism to signal congestion problems, have persisted in the literature. New questions, such as whether privatization would induce better capacity investment or not, have appeared. And as far as we know, there has been no paper devoted to put together all the questions and answers that have been obtained in the literature. We attempt to do that here; in this paper, we review the airport pricing literature, with a focus on analytical papers. Indeed, we are narrowing the scope of our work, by leaving aside a number of important empirical papers. We do not mean by this that the empirical work is irrelevant, but as it will be seen, a comprehensive survey of the analytics of airport pricing easily use the space in a paper, and we believe that a good command of theoretical results helps to better grasp empirical results. Also, we will focus on papers on the last 20 years. We believe that this is enough to understand what is known today about the theory of airport pricing, since earlier contributions such as Levine (1969), Carlin and Park (1970) and Morrison (1983) have been incorporated into the papers we review.

We summarize the findings and provide directions of what we think should be future research. In order to do this in an orderly manner, we group the papers in 'approaches'. Papers within one approach share many features regarding the analytical modeling, which makes it easier to explain what characterize them, while also enabling a better description of the contributions of each of individual article. Therefore, Sections 2.2 and 2.3 will be devoted to explain what we have called '*the traditional approach*' and '*the vertical structure approach*' to airport pricing respectively and, within each approach, what we have learned from individual articles. Later, in Section 2.4, we attempt to connect the approaches as a mean to better understand how the results stemming from the two

¹ A version of this chapter is forthcoming as: Basso, L.J. and Zhang, A. (2006a) A Survey of Analytical Models of Airport Pricing, in *Advances in Airline Economics, Volume 2: The Economics of Airline Institutions, Operations and Marketing*, Darin Lee, ed., Elsevier.

approaches relate to each other. In short, whether results from one approach can be transferred to the other. Sections 2.2 and 2.3 deal with a single airport's decision or, at most, two (complementary) airports, given the complexity of the economics of airport pricing. However, there are a few articles that looked at the pricing of airport networks, i.e., three or more connected airports, something that indeed bears more relationship with reality. We discuss these papers in Section 2.5, noting that the previous classification may still be applied. We conclude in Section 2.6 by providing what we think should be the lines of future research.

2.2 The Traditional Approach to Airport Pricing

The main characteristic of this approach is that it typically considers a single airport decisions and follows a partial equilibrium analysis in which the airport's demand is directly a function of its own decisions; airlines decisions (and competition) are not directly considered and hence, the derived characteristic of an airport's demand is not formally recognized. In this Section, we consider papers by Morrison (1987), Morrison and Winston (1989), Oum and Zhang (1990), Zhang and Zhang (1997; 2001; 2003), Carlsson (2003), Oum et al. (2004), Lu and Pagliari (2004) and Czerny (2006). Most of these papers follow essentially the same model: they assume that the demand for the airport is a function of a full price. This full price includes the airport charge and, in an additive fashion, some cost measure of the delay caused by congestion. Delay functions have always been measured through some non-linear function of traffic and capacity, although the modeling has not been unique: the main discrepancy has been whether the function should or should not be homogenous of degree one in the traffic to capacity ratio. Delay is assumed to affect both airlines and passengers, and consumers' surplus is measured by integration of the airport's demand. When the airport capacity is variable, the cost function has been usually assumed to be separable in operating and capacity costs.

This approach has been used to analyze many different issues regarding airport pricing and capacity decisions and under many different sets of assumptions, as can be seen in Table 2.1. Initially the focus was on deriving optimal prices and capacities on the presence of congestion but, lately, it has been used to assess the effects of privatization and regulation as well.

The basics of the traditional approach can be synthesized in a fairly concise analytical manner, which we present below. Certainly, not all the papers can directly be assimilated to this presentation –particularly Lu and Pagliari (2004) and Czerny (2006) may seem more distant–, but most of them fit through slight adjustments, which are indicated where relevant. The approach is as follows: in order to provide aviation services, an airport incurs both operating and capital expenses. It collects user charges to cover these costs and, in the private case, to make a return on capital investments. For a given capacity, congestion will start to build up as demand grows, inducing delays and therefore extra costs on passengers and airlines. It is usually assumed that airlines fully pass whatever airport charge they face

to passengers; the same is assumed for airlines delay costs.² Therefore, passengers will perceive a full price consisting of the airport charge, the airline delay cost, passengers delay cost, travel-time costs plus other charges (air ticket). It has been argued that since other airline charges are exogenous as far as the airport is concerned, the demand an airport faces may be considered to be a function only of the airport charge P and the flight delay costs D (which includes both airlines and passengers). The variables in the model would be:

$Q_t(\rho_t)$	= Demand for airport facilities in period t , measured in number of
	flights, which is a function of a full price ρ_t
$ ho_t$	= $P_t + D_t$, full price in period t, which determines airport's demand
	(and which is a part of the full price perceived by passengers)
P_t	= airport charge per flight in period t
$D_t = D(Q_t, K)$	= flight delay costs (airlines plus passengers) experienced by each
	flight in period t, which depends on traffic Q_t and airport capacity K
K	= capacity of the airport
$C(Q_t)$	= operating costs of the airport in period t
r	= cost of capital

In most cases, when many periods are considered, demands are assumed to be independent (the exception is Oum and Zhang, 1990). When capacity is assumed to be continuously adjustable, this is has been justified because capacity would be defined not only by the number of runways –which can only be increased discretely– but also by air navigation systems and other infrastructure –which can be increased or enhanced continuously.

One of the first issues that was analyzed using the *traditional approach* was airport's choices of capacity, K, and user charges, P_i , for the benchmark case in which social welfare is maximized subject to a budget constraint –the *public airport* case. Thus, this case corresponds to Ramsey-Boiteux pricing. The problem the public airport faces is given by:

$$\max_{P_t,K} \sum_{t} \left(\int_{\rho}^{\infty} Q_t(\rho_t) d\rho_t + P_t Q_t - C(Q_t) \right) - rK$$
(2.1)

s.t.
$$\sum_{t} P_{t}Q_{t} - C(Q_{t}) - rK = 0$$
 (2.2)

The first term in the objective function would correspond to consumer surplus; the remaining terms are airport's profits. Forming the Lagrangean and taking derivatives with respect to P_t and K, first-order conditions are obtained. From them, the following pricing and capacity investment rules follow

² Morrison (1987) make this assumption by equating airlines' elasticity of demand for airport services to the elasticity of passengers' demand with respect to full price times the fraction that airport charges and congestion costs represent in total flight costs (p.48, see also Raffarin p. 115). Oum et al. (2004) make this assumption explicitly, arguing that this will be the case under perfect competition. However, this is going to be so only in the case of constant marginal costs for the airlines. As a tax, the proportion of the charge that is actually passed to consumers will depend on the relative slopes of demand and supply curves in the case of perfect competition.

$$P_{t} = C' + Q_{t} \frac{\partial D}{\partial Q_{t}} + \frac{\lambda}{1 + \lambda} \left(\frac{\rho_{t}}{\varepsilon_{t}} \right)$$
(2.3)

$$-\sum_{t} Q_{t} \frac{\partial D}{\partial K} = r$$
(2.4)

where λ denotes the Lagrange multiplier of the budget constraint. According to Morrison (1987) and Zhang and Zhang (1997, 2001), the interpretation of the pricing rule is as follows: the first two terms on the right hand side of equation (2.3) represent the social marginal cost of one flight (operational marginal cost plus marginal cost of congestion) while the third term represents a markup that is inversely related to the (positive) elasticity of demand with respect to the full price, ε_t , and that depends on the severity of the budget constraint. Hence, the difference with the usual Ramsey-Boiteux pricing is that the pricing rule has to take into account the congestion that a new flight imposes on others.

Regarding the optimal capacity rule –Equation (2.4)–, Zhang and Zhang (1997) note that it does not depend on λ and hence it is identical to the one obtained when a budget constraint is not imposed, as in Morrison and Winston (1989). Therefore, airport authorities which adopt Ramsey pricing should still pursue the same optimal policy of capacity investment. In this policy, the socially optimal level of capacity is set such that *the marginal benefit of capacity* in terms of reduction of delays, equates the marginal cost of capacity (Morrison and Winston, 1989; Zhang and Zhang, 1997).

Now that we have set up the basics of the approach, we can look into the modifications that have been made to it and the insights that have been gained in each case.

On weight-based airport charges

Because in general aircrafts are not charged by the contribution they make to congestion but by their weights, Morrison (1987) wanted to uncover the importance regulators give to each type of aircraft when choosing the aeronautical charges. For this, he assumed that the demand in each period is Q_{ii} , where *i* denotes a class of airport user, that is, a type of aircraft. Then, assuming that capacity is fixed, he put weights on the contribution of each class of users at each period, to the social welfare function (2.1):

$$\sum_{i,t} \eta_{it} \left(\int_{\rho}^{\infty} Q_{it}(\rho_{it}) d\rho_{it} + P_{it} Q_{it} - C(Q_{it}) \right)$$
(2.5)

where η_{ii} is the weight of user *i* at period *t*. With this social welfare function, the optimal pricing rule, (2.3), turns into:

$$P_{it} = C' + Q_{it} \frac{\partial D}{\partial Q_{it}} + \frac{\lambda + 1 - \eta_{it}}{1 + \lambda} \left(\frac{\rho_{it}}{\varepsilon_{it}} \right)$$
(2.6)

Morrison then asks the following question: what set of weights is implied by actual airport charges? To uncover the weights, he follows Ross (1984) and solves for the weights η_{it} in (2.6). Hence, using actual data, those weights can be obtained up to a multiplicative constant. Morrison's main result was that when the airport is non-congested, weight-based landing fees imply welfare weights (the η_{it}) that are very similar. But when congestion increases, the dispersion in the weights also increases, implying that the weight-based landing fees would be less appropriate when there is congestion. He argues that this happens because, while weight is a reasonable proxy for elasticity of demand, it is a poor proxy for congestion costs.

Lumpy capacity and cost recovery

Oum and Zhang (1990) and Zhang and Zhang (2001) were interested on how would budget adequacy be affected by the fact that, in reality, capacity can be increased only discretely. The conjecture was that the lumpy nature of capacity expansions would make social marginal congestion pricing lead to alternating periods of airport surplus and deficit. Oum and Zhang (1990) considered lumpy capacity expansions, i.e. that K can be increased only by a minimum amount ΔK , incorporating a positive time trend to the airport's demand to capture the fact that demand would increase on time together with the economy. They focused on the timing of capacity expansions rather than in a steady-state as above, but they did not consider a budget constraint though. They concluded that, when capacity is indivisible, optimal congestion pricing (as in equation 2.3 with λ =0) and optimal capacity expansion would lead to alternating periods of excess capacity and capacity shortage. During capacity shortage, the congestion toll would exceed annualized capacity costs but during excess capacity, the congestion toll would fall short of annualized capacity costs. This implies that budget adequacy would depend entirely on the number of shortage/excess capacity periods between capacity expansions. And the number of periods in each case depends on the pattern of traffic growth.

Oum and Zhang (1990) concluded that, when capacity is indivisible, the cost recovery status of an airport cannot be predicted without reference to the time path of the traffic growth and, therefore the *cost recovery theorem* for investment in transportation infrastructure would note hold. This theorem states that (see e.g. Mohring, 1976) when operational costs are separable from capacity costs and exhibit constant returns to scale, and the delay function is homogenous of degree one in the traffic to capacity ratio, congestion pricing lead to exact cost recovery theorem would fail for airports though. Even if capacity is divisible, as in model (2.1)-(2.3), Zhang and Zhang (1997) showed that, without a budget constraint, social-marginal-cost pricing would always give rise to a financial deficit to the airport because the delay function D would not be homogenous of degree one in the traffic to capacity financial deficit to the airport because the delay function D would not be homogenous of degree one in the traffic to capacity rise to a financial deficit to the airport because the delay function D would not be homogenous of degree one in the traffic to capacity ratio (Lave and De Salvo, 1968; US Federal Aviation Administration, 1969; Horonjeff and McKelvey, 1983). Furthermore, the deficit would increase the more congested the airport is.

Given all this, Zhang and Zhang (2001) were interested in the case in which delays are nonhomogenous of degree one, capacity is indivisible, traffic grows in time, but airports are required to recover their costs, both from operations and capacity investments. The question they asked was: should public airports be asked to break even in the short-run, or in the long run, that is, taking losses in early years but surplus in later years? For this, they modified the program (2.1)-(2.2) so as to consider that the airport would now maximize social welfare over a period of time S, while achieving cost recovery over the entire period. Capacity was assumed to be fixed during the period, owing to its indivisibility. The new long-run problem faced by the public airport is:

$$\max_{P_t,K} \int_{0}^{s} \left\{ \sum_{t} \left(\int_{\rho}^{\infty} Q_t(\rho_t, s) d\rho_t + P_t Q_t - C(Q_t) \right) - rK \right\} e^{-rs} ds$$

$$s.t. \quad \int_{0}^{s} \left\{ \sum_{t} P_t Q_t(\rho_t, s) - C(Q_t) - rK \right\} e^{-rs} ds = 0$$

$$(2.7)$$

Now, the airport's demand increases in time, that is $\partial Q_t / \partial s > 0$ and future revenues are discounted using the cost of capital, r. The short-run problem is as in (2.1). Not surprisingly, Zhang and Zhang found that the short-run financial break-even constraint lead to smaller social welfare levels than a long-run break-even constraint. This was expectable, since short-run budget adequacy implies long-run budget adequacy. The slackness won with the latter though, explains the increase in social welfare. In fact, Zhang and Zhang showed that the two will be equal only when the airport's demand does not change on time, that is $\partial Q_t / \partial s = 0$. This directly speaks of the importance of the time path of the traffic growth, as pointed out by Oum and Zhang (1990): to maximize social welfare, airports should be allowed to take losses or make profits at different times, seeking cost recovery only in the long run.

What is perhaps more interesting is that they found that, under short-term cost recovery, airport charges are high when the demand is low and when there is excess capacity. However, when the demand is high, and there is congestion, airport charges would be low. This seems to be undesirable. On the other hand, under long term cost recovery, airport charges grow together with the demand.

Airport concessions and their effects on public and private airports

Given the increasing pressure on public airports to self finance, airports have been increasingly depending on revenues generated by non-air related business, such as parking, in-airport stores and so on. The demand for these *concession services* is complementary to the demand for *aeronautical services* in that, the more people there are using the airport, the higher the concession revenues will be. Zhang and Zhang (1997) wanted to analyze what would be the socially optimal balance between aeronautical revenues and concession revenues given the cost recovery constraint, and how would the pricing practices involved look like. For this, they modified the problem of the public airport in (2.1)-(2.2), to incorporate the fact that concession demand is complementary to aeronautical demand:

$$\max_{P,K,p} \int_{\rho}^{\infty} Q(\rho)d\rho + PQ - C(Q) - rK + Q\left(\int_{p}^{\infty} X(p)dp + pX - c(X)\right)$$
s.t.
$$PQ - C(Q) - rK + Q(pX - c(X)) = 0$$
(2.8)

In (2.8), p represents the price for concession goods or non-aeronautical services provided in the airport, X(p) is the demand for concession services per flight, and c(X) are the costs of providing the concession services, which are assumed to feature constant returns to scale. For notational simplicity, we present hereafter the models without intraday variation in demand, i.e., there is only one t.

There are two important things two note in Zhang and Zhang (1997) setup. First, that the complementarity between the demands is unidirectional, that is, consumers decide to fly or not based on the full price of the aeronautical service; they do not take into account the price of the concessions in their travel decisions. Only after arrival at the airport, passengers observe concession prices and make purchasing decisions. Second, note that the budget constraint includes both revenues, from aeronautical and concession services, which effectively enables cross-subsidies.

Without the budget constraint, the (first best) optimal solution obviously involves marginal cost pricing on the concessions side, i.e. p = c'(X). On the aeronautical side, the social-marginal-cost pricing of equation (2.3) would have an additional markdown; this happens because now, a smaller aeronautical charge increases the demand for both, aeronautical services and concessions services. Hence, the optimal charge is smaller. This, however, would lead to deficits if the delay function is non-homogenous of degree one in the traffic to capacity ratio, as discussed above. With the budget constraint, and assuming that the delay function is non-homogenous of degree one, Zhang and Zhang showed that, at the (second best) optimal solution of problem (2.8), concession operations would subsidize aeronautical operations, i.e. p > c'(X) showing that profits would be made in concession services. If the airport were not allowed to make profits from its concessions, but was still asked to self finance, then this would obviously lead to smaller levels of social welfare. Further, they showed that the cross subsidy from concessions does not in general restore social-marginal-cost pricing on the aeronautical side (equation 2.2), unless the demands and costs fulfill a very particular condition.

The attention to concession revenues, however, does not stop at the cost-recovery issue of public airports. It has also been suggested that the complementary nature of the concessions demand would give incentives to private airports to diminish the price they charge for aeronautical services in order to maximize the number of travelers *in the airport* using the concessions. This may imply that price regulations may be unnecessary (see e.g. Condie, 2000; Starkie, 2001). In order to assess whether the argument holds, Zhang and Zhang (2003) and Oum et al. (2004) use Zhang and Zhang (1997) model, but also look at the decisions a private unregulated airport would make. The profit maximization problem faced by a private unregulated airport is

$$\max_{P,K,p} PQ - C(Q) - rK + Q(pX - c(X))$$
(2.9)

Zhang and Zhang (2003) and Oum et al. (2004) found that, while airside private prices diminish as it was conjectured by Condie (2000) and Starkie (2001), they decrease less than prices in a public airport that also has concessions, and that this is the case for both, first-best pricing (unconstrained public airport) and second best-pricing (budget constrained public airport). Therefore, concession revenues would not be a valid argument to deregulate. The intuition of the result is simple: a private airport would care about the extra profits it can make from concession activities; a public airport maximizing social welfare, however, would care about concession profits but also about the consumer surplus induced. Consequently, the decrease in the aeronautical charge would be larger in the public case: concession revenues would actually increase the gap between private and public airside charges.

Another important result obtained by Oum et al. (2004) was that the capacity investment rule of the private airport would be the same as the one the public airports follow, as in equation (2.3). Hence, they argued that, if the capacity can be adjusted continuously, the capacity investment decision of the private unregulated airport would be efficient from a social view point. However, since price and capacity decisions are simultaneously determined, and pricing rules are different, so will be actual traffic levels and capacity. In fact, since a private airport charges more, the actual capacity of the private airport would be smaller. But, their main point was that, conditional on traffic level Q, the capacity K determined by (2.4) would be efficient because marginal benefit equals marginal cost. On line with the actual capacity of private airports being smaller when capacity can be adjusted continuously, Zhang and Zhang (2003) found that, when capacity is indivisible, a private airport would make the (lumpy) addition of capacity later than a public airport. Note that none of these two results imply anything about the level of actual delays, because traffic levels will be different as well.

Czerny (2006) also looks at the effects of concession revenues on airside charges. There are two important differences between his model and Oum et al. (2004): First, he considers an airport that is non-congestible and which has spare capacity, making the reasons for crosssubsidization discussed above to vanish. Second, in Oum et al. (2004) the number of actual flyers would depend only on the full price ρ and not on the price for concession services. The concession services price would only determine how many of those already flying buy concession services. Czerny (2006) however, considers that both, airport and concession charges affect the number of flyers, and that the complementarity arises because only people actually flying will be able to purchase concession goods. Hence, in Czerny's setting it may happen that the airport charge is higher than a consumer's willingness to pay for flying, but that negative payoff is compensated by positive benefits arising from consumption of commercial services. These differences are material. Czerny shows that in this setting, the monopoly charge for aeronautical activities is actually higher with concession revenues than without, rejecting Condie (2000) and Starkie (2001) conjecture. The intuition is as follows: when the airport has concession services, and since these influence the number of flyers, the airport may increase its revenues in two ways. It may increase the price for aeronautical services, using a low concessions charge to mitigate the decrease in demand, or it may decrease its aeronautical charge, hoping to make revenues on the concessions side. But since only passengers can buy commercial services, the demand for the latter is a subset of the demand for flights. Therefore, and increase in aeronautical charges increases revenue more than an increase in the concession services charge.

Efficiency implications of alternative forms of regulation

Although traditionally airports have been owned by governments (national or local) a wave of privatizations started in the late 1980s. These privatized airports have been regulated out of the concern that they would exert market power given their monopoly nature. However, many economists have argued that the regulation mechanisms have failed to provide the airport with the correct incentives for pricing and capacity investments. Oum et al. (2004), Lu and Pagliari (2004) and Czerny (2006) analyze the effects of alternative mechanisms of regulation on the performance of private airports, with a particular focus on how revenues from concession services should be dealt with.

Oum et al. look into four different regulation mechanisms: single-till rate of return (ROR), dual-till rate ROR, single-till price cap and dual-till price cap. Let us discuss first the ROR mechanisms. Under the single-till ROR, airport charges (both for airside and concession operations) are set for cost recovery plus a *fair* return on invested capital. If u is the allowed ROR, then the new problem the private airport solves is:

$$\max_{P,K,p} PQ - C(Q) - rK + Q(pX - c(X))$$

s.t. $PO - C(Q) + O(pX - c(X)) = uK$ (2.10)

The well-know problem with ROR is that, if the allowed return is greater than the cost of capital, i.e. u>r, the airport has an incentive to over-invest in capital, a problem know as the Averch-Johnson effect. However, if the regulators get the allowed return right, the problem vanishes. It has been argued though that, even if the allowed return is chosen correctly, the single-till ROR would still misplace the incentives on terms of the productive efficiency, because it is essentially a cost-based mechanism. While the argument is sensible and has been detected empirically in several industries, it does not flow analytically from model (2.10).

Under the dual-till ROR, the allowed return applies only to aeronautical operations. If the regulators get the allowed rate right, the new restriction is PQ - C(Q) = rK. In this case, then, the airport would make no profits in airside operations and therefore, would try to maximize the profits coming from concession operations: Q(pX - c(X)). Given the complementary nature of the concessions demand, the airport will, in fact, try to maximize traffic, which is equivalent to minimize the full price ρ . Hence, this regulation mechanism would lead to a capacity rule as in the public case, that is (2.3), and to average cost pricing, that is P = (C(Q) + rK)/Q. Note however, that if u > r, the Averch-Johnson effect reappears.

We now turn to price-cap regulation, mechanism in which the regulator sets a ceiling for the aeronautical charge, that is $P \leq P^*$. Theoretically, the cap is set to limit the airport's market power, while ensuring its financial viability (this may include a fair rate of return on capital investment). The difference between the single-till and dual-till is, again, related to whether concession revenues will be lumped together with airside revenues or not; to be perfectly clear, the debate is not about regulating concession activities. Under single-till price cap regulation, the cap P^* will be set considering the fact that the airport will make revenues from concession activities. According to Oum et al. (2004), this implies that most of the revenues will have to come from concessions and, therefore there will be a crosssubsidy, just as in the case of a public airport subject to budget constraint (Zhang and Zhang, 1997). However, a problem is that the more revenues the airport makes from concessions, the smaller the allowed price would be in future revisions of the cap, even if traffic growths and congestion builds. Because of this, single-till cap regulation for the case of congested airports has been criticized (e.g. Starkie, 2001): the airport charge would not be a useful signal to users regarding congestion. Moreover, Oum et al. (2004) also showed that a price cap induces under-investment in capacity, worsening the problem. This result is in fact very robust. Spence (1975) showed that if a monopolist who initially can choose both price and quality of its product, is constrained to charge below some price ceiling, the quality it chooses will be always below what is socially optimal for that price.

On the other hand, under dual-till price cap, that is, when concessions revenues are not considered in establishing the cap, Oum et al. (2004) showed that the cap would not be set as low as in the single-till, something that seems desirable. However, the problem of under-investment in capacity would worsen. Hence, overall, Oum et al. (2004) concluded that the presence of the concession revenues make the dual-till ROR approach a quite interesting mechanism as it would induce the airport to invest optimally in capacity, while minimizing its costs and congestions delays, since it would try to minimize the full-price. Indeed, Spence (1975) suggested that ROR has nice properties when regulating both quantity and quality.

Lu and Pagliari (2004) also looked at the effects of single-till and dual-till price cap regulation. They considered a social welfare function as in (2.1), but considered that more traffic caused no congestion, that is D=0. The difference is that in a model with a delay function non-homogenous of degree one, congestion is essentially a cost. And given that the cost increases importantly as the traffic gets closer to capacity, equilibrium levels of traffic would never surpass capacity. In Lu and Pagliari's case however, if the aeronautical charge is too low, demand may well exceed capacity, particularly because in their model, capacity is assumed to be fixed. Lu and Pagliari found that a single-till would be appropriate when the average cost of the airport is larger than the market clearing price (for the given capacity), because cross-subsidies from concession revenues would be needed to decrease the airside charge and restore full capacity use. In other cases, however, they found that a dual-till would be better: under the single-till the price cap may be set 'too low' because of the concession revenues and hence dead-weight losses would occur because of excess demand.

Czerny (2006) also compared single-till to dual-till price cap regulation. As discussed previously, he considers an airport that is non-congestible and which has spare capacity, and considers that both, airport and concession charges affect the number of flyers. Under these conditions, he founds that the single-till dominates the dual-till social welfare wise, a result which is similar to what Lu and Pagliari found when the airport does not suffer from excess demand. The intuition is that with the single-till, the regulator has better control of the overall profits of the airport, which is not the case with dual-till regulation, which therefore helps to limit market power.

Hence, overall, when the airports are not congested, a single-till price cap seems like a reasonable approach to control market power. However if congestion actually occurs, the single-till would induce incorrect signals regarding congestion, while the dual-till would distort capacity investments. Further, if there are delays as traffic levels approach capacity (as in the original setup), the socially optimal pricing structure would require cross-subsidization (Zhang and Zhang, 1997), but this is precluded in the dual-till. Hence, in congested airports, dual-till ROR regulation may be a better option: the incentives for capacity investments would be well placed, while the regulated airport would pursue average cost pricing.

Airport pricing considering environmental costs

Carlsson (2003) developed a model of airport pricing that, in addition to congestion, also included environmental damages considerations (noise, emissions). For this, he modified the social welfare function in (2.1) to include environmental costs, as follows:

$$\max_{P} \int_{\rho}^{\infty} Q(\rho) d\rho + PQ - C(Q) - rK - QE(D(Q, K))$$
(2.11)

where E is the average environmental cost per flight. It depends on the level of congestion because, for example, delays increase fuel consumption and hence increase emissions. Carlsson considered many periods throughout the day and allowed for the environmental costs to vary according to the type of aircraft. For simplicity we do not do so here; the intuition of the results remains unchanged. The optimal pricing obtained has now two more terms than the *only congestion* social-marginal-cost pricing in (2.3) when $\lambda=0$:

$$P_{t} = C' + Q \frac{\partial D}{\partial Q} + E + Q \frac{\partial E}{\partial D} \frac{\partial D}{\partial Q}$$
(2.12)

The last two terms in the left-hand side represents marginal environmental cost: in addition to the airport's marginal cost and the marginal cost of congestion, each aircraft would have to pay the environmental cost it produces, plus another sum, owing to the fact that the extra delay a new flight imposes in existing flights increases the average environmental cost of all flights.³ These last two terms are obviously positive, which shows that, when

 $^{^{3}}$ The optimal charge is differentiated between type of aircraft and time of the day when these are differentiated.

environmental costs are considered, the airside charge is higher. Carlsson then points out that, if the proceeds from the environmental charge accrue to the airport, then cost recovery may be feasible. Whether this is the case or not, however is an empirical matter, as it depends heavily on the shapes of the delay function and the average environmental cost.

As for the capacity decision, although Carlsson did not look into it, it is fairly evident the direction in which it would change. Since now more capacity is beneficial not only because smaller delays decrease full price, but also because smaller delays diminish average environmental costs, the capacity investment rule would induce a larger investment in capacity.

2.3 The Vertical Structure Approach to Airport Pricing

This approach is newer and, hence, there are fewer papers. Here we review Brueckner (2002), Raffarin (2004), Pels and Verhoef (2004), Basso (2005) and Zhang and Zhang (2006). In this approach, the airline market is formally modeled as an oligopoly, which takes airport charges and taxes as given; these are two-stage games. Airports however, are not always considered integrally; in some cases, only airport authorities, who have to set a tax to be paid in addition to the airport charge –implicitly assumed to be marginal cost–, are considered. In these cases, airport profits do not enter the social welfare function. This approach has been driven by increasing levels of delays at hubs throughout the world, and therefore the focus has mainly been on optimal (public) runway pricing under congestion and airline market power, as can be see from Table 2.2. Until recently, capacity was assumed to be fixed and thus was not a decision variable of the airport or the airport congestion pricing and *road* congestion pricing.

Brueckner (2002) should undoubtedly be credited for starting this stream of literature; his work has been very influential. He considers N airlines that are seen as homogenous by consumers and which compete in Cournot fashion. He allows for peak and off-peak demand, which are interrelated, and where the peak period consists of a set of relatively short time intervals containing the daily most desirable travel times. Only the peak is congested. In this sense, it would seem that peak and off-peak travel are vertically differentiated in that, other considerations such as income and congestion levels absent, consumers would prefer to travel in the peak. In fact, he does not directly assume downward sloping demands, but starts with a continuum of consumers who would decide to use the peak or the off-peak depending on the full prices they perceive from airlines, that is, airfare, plus congestion costs caused by delays at the airports. However, Brueckner also adds a 'tendency to flight in business', which correlates to travel in the peak, as a device that would enable simpler (non-corner) solutions. The problem with this is that it actually imposes that, in terms of pure utility, with no income or congestion effect whatsoever, some consumers would *prefer* to travel in the off-peak. This seems to contradict the idea of 'the most desirable travel times'.

The airlines, observing the demands and understanding how consumers' decisions are taken, make their quantity decisions. An important aspect here is that congestion also affects airlines: there are externalities in production in that, the more a rival produces, the higher a firm's marginal and average cost will be. The delay function is a non-necessarilylinear function of traffic. In equilibrium then, the sorting of consumers towards peak and off-peak happen through the (airlines) quantity decisions. Brueckner then looks at what should be the optimal additional tax that should be charged to airlines in the peak period, in order to adequately account for the congestion externality. In this sense, he looks at the regulator case in that the airport is not formally incorporated into the analysis: its profits do not enter the social welfare function, which is compound only by consumer surplus and airlines profits, and there is no consideration of cost recovery, something that has drawn important attention within the traditional approach (see Section 2.2). Brueckner's main conclusion, and the one that since has drove research in the area, is that with Cournot oligopoly, each airline internalizes only the congestion is imposes on itself and its passengers, which enables a role for congestion pricing. In a symmetric airlines case, the optimal toll that should be charged during congested periods is equal to the congestion cost from an extra flight times one minus a carrier's share. In particular, a monopoly airline would perfectly internalize all the congestion it produces and hence there would be no space for congestion pricing. This shows the difference with the road case: with market power, the degree of internalized congestion is usually sizeable.

Pels and Verhoef (2004) wanted to expand Brueckner's work in two directions: first, they wanted to explicitly consider market power distortions and its effects on optimal congestion tolls. Second, they wanted to consider the fact that, at an origin-destination (OD) pair, the airports may not collaborate to maximize overall social welfare, but may maximize local measures of it. Their model is as follows: they consider an OD pair in which the airports decide charges prior to competition in the airline market. The capacities of the airports are assumed to be fixed. In this OD pair, two homogenous and symmetric airlines compete in Cournot fashion, taking airport charges and taxes as given when they choose their quantities (frequencies). Congestion delays affect airlines costs; the delay function is a linear function of total traffic at an airport. Passengers have only demand for roundtrips, and choose airlines based on a generalized cost which is the addition of the air ticket and congestion delay costs. The model is solved by backward induction to obtain sub-game perfect equilibrium. Hence, the first step is to solve the airlines' oligopoly, in order to obtain a sub-game equilibrium which will be parametrically dependent on the congestion tolls charged at each airport. With that sub-game equilibrium at hand, the authors looked for the optimal taxes that should be charged at each airport in order to adequately account for congestion. Initially, they consider that a single authority handles both airports and consequently tries to maximize the sum of consumer surplus and airlines profits. Hence, as Brueckner, Pels and Verhoef look at the regulator case, in that the airports profits do not enter the social welfare function.

Their main result indicates that the optimal toll would have two components: a congestion effect (which is positive) and a market power effect (which is negative). The first part is the one identified by Brueckner: since airlines only internalize the congestion they imposed in themselves, the uninternalized congestion should be charged. The second term, which

decreases the toll, arises because of market power at the airline level. What happens is that the regulator, in maximizing social welfare, would need to subsidize the airlines to induce them the produce more. The sign of the optimal toll is therefore undetermined; when the market power effect exceeds the congestion effect, a subsidy is necessary. The toll would be positive if the congestion effect dominates. They pointed out, for example, that this would undoubtedly be the case for a monopoly airline.

Pels and Verhoef compared their toll to the *pure congestion* toll suggested by Brueckner (2002). They found that, when the market power effect is strong, a pure congestion toll may actually be harmful for social welfare, since airlines are charged with a tax when in fact they should be receiving a subsidy. Brueckner did acknowledge this, though, by stating in his proposition that, "since congestion pricing corrects one distortion but leaves the residual market-power effect in place, tolls are guaranteed to be welfare improving only if that effect is sufficiently small. Otherwise, a negative welfare effects is possible." (p. 1367). Pels and Verhoef argued that, if a negative toll (subsidy) is optimal but unfeasible (for example for political reasons), the regulator should charge a zero toll.

Finally Pels and Verhoef considered the case in which, at each airport, different regulators only maximize consumer surplus of passengers that live in the airport's region, plus the profits of the *home airline*. The non-cooperative behavior of airports obviously hints that the result will be inferior to the single regulator case. In fact, the authors show, both numerically and analytically, that in the non-cooperation case, tolls at each of the two airports would always be positive.

Raffarin (2004), as Brueckner (2002) and Pels and Verhoef (2004), is interested in the optimal airport toll. But rather than considering a two-stage model, she considers three stages. In the first stage of her model, the airport chooses its price; its capacity is fixed. But then, conditional on the airport charge, duopoly airlines sequentially decide frequencies and then prices. The difference with Brueckner and Pels and Verhoef is that, in their case, airlines only decided frequencies; price is determined in equilibrium by the Cournot assumption. Raffarin, however, has a system of differentiated demands (obtained from a representative consumer framework), such that an airline's demand increase when its frequency increases or its price decreases, and decreases when its rival's frequency increases.

Raffarin's model has three key assumptions that determine her results: first, she assumes that, even though frequencies are airlines' decisions, any demand will always be fulfilled. And this is not ensured by airlines' choice of aircraft size, k, because k is an exogenous parameter in the model (that is, equilibrium results will be dependent on k). Hence, there is no real connection between the number of passengers and the number of flights, other than the assumption that there will be enough space. Both Brueckner (2002) and Pels and Verhoef (2004) made a *fixed-proportions* assumption, by which the number of passengers in a flight was a fixed constant. This make it easier, yet transparent, to transform the demand in terms of passengers, into an airport's demand in terms of flights. The second assumption is that congestion delays –which as in Pels and Verhoef (2004) increase linearly with total traffic– do not affect consumers' or airlines' decisions. Instead, congestion costs

are subtracted in the social welfare function, which, interestingly, explicitly considers the airport's profits. Hence, in this case, airlines do not internalize any of the congestion they cause because it does not directly affect them (it is not a cost to them), and its customers do not care about congestion either. Finally, the third important assumption is that an airline's operational cost per flight, z, depends on the aircraft size in an increasing fashion, that is dz(k)/dk > 0. Hence, even though using larger aircrafts means fewer flights, which saves on costs, each of those flights will be individually more costly. Aircraft size however is not a decision variable but a parameter. The implication is then that for given airport charges, equilibrium frequencies increase as the aircraft size diminishes.

Rafffarin then maximizes social welfare –which is the sum of airport's profits, airlines profits, and consumer surplus, minus congestion costs– in order to find what the optimal frequencies are, that is, the optimal level of airport's demand. The optimal airport charge is then obtained as the price that would induce the optimal frequencies. The optimal charge she obtains has three components (which she did not recognize): airport's marginal cost, plus the costs of congestion (recall that airlines do not internalize any fraction of congestion in this model), plus a third term, which is negative, and that could be assimilated to Pels and Verhoef's market power effect. The interesting twist however, is that this term depends on the aircraft size, k, and diminishes the higher k is. That is, the airport charge should be larger for smaller aircraft. And since aircraft size and weight are positively correlated, this implies that the airport charge should decrease with the aircraft weight, rather than increase as it is usually the case. The airport would reward airlines that use larger aircrafts because that implies smaller frequencies and hence smaller congestion costs. The choice of k however, is not endogenous for the airlines in the model.

The papers we have reviewed so far have in common two important features: they all consider maximization of social welfare and in all three cases the airport capacities were fixed. In close but independent work, Basso (2005) and Zhang and Zhang (2006) generalized these two aspects. Both of this papers considered that the airport decides on price and capacity in the first stage, and in the second stage N airlines choose quantities (frequencies) in Cournot Fashion. The airlines have identical cost functions, but are insensitive to congestion costs in Zhang and Zhang (2006) while they do have extra costs owing to congestion in Basso (2005). Passengers, as usual, are sensitive to the full price of travel, that is airline ticket plus congestion delay costs.⁴ Both used congestion delay functions that are not homogenous of degree one in the traffic to capacity ratio, that is, congestion increases more than linearly with total traffic (for a given level of capacity). Other differences between the two papers are: Zhang and Zhang considered that airlines are homogenous in the eyes of the consumers, while Basso allowed them to be horizontally differentiated (in non-address fashion). Basso also considered in the full price perceived by the passengers another time cost, namely, schedule delay cost. This time cost arises because flights do not depart at a consumer's will but have a schedule. Hence, schedule delay costs are a sort of waiting time, which decreases with higher airlines' frequencies. On the other hand, Zhang and Zhang considered a general demand function (of the full price) while Basso considered a more restrictive system of demands: linear in the full-prices of airlines.

⁴ This last point is enough for the internalization of own congestion by an airline to arise in oligopoly, as discussed before. It is not needed for both, airlines and consumers, to be sensitive to congestion costs.

Both, Basso (2005) and Zhang and Zhang (2006) solved by backward induction, characterizing the shape of the derived demand for the airport through comparative statics. Then, they both considered three different objective functions (Basso considered two more which are discussed later): unregulated profit maximization, unconstrained social welfare maximization, and social welfare maximization subject to cost recovery. Let us discuss first the pricing rules obtained. In the case of unconstrained maximization of social welfare, they both considered a social welfare function in which the airport profits were included. They find that, in their more general settings, Pels and Verhoef's insight go through: the optimal pricing rule is the sum of airport's marginal cost, plus a congestion effect (positive) and a market-power effect (negative). When capacity is fixed, this pricing rule shows that with large values of N, the congestion effect is large while the market power effect weakens. Smaller values of N imply a weaker congestion effect but a stronger market power effect. With this pricing rule, the airport manages to obtain a first-best outcome in the airline market. Note that, in this setting, rather than a regulator setting the toll, is the airport who would distort marginal cost pricing to account for both, uninternalized congestion and market power. Since the optimal airport charge may be below marginal cost and even below zero, the airport may run a deficit.

In the case of unregulated profit maximization, Basso (2005) and Zhang and Zhang (2006) clearly found in the pricing rule of the airport the double marginalization problem that affects uncoordinated vertical structures. For a given capacity, the airport charge will decrease with the number of airlines downstream. On the other hand, and in a somewhat expectable result, an airport that maximizes social welfare subject to cost recovery will have a charge that is in between the previous two. The balance will be given by the severity of the budget constraint.

Turning to capacity decisions, Basso (2005) and Zhang and Zhang (2006) found that an unconstrained welfare maximizing airport will provide capacity until the marginal cost of capacity equates the marginal benefits in saved delays (to airlines and passengers in the case of Basso, to passengers only in the case of Zhang and Zhang). Interestingly, Zhang and Zhang (2006) proved that when both price and capacity are decision variables, in their setting, the market structure (i.e. N) has no impact on airport's actual demand and capacity. Consequently, delay levels will be independent of market structure. This however does not hold in Basso's setting, in which airlines are differentiated and/or passengers care about schedule delay cost. The explanation has to do with the 'preferred N' of a welfare maximizing airport. Basso showed that there are two opposing effects. With congestion and market power effect controlled, as it is the case here, fewer airlines in oligopoly would provide –each of them–, higher frequencies than more airlines, thus delivering smaller schedule delay costs which increases social welfare. Smaller N would be preferable. On the other hand, differentiation brings about new demand when N increases, so a larger N is preferable.

A private airport, however, would increase its capacity until the marginal revenue of doing so equate its marginal revenue. Clearly, this capacity rule is different that the previous one. Basso (2005) note then, that this is different than what happened in the traditional approach

(e.g. Oum et al. 2004), in which the capacity rules of private and public (unconstrained) airports were the same. However, when N goes to infinity, i.e. airlines become perfectly competitive, the pricing rules become the same. The explanation for this is given in the next Section. Further, Zhang and Zhang, and Basso, showed that *conditional on the level of traffic*, a private airport would oversupply capacity. However, that capacity would most likely be too small in a second best sense. That is, a public airport forced to charge using the private airport pricing rule, would most likely supply more capacity than the actual capacity offered by the private airport (Basso, 2005). As with price, a budget constrained airport would, *conditional on the level of traffic*, choose a capacity that is in between the private capacity and the unconstrained public capacity.

Basso looks at two other types of ownerships as well. First, he looks at the case in which airports and airlines vertically integrate. The reason to look at this is because it has been often argued that more strategic collaboration between airlines and airports would solve incentive problems, particularly regarding capacity expansions. Basso finds that the airport charge would include marginal cost, a term equal to the uninternalized congestion cost of each carrier, but would also include a third term, which is positive. This mark-up is put in place to fight the *business-stealing effect*, a horizontal externality typical of oligopoly: firms do not take into account profits lost by competitors when expanding their output. By increasing the airlines' marginal cost, the airport would be able to induce a profitable (for the whole structure) contraction of total output. In fact, the final outcome is indeed that of cooperation between competitors in the airline market. The intuition is that airlines would 'capture' an input provider to run the cartel for them, given that they are unable to collude on their own. As for capacity, the vertically integrated structure would have the same capacity rule than the unconstrained public airport. The actual capacity however would be below the 2nd best capacity (i.e. a public airport forced to charge using the vertical integration pricing rule would supply more capacity). Basso also showed that, depending on how strong airlines' differentiation is, and how strong schedule delay effects are, profits may be higher when the airports integrate with a single airline. A non-integrated private airport though will always prefer a larger N^{5}

Basso (2005) also looked at the case in which two distant airports are privatized separately. Social welfare wise, results worsen because when airports are both, origin and destinations of trips, their demands are perfect complements and therefore 'competition' between airports induces a horizontal double marginalization problem. This horizontal double marginalization arises in both integrated and non-integrated vertical structures.

2.4 Relationship between Approaches

It is clear that the two approaches –which we have called *traditional approach* and *vertical structure approach*–, are rather different and that the questions examined with each of them have not perfectly overlapped. This raises questions about the transferability of results

⁵ Both Brueckner (2002) and Zhang and Zhang (2006) had N airlines downstream. Public airports and vertically integrated airports would have no particular preference for N in their settings though, because airlines are homogenous and there are no schedule delay effects.

which we attempt to answer next; most of this Section is a simplified exposition of some materials available in Basso (2005).

In the *traditional* approach, the airline market is not formally modeled, under the assumption that the airport charge would be completely passed to consumers, and that airline tickets and other charges would be exogenous to the airport. Oum et al. (2004) argue that this would be the case under perfect competition. In the *vertical structure* approach, on the other hand, it is recognized that airports provide an essential service that is required by airlines to move passengers; therefore, airports are viewed as providing a necessary input for the production of and output: travel. In fact, some authors using the *vertical structure* approach have been somewhat critic of the *traditional* approach on the grounds that it does not properly consider all actors involved. For example, Raffarin (2004) say that it is rather strange that the pricing rules obtained from the *traditional* approach do not consider passengers' utility. This is not completely accurate though. Passengers are indeed somehow considered in the approach, as delay costs affect them as well, something that Raffarin missed.⁶ But what it is true is that a vision of the problem that recognizes that airports provide a necessary input for the production of an output for the production of an output that is sold at another market, seems more complete.

Using the notation of Section 2.2, what the papers in the vertical structure approach have shown, is that for any given airport charge, P, and airport capacity, K, the airline market – the downstream market- will reach some equilibrium. This equilibrium is constituted not only by equilibrium traffic but also by equilibrium delays and air ticket prices. By stressing this fact, three things become apparent. First, that as long as the airport is concerned, its demand is going to be some direct function of P, K and of the (exogenous) airline market structure, which in most papers is represented by the number of airlines N. Hence, the airport's derived demand would be Q(P, K; N). Delays enter the picture through the equilibrium of the downstream market. How does this demand faced by the airport respond to changes in P and K is something that formal analysis of the airline market can unveil. Second, that how airport charges and airlines' delay costs are passed to consumers is built inside the demand faced by the airport and depends on the nature of the equilibrium reached in the airline market. In this sense, it would seem that a full price model pertains more to the airline market stage than the airport market stage. And third, that other airline charges are not exogenous to the airport because the downstream equilibrium -- that is, the airport demand- depends on P and K, which are decided by the airport. Airport managers with foresight will take this into account and decide user charges and capacity accordingly.

We can then go back to the *traditional approach* and contrast its basic setting with what we have described above. Two important questions arise:

1. Is it reasonable to use the full price idea at the airport, rather than at the airline market level? That is, under what conditions it would be legitimate to assume that

 $^{^{6}}$ The problem may lie in that Morrison (1987) states that the final consumers of airports services are airlines, even though in his model congestion explicitly affects passengers. In Ourn et al. (2004), passengers are said to be the final consumers.

the airport demand can be written as $Q(\rho)$ – with $\rho = P + D(Q, K)$ – rather than as Q(P, K; N)?

2. If under some conditions the airport demand can reasonably be written as $Q(\rho)$, would its integration give a correct measure of *consumer* surplus? We have learned that the consumers of airports are both airlines and passengers. Hence, a social welfare function should include both airlines profits and passenger surplus. This is in fact what is explicitly done in the vertical structure approach when analyzing the maximization of social welfare. In the *traditional approach*, however, that is not the case. There, *consumer* surplus has been obtained through integration of the airport demand function with respect to a full price. Under what conditions does the derived demand for the airport carry enough information about the downstream market so that its integration gives a correct measure of airlines profits and passenger surplus?

In short, these two question attempt to clarify how the two approaches are related. This is what we undertake in this Section. We start by concisely describing a simple oligopolistic airline market, which takes airport charges and capacities as given. We use this model to derive the demand for the airport. Consider N homogenous and symmetric airlines servicing a congestible airport. Airlines are assumed to choose quantities (frequencies) in Cournot fashion. Let Q_i be airline *i*'s number of flights, and $Q = \sum_{i=1}^{N} Q_i$ be the airport's demand. q_i denotes the demand, in terms of passengers, of airline *i*, while $q = \sum_{i=1}^{N} q_i$ is the aggregate demand for airline services. Assuming that aircraft size times load factor equals S for all carriers, we have that:

$$Q_i S = q_i , \qquad QS = q \tag{2.13}$$

We assume that the aggregate demand for airline services is linear in the full price. The full price, θ , will be the sum of the air ticket, t, plus passenger delay costs, D. It is, therefore, a full price at the airline market level. Thus we have:

$$q(\theta) = a - b\theta \tag{2.14}$$

$$\theta = t + D(Q, K) \tag{2.15}$$

where a and b are positive. Inverting (2.14) and using (2.13) we obtain

$$\theta(Q) = A - SBQ \tag{2.16}$$

where A and B are obviously defined. Further, replacing (2.15) in (2.16) gives us t(Q) = A - SBQ - D(Q, K), which is the aggregate inverse demand faced by the airlines. Assuming for simplicity that airlines have constant operational marginal costs, c, and that they do not suffer from congestion⁷, an airline's profit is:

⁷ This assumption does not influence our results below and could be easily relaxed, see Basso (2005).

$$\pi^{i}(Q_{i}, \mathbf{Q}_{-i}, P, K) = [A - SBQ - D(Q, K)]SQ_{i} - (c + P)Q_{i}$$
(2.17)

where *P* is the airport charge. To obtain the derived demand for the airport, we need to find the equilibrium of the airline market. The Cournot-equilibrium is characterized by first order conditions $\partial \pi^i / \partial Q_i = 0$. Computing these and imposing symmetry, we obtain:

$$\Omega(Q, P, K; N) = D(Q, K) + \frac{Q}{N} D_Q(Q, K) + Q \frac{(N+1)}{N} BS^2 + c + P - AS = 0$$
(2.18)

Equation (2.18) implicitly defines a function Q(P, K; N), which is airports' demand as a function of airport charges, capacities and airline market structure, N. An explicit expression of the airports' inverse demand P(Q, K; N) is also obtainable. From (2.18), comparative statics would enable a characterization of the shape of the airport's demand function (see Basso, 2005; Zhang and Zhang, 2006). Here, however, we will not follow that path because we are interested in different questions. We can in fact now answer the first question we posed before. Forming $\rho = P + D(Q, K)$ in (2.18), we can re-write this equation as:

$$Q\frac{(N+1)}{N}BS^{2} + \rho + c - AS + \frac{Q}{N}D_{Q} = 0$$
(2.19)

Hence, in general, Q would not depend only on ρ but also on D_Q and N: the (implicit) demand for airports then should be $Q \equiv Q(\rho, D_Q, N)$ and not $Q(\rho)$. However, in the perfect competition case, i.e. when $N \to \infty$, (2.19) leads to $Q(N \to \infty) = (AS - c - \rho)/(BS^2)$, which implies that $Q(\rho, D_Q, N \to \infty) \equiv Q(\rho)$. Thus, under perfect competition, a full price as defined by ρ , can in fact be used directly at the airport market level. It does summarize well the equilibrium of the downstream market.

We can thus turn to second question. If we assume that there is perfect competition, would integration of $Q(\rho)$ correspond to airlines profits plus passenger surplus? This question is related to the more general subject of the relation between input and output market surplus measures (see Jacobsen, 1979; Quirmbach, 1984; Basso, 2006). To answer it, we first compute the surpluses –in sub-game equilibrium and when N goes to infinity– of airlines and passengers. Passenger surplus is given by $PS = S\left(\int_{0}^{Q} \theta(Q)dQ - \theta(Q)Q\right)$, where $\theta(Q)$ is given by (2.16). Straightforward calculations lead to

$$PS = \frac{BS^2 Q^2}{2} \tag{2.20}$$

The aggregate (equilibrium) profit of airlines, $\Pi = \sum_{i} \pi^{i}$, is easily obtainable from an individual carrier's profit (2.17) and the imposition of symmetry, that is, $Q_{i} = Q/N$. Taking the limit when $N \to \infty$ and regrouping terms to form ρ , we obtain:

$$\Pi(Q,\rho) = QS[A-QSB] - Q[c+\rho]$$
(2.21)

Consider now the total derivative of Π with respect to ρ . Noticing that $AS - c - \rho = QS^2B$, we obtain

$$\frac{d\Pi}{d\rho} = -Q(\rho) - QBS^2 \frac{\partial Q}{\partial \rho}$$
(2.22)

Integrating from ρ to ∞ , reordering, and using equation (2.20) we finally get⁸

$$\int_{\rho}^{\infty} Q(\rho) \, d\rho = \Pi + PS \tag{2.23}$$

Therefore, equation (2.19) shows that, when $N \to \infty$, one can reasonably write the airport demand as $Q(\rho)$, with $\rho = P + D(Q, K)$. Equation (2.23) further show that, in that case, integration of the airport demand with respect to the full price ρ , will deliver a correct measure of consumer surplus, that is, airlines' profits plus passenger surplus. Basso (2005) shows that with differentiated demands, this result holds.

Perfect competition in the airline market was in fact, the maintained assumption of Oum et al. (2004). Hence, we have provided theoretical support for their claim. But we have also provided boundaries for the use of the *traditional approach*: it would be reasonable to use it only if market power at the airline level is absent. If market power is present however, modeling the demand for the airport as $Q(\rho)$ would be incorrect. Furthermore, its integration with respect to ρ would fall short of giving the sum of airlines' profits and passenger surplus (see Basso, 2005). In this case, a full model that formally considers the three actors involved, as in the vertical structure approach, is necessary.

Lastly, since $Q(\rho)$ is not an approach that can be used when there is market power downstream, one may well wonder whether by using the demand function Q(P, K; N) – which may be estimated empirically for instance–, and by integrating it with respect to P, one can adequately capture consumer surplus. This is not the case unfortunately. Using results in Basso (2006) it can be shown that integration of the airport demand would give:

$$\int_{P}^{\infty} Q(P,K;N) dP = \Pi + \frac{N-1}{N} PS - \frac{(N-1)}{N} \int_{P}^{\infty} Q \frac{\partial Q}{\partial P} D_{Q} dP$$
(2.24)

⁸ Here, we used the fact that $Q(\rho = \infty) = 0$ and therefore $\Pi(\rho = \infty) = 0$.

Hence, there is no value of N for which the integral of the airport demand with respect to P equals airlines profits plus passenger surplus. Not even if N is very large.

2.5 **Pricing of Airport Networks**

The papers we have reviewed, in both approaches, are papers that do not really deal with airport networks. In most cases they deal with an airport in isolation. The exceptions, so far, have been Pels and Verhoef (2004) and Basso (2005) who consider a 'network' of two airports. Yet, real air networks are obviously more complex than that, and it is fairly clear that in these real networks other issues arise. We review here three papers that have dealt with networks of airports, that is, three or more airports.

Oum et al (1996) argue that in hub and spoke networks (HS), airports' demands are complementary because any take-off at a spoke airport will generate a landing at the hub. This complementarity is of different nature that the complementarity that arises in two-airport networks because the presence of a hub introduces asymmetries. As in the two-airport cases, failure to consider the complementarities when looking for optimal pricing policies will result in social welfare losses. But in a HS network, congestion at the hub will build up more rapidly than at spoke airports. And when budget adequacy is an issue, this may imply the need for cross-subsidizations between airports. Depending on the type of ownership however, cross subsidies may be unfeasible. Oum et al. (1996) study how ownership and cost recovery constraints affect airport pricing in a HS network and, consequently, social welfare.

Oum et al. (1996) consider *n* airports in a hub and spoke system: *n*-1 airports are spoke airports and there is one hub. All the airports have constant operational marginal costs and fixed capacity, but capacity maintenance costs are positive. The demands for spoke airports depend on their own charge and the hub charge. Demand for the hub airport depends on its own charge and the charge of all spoke airports. All airports are congestible, but congestion is an external cost that the airport authority will include in the social welfare function; it does not affect the demands (as in Raffarin, 2004; see Section 2.3). This set up shows two things: First, that the spoke airports' demands are indeed complementary with the hub's demand, but they are not directly complementary between them. Second, that this paper is ascribable to the *traditional* approach, since the airline market is not formally included. Indeed, consumer surplus is measured as the integrals of the airports demand.

Oum et al. first analyze the case in which all the airports are publicly owned and under the control of a single authority: this is the 'federal' case. The authority will maximize the sum of airports profits plus consumer surplus –the sum of the integrals of airports demands– minus external congestion costs. The optimal pricing policy would have all airports charging social marginal cost (SMC), that is, operational marginal cost plus the external costs of congestion. Since the hub is more likely to be heavily utilized, congestion will be larger there than at spoke airports. Hence, they assume that SMC pricing would lead to cost recovery at the hub but to deficit at spoke airports. The first best federal case then would
require cross subsidies from the hub to the less utilized spoke airports. If a budget constraint is set in place, the question becomes whether the hub makes enough profits to cover for the spoke airports deficits. If it does, where back in the first best case. If it does not, then Ramsey pricing is called for: the charge at the hub will increase. Cross subsidization will be obviously still needed and this alternative will be welfare inferior to SMC pricing.

They then look at the case in which each airport is under the control of a different authority who, subject to cost recovery, maximizes own social welfare, that is, the integral of own demand plus own profits, minus congestion costs. This is the 'de-federalized' or 'local government' case. Given the assumption about SMC not covering costs in spoke airports, in this case, the hub will price at SMC, but the spoke airports would charge average costs to ensure cost recovery. Since individual cost recovery implies overall cost recovery, this case will be inferior, social welfare wise, to the previous Ramsey pricing case. In general, in the federal case, and independently of whether SMC or Ramsey prices are used, charges at the hub will be larger and charges at the spoke airports will be smaller than in the local government case. Oum et al. (1996) conclude that de-federalization of airports may imply social welfare losses: by not jointly pricing the airports, the local airport authorities will not take into account that demands are complementary and cross-subsidies will likely become unfeasible. The welfare losses, though, would have to be balanced against possible X-inefficiencies gains that de-federalization may bring about.

Recall, however, that the conclusion of Section 2.4 was that a *traditional* approach would be justified only when air carriers are atomistic. What would happen if carriers have market power? In this case, we would need a *vertical structure* type of approach. This is what Brueckner (2005) analyzes. The main point here has to do with the meaning of *market power*. One of the conclusions in Section 2.3 was that congestion tolls would be decreasing on an airline's share of flights at the airport, because an airline only internalizes the congestion caused on own flights. Since in that Section, only one or two airports were considered, the share of flights at the airport was identical to the share of flights at the city-pair market level. However, when one considers even a simple network of airports in which airline competition exists, it is no longer true that the share of flights at the airports will be necessarily equal to the share of flights at the city-pair market level. Hence, the relevant question becomes, what is the relevant flight share for congestion internalization? Brueckner consider the following network, in which two airlines compete:



Figure 2.1: Network Structure and airline competition (Brueckner, 2005)

In this network, airport H is airline 1's hub, while airport K is airlines 2's hub. Airline 1 serves four city-pair markets: AH, KH, BH and AB (two legs). Airline 2 also serves four city-pair markets. The airlines compete in two markets, KH and AB, while each of them is a monopolist in other two markets. It can be easily recognized –for example under full symmetry– that airline 1's share of departures and take-offs at hub H is larger than airline 2's share. Similarly, airline 2 dominates hub K in terms of departures and take-offs. However, in the two markets where airlines compete, they would both have a 50% share of flights under symmetry. This nicely shows the difference between shares of flights at airports and share of flights in city-pair markets, which justify the research question.

To analyze what would be the optimal congestion toll, Brueckner uses a setup which is essentially the same as in his single airport paper (Brueckner 2002, see Section 2.3 for a description), but considering each of the various markets. Airports are assumed to have a fixed exogenous capacity while only the hubs are prone to congestion. The derivation of optimal congestion tolls is involved so it is omitted here, but the conclusion is simple and important: regardless of the degree of market power that an airline has in the city-pair markets it serves, the amount of congestion it internalizes depends only on its flight share at the congested airport. Hence, "the appropriate airport congestion tolls are carrier-specific and equal to the congestion damage from an extra flight times one minus the carrier's airport flight share" (Proposition 1, p. 612).

An important final point that Brueckner raise has to do with the market power effect we discussed in Section 2.3. There, we saw that, while a congestion toll is justified when carriers are oligopolistic, from a first best point of view a subsidy was also justified as a mean to fight market power at the airline level, and hence reduce allocative inefficiencies. In the simple settings of one or two airports, both, the congestion effect and the market power effects depended on a carrier's flight share. But, again, in that case the airport share and the city-pair market share were the same. Brueckner (2005) showed that, in a network setting, the congestion tolls are airport specific, while the subsidies required are city-pair specific. Hence, an airport regulator would need to calculate appropriate airport specific congestion tolls to be combined with city-pair specific subsides to obtain, finally, the

optimal charge, which would be positive if congestion effects dominate market power effects. Brueckner argues that, since market-level subsides are impractical to implement, only airport congestion tolls would be used, an approach that would be welfare improving, yet not first-best, if congestion effects dominate.

Now, both Oum et al. (1996) and Brueckner (2005) assumed that the route structure of airlines, that is, the way airlines move passengers between origin and destinations, remain unchanged and were independent of the pricing practices of airports. But, what would happen in the long run if route structures were adjustable? For example, it has been often argued that economies of density drive the selection of HS networks. But if congestion at hubs is too important, airlines may decide to by-pass them, offering direct connections in some city-pair markets.⁹ Would congestion pricing affect the timing of such a decision? May airports use their pricing practices as a way to compete for connecting passengers, that is, compete to become hubs? A model including all these elements would be indeed very complicated and has not been proposed yet, as far as we know. However, there is one paper that, even though in a context of non-congestible airports and a monopoly airline, does look at how airport pricing and airline's choice of route structure are related. Pels et al. (1997) consider a model with three public and non-congested airports, and a monopoly airline. Airport charges are directly made to passengers. Thus, demands for airports and for the airline depend on both air tickets and airport charges. The airports and the monopoly airline play a simultaneous game in which each airport choose its per-passenger charge, while the airline choose a route structure and its prices. The objective function of the airports is to maximize own social welfare (as in the de-federalized case in Oum et al., 1996), which is measured as the integral of the airport's demand. They are subject to a budget constraint though. The airline seeks to maximize profit. There are some key assumptions in the model, which are more easily explained using figure 2.2:



Figure 2.2: Possible route structures

First, it is assumed that node A has more passenger generating capacity. That is, if airports and airline charges were zero, demands in the OD pairs AB and AC would be α , while in the OD pair BC it would be $\sigma \alpha$, where $\sigma < 1$. Second, consumers only care about the monetary charges (from the airports and the airline) but would not care about travel times (which are higher in a HS route structure) or whether they have to make connections or not. Third, the marginal cost of carrying a passenger is constant and equal across links; hence,

⁹ For a paper related to this issue, see Basso and Jara-Diaz (2006)

in a HS route structure, a passenger traveling from B to C would cost to the airline 2c while with a FC route it would cost only c. Finally, if a link is used, it has a fixed cost c_0 . Hence, a HS route structure is cheaper in terms of fixed costs, as it only uses two links, but is more expensive in term of operational costs.

Pels et al (1997) show that, in this setup, if the airport charges are zero (or, if they are equal but chosen non-strategically, i.e. without considering what the airline does), a HS route structure will be preferred by the monopoly airline if $\sigma < \overline{\sigma}$, in other words, if the demand in the BC market is much smaller than the demands in the AB and AC markets. The limit $\overline{\sigma}$ increases with c and with c_0 and decreases with α . Further, they show that the airline will always choose to place its hub at the node with the highest level of demand, in this case, node A.¹⁰

When airports choose their prices simultaneously with the airline's choice of route structure and prices, Pels et al. show that that the airport charges will be increasing in airline prices, but airline prices will be decreasing in airport charges. The 'dynamics' of equilibrium would be: the monopoly, who is a profit maximizer, would increase its prices depressing demands. Since the airport has to break even, it would raise its own charges, but that would induce the airline to decrease its prices. This in turn would increase demand, inducing a decrease in the airports charge, which would induce the airlines to increase price. Eventually, this loop may reach equilibrium, although Pels and Verhoef show that nonexistence of equilibrium is a possible outcome. Since analytical solution of the equilibrium was unfeasible, they rely on a numerical simulation to extract more conclusions. They found that, only if σ is small enough, the airline would choose a HS route structure. The higher the σ , however, the better for the hub. More importantly, price competition between airports seems to have little effect on the airline's choice of a hub; the choice would still be made based on passenger generating capacity. Obviously, one can foresee that the actual geographic position of the airports would be important as well. A hub would not be placed really far away from all its spoke airports. But in this model, distances, that is the topology of the network, does not play a role. This is reasonable under the assumption that all airports are located fairly close or equidistant from each other.

2.6 Conclusions and Further Research

Airport pricing has been widely analyzed in the economics literature. In this survey paper, we have focused on analytical models of airport pricing, from 1987 on. We claimed that the models in the literature con be grouped into two broad *approaches*. Roughly, the *traditional approach* has used a classical partial equilibrium model where the demand for airports depends on airport charges and on congestion costs of both passengers and airlines; the airline market is not formally modeled, in several cases under the assumption that airline competition is perfect and hence airport charges and delay costs are completely passed to passengers. The *vertical structure approach* has instead recognized that airports provide an input for the airline market –which is modeled as a rather simple oligopoly– and

¹⁰ For more on the choice of route structure, see Hendricks et al. (1999), Pels et al. (2000) and Jara-Díaz and Basso (2003).

that it is the equilibrium of this downstream market which determines the airports' demand: the demand for airports is therefore a derived demand.

The questions examined with the approaches have not perfectly overlapped. The traditional approach has been wider in scope, having been used to analyze issues such as optimal capacity investments, the effect of concession revenues, privatization, the efficiency of alternative regulation mechanisms, cost recovery when capacity cannot be increased continuously, or the efficiency of weight-based airport charges. The vertical structure approach on the other hand, has mainly focused on calculating the additional toll that airlines should be charged to attain maximization of social welfare. It is only recently that vertical structure models have been used to assess issues such as optimal capacity levels or the effects of privatization on airport charges. However, the two approaches have not only examined different questions, but also grew somewhat disconnected, which raise the questions of transferability of results. Drawing from results in Basso (2005), we showed here that abstracting from the airline market, as is done in the *traditional approach* is a reasonable approximation only when airlines behave competitively, but it is not when airlines have market power. In the latter case, the derived demand for the airport would not be dependent only on its *full-price*, as it is assumed. As a result, the integration of the airport demand with respect to the full price, which is said to capture *consumer surplus*, would not adequately capture the surpluses of passengers and airlines because market power and congestion effects preclude it.¹¹

The fact that the airline market cannot be ignored if airlines have market power implies, on one hand, that future research would have to use vertical structure models to re-examine some of the questions that have been addressed only with the *traditional* approach, for example, effect of concession revenues on airport charges, the efficiency of regulation mechanisms and what happens if capacity is lumpy. But on the more practical side, the fact that the airline market has to be included in the models is also bad news for managers of public airports and regulators: to take optimal decisions, the amount of information required would be massive even in simple settings, which undoubtedly complicates the problem.

In the models we have reviewed in this survey, authors have resorted to a number of simplifications, which was the price to pay to preserve tractability. In the airline market of vertical structure models, two usual simplifications have been assuming fixed proportions and symmetric airlines. The fixed proportions assumption is made when authors assumed as constant the product between aircraft size and load factor (or both). Yet, it has been widely accepted that airlines enjoy what is called 'economies of density' –decreasing average cost on nonstop connections– because of economies of aircraft size. These economies are not considered under the fixed-proportions assumption. A variable proportions case would arise because, if the charge per flight is too high, airlines would have an incentive to change to larger airplanes, independently of existing or exhausted economies of aircraft size. So, with privatization for example, not only capacities and traffic level would be distorted downwards, but aircraft size would be distorted as well.

¹¹ This result in fact applies not only to airports but to any other type of transport terminal, or even railroad tracks, since the situation is essentially the same.

Modeling this effect is an interesting area of future research albeit a complex one, as larger aircrafts imply smaller frequencies, which directly affects congestion and demand through schedule delay costs.¹² Regarding asymmetric airlines, certainly insights would be gained, as the model would depict a more realistic case. Brueckner (2002, p.1368) stated that "cost differences across firms may not be a useful source of asymmetry however, because a planner would not allow high-cost firms to operate at the social optimum". In Basso (2005) and Zhang and Zhang (2006), however, there was no social planner but rather managers of public airports maximizing social welfare, who probably would not have the power to preclude less efficient airlines to operate. But they did not consider asymmetries. It seems to us that this would be unfeasible analytically. Numerical simulations for the case of asymmetric airlines would be required. Pels and Verhoef (2004) did present numerical simulation for the case of an asymmetric duopoly (see Table 2.2 for a description of their setting).

The papers in Section 2.3 looked at airports in isolation or at round trip travel between two airports at the most. In the latter case, that implied that airports have complementary demands, from where it followed that public airports that are priced independently would not achieve a first best (Pels and Verhoef, 2004), and private airports would ended up with a horizontal double marginalization (Basso, 2005). However, airport networks are more complex than that; and on this subject, the papers presented in Section 2.5 are good progress in understanding the main issues. But there is still work to do because in Oum et al. (1996) and in Brueckner (2005) there is no route structure decisions from the part of airlines, and it is through route structure decisions that airports may actually compete: they would be competing for connecting passengers. Pels et al. (1997) on the other hand, although having route structure decisions, do not include congestion, capacity choices, or airline competition. Further work in the pricing of airport networks–including effects of privatization and regulation mechanisms– is, in our view, a clear line of future work.

There may be also geographic competition; airports competing for costumers in the same origin, i.e. with overlapping catchment areas, as in the case of New York. There has been some empirical work on this issue (e.g. Ishii et al. 2005), but not too much work on the analytical side. Some papers have looked at competition between congestible Bertrand oligopolists (e.g. De Borger and Van Dender, 2006) but, as it has been discussed, that overlooks the vertical structure in which airports are inserted. A simple model of geographic competition between two airport-airline structures is Gillen and Morrison (2003); but they consider only one airline per airport, that the airport and the airline maximize joint profits, and that there are no congestion and capacity choices. We think this is another interesting area of research, in which we have been doing some work.

Another aspect that we think is important and that we are currently pursuing (Basso and Zhang, 2006) has to do with peak-load pricing, in addition to congestion pricing. Most of

¹² Note that Raffarin (2004) is not an analysis of variable proportions case because, although she did considered different aircraft sizes, the airlines where not free to decide about their preferred aircraft size. Rather, the aircraft size was exogenously given through a parameter, which thus showed up in the final pricing rules of the airport. Also, in her model congestion did not directly affect passengers or airlines but was an external cost to be minimized by the airport authority and capacity was fixed.

the models we revised are about congestion pricing and not peak-load pricing, in the sense that, even if there is more than one period, the demands between periods are not interdependent. Hence, the only way to fight excess usage is to dampen the demand. When periods are interdependent, pricing can be used not only to dampen the demand but to redistribute consumers through the periods, 'flattening' the demand curve. This is the case of peak-load pricing. Brueckner (2002) allows for endogenous sorting to periods, but this is done mainly through airlines decisions. If the airlines use peak-load pricing, then that would deliver a different demand to the airport, which would probably also have peak and off-peak periods. The airport would then have an incentive to choose prices for both periods, using peak-load pricing as well, in order to maximize its objective function, which would be dependent on the type of ownership and regulation. Hence, we would be in a situation of *sequential* peak-load pricing, which is quite particular the case of airports and airlines.

We have highlighted some of the issues that we think should be examined in the future, but perhaps one of the most important aspects of future research has to do with actual policies. It is seldom true that airports are priced as in a system, and it is seldom true that airport managers have access to all the information that they would need to do what is best. Hence, how should public airports be priced when they are not in a system, and when information is incomplete? And given this, what are the costs and gains of privatization? And what would be a good and feasible regulation mechanism?

Table 2.1: Summary of papers using the <i>traditional approach</i> (from 1987 on	I)

Author	Goal of the Paper	Objective Functions	Capacity	Delay	Observations
Morrison (1987)	Uncover the importance regulators give to each type of aircraft when they max SW	Max SW st BC	Fixed	NHDO	Many periods with independent demands
Morrison and Winston (1989)	Efficient pricing and capacity with congestion	Max SW	Variable and continuous	HDO	Many periods with independent demands
Oum and Zhang (1990)	Analyze budget adequacy under congestion pricing when capacity investments are lumpy	Max SW	Variable and lumpy	NHDO	Many periods with dependent demands
Zhang and Zhang (1997)	Effects of concessions. Should the BC be common to both concessions and airside activities or separate?	Max SW st BC	Variable and continuous	NHDO	Many periods, independent demands. First model to formally incorporating concessions
Zhang and Zhang (2001)	Analyze whether public airport should have a strict brake-even constraint (short run) or a longer run constraint	Max SW st BC	Lumpy	NHDO	Many periods, independent demands
Carlsson (2003)	Efficient pricing and capacity with congestion and emissions	Max SW	Variable and continuous	NHDO	One period. Social cost for emissions added to SW
Zhang and Zhang (2003)	Analyze privatization and the effects of concessions on pricing and capacities	Max SW Max profits (private case) Max SW st BC	Variable and lumpy	NHDO	One period. They include concession operations. BC is in the long run.
Lu and Pagliari (2004)	Regulation and concessions: single-till versus dual-till cap	Max profits st two different forms of regulation	Fixed	No delays	Rather than having delays, they assumed that capacity is a restriction on feasible output: potential for excess demand
Oum, Zhang and. Zhang (2004)	Efficiency implications of alternative forms of regulation	Max SW st BC Max profits (private case) Max profits st four different forms of regulation	Variable and continuous.	NHDO	One period. They include concession operations. BC is in the long run
Czerny (2006)	Effects of concessions on aeronautical charges. Regulation: single-till versus dual-till cap	Max SW Max Profits st two different forms of regulation	Fixed but large: no excess demand	No delays	Both airside and concession charges determine the number of consumers.

SW: Social Welfare; BC: Budget constraint; NHDO: The delay function is Non-homogenous of degree one in the traffic to capacity ratio; HDO: The delay function is homogenous of degree one in the traffic to capacity ratio.

.

Table 2.2: Summ	nary of papers	using the vertica	al structure approach
-----------------	----------------	-------------------	-----------------------

Author	Goal of the Paper	Oligopoly model	Objective Function and airport modeling	Observations
Brueckner (2002)	Optimal tax (additional to airport charges) to account for congestion	N airlines in homogenous Cournot	Max SW=CS+ Φ No formal modeling of the airport, only a <i>regulator</i> .	There are peak and off peak periods (peak-load pricing). Sorting to periods is endogenous through airlines decisions. Only the peak is congested Congestion is a non-linear function of traffic and affects both airlines and passengers.
Pels and Verhoef (2004)	Optimal tax (additional to airport charges) to account for congestion and market power	Duopoly in homogenous Cournot	Two airports not formally modeled, only two regulators. Max SW=CS+Φ Also analyze Individual Max SW.	One period (congestion pricing) Delay is a linear function of traffic and affects both airlines and passengers.
Raffarin (2004)	Efficient congestion pricing	Differentiated duopoly competing in prices and frequencies	Max SW= CS+ Φ + π - congestion costs. Single airport.	One period (congestion pricing). Three stage game: airport pricing, frequencies, prices. Congestion does not affect airlines or demand. They are only an external social cost. Delay is a linear function of traffic.
Basso (2005)	Effects of ownership on prices and capacity	N airlines in differentiated Cournot	Two airports (roundtrips) Max SW= CS+ Φ + π Max airports' profits Max airport-airlines joint profits Max SW st BC Max individual airport profits	One period (congestion pricing). Congestion is a non-linear function of traffic and affects both airlines and passengers. Consumers are also affected by schedule delay cost.
Zhang and Zhang (2006)	Optimal pricing to account for congestion and market power when there are N airlines and capacity is variable	N airlines in homogenous Cournot	Max SW= CS+ Φ + π Max airports' profits Max SW st BC	One period (congestion pricing). Congestion is a non-linear function of traffic affecting only the passengers. The demand function is general.

SW: Social Welfare; CS: Consumers' surplus; Φ : Airlines profits (industry wide); π Airport profits; BC: Budget constraint

2.7 References

- Basso, L. J. (2005) Airport Ownership: Effects on Pricing and Capacity. Working Paper, Sauder School of Business, The University of British Columbia (http://ssrn.com/abstract=849584).
- Basso, L. J. (2006) On Input Markets Surplus and its relation to The Downstream Market Game. Working Paper, Sauder School of Business, The University of British Columbia.
- Basso, L. J. and Jara-Diaz, S. R. (2006) Distinguishing Multiproduct Economies of Scale from Economies of Density on a Fixed-Size Transport Network. *Networks and Spatial Economics*, 6, 149-162.
- Basso, L. J. and Zhang, A. M. (2006) Sequential Peak-Load Pricing in a Vertical Setting: the case of airports and airlines. *Working Paper, Sauder School of Business, The University of British Columbia.*
- Brueckner, J. K. (2002) Airport congestion when carriers have market power. American Economic Review, 92(5), 1357-1375.
- Brueckner, J. K. (2005) Internalization of airport congestion: A network analysis. International Journal of Industrial Organization, 23(7-8), 599-614.
- Carlin, A. and Park, R.E. (1970) Marginal Cost pricing of Airport Runway Capacity. American Economic Review, 60, 310-319
- Carlsson, F. (2003) Airport Marginal Cost Pricing: discussion and an application to Swedish Airports. *International Journal of Transport Economics*, 30, 283-303.
- Condie, S. (2000) Whither Airport Regulation? In Smith, H. L. and Bradshaw, W. (Eds.), *Privatization and Deregulation of Transport:* St. Martin's Press, Inc, New York 364-393.
- Czerny, A.I. (2006) Price-cap Regulation of Airports: single-till versus dual-till. *Journal of Regulatory Economics*, 30, 85-97.
- De Borger, B. and Van Dender, K. (2006) Prices, Capacities and Service Quality in a Congestible Bertrand Duopoly. *Journal of Urban Economics, forthcoming.*.
- Doganis, R. (1992) The Airport Business, Routledge, London.
- Hendricks, K., Piccione, M., and Tan, G. F. (1999) Equilibria in networks. *Econometrica*, 67(6), 1407-1434.
- Horonjeff, R. and McKelvey, F. X. (1983) Planning and Design of Airports. McGraw-Hill.

- Ishii, J., K. Van Dender and S. Jun (2005) "Air travel choices in multi-airport markets," Working Paper, Department of Economics, University of California at Irvine.
- Jacobsen, S. E. (1979) Equivalence of Input and Output Market Marshallian Surplus Measures. *American Economic Review*, 69(3), 423-428.
- Jara-Diaz, S. R. and Basso, L. J. (2003) Transport cost functions, network expansion and economies of scope. *Transportation Research Part E-Logistics and Transportation Review*, 39(4), 271-288.
- Lave, L. B. and De Salvo, J. S. (1968) Congestion, Tolls and the Economic Capacity of a Waterway. *Journal of Political Economy*, 76, 375-391.
- Levine, M.E. (1969) Landing fees and the airport congestion problem. *Journal of Law and Economics*, 12, 79-108.
- Lu, C. C. and Pagliari, R. I. (2004) Evaluating the potential impact of alternative airport pricing approaches on social welfare. *Transportation Research Part E-Logistics and Transportation Review*, 40(1), 1-17.
- Mohring, H. (1976) Transportation Economics. Ballinger, Cambridge MA.
- Morrison, S.A. (1983) Estimation of long-run prices and investment levels for airport runways. *Research in Transportation Economics*, 1, 103-130
- Morrison, S. A. (1987) The Equity and Efficiency of Runway Pricing. Journal of Public Economics, 34(1), 45-60.
- Morrison, S. A. and Winston, C. (1989) Enhancing the Performance of the Deregulated Air Transportation System. *Brookings Papers on Economic Activity*, 61-112.
- Oum, T. H., Zhang, A. M., and Zhang, Y. M. (1996) A note on optimal airport pricing in a hub-and-spoke system. *Transportation Research Part B-Methodological*, 30(1), 11-18.
- Oum, T. H. and Zhang, Y. M. (1990) Airport Pricing Congestion Tolls, Lumpy Investment, and Cost Recovery. *Journal of Public Economics*, 43(3), 353-374.
- Oum, T. H., Zhang, Y. M., and Zhang, A. M. (2004) Alternative forms of economic regulation and their efficiency implications for airports. *Journal of Transport Economics and Policy*, 38(2), 217-246.
- Pels, E., Nijkamp, P., and Rietveld, P. (1997) Substitution and complementarity in aviation: Airports vs airlines. *Transportation Research Part E-Logistics and Transportation Review*, 33(4), 275-286.
- Pels, E., Nijkamp, P., and Rietveld, P. (2000) A note on the optimality of airline networks. *Economics Letters*, 69(3), 429-434.

- Pels, E. and Verhoef, E. T. (2004) The economics of airport congestion pricing. *Journal of Urban Economics*, 55(2), 257-277.
- Quirmbach, H. C. (1984) Input Market Surplus the Case of Imperfect Competition. Economics Letters, 16(3-4), 357-362.
- Raffarin, M. (2004) Congestion in European airspace A pricing solution? Journal of Transport Economics and Policy, 38, 109-125.
- Ross, T. W. (1984) Uncovering Regulators Social-Welfare Weights. Rand Journal of Economics, 15(1), 152-155.
- Spence, A. M. (1975) Monopoly, Quality, and Regulation. *Bell Journal of Economics*, 6(2), 417-429.
- Starkie, D. (2001) Reforming UK airport regulation. Journal of Transport Economics and Policy, 35, 119-135.
- US Federal Aviation Administration (1969) Airport Capacity handbook. Government Printing Office, Washington DC.
- Zhang, A. M. and Zhang, Y. M. (1997) Concession revenue and optimal airport pricing. Transportation Research Part E-Logistics and Transportation Review, 33(4), 287-296.
- Zhang, A. M. and Zhang, Y. M. (2001) Airport Charges, Economic Growth and Cost Recovery. *Transportation Research Part E*, 37, 25-33.
- Zhang, A. M. and Zhang, Y. M. (2003) Airport Charges and Capacity Expansion: effects of concessions and privatization. *Journal of Urban Economics*, 53, 54-75
- Zhang, A. and Y. Zhang (2006) Airport capacity and congestion when carriers have market power. *Journal of Urban Economics, forthcoming.*

3 ON INPUT MARKETS SURPLUS AND ITS RELATION TO THE DOWNSTREAM MARKET GAME¹

3.1 Introduction

Evaluating the welfare effects on buyers and sellers of price changes is a familiar task for policy analysts in a variety of circumstances. When the market in question is one for a final good, the analysis is generally straightforward; indeed, it is a familiar problem put to undergraduate economics students to teach them how to measure consumer and producer surplus. When the prices changing are that for inputs which are subsequently used by its downstream purchasers to produce their own outputs, a full analysis of the welfare effects of price changes must take into account the effects on downstream consumers. A common example involves attempts to measure the harm caused by price fixing. When the price-fixed goods are purchased by final consumers, the impact of the higher prices is typically measured by changes in the consumers' surplus under the products demands curve. When the price-fixed goods are inputs, however, the area under the input demand curves will be determined by the benefits received by the direct inputs purchasers and their final consumers. Just how "consumers' surplus" measures coming from areas under the input demand curves capture the effects of direct purchasers and downstream final consumers then becomes a critical question.

The relation between the measures of input and final markets surplus has been addressed by Schmalensee (1971, 1976), Wisecarver (1974), Anderson (1976), Carlton (1979), Jacobsen (1979) and Quirmbach (1984). The latest result indicates that the input markets surplus is equal to downstream industry profits, plus a fraction –which depends on the degree of downstream competition– of consumer surplus. Yet, this result hinges on a number of simplifying assumptions which we would like to relax. In this paper, it is argued that this is possible by linking the input markets surplus literature to results from a second, seemingly unrelated, literature.

This second literature involves the search for a function, the maximization of which generates the same choices as would be obtained in the play of the standard non-cooperative game (Samuelson, 1947; Spence, 1976; Bergstrom and Varian, 1985; Slade, 1994; and Monderer and Shapley, 1996). This class of functions succinctly, and more simply, describes the equilibria of the game by transforming a fixed point problem into a maximization problem. The connection between the two streams is as follows: I show that the input markets surplus measure, obtained by integration of the input demands derived from the equilibrium of a downstream oligopoly game, is equal to the change of a multivariate function which critical points coincide with the equilibria of the downstream game. In particular, for the case of *potential* oligopoly games, the input market surplus is shown to be equal to the change in the exact potential function. A characteristic of a potential function is that its Jacobian coincides with the system of first order conditions of

¹ A version of this chapter has been submitted for publication. Basso, L.J. (2006) On Input Markets Surplus and its Relation to the Downstream Market Game.

the game, which implies nice stability properties of equilibria. The relation between input market surplus and potential functions is helpful because finding out if a game is potential or not is simple, and potential functions have a well-defined way to be computed.

The results regarding the relationship between input and final markets surplus measures available up to now in the literature, are particular cases of the Proposition proved in this paper. Therefore, the Proposition synthesizes and generalizes them, relaxing the strong simplifying assumptions about the nature of the downstream competition that they needed, and further providing the analyst with a tool to understand what does the input market surplus captures in a wide variety of situations.

The plan of the paper is as follows. In Section 3.2 the downstream oligopoly game is introduced. I use this game and its notation to briefly review both streams of the literature. I then discuss their connection in Section 3.3, proving the main Proposition and presenting some examples of cases that could not have been handled with the current literature. In Section 3.4 I discuss the particular, yet interesting, case of potential games and potential functions. Formal definitions and properties are provided here. Section 3.5 concludes.

3.2 The Downstream Game and the Literature

The game

I start by defining the game I look into in this paper. This game is also used next to present both streams of the literature. Let N be the set of n firms competing in a final market, $N = \{1, .., n\}$. A strategy for firm *i* will be the choice of a level of production q_i from some compact subset A^i of the real line, \Re . I denote **q** the vector of choices $(q_i, ..., q_n)$, define $Q = \sum_{i=1}^{n} q_i$ and let A be the strategy space, $A = \times_{i \in N} A^i$. For $S \subseteq N$, -S denotes the complementary set of S , and A^S denotes the Cartesian product $\times_{i \in S} A^i$. For singleton sets $\{i\}, A^{-\{i\}}$ is denoted A^{-i} so that $\mathbf{q}_{-i} \in A^{-i}$ is $\mathbf{q}_{-i} = (q_1, \dots, q_{i-1}, q_{i+1}, \dots, q_n)$. As is usual, with a slight abuse of notation I let $\mathbf{q} = (q_i, \mathbf{q}_{-i})$. Let w be the vector of prices of m inputs used by all firms, where $w_z \in \mathfrak{R}^+$, with $z \in Z = \{1, ..., m\}$ and $\mathbf{w} \in \overline{\mathfrak{R}}^+ \equiv \times_{z \in Z} \mathfrak{R}^+$. How these input prices are determined is not important. What matters, is that downstream firms take these prices as given. Firm *i*'s profit, $\pi^i : A \times \overline{\mathfrak{R}}^+ \to \mathfrak{R}$, is given by $\pi^i(\mathbf{q}, \mathbf{w}) = R^i(\mathbf{q}) - C^i(\mathbf{q}, \mathbf{w})$, where R^i are revenues and C^i are costs. I assume π^i to be twice continuously differentiable. The final market game is then given by $\Gamma < N, \{A^i\}, \{\pi^i\} > .$ As for the inputs market, let $x_{iz}(\mathbf{q}, \mathbf{w})$ be firm *i*'s conditional demand for input $z \in \mathbb{Z}$, and $X_z = \sum_{i=1}^n x_{iz}$ input z's total demand. $\mathbf{q}^*(\mathbf{w})$ denotes an interior Nash-equilibrium of the downstream game as function of w.

Equivalence of input and output markets Marshallian surpluses: the literature

As explained in the introduction, the question this literature has attempted to answer is, "given a change in the vector of unit factor prices, to what is the resulting change in Marshallian surplus (in the factor markets) equal?" (Jacobsen, 1979, p.423). Formally, the idea is to try to unveil what is exactly being calculated when one computes $\int_{w_0}^{w_1} \sum_{z \in \mathbb{Z}} X_z(\mathbf{w}) dw_z$. This issue was addressed by Schmalensee (1971, 1976), Wisecarver (1974), Anderson (1976) and Carlton (1979) under different sets of simplifying assumptions such as, perfect competition downstream, change in only one factor price and/or linear homogenous output production. Jacobsen (1979), in an elegant paper, obtained more general results. He showed that, regardless of how many factor prices are changed, with a monopoly downstream, $\int_{w_0}^{w_1} \sum_z X_z(\mathbf{w}) dw_z$ will be exactly equal to the change in monopoly profits downstream that occurs because of the factor-price vector markets is precisely equal to the change of final consumers' surplus plus the change in downstream producers' profits.

An obviously unanswered question was how this extended to imperfect competition cases. Quimbarch (1984) looked into this. He considered a simplified version of the game described previously: firms had identical cost functions dependent only on own production, q_i , used only one input, and were homogenous in that they faced an inverse demand function given by P(Q). Letting $\beta_i = \beta = (q_i/Q)(dQ/dq_i)$ be the firms' conjectures about the elasticity of total production before a change in own production, and using the symmetry of the game, Quirmbach showed that:

$$\int_{w_0}^{w_1} X(w) dw = -\left| \sum_i \pi^i + (1 - \beta) CS \right|_{w_0}^{w_1}$$
(3.1)

where CS is Marshallian consumer surplus (in the final market). Since $\beta=0$ corresponds to perfect (Bertrand) competition, $\beta=1/n$ to Cournot competition and $\beta=1$ to monopoly, Quirmbach's result showed that the area under the demand curve for the input, corresponds to the sum of industry wide profits of downstream producers plus a fraction –dependent on the degree of competition– of consumer surplus. This generalized Jacobsen's results to a homogenous and symmetric oligopoly, but with the qualification that Quirmbach considered only one factor price.

Transformation of a game into a maximization problem: the literature

The second stream of literature, which I argue can be linked to the first, asks whether there exists a function that firms in an oligopoly collectively, yet non-cooperatively, maximize, and what would be its characteristics. The main references are Samuelson (1947), Spence (1976), Bergrstrom and Varian (1985), Slade (1994) and Monderer and Shapley (1996); I will omit the factor prices here since inputs were not an issue in these papers. Samuelson

asked, under what conditions the first order conditions of a game can be regarded as the solution of an extremum problem so that the equilibrium loci correspond to the vanishing of the partial derivatives of some function. In terms of Γ , he was looking for $H: A \to \Re$ such that $\partial H(\mathbf{q}^*)/\partial q_i = \partial \pi^i(\mathbf{q}^*)/\partial q_i = 0$, $\forall i \in N$. I shall call this type of functions Samuelson functions hereafter. Spence (1976) was interested in studying the process of product choice in a monopolistic environment. He argued that, by looking at the function that is implicitly maximized by the monopolistic competitors –if one exists– and by comparing it with the total surplus function, biases in product choice under the market mechanism could be found. Spence found that a broad class of utility functions led to a system of differentiated inverse demands and to profit functions that can be written as $\pi^i(\mathbf{q})/\alpha_i = \tilde{F}(\mathbf{q}) + \Theta^i(\mathbf{q}_{-i})$, where α_i is a firm-specific constant. Hence, when firm *i* maximizes π^i with respect to q^i , it maximizes $\tilde{F}(\mathbf{q})$ with respect to q^i , making $\tilde{F}(\mathbf{q})$ the function sought. Spence called \tilde{F} the wrong surplus function.

Bergstrom and Varian (1985) wanted to address the same question but for a standard (homogenous and symmetric) Cournot: what does a Cournot equilibrium maximize? They found that the following function of total output, $\tilde{W}(Q) = \sum_{i \in N} \pi^i(Q) + ((1-n)/n)CS(Q)$, had a first-order condition with respect to Q which led to the same first-order condition of the Cournot game *after imposition of symmetry*. They concluded that the industry output in a symmetric Cournot equilibrium maximizes \tilde{W} . Slade (1994) considered more general oligopoly games. She asked when it is true that individuals firms pursing selfish objectives act as if a single agent was maximizing a well defined *fictitious-objective function*. She found necessary and sufficient conditions under which a function $F: A \to \Re$, such that $\partial F(\mathbf{q})/\partial q_i = \partial \pi^i(\mathbf{q})/\partial q_i$, $\forall i \in N$, exists. Slade's fictitious-objective function, which was later defined to be an *exact potential function* by Monderer and Shapley (1996), has nice properties that facilitate solving and characterizing the game. Among others, local maxima of F are a subset of the local-Nash equilibria of the game, while Nash equilibria that are not local maxima of F are unstable. I further discuss exact potential functions in Section 3.4.

3.3 Input Market Surplus and the Downstream Game

So, how are the two issues connected? The more general result available in the literature regarding the relation between input and output markets surplus measures is Quirmbach (1984). Yet, it hinges on a number of assumptions which we would like to relax. I argue that this is possible by relating the input markets surplus measure to one of the functions that describe the equilibrium of the downstream game, and that Quirmbach and Jacobsen's result are particular cases of this stronger theoretical relation. That such relation exists is hinted by the following simple observation. In Quirmbach's setting, when the downstream game is Cournot, i.e. $\beta = 1/n$, (3.1) becomes $\int_{w_0}^{w_1} X(w)dw = -\left|\sum_i \pi^i + ((1-n)/n)CS\right|_{w_0}^{w_1}$. But Bergstrom and Varian (1985) showed that if a social planner maximizes a pseudo-social

welfare function given by $\widetilde{W} = \sum_{i} \pi^{i} + ((1-n)/n)CS$, then the Cournot outcome is obtained. Hence, it is true that $\int_{v_0}^{v_1} X(w) dw = -\left|\widetilde{W}\right|_{w_0}^{w_1}$.

If a more general relation is to be established, the first issue that needs to be addressed is which function describing the equilibria of the downstream game we should use. As explained in the previous Section, many authors have addressed the question of finding and characterizing the equilibria of an oligopoly game through a single function. Yet, the ways they have answered it differ. Samuelson's analysis was done under the heading of convertibility into a maximum problem, but, as can be seen, his question was actually related to the more general problem of existence of a function whose critical values are equilibria of the game. Bergstrom and Varian were close to Samuelson in that they were interested in the equilibrium point, but they derived their function with respect to the total output and not with respect to each firm's output, which is why they had to compare against the first-order conditions of the game after imposing symmetry. They did however check the second-order condition to ensure that the critical Q value was in fact a maximum. Spence on the other hand was indeed interested in the maximization of a multivariate function and he obtained something stronger that the Samuelson condition: from what he found, it flows that $\alpha_i \partial \tilde{F}(\mathbf{q}) / \partial q_i = \partial \pi^i(\mathbf{q}) / \partial q_i$, a condition which implies Samuelson's but which holds for points other than the equilibria. Slade explicitly acknowledged that the fictitious-objective function (or exact potential) she was looking for was different from the functions Samuelson was looking for, because she was concerned with stability, a notion that deals with behavior both on and off equilibrium points. For the case of monopolistic competition, her functions would be a subset of Spence's, since she requires α_i to be 1 for all i. In fact, exact potential -or fictitious objective- functions (Slade, 1994; Monderer and Shapley. 1996) are a subset of Spence's wrong surplus functions, which in turn are a subset of Samuelson functions. Because of this, it is better to establish the relation between input markets and final market surplus measures using the Samuelson functions.

Proposition 3.1: Consider the game described in Section 3.2. If in the producers' game there exists a differentiable function $H(\mathbf{q}, \mathbf{w}): A \times \overline{\mathfrak{R}}^+ \to \mathfrak{R}$ such that:

$$\partial H(\mathbf{q}^*, \mathbf{w}) / \partial q_i = \partial \pi^i (\mathbf{q}^*, \mathbf{w}) / \partial q_i, \qquad i \in N$$
 (3.2)

$$\partial H(\mathbf{q}^*, \mathbf{w}) / \partial w_z = -\sum_i \partial C^i(\mathbf{q}^*, \mathbf{w}) / \partial w_z, \quad z \in \mathbb{Z}$$
 (3.3)

then

$$\int_{\mathbf{w}_0}^{\mathbf{w}_1} \sum_{z \in Z} X_z(\mathbf{w}) dw_z = - \left| H(\mathbf{q}^*(\mathbf{w}), \mathbf{w}) \right|_{\mathbf{w}_0}^{\mathbf{w}_1}$$
(3.4)

<u>*Proof*</u>: Take the total derivative of $H(\mathbf{q}^*(\mathbf{w}), \mathbf{w})$ with respect to w_z . Since $\mathbf{q}^*(\mathbf{w})$ is a Nash equilibrium, it is true that $\partial \pi^i(\mathbf{q}^*)/\partial q_i = 0$. Use conditions (3.2) and (3.3) to get

 $dH(\mathbf{q}^*(\mathbf{w}), \mathbf{w})/dw_z = -\sum_i \partial C^i / \partial w_z$. Then use Shepard's Lemma, that is $\partial C^i / \partial w_z = -x_{iz}$, sum over z and integrate between \mathbf{w}_0 and \mathbf{w}_1 .

Proposition 3.1 shows that input markets surplus correspond to the change of a Samuelson function of the downstream game; that we are considering a Samuelson function is apparent from condition (3.2). Although it may be tempting to say that the input markets surplus corresponds to what is maximized in the downstream game, this is not true; condition (3.2) pertains to critical values and not to extreme values necessarily. This would be the case only if we narrow the class of functions we use, as I discuss in the next Section. Condition (3.3) may appear to be stringent but it is not really; it will be apparent that it can be always made to hold when $C^i(\mathbf{q}, \mathbf{w}) \equiv C^i(q_i, \mathbf{w})$.

Proposition 3.1 allows us to find the relation between input markets surplus and the surpluses of downstream agents for a large variety of downstream games. For example, it is direct to verify that Bergstrom and Varian's function is a Samuelson function of the Cournot game² and that condition (3.3) holds, which explains its link to Quirmbach's result. However, this result can be expanded to other conjectures and change in more than one input price: consider symmetric and homogenous firms as before and the following candidate function: $H_1(\mathbf{q}, \mathbf{w}) = \sum_i \pi^i(\mathbf{q}, \mathbf{w}) + ((n-1-\nu)/n)CS(Q)$, where $\nu = \partial \sum_{j \neq i} q_j / \partial q_i$,

 $\forall i \in N$, is firms' conjecture about the change in total rivals' output when they individually change their own, $\pi^i(\mathbf{q}, \mathbf{w}) = q_i P(Q) - C(q_i, \mathbf{w})$ and $CS(Q) = \int_0^Q P(Z) dZ - QP(Q)$. Differentiating $H_1(\mathbf{q}, \mathbf{w})$ with respect to q_i , we obtain (see appendix A.1): $\partial H_1 / \partial q_i = P(Q) + ((1+\nu)/n)QP'(Q) - C_q(q_i, \mathbf{w})$. This is identical to the first order conditions of the producers' game after imposing symmetry, which is true in equilibrium. This shows that H_1 is a Samuelson function and hence condition (3.2) is satisfied. To prove that H_1 is what is captured by integration of the input demands, we only need to verify that condition (3.3) holds, which is direct.³ Thus, if we consider only one input, we recover Quirmbach's result. If we let ν be equal to -1 (perfect competition) or ν be equal to 0 and nequal to 1 (monopoly), we recover Jacobsen's results. Furthermore, we can see that under collusion, i.e. $\nu = n - 1$, it is also true that only downstream profits are captured (collusion and monopoly profits will be equal only if marginal costs are constant). Hence, as argued, the two most important results in the input markets surplus literature are particular cases of Proposition 3.1.

² Hence, although Bergstrom and Varian took the derivative only with respect to Q, they could have done it as Samuelson and in condition (3.2), that is, with respect to each firm's output. Bergstrom and Varian's function has an extra feature though: the partial derivative of their function with respect to q_i is equal to the partial derivative of the profit function of firm *i* at any symmetric point (i.e., when all *q*'s are equal). In particular, they are equal for the Nash equilibrium –the Samuelson condition–, which is symmetric because the game is symmetric.

³ Note that this show that Proposition 3.1 is not reduced to quantity competition only, as the Bertrand case is captured here by v=-1.

The advantage of Proposition 3.1 though, is that Samuelson functions exist for many more cases than the previous simple ones, so we can depart from these settings and examine how results change. Consider for example homogenous Cournot competition downstream, but with externalities in production. This would be the case, for example, of firms using congestible upstream facilities in order to produce a final good (e.g. airlines using airports to provide travel or telephone companies using upstream networks to provide phone calls). Externality in production, caused by congestion, may be modeled as $C(q_i, Q, w)$; I consider the single input case for simplicity. The symmetric Nash equilibrium is given by $\pi_i^i(\mathbf{q}^*, w) = P(Q) + (Q/n)P'(Q) - C_q(Q/n, Q, w) - C_Q(Q/n, Q, w) = 0$. Next, consider the function

$$H_{2}(\mathbf{q},w) = \sum_{i} \pi^{i}(\mathbf{q},w) + \frac{n-1}{n} CS(Q) + (n-1) \int_{0}^{Q} C_{Q}\left(\frac{Z}{n}, Z, w\right) dZ$$
(3.5)

It is easy to check that H_2 fulfils (3.2), i.e. $\partial H_2(\mathbf{q}^*, w)/\partial q_i = \pi_i^i(\mathbf{q}^*, w)$ -see appendix A.2so, to prove that $\int_{w_0}^{w_1} X(w) dw = -|H_2(\mathbf{q}^*, w)|_{w_0}^{w_1}$, we need to verify that (3.3) holds, which is the case if $C_{Qw} = 0$.⁴ Here, then, the integral of the demand for the input is not equal to producers' profits plus a fraction of consumer surplus. There is a third term, which captures uninternalized congestion; and as *n* grows, $\int_{w_0}^{w_1} X(w) dw$ does not approach industry profits plus consumer surplus as before. This analysis easily extends to conjectures other than Cournot.

3.4 Potential Function and Input Market Surplus

As explained, one of the advantages of Proposition 3.1 is that, for any given downstream game, there are probably many Samuelson functions fulfilling condition (3.3). As stated by Slade (1994), it is almost always true that a function which first order condition have the same zeroes as the first order conditions of the game, exists (it is enough to square and sum the first order conditions of the game). This, however, may be also an embarrassment of riches; without further structure, it may be hard to obtain economic intuition from such functions, as pointed out by Slade and Samuelson himself. A way to circumvent this is to look at some more economically meaningful subset of the Samuelson functions. This is the case of Slade's fictitious-objective functions, are in fact the same in the class of infinite

⁴ A cost function featuring this characteristic is used, for example, by Pels and Verhoef (2004) in their analysis of airport markets. They consider $C(q_i, Q, w) = q_i(c + w + \alpha D(Q))$, where c and α are constants and D(Q) is the *delay* function.

differentiable games, as summarized in Lemma 3.1, and they have many nice properties, as Lemma 3.2 shows:⁵

Lemma 3.1: (Slade, Monderer and Shapley) Consider the game $\Gamma < N, \{A^i\}, \{\pi^i\} > \text{ and a}$ function $F(\mathbf{q}, \mathbf{w}) : A \times \overline{\mathfrak{R}}^+ \to \mathfrak{R}$. Then, the following statements are equivalent:⁶

1. Γ is a potential game and F is a potential function, i.e. $\forall q_i, r_i \in A^i$, $\forall \mathbf{q}_{-i} \in A^{-i}$,

$$F(q_i, \mathbf{q}_{-i}, \mathbf{w}) - F(r_i, \mathbf{q}_{-i}, \mathbf{w}) = \pi(q_i, \mathbf{q}_{-i}, \mathbf{w}) - \pi(r_i, \mathbf{q}_{-i}, \mathbf{w})$$

- 2. $\partial F(\mathbf{q}, \mathbf{w}) / \partial q_i = \partial \pi^i(\mathbf{q}, \mathbf{w}) / \partial q_i$, $\forall i \in N$
- 3. There exist functions $\Theta^{i}(\mathbf{q}_{-i}, \mathbf{w})$ such that $\pi^{i}(\mathbf{q}, \mathbf{w}) = F(\mathbf{q}, \mathbf{w}) + \Theta^{i}(\mathbf{q}_{-i}, \mathbf{w})$.

Lemma 3.2: (Slade, Monderer and Shapley)

1.
$$\Gamma$$
 has a potential function if and only if $\frac{\partial^2 \pi^i(\mathbf{q}, \mathbf{w})}{\partial q_i \partial q_j} = \frac{\partial^2 \pi^j(\mathbf{q}, \mathbf{w})}{\partial q_i \partial q_j}, \forall i, j \in N$

(Monderer and Shapley).

- 2. The local maxima of F are a subset of the local Nash-Equilibria of Γ (Slade).
- 3. Let q^* be an interior local pure-strategy Nash Equilibrium of Γ . Then q^* is (not) an interior local maximum of F if and only if q^* is locally asymptotically (un)stable in a myopic learning sense (Slade).
- 4. The potential function is defined up to an additive constant (Monderer and Shapley).
- 5. A potential function for a potential game can be obtained as: $\Sigma(\pi \pi) = \sum_{i=1}^{n} \frac{\partial \pi^{i}}{\partial x^{i}} (\pi(x) - \chi(x)) (x) = 0$

$$F(\mathbf{q}, \mathbf{w}) = \sum_{i \in N} \int_{0}^{1} \frac{\partial q^{i}}{\partial q^{i}} (\mathbf{x}(t), \mathbf{w}) (x^{i})'(t) dt, \text{ where } \mathbf{x} : [0,1] \to A \text{ is a piecewise}$$

continuously differentiable path in A that connects an arbitrary but fixed strategy profile in A to q (Monderer and Shapley).

From Lemma 3.1.2, it is apparent that potential functions fulfill condition (3.2) in Proposition 3.1. Their requirement is stronger though, since the equality between first order conditions have to occur not only at equilibria points but also off the equilibrium path. The flip side is that potential functions are economically more meaningful, as Lemma 3.2.2 and 3.2.3 show; the potential function can unmistakably be considered as what the oligopoly maximize and in fact, it helps refine the set of Nash equilibria. Next, while Lemma 3.2.1 provides an easy way to test whether a game is potential or not, Lemmas 3.2.4 and 3.2.5 show two other important features. The former indicates that if we use a potential function instead of a Samuelson function –and condition (3.3) in Proposition 3.1 holds–, the input market surplus measure is univocally defined by a single expression; in equation (3.4), the

⁵ Monderer and Shapley discussed several notions of potentials functions for games in strategic form for both finite and infinite games, and for both differentiable and non-differentiable games. I present here only the results that pertain to the class that interest me, infinite differentiable games. I included the vector of factor prices in the notation because it is central for this paper, but their presence is immaterial at this point.

⁶ The equivalence between (3.1.1) and (3.1.2) is in Monderer and Shapley (1996), the equivalence between (3.1.2) and (3.1.3) in Slade (1994).

input market surplus measure is equal to the change of the potential function and therefore constants do not matter.⁷ Lemma 3.2.5 provides us with a way to actually compute the potential function, which significantly strengths the applicability of the Proposition.

The trade-off is clear, then. Potential functions are more meaningful economically, but exist in fewer cases than Samuelson functions. For example, in homogenous Cournot, a potential function exists if and only if the demand is linear (see Lemma 3.2.1). However, if products are differentiated, things greatly improve. If we consider Spence's example, the following class of demands are sufficient (yet not necessary) for a potential function to exist:

$$p^{j}(\mathbf{q}) = (\partial h(\mathbf{q}) / \partial q_{j})(1/\alpha_{j})$$
, where $h(\mathbf{q}) = \sum_{k=1}^{2} H_{k}(\prod_{i \in \Gamma_{k}} q_{i}^{\alpha_{i}})$; Γ_{k} are the 2ⁿ possible subsets

of the first *n* positive integers, H_k are 2^n scalar functions and $\alpha_1,..., \alpha_n$ are *n* positive numbers. The potential function is then⁸ (see appendix A.3 for the derivation): $F(\mathbf{q}, \mathbf{w}) = \sum_{k=1}^{2^n} \left[H_k (\prod_{i \in \Gamma_k} q_i^{\alpha_i}) \prod_{i \in \Gamma_k} q_i^{\alpha_i} \right] - \sum_j C^j (q_j, \mathbf{w})$. As noted by Spence, $h(\mathbf{q})$ contains 2^n arbitrary functions H_k and *n* arbitrary weights α_i , so it provides great flexibility and therefore, potential functions exist for a broad class of demand systems. Since condition (3.3) obviously holds, we can use Proposition 3.1 to obtain an expression for $\int_{\mathbf{w}_0}^{\mathbf{w}_1} \sum_{z} X_z(\mathbf{w}) dw_z$ whenever the downstream game belongs to this broad class of games. For example, the demand systems defined by $p^j(\mathbf{q})$ include linear demands. For n=2, $p^i(\mathbf{q}) = a - bq_i - eq_j$ is obtained by considering $\Gamma_1 = \{1\}, \Gamma_2 = \{2\}, \Gamma_3 = \{1,2\}, \Gamma_4 = \emptyset$; $\alpha_1 = \alpha_2 = 1$; $H_1(x) = H_2(x) = ax - (1/2)bx^2$ and $H_3(x) = -ex$. For the *n* products case, $p^i(\mathbf{q}) = a - b \cdot q_i - \sum_{j \neq i}^n e \cdot q_j$, functions and weights are analogous but less simple to list. The resulting potential function however is not complex: it is given by $F_1(\mathbf{q}, \mathbf{w}) = \sum_{i=1}^n [(a - bq_i)q_i - C^j(q_i, \mathbf{w})] - e\sum_{k=1}^{n-1} \sum_{i=k+1}^n q_k q_j$. Now, ideally, we would like to

relate F_1 to downstream profits and consumer surplus; it happens that finding the relation is easy. First, straightforward algebra allow us to get (see appendix A.4):

⁸ Slade stated that $p^{i}(\mathbf{q}) = \partial h(\mathbf{q})/\partial q_{j}$ would lead to the existence of an exact potential function. In fact, that leads to the existence of a function \tilde{F} such that $\alpha_{i}\partial\tilde{F}(\mathbf{q})/\partial q_{i} = \partial \pi^{i}(\mathbf{q})/\partial q_{i}$, which is different than the statement in Lemma 3.1.2. \tilde{F} would be what Monderer and Shapley defined as a *weighted* potential function. This function has useful properties in finite games but is not really helpful in infinite games. For example, the nice stability property of local Nash equilibria that are also local maxima of F (Lemma 3.2.3) does not hold for \tilde{F} . This is so because the proof hinges on the fact that the Hessian of F and the Jacobian of the system of first order conditions of the game are equal, making their characteristics roots equal. With \tilde{F} , this is no longer true. It is still obviously true though, that \tilde{F} is a Samuelson function.

 $^{^{7}}$ This is not to say that if we use Samuelson functions, the input market surplus measure would have different values depending on the function used. The expressions in (3.4) may differ, but when evaluated they would lead to the same number.

$$F_{1}(\mathbf{q}, \mathbf{w}) = \sum_{j=1}^{n} \pi^{j}(\mathbf{q}, \mathbf{w}) + e \sum_{k=2}^{n} \sum_{j=1}^{k-1} q_{k} q_{j}$$
(3.6)

From (3.6) it can be seen that it is still true that, as the degree of competition decreases, that is, as substitutability diminishes (*e* goes to zero), the integral of the input demand captures only producers profits. Next, consumer surplus is given by the line integral $\int_{0}^{a} \sum_{i} q_{i}(\mathbf{p}) dp_{i}$, which is straightforward to compute since direct demands exist (as long as b > e > 0) and the solution is path independent (see appendix A.5). A linear integration path leads to $CS = (b/2)\sum_{k} q_{k}^{2} + e\sum_{k=2}^{n} \sum_{j=1}^{k-1} q_{k}q_{j}$, from where we get:

$$\int_{\mathbf{w}_{0}}^{\mathbf{w}_{1}} \sum_{z} X_{z}(\mathbf{w}) dw_{z} = -\left|F_{1}(\mathbf{q}^{*}, \mathbf{w})\right|_{\mathbf{w}_{0}}^{\mathbf{w}_{1}}$$

$$= -\left|\sum_{j=1}^{n} \pi^{j}(\mathbf{q}^{*}, \mathbf{w}) + CS(\mathbf{q}^{*}, \mathbf{w}) - (b/2)\sum_{k=1}^{n} (q_{k}^{*})^{2}\right|_{\mathbf{w}_{0}}^{\mathbf{w}_{1}}$$
(3.7)

Hence, Proposition 3.1 allowed us to obtain yet another extension to Quimbarch's result. Equation (3.7) is valid when the downstream game is asymmetric Cournot, when demands are differentiated (linear) and when many input prices change. Here, we no longer have a fraction of consumer surplus; now, $\int_{w_0}^{w_1} \sum_z X_z(\mathbf{w}) dw_z$ fails to exactly capture profits and consumer surplus by an amount that is a function of the squares of firms' production in equilibrium, which are not (necessarily) equal.

3.5 Conclusions

In this paper, I have proved a Proposition that states that the input markets surplus measure is equal to the change of a single multivariate function, which main characteristic is that it can be used to synthetically describe the equilibria of the downstream game. This Proposition generalizes the two more important results of the literature on input markets surplus and its relation to the output market (i.e. Jacobsen, 1979, and Quirmbach, 1984). Furthermore, it provides more significance to a class of functions, which I called Samuelson functions, which were said to be 'mere technical devices'. Yet, in my opinion, the more important aspect of the Proposition is that for downstream potential games, the input market surplus measure was shown to be equal to the change of the potential function. This is significant because, today, potential games are well-understood: we have simple ways to test whether a game is potential or not; when it is, we can easily learn a great deal about its equilibria points; and we have a well-defined way to compute the potential function (up to an additive constant, which in this case is irrelevant). Hence, whenever the downstream game is potential, the problem of calculating the input markets surplus is, in theory, completely solved. The applicability of Proposition 3.1 for the analysis of input markets is important. It allows us to assess how much information about the downstream market is carried in the input demand functions or, in other words, how far-off the line integral of input demands is from exactly capturing producers' profits and consumer surplus. For example, it should be useful to adequately estimate damages on price-fixing cases on input markets (see Brander and Ross, 2005, for a discussion of this issue). Or, to estimate the welfare effects of input prices changes –which was Carlton (1979)'s motivation. What we would like to know is when it is legitimate to perform cost-benefit analysis *considering only the input markets*, since input demands are the ones that are derived from equilibrium in a related market. It will never be incorrect to use the downstream information; by adding to the upstream firms' profits, the downstream profits and consumer surplus, one always get a proper social welfare function. But whether this is the case if one adds upstream firms' profits and input markets surplus, is not as simple. And here is where Proposition 3.1 comes in handy, as it helps us find out how far the input market surplus is from capturing the sum of final consumer surplus and downstream producer profits.

The advantage of Proposition 3.1 over the particular cases previously proved in the literature is its flexibility and generality, which I showed with some examples. The Proposition should be helpful to better analyze intermediate markets in a wide range of situations, such as non-symmetric firms or differentiated demands.

3.6 References

- Anderson, J. (1976) The Social Cost of input Distortions: a comment and a generalization. American Economic Review, 66, 235-238.
- Bergstrom, T. and Varian, H. (1985) Two remarks on Cournot equilibria. *Economic Letters*, 19, 5-8.
- Brander, J. and Ross, T. (2006) Estimating Damages from Price-Fixing, Canadian Class Action Review, 3, 335-369
- Carlton, DW. (1979) Valuing market benefits and Costs in related Output and Input Markets. *American Economic Review*, 69, 688-696.
- Jacobsen, S.E. (1979) On the equivalence of input and output market Marshallian surplus measures. *American Economic Review*, 69, 423-428.
- Monderer, D. and Shapley, L.S. (1996) Potential Games. Games and Economic Behavior, 14, 124-143.
- Pels, E. and Verhoef, T. (2004) The economics of airport congestion pricing. Journal of Urban Economics, 55, 257-277.
- Quirmbach, H.C. (1984) Input market surplus: the case of imperfect competition. *Economic Letters*, 16, 357-362.
- Samuelson, PA. (1947) Foundations of Economic Analysis, Harvard University Press.
- Schmalensee, R. (1971) Consumer's Surplus and Producer's Goods. American Economic Review, 61, 682-687.
- Schmalensee, R. (1976) Another look at the Social Valuation of input Price Changes. American Economic Review, 1976, 239-243.
- Slade, M.E. (1994) What does an Oligopoly Maximize?. The Journal of Industrial Economics, 42, 45-71.
- Spence, M. (1976) Production selection, fixed costs, and monopolistic competition. *Review* of Economic Studies, 43, 217-235.
- Wisecarver, D. (1974) The Social Costs of Input Market Distortions. American Economic Review, 64, 359-372.

4 AIRPORT OWNERSHIP: EFFECTS ON PRICING AND CAPACITY¹

4.1 Introduction

In the last decades, some industry watchers, commentators and economists have argued in favor of the privatization of airports. They have given many reasons; among others, government revenues, financing aspects and private enterprise creativity and drive. On efficiency grounds, which is the focus of the present paper, it has been argued that private airports would charge more efficient congestion and peak-load prices and that they will respond to market incentives for capacity expansions (see e.g. Craig, 1996). These last points are important because, in the literature, congestion is often mentioned as the most important problem major airports face.

In 1987, the three airports in the London area and four other major airports in the UK were privatized. Following the example of the UK, many countries moved -or are movingtowards privatization of some of their public airports (among others, Austria, Denmark, New Zealand, Australia, Mexico and many Asian countries). Out of the concern that the privatized airports would exert market power -they would be local monopolies by having a captive market- most of the newly privatized airports have been subject to economic regulation, either in the form of price caps (as London Heathrow) or rate-of-return (as Flughafen Düsseldorf). Lately, however, many authors have argued that the regulation mechanisms fell short of being optimal; in particular, privatization has not been as successful as expected because the regulation mechanisms would misplace the incentives regarding capacity: price caps would lead to underinvestment while rate-of-return would lead to overinvestment in capacity.² Moreover, some authors and government agencies have argued that *ex-ante* regulation could be unnecessary altogether so it should be either completely divested or replaced by *ex-post* price monitoring. Why? Some of the reasons that have been put forward are the following (see e.g. Beesley, 1999; Condie 2000; Forsyth, 1997, 2003; Starkie, 2000, 2001, 2005; Productivity Commission, Australia, 2002; Civil Aviation Authority UK, 2004): (i) airports have low price elasticity of demand so price levels will not have large implications for allocative efficiency; (ii) airlines have countervailing power that will put downward pressure on airport prices; (iii) alternatively, most of the problems would be solved if deeper collaboration between airlines and airports was allowed and encouraged; and (iv) demand complementarities between aviation and concession activities would induce the airport to charge below monopoly prices on the aeronautical side (particularly when concession revenues are larger than airside revenues). In fact, the move towards divestment of regulation or the less-stringent price monitoring has already started in some countries (e.g. New Zealand and Australia).

However, as important as this may appear, there have been, to our knowledge, only two papers that have analytically examined what the outcomes of privatization or divestment of regulation may be (Zhang and Zhang, 2003; Oum et al., 2004). And, although there are

¹ A version of this chapter has been submitted for publication. Basso, L.J. (2005) Airport Ownership: effects on pricing and capacity.

² For a list of papers that discuss country-specific experiences with regulation see Oum et al. (2004).

many analytical papers that examine optimal pricing of public airports, most of the papers that do deal with privatization and divestment of regulation issues are fairly descriptive. Forsyth (2003) acknowledges this: "The shift to price monitoring has been a response to these problems [the problems with regulation], though the content and likely impact of monitoring has yet to be determined". What this paper does, precisely, is to analyze the effects of airport ownership on prices and capacities using a formal model, since the suggested move towards private unregulated airports is fairly new and hence empirical analyses are not feasible. The idea is to make an analytical examination of some of the assertions that have been put forward in the literature regarding privatization and regulation of airports, and to gain insights about other issues that have yet been discussed.

What makes this paper different from the previous two is the way the airline market enters the picture. Zhang and Zhang (2003) and Oum et al. (2004) essentially abstract from it, assuming that an airport's demand is a function of a full price –which includes airport charges and congestion costs–, and measuring *consumer surplus* through the integration of the airport's demand. In this paper, we formally model the airline market as an oligopoly, which takes airport charges and capacities as given, recognizing that this is a vertical setting: airports provide an input –airport service–, which is necessary for the production of an output –movement– that is sold at a downstream market. Hence, the demand for airports services is a derived demand. Indeed vertical settings similar to the one considered here have been proposed before (Brueckner, 2002; Pels and Verhoef; 2004; Raffarin, 2004), but they have been mainly used to study optimal congestion pricing. Optimal capacity or the effects of privatization have not been analyzed (capacity has always been assumed to be fixed).

In this paper we look into private ownership and allow capacity to be a decision variable. We consider both system and individual privatization of airports, and the case of joint maximization of airports' and airlines' profits, comparing these cases against both the firstbest benchmark and budget constrained public airports. Analytical and numerical results show a rather unattractive picture for privatization when compared to the first-best. First, the idea that low elasticities of demand for airports would induce small allocative inefficiency would be true only if the elasticity was constant, something rather improbable. Observed elasticities from public airports or regulated airports are not evidence that this would be the case in a private unregulated airport; monopolies price in the elastic range of the demand. What is obtained here is that important allocative inefficiencies may well arise. Besides, the low price elasticity argument overlooks the fact that private and socially optimal capacities will be different. When capacity becomes a decision variable of the airport -an idea that dominates the airport privatization and regulation literature- private airports would tend to be fairly small in terms of capacity, which further decreases traffic. Results worsen when privatization is done on an airport by airport basis rather than in a system because when airports are both origin and destinations of trips, their demands are perfect complements and therefore 'competition' between airports induces a horizontal double marginalization problem. On the other hand, the maximization of joint profits benchmark shows that the arguments regarding airlines countervailing power or an increased scope for cooperation between airlines and airports are probably overstated. The outcome does improve but still falls far off from the first-best. When privatization is compared against public airports that have a budget constraint, its performance depends on whether public airport are able to use two-part tariffs or not. If they are, the results are essentially unchanged. If they are not, the gap diminishes but remains large.

The airline oligopoly model we use expands on previous work in several ways: airlines' demands are sensitive to schedule (frequency) delay cost in addition to flight delay caused by congestion at the airport, airlines services are not necessarily perfect substitutes, and the impact of the number of firms on airport demand (a proxy for market structure) is highlighted. Evidently, the idea is to better understand how these three aspects of the downstream market influence the performance of the airport market. It is shown that they have an important role on the incentives an airport has with respect to the dominance by a single airline.

The plan of the paper is as follows: Section 4.2 contains formal modeling of the downstream airline market. The derived demand for airports is obtained and characterized here. We analyze whether this derived demand carries enough information about the downstream market so that it would be possible to focus only on the airport market, abstracting from what happens downstream. Section 4.3 uses the results obtained in the previous Section to analyze airport pricing, capacity and incentives under private and public ownership. Since most of these analyses rely on comparative statics, Section 4.4 provides numerical simulations that allow a better assessment of the differences. Section 4.5 concludes.

4.2 The Airline Market

4.2.1 *The airline oligopoly model*

The oligopoly model presented here is used to obtain the derived demand for airports and to characterize it. We start by making two simplifying assumptions: First, we abstract from network and route structure decisions by having only two national airports. Second, we assume there is only demand for round trips, not for unidirectional trips.³ Having two airports enable comparisons between system and individual privatization later. The game we analyze is a three stage game: first, airports choose their capacities, K_h ; second, they choose the charge per flight, P_h ; finally, airlines choose their quantities. We look for subgame perfect equilibria through backward induction, so we focus first, in this Section, on the Nash equilibria of the airlines' sub-game. We consider N airlines with identical cost functions, facing differentiated demands in a non-address setting with fixed variety. Thus, differentiation is horizontal and N is exogenous and represents the main airline industry

³ These assumptions are consistent with Pels and Verhoef (2004) and are, in fact, a generalization of Brueckner (2002) and Raffarin (2004), who consider a single airport. Zhang and Zhang (2003) and Oum et al. (2004), also consider an airport in isolation.

structure indicator in the model. Each firm's demand is dependent on the vector of full prices,⁴ θ :

$$q_i(\mathbf{\theta}) = q_i(\theta_i, \mathbf{\theta}_{-i})$$

$$\theta_i = t_i + G(\tau_i) + \alpha \left(D(Q, K_1) + D(Q, K_2) \right)$$
(4.1)

where q_i is the demand faced by airline *i*, θ_i is its full price, *Q* is the total number of flights of all airlines, t_i is the ticket price for the round trip, $G(\tau_i)$ is schedule delay cost,⁵ τ_i is the expected gap between passengers' actual and desired departure time, $D(Q,K_h)$ represents flight delay because of congestion at airport *h* and α represents the passengers' value of time. Note that τ_i depends on the frequency chosen by airline *i*; the higher the frequency, the smaller the gap. Thus, schedule delay cost can be written as $g(Q_i) \equiv G(\tau_i(Q_i))$ where Q_i is the number of flights of airline *i*, $g'(Q_i) < 0$ while $g''(Q_i)$ has no evident sign a priori. The delay function considered is

$$D^{h}(Q, K_{h}) = \frac{Q}{K_{h}(K_{h} - Q)}$$
(4.2)

This convex function of Q was proposed by the US Federal Aviation Administration (1969) and is further discussed in Horonjeff and McKelvey (1983).⁶ D^h is the total delay of both take-off and landing at airport h, which requires to assume that take-off and landing capacities are equal.

Assuming that demands are linear, symmetric and airlines' outputs are substitutes:

$$q_i(\mathbf{\theta}) = a - b\theta_i + \sum_{j \neq i}^N e\theta_j$$
(4.3)

where a, b and e are positive. Inverting the system and re-labeling we get

$$\theta_i = A - B \cdot q_i - \sum_{j \neq i}^N E \cdot q_j \tag{4.4}$$

⁴ The use of full prices is a common feature of airline and transport economics papers. They are indeed used in the previous airport privatization papers but there they directly determine airport's demand (e.g. Oum et al. 2004).

⁵ Schedule delay cost represents the monetary value of the time between the passenger's desired departure time and the actual departure time. It was first introduced by Douglas and Miller (1974) and there, it was the addition of two components: frequency delay cost and stochastic delay cost. The former is a cost induced by the fact that flights do not leave at a passengers' request but have a schedule. Stochastic delay has to do with the probability that a passenger cannot board her desired flight because it was overbooked. Overbooking arises in the presence of stochastic demands, which is not the case here; hence our schedule delay cost corresponds only to frequency delay cost. For more on schedule delay cost, see also Small (1992).

⁶ This delay function has been used by Morrison (1987), Zhang and Zhang (1997) and Oum et al. (2004). Pels and Verhoef (2004) and Raffarin (2004) considered delay functions that were linear on the traffic level.

where A, B and E are positive. As in Vives (1985), A, B and E are assumed to be fixed and B>E, that is, outputs are imperfect substitutes. It is easy to verify that B>E is equivalent to b>(N-1)e: if all full prices increase by the same amount, the demand for airline *i* will decrease (a condition sometimes called *diagonal dominance*). We assume that airlines behave as Cournot oligopolists in that they choose quantities, an assumption that is backed by some empirical evidence (Brander and Zhang, 1990; Oum et al., 1993). In the absence of congestion and schedule delay cost –i.e. when $\theta_i \equiv t_i$ – and with constant marginal cost, the game is well behaved in that a unique equilibrium exists. Furthermore, market power decreases with the number of firms (Vives, 1985). So, in principle, the model is useful to assess the importance of airline industry structure (N) on airport pricing.

Three more comments about the demand model are important. First, note that homogeneity in the Cournot competition, the usual case in airline oligopoly models⁷, is a special case of our model (it will suffice to replace E by B in the results). This enables an assessment of the importance of (horizontal) airline differentiation in airport decisions. Second, we incorporated the schedule delay cost, an important aspect of service quality which has sometimes been considered in pure airline oligopoly models but never in airport markets analysis.⁸ Finally, we chose to have N as an exogenous parameter because airports may have preferences regarding N that are different than the pure free entry equilibrium, and they may indeed have a sizeable influence in the number of active firms. Airports' preferred N under different ownership and pricing schemes is analyzed in Section 4.3. In any case, the equations that define the free entry N are easy to identify.

Using (4.4) and (4.1), the following system of *inverse demands* faced by the airlines can be obtained: $t^i = A - B \cdot q_i - \sum_{j \neq i}^{N} E \cdot q_j - g(Q_i) - \alpha(D(Q, K_1) + D(Q, K_2))$. This can be simplified though, by recognizing that $q_i = Q_i \times \text{Aircraft Size} \times \text{Load Factor.}$ Here, we assume that the product between aircraft size and load factor, denoted by *S*, is constant and the same across carriers, making the vertical relation between airports and airlines of the fixed proportions type.⁹ Thus

$$t^{i}(Q_{i}, \mathbf{Q}_{-i}) = A - SBQ_{i} - \sum_{j \neq i}^{N} SEQ_{j} - g(Q_{i}) - \alpha (D(Q, K_{1}) + D(Q, K_{2}))$$
(4.5)

It can be noted that linear demands in full prices do not lead to *inverse demands* that are linear in output, as D is not linear and there is no reason to think that g is.¹⁰ In fact, we make now the following useful assumptions regarding schedule delay costs:

⁷ See e.g. Brueckner and Spiller (1991), Oum et al. (1995), Brueckner (2002), and Pels and Verhoef (2004).

⁸ See e.g. Douglas and Miller (1974) and Oum et al. (1995). Morrison and Winston (1989) empirically measured schedule delay cost in their estimation of an airline choice model.

⁹ This assumption was also made by Brueckner (2002) and Pels and Verhoef (2004). A variable proportions case arise if, before a change in airport charges, airlines decide to change S (aircraft size, load factor or both).

¹⁰ Pels and Verhoef (2004) obtain linear inverse demands because they assumed a linear delay function and no schedule delay cost.

- (a) The monetary cost of the gap between the actual and desired departure times, τ , is proportional to its length.
- (b) τ is inversely proportional to the frequency of flights.

Assumption (a) is similar to what has been already assumed regarding congestion delay costs: (b) is equivalent to say that τ is directly proportional to the interval between flights (inverse of the frequency). Hence. under (a) and (b) we get $g(Q_i) = G(\tau_i(Q_i)) = \gamma \cdot \tau_i(Q_i) = \gamma \cdot \eta \cdot Q_i^{-1}$, where γ is the constant monetary value of a minute of schedule delay and η is a constant.¹¹ Thus, the residual inverse demand is negative and upward-sloping first; it then becomes positive, and then downward sloping, when the linear part of the function starts dominating schedule delay cost. Finally, for higher values of Q_i , congestion starts to kick and t^i decrease faster than linearly. This particular feature of this demand system is not troublesome though: the main insight is that schedule delay cost put by itself, and regardless of other technological considerations such as a fixed cost, a limit to the number of firms that can be active in the industry: there is a minimum scale of entry (see appendix B.2). What it does imply is that perfect competition is not consistent with this model.

The final ingredient necessary before analyzing equilibria is costs. Airline costs are

$$C_{A}^{i}(Q_{i}, \mathbf{Q}_{-i}, P_{h}, K_{h}) = Q_{i} \left[c + \sum_{h=1,2} (P_{h} + \beta D(Q, K_{h})) \right]$$
(4.6)

The term in square brackets is the cost per flight, which includes pure operating costs c, airports charges P_1 and P_2 , and congestion delay costs.¹² Using the expression for delay in (4.2), it can be verified that marginal costs are strictly increasing and larger than average cost (except at $Q_i=0$). Cost, marginal cost and average cost functions are strictly convex.

Airline *i*'s profits are obtained from (4.5), (4.6) and the fact that revenues are $t^i q_i = t^i Q_i S$. We get

¹¹ If passengers' desired departure time is uniformly distributed along the day, then assumption (b) holds and $\eta=1/4$. Note that we are assuming only one period here and not peak and off-peak periods: this is a model of congestion pricing and not peak-load pricing. If we were to assume more than one period (e.g. Zhang and Zhang, 1997), it would still be a reasonable assumption that, within each period, desired departure times are uniformly distributed. The results in this paper extend trivially to the case of many periods as long as demands in each period are independent of each other.

¹² Note that here it is assumed that c does not depend on aircraft size or load factor, which may appear as a strong simplification. Given that Q_i is directly proportional to q_i , this essentially says that the cost per passenger is fixed, something that has been assumed elsewhere (e.g. Brander and Zhang, 1990, Pels and Verhoef, 2004). Alternatively, one could assume that a single aircraft size and load factor prevail, as in Brueckner (2002). Also, the cost function should depend on a vector w of other input prices. That dependence could be modeled through c(w). Since it is assumed that input prices other than airport charges remain constant throughout, vector w will be suppressed for notational simplicity.

$$\phi^{i}(Q_{i}, \mathbf{Q}_{-i}, P_{h}, K_{h}) = \left[AS - BQ_{i}S^{2} - \sum_{j \neq i} EQ_{j}S^{2} - c - \sum_{h=1,2} P_{h}\right]Q_{i} - SQ_{i}g(Q_{i}) - (\alpha S + \beta)\sum_{h=1,2} Q_{i}D(Q, K_{h})$$

$$(4.7)$$

4.2.2 Equilibrium in the airline market: derived demand for airports and its characteristics

To obtain the derived demand for airports, we need to find the equilibrium of the airline market. Using (4.7), it can be shown that under assumptions (a) and (b) there exists a unique, interior and symmetric Cournot-Nash equilibrium of the sub-game, as long as N is smaller than the free-entry number of firms which should always hold¹³ (see appendix B.1; appendix B.2 contains the derivation of the free-entry equilibrium). Thus, $\partial \phi^i / \partial Q_i = 0$ gives us the unique and symmetric Cournot-Nash equilibrium of the game. Calculating this and imposing symmetry, we obtain the following important equation

$$\Omega(Q, P_h, K_h, N) = (\alpha S + \beta) \sum_{h=1,2} \left(D^h(Q, K_h) + \frac{Q}{N} D_Q^h(Q, K_h) \right) + S\left(g(\frac{Q}{N}) + \frac{Q}{N} g'(\frac{Q}{N})\right) + S^2 \left(2B + (N-1)E\right) \frac{Q}{N} + c + \sum_{h=1,2} P_h - AS = 0$$
(4.8)

Equation (4.8) implicitly defines a function $Q(P_h, K_h; N)$, which is airports' demand as a function of airport charges, capacities and airline market structure N (the implicit function theorem holds). Two observations are worthy to be made: first, under assumptions (a) and (b), g(x) + xg'(x) = 0 and so the second term would be zero. Second, one can define, without loss of generality, $P=P_1+P_2$; if airports were to be priced jointly then an explicit expression of the airports' inverse demand $P(Q, K_h; N)$ is obtainable.

We now characterize the demand for airports. We are interested first in learning how airports' demand changes with P_h , K_h and N or, alternatively, how the inverse demand $P(Q, K_h; N)$ changes with Q, K_h and N. Consider first changes of Q with N. If assumptions (a) and (b) hold,

¹³ As for Cournot (or *tatonnement*) stability, a sufficient condition is that the best reply mapping is a contraction. In homogenous Cournot, it is known that this condition holds for N small. Here, while differentiation (E < B) gives some latitude, congestion works in the opposite direction. In fact, the contraction condition was checked in the numerical application of Section 4.4 and results indeed show that it holds only for N=2.

$$\frac{dQ}{dN} = -\frac{\Omega_N}{\Omega_Q} = \frac{Q}{N} \left(\frac{(\alpha S + \beta) \sum D_Q^h + S^2 (2B - E)}{(\alpha S + \beta) \sum D_Q^h + S^2 (2B - E) + (\alpha S + \beta) \sum (ND_Q^h + QD_{QQ}^h) + S^2 EN} \right)$$
(4.9)

It can be checked that dQ/dN > 0, $d^2Q/dN^2 < 0$ and $dQ^i/dN < 0 \quad \forall E \in (0,B]$, so total flights increase with N at a decreasing rate, while each firm's number of flights decrease. In the absence of congestion, when $E \rightarrow 0$, Q^i becomes independent of N: each firm has its own turf. With congestion, Q^i decreases even when substitutability is very low ($E \rightarrow 0$) because the congestion externality causes marginal costs to increase. Note that, in this Section, this and any other examination of changes with respect to N are valid in the sub-game only: P and K are fixed and not functions of N yet as they will be in the equilibrium of the full dynamic game. All other derivatives are obtained in a similar fashion as above. In summary

$$\frac{\partial P}{\partial Q} < 0, \quad \frac{\partial P}{\partial N} > 0, \quad \frac{\partial^2 P}{\partial Q \partial N} > 0, \quad \frac{\partial^2 P}{\partial Q^2} < 0, \quad \frac{\partial Q}{\partial N} > 0,$$
$$\frac{\partial Q}{\partial P_h} < 0, \quad \frac{\partial Q}{\partial K_h} > 0, \quad \frac{\partial^2 Q}{\partial P_h^2} < 0, \quad \frac{\partial^2 Q}{\partial P_h \partial K_h} > 0 \tag{4.10}$$
$$\frac{\partial P}{\partial K_h} > 0, \quad \frac{\partial^2 P}{\partial Q \partial K_h} > 0, \quad \frac{\partial^2 P}{\partial K_h^2} < 0, \quad \frac{\partial^2 P}{\partial K_1 \partial K_2} = 0, \quad \frac{\partial^2 P}{\partial K_h \partial N} < 0$$

Results in the first two rows of (4.10) require assumptions (a) and (b) regarding schedule delay cost, while those in the third row do not.¹⁴

Having characterized the shape of the demand function, we can now compute the surpluses (in sub-game equilibrium) of airlines and passengers. Passenger surplus is given by $PS = \int_{(P,K,N)}^{\Lambda} \sum_{i}^{N} q_{i}(\theta) d\theta_{i}$. Since $\partial q_{i} / \partial \theta_{j} = \partial q_{j} / \partial \theta_{i}$, the line integral has a solution that is path independent (*PS* is equal to both Hicksian measures). Using a linear integration path, straightforward calculations lead to (see appendix B.3)¹⁵

¹⁴ Regarding how market power (air tickets) change with N, the result by Vives (1985) that $dt^i / dN < 0$ when $E \in (0,B]$ is the normal case here, but the opposite case may also arise. The intuition is: when N increases, prices tend to go down because of two effects: Q increase, so substitutability will put downward pressure on prices, and demands shift inwards because of increased congestion. Marginal costs of each firm go up though, because of congestion, which makes Q_i decrease, putting upward pressure on prices. When E is close to zero and if B is large enough (i.e. substitutability is weak), the first effect is not important, while the third effect may dominate the second, resulting in prices that actually increase with N (this is confirmed by numerical simulations). With no externalities and no substitutability, a change in N does not affect a firm's marginal cost or demand: N does not affect prices.

¹⁵ It can be checked that, when there is no congestion, $dPS / dN > 0 \forall E \in (0,B]$. But, as with t^i , when there is congestion and substitutability is small, the opposite case may occur. However, the somewhat strange case dPS / dN < 0 is not necessarily tied to increasing prices because, as N increase, not only quantities and prices changes but demands shift as well due to increased congestion. Thus, although prices may be smaller (larger), the area under the demand curves may have decreased (increased).

$$PS(P_h, K_h, N) = (B + (N - 1)E)S^2Q(P_h, K_h, N)^2 / 2N$$
(4.11)

The aggregate (equilibrium) profit for carriers, Φ , is easily obtainable from an individual carrier's profit (4.7) and the imposition of symmetry, that is, $Q_i = Q(P, K_h, N)/N$. We obtain:

$$\Phi(P, K_h, N) = QS \left[A - \frac{QS}{N} (B + (N - 1)E) - g\left(\frac{Q}{N}\right) - \alpha \sum D(Q, K_h) \right]$$

$$-Q \left[c + P + \beta \sum D(Q, K_h) \right]$$
(4.12)

We can now look at how much information about the downstream market is captured by the derived demand for airports. This is important because of the following: the airline market model was useful to derive and characterize the demand for airports (equations 4.8 and 4.10). It would be simple if we could directly use this demand function to fully analyze the airports markets, because this function may be estimated only with airport level information. In the private airports case, we will indeed use this demand function to setup the maximization of profits problem. Things are less obvious with the maximization of social welfare case, though. What is needed is a measure of *consumer* surplus. But as it is clear from this vertical setting, *consumers* of airports are both final consumers (passengers) and airlines. What we need then is a measure of the sum of passenger surplus and airlines profits. What has been assumed in previous papers about privatization, where the airline market is not formally incorporated (Zhang and Zhang, 2003; Oum et al. 2004), is that the airport demand does carry enough information so that its integration gives *consumer* surplus. We investigate now under which conditions this is true.

In Zhang and Zhang (2003) and Oum et al. (2004), the demand for the airport, Q, is assumed to be dependent on a full price ρ , which includes flight delay costs and the airport charge. They argue that, under perfect competition, the airport charge would be passed entirely to consumers. Using the notation of this paper, the demand for the airport would be $Q \equiv Q(\rho)$, where:

$$\rho = \sum_{h=1,2} P_h + (\alpha S + \beta) \sum_{h=1,2} D^h(Q, K_h) \equiv P + (\alpha S + \beta) \sum_h D^h$$
(4.13)

Indeed, $Q(\rho)$ defines a fixed-point rather than a closed form demand. Other charges to passengers, such as the flight ticket, are assumed to be exogenous as far as the airport is concerned. However, when one considers the full vertical structure and the associated subgame equilibrium, $Q^i(P_h, K_h; N) = Q(P_h, K_h; N)/N$, both delays (equation 4.2) and ticket prices (equation 4.5) will directly depend on airport charges and capacities, which are the decision variables of the airports. So, the first question is, is it reasonable to use the full price idea at the airport, rather than at the airline market? A clearer picture can be obtained by looking at equation (4.8). Using (4.13) to form ρ , and abstracting from schedule delay cost effects (i.e., making g=0), so that we can take $N \to \infty$, (4.8) can be written as:

$$QS^{2}\left(\frac{2B}{N} + \frac{(N-1)}{N}E\right) + \rho + c - AS + (\alpha S + \beta)\frac{Q}{N}\sum_{h}D_{Q}^{h} = 0$$
(4.14)

Hence, in general, Q would not depend only on ρ but also on D_Q and N; the (implicit) demand for airports should be $Q \equiv Q(\rho, D_Q, N)$. However, in the perfect competition case, that is when $N \to \infty$, (4.14) leads to $Q(N \to \infty) = \frac{AS - c - \rho}{S^2 E}$, which implies that $Q(\rho, D_Q, N \to \infty) \equiv Q(\rho)$. Thus, under perfect competition, a full price as defined by ρ , can in fact be used directly at the airport market level. It does summarize well the equilibrium of the downstream market.

Next, what has been (implicitly) assumed previously is that the integration of the airport demand with respect to the full price would capture the *consumer* surplus. Let us study this, using the general (implicit) demand function $Q \equiv Q(\rho, D_Q, N)$. We are thus interested in unveiling how

$$\int_{\rho}^{\infty} Q(\rho, D_{\varrho}, N) \, d\rho \tag{4.15}$$

is related to airlines profits and passenger surplus. For this consider the aggregate (equilibrium) profit for carriers, Φ , in equation (4.12). Regrouping terms to form ρ , and assuming away schedule delay effects, we obtain that:

$$\Phi(Q,\rho) = QS\left[A - \frac{QS}{N}(B + (N-1)E)\right] - Q[c+\rho]$$
(4.16)

Consider now the total derivative of Φ with respect to ρ . Using (4.16) the following results (see appendix B.4)

$$\frac{d\Phi}{d\rho} = -Q(\rho, D_Q, N) - \frac{(N-1)ES^2Q}{N} \frac{\partial Q}{\partial \rho} - (\alpha S + \beta) \frac{Q}{N} \sum_h D_Q^h \frac{\partial Q}{\partial \rho}$$
(4.17)

Reordering, integrating from ρ to ∞ , and using equation (4.14) we finally get¹⁶

$$\int_{\rho}^{\infty} Q(\rho, D_{Q}, N) \, d\rho = \Phi + PS - \frac{BS^{2}Q^{2}}{2N} - \frac{1}{N} (\alpha S + \beta) \int_{\rho}^{\infty} Q \sum_{h} D_{Q}^{h} d\rho \tag{4.18}$$

¹⁶ Here, we used the fact that $Q(\rho = \infty, D_0, N) = 0$ and therefore $\Phi(\rho = \infty, D_0, N) = 0$.

From equation (4.18) it is obviously clear that integration of the airports demand with respect to the full price, will deliver a correct measure of consumer surplus if and only if the airline market is perfectly competitive $(N \rightarrow \infty)$, which was in fact the maintained assumption of Zhang and Zhang (2003) and Oum et al. (2004). Hence, we have provided theoretical support for their modeling. When the airline market is imperfectly competitive though, the integral of Q with respect to ρ , does not capture airlines profits plus passenger surplus because market power induces losses of consumer surplus and partial internalization of congestion (third and fourth terms in equation 4.18 respectively).¹⁷

The main conclusion of the previous analysis is that to fully analyze airport markets, one cannot abstract from the airline market if competition is imperfect there. Formal modeling is required to adequately set up the social welfare maximization problem. The simplest way to do this is by considering directly the three actors involved, although one could also add the missing terms to the integral of airport's demand. At the practice level, the conclusion is bad news for managers of public airports or airport regulation authorities: even in a setting of complete information, optimal pricing and capacity require detailed knowledge about the market structure and demand of the airline market; information on costs and demand for airports alone is not enough. This unquestionably complicates the problem. Thus, fully modeling the vertical setting stands as a more correct way to study airport markets in general and the effects of airport ownership on prices and capacities in particular. The latter, which would have been new in any case, has now a stronger *raison d'être*.

4.3 The Airports Market

4.3.1 General Features

In this Section, we look at the first two stages of the game –airports' capacities and pricestaking as known the equilibrium in the third stage. We compare prices and capacities of private airports against first-best 'public' airports that maximize social welfare, but later address the problem of budget adequacy of public airports. We assume that both public and private airports impose per-flight charges, although this may not be a reflection of actual practice. The point is that we would like to shed light on what are the differences induced by ownership and not by failure to implement the right policies.¹⁸ It is further assumed that both type of airports have the same cost structure, despite arguments by some authors that public airports would not be as cost efficient as private ones (e.g. Condie, 2000). This

¹⁷ Jacobsen (1979) and Quirmbach (1984) are the classic references on the relation between input market surplus and the downstream market. Basso (2006) synthesize and generalize their results. Among other things, he shows that, in a case like this, the integral of the input demand with respect to P-as opposed to ρ - would never adequately capture downstream firms' profits plus final consumer surplus, not even under perfect competition.

¹⁸ Many airports actually have weight-based charges, but this has been starkly criticized on efficiency grounds by economists for at least three decades (e.g. Carlin and Park, 1970). And while in public airports efficient prices –which we show are not simply marginal cost– have certainly not been the norm (see discussions in Morrison, 1987, and Borenstein, 1992), private airports have not really moved to congestion pricing either (Forsyth, 2003; Starkie, 2005).Per-flight charges have been assumed in most analytical airport pricing papers.

assumption however, has some empirical support: Oum et al. (2004) found no significant differences between private and public airports in productivity terms. Overall, the idea is to understand the differences in a stylized benchmark model, leaving for future research the analysis of whether the alleged inefficiencies of public airports, if unsolvable, are important enough to make differences small.

As explained in the introduction, it has been argued that concession revenues would put downward pressure on airside charges of private airports, making regulation less needed (Condie, 2000; Starkie, 2001; Forsyth, 2003). We do not consider concession revenues here though, first, because probably the demand for concession services is not a Marshallian demand but a derived one, just as in the aeronautical side. Hence, when analyzing the public case, we would need to device a way to adequately consider the surpluses of all actors involved: profits of the companies that run the concessions and final consumers. Modeling and incorporating this would increase the complexity of the model while obscuring the insights we look for. The second reason to abstract from the concession revenues issue is that the outcome of incorporating them is probably known. Zhang and Zhang (2003) find that, while airside private prices diminish as expected, they decrease less than prices in a public airport that also has concessions. The intuition is simple: a private airport cares only about the profits from the concession activities while a public airport, maximizing social welfare, cares about consumer surplus as well. Consequently, the decrease in prices is stronger in the public case: concession revenues actually increase the gap between private and public airside charges. And while Zhang and Zhang modeled this without considering the vertical feature of the problem, it is very likely that their insight goes through. We see the confirmation of this, though, as future work.

4.3.2 System of Private Airports

We examine first the decisions of a System of Private Airports (SPA): pricing and capacity decisions at both airports are made by a single entity which maximizes profits. This is truly a monopoly situation quite comparable to the analysis of an airport in isolation (the most common case in the airport pricing literature). $Q(P_h, K_h; N)$ represents the demand for both airports as a function of prices, capacities and the (exogenously given) number of airlines. Decision variables are Q, P (which is the sum of P_1 and P_2), K_1 and K_2 , but Q and P are related through the demand function. We use P and K_h as decision variables –i.e. we use the inverse demand function $P(Q, K_h; N)$ but obviously results do not vary if we choose otherwise. In this setup, the three-stage game is identical to a two-stage game where Q and K_h are chosen simultaneously. As it is usual in the literature, it is assumed that an airport costs are given by C(Q) + rK, where, C are operating costs and the second term represents capital costs. The problem the SPA faces is given by

$$\max_{Q,K_1,K_2} \pi(Q,K_h;N) = P(Q,K_h;N)Q - 2C(Q) - (K_1 + K_2)r$$
(4.19)

where π is the sum of profits of both airports. First-order conditions lead to the following pricing and capacity rules:
$$P = 2C' + \frac{P}{\varepsilon_P} \tag{4.20}$$

$$Q\frac{\partial P}{\partial K_h} = r , \qquad h = 1,2 \qquad (4.21)$$

where ε_P is the (positive) price elasticity of airports' demand. As for second-order conditions, it cannot be proved that they hold globally but simulation show that they do hold for a large range of parameter values, particularly for the numerical applications in Section 4.4. A necessary condition though, is that *C* is not *too* concave.¹⁹ It is also easy to prove that at the optimum, $K_I = K_2 = K$ (see appendix B.5). Equation (4.20) is the familiar market failure in which monopolies set price above marginal cost. Equation (4.21) shows that private airports increase capacity until the marginal revenue of doing so equals the marginal cost of providing that extra capacity (recall that it was found that the marginal value of capacity, $\partial P/\partial K_h$, is positive). The monopoly system of airports only cares about the last or marginal *consumer*: increasing capacity by ΔK allow the airport to charge an extra ΔP , without loosing the marginal consumer (recall that a consumer lost for the airport is equivalent to a change in the equilibrium quantity in the downstream market). The extra charge however can be passed to all inframarginal consumers. What is important to note is that the marginal revenue perceived by the airport is not necessarily a measure of the social benefit of an increase in capacity (Spence, 1975).

Interesting as well, is to see how optimal Q, P and K_h change with N. Unfortunately, comparative statics are not definitive: derivatives cannot be signed a priori (see appendix B.6) so we will need to wait until the numerical simulation to have a better idea (the same goes for final outcomes in the airline market of course). What it is easy to show, however, is that as N increases, profits increases. To see this, simply differentiate profits evaluated at optimal Q and K with respect to N and apply the envelope theorem:

$$d\pi/dN = \pi_{Q}Q_{N}^{SPA} + \sum \pi_{K_{h}}K_{N}^{SPA} + \pi_{N} = \pi_{N} = Q^{SPA}P_{N} > 0.$$

4.3.3 System of Public Airports

Consider now a system of public airports that maximizes social welfare. This case will be denoted by W. According to the discussion in Section 4.2, with imperfect competition in the airline market, the social welfare (SW) function is not simply the integral of airports' demand plus airports' profits. The correct SW function can be obtained by directly adding passenger surplus (4.11), total airlines profits in the sub-game equilibrium (4.12)-, and airports profits: $SW = PS + \Phi + \pi$, that is,

¹⁹ Evidence suggests that airports' operational economies of scale are exhausted at fairly small amounts of traffic (e.g. Doganis, 1992).

$$\max_{Q,K_1,K_2} SW(Q,K_h;N) = P(Q,K_h;N)Q - 2C(Q) - (K_1 + K_2)r + \frac{(B + (N-1)E)S^2Q^2}{2N} + QS\left[A - \frac{QS}{N}(B + (N-1)E) - g\left(\frac{Q}{N}\right) - \alpha\sum D^h\right] - Q[c + P + \beta\sum D^h]$$
(4.22)

First-order conditions lead to

$$P = 2C' + \frac{(N-1)}{N} (\alpha S + \beta) Q \sum_{h} D_{Q}^{h} - \frac{BS^{2}Q}{N}$$
(4.23)

$$-Q(\alpha S + \beta) \frac{\partial D(Q, K_h)}{\partial K_h} = r, \qquad h = 1,2$$
(4.24)

Again, second-order conditions do not hold globally but do in the numerical simulation, at the optimum $K_1 = K_2 = K$, and results do not change if P and K_h were taken as the decision variables. As advanced, we do not impose a budget constraint here, so budget adequacy is not ensured. The discussion about this issue is delayed for Section 4.3.5.

The public airports' total charge has three components: marginal cost, a charge that increases price and is equal to the uninternalized congestion of each carrier, and a term that decreases price, which countervails airlines' market power. In fact, this system of public airports' manage to induce the outcome of social welfare maximization in the airline market: if one directly maximizes social welfare in the airline market, the equation that describe the final (symmetric) outcome is, precisely, $\Omega(Q, P, K_h, N) + ((N-1)/N)(\alpha S + \beta)Q\sum_h D_Q^h - BS^2Q / N = 0$. As can be seen, whether describe the final charge will be above or below marginal cost depends on whether the congestion effect or the market power effect dominates. For the monopoly airline case, congestion is perfectly internalized and the airports charge will be below marginal cost (and probably below zero). The third term in fact amounts to subsidize firms with market power in order to increase social welfare by diminishing allocative inefficiency. The implicit assumption is, evidently, that there is no other mechanism in place to control this market power. Note that if K is fixed, the market power effect decreases as N grows while the congestion effect increases. When K is not fixed, this is expectable but not clear cut, because K will change with N as well. In fact, the signs of dQ^{W}/dN , dK^{W}/dN and dP^{W}/dN cannot be determined a priori.

The congestion term was first found by Brueckner (2002). Pels and Verhoef (2004) later pointed out that the market power term was also needed (Brueckner acknowledged this, though, by stating that if market power was strong, the pure congestion charge may actually be harmful). There are some differences between Pels and Verhoef's result and the result here, however: (i) in Pels and Verhoef's model (and in Brueckner's), a regulator would charge a toll equal to the second and third terms in (4.23). Here, it is the public airport that distorts marginal cost pricing by an amount equal to that toll; (ii) they only considered a duopoly in a homogenous Cournot setting while here there are N firms in a differentiated Cournot setting; (iii) they assumed a delay function that was linear in traffic while here it is not; (iv) they assumed a fixed capacity while here capacity is not fixed. Hence, it can be seen that their main insight regarding price expands to a more general case.

As for capacity, public airports will add capacity until the costs of doing so equate the benefits in saved delays to passengers and airlines. Clearly, this capacity decision is different from the decision (a system of) private airports make, as they care about extra revenues and not extra social benefits (Spence, 1975, provided this insight). This result differs from what was obtained by Oum et al. (2004) as they found that private and public airports followed the same capacity rule, and hence it was concluded that private airports set capacity levels efficiently for the traffic they induced through pricing. The divergence is caused by the fact that their set-up only holds for a perfectly competitive airline market, as discussed in Section 4.2.3. In effect, if one replaces in the private airport capacity rule, marginal value of capacity by (4.21), the its full expression, i.e. $\frac{\partial P}{\partial K_h} = -(\alpha S + \beta) \left(\frac{Q}{N} D_{QK}^h + D_K^h \right), \text{ one can see that, if } N \to \infty, \text{ then the capacity rules}$ (4.21) and (4.24) do coincide.

How does social welfare change with N? Differentiating SW evaluated at optimal Q and K with respect to N, and applying the envelope theorem we get:

$$\frac{dSW}{dN} = \frac{\partial SW}{\partial dN} = \frac{(B-E)S^2Q^2}{2N^2} + Sg'\left(\frac{Q}{N}\right)\frac{Q^2}{N^2}$$
(4.25)

The first term on the right hand side is non-negative while the second is negative. It can be seen that when differentiation is weak, (4.25) may be negative implying that it would be better, in a social welfare sense, to have one airline. This may appear surprising but the explanation is simple: with both market power and the congestion externality controlled, as it is the case here, a monopoly airline provides a higher frequency than each airline in oligopoly, thus diminishing schedule delay cost, which increases demand. When differentiation is strong, (4.25) would probably become positive. In that case, the expansion of demand generated by a new firm will overweight the increased schedule delay cost due to reduced frequencies. The notable thing is that, in this model, with 'enough homogeneity' a monopoly airline is optimal but there is no need to regulate it: the public airports system would subsidize the airline to induce the optimal quantity (but there is still the issue of budget adequacy). These results were not obtainable in Pels and Verhoef's model because they only considered a homogenous duopoly and no schedule delay cost. Brueckner did considered N firms, but (4.25) would have always been zero in his case because his model featured homogeneity and no schedule delay cost.

We compare now the system of private airports and the system of public airports. Regarding price, we know the SPA price will be above marginal cost; the W price may be above or below marginal cost depending on whether the congestion effect or the market power effect dominates. May it happen that private airports charge less than public airports, actually inducing more traffic? The problem is that comparisons are complex because quantity (prices) and capacities are chosen simultaneously. A way to make comparisons feasible is to assume first fixed capacity.

Proposition 4.1: For a given *K*, the system of private airports will induce fewer flights than the system of public airports or, equivalently, it will charge a higher price.

Proof: See appendix B.7.

Proposition 4.1 indicates that it will never happen that the congestion effect leads to public airports that have smaller traffic than private airports; allocative inefficiency will always be in the form of restricted output. As explained in Section 4.1, it has been argued that this inefficiency would not be too important because the price elasticity of airports' demand is low. So, even if the price increases importantly, the actual quantity would not decrease as much. This assertion cannot be confirmed or negated with Proposition 4.1, but something can be said even at this point: observed price elasticities are not necessarily good forecasts of the value the price elasticity will attain under other circumstances. The contention would be true only if the price elasticity of airports demand is constant, something rather unlikely. For example, the efficient pricing rule in (4.23) has probably not been implemented in any airport, so we can hardly know what price elasticity value it would induce. More importantly, monopolies price in the elastic range of the demand. Thus, while it may be true that the price elasticity is low under the current pricing system (say, pure marginal cost, which as seen is *not* the efficient price), the system of private airports would price so to get into the elastic range of the demand, something that may indeed induce important allocative inefficiency. This issue will be further discussed under the light of the numerical simulation.

The reasoning regarding the price elasticity and the allocative inefficiency, also fails to take into account that a private airport would choose a different capacity than a public airport would. How can capacity decisions be compared? Various cases can be distinguished. First, quantity and capacity are defined simultaneously in a system of equations. We could therefore compare actual capacities and quantities. A more interesting question is, what distortions, if any, arise on the capacity side when the (well known) monopoly pricing distortion is taken into account. How would the SPA capacity compare to constrained social welfare maximization where monopoly pricing is taken as given? Is the distortion in capacity a mere byproduct of monopoly pricing? To analyze these two cases, we first examine the transposed of Proposition 4.1, i.e. what happens with K when Q is given (e.g. the airline market is frequency regulated). In these analyzes, the reader will find strong similarities with Spence (1975) examination of the provision of quality by a monopolist. Indeed, under the current modeling, K can be seen as a measure of quality. Spence's Proposition 4.1 states that, if $\partial^2 P / \partial Q \partial K > 0$, then a monopoly would oversupply capacity. However, Spence's insights -although pervasive- do not apply here directly even though we are in a case in which the above derivative is indeed positive (see 4.10). The problem is that in this case, the firm that has to choose quality provides an input to a downstream oligopoly and not a final product. Moreover, in the downstream (final) market, there are externalities in both production and consumption. Hence, a new proof is required.²⁰

Proposition 4.2: For a given Q, the system of private (SPA) airports will oversupply capacity with respect to the system of public airports (W).

Proof: See appendix B.8.

As for actual capacities and quantities, from Proposition 4.2 it is clear that, if for a given capacity the output restriction of the system of private airports is not too important, i.e. $Q^{SPA}(K) \approx Q^{W}(K)$ (these denote quantity rules for given K), then private airports' capacities will be higher than the W ones. If the output restriction is severe, $Q^{SPA}(K) \ll Q^{W}(K)$, then the result is reversed. The *low price elasticity* reasoning then is also important here because it has a counterpart in terms of capacity.

To analyze optimal social welfare capacities under monopoly pricing, consider the following constrained SW function, $\tilde{S}W(K) \equiv SW(Q^{SPA}(K))$, and maximize it with respect to K (recall that $K_1=K_2=K$). How does the second constrained social welfare capacity, \tilde{K}^w , compare to K^{SPA} ? Differentiating and evaluating at K^{SPA} we get

$$\frac{d\widetilde{S}W}{dK}\Big|_{K^{SPA}} = \frac{\partial SW}{\partial Q}\Big|_{Q^{SPA}(K^{SPA}), K^{SPA}} \cdot \frac{\partial Q^{SPA}(K)}{\partial K}\Big|_{K^{SPA}} + \frac{\partial SW}{\partial K}\Big|_{Q^{SPA}(K^{SPA}), K^{SPA}}$$
(4.26)

We are interested on the sign of (4.26). If it is positive, then constrained social welfare capacities are larger than the SPA ones. The first derivative on the right hand side is always positive by Proposition 4.1; the second one also is.²¹ The third derivative can be rewritten as $\partial SW / \partial K |_{Q^{SPA}, K^{SPA}(Q^{SPA})}$ —where $K^{SPA}(Q)$ is the SPA capacity rule— showing that it is negative by virtue of Proposition 4.2. Therefore, the sign of (4.26) is not determined a priori: we cannot say whether \tilde{K}^W is below or above K^{SPA} . However, if $Q^{SPA}(K) \approx Q^W(K)$, then $K^{SPA} > \tilde{K}^W$ because the first derivative on the right hand side of (4.26) would be close to zero by first-order condition in the (unrestricted) max SW case, and (4.26) would be negative (of course, we would also have $K_h^{SPA} > K_h^W$). So, if it was

 $^{^{20}}$ Something else that is worth noting is what is required to obtain the positive sign of the cross derivative of *P*. According to Spence, one way such a result may arise is consumers' heterogeneity; they would differ in their marginal willingness to pay for quality. A positive sign for the derivative shows that the marginal willingness to pay for quality of the marginal consumer is higher than for the average consumer. Here, however, things are different; both passengers and airlines are identical –they all care in the same way about congestion and therefore about capacity–, but the *marginal consumer* is not defined directly by differences in willingness to pay but by an equilibrium in the downstream market.

²¹ We have that $\partial Q^{SPA}(K) / \partial K = -\pi_{QK} / \pi_{QQ}$, but $\pi_{QK} = P_{QK}Q + P_K > 0$ (see equation 4.10) and $\pi_{QQ} < 0$.

true that private airports induce small allocative inefficiencies, this would mean that private capacities would be too large, even in a second best sense. If $Q^{SPA}(K) \ll Q^{W}(K)$, then the positive terms in (4.26) are more likely to dominate the negative one, and \tilde{K}^{W} may be above K^{SPA} : if the system of private airports restrict output severely, and therefore has smaller capacities, the public airports, when forced to price as the private system, would increase its capacity departing from K^{SPA} , as this directly benefits the other two parties, airlines and passengers. Overall, we can only say that, probably, the monopoly of private airports does induce distortions in capacity, which are in addition to pricing distortions. But whether this distortion is under or overinvestment can be unveiled only through numerical simulation, which is done in Section 4.4.

4.3.4 Maximization of Joint Profits: Airlines and Airports

There are at least two reasons why it is interesting to look at this case. First, it has been argued that regulation may be unnecessary -in that airport charges may be kept down and capacity investments may be more efficient- if, on one hand, airlines were allowed to have a stake at the airport or if deeper collaboration between airlines and airports was allowed and encouraged, or, on the other hand, if airlines had enough countervailing power (Beesley, 1999; Condie 2000; Forsyth, 1997, 2003; Starkie, 2000, 2001, 2005; Productivity Commission, 2002; Civil Aviation Authority UK, 2004). The maximization of joint profits emerge then as an obvious way to analyze these assertions. It would be the best that can be achieved if collaboration was allowed, while countervailing power would have an effect only on the division of profits. There might be a myriad of implementation problems though, as recognized in the literature (e.g. Condie, 2000; Starkie, 2005). We do not intend to model these problems here but, instead, to use the maximization of joint profits as a benchmark. If the outcome of the benchmark is deemed acceptable, then it can be later discussed how actual implementation would deviate from it.²² A second reason why it is interesting to look at the maximization of joint profits is because through a simple pricing scheme -two part tariffs-, that outcome is obtained in a non-cooperative fashion. With twopart tariffs, airports not only charge a per-flight price but they also charge a fixed-fee to each airline. Airlines then compete as in Section 4.2 but with this fee added to the cost function, which does not affect their quantity decisions but only whether they operate or not. The outcome is exactly that of maximization of the sum of profits: the system of private airports tries to maximize profits of the chain and then captures airlines' profits through the fixed fee. This is well-known in the vertical control literature and is somewhat surprising that almost no author has mentioned it (the only exception we are aware of is Borenstein, 1992). The difference with the usual setting is that here the upstream company has a quality (capacity) that matters.

This case is denoted JP, for joint profits. Using airlines' aggregate profit in (4.12), we set up the problem as:

 $^{^{22}}$ For a discussion about potential strategic coalitions between airlines and airports, see Albers et al. (2005).

$$\max_{Q,K_1,K_2} \pi + \Phi = P(Q,K_h;N)Q - 2C(Q) - (K_1 + K_2)r + QS\left[A - \frac{QS}{N}(B + (N-1)E) - g\left(\frac{Q}{N}\right) - \alpha\sum D^h\right] - Q[c + P + \beta\sum D^h]$$
(4.27)

First-order conditions yield

$$P = 2C' + \frac{(N-1)}{N}(\alpha S + \beta)Q\sum_{h}D_{Q}^{h} + \frac{(N-1)ES^{2}Q}{N}$$
(4.28)

$$-Q(\alpha S + \beta) \frac{\partial D(Q, K_h)}{\partial K_h} = r, \qquad h = 1,2$$
(4.29)

Again, second-order conditions do not hold globally and $K_1=K_2=K$ at the optimum. The price charged by the system of private airports –the variable part in the case of a two-part tariff–, has three components, each one related to a different *externality*. First, it has marginal cost to avoid the vertical double marginalization problem –a vertical externality to the vertical structure–, which arises in the SPA case. Second, it adds a charge equal to the uninternalized congestion cost of each carrier, a horizontal externality. Third, it adds a term to fight the *business-stealing effect*, a horizontal externality typical of oligopoly: firms do not take into account profits lost by competitors when expanding their output. The first two components are on line with maximization of social welfare while the third moves in the opposite direction; it destroys competition downstream instead of attacking airlines' market power. The final outcome is indeed that of cooperation between competitors in the airline market.

This result, which has not been obtained in the airport pricing literature before, has different intuitions depending on why the maximization of joint profits was the relevant case. With two-part tariffs, the private airports use the variable price to destroy competition downstream in order to maximize the profits of airlines, which are later captured through the fixed fee. The process is known: the fixed fee allows the marginal price to act only as an aligner of incentives, relieving it from the duty of transferring surplus as well. When the max joint profits case arises because of collaboration between airlines and airports, what happens is that airlines would like to collude in order to increase profits, but fail to do so because of the incentives to defect on any possible agreement. What they manage to de here, however, is to 'capture' an input provider to run the cartel for them. By increasing the price of the input, the input provider induces the collusion level of output. Here, the price increase takes into account both, the congestion externality and the business-stealing effect. Note that with N=1, there is no business-stealing effect and congestion is perfectly internalized by the monopolist; consequently, the last two terms vanish. Also, if airlines were completely differentiated, i.e. E=0, there would not exist the business-stealing effect but congestion would still need to be internalized. The upstream firm is rewarded with part of the profits, which is where bargaining power enters the picture.²³ Now, despite the fact

²³ This idea of an upstream firm running the cartel for the downstream firms has been discussed in the vertical control literature and, particularly, in the input joint-venture case. For example, Shapiro and Willig

that the result is as if airlines collude, this is not necessarily worse for social welfare than a system of private airports charging linear prices as in SPA because, here, two other harmful externalities are dealt with, the vertical double marginalization and the congestion externality. The final outcome is indeed closer to the public case as shown below. As for capacity decisions, it can be seen that the rule is the same as in the public case. This happens because this is the capacity that maximizes downstream profits as well (for a given Q).

The signs of $dQ^{J^{p}}/dN$ and $dK^{J^{p}}/dN$ cannot be determined a priori but we can know how, in equilibrium, joint profits change with N. For this, differentiate $\pi+\Phi$, evaluated at optimal Q and K, with respect to N and apply the envelope theorem:

$$\frac{d(\pi + \Phi)}{dN} = \frac{(B - E)S^2Q^2}{N^2} + Sg'\left(\frac{Q}{N}\right)\frac{Q^2}{N^2}$$
(4.30)

The analysis is similar to the social welfare case. When substitutability is weak, (4.30) may be negative so joint profits would be maximized with a monopoly airline: airports would have an incentive to let a single airline dominate. This may be facilitated if airlines and airports are encouraged to collaborate, as the airports may try to deal with only one airline and, together, foreclose entry to other airlines. In the two-part tariff case, the airport would extract all the profits of the monopoly airline through the fixed fee. What is remarkable is that for the SPA case, the larger the N the better, irrespective of the degree of substitutability. This was Borenstein's (1992, p.68) insight: he was critic about privatization of airports because, among other things, "without competition from other airports, an operator's profits would probably be maximized by permitting dominance of the airport by a single carrier and then extracting the carrier's rents with high facility fees". His comment is supported by these results but, in this model, airport domination by a single airline is not necessarily harmful. Social welfare may actually increase because, for N>1, it is still true that the congestion externality is internalized and that there is no competition, as with monopoly. But a monopoly will offer a frequency even higher than the frequency offered by each airline in the coordinated case, reducing schedule delay cost.

When airports are relatively indifferent between N=1 or higher, the implementation problems mentioned before may play a role. In the case of collaboration between airport and airlines, it may be easier for the airports to coordinate actions with only one airline, but it may be also true that this could increase the airline's countervailing power. With two-part tariffs, however, airports may still prefer to let a single airline dominate even if (4.30) is slightly positive because the pricing rule becomes simpler: (i) airports do not need to estimate the second and third terms of the pricing rule (something indeed difficult); (ii) they would need to worry about assessing the right fixed fee for only one firm. This shows that

⁽¹⁹⁹⁰⁾ conjecture that input joint-ventures can facilitate collusion and push a market toward the monopoly outcome. Chen and Ross (2003) formalize this. If airport provision was seen as an input joint-venture by the airlines, our results show two things in addition to what Chen and Ross found. First, that if there are externalities, the input price is, additionally, used to force their internalization by downstream competitors. Second, that when marginal costs are not constant downstream, the outcome is not as in monopoly or a downstream merger, but as in a cartel.

recognizing the scope for vertical control in airport pricing is important. Two-part tariff is the simplest form of vertical control and even this pricing mechanism has important and rather unexplored consequences on the airline market.

We can now turn to comparisons. They are summarized through the following Propositions:

Proposition 4.3: For a given K the JP airports will: (i) induce fewer flights than the W ones (ii) Induce more flights than the SPA ones

<u>*Proof*</u>: Part (i) is direct because $P^{JP}(K) > P^{W}(K)$ and $\partial Q/\partial P < 0$. The proof of (ii) is analogous to the proof of Proposition 4.1 (in appendix B.7).

Thus, for a given capacity, JP airports induce a smaller allocative inefficiency than SPA airports, showing that the proposal of increased collaboration is an improvement. How strong this allocative inefficiency would be cannot be unveiled until a parameterization is chosen; however, it can be easily pictured that it may not be small since in this case competition downstream is absent while in the public case, market power downstream is controlled.

Proposition 4.4: For a given Q, the JP airports will: (i) have the same capacity as W airports (ii) Have less capacity than SPA airports.

<u>*Proof:*</u> (i) is direct as they have the same capacity rule. (ii) follows from Proposition 4.2 and (i).

Proposition 4.5: As for actual capacities and quantities, JP airports will induce fewer flights and will have smaller capacities than W airports.

Proof: See appendix B.9

As before, whether actual JP capacities are below or above SPA capacities will depend on whether the output restriction of SPA airports, with respect to JP, is severe or not.

Next, it has been argued before that a capacity rule such as the one JP airports follow would be efficient because it is identical to the public one so, for a given Q, capacity will be set efficiently (Oum et al. 2004). The question we ask now is different: do JP airports induce distortions in capacity that go beyond what is induced only by pricing? To analyze this we look for constrained optimal capacities, by maximizing social welfare subject to the restriction of JP pricing. It can be shown that.

Proposition 4.6: The JP airports undersupply capacity with respect to optimal social welfare capacities under JP pricing (despite having the same capacity rule).

Proof: See appendix B.10.

4.3.5 On the budget adequacy of public airports

The comparison between private and first-best public airports is very useful as a benchmark, yet budget adequacy of public airports is evidently important for policy making. The issue of budget adequacy was explicitly considered by Zhang and Zhang (2003) and Oum et al. (2004), but in models that only looked at the airport market, with social welfare functions that are valid only if the airline market is perfectly competitive. On the other hand, in the airport pricing literature that takes into account the vertical relation between airports and airlines, airports profits are usually not considered in the social welfare function; only passengers' surplus and airlines profits are included. For example, Brueckner (2002) and Pels and Verhoef (2004) were interested in the toll that some airport authority has to charge to make efficient use of installed capacity, so whether revenues would cover costs or not was not examined.

As is evident from the analysis of first-best practice in Section 4.3.3, when N is small it is very likely that public airports would run a deficit because, in this case, it would be optimal for airports to subsidize the airlines (the market power effect dominates the congestion effect). Pels and Verhoef argued that when subsidies are optimal but unfeasible, it would be optimal to set the toll to zero, which in this model is equivalent to airports charging marginal cost. However, the analysis of the joint maximization of airports' and airlines' profits in the previous Section contained and important lesson: budget adequacy of public airports may be achievable through a fixed fee. Since lump-sum transfers will not affect airlines' marginal decisions, the airports may use the efficient pricing and capacity rules which may include actually paying airlines to land-, and then collect the money necessary to cover their expenses through a monthly facility fee. This would be a sort of Loeb-Magat mechanism, which has also been suggested for the access problem to telecommunications networks (Laffont and Tirole, 2000). Yet, as appealing as the mechanism may be, there is still no guarantee that two-part tariffs would enable cost recovery, because airlines may not make enough money to actually cover the airports' expenses. To be sure that this would be the case, the restriction $\pi + \Phi \ge 0$ would need to be included. The cost recovery two-part tariff pricing and capacity rules, case that we denote by CRT, are easy to obtain. The capacity rule would be the same as in W and the JP cases (equations 4.24 and 4.29), while the pricing rule is:

$$P^{CRT} = \frac{\mu}{1+\mu} P^{JP} + \frac{1}{1+\mu} P^{W}$$

$$= 2C' + \frac{(N-1)}{N} (\alpha S + \beta) Q \sum_{h} D_{Q}^{h} + \frac{S^{2} Q (\mu (N-1)E - \mu B)}{(1+\mu)N}$$
(4.31)

Where $\mu \ge 0$ is the Lagrange multiplier of the restriction, which captures the severity of the constraint. Here, μ balances the charge between the efficient first-best price and the JP price, enabling the airlines to make enough money to cover the airport costs through the fixed fee.

But, what if two-part tariffs are unfeasible? Note that setting the toll to zero, that, is charging marginal cost, would not be enough to cover airports costs even if the marginal cost function is flat, because the airport has to pay for the capacity (see Zhang and Zhang, 1997). In this case, the less efficient alternative of Ramsey-Boiteux pricing is called for. Formally, to ensure cost recovery using a linear price, the restriction that has to be considered is $\pi \ge 0$. This case, which we denote by CRL, is characterized by the following pricing and capacity rules:

$$P^{CRL} = \frac{\lambda}{1+\lambda} P^{SPA} + \frac{1}{1+\lambda} P^{W}$$

$$= 2C' + \frac{\lambda}{1+\lambda} \frac{P}{\varepsilon_{P}} + \frac{1}{1+\lambda} \left(\frac{(N-1)}{N} (\alpha S + \beta) Q_{\sum_{h}} D_{Q}^{h} - \frac{BS^{2}Q}{N} \right)$$

$$- \frac{Q(\alpha S + \beta)}{1+\lambda} \frac{\partial D(Q, K_{h})}{\partial K_{h}} + \frac{\lambda}{1+\lambda} Q \frac{\partial P}{\partial K_{h}} = r, \qquad h = 1,2 \qquad (4.33)$$

Where $\lambda \ge 0$ is the Lagrange multiplier of the restriction and, obviously, $\lambda \ge \mu$.

4.3.6 Independent Private Airports

So far, there has been no apparent need to have two airports in the model. We have them because in many cases the idea is to privatize airports independently and not in a system, and we would like to know what the outcome of this may be. However, there are some aspects here that were not present before and that need to be defined. First, do airports choose prices or quantities? This made no difference before but now it does. Given that airports' direct demands are $Q^1(P_1, P_2, K_1, K_2) = Q^2(P_1, P_2, K_1, K_2) = Q(P_1 + P_2, K_1, K_2)$, we take prices as tactical variables, so airports behave as Bertrand oligopolists with complement products. Second, are capacities and prices chosen simultaneously or sequentially, K_h first and then P_h ? The first case is usually called open-loop, the second closed-loop. In the closed-loop, the overall game has three stages as originally defined; in the open loop it has two. Let us look first at linear prices in the open-loop case. We denote this case IPA, for independent private airports. Airports choose P_h and K_h simultaneously in a non-cooperative game. Each airport's program is

$$\max_{P_h, K_h} \pi^h = Q_h(P_1, P_2, K_1, K_2) P_h - C(Q_h) - K_h r, \qquad h = 1,2$$
(4.34)

A necessary condition for existence of equilibria is that C is not *too* concave, something that has been assumed throughout. If this is the case, it can be shown that prices are strategic substitutes. We look for symmetric equilibrium. Interest lies on the sum of airport charges, P, rather than individual charges. First-order conditions and imposition of symmetry leads to

$$P = 2C' + 2\frac{P}{\varepsilon_P} \tag{4.35}$$

$$Q\frac{\partial P}{\partial K_h} = r, \qquad h = 1,2 \tag{4.36}$$

(4.35) is to be compared with the SPA case in (4.20); clearly $P^{IPA} > P^{SPA}$. This was expected: it is the result of the *horizontal* double marginalization problem that arises in oligopoly when outputs are complements. In these cases, competition is harmful for social welfare. Capacity rules are the same but obviously actual capacities will be different. Hence, independent private airports induce fewer flights and have smaller capacities than a system of private airports. From Propositions 1 to 4, we have that:

- For given K, $Q^{W}(K) > Q^{JP}(K) > Q^{SPA}(K) > Q^{IPA}(K)$.
- For given Q, we will have that, $K^{JP}(Q) = K^{W}(Q) < K^{SPA}(Q) = K^{IPA}(Q)$.
- For actual capacities and prices, $Q^W > Q^{TPT}$, $Q^{SPA} > Q^{IPA}$, $K^{JP} < K^W$ and $K^{IPA} < K^{SPA}$.

In the closed-loop game, where airports first choose capacities (simultaneously) and then prices, airports over-invest in capacity par rapport to the open loop. Qualitatively (a full derivation is in appendix B.11), what happens is that, in the three stage game, investment in capacity makes an airport tough: it leads to an own price increase, which hurts the other airport. Since in addition prices are strategic substitutes, increasing capacity increases own profits. Using the terminology of Fudenberg and Tirole (1984), airports over-invest in capacity following top-dog strategies. This leads to higher prices than in the open loop, but the overall effect on traffic is unclear.

What if the independent private airports collaborate with the airlines? In this case, the relevant problem is each airport maximizing its profit plus the profits of airlines, given that the other airport is doing the same. The outcome of this is the same as if airports, individually, charge two part tariffs (in an open-loop setting). We denote this case IJP. Solving the game, we get the following pricing and capacity rules (see appendix B.12)

$$P = 2C' + 2\frac{(N-1)}{N}(\alpha S + \beta)Q\sum_{h} D_{Q}^{h} + 2\frac{(N-1)ES^{2}Q}{N}$$
(4.37)

$$-Q(\alpha S + \beta) \frac{\partial D(Q, K_h)}{\partial K_h} = r \qquad , h = 1,2$$
(4.38)

Jointly, individual airports using two part tariffs or collaborating with airlines charge more than a system of private airports using a two-part tariff or collaborating with airlines (except when N=1). The horizontal double marginalization also arises here: each airport tries to correct externalities on their own and, as a result, they jointly overcharge for congestion and the business stealing effect. Capacity rules on the other hand are as in JP, therefore comparisons between this case and the JP case is analogous to the comparison between JP and W. Finally, whether there is over or under-investment in the close-loop cases cannot be determined analytically.

4.4 Numerical Simulations

The need for numerical simulations arises from three facts. First, since the move towards unregulated private airports is only a proposed move, which has not been implemented at a large scale, there is no real data to conduct an empirical analysis. Second, in this model, comparative statics and analytical comparisons were not conclusive in all cases. For example, it was not possible to know how Q, K and P change with N, or how SPA actual capacities and prices compare to W ones. And third, even when analytical results were obtainable, they were necessarily qualitative. For example, JP capacities and prices are below W ones, but by how much? We resort to simulation to shed light on these types of questions. We use the parameter values in Table 4.1.

	De	mand			Airlines	Airports		
α	40	A	2000	S	100	r	10000	
β	3000	В	0.15	Ν	varies	C'	2000	
γ	4	Е	0.13	c	36000			

Table 4.1: Parameter values for the numerical simulation

For the schedule delay cost, it is assumed that (a) and (b) in Section 4.2 hold, so that the schedule delay cost function is only defined by γ and η ; we impose η equal to one.²⁴ We consider a constant airport operational marginal cost, implying that economies of scale (if any) arise from the presence of fixed costs. We do not define a value for these so that airports' profits below are net of the fixed costs. It is not our intention to portray real aviation cases with these parameters but, rather, to obtain insights about what are the consequences of different ownership and pricing schemes; in fact, a two airports system is rarely the usual case. We did try, however, to be as reasonable as possible with the parameterization, by drawing data and values from other studies.²⁵ Relative comparisons

²⁴ As explained in footnote 11, if passengers' desired departure time is uniformly distributed along the day, then assumption (b) holds and $\eta=1/4$. We chose a larger η because we wanted to capture the fact that, in some cases, passengers cannot take the scheduled flight they would like to since they are already sold out. Taking $\eta=1/4$ or $\eta=1$ though, will analytically only affect the value of the air ticket *t*, not *P*, *Q* or *K*.

²⁵ The values of some of the parameter may be justified as follows: For α , Morrison and Winston (1989, p. 90) empirically found a value of \$45.55 an hour in 1988 dollars; for γ , they found a value of \$2.98 an hour in 1983 dollars (p. 66). For β , Morrison (1987, p. 51 footnote 20), finds that the hourly extra cost for an aircraft due to delays is approximately \$1,700 (resulting from 3,484 – 18*100) in 1980 dollars. For S, recall that it reflects the product between aircraft size and load factor. In North America, the average plane size in 2000 was 159 (see Swan 2002, Table 2); considering in addition an average load factor of 65% (see Oum and Yu, 1997, p.33) we obtain a value for S of 103.35. Regarding airlines' operational per flight cost c, Brander and Zhang (1990) proposed the following formula for the marginal cost per passenger in a direct connection:

between results are probably more enlightening that the individual results by themselves, although previous results in the literature (discussed below) confirm the plausibility of the parameterization.

Table 4.2 summarizes some of the results obtained. It has both, variable-capacity and fixedcapacity cases; the latter, in order to better see whether the argument that says that allocative inefficiency would be small with privatization holds or not. When capacity is fixed, it was set at the socially optimal level but choosing it otherwise does not change the qualitative conclusions. When airports are independent (IPA and IJP cases), results are for open-loop games. Social welfare is presented in terms of percentages rather than dollars. Second order conditions hold in all cases. As reading directly from Table 4.2 is a rather difficult task, we highlight below what we deem are the main insights gained from the numerical simulation.

- 1. In the system of private airports (SPA) case, both Q and K increases with N; delays also increase. Airports charges, P, almost do not change, leading to a t that decreases with N because of increased competition downstream. Airports profits increase with N as analytically showed, but social welfare also does. P is fairly large in all cases and way above marginal cost. This, however, is on line with a previous result: Morrison and Winston (1989) found that the difference between the monopoly and the *efficient* perpassenger landing fee was \$498.4. Multiplying this by S=100 lead to a difference of \$49,840 dollars per flight. Since they did not formally consider the airline market, their results are valid for perfect competition in the airline market. In our case, the difference between SPA and W when N=10 is \$43,505 per landing (recall that P is the sum of charges at both airports), which is comparable to theirs.²⁶
- 2. For public airports (W), Q and K increase with N but delays decrease, as opposed to the SPA case. P increases with N: as N grows, the congestion effect starts dominating the market power effect. When N=1, congestion is perfectly internalized by the airline without the need for correction from the part of the airports, and market power is at its ceiling; the need for subsidy is hence at its maximum, as is evident from the negative and large value of P. The charge increases with N and for N large enough subsidies are no longer required. Subsidies required when N is small may appear large but are consistent with Pels and Verhoef (2004) results.²⁷ It can also be seen that SW increases with N: differentiation dominates the schedule delay cost effect in equation (4.25). When homogeneity is increased (not shown), the result reverses as explained. Finally, air tickets decrease marginally with N, due to increased schedule delay costs because of depressed frequencies.

 $cpm(D/AFL)^{-\theta}D$; where cpm is the cost per passenger-mile, D is the origin-destination distance, AFL is the average flight length of the airline and θ is the cost sensitivity to distance. The following were the average values for American and United Airlines in the period 1981-1988 (see Oum et al., 1993): cpm=\$0.12/pax/mile, AFL=775 miles and $\theta=-0.43$. If we use AFL=800, cpm=\$0.20 and D=1000 (e.g. Chicago-Austin), and multiply the result by 2S to reflect the operational cost of a return flight, we obtain a value for c of \$36,340.

²⁶ This, despite the fact that they used a different delay function (theirs was estimated and homogenous of degree one on Q and K), and that their airport's demand was actually estimated.

²⁷ In the monopoly airline case, they obtained a toll (congestion plus market power components in the pricing rule 21) which was negative and equal to 340,000, and an air ticket of 1,393. Here, the subsidy equals 130,263 (marginal cost was deducted to obtain their *toll*) and the air ticket is 608.14.

N	Туре	Q	K	Р	D	t	PS	Φ	π	$\pi + \Phi$	SW
1	SPA	21.92	31.21	93,629	0.076	1,665	360,368	798,224	1,340,397	2,138,621	40.15
	W	92.18	100.18	-134,263	0.115	608	6,372,189	14,599,332	-14,748,019	-148,686	100.00
	JP	45.85	51.48	4,000	0.158	1,300	1,576,519	4,080,700	-1,029,579	3,051,121	74.36
	CRL	42.92	53.71	29,024	0.074	1,350	1,381,798	2,985,054	0	2,985,054	70.17
	CRT	91.09	99.05	-131,008	0.116	624	6,222,620	14,278,419	-14,278,419	0	99.99
	IPA	11.62	17.78	123,352	0.106	1,817	101,304	252,124	822,021	1,283,695	22.25
	IJP	45.85	51.48	4,000	0.158	1,300	1,576,519	4,080,700	-1,029,579	3,051,121	74.36
3	SPA	36.10	45.52	93,533	0.084	1,500	890,628	719,011	2,321,856	3,040,867	57.44
	W	101.23	109.62	-33,201	0.110	608	7,001,866	5,800,933	-5,958,055	-157,122	100.00
	JP	50.36	56.27	61,129	0.152	1,299	1,733,133	1,606,429	1,751,761	3,358,190	74.38
	CRL	71.11	81.31	26,867	0.086	1,021	3,455,375	2,754,159	0	2,754,159	90.72
	CRT	100.08	108.42	-31,065	0.111	623	6,843,847	5,677,648	-5,677,648	0	99.99
	IPA	19.80	26.23	123,174	0.117	1,719	267,820	238,905	1,518,049	2,073,541	34.21
	IJP	34.34	39.21	90,607	0.180	1,516	805,821	820,940	2,189,990	3,010,930	55.76
10	SPA	46.41	54.67	93,389	0.103	1,378	1,421,451	363,240	3,055,011	3,418,251	68.27
	W	104.83	113.37	6,379	0.108	607	7,252,401	1,855,122	-2,017,987	-162,865	100.00
	JP	52.16	58.17	83,218	0.149	1,299	1,795,457	509,499	2,968,472	3,477,971	74.38
	CRL	91.84	100.71	25,934	0.103	779	5,566,250	1,410,818	0	1,410,818	98.41
	CRT	103.64	112.13	8,121	0.109	623	7,088,606	1,815,514	-1,815,514	0	99.99
	IPA	25.87	31.67	122,926	0.141	1,646	441,640	124,213	2,054,761	2,567,223	42.44
	IJP	31.16	35.79	113,528	0.188	1,572	640,840	205,027	2,697,114	2,902,142	49.97
Fixed capacity (at the socially optimal level)											
3	SPA	42.76	109.62	84,054.67	0.0058	1,414.86	1,249,429	914,925	1,230,779	2,145,705	49.60
	W	101.23	109.62	-33,200.87	0.1100	607.66	7,001,866	5,800,933	-5,958,055	-157,122	100.00
	JP	58.37	109.62	54,794.35	0.0104	1,201.25	2,328,119	1,708,355	772,470	2,480,825	70.26
	CRL	68.31	109.62	36,095.95	0.0151	1,065.09	3,188,313	2,344,470	0	2,344,470	80.83
	CRT	100.05	109.62	-29,206.49	0.0954	624.85	6,840,677	5,514,806	-5,514,806	0	99.94
	IPA	28.53	109.62	110,686.70	0.0032	1,609.47	556,042	406,238	850,949	1,257,186	26.49
	IJP	44.35	109.62	81,072.63	0.0062	1,393.08	1,344,236	984,543	1,226,018	2,210,560	51.93

Table 4.2: Results of numerical simulation

SPA: system of private airports (linear price). W: system of public airports maximizing social welfare. **JP**: maximization of joint profits: airports and airlines (or two-part tariff). **CRL**: public airports with linear cost recovery pricing. **CRT**: public airports with cost recovery two-part tariff pricing. **IPA**: independent private airports using linear prices. **IJP**: independent airports maximizing own profit plus airlines' profits (or each airport with two-part tariff).

Q: traffic K: capacity P: Airport charges D: delay t: air ticket PS: Passenger surplus Φ : Airlines profits π : Airports profits SW: Social Welfare

- 3. Regarding the JP case, Q and K increase with N while delays decrease. P increases with N: as N grows, both the uninternalized congestion and the business stealing effects are more important, and they are not countervailed by changes in capacity. When N=1, the monopoly airline perfectly internalizes congestion and it obviously produces at the profit maximizing level so there is no need for corrections from the part of the airports: P is thus equal to marginal cost. Air fares decrease marginally with N, due to increased schedule costs because of depressed frequencies (recall that competition is destroyed in this case). Joint profits increase with N: differentiation dominates the schedule delay cost effect in equation (4.30). When homogeneity is increased though, the result reverses so it would be better for the airports to let a single airline dominate. This is not harmful for society however as social welfare actually increases (results not shown).
- 4. It can be seen that actual SPA capacities and quantity are way below social welfare ones. We are therefore in the case in which the SPA output restriction is severe: $Q^{SPA}(K) << Q^{W}(K)$, implying that $K^{SPA} < K^{W}$. This is confirmed by the fixed-capacity simulations. The main insight here is that the allocative inefficiency of private airports, if capacity is exogenously decided, is by no mean small, leading to important dead-weight losses. The argument was that price elasticities of demand are low, but the problem with that assertion is that it assumed the elasticity is constant. Observed elasticities however, are not the elasticities that would arise under private (unregulated) ownership, or with the efficient prices derived in (4.23) because efficient prices have not been the rule. In fact, it is true that the price elasticity of demand when P is equal to the linear cost-recovery price is fairly low (around 0.14 in absolute value) but, still, allocative inefficiency is very strong. Traffic is even smaller when the private airports can choose capacities. But since there is less waste of resources, social welfare is higher when capacity is a decision variable.
- 5. Comparisons between the SPA and the JP cases were not analytically simple. We can now see that, in general, JP airports are less harmful for social welfare than SPA airports. They induce higher joint profits and consumer surplus –and therefore SW–, more traffic, higher capacities and smaller airfares. The differences decrease with N though, because in the JP case competition downstream is destroyed for every N while in SPA it is not. These findings support the idea that collaboration between airlines and a system of private airports leads to a better outcome. There are two important things to note however. First, the same outcome is obtained through vertical control by the airports; two-part tariffs are enough in this case. Second, and more importantly, JP airports' traffic and capacities, while higher than the SPA ones, still fall way off optimal ones, which reflects in large deadweight losses. It would be adventurous, to say the least, to conclude that with privatization and collaboration –or strategic agreements– between airlines and airports, regulation becomes unnecessary. If anything, the outcome is closer to private unregulated airports rather than optimal ones.
- 6. What about budget adequacy of public airports? As expected, public airports charging linear prices (W cases) would run deficits and, although these diminish as the number of airlines increase, they are still sizeable when N=10. In Section 4.3.5 we argued that, perhaps, two-part tariff would solve the problem. The simulation however shows that this is not the case: airlines do not make enough profits. Hence, the cost-recovery cases gain importance. It can be seen that *cost-recovery two part tariff* (CRT) falls extremely close to the first best, showing that if two part-tariffs are feasible, even when one

considers budget adequacy, privatization induces important deadweight losses. If twopart tariffs are unfeasible, then the relevant cases to be compared are SPA and CRL. Obviously the performance of private unregulated airports is better here, but they still induce about half the traffic they should, generating deadweight losses of about 30 to 40%.

- 7. When comparing delays, it can be seen that in almost all cases, SPA airports have the smallest delays. This issues a warning: congestion has been one of the main drivers of research in this area and proponents of privatization have argued that private airports would charge efficient congestion and peak load prices and would respond to market incentives for expansion. If one measures the result of privatization only by its effects on congestion, privatization may appear as a better idea than it actually is. Despite the smaller delays, we have seen that the private airports themselves would be substantially smaller both in terms of traffic and capacity. More importantly, social welfare would be substantially smaller. JP airports on the other hand, would have larger delays than the public airports.
- 8. When airports are privatized individually (IPA and IJP cases), the horizontal *double* marginalization problems visibly arise. Independent private airports charging linear prices (IPA), while still performing better than public airports congestion wise, are very small and induce the larger airfare and the smaller social welfare. For individual airports collaborating with airlines –or charging a two-part tariff–, something stranger happens: as N increases, the double-charging problem worsens: total airport charges increase with N faster than in the system case, making traffic and capacities actually get smaller as the number of airlines increase (resulting in important reductions of social welfare as N grows).
- 9. Regarding constrained social welfare capacities (results not shown in Table 4.2), when N=3, if public airports are forced to price as SPA, they will increase capacity from 45.5 to 59.1, which will lead to a traffic of 41.2 instead of 36.1. It will still be far away from the first-best capacity though, which was 109.6. Hence, SPA does induce an extra distortion: given that their restriction of output is severe, they undersupply capacity with respect to the second best.
- 10. Finally, the insights do not qualitatively change with changes in the value of the parameters, although some numbers do. Specifically, different values for the demand parameters (A, B and E) and for r and for C' were tried, since for these parameters there was less external information. It was found that the impact of changes in r and C' are quite small, while demand parameters impact on the levels but not on the order of the results. For example, taking A=5,000 and B=1, as in Pels and Verhoef (2004), and then taking E=0.8 and N=3, SPA traffic decreases from 36 to 18 and SPA capacity decreases from 45 to 24, while W traffic decreases from 101 to 50 and W capacity decreases from 110 to 56.

4.5 Final Comments

Privatization of airports has been argued for on the grounds that private airports would implement more efficient congestion and peak-load prices, and would have better incentives to invest in capacity. Privatized airports have been subject to economic regulation though, out of the concern that they would exert market power. But it has been argued that regulation may be unnecessary because a private unregulated airport would not induce large allocative inefficiencies since price elasticities are low, because potential collaboration between airlines and airports –or, alternatively, airlines countervailing power– would put downward pressure on market power, and because concession revenues would induce the airports to charge less on the aeronautical side. The aim of this paper was to build an analytical model where these ideas could be tested and other insights gained, since most of the literature on airport privatization has been essentially descriptive and empirical analysis are unfeasible because of absence of real data.

A vertical setting was used to analyze airport privatization, both analytically and numerically. In the model, airports are input providers for the downstream airline market, in which airlines take airport prices and capacities as given. Our airline oligopoly model expanded on previous models on three aspects: it featured demand differentiation, schedule delay cost was included in the full price perceived by passengers, and had a particular emphasis on the importance of the number of airlines in the market. It was shown that these aspects have an important role on the incentives an airport has with respect to the dominance by a single airline. At the airport level, the results showed a rather unattractive picture for privatization when compared to both the first- and second best. First, the idea that low price elasticities of demand for airports would induce small allocative inefficiency failed to take into account the fact that observed elasticities may be poor forecasters of elasticities in other settings, and that capacity would be chosen by a private airport in a different way than a public airport. Our results showed that private airports would be much smaller than efficient public airports in terms of both traffic and capacities, which was reflected in important deadweight losses. Second, the arguments that airlines countervailing power or increased cooperation between airlines and airports may make regulation unnecessary are, most likely, overstated. The benchmark of maximization of joint profits showed, on one hand, that airports exerting vertical control on airlines (two-part tariffs in this model is enough) leads to the same outcome. More importantly, while the vertical double-marginalization problem is solved and the incentives for investment in capacities are better aligned, competition at the airline level is destroyed. So, while the outcome is indeed better in terms of traffic, capacities and social welfare, it is still closer to the pure private case than to the public one. It seems bold to conclude from here then, that regulation is unnecessary, especially because any implementation problem, which would only worsen the outcome, was assumed away. The analysis of budget adequacy showed that two-part tariffs alone may not be enough to avoid deficit of public airports. However, cost recovery two-part tariffs fell very close to the first-best. In the event that two-part tariffs were unfeasible, we found that a cost recovery linear price (i.e. a Ramsey price) would still lead to a superior outcome. Finally, it was shown that things deteriorate further when privatization is done on an airport by airport basis rather than in a system, because airports' demand complementarities induce horizontal double marginalization problems. These arise with simple linear prices, two-part tariffs, and when airports strategically collaborate with airlines.

We note that our model and its results apply to many other cases such as other transportation terminal (seaports; container terminals), railroad tracks and any vertical

setting where upstream quality (here measured by capacity) matters, although we resorted to a number of simplifications to preserve tractability. A few that could be relaxed in future research are: the fixed-proportions assumption, absence of concession revenues, continuous capacity and a single demand period (several independent periods is a trivial expansion so the relevant case to study would be interdependent periods). Another important simplification was assuming (round trip) travel between only two airports (notwithstanding that the literature usually focuses on a single airport's decisions). It allowed us to abstract from airlines' route structure choices, an endogenous decision which is central for cost minimization and strategic aspects of competition (See Oum et al, 1995; Jara-Díaz and Basso, 2003). The assumption has a direct effect on the analysis of privatization: the two airports in this paper do not face any kind of competition. In fact, we presented what can be seen as the worst case scenario, social welfare wise, for private airports: the two airports have demands that are perfect complements. Real competition between airports can emerge in two ways though. First, there may be Geographic Competition; airports in the same city area -such as the three San Francisco Bay area airports-compete for consumers in the same origin. Second, there may be competition for connecting passengers. When there is a network of airports (three or more distinct origin-destinations pairs), airlines can partly offset airports' market power through routing, something that would be taken into account by private airports when making decisions. Modeling these two types of competition seems to us the most important directions for future research, albeit they are complex ones. In our opinion, only with results from such models at hand, we will have a better and more complete picture about the economic effects of privatization and how we should go about regulation.

4.6 References

- Albers, S. Koch, B. and Ruff, C. (2005) Strategic Alliances between Airlines and Airports -theoretical assessment and practical evidence, *Journal of Air Transport Management*, 11, 49-58.
- Basso, L.J. (2006) On Input Markets Surplus and its Relation to the Downstream Market Game, working paper, Sauder School of Business, The University of British Columbia.
- Beesley, M.E. (1999) Airport Regulation, in *Regulating Utilities: a New Era?*, Beesley, M.E. ed., Institute of Economic Affairs, London.
- Borenstein, S. (1992) The Evolution of US Airline Competition, Journal of Economic Perspectives, 6, 45-73.
- Brander, J. and Zhang, A. (1990) Market Conduct in the Airline industry: an Empirical Investigation, *Rand Journal of Economics*, 21, 567-583.
- Brueckner, J.K. and Spiller, P.T. (1991) Competition and Mergers in Airline Networks, International Journal of Industrial Organization, 9, 323-342.
- Brueckner, J.K. (2002) Airport Congestion when Carriers Have Market Power. American Economic Review, 92, 1357-1375.
- Carlin, A. and Park, R.E. (1970) Marginal Cost pricing of Airport Runway Capacity. American Economic Review, 60, 310-319.
- Chen, Z. and Ross, T.W. (2003) Cooperating Upstream while Competing Downstream: a Theory of Input Joint Ventures, *International Journal of Industrial Organization*, 21, 381-397.
- Civil Aviation Authority (2004) Airport Regulation: Looking to the Future, Learning from the Past. CAA, London.
- Condie, S. (2000) Whither Airport Regulation? in *Privatization and Deregulation of Transport*, Bradshaw, W. and Smith, H.L. eds., St. Martin's Press, Inc, New York, NY, US. 364-393.
- Craig, E.D. (1996) The Benefits and Costs of Airport Privatization, in *Privatizing Transportation Systems*, Hakim, S., Seidenstat P. and Bowman, G. eds, Praeger Publishers, Westport US. 87-96.

Doganis, R. (1992) The Airport Business, Routledge, London.

- Douglas, G. and Miller III, J. (1974) Quality Competition, Industry Equilibrium and Efficiency in the Price-Constrained Airline Market, *The American Economic Review*, 64, 657-669.
- Forsyth, P. (1997) Price Regulation of Airports: principles with Australian Applications, *Transportation Research E*, 33, 297-309.
- Forsyth, P. (2003) Regulation under Stress: Developments in Australian Airport Policy, Journal of Air Transport Management, 9, 25-35.
- Fudenberg, D. and Tirole, J. (1984) The Fat-cat Effect, the Puppy-dog Play, and the Lean and Hungry Look, American Economic Review (Papers and Proceedings), 74, 450-467.
- Horonjeff, R. and McKelvey, F.X. (1983) *Planning and Design of Airports*, third edition, McGraw-Hill.
- Jacobsen, S.E. (1979) On the Equivalence of Input and Output Market Marshallian Surplus Measures. *American economic Review*, 69, 423-428.
- Jara-Díaz, S.R. and Basso, L.J. (2003) Transport Cost Functions, Network Expansion and Economies of Scope, *Transportation Research E*, 39(4), 271-288.
- Laffont, J.J. and Tirole, J. (2000) Competition in Telecommunications, Cambridge, MA: MIT Press.
- Morrison, S.A. (1987) The Equity and Efficiency of Runway Pricing. Journal of Public Economics, 34, 45-60.
- Morrison, S.A. and Winston, A. (1989) Enhancing the Performance of the Deregulated Air Transportation System. *Brookings Papers on Economic Activity: Microeconomics*, 1989, 61-112.
- Oum, T., Zhang, A. and Zhang, Y. (1993) Inter-Firm Rivalry and Firm Specific Price Elasticities in Deregulated Airline Markets, *Journal of Transport Economics and Policy*, 27, 171-192.
- Oum, T., Zhang, A. and Zhang, Y. (1995) Airline Network Rivalry, Canadian Journal of Economics, 4, 836-857.
- Oum, T., Zhang, A. and Zhang, Y. (2004) Alternative forms of Economic Regulation and their Efficiency Implications for Airports. *Journal of Transport Economics and Policy*, 38, 217-46.
- Oum, T. and Yu, C. (1997) Winning Airlines: Productivity and Cost Competitiveness of the World's Major Airlines, Kluwer Academic Publishers.

- Pels, E. and Verhoef, T. (2004) The Economics of Airport Congestion Pricing, Journal of Urban Economics, 55, 257-277.
- Productivity Commission (2002), Price Regulation of Airport Services, Report No. 19, AusInfo Canberra, Australia.
- Quirmbach, H.C. (1984) Input Market Surplus: The Case of Imperfect Competition, *Economic Letters*, 16, 357-362.
- Raffarin, M. (2004) Congestion in European Airspace: A Pricing Solution? Journal of Transport Economics and Policy, 38, 109-126.
- Shapiro, C. and Willig, R. (1990) On the Antitrust Treatment of Production Joint Ventures, Journal of Economic Perspectives, 4, 113-130.
- Small, K. (1992) Urban Transportation Economics, Harwood Academic Publishers.
- Spence, M. (1975) Monopoly, Quality and Regulation, *The Bell Journal of Economics*, 6, 417-429.
- Starkie, D. (2000) Allocating Airport Slots: A Role for the Market? in *Privatization and Deregulation of Transport*, Bradshaw, W. and Smith, H.L. eds., St. Martin's Press, Inc, New York, NY, US, 352-363.
- Starkie, D. (2001) Reforming UK Airport Regulation, Journal of Transport Economics and Policy, 35, 119-135.
- Starkie, D. (2005) Making Airport Regulation less Imperfect, Journal of Air Transport Management, 11, 3-8.
- US Federal Aviation Administration (1969) Airport Capacity Handbook, second edition, US Government Printing Office, Washington DC.
- Vives, X. (1985) On the Efficiency of Bertrand and Cournot Equilibria with Product Differentiation, *Journal of Economic Theory*, 36, 166-175.
- Zhang, A. and Zhang, Y. (1997) Concession Revenue and optimal Airport Pricing, *Transportation Research E*, 33, 287-296
- Zhang, A. and Zhang, Y. (2003) Airport Charges and Capacity Expansion: effects of concessions and privatization, *Journal of Urban Economics*, 53, 54-75

5 SEQUENTIAL PEAK-LOAD PRICING IN A VERTICAL SETTING: THE CASE OF AIRPORTS AND AIRLINES¹

5.1 Introduction

During the last several years airlines and passengers have been suffering from congestion and delays at the majority of busy airports, and airport delays have become a major public policy issue. Since the early work of Levine (1969), Carlin and Park (1970) and Borins (1978), economists have approached airport runway congestion by calling for the use of price mechanism, under which landing fees are based on a flight's contribution to congestion. While congestion pricing is economically desirable in that it would induce a better use of existing runway capacity and hence reduces congestion, it has not really been practiced. The existing landing fees depend on aircraft weight, and the fee rates are based on the accountancy principle of cost recovery required usually for a public enterprise.² Airports have traditionally been owned by governments, national or local. This is changing, however. Starting with the privatization of seven airports in the UK to BAA plc. in 1987, many airports around the world have recently been, or are in the process of being, privatized.³ One of the leading arguments for airport privatization is that privatised airports might well shift toward peak-load congestion pricing of runway services they provide to airlines, thus reducing delays in peak travel times (Poole, 1990; Gillen, 1994; Vasigh and Haririan, 1996). For example, Gillen (1994) argues that privatization does a better job of producing efficient runway pricing mechanisms compared to public ownership.

Taken together, today's shortage of airport capacity has revived much of the recent discussions about peak-load congestion pricing and airport privatization. In this paper we carry out an analysis of peak-load congestion pricing for a private, unregulated airport, as well as for a public airport that maximizes social welfare. The comparison of the two cases then allows us to shed some light on their pricing policies and traffic allocations to the peak and-off peak periods. We find that compared to a public, welfare-maximizing airport, a private, profit-maximizing airport would charge both higher peak and off-peak runway prices, as well as a higher peak/off-peak price differential. As a consequence, privatisation

¹ A version of this chapter has been submitted for publication. Basso, L.J. and Zhang A. (2006b) Sequential Peak-Load Pricing on a Vertical Setting: the case of airports and airlines.

² Airport charges include landing fees, aircraft parking and hangar fees, passenger terminal fees and air traffic control charges (if the service is provided by the airport authority), with landing fees being most dominant. The revenues derived from these charges are referred to as aeronautical revenues. In addition, busy airports derive significant revenues from non-aeronautical business, such as concessions and other commercial activities. As Daniel (2001) pointed out, landing fees in the U.S. traditionally recovered the "residual" costs –those remaining after all other revenue sources are full exploited, with the fee rate equalling the annual residual costs divided by the weight of all aircraft landing during the year.

³ A number of major airports in Europe, Australia and New Zealand have recently been privatized, and airports in several Asian countries are in the process of being privatized. In the U.S., on the other hand, the airports that are used by scheduled airlines are virtually all publicly owned facilities run by an agency on behalf of the state or local government. No major U.S. airports have been privatized to date. Canada may represent a middle-of-the-road case in which airports recently devolved from direct Federal control to become autonomous entities and major airports, though still government-owned, are now managed by private not-for-profit (but subject to cost recovery) corporations.

would lead to both fewer total air passengers and fewer passengers using the premium peak hours of the day for their travel, both of which reduce social welfare. Although those passengers who still use the peak period benefit from privatization, owing to less congestion delays, overall it is not economically efficient to have such a lower level of peak congestion, suggesting that airport privatization cannot be judged based on its effect on congestion delays alone. Our analysis also shows that whilst private, profit-maximizing airports will always use peak-load pricing, somewhat surprisingly, a public airport may actually charge a peak price that is *lower* than the off-peak price. Here the public airport, on the surface, is not practicing the peak-load pricing, but such pricing structure is nevertheless socially optimal.

We further investigate a case where a private airport strategically collaborates with the airlines so that it maximizes the joint airport-airline profits, since it has been often argued that greater airlines' *countervailing power* or more strategic collaboration between airlines and airports may improve efficiency of privatized airports by allowing a better alignment of incentives. The analysis shows that while its pricing practices would induce a collusive outcome in the airline market, it would induce greater total traffic and greater peak traffic than a pure (no-collaboration) private airport. Nevertheless, both figures will still be smaller than those for a public airport.

As indicated above, the present paper investigates airport peak-load pricing (PLP) and analyzes both the price level and price structure (peak vs. off-peak). This is in contrast to the majority of airport pricing studies which did not address inter-temporal pricing across different travel periods. In these congestion-pricing studies, there is only one demand function (i.e., a single-period model) for the airport, which is obtained by aggregating the demands of many agents -in this case, the airlines. Since airport runways are congestible, when an airline decides to schedule a new flight, it induces extra-delays on every other flight. The airline however would only internalize the delays it imposes on its own flights and not others. Congestion pricing then looks at the price the airport, or a regulatory authority, should charge to the airline for the new flight, in order for the airline to internalize all the congestion it produces (e.g., Morrison, 1987; Zhang and Zhang, 2003, 2005; Pels and Verhoef, 2004; Basso, 2005). Notice that under congestion pricing, since time-varying congestion is absent, there is only one way for either the airport or the airlines to internalize congestion: raising prices to suppress the demands. In a PLP framework, on the other hand, excess demand problems arise because of the variability of demands during the reference times of the day. If the same price was charged throughout the day, there would be peak periods at which the demand would be much higher than at off-peak periods. Peak-load pricing looks at the optimal time-schedule of prices so as to flatten the demand curve and make better use of existing capacity. As discussed below, both airports and airlines may engage in such demand spreading by using peak-load pricing. Note that in this PLP framework, the airport is still a congestible facility, which implies that in the resulting optimal price-schedule, prices at peak periods would still have to correct for uninternalized congestion: peak-load prices will have a congestion pricing component.⁴

⁴ As demonstrated in the text, the PLP-congestion pricing distinction is also important in that a singleperiod congestion toll is not optimal unless it is charged on top of the optimal charge in the off-peak period,

Another major feature of our analysis lies in the basic model structure used, which has strong implications for peak-load pricing. Here an airport, as an input provider, makes its price decisions prior to the airlines' output decisions. This vertical structure gives rise to *sequential* PLP: The PLP schemes implemented by the downstream airlines induce a different periodic demand for the upstream airport, with the shape of that demand depending on the number of downstream carriers and the type of competition they exert. The airport then would have an incentive to use PLP as well, which in turn affects the way the downstream firms use PLP. Although several very useful models of airport peak-load congestion pricing have been developed (e.g., Morrison, 1983; Morrison and Winston, 1989; Oum and Zhang (1990); Arnott, de Palma and Lindsey, 1993; Daniel, 1995, 2001), these studies considered PLP primarily at the airport level. Further, most of these studies considered only a public airport that maximizes social welfare, making no assessments about the effects of privatization on airport price structures.

There is an extensive body of literature on peak-load pricing. The classical papers (Boiteux 1949; Steiner 1957; Hirschleifer 1958; Williamson 1966) focused on normative rules for pricing a public utility's non-storable service subject to periodic demands. Some of the usual assumptions were: (i) demand is constant within each pricing period; (ii) demand in one period is independent of demand in other periods; (iii) constant marginal costs; (iv) the length of pricing periods is exogenous; (v) the number of pricing periods is exogenous; and (vi) peak time is known. Many authors have since contributed to the generalization of PLP results by relaxing one or a group of these assumptions, including Pressman (1970), Panzar (1976), Dansby (1978), Craven (1971, 1985), Crew and Kleindorfer (1986, 1991), Gersten (1986), De Palma and Lindsey (1998), Dana (1999), Laffont and Tirole (2000), Shy (2001) and Calzada (2003).⁵ However, the case of sequential peak-load pricing, be it for public or private utilities, has yet been analyzed. In the telecommunications research, for instance, Laffont and Tirole looked at PLP only at the upstream level (the network access charge) whilst Calzada considered PLP only at the downstream level. Because of this, we think our paper could be a contribution to the general peak-load pricing literature as well.

The paper is organized as follows. In the next Section we set out the basic model. In Section 5.3 we analyze and characterize the output-market equilibrium, paying particular attention to the peak and off-peak derived demands for airport services. Section 5.4 examines the airport's decisions and discusses how the airport ownership influences the peak and off-peak runway prices, traffic volumes, delays and welfare. Section 5.5 examines the case where a private airport maximizes the joint airport-airline profits. Section 5.6 contains concluding remarks.

which may not be the marginal cost. In other words, restricting the analysis to the toll that should be charged during the peak hours offers only a partial view of the problem.

⁵ See Crew, Fernando and Kleindorfer (1995) for an excellent survey of the peak-load pricing literature.

5.2 The model

We consider a two-stage model of airport and airline behavior, in which N air carriers service a congestible airport. In the first stage the airport decides on its charges on airlines for runway services, and in the second stage each carrier chooses its output in terms of the number of flights.

There is a continuum of consumers labelled by θ . Denoting $B_h(\theta)$ the (gross) benefits for consumer θ from traveling in period *h*, consumers' utility function may be written as:

$$U(x, B_h(\theta), -D_h)$$

where x is a vector of products and D_h denotes the flight delay associated with travel in period h. We shall consider a discrete choice model in which the consumer chooses between three mutually exclusive alternatives, namely: h=p, peak period travel; h=o, off-peak period travel; and h=n, not traveling. We assume that for any given θ ,

$$B_{p}(\theta) > B_{o}(\theta) > B_{n}(\theta) = 0$$
(5.1)

where $B_n(\theta)$ is, for convenience, normalized to be zero. The inequalities say that if travel was free and without congestion, the consumer would always prefer traveling to nontraveling. Furthermore, with identical airfares and delays, consumers would always prefer traveling in peak hours of a day to off-peak travelling. Thus, peak travel and off-peak travel are *vertically differentiated*: Controlling for fares and delay costs, passengers regard a peak flight as a better product than an off-peak flight. This vertical-differentiation feature of air travel can arise if the peak period represents the day's more desirable travel times than the off-peak period. Since people want to travel in those "popular" hours, the (unfettered) demand approaches or exceeds the capacity of the existing infrastructure, thereby resulting in (potential) congestion in the peak hours. Note that our formulation (5.1) is different from the one used in Brueckner (2002, 2005), in which he assumed a "single crossing property" in the sense that the benefit functions intersect at an intermediate value of θ , thus indicating that $B_p(\theta) > B_o(\theta)$ for large values of θ but $B_p(\theta) < B_o(\theta)$ for small values of θ . The single-crossing condition was imposed to avoid a degenerate (corner) equilibrium in his analysis.

As is usual in the discrete-choice models, we solve consumers' optimization problem in two steps:

$$\max_{x} \left\{ \max_{h \in p, o, n} U(x, B_{h}(\theta), -D_{h}) \right\}$$

s.t. $P_{x} \cdot x + t_{h} \leq I(\theta)$ (5.2)

where P_x is the vector price of products x, t_h is the ticket price (airfare) of travelling in period h, and $I(\theta)$ is consumer θ 's income. The first-step maximization leads to the

conditional Marshallian demands, $x^*(P_x, I(\theta) - t_h, B_h(\theta), -D_h)$. Replacing these demand expressions in U we obtain the conditional indirect utility function, $U(x^*, B_h(\theta), -D_h) \equiv V_h(P_x, I(\theta) - t_h, B_h(\theta), -D_h)$. For simplicity we assume that V_h is linear and further, $B_h(\theta)$ takes the simple linear form of $B_h \cdot \theta$, with B_h being constant, leading to:

$$V_{h} = \psi P_{x} + \lambda (I(\theta) - t_{h}) + B_{h} \cdot \theta - \alpha D_{h}$$

Condition (5.1) then implies $B_p > B_o > 0$, indicating that consumers differ in their travel benefits, with small θ indexing consumers with low benefit values. For simplicity, we assume θ is distributed uniformly on $\Theta = [\theta; \overline{\theta}]$ and normalize the number of total consumers to $\overline{\theta} - \underline{\theta}$, so the number of passengers with type belonging to $[\theta_1, \theta_2]$ is directly given by $\theta_2 - \theta_1$.

For the second-step maximization –comparisons of V_h for different h– we focus on the elements that determine the discrete choice (note that $\psi P_x + \lambda I(\theta)$ plays no role), obtaining a truncated conditional indirect utility function. Dividing by the marginal utility of income, λ , and redefining θ and α we then obtain:

$$V_h(\theta) = B_h \theta - \alpha D_h - t_h \tag{5.3}$$

In (5.3), while θ indexes consumers according to their (gross) travel benefits, the (positive) parameter α represents the subjective value of time savings and so αD_h represents monetary costs of delays to passengers. Note that our demand problem is identical to the one that will result if we fix θ but allow the value of time α to have a distribution among consumers (in simple models with endogenous hours of work, the consumers' "opportunity cost" of time lost in delays is proportional to their wages). One could also argue that θ and α are related (Yuen and Zhang, 2005), but we do not do this here.

As for the delay itself, the flight delay at period h, for h=p,o, may be given by $D_h = D(Q_h; L_h, K)$, where Q_h is the total number of flights in the period, L_h is the length (duration) of the pricing period, and K is the airport's runway capacity (measured in terms of the maximum number of flights that the airport's runways can handle per hour). In this paper we consider that K and L_h are exogenously given.⁶ We further assume L_o is sufficiently long so that $D(Q_o; L_o, K) = 0$ throughout the relevant range of our analysis. In other words, whilst the narrow peak period is congestible, congestion never arises in the broader off-peak period.⁷ For the peak delay function, we make the standard assumption that $D_p = D(Q_p)$ is differentiable in Q_p and

⁶ The case of variable and endogenous capacity is examined in Basso (2005) and Zhang and Zhang (2006) in a congestion-pricing framework.

⁷ This is similar to the two-period (peak/off-peak) formulation developed in Brueckner (2002).

$$D'_{p} = \frac{dD}{dQ_{p}} > 0, \qquad D''_{p} = \frac{d^{2}D}{dQ_{p}^{2}} \ge 0$$
 (5.4)

This assumption is quite general, requiring only that for given airport capacity, increasing peak traffic will increase congestion of the peak period and the effect is more pronounced when there is more congestion; that is, for given capacity, the peak delay is convex in traffic volume. The assumption is certainly satisfied under a linear delay function, $D(Q_p; L_p, K) = \delta \cdot Q_p / (L_p K)$, –which has been used by, e.g., Pels and Verhoef (2004)– or the functional form suggested by Lave and de Salvo (1968), $D(Q_p; L_p, K) = Q_p [L_p K(K - (Q_p / L_p))]^{-1}$.⁸

To obtain the consumer demands for peak and off-peak travel, we first note the following characteristics about the allocation of consumers: (i) if consumer θ_1 flies, then consumers $\theta \ge \theta_1$ fly; (ii) if consumer θ_1 does not fly, then consumers $\theta < \theta_1$ do not fly; and (iii) if θ^* denotes the consumer who is indifferent between traveling in the peak and off-peak periods, then passengers $\theta \ge \theta^*$ choose peak travel whereas passengers $\theta < \theta^*$ choose off-peak travel.⁹ Hence, if we denote θ^f the consumer who is indifferent between flying and not flying, (i), (ii) and (iii) above imply, in the case of an interior solution, that $\underline{\theta} < \theta^f < \theta^* < \overline{\theta}$. We assume interior allocations for now but later shall find conditions on the parameters for this to hold. Using q_h to denote the total number of passengers in period h(h=p,o), then $q_p = \overline{\theta} - \theta^*$ and $q_o = \theta^* - \theta^f$.

Since runway fees are imposed by aircraft (flight), we need to transform the passengerbased demands q_p and q_o into per-flight demand functions. As in Brueckner (2002), Pels and Verhoef (2004) and Basso (2005), we make a "fixed proportions" assumption that S =Aircraft Size × Load Factor, is constant and the same across carriers.¹⁰ It then follows immediately that $q_p = Q_p S = \overline{\theta} - \theta^*$ and $q_o = Q_o S = \theta^* - \theta^f$, or equivalently,

$$\boldsymbol{\theta}^* = \boldsymbol{\overline{\theta}} - \boldsymbol{Q}_p \boldsymbol{S}, \qquad \boldsymbol{\theta}^f = \boldsymbol{\theta}^* - \boldsymbol{Q}_o \boldsymbol{S} \tag{5.5}$$

⁹ Proof: (i) If θ_1 flies, then $\theta_1 B_h - \alpha D_h - t_h \ge 0$ for h=p,o. If $\theta \ge \theta_1$, $\theta B_h - \alpha D_h - t_h \ge \theta_1 B_h - \alpha D_h - t_h \ge 0$ and so θ flies. (ii) is analogous. (iii) Let $\Delta \overline{V}(\theta) \equiv \overline{V}_p(\theta) - \overline{V}_o(\theta) = (B_p - B_o)\theta - \alpha (D_p - D_o) - (t_p - t_o)$ and suppose θ flies. Then if $\Delta \overline{V}(\theta) \ge 0$, θ chooses to fly in the peak period. If $\Delta \overline{V}(\theta) < 0$, θ chooses to fly in the off-peak period. Now, suppose that there exists θ^* such that $\Delta \overline{V}(\theta^*) = 0$ (interior solution). Then it follows, since $\Delta \overline{V}'(\theta) > 0$, that if $\theta \ge \theta^*$, θ chooses peak travel and if $\theta < \theta^*$, θ chooses off-peak travel.

⁸ This functional form was previously estimated from steady-state queuing theory and is further discussed in U.S. Federal Aviation Administration (1969) and Horonjeff and McKelvey (1983). It has been used by, e.g., Morrison (1987), Zhang and Zhang (2003), and Basso (2005).

¹⁰ This assumption also allows us to abstract away from the issue of weight-based pricing as aircraft here have the same weight, and focuses on the main issue of peak-load congestion pricing.

In (5.5) the indifferent flyer θ^* is determined, using $D_o = 0$, by $\theta^*(B_p - B_o) - \alpha D_p = t_p - t_o$, which says that a passenger's gain $\theta^*(B_p - B_o)$, net of the delay cost αD_p when shifting from the off-peak to peak periods, is balanced by the fare differential $t_p - t_o$. Further, since the final flyer θ^f is determined by $\theta^f B_o = t_o$, (5.5) can be rewritten as:

$$t_o(Q_o, Q_p) = B_o \theta - B_o SQ_o - B_o SQ_p$$
(5.6)

$$t_p(Q_o, Q_p) = B_p \theta - B_o SQ_o - B_p SQ_p - \alpha D(Q_p)$$
(5.7)

Equation (5.6) is the (inverse) consumer demand function faced by the airlines for the offpeak period, whereas (5.7) is the consumer demand function for the peak period. Note that this demand system is not linear if D is not. Further, the peak and off-peak flights are substitutes for the final passengers, which gives the room for airlines to "spread the demand" across the peak and off-peak travel periods by using peak-load airfares. The analytical expressions of the cross elasticity of demand between peak travel and off-peak travel can be obtained from (5.6) and (5.7).

We now turn to the airlines. They have identical cost functions, given by:

$$c_{A}^{i}(Q_{h}^{i}, \mathbf{Q}_{h}^{-i}, P_{h}) = \sum_{h \in p, o} [c + P_{h} + \beta D(Q_{h})]Q_{h}^{i}$$
(5.8)

where Q_h^i is the number of airline *i*'s flights in period *h*, \mathbf{Q}_h^{-i} denotes the vector of flights of airlines other than *i*, *c* is the airline's operating cost per flight, and P_h is the airport landing fee in period h.¹¹ Further, parameter β (>0) measures the delay costs to an airline per flight, which may include wasted fuel burned while taxiing in line or holding/circling in the air, extra wear and tear on the aircraft, and salaries of flight crews. Airlines' profit functions can then be written as:

$$\phi^{i}(Q_{h}^{i}, \mathbf{Q}_{h}^{-i}, P_{h}) = \sum_{h \in p, o} t_{h}(Q_{o}, Q_{p})Q_{h}^{i}S - c_{A}^{i}(Q_{h}^{i}, \mathbf{Q}_{h}^{-i}, P_{h})$$
(5.9)

With these functions at hand, we shall investigate the subgame perfect equilibrium of our two-stage airport-airlines game.

¹¹ As indicated earlier, airport charges usually include landing and terminal charges (charges for aircraft parking are minor). While landing fees are based on aircraft movements, terminal charges are typically per-passenger based. Since the present paper is concerned with runway congestion, we shall focus on landing fees.

5.3 Analysis of output-market equilibrium

To solve for the subgame perfect equilibrium we start with the analysis of the second-stage airline competition. Given the airport's runway charges P_p and P_o , the N carriers choose their quantities to maximize profits, and the Cournot equilibrium is characterized by the first-order conditions,

$$\partial \phi^i / \partial Q_h^i = 0$$
, $h = p, o$

(note that the second-order conditions are satisfied).¹² Imposing symmetry $Q_h^i = Q_h / N$ and re-arranging, the first-order conditions can be expressed as:

$$\Omega^{o}(Q_{o}, Q_{p}, P_{o}, N) \equiv (B_{o}\overline{\theta}S - c - P_{o}) - Q_{o}\frac{B_{o}S^{2}(N+1)}{N} - Q_{p}\frac{B_{o}S^{2}(N+1)}{N} = 0$$
(5.10)

$$\Omega^{p}(Q_{o},Q_{p},P_{p},N) \equiv (B_{p}\overline{\theta}S - c - P_{p}) - Q_{o}\frac{B_{o}S^{2}(N+1)}{N} - Q_{p}\frac{B_{p}S^{2}(N+1)}{N} - (\alpha S + \beta)\left(D(Q_{p}) + \frac{Q_{p}}{N}D'(Q_{p})\right) = 0$$
(5.11)

As demonstrated in Appendix C.1, there exist conditions on the parameters that guarantee interior solutions, that is, $\underline{\theta} < \theta^f < \overline{\theta}^* < \overline{\theta}$ or equivalently, $Q_n, Q_n, Q_n > 0$. For example, the peak period is used if the per-passenger airport peak/off-peak price differential is smaller than the incremental gross benefit, for the highest consumer type $\overline{\theta}$, of shifting from off-peak travel to peak travel. In particular, when the airport does not practice peakload pricing (so $P_p = P_q$), the peak period is always used. The proof also reveals that a smaller airport peak/off-peak price differential increases the likelihood of both the peak and off-peak periods being used. The result suggests that Brueckner (2002, 2005)'s single crossing property, which was introduced to guarantee the existence of a non-empty peak/off-peak interior solution, may not be needed. This may be desirable because the property implies, if using Brueckner's interpretation of θ as an index of a passenger's tendency to travel on business (as opposed to leisure travel), that peak benefits are higher than off-peak benefits for business travelers, but are lower than off-peak benefits for leisure travelers. This appears contradictory with the idea that the peak and off-peak periods are vertically differentiated. In the remainder of the paper we shall restrict our attention to interior allocations.

A useful equation obtained from (5.10) and (5.11) is:

¹² We have assumed a Cournot game in the output-market competition. Brander and Zhang (1990, 1993), for example, find some empirical evidence that rivalry between duopoly airlines is consistent with Cournot behaviour.

$$-\Omega^{p} + \Omega^{o} = Q_{p} \frac{(B_{p} - B_{o})S^{2}(N+1)}{N} + (\alpha S + \beta) \left(D(Q_{p}) + \frac{Q_{p}}{N} D'(Q_{p}) \right) + (P_{p} - P_{o}) - \overline{\theta}S(B_{p} - B_{o}) = 0$$
(5.12)

Since equation (5.12) depends on Q_p but not on Q_o , it implicitly defines Q_p as a function of P_o , P_p and N. Substituting this function into (5.10), equation (5.10) then implicitly defines Q_o as a function of P_o , P_p and N, leading to:

$$Q_p = Q_p(P_o, P_p; N), \qquad Q_o = Q_o(P_o, P_p; N)$$
 (5.13)

which are the airport's demands for the use of its peak and off-peak periods, respectively. Here it is worth stressing that while $t_o(Q_o, Q_p)$ and $t_p(Q_o, Q_p)$, defined by equations (5.6) and (5.7), capture the *final* consumer (inverse) demands for air travel, $Q_o(P_o, P_p; N)$ and $Q_p(P_o, P_p; N)$ are the *derived* demands faced by the airport.

We now characterize the airport's demands $Q_p(P_o, P_p; N)$ and $Q_o(P_o, P_p; N)$. Totally differentiating (5.10) and (5.12) with respect to P_p yields:

$$\frac{\partial Q_p}{\partial P_p} = -\frac{\partial (-\Omega^p + \Omega^o)/\partial P_p}{\partial (-\Omega^p + \Omega^o)/\partial Q_p} = -\frac{N}{(B_p - B_o)S^2(N+1) + (\alpha S + \beta)((N+1)D'(Q_p) + Q_pD''(Q_p))} < 0$$
(5.14)

where the inequality follows from conditions (5.1) and (5.4). So the airport's demand for the peak period is, as expected, downward-sloping in peak charge. Similarly, we can obtain:

•

We can see that, *ceteris paribus*, the airport peak price does not influence total traffic but only the allocation of traffic to the peak and off-peak periods. Furthermore, from (5.12) and (5.14) we get:

$$\frac{\partial Q_p}{\partial \Delta P_{p-o}} = \frac{\partial Q_p}{\partial P_p} < 0$$
(5.16)

where $\Delta P_{p-o} \equiv P_p - P_o$. The above results (5.15) and (5.16) lead to:

Proposition 5.1: The airport's demands $Q_p(P_o, P_p; N)$ and $Q_o(P_o, P_p; N)$ have the following properties:

- (i) They are downward-sloping in own prices;
- (ii) The peak and off-peak periods are gross substitutes;
- (iii) The off-peak runway fee (P_o) determines the amount of total traffic, while the difference between the peak and off-peak charges (ΔP_{p-o}) determines the partition of that traffic into the two periods, with peak traffic declining with the charge differential.

Notice that Part (ii) of Proposition 5.1 shows that the airport has room to "spread the flights" across the peak and off-peak periods by using peak-load landing fees. Together with the discussion following equations (5.6) and (5.7), therefore, our vertical airport-airline structure gives rise to a possible *sequential* PLP: the PLP schemes implemented by the downstream airlines induce a different periodic demand for the upstream airport. As shown in Proposition 5.2 below, the shape of that demand depends on the number of downstream carriers and the type of competition they exert. Our analysis conducted later in Section 5.4 shows that indeed, the airport then would have an incentive to use PLP as well, which in turn affects the way the downstream firms use PLP.

Next we examine the airport demands change with the number of airlines, *N*. We obtain the following comparative static results:

Proposition 5.2: At the sub-game Cournot equilibrium,

- (i) $0 < \partial Q_p / \partial N < Q_p / (N(N+1))$, so the number of passengers traveling in the peak period increases with N;
- (ii) $\partial (Q_o + Q_p) / \partial N = (Q_o + Q_p) / (N(N+1)) > 0$, so the total number of passengers traveling increases with N;
- (iii) $\partial Q_o / \partial N > Q_o / (N(N+1))$, so if the off-peak period is used $(Q_o > 0)$ then the number of passengers traveling in the off-peak period increases with N.

The proof of Proposition 5.2 is given in Appendix C.2. Given that we consider interior solutions, conditional on runway fees P_p and P_o , both the peak and off-peak traffic volumes increase with the number of firms in the output market. The Proposition also shows that the (positive) elasticity of total traffic with respect to N is equal to 1/(N+1), whereas the elasticities of peak traffic and off-peak traffic (with respecto to N) are, respectively, smaller and larger than 1/(N+1). All three elasticities become smaller as N increases.

The final ingredient to characterize the Cournot equilibrium in the output market is related to the important issue of airfares: For given airport charges, how do the peak and off-peak airfares compare with each other? From (5.6) and (5.7) it follows that

$$\Delta t_{p-o} \equiv t_p - t_o = \overline{\theta}(B_p - B_o) - Q_p S(B_p - B_o) - \alpha D(Q_p)$$
(5.17)

From the equilibrium condition (5.12) we obtain an expression for $\overline{\theta}(B_p - B_o)$. Replacing that expression in (5.17) gives rise to the following airfare-differential formula, evaluated at the Cournot equilibrium:

$$\Delta t_{p-o}\Big|_{\text{Cournot eq}} = \frac{P_p - P_o}{S} + \frac{\beta}{S} D(Q_p) + \frac{\beta}{S} \frac{Q_p}{N} D'(Q_p) + \alpha \frac{Q_p}{N} D'(Q_p) + Q_p \frac{(B_p - B_o)S}{N}$$
(5.18)

It is clear from (5.18) that if $P_p \ge P_o$, then $\Delta t_{p-o} > 0$, that is, if the airport uses peak-load pricing, airlines will also use it in equilibrium. More interesting perhaps is the fact that, even if the airport price the periods *backwards*, i.e., $P_p < P_o$, airlines may still use peak-load pricing in equilibrium.

To further interpret (5.18), first note that holding P_p and P_o constant, $\partial \Delta t_{p-o} / \partial N$ is negative, which can be seen by differentiating (5.17) and recalling, from Proposition 5.2, that sub-game equilibrium Q_p and Q_o increase in N. This implies that a monopoly airline would have the largest airfare differential. Since, from (5.6), $\partial t_a / \partial N$ is also negative, the lower the N, the larger the off-peak fare. These two observations are consistent with what we have already shown in Proposition 5.2 with respect to total peak and off-peak traffic. Next, it can be seen that for very large N, the airfare differential approaches to the difference between an airline's peak and off-peak per-passenger average costs, i.e., the first and second terms on the right-hand side (RHS) of (5.18). When there is an oligopoly, however, three extra terms are added. Specifically, the third term on the RHS of (5.18) is the cost of extra congestion on an airline's own flights and caused by an additional passenger flying in the peak period. Thus, the first three terms on the RHS of (5.18) represent the difference between an airline's peak and off-peak marginal costs. The fourth term represents the money value of extra congestion to an airline's passengers when a new passenger chooses to fly in the peak period, whereas the fifth term is the mark-up term that arises from the oligopoly airlines' exploitation of market power. Hence, as it is now known, the airlines in oligopoly only internalize (charge for) the congestion they impose on their own flights, which has two cost components: extra operating costs for the airline, and extra delay costs for its passengers (Brueckner, 2002). When there is a monopoly airline, congestion is perfectly internalized but exploitation of market power is at its highest degree. When N is large, exploitation of the market power is small but congestion is imperfectly internalized.

These points can be made more clearly if the Cournot case is compared to the case in which a social planner maximizes total surplus in the second-stage of the game. To do this we first need a measure of consumer surplus (CS). Given the linearity of our conditional indirect utility function, CS is given by:

$$CS = \int_{\theta^{\star}}^{\overline{\theta}} \left[\theta B_p - \alpha D(Q_p) - t_p(Q_p, Q_o) \right] f(\theta) d\theta + \int_{\theta^{\star}}^{\theta^{\star}} \left[\theta B_o - t_o(Q_p, Q_o) \right] f(\theta) d\theta$$
(5.19)

where $f(\theta)$ is the density function. Using (5.6) and (5.7) for t_o and t_p , solving the integrals and replacing θ^* and θ^f with (5.5), we finally obtain:

$$CS = \frac{S^{2}}{2} \left(B_{o} Q_{o}^{2} + 2B_{o} Q_{o} Q_{p} + B_{p} Q_{p}^{2} \right)$$
(5.20)

We are then interested in the case where the planner maximizes, for given airport charges, the sum of consumer surplus and airline profits:

$$CS + \Phi \equiv CS + \sum_{i=1}^{N} \phi^{i}$$
(5.21)

where Φ denotes the aggregate airline (equilibrium) profits. The first-order conditions of (5.21) with respect to airline outputs, together with the imposition of symmetry, then lead to two equations, analogous to (5.10) and (5.11), which characterize the optimum. Subtracting the two equations from each other yields:

$$Q_{p}(B_{p} - B_{o})S^{2} + (\alpha S + \beta)(D(Q_{p}) + Q_{p}D'(Q_{p})) + (P_{p} - P_{o}) - \overline{\theta}S(B_{p} - B_{o}) = 0$$
(5.22)

Using (5.22) to obtain a new expression for $\overline{\theta}(B_p - B_o)$ and replacing the term in (5.17), we get:

$$\Delta t_{p-o}\Big|_{\text{efficient}} = \frac{P_p - P_o}{S} + \frac{\beta}{S} D(Q_p) + \frac{\alpha S + \beta}{S} Q_p D'(Q_p)$$
(5.23)

Conditional on the airport charges and the airline market structure, (5.23) gives the socially efficient difference between the peak and off-peak airfares. This fare differential is equal to the difference between an airline's peak and off-peak average costs (the first and second terms on the RHS of (5.23)), plus all the external costs associated with a new flyer in the peak period, with the latter being the extra congestion cost of *all* the airlines and passengers, not just that of the airline that carries the new peak passenger. Obviously, the last two terms represent the portion of the optimal airfare differential that is not directly affected by the airport's pricing practices.

It is also insightful -- and useful later in Section 5.5- to compare the Cournot case with the "cartel" case in which the airlines choose outputs to maximize their joint profit

 $\Phi = \sum_{i=1}^{N} \phi^{i}$. The first-order conditions in the cartel case, together with the imposition of symmetry, lead to two equations, analogous to (5.10) and (5.11), which characterize the cartel optimal solution. Subtracting the two equations then yields:

$$2Q_{p}(B_{p} - B_{o})S^{2} + (\alpha S + \beta)(D(Q_{p}) + Q_{p}D'(Q_{p})) + (P_{p} - P_{o}) - \overline{\theta}S(B_{p} - B_{o}) = 0$$
(5.24)

Using (5.24) to obtain a new expression for $\overline{\theta}(B_p - B_o)$ to replace the term in (5.17), we find the difference between the peak and off-peak airfares in the cartel case:

$$\Delta t_{p-o}\Big|_{\text{cartel output}} = \frac{P_p - P_o}{S} + \frac{\beta}{S}D(Q_p) + \frac{\alpha S + \beta}{S}Q_pD'(Q_p) + Q_p(B_p - B_o)S$$
(5.25)

Here, the fare differential is equal to the difference between an airline's peak and off-peak average costs (the first and second terms on the RHS of (5.26)), plus all the external costs associated with a new flyer in the peak period. Comparing (5.25) with (5.23) reveals that the cartel, as in the social-planner case, internalizes the congestion costs of all the carriers and passengers. Here however, there is a fourth term which will increase the fare differential. This term is related to the "business stealing" externality: Since oligopoly carriers behave in non-cooperative fashion, they produce too much with respect to the optimum for the airlines as a whole. This is so because they fail to consider the profits lost by the other airlines when they increase own output, depressing airfares: the fare differential of an oligopoly is insufficiently large from the cartel's point of view, and this problem worsens, the "looser" the oligopoly is (i.e., the larger N is). Consequently, the cartel, as a monopoly, is interested in having a less used peak period. In effect, the cartel airfare differential is identical to the monopoly's; see (5.18) by imposing N=1. The cartel and monopoly traffic volumes will differ however, since cost functions are convex and not flat.

5.4 Airport pricing, traffic, delay and welfare comparisons

We have shown that the airport decisions, namely, P_p and P_o , can influence the subsequent output-market competition among airlines. When deciding its runway fees in the first stage, therefore, the airport will take the second-stage equilibrium outputs into account. These decisions may in reality be set by a public airport or a privatized airport. Consequently, the objective of an airport may be to maximize social welfare or to maximize profit. In this Section, we compare airport charges and consequent airfares for these two airport types.

Consider first a private, unregulated airport. The airport's profit may be written as:

$$\pi(P_o, P_p; N) = P_o Q_o + P_p Q_p - C \cdot (Q_o + Q_p) - rK$$
(5.26)

where $Q_o = Q_o(P_o, P_p; N)$ and $Q_p = Q_p(P_o, P_p; N)$ are the airport's demands for the peak and off-peak periods, respectively, and are given by (5.13). In addition, *C* is the unit runway operating cost of the airport and *r* is its unit cost of capital. Note that we have assumed, as is common in the literature, that the operational and capital costs are separable and that the marginal operating cost is constant. For the latter, the estimation of cost function showed that airport runways have relatively constant return to scale (e.g., Morrison, 1983). The airport will choose P_p and P_o to maximize its profit, and the firstorder conditions lead to (P_o^{π}, P_p^{π}) denoting the profit-maximizing airport charges):

$$P_o^{\pi} - C = \frac{P_o^{\pi}}{\varepsilon_{oo}} + \frac{(P_p^{\pi} - C)Q_p\varepsilon_{po}}{Q_o\varepsilon_{oo}}$$
(5.27)

$$P_{p}^{\pi} - C = \frac{P_{p}^{\pi}}{\varepsilon_{pp}} + \frac{(P_{o}^{\pi} - C)Q_{o}\varepsilon_{op}}{Q_{p}\varepsilon_{pp}}$$
(5.28)

where $\varepsilon_{oo} \equiv -(\partial Q_o / \partial P_o)(P_o / Q_o)$ is the (positive) elasticity of airport demand, $\varepsilon_{po} \equiv (\partial Q_p / \partial P_o)(P_o / Q_p)$ is a cross-price elasticity, and ε_{pp} and ε_{op} are defined analogously. Since $\partial Q_p / \partial P_o > 0$ and $\partial Q_o / \partial P_p > 0$ -see (5.15) or Proposition 5.1– both ε_{op} and ε_{po} are positive, implying that the airport charges are higher than would be if the peak and off-peak charges were chosen independently (in which case the mark-ups would be proportional to the inverse of demand elasticities only). This is a well-known result for multi-product monopolies that produce substitutes.

We can simplify the pricing equations and show that $P_p^{\pi} > P_o^{\pi}$. To do this, replace the elasticities' definitions and simplify, using the fact that $\partial Q_o / \partial P_p = -\partial Q_p / \partial P_p$ in (5.15) and then using equation (5.14). We obtain the following charging formulas:

$$P_o^{\pi} = C + \frac{Q_o S^2 B_o (N+1)}{N} + \frac{Q_p S^2 B_o (N+1)}{N}$$
(5.29)

$$\Delta P_{p-o}^{\pi} = \frac{\alpha S + \beta}{N} Q_p \Big[(N+1)D'(Q_p) + Q_p D''(Q_p) \Big] + \frac{Q_p (B_p - B_o)S^2 (N+1)}{N}$$
(5.30)

The RHS of (5.30) is, by (5.1) and (5.4), positive and hence $P_p^{\pi} > P_o^{\pi}$: The private airport charges higher runway fees in the peak period than in the off-peak period, and this is true for any *N*. Thus, a profit-maximizing airport has an incentive to use peak-load pricing. Further, notice from (5.29) that the off-peak charge, which determines the amount of total traffic, is above marginal cost. This is a result of monopoly power on the part of the airport. There is, therefore, a "double marginalization" problem, which is typical of an uncoordinated vertical structure. The discussion leads to the following Proposition:
Proposition 5.3: A private, profit-maximizing airport would use peak-load pricing. Further, it would charge an off-peak runway fee that is above its marginal cost.

Next, consider a public airport that chooses P_p and P_o to maximize social welfare. In the present situation with three agents (airport, airlines, and passengers), social welfare (SW) is the sum of their payoffs:

$$SW(P_o, P_p; N) = \pi(P_o, P_p; N) + CS + \Phi$$
 (5.31)

where the airport's profit, π , is given by (5.26), consumer surplus, CS, is given by (5.19), and the airlines' profit, $\Phi = \sum_{i=1}^{N} \phi^{i}$, is introduced in (5.21). Since the downstream equilibrium is symmetric, $\phi^{i}(Q_{h}^{i}, \mathbf{Q}_{h}^{-i}, P_{h}) = \phi^{1}(Q_{o}(P_{o}, P_{p}; N), Q_{p}(P_{o}, P_{p}; N), P_{h})$ is each airline's equilibrium profit. We can then easily calculate Φ as $\Phi(P_{o}, P_{p}; N) = N \cdot \phi^{1}(P_{o}, P_{p}; N)$, that is,

$$\Phi(P_o, P_p; N) = \overline{\theta}S(B_pQ_p + B_oQ_o) - S^2(B_oQ_o^2 + 2B_oQ_oQ_p + B_pQ_p^2) - (\alpha S + \beta)Q_pD(Q_p) - (c + P_o)Q_o - (c + P_p)Q_p$$
(5.32)

We do not include a budget constraint in the public airport problem, noting that fixed fees may solve the problem of budget adequacy. If lump-sum transfers are not feasible, then Ramsey-Boiteux prices should be considered (see Basso, 2005, for more discussion on this).

Replacing Φ from (5.32) (and π , CS from their earlier equations) in the social-welfare function, we obtain:

$$SW = \overline{\theta}S(B_pQ_p + B_oQ_o) - c \cdot (Q_p + Q_o) - C \cdot (Q_p + Q_o) - rK -S^2(B_oQ_o^2 + 2B_oQ_oQ_p + B_pQ_p^2)/2 - (\alpha S + \beta)Q_pD(Q_p)$$
(5.33)

Derivation of pricing formulas follows from the first-order conditions. Specifically, we obtain, using equations (5.10), (5.12) and (5.15) (P_o^W, P_p^W denoting the welfare-maximizing runway fees):

$$P_{o}^{W} = C - \frac{Q_{o}S^{2}B_{o}}{N} - \frac{Q_{p}S^{2}B_{o}}{N}$$
(5.34)

$$\Delta P_{p-o}^{W} = \frac{N-1}{N} (\alpha S + \beta) Q_{p} D'(Q_{p}) - \frac{Q_{p} S^{2}(B_{p} - B_{o})}{N}$$
(5.35)

The above welfare-maximizing airport pricing may be seen as if the fees were determined in two phases. First, choice of an off-peak price P_a^W induces the (socially) right amount of total traffic; as can be seen from (5.34), P_o^W is below the airport's marginal cost. This is needed because exploitation of market power in the airline market would induce allocative inefficiencies by producing too little output. A welfare-maximizing airport fixes this inefficiency by providing a "subsidy" to the airlines and hence lowering their marginal costs in the off-peak period. The exact amount of the subsidy depends in part on the extent of market power, which here is captured by N. Once the total traffic is set to its optimal level, the next phase is concerned with the optimal allocation of this traffic to the peak and off-peak periods, which is, as indicated earlier, determined by ΔP_{p-o} . In particular, the public airport sets the peak/off-peak price differential to ΔP_{p-o}^W that will induce the optimal airfare differential downstream. This is apparent from substituting (5.35) into (5.18), which yields

$$\Delta t_{p-\sigma}^{W}\Big|_{\text{Cournot eq}} = \frac{\beta}{S} D(Q_p) + \frac{\alpha S + \beta}{S} Q_p D'(Q_p) > 0$$
(5.36)

The RHS of (5.36) is equal to the optimal airfare differential that is not directly affected by the airport's pricing practices, as discussed in (5.23). Hence, the outcome is the same as if the airport were to set $P_o = P_p$, which is optimal because there are no differences in costs, and then social welfare is maximized in the airline market.

It is worth examining further the welfare-maximizing charge differential, ΔP_{p-o}^{W} , given in (5.35). This differential is not signed a priori; hence, it may happen that the airport charge is smaller in the peak period than in the off-peak period. More specifically, the airport charge differential will be negative for small N. This is so because a "tight" airline oligopoly has an airfare differential that is too large due to strong market power, while congestion is reasonably internalized. As a consequence, the airport price differential is driven predominantly by the market-power effect (the second term on the RHS of (5.35)). When N is large, on the other hand, the airport price differential that is too small due to uninternalized congestion, whereas market power is relatively weak. The airport charge differential is then driven by the congestion effect (the first term in (5.35)). Note from (5.36), that although P_p^W (the welfare-maximizing peak charge) may be less than P_o^W , (final) passengers will, nevertheless, always pay higher peak airfare than off-peak airfare.

The above discussion may be summarised in the following Proposition:

Proposition 5.4: For a public, welfare-maximizing airport, (i) the off-peak runway fee is below its marginal cost; (ii) for small N, the off-peak runway fee may be *greater* than its peak runway fee; in this sense, it appears that the airport does not use peak-load pricing; (iii) although the airport's peak charge may be less than its off-peak charge, final passengers will nevertheless always pay higher peak airfare than off-peak airfare.

Brueckner (2002) identified the first term in (5.35) as the per-flight toll that should be charged by the airport authorities to address the problem of uninternalized congestion (note that when N=1, this toll is equal to zero). Pels and Verhoef (2004) and Basso (2005)

pointed out that the optimal toll should also include the second term, the market-power effect;¹³ they did this, however, using models of congestion pricing (one period), while Brueckner (2002) and the present paper use a model of peak-load pricing. This distinction is important because a toll equal to the two terms, thereby capturing both the congestion and market power effects, will not be optimal unless it is charged on top of the optimal charge in the off-peak period, *which is not the marginal cost*. In other words, restricting the analysis to the toll that should be charged during the peak hours offers only a partial view of the problem.

Note that if lump-sum transfers (two-part tariffs) are unfeasible, the pricing rules previously discussed may lead to airport's budget inadequacy. If budget adequacy has to be ensured but lump sums are not feasible, then the first best may not be attainable: Marginal prices would have to do both –namely, aligning incentives and transferring surplus– making the airport fall short of "control instruments" (Mathewson and Winter, 1984).

Having derived and characterized the pricing structures for both the public and private airports, we now want to compare them. To have a clearer picture of the differences in performances, we shall compare not only the off-peak runway fees and the peak/off-peak fee differentials, but also the induced traffic levels, delays and total surplus levels. Moreover, we want to assess how these differences (if any) change with the number of airlines, N, which is exogenously given and may be used as a proxy for airline market structure. We summarize our findings in the following Proposition (the proof is provided in Appendix C.3):

Proposition 5.5: Comparisons of airport pricing, traffic, delay and welfare between the private and public airports are as follows:

(i)
$$P_o^W < P_o^{\pi}$$
 and $\frac{dP_o^W}{dN} > \frac{dP_o^{\pi}}{dN} = 0$;

(ii)
$$\Delta P_{p-o}^{W} < \Delta P_{p-o}^{\pi}$$
 and $\frac{d\Delta P_{p-o}^{\pi}}{dN} > 0$. If the delay function is linear, then $\frac{d\Delta P_{p-o}^{\pi}}{dN} = 0$;

(iii)
$$Q_p^W > Q_p^\pi$$
 and $\frac{dQ_p^\pi}{dN} > \frac{dQ_p^\pi}{dN} = 0;$

· ~ #

I O W

(iv)
$$Q_t^W > Q_t^{\pi}$$
 and $\frac{dQ_t^{\pi}}{dN} > \frac{dQ_t^{''}}{dN} = 0$, where $Q_t \equiv Q_p + Q_o$ denotes total traffic volume;

(v)
$$D_p^W > D_p^\pi$$
 and $\frac{dD_p^\pi}{dN} > \frac{dD_p^W}{dN} = 0;$

(vi)
$$SW^{W} > SW^{\pi}$$
 and $\frac{dSW^{\pi}}{dN} > \frac{dSW^{W}}{dN} = 0$.

¹³ To be fair, although Brueckner did not formally consider the second term in the toll to be charged, he did point out that, depending on the size of the market-power term, a pure congestion toll could be detrimental for social welfare.

To help better understand this Proposition we also offer a schematic representation of the findings in Figure 5.1. From Proposition 5.5 we see that a private, profit-maximizing airport would induce too small total traffic as compared to the first-best outcome, thereby resulting in allocative inefficiencies. Additionally, a private airport has a greater peak/offpeak runway charge differential than a public airport. Hence, with a private airport, the peak period would be underused not only because the airport has smaller total traffic, but also because its charge differential is too large. This reduction in peak traffic volume is welfare-reducing in that passengers view traveling in the peak times as, other things being equal, a higher quality product than traveling in the off-peak times. Less peak traffic then means fewer consumers will enjoy a premium product. And although those passengers who still use the peak period benefit from less congestion delays as part (v) enounces, overall it is not economically efficient to have such a lower level of peak congestion because total welfare is in fact reduced as shown in part (vi) of Proposition 5.5. In effect, in terms of overall welfare level, a numerical simulation has shown that $SW^{\pi}(N=1) = 49\%$, $SW^{\pi}(N=3) = 65\%$, $SW^{\pi}(N=5) = 69\%$ and $SW^{\pi}(N=50) = 77\%$, where the percentages are par rapport to the public case, that is, $SW^{W} = 100\%$ (as can be seen from part (vi), SW^w does not depend on N).¹⁴ This is an important point: one of the main ideas behind airport privatization has been that it would allow airports to use peak-load pricing and thus help solve the congestion problems. But if privatization is measured solely by its effect on congestion delays, it may be seen as a better idea than it actually is, and important deadweight losses may be overlooked. This result, which holds here for a fixed capacitypeak-load pricing model, was also found by Basso (2005) in a variable (endogenous) capacity-congestion pricing model.



Figure 5.1: Schematic representation of the results in Proposition 5.5

We have seen that the public airport is indifferent between values of N –although Basso (2005) showed that this may not be the case if airlines are not homogenous or if passengers are affected by schedule delay cost. Given that the (welfare) performance of a private airport improves as the number of airlines rises (see Proposition 5.5 and Figure 5.1), it

¹⁴ The parameters used in the simulation are presented in Table c.1, in Appendix C.4

seems important to know what would be the preferred N of a private airport. From (28) we have:

$$\left|\frac{d\pi}{dN}\right|_{P_o^{\pi}, P_p^{\pi}} = \left|\frac{\partial\pi}{\partial N}\right|_{P_o^{\pi}, P_p^{\pi}} = \left(P_o^{\pi} - C\right)\frac{\partial Q_o}{\partial N} + \left(P_p^{\pi} - C\right)\frac{\partial Q_p}{\partial N} > 0$$

where the first equality follows from the envelope theorem and the inequality follows from Proposition 5.2 and the first-order conditions (5.27) and (5.28), which indicate that prices are above marginal costs. Thus, the private airport prefers a large N, which is a desirable property, given the findings of Proposition 5.5.

5.5 Airport-airline joint profit maximization

In this Section, we shall consider an airport that has some sort of strategic agreement with the airlines using it. The reasons why it is interesting to look at this case are two-fold: on one hand, a simple pricing mechanism, two-part tariff, may be enough for the outcome of joint profit maximization to arise. On the other hand, it has been often argued that greater airlines' *countervailing power* or more strategic collaboration between airlines and airports may improve efficiency of privatized airports by allowing a better alignment of incentives, and even may make price regulation unnecessary (see, e.g., Beesley, 1999; Condie 2000; Forsyth, 1997; Starkie, 2001; Productivity Commission, 2002; Civil Aviation Authority UK, 2004). The analysis of joint profit maximization may then serve as a benchmark case.

The objective faced by this airport is to maximize the sum of the airport's profit and airlines' profits. Using π in (5.26) and Φ in (5.32), the problem can be re-written as:

$$\max_{P_o, P_p} \pi + \Phi = \overline{\theta} S(B_p Q_p + B_o Q_o) - c \cdot (Q_p + Q_o) - C \cdot (Q_p + Q_o) - rK$$
$$-S^2(B_o Q_o^2 + 2B_o Q_o Q_p + B_p Q_p^2) - (\alpha S + \beta)Q_p D(Q_p)$$

Derivation of the pricing formulas follows from the first-order conditions of the above problem, using equations (5.10), (5.12) and (5.15) and rearranging (P_o^{JP}, P_p^{JP}) denoting the joint profit-maximizing airport charges),

$$P_o^{JP} = C + \frac{Q_o S^2 B_o (N-1)}{N} + \frac{Q_p S^2 B_o (N-1)}{N}$$
(5.37)

$$\Delta P_{p-o}^{JP} = \frac{N-1}{N} (\alpha S + \beta) Q_p D'(Q_p) + \frac{Q_p S^2 (B_p - B_o)(N-1)}{N}$$
(5.38)

The interpretation of the JP airport's pricing rules (5.37) and (5.38) is as follows. As before, this airport may be seen as deciding its runway fees in two phases; first, it induces a contraction of total traffic by choosing an off-peak price P_a^{JP} above its marginal cost. It

does so because the failure of coordination among the airlines results in them producing too much with respect to what would be best for them as a whole. The amount of excess production depends on how tight the oligopoly is, which is why the off-peak mark-up decreases with N. In particular, when the airline market is monopolized, (5.37) shows that $P_o^{JP} = C$ so the airport does not need the mark-up at all. Note, however, that the total traffic contraction in the JP case is smaller than that in the pure private case.

In the second phase, the airport chooses the (non-negative) price differential ΔP_{p-o}^{JP} that will induce the airlines' cartel outcome, destroying airline competition downstream. This is apparent from substituting (5.38) into (5.18), and noting that the result is equal to the cartel's airfare differential not directly affected by the airport pricing practice. Hence, the outcome is the same as if the airport were to set $P_{q} = P_{p}$ and a cartel were running the airline market This result, which was obtained by Basso (2005) in a congestion pricing setting, has different intuitions depending on why the maximization of joint profits is the relevant case. With two-part tariffs, the private airport use the variable prices, peak and offpeak, to destroy competition downstream in order to maximize the profits of airlines, which are later captured by the airport through the fixed fee. When the joint profit maximization arises because of collaboration between airlines and the airport, what happens is that the airlines would like to collude in order to increase profits, but are unable to do so themselves because of the incentives to defect on any possible agreement. What they manage to do, however, is to "capture" an input provider to run the cartel for them. By altering the prices of the inputs (runway services) and hence the downstream marginal costs in both the peak and the off-peak periods, the input provider (airport) induces both the collusive total output and the "right" (to the airlines) allocation of passengers to the peak and off-peak periods. The upstream firm is then rewarded with part of the collusive profits, which is where bargaining power enters the picture.

Note also that the airport pricing rules (5.37) and (5.38) take into account both the congestion externality and the business-stealing externality, at both pricing phases: the airport's price differential has two parts.¹⁵ When N=1, there is no business-stealing effect and congestion is perfectly internalized by the monopolist. Consequently, both terms vanish: with a monopoly airline, the airport will not use peak-load pricing.

Now, despite the fact that the result is as if airlines were colluding, this case is not worse, in terms of social welfare, than a private airport charging linear prices as before. This is because, here, the two other harmful externalities, namely, the vertical double marginalization and the congestion externality, have been dealt with. In effect, we can show that the JP case (where the airport has strategic agreements with airlines) represents a

¹⁵ This idea of an upstream firm running the cartel for the downstream firms has been discussed in the vertical control literature and, particularly, in the input joint-venture case. For example, Chen and Ross (2003) formalized the conjecture that input joint-ventures can facilitate collusion and push a market toward the monopoly outcome. If airport provision was seen as an input joint-venture by the airlines, our results show three things in addition to what Chen and Ross have found. First, the results hold even in a peak-load pricing setting, i.e., when demand is periodic. Second, if there are externalities, the input prices are, additionally, used to force their internalization by downstream competitors. Third, when marginal costs downstream are not constant, the outcome is not as in a monopoly or a downstream merger, but as in a cartel.

middle-of-the-road case: in Proposition 5.5, the runway fees, traffic volumes, delays and social welfare will be in between those of the private and public airport cases. And in Figure 5.1, the curves pertaining to the JP case would be parallel displacements of the public airport curves, lying in between the two existing public and private curves. Strategic collaboration between the airport and the airlines smoothes the airport-charge problem. But recall that the downstream airfares would be as if the airlines were engaging in collusion, so we cannot expect to end up too close to the first best. In effect, a numerical simulation has shown that deadweight losses would correspond to 22% of the maximum social welfare attainable.

5.6 Concluding remarks

In this paper, we have analyzed the sequential peak-load pricing (PLP) problem that arises when the airport is recognized as an input provider for a final consumer market facing a periodic demand. We have analyzed this PLP problem for a private unregulated airport, for a public airport that maximizes social welfare, and for an airport that strategically collaborates with the airlines and hence maximizes their joint profits. We found that privatization would not induce efficient peak-load pricing schemes as it has been argued in some studies. While a private airport has always an incentive to use PLP -higher runway fees in the peak than off-peak periods, even when the airlines have used PLP themselves and irrespective of the number of airlines servicing the airport- its pricing structure would induce insufficient total traffic and insufficient peak traffic. Somewhat surprisingly, depending on the degree of market power (captured here by the number of carriers at the airport), a public airport may choose a peak runway charge that is lower than the off-peak charge, so as to offset the market power downstream at the airline level. Here, the public airport, on the surface, is not practicing the peak-load pricing, but such pricing structure is nevertheless socially optimal. Finally, a private airport that strategically collaborates with the airlines would induce greater total traffic and greater peak traffic than a pure private airport, but both figures will still be smaller than those for a public airport. If the airport collaborates with a monopoly airline, it would not use peak-load pricing.

Although the airline industry is chosen for analysis, our basic model structure, in which airports, as input providers, make their pricing decisions prior to airlines' strategic interactions in the final output market, is highly relevant to several other industries including electricity, telecommunications, and transport terminals (e.g., the vertical chain of ports-sea carriers-shippers). In telecommunications, for example, at the upstream level there are the network owners, while downstream there are carriers who must use the network in order to produce the final good (telephone calls). Like airports, these industries are undergoing privatization in a number of countries. We note that the sequential PLP method used in the present paper may be useful in examining similar issues in those sectors as well.

5.7 References

- Arnott, R., de Palma A. and Lindsey, R. (1993) A Structural Model of Peak-Load Congestion: A Traffic Bottleneck with Elastic Demand, American Economic Review, 83, 161-179.
- Basso, L. J. (2005) Airport Ownership: Effects on Pricing and Capacity, Working Paper, The University of British Columbia (<u>http://ssrn.com/abstract=849584</u>).
- Beesley, M.E. (1999) Airport Regulation, in: M.E. Beesley (ed.), *Regulating Utilities: a New Era?*, Institute of Economic Affairs, London.
- Boiteux, M. (1949) La Tarification des Demandes en Point: Application de la Theorie de la Vente au Cout Marginal, *Revue Generale de L'Electricite*, 58, 321-340; translated as Boiteux, M. (1960) "Peak Load Pricing," *Journal of Business*, 33, 157–179.
- Borins, S.F. (1978) Pricing and Investment in a Transportation Network: The Case of Toronto Airport, *Canadian Journal of Economics*, 11, 680-700.
- Brander, J.A. and Zhang, A. (1990) Market Conduct in The Airline Industry, *Rand Journal* of Economics, 21, 567-583.
- Brander, J.A. and Zhang, A. (1993) Dynamic Behaviour in The Airline Industry, International Journal of Industrial Organization, 11, 407-435.
- Brueckner, J.K. (2002) Airport Congestion when Carriers Have Market Power, American Economic Review, 92: 1357-1375.
- Brueckner, J.K. (2005) Internalization of Airport Congestion: A Network Analysis, International Journal of Industrial Organization, 23, 599-614.
- Calzada, J. (2003) Capacity Charge and Peak-Load Pricing, Working paper, Universitat de Barcelona.
- Carlin, A. and Park, R.E. (1970) Marginal Cost Pricing of Airport Runway Capacity, American Economic Review, 60, 310-319.
- Chen, Z.Q. and Ross, T.W. (2003) Cooperating Upstream While Competing Downstream: A Theory of Input Joint Ventures, *International Journal of Industrial Organization*, 21(3), 381-397
- Civil Aviation Authority (2004) Airport Regulation: Looking to the Future, Learning from the Past, Civil Aviation Authority, London.
- Condie, S. (2000) Whither Airport Regulation?, in W. Bradshaw and H.L. Smith (eds.), *Privatization and Deregulation of Transport*, St. Martin's Press, Inc, New York.

- Craven, J. (1971) On the Optimal Time Periods for a Surplus-Maximising Utility Subject to Fluctuating Demand, *Bell Journal of Economics*, 2, 495-502.
- Craven, J. (1985) Peak-Load Pricing and Short-Run Marginal Cost, *Economic Journal*, 95, 778-780.
- Crew, M. and Kleindorfer, P. (1986) *The Economics of Public Utility Regulation*, Cambridge, Mass., MIT Press.
- Crew, M. and Kleindorfer, P. (1991) Peak-Load Pricing of Postal Service and Competition, in M. Crew and P. Kleidorfer (eds.) Competition and Innovation in Postal Sciences, Kluwer Academic Publisher, pp. 37–56.
- Crew, M., Fernando, C.S. and Kleindorfer, P. (1995) The Theory of Peak-Load Pricing: A Survey, *Journal of Regulatory Economics*, 8, 215-248.
- Dana, J.D. (1999) Using Yield Management to Shift Demand When the Peak Time Is Unknown, RAND Journal of Economics, 30, 456-474.
- Daniel, J.I. (1995) Congestion Pricing and Capacity of Large Hub Airports: A Bottleneck Model with Stochastic Queues, *Econometrica*, 63, 327-370.
- Daniel, J.I. (2001) Distributional Consequences of Airport Congestion Pricing, Journal of Urban Economics, 50, 230-258.
- Dansby, R. (1978) Capacity Constrained Peak-Load Pricing, *Quarterly Journal of Economics*, 92, 387-98.
- De Palma, A. and Lindsey, R. (1998) Information, Welfare and Usage of Congestible Facilities under Different Pricing Regimes, *Canadian Journal of Economics*, 31, 666-692.
- Forsyth, P. (1997) Price regulation of airports: Principles with Australian applications, *Transportation Research E*, 33, 297-309.
- Gersten, J. (1986) Peak-Load Pricing in Competitive Markets, *Economic Inquiry*, 24, 349-361.
- Gillen, D.W. (1994) The Market and Aviation Infrastructure: Pricing, Productivity and Privatization, *Transportation Research Circular*, No. 425 (App): 82-86.
- Hirschleifer, J. (1958) Peak-Loads and Efficient Pricing: A Comment, Quarterly journal of Economics, 72, 451-462.

- Horonjeff, R. and McKelvey, F.X. (1983) Planning and Design of Airports, third edition, McGraw-Hill.
- Laffont, J.J. and Tirole, J. (2000) Competition in Telecommunications, Massachusetts, MIT Press.
- Lave, L.B. and J.S. DeSalvo (1968) Congestion, Tolls, and the Economic Capacity of a Waterway, Journal of Political Economy, 76, 375-391.
- Levine, M.E. (1969) Landing Fees and the Airport Congestion Problem, Journal of Law and Economics, 12, 79-108.
- Mathewson, G.F. and Winter, R.A. (1984) An Economic Theory of Vertical Restraints, *The Rand Journal of Economics*, 15, 27-38.
- Morrison, S.A. (1983) Estimation of Long-Run Prices and Investment Levels for Airport Runways, *Research in Transportation Economics*, 1, 103-130.
- Morrison, S.A. (1987) The Equity and Efficiency of Runway Pricing, Journal of Public Economics, 34, 45-60.
- Morrison, S.A. and Winston, C. (1989) Enhancing the Performance of the Deregulated Air Transportation System, *Brookings Papers on Economic Activity, Microeconomics* (1989) Special Issue, 61-123.
- Oum, T.H. and Zhang, Y. (1990) Congestion Tolls, Lumpy Investment and Cost Recovery, Journal of Public Economics, 43, 353-374.
- Panzar, J. (1976) A Neoclassical Approach to Peak-Load Pricing, Bell Journal of Economics, 7, 521-530.
- Pels, E. and Verhoef, E.T. (2004) The Economics of Airport Congestion Pricing, Journal of Urban Economics, 55: 257-277.
- Poole, R.W. (1990) Airport Privatization: What the Record Shows, The Reason Foundation Policy Study No. 124, The Reason Foundation, Los Angeles, CA.
- Pressman, I. (1970) A Mathematical Formulation of the Peak-Load Pricing, *Bell Journal of Economics*, 1, 304–326.
- Productivity Commission (2002), Price Regulation of Airport Services, Report No. 19, AusInfo Canberra, Australia.
- Shy, O. (2001) Dynamic Peak-Load Pricing, Working paper, University of Haifa and Stockholm School of Economics.

- Starkie, D. (2001) Reforming UK Airport Regulation, Journal of Transport Economics and Policy, 35, 119-135.
- Steiner, P. (1957) Peak-Loads and Efficient Pricing, Quarterly Journal of Economics, 71, 585-610.
- U.S. Federal Aviation Administration (1969) Airport Capacity Handbook, 2nd edition, Washington, D.C.
- Vasigh, B. and M. Haririan (1996) Airport Privatization: A Simple Welfare Analysis, Journal of Economics and Finance, 20: 89-95.
- Williamson, O. (1966) Peak-Load Pricing and Optimal Capacity Constraints, American Economic Review, 56, 810–827.
- Yuen, A. and A. Zhang (2005) Airport Congestion and Market Structure: The Case of Variable Passenger Time Costs, Working paper, University of British Columbia.
- Zhang, A. and Y. Zhang (2003) Airport Charges and Capacity Expansion: Effects of Concessions and Privatization, *Journal of Urban Economics*, 53, 54-75.
- Zhang, A. and Y. Zhang (2006) Airport Capacity and Congestion when Carriers Have Market Power, *Journal of Urban Economics*, forthcoming.

6 CONGESTIBLE FACILITY RIVALRY IN VERTICAL STRUCTURES¹

6.1 Introduction

In this paper we investigate competition between congestible facilities –such as airports and seaports- and the associated effects on facility charges, capacities and congestion delays. In the case of airports, most of the scholarly research is concerned with the case of a single airport. This is understandable given the local monopoly nature of an airport. This is changing, however. The world has experienced a rapid growth in air transport demand since the 1970s, and many airports have been built or expanded as a result. This has led to a number of multi-airport regions such as greater London and the San Francisco Bay Area, within which airports may compete for air travelers.² At the same time, the dramatic growth of low cost carriers (e.g., Southwest and Jet Blue in the U.S.) has enabled some smaller and peripheral airports to cut into the catchment areas of large airports (see, e.g., Mason, 2000; Gillen and Lall, 2004). Taken together, these two developments have significantly increased the degree of competition between certain airports. Furthermore, airports susceptible to competition are usually prime candidates for congestion. In the United States, for example, the three multi-airport markets -Chicago, New York, and Washington metropolitan areas- contain the four airports that are officially designated by the FAA as 'slot controlled'. The description also applies to several of the 23 airports identified by the FAA as 'delay-problem airports' -these airports are in the metropolitan areas containing one or more other airports with airline service (e.g., Dallas, Detroit, Huston, Los Angeles, and San Francisco).

Our first objective is to compare our results –regarding facility charges, capacities, congestion delays, and so on– with the results derived from the single-facility case where competition is absent, and with the socially optimal outcomes. To illustrate this objective, consider that the facility is an airport. With airport competition, landing fees (facility charges) are lowered, which is beneficial to the airlines that use runways to produce air transport service and, ultimately, to the passengers who consume the transportation service. On the other hand, lower airport charges will stimulate demand and will, holding the capacity constant, increase congestion, which is 'bad' for both the airlines and passengers. However, since capacity is an endogenous decision variable, competition might lead to increased capacity for airports, leading to less congestion. The net effects on airport congestion, airfare, airlines' profit and consumer surplus are, *a priori*, unclear. Furthermore, it is useful, as to be seen below, to compare the monopoly provision of service quality (congestion delays) with the socially optimal level

A major feature of our investigation is that it contains a vertical structure within a facility: Each facility is an upstream firm that provides an input service to downstream firms (to be

¹ A version of this chapter has been submitted for publication. Basso, L.J. and Zhang A. (2006c) Congestible Facility Rivalry in Vertical Structures.

 $^{^{2}}$ De Neufville (1995) identified 26 multi-airport regions in different parts of the world as of the early 1990s. These multi-airport regions cover large territorial size, with some spanning over 100 kilometres, and have high passenger generating capacity (10 million or more annual originating air passengers).

referred as 'carriers' hereafter), which in turn produce outputs for final consumers. In particular, we shall explicitly recognize that the downstream carriers may possess market power in the output market: as argued by Brueckner (2002), Pels and Verhoef (2004) and others, airlines at congested airports usually are not atomistic and hence price-takers. More specifically, we shall employ a three-stage model of facility (e.g., airport) and carrier (e.g., airline) behavior. In the first stage, duopoly facilities choose their capacities and prices. Initially, we shall look into a 'closed loop' game, that is, capacities are decided strictly prior to prices. We later compare the results with the 'open loop' case in which capacities and prices are decided at the same time. In the second stage, oligopolistic carriers compete with one another in homogenous Cournot fashion. In the third stage, final consumers decide which facility to go –if they decide to consume the product– by comparing the 'full prices' they face, which include the price of the final product, congestion at each facility and transportation costs (from their own location to the facility).

Very few papers in the literature have examined the case of competing airports analytically. Several recent papers, including Pels and Verhoef (2004), Brueckner (2005) and Basso (2005), considered multiple airports but these airports are complementary to each other: passengers travel from one airport to another (and back) so the airports produce complements, not substitutes.³ A number of authors have examined more general duopolistic interactions between congestible facilities: Braid (1986) and Van Dender (2005) examined competition between fixed-capacity facilities, while De Palma and Leruth (1989), Baake and Mitusch (2004) and De Borger and Van Dender (2006) examined the competition between facilities that are able to adjust capacities. However, all of them have considered the facilities as final providers rather than as input providers in an intermediate market that may be imperfectly competitive. Rivalry between congestible vertical structures with imperfect competition has, as far as we know, not yet been studied in the literature.⁴

De Borger and Van Dender (2006), hereafter DBVD, is the closest paper to ours. They studied duopolistic interaction between congestible facilities that first decide on capacities and then on prices (i.e., a closed-loop game). They found that (i) the duopolists offer lower service quality, in terms of longer delays, than the monopolist, who provides exactly the same service quality as the social optimum; and (ii) for constant returns to scale and linear delay functions, the optimal pricing and optimal provision of capacity lead to exact cost recovery. DBVD indicated that their analysis may apply to seaports, airports, internet access providers and roads. However, whilst roads and internet access providers may provide services directly to final consumers, seaports and airports are input providers that reach final consumers only through downstream firms (carriers): these facilities are in an

³ One exception is Gillen and Morrison (2003), who examined two competing airports in the context of a full-service carrier and a low cost carrier. But they did not address the issue of congestion and capacity decisions, nor airline competition within each airport.

⁴ There are many papers that analyze competition between vertical structures and how vertical integration affects the rivalry, e.g., Abiru (1988) and Lee and Staelin (1997). But a case in which the upstream firms are spatially differentiated, congestible, and their congestion affects both the downstream oligopolists and the consumers, has yet been studied. Next, to be perfectly clear, in this paper we do not investigate the incentives firms would have to vertically integrate, as it has been done elsewhere (e.g. Salop and Scheffman, 1987; Hart and Tirole, 1990). Instead, we just assume those two different ownership structures (as in Abiru, 1988).

intermediate market and not in the final market. Thus, the use of vertical structures may be better suited for airports, seaports and other facility markets, such as telephone networks for example. Hence, our second objective in this paper is to explicitly incorporate downstream carriers –both their decision-making and market structure –into the analysis, in order to compare the results with those of the existing literature in which, while facility competition is present, the vertical structure is absent.⁵ In this sense, our paper is complementary to DBVD and the existing literature.

Another major departure from DBVD lies in our treatment of facilities' outputs. DBVD considered that facilities supply perfect substitutes in the eyes of final consumers. This assumption may be unrealistic for airports and seaports given that they are actually located at different points in space, with potential consumers being distributed in between and around them and consumers' travel to/from them being costly. Within a multi-airport region, for instance, some passengers may not necessarily choose an airport with cheapest fare, but may go to an airport that is nearer and has shorter total travel time (which includes delay times due to airport congestion), while other passengers may find the competing airport *just too far away*. In other words, airports may no longer have exclusive (monopoly) and clear-cut catchment areas in their vicinity. Instead, their market areas may be compound of one portion in which they face no competition, and another portion in which they do face competition. Indeed, the access time has been shown to be one of the main determinants of airport choice (e.g., Pels, et al., 2001; Ishii, et al., 2005; Fournier, et al., 2006). In this paper, we shall consider competing facilities providing differentiated services.

We find, among other results, that: (i) the duopolists' equilibrium prices increase with both the consumers' value of time and the carriers' cost sensitivity to congestion delays; entrance of a new carrier to any of the facilities depresses the prices charged by both facilities; and lower marginal cost of the carriers at one facility will induce a higher facility price at that facility but a lower facility price at the other facility. (ii) The duopoly facilities provide longer congestion delays than a monopolist only if capacity decisions are made prior to the facility pricing decisions. When the capacity and pricing decisions are made simultaneously, or when capacity investments are not observable prior to the pricing decisions, the duopolists would provide the same level of service quality (congestion delays) as the monopolist. (iii) A closed-loop duopoly invests less in capacity, has higher facility prices and longer delays than an open-loop duopoly. (iv) Conditional on facility charges, the monopoly capacity rules are the same as the socially-optimal capacity rules if and only if the downstream carriers' markets are perfectly competitive. When the downstream markets are imperfectly competitive, the monopoly capacity rules will be different from the social capacity rules. (v) When capacities are endogenous and adjustable, monopoly pricing and capacity choices result in a higher level of service quality than the social optimum. (vi) When the monopolist vertically integrates with the carriers at the facilities, it would provide the same congestion level as the social optimum. Nevertheless, the monopoly service level is not socially optimal in a second-best sense. In effect, in the fully ex-ante symmetric case, it is too low with respect to the second best. (vii) Finally,

⁵ Furthermore, as demonstrated by Basso (2005, 2006), this distinction is relevant and important in the derivation of welfare-maximization results. Thus, it may be necessary to incorporate carriers into the analysis.

despite the fact that we assume constant returns to scale and a linear delay function, the socially optimal pricing and capacity will probably not lead to cost recovery, owing to market power at the carriers' level.

The paper is organized as follows. Section 6.2 sets up the basic model and characterizes the equilibria for both the final and intermediate markets. Section 6.3 examines the rivalry between two profit-maximizing facilities and compares the open-loop game to the closed-loop game. In Section 6.4, we compare the results from the duopoly rivalry with the monopoly outcome and the social optimum. Section 6.5 examines a vertical-integration case in which a single owner of the facilities maximizes the sum of the facilities' profit and the carriers' profit. Section 6.6 contains concluding remarks.

6.2 The Model

We consider an infinite linear city, where consumers are distributed uniformly with a density of one consumer per unit of length. There are two congestible facilities, located at 0 and 1 respectively and there are N_0 and N_1 carriers offering *services* at each facility. The locations of the facilities, the number of carriers and the facility from which they produce are exogenous. At each facility, carriers are ex-ante symmetric and they all offer a homogenous good/service, which is to be referred to as a 'product' hereafter. We will use the term 'fare' to indicate the price of the final product, reserving the terms 'price' and 'charge' for the facilities' price. Given the homogeneity and symmetry at each facility, the fare will, in equilibrium, be the same for every carrier at each facility.

The timing of our facility-rivalry game is as follows:

- (i) Facility market competition: The facilities choose capacities and prices for the *input*. Initially, we will investigate a closed-loop game where capacities are decided prior to prices. We later compare this case with an open-loop game in which capacities and input prices are decided simultaneously, or capacity investments are not observable (DBVD looked at a closed-loop game).
- (ii) Output (carriers) market competition: Given the facilities' decisions, carriers at each facility make their decisions in homogenous Cournot fashion.⁶
- (iii) Final consumers decide whether to consume the product and if so, which facility to go.

We shall solve the game by backward induction: First we obtain the demands faced by the carriers at each facility by solving the consumers' problem, taking the fares and delay levels as given. We then solve the carriers' maximization/competition problem, taking as given the facility charges and capacities. Finally, we solve the facilities' problem for various cases including duopoly, monopoly, the social optimum, and the case in which the facilities vertically integrate with the carriers.

⁶ We assumed this sequential, or Stackelberg, structure because it is what seems more reasonable for the industries we have used as examples (airports, seaports and telecommunication networks). See Young (1991) for a discussion of the effects of simultaneous choice of prices in a vertical structure.

6.2.1 The consumer problem

Consumers have unit demands for the final product, and they care for its 'full price.' If the product is consumed, the consumer derives a gross benefit of V. A consumer located at $0 \le z \le 1$, derives the following net benefit (utility) if she goes to facility 0:

$$U_0 = V - f_0 - \alpha D(Q_0, K_0) - \frac{t}{4}z$$

where f_0 is the (equilibrium) fare of the carriers' product at facility 0, D is congestion delay time at the facility, α represents the consumers' value of time, and t/4 is the parameter representing consumers' transportation cost.⁷ Thus congestion affects consumers negatively, and the product's full price, denoted ρ_0 , is the sum of fare f_0 , consumers' delay cost incurred at the facility αD , and transportation cost tz/4. Note that the congestion delay at facility h (h = 0,1) depends on Q_h , which denotes the total output produced at h, and K_h , the capacity of facility h. Similarly, if the consumer goes to facility 1, then she derives a net benefit:

$$U_1 = V - f_1 - \alpha D(Q_1, K_1) - \frac{t}{4}(1 - z)$$

with the full price being $\rho_1 = f_1 + \alpha D_1 + t(1-z)/4$. This is an 'address model' with linear transportation costs, and the differentiation of the two facilities is captured by consumer transportation cost (i.e., positive t). We note that in addition to distance, other aspects of facility differentiation may be captured by extending the present model.⁸ For instance, we could further address differential access costs to the two facilities (e.g., ground access to airports) by introducing a parameter to the net-benefit function such that

$$U_1 = V - f_1 - \alpha D(Q_1, K_1) - t\lambda_1(1-z)/4$$

where $\lambda_1 > 1$ ($0 < \lambda_1 < 1$, respectively) if facility 1 has a higher (lower, respectively) access cost for consumers than facility 0. On the other hand, while we allow facility differentiation by final consumers, the facilities may be differentiated also in the eyes of

⁷ This transportation cost may seem awkward but, of course, it has no influence on the results: one could always define t'=t/4 to have a more usual transportation cost. It is chosen because it will simplify most of the equations in the paper (see, e.g., equations (4)).

⁸ When distances among competing airports in a multi-airport region are not great, access time to airports and other airport service factors are particularly important in determining airport patronage. For example, using revealed preference data Ndoh, et al. (1990) examined air travelers' choice of departure airports among four Central England airports. Access time to airports, flight frequency of both direct and connecting routes, average journey time, average connection time to hub airports, and weekly available aircraft seats were found to be significant factors in affecting travelers' choice. Similarly, Bradley (1998) showed, based on stated preference surveys, that air travelers' choice among competitive departure airports in Europe is affected by twelve factors, including air fare, travel time to the airport, and airport congestion. Using a hypothetical example and later the San Francisco Bay Area case study, Pels, et al. (2000, 2001, 2003) showed that ground accessibility of an airport is the most important factor in affecting airport choices in a multi-airport market.

downstream firms. In this paper we shall for simplicity abstract away from this aspect of facility differentiation by assuming that the carriers are specific to a facility. This assumption may be reasonable for airports, in that given the high specific investment by an airline in serving an airport, it is hard for the airline –certainly harder than for travelers– to switch from one airport to another, especially in the short run.⁹

Assuming that everyone in the [0,1] interval consumes (for this we need V to be sufficiently large; see the analysis below) and both facilities receive consumers from [0,1] (i.e., an interior solution), then the indifferent consumer $\tilde{z} \in (0,1)$ is given by

$$\rho_0 = f_0 + \alpha D(Q_0, K_0) + \frac{t}{4} \tilde{z} = f_1 + \alpha D(Q_1, K_1) + \frac{t}{4} (1 - \tilde{z}) = \rho_1$$

that is,

$$\tilde{z} = \frac{1}{2} + \frac{f_1 + \alpha D(Q_1, K_1) - f_0 - \alpha D(Q_0, K_0)}{t/2}$$
(6.1)

Thus, \tilde{z} , or the number of consumers (who are between the facilities) going to facility 0 rather than facility 1, increases in $f_1 + \alpha D(Q_1, K_1)$ while decreasing in $f_0 + \alpha D(Q_0, K_0)$. Taking into account the fact that facilities 0 and 1 also capture the consumers at their immediate left and immediate right sides respectively, we thus identify four types of consumers, defined by three different locations in the city:

- $z = z^{l}$: the last consumer on the left side of the city, who consumes and goes to facility 0
- $z = \tilde{z}$: the indifferent consumer who belongs to (0,1)
- $z = z^r$: the last consumer on the right side of the city, who consumes and goes to facility 1

These points also define the catchment areas of each facility as shown in Figure 6.1:



Figure 6.1: Consumer distribution and facilities' catchment areas

⁹ See Pels et al. (1997) for an analysis on airport pricing and competition in attracting airlines, and Adler and Berechman (2001) for an empirical study of airport quality from the airlines' viewpoint.

With the assumptions about uniformity and unit density of consumers, we obtain that z^{t} and z^{r} are given by

$$z' = -\frac{V - f_0 - \alpha D(Q_0, K_0)}{t/4}, \qquad z' = 1 + \frac{V - f_1 - \alpha D(Q_1, K_1)}{t/4}$$
(6.2)

Hence, the consumer demands for products at each facility are then given by:

$$Q_0 = \tilde{z} + |z^l|, \qquad Q_1 = (1 - \tilde{z}) + (z^r - 1)$$

Replacing \tilde{z} from (6.1) and z^{l} , z^{r} from (6.2) yields:

$$Q_{0} = \frac{(t/4) + 2V}{t/2} + \frac{f_{1} + \alpha D_{1} - 3(f_{0} + \alpha D_{0})}{t/2}$$

$$Q_{1} = \frac{(t/4) + 2V}{t/2} + \frac{f_{0} + \alpha D_{0} - 3(f_{1} + \alpha D_{1})}{t/2}$$
(6.3)

where $D_h \equiv D(Q_h, K_h)$ for h = 0,1. It is clear that the consumer demands depend not only on product prices (f_0, f_1) of the two facilities, but also on their delay levels (D_0, D_1) . Note that in order to have both facilities receiving consumers from [0,1], we need $|f_1 + \alpha D_1 - f_0 - \alpha D_0| < t/4$, whereas in order to have everyone in the [0,1] interval consuming the product we need $2V \ge f_1 + \alpha D_1 + f_0 + \alpha D_0 + (t/4)$, both of which are our maintained assumptions.

If the facilities served the final consumers directly and had identical constant marginal costs, then it is easy to obtain the Bertrand-Nash equilibrium in prices.¹⁰ Here, however, we are interested in the case where the facilities are upstream input providers to carriers, who are the ones that provide the product to final consumers. Since we are assuming that the carriers compete in Cournot fashion, we are interested in obtaining the inverse demand functions.¹¹ Inverting the 'direct' demand system (6.3) in (f_0, f_1) we obtain the inverse demand functions that the carriers at each facility face: ¹²

¹⁰ It would be given by $p_0^* = ((7t/4) + 14V + 21c - 17D_0 + 3\alpha D_1)/35$, and analogously for p_1 . Further, if

the facilities were non-congestible, the fares would be equal and given by $p^* = ((t/4) + 2V + 3c)/5$

¹¹ Earlier studies that have explicitly incorporated imperfect competition in the carriers' market at a congestible airport (e.g., Brueckner, 2002, 2005; Pels and Verhoef, 2004; Basso, 2005; Basso and Zhang, 2006; Zhang and Zhang, 2006) have all assumed a Cournot game in the output-market competition. Brander and Zhang (1990, 1993), for example, find some empirical evidence that rivalry between duopoly airlines is consistent with Cournot behaviour.

¹² Perhaps a more sensible assumption would have been to consider a non-constant density of consumers to indicate that towards the center, that is, in between the facilities, the density may be greater than outside. Such an assumption would complicate the derivations but should not affect the main insights. What is important is the 'infinity' of the city or, more precisely, that consumers far away from the facilities find it

$$f_0(Q_0, Q_1, K_0) = 2t + V - 3tQ_0 - tQ_1 - \alpha D_0(Q_0, K_0)$$

$$f_1(Q_0, Q_1, K_1) = 2t + V - 3tQ_1 - tQ_0 - \alpha D_1(Q_1, K_1)$$
(6.4)

With these inverse demand functions, we can now model the carriers' competition which occurs at stage 2 of the game.

6.2.2 Output (carriers) market competition

In the output (carriers) market, although any given carrier faces direct competition from the other carriers at the same facility, it would also take into account what happens at the other facility: the demands depend on both Q_0 and Q_1 . From (6.4) it may also seem that carriers would only care about the congestion at their own facility, but this is not the case. Recall that in the direct demand system (6.3), the demands depend on the delays at both facilities.

We consider ex-ante symmetric carriers at each facility. The cost function of carrier i at facility h is given by:

$$C^{ih}(Q_h^i, \mathbf{Q}_h^{-i}) = (c_h + P_h + \beta_h D(Q_h, K_h))Q_h^i, \quad h = 0,1$$
(6.5)

where P_h is the facility charge (the input price) at facility *h*. Notice that both the carriers' pure operational cost, denoted c_h , and the carriers' delay cost parameter, β_h , are facility-specific. Thus, congestion at each facility affects not only the final consumers as indicated earlier, but the carriers as well. Further, the cost function C^{ih} depends on the production of other carriers at the facility, \mathbf{Q}_h^{-i} , through the congestion term, as well as its own production level Q_h^i (and $Q_h = \sum_i Q_h^i$). It does not depend on the output of carriers at the other facility, however.

Given the demand and cost specifications, the profit for carrier i at facility 0 is:

$$\phi^{i0}(Q_0^i, \mathbf{Q}_0^{-i}, Q_1, K_0) = f_0(Q_0, Q_1, K_0)Q_0^i - (c_0 + P_0 + \beta_0 D(Q_0, K_0))Q_0^i, \quad i=1,\dots,N_0 \quad (6.6)$$

and the profit at facility 1 can be similarly written. As can be seen from (6.6), these profit functions depend on the outputs of carriers at both facilities.¹³ The carriers' first-order conditions at facility 0 are given by

optimal not to consume. For the case of airports, De Neufville (1995) has indicated that the competing airports can be separated as far as 135 kilometres apart, which is consistent with our infinite-city modeling. If the city was finite and the whole market was covered, both direct demands would be dependent on the difference $f_1 + \alpha D_1 - (f_0 + \alpha D_0)$. The system of direct demands would not be invertible in prices as a result, making it impossible to model Cournot competition.

¹³ Since each consumer in the model consumes one unit of the carriers' products, this would imply, in the airport-airlines case, one flight per person, which is indeed indefensible. This is easily solved however, through a 'fixed proportions' assumption: i.e., assuming that $S = \text{Aircraft Size} \times \text{Load Factor, is constant and}$

$$\partial \phi^{i0} / \partial Q_0^i = 0, \qquad i=1,\ldots,N_0$$

Calculating these conditions, imposing symmetry (i.e., $Q_0^i = Q_0 / N_0$) and re-arranging, and doing the same for facility 1, we obtain:

$$\Omega^{h} = (2t + V - c_{h} - P_{h}) - Q_{h} \frac{3t(N_{h} + 1)}{N_{h}} - Q_{-h} t$$

$$- (\alpha + \beta_{h}) \left(D(Q_{h}, K_{h}) + \frac{Q_{h}}{N_{h}} D_{Q}(Q_{h}, K_{h}) \right) = 0$$
(6.7)

where h, -h = 0,1, $h \neq -h$, and $D_Q \equiv \partial D / \partial Q_h$. From the two equations in (6.7) it is straightforward to obtain the *derived* demand functions:

$$P_h(Q_0, Q_1, K_h; N_h)$$

which are the inverse demand functions faced by the facilities. However, the direct demands –necessary to model price competition between the facilities– do not necessarily have closed expressions and may be defined only *implicitly* by the system of equations (6.7). Working with implicit demand functions does not lead to useful results in this case however, because comparative statics will be too complicated. For example, the changes in Q_h with respect to own price and the other price, namely, $\partial Q_0 / \partial P_0$ and $\partial Q_1 / \partial P_0$, are obtained as the solution of the following system:

$$\frac{\partial \Omega^{0}}{\partial P_{0}} + \frac{\partial \Omega^{0}}{\partial Q_{0}} \frac{\partial Q_{0}}{\partial P_{0}} + \frac{\partial \Omega^{0}}{\partial Q_{1}} \frac{\partial Q_{1}}{\partial P_{0}} = 0, \qquad \qquad \frac{\partial \Omega^{1}}{\partial Q_{0}} \frac{\partial Q_{0}}{\partial P_{0}} + \frac{\partial \Omega^{1}}{\partial Q_{1}} \frac{\partial Q_{1}}{\partial P_{0}} = 0$$

leading to expressions that are quite unmanageable.

To have the explicit demand functions, we shall specify a linear delay function. The use of a linear delay function also enhances comparability between our results and those of De Borger and Van Dender (2006) in which a linear specification was employed. Linear delay functions may, nevertheless, lead to the problem of the first-order condition approach prescribing a solution in which capacity is exceeded, something that does not happen when delay functions are convex enough (e.g., when $D(Q, K) = Q[K(K - Q]^{-1})$, delays approach infinity when demand approaches capacity). There are two ways around this problem: (i) we can assume an interior solution and later find conditions for this to be true; or (ii) we can impose *a priori* a capacity-rationing rule for the case in which capacity is reached. We shall take the first approach.

the same across the airlines, and that S consumers make up one flight. Then the only change in our results would be that a parameter S would be included. Fixed proportions has been assumed in Brueckner (2002; 2005), Pels and Verhoef (2004), Basso (2005), Zhang and Zhang (2006), and Basso and Zhang (2006).

In what follows we assume a linear delay function,

$$D(Q,K) \equiv a \frac{Q}{K} \tag{6.8}$$

With (6.8) and letting $\mathbf{P} \equiv (P_0, P_1)$, $\mathbf{K} \equiv (K_0, K_1)$ and solving system (6.7), we obtain the following derived demands for the two facilities (the expressions for facility 1 are analogous):

$$Q_0(\mathbf{P}, \mathbf{K}) = \frac{t(c_1 + P_1 - 2t - V) - g_1(c_0 + P_0 - 2t - V)}{g_0 g_1 - t^2}$$
(6.9)

where

$$g_{0} = \frac{N_{0} + 1}{N_{0}} \left(3t + a \frac{\alpha + \beta_{0}}{K_{0}} \right).$$
(6.10)

Thus, the demands faced by the facilities depend directly on facility prices and capacities, and they are linear in P_0 and P_1 . Notice that g_0 has two parts: the first part is related to t (the consumers' transportation cost) which leads to the two facilities being differentiated and having associated market power; and the second part is related to K_0 , reflecting the congestion effect. Further, g_0 depends on the carrier market structure: it decreases in N_0 . In particular, g_0 is largest with a monopoly carrier ($N_0 = 1$) while smallest in the competitive case (when $N_0 \rightarrow \infty$). It can be easily shown that $g_h > t$ and, consequently, the denominator of (6.9) is positive. Furthermore, g_h increases with α and β_h , and decreases with K_h and N_h .

Taking the perspective of facility 0 we can, from (6.9) and (6.10), characterize these facility demands through the following comparative statics:

$$\frac{\partial Q_0}{\partial P_0} = \frac{\partial Q_0}{\partial c_0} = -\frac{g_1}{g_0 g_1 - t^2} < 0 \tag{6.11}$$

$$\frac{\partial Q_0}{\partial P_1} = \frac{\partial Q_0}{\partial c_1} = \frac{t}{g_0 g_1 - t^2} > 0$$
(6.12)

$$\frac{\partial Q_0}{\partial \eta} = -\frac{Q_0 g_1 (\partial g_0 / \partial \eta)}{g_0 g_1 - t^2} > 0, \quad \text{for } \eta \in \{K_0, N_0, -\beta_0\}$$
(6.13)

$$\frac{\partial Q_0}{\partial \eta} = \frac{Q_1 t \left(\partial g_1 / \partial \eta \right)}{g_0 g_1 - t^2} < 0, \quad \text{for } \eta \in \{K_1, N_1, -\beta_1\}$$
(6.14)

$$\frac{\partial Q_0}{\partial \alpha} = \frac{Q_1 t (\partial g_1 / \partial \alpha) - Q_0 g_1 (\partial g_0 / \partial \alpha)}{g_0 g_1 - t^2}$$
(6.15)

All the signs in (6.11) to (6.14) are as expected: e.g., inequality (6.11) is equivalent to the demand functions being downward sloping, and (6.12) shows that the facilities are 'gross' substitutes. From (6.11), (6.12) and $g_h > t$, it follows that

$$\frac{\partial Q_h}{\partial P_h} + \frac{\partial Q_h}{\partial P_{-h}} = -\frac{g_{-h} - t}{g_0 g_1 - t^2} < 0$$

that is, own-price effects on demand dominate cross-price effects. This condition is equivalent to the stability condition for price equilibrium in the duopoly facility game and, together with downward-sloping demands, further implies the uniqueness of the price equilibrium (see, e.g., Dixit, 1986).

The demand for a facility increases in own capacity, but decreases in the competitor's capacity. The effects of carriers' marginal costs on the facility demands are the same as those of prices; after all, for the carriers, the facility charge is part of its marginal cost. Moreover, a facility's demand rises the greater the number of carriers it has, and the less its carriers care about congestion (i.e., the lower the value β_h is). Its demand also rises the fewer carriers there are at the other facility, and the more they care about congestion.

As for the effect of consumers' time value α on facility demands, equation (6.15) shows, by $g_h > t$, that in a fully symmetric case (facilities and carriers), a larger time value will reduce the demand for both facilities. However, it may occur that facility 0's demand increases with α in asymmetric cases. For example, if everything is symmetric except for carriers' marginal cost, and $(c_0 - c_1)$ is large enough, then higher c_0 implies that carriers at facility 0 would have, *ceteris paribus*, higher prices and hence the facility will have a smaller demand but also less congestion. A marginal increase in the value of time would induce a shift by consumers towards the less congested facility, increasing the demand. Finally, the analysis of the demand effect of transportation cost t is similar to that of α , although the expression is slightly more complicated than (6.15). In a fully symmetric case, a smaller transportation cost will increase the demand. However, this may be reversed with strong asymmetry.

6.3 Equilibria of Duopoly Facilities

Now that we have characterized the final market equilibrium and the facilities' demands, we can analyze the facilities' market –i.e., stage 1 of the game– for various cases. In this stage, facilities take into account what will happen downstream. This Section investigates the rivalry between two profit-maximizing facilities; in Section 6.4, we will compare the results from this duopoly rivalry with those of monopoly and the social optimum.

6.3.1 Closed-loop duopoly

Consider first a 'closed loop' game in which the facilities first choose capacities and then, given the capacities, they choose prices. We assume that the facilities' operational and capacity costs are separable and their marginal costs are constant. Without loss of generality, we set the operational marginal costs to zero, so the profit of facility h can be written as:

$$\pi^{h}(\mathbf{P}, \mathbf{K}) = Q_{h}(\mathbf{P}, \mathbf{K})P_{h} - m_{h}K_{h}, \qquad h = 0,1$$
 (6.16)

where m_h denotes the marginal cost of capacity.

Pricing

In the pricing stage the capacities are given, and the two facilities simultaneously choose their prices P_h to maximize profit (6.16) taking into account the carriers' competition and consumers' behaviour. The first-order conditions and equation (6.11) lead to the following pricing rules:

$$P_0 = Q_0 g_0 - \frac{Q_0 t^2}{g_1}, \qquad P_1 = Q_1 g_1 - \frac{Q_1 t^2}{g_0}$$
(6.17)

The pricing rule for each facility has two parts. Consider facility 0, for example. Using (6.8) and (6.10), the first part, Q_0g_0 , can be further written as:

$$Q_0 g_0 = (1 + s_0) (D_0 (\alpha + \beta_0) + 3t)$$
(6.18)

where $s_0 = 1/N_0$ is a carrier's market share at facility 0. The first term on the RHS of (6.18) is a congestion charge, but the duopoly facilities charge more than just the pure uninternalized congestion of each carrier (which would have been $1 - s_0$, rather than $1 + s_0$): this is caused by the failure of coordination in the vertical structures.¹⁴ The second term in (6.18) is a mark-up from exploitation of market power that arises given the locational preferences of consumers. Going back to (6.17), the second part, $-Q_0t^2/g_1$, is a mark-down that arises owing to facility competition. As the other facility becomes less attractive –i.e., as K_1 decreases and hence g_1 increases– the mark-down falls, thereby making the price of facility 0 larger.¹⁵

¹⁴ As is to be seen in Sections 6.4 and 6.5, this coordination failure is not solved if facilities are singleowned but would be solved by a vertical integration of facilities and carriers.

¹⁵ This expression has a different flavour than the one obtained in a non-vertical setting (e.g., De Borger and Van Dender, 2006, equation (9)). In the non-vertical case, the second part would be positive, although a less attractive facility 1 would reduce the mark-up.

Notice that each pricing rule in (6.17) does not represent a 'best reply' function because Q_0 still depends on P_0 . Instead, the best-reply function $P_0^B(P_1, \mathbf{K})$ is implicitly determined by the first-order condition,

$$\frac{\partial \pi^0(P_0^B(P_1,\mathbf{K}),P_1,\mathbf{K})}{\partial P_0} = 0$$

and, therefore, its slope is given by $\partial P_0^B / \partial P_1 = -(\partial^2 \pi^0 / \partial P_0 \partial P_1) / (\partial^2 \pi^0 / \partial P_0^2)$, leading to:

$$\frac{\partial P_0^B}{\partial P_1} = \frac{t}{2g_1}, \qquad \qquad \frac{\partial P_1^B}{\partial P_0} = \frac{t}{2g_0}$$

It is easy to see that the slope of the best-reply function is positive but smaller than 1, which indicates that the prices (P_0, P_1) are strategic complements and the Nash equilibrium (NE) in prices will be unique and stable. This pricing NE –which we will denote by a *– is at the intersection of the two best-reply functions. Letting $\mathbf{x} \equiv (N_0, N_1, c_0, c_1, \beta_0, \beta_1, \alpha, t)$ be the vector of exogenous parameters, the NE, $\mathbf{P}^* = (P_0^*, P_1^*)$ is found to be

$$P_0^*(\mathbf{K};\mathbf{x}) = \frac{g_0 t (c_1 - 2t - V) - (c_0 - 2t - V)(2g_1g_0 - t^2)}{4g_0g_1 - t^2}$$
(6.19)

Next, we conduct comparative statics to see how capacities and features of the consumer demand and the carriers' market affect the equilibrium prices. For capacities, we obtain:

$$\frac{\partial P_0^*}{\partial K_0} = \frac{t^2 (Q_0 g_0 + Q_1 t)}{g_0 (4g_0 g_1 - t^2)} \frac{\partial g_0}{\partial K_0} < 0, \qquad \frac{\partial P_0^*}{\partial K_1} = \frac{2g_0 t (Q_1 g_1 + Q_0 t)}{g_1 (4g_0 g_1 - t^2)} \frac{\partial g_1}{\partial K_1} < 0 \qquad (6.20)$$

This shows that higher capacities imply smaller equilibrium prices. In other words, the more congestible the system is, the higher the equilibrium prices are. This result was also found in the non-vertical setting of DBVD, which is not surprising because our derived demands for the facilities react to changes in prices and capacities in the same fashion as the demands they assumed in their final market.

Another interesting parameter to look at is the time value α , a characteristic of the final market. We find that $\partial P_0^* / \partial \alpha > 0$, which is interesting because, despite the fact that the demand for facility 0 may increase or decrease with higher α , the equilibrium prices will always increase. Just as a more congestible system leads to higher prices, consumers that are more sensitive to congestion also induce higher facility prices.

But perhaps more interesting than the previous comparative statics –which pertain to the final market– is to see how the equilibrium prices change with characteristics of the *intermediate* market (such as variables N_h , β_h or c_h) and with the facility differentiation, as these two aspects are our main departures from the literature. The results are reported in Proposition 6.1:

Proposition 6.1: For given capacities, (i) the duopolists' equilibrium prices increase with both the consumers' value of time and the carriers' cost sensitivity to congestion delays. (ii) Entrance of a new carrier to *any* of the facilities depresses the prices charged by *both* facilities. (iii) Lower marginal cost of the carriers at one facility will induce a higher facility price at that facility but a lower facility price at the other facility.

Proof: (i), (ii) The first part of (i) has been indicated by in the text (i.e., $\partial P_0^* / \partial \alpha > 0$). To prove the second part and (ii), we can, from (6.19), show:

$$\frac{\partial P_0^*}{\partial \beta_0} > 0, \qquad \frac{\partial P_0^*}{\partial \beta_1} > 0, \qquad \frac{\partial P_0^*}{\partial N_0} < 0, \qquad \frac{\partial P_0^*}{\partial N_1} < 0.$$

Hence, the more sensitive the carriers are to congestion, the higher the prices will be, which is consistent with the comparative statics with respect to capacities and α . Also, the entrance of a new carrier to *any* of the facilities would depress the prices charged by *both* facilities.

(iii) This part follows from the following inequalities:

$$\frac{dP_0^*}{dc_0} = -\frac{2g_0g_1 - t^2}{4g_0g_1 - t^2} < 0, \qquad \frac{dP_0^*}{dc_1} = \frac{g_0t}{4g_0g_1 - t^2} > 0.$$

Proposition 6.1 suggests that, when capacity is fixed, lower marginal cost of the carriers at one facility would induce a higher price at that facility but a lower price at the other facility. For example, if we start from a case in which airlines have the same marginal costs $c_0 = c_1$ and we replace the airlines of one facility by lower marginal-cost carriers, then the airport charge would increase at the airport with low-cost carriers, while the charge at the other airport would fall. This may serve as a testable implication for empirical studies. Note, here, that the number of airlines at each airport would not need to be the same.

Capacity Decisions

The comparative statics in (6.20) show that the more congestible the system is, the higher the equilibrium prices are. Nevertheless, this result does not necessarily imply that the facility firms would prefer a more congestible system. When deciding their capacities, these

firms are able to foresee what is going to happen in the pricing stage and hence take the price equilibrium $\mathbf{P}^{*}(\mathbf{K})$ into account. The reduced-form profits for the capacity game are:

$$\Pi^{h}(\mathbf{K}) = \pi^{h} \left(\mathbf{P}^{*}(\mathbf{K}), \mathbf{K} \right), \qquad h = 0, 1$$
(6.21)

where $\pi^{h}(\mathbf{P}, \mathbf{K})$ is given by (6.16).

Consider facility 0. Taking the derivative of its profit with respect to capacity, we obtain the following first-order condition:

$$\Pi_0^0 = \frac{\partial \Pi^0(\mathbf{K})}{\partial K_0} = \frac{\partial \pi^0}{\partial P_0} \frac{\partial P_0^*}{\partial K_0} + \frac{\partial \pi^0}{\partial P_1} \frac{\partial P_1^*}{\partial K_0} + \frac{\partial \pi^0}{\partial K_0} = \frac{\partial \pi^0}{\partial P_1} \frac{\partial P_1^*}{\partial K_0} + \frac{\partial \pi^0}{\partial K_0} = 0$$
(6.22)

where the second equality follows from the envelope theorem. Hence, when a facility decides to marginally increase or decrease its capacity, it considers two effects on profit: a direct effect $\frac{\partial \pi^0}{\partial K_0}$, and an indirect effect $\frac{\partial \pi^0}{\partial P_1} \frac{\partial P_1^*}{\partial K_0}$. Making the calculations, (6.22) becomes:

$$P_0^* \frac{\partial Q_0}{\partial P_1} \frac{\partial P_1^*}{\partial K_0} + P_0^* \frac{\partial Q_0}{\partial K_0} - m_0 = 0$$
(6.23)

The second and third terms on the left-hand side (LHS) of equation (6.23) are the direct effect of a marginal increase in capacity. Its sign obviously depends on the extent by which own demand increases when capacity is increased, and on the marginal cost of the capacity expansion. The indirect, strategic effect –the first term on the LHS of (6.23)– indicates that a marginal increase in own capacity will lead to a reduction in the rival facility's price (recall comparative statics (6.20)), which in turn will decrease own demand and own profit. The optimal level of capacity is then chosen such that the marginal gain from the direct, positive demand effect equates the marginal loss from both the capacity cost, m_0 , and the negative strategic effect.

Open-loop duopoly 6.3.2

In an 'open-loop' game, capacities and prices are chosen simultaneously by the facilities or, at least, if capacities are chosen earlier, they are not observable by the rival facility. The problem faced by facility h is:

$$\max_{P_h, K_h} \pi^h(\mathbf{P}, \mathbf{K})$$

where π^{h} is given by (6.16). The first-order conditions will give rise to the pricing and capacity rules. The pricing rules remain the same as those given in (6.17), whereas the capacity rules can be derived, using equation (6.13), as:

$$K_{h} = \left(\frac{a(\alpha + \beta_{h})(N_{h} + 1)}{m_{h} N_{h}}\right)^{1/2} Q_{h}, \qquad h = 0,1$$
(6.24)

The capacity rules (6.24), which obviously equates the marginal benefit of an increment in capacity to marginal cost (m_h) for each facility, shows that at the subgame perfect Nash equilibrium –denoted by do, for the duopoly open-loop game– congestion delays at facility h will be equal to:

$$D(Q_h^{do}, K_h^{do}) = a \frac{Q_h^{do}}{K_h^{do}} = \left(\frac{a m_h N_h}{(\alpha + \beta_h)(N_h + 1)}\right)^{1/2}, \qquad h = 0,1$$
(6.25)

Thus, in equilibrium, the delay times at a facility increase with the number of carriers at the facility (N_h) and the capacity cost (m_h) . They decrease as the time costs –either consumers' cost (α) or carriers' cost (β_0) – increase. Equations (6.25) also define a sufficient condition for an 'interior solution' in the sense that $Q_h \leq K_h$:

$$m_h \le a(\alpha + \beta_h), \qquad h = 0,1 \tag{6.26}$$

Hence, if the capacity costs are *low enough*, or if the time costs are *high enough*, we are guaranteed to have an interior solution.

6.3.3 Comparisons: closed- vs. open-loop duopoly

We first show that the facilities will invest less in capacity in the closed-loop (sequential) game than in the open-loop (simultaneous) game. To do this, we take the capacity first-order condition of the closed-loop game, equation (6.22), and evaluate it at the open-loop capacity:

$$\Pi_{0}^{0}\Big|_{\text{loop}} = \left(\frac{\partial \pi^{0}}{\partial P_{1}}\frac{\partial P_{1}^{*}}{\partial K_{0}} + \frac{\partial \pi^{0}}{\partial K_{0}}\right)\Big|_{\text{loop}} = \left(\frac{\partial \pi^{0}}{\partial P_{1}}\frac{\partial P_{1}^{*}}{\partial K_{0}}\right)\Big|_{\text{loop}} = \underbrace{P_{0}^{*}}_{+}\underbrace{\frac{\partial Q_{0}}{\partial P_{1}}\frac{\partial P_{1}^{*}}{\partial K_{0}}}_{+} < 0 \quad (6.27)$$

The second equality follows from the capacity first-order condition in the open-loop game. The signs of the derivatives follow from inequalities (6.12) and (6.20). Hence, in the closed-loop game, according to the nomenclature of Fudenberg and Tirole (1984), firms invest less in capacity following 'puppy dog' strategies: Investment in capacity would make a facility tough, in that it decreases the facility's price hurting the rival, but that triggers a

harsh pricing reaction from the rival facility, since the prices are strategic complements. Hence, the facilities will try to soften the price competition by committing to *small* capacities in the first stage: facilities want to look small and inoffensive. This directly leads to higher prices.

However, that the fact that capacities are smaller when they are chosen prior to prices does not directly imply that delays will be longer. This is because, on one hand, capacity levels directly affect demands and, on the other hand, we now have higher prices. Yet, it can be shown that delays do increase. From (6.27) and (6.23) it can be seen that at the closed-loop duopoly equilibrium –which we denote by dc– it happens that:

$$\left.\left(P_0^*\frac{\partial Q_0}{\partial K_0}-m_0\right)\right|_{dc}>0$$

Since in the equilibrium, equation (6.17) must hold, we can replace P_0^* with it. Further replacing $\partial Q_0 / \partial K_0$ with (6.13) and rearranging we obtain that

$$a\frac{Q_0^{dc}}{K_0^{dc}} > \left(\frac{a\,m_0\,N_0}{(\alpha + \beta_0)(N_0 + 1)}\right)^{1/2} \implies D(Q_0^{dc}, K_0^{dc}) > D(Q_0^{do}, K_0^{do})$$
(6.28)

The above discussions lead to:

Proposition 6.2: A closed-loop duopoly (in which capacities are chosen prior to prices) invests less in capacity, has higher facility prices and higher congestion delays than an open-loop duopoly (in which prices and capacities are chosen simultaneously, or capacity investments are not observable).

6.4 Monopoly and the Social Optimum

Having examined the duopoly case, we shall in this Section investigate both the monopoly case and the social optimum, emphasizing comparisons among the three cases.

6.4.1 Monopoly

Here a monopolist owns and operates both facilities, so its problem is:¹⁶

$$\max_{\mathbf{P},\mathbf{K}} \pi^0(\mathbf{P},\mathbf{K}) + \pi^1(\mathbf{P},\mathbf{K}) = \max_{\mathbf{P},\mathbf{K}} \sum_h (Q_h(\mathbf{P},\mathbf{K})P_h - m_hK_h).$$

¹⁶ Note that the results will remain the same in the monopoly case (and the social optimum) whether capacity and price are made simultaneously or sequentially.

The first-order conditions with respect to prices are:

$$P_0 \frac{\partial Q_0}{\partial P_0} + Q_0 + P_1 \frac{\partial Q_1}{\partial P_0} = 0, \qquad P_1 \frac{\partial Q_1}{\partial P_1} + Q_1 + P_0 \frac{\partial Q_0}{\partial P_1} = 0$$

Using (6.11) and solving for P_0 and P_1 we obtain:

$$P_0 = Q_0 g_0 + Q_1 t, \qquad P_1 = Q_1 g_1 + Q_0 t \tag{6.29}$$

where g_0 and g_1 are given by equation (6.10). These pricing rules can be compared to the duopoly pricing rules (6.17). Contrary to what happens with duopoly, the monopoly has a mark-up term. This occurs because of the monopolist's internalization: When increasing the price for one facility, the monopolist takes into consideration that it is actually increasing the demand for the other facility, with the resulting profit accruing to itself. Moreover, replacing g_0, g_1 with (6.8) and (6.10), (6.29) becomes, for facility 0:

$$P_0 = D_0(\alpha + \beta_0)(1 + s_0) + 3tQ_0(1 + s_0) + Q_1t$$
(6.30)

and the expression for facility 1 is analogous. The first term on the RHS of equation (6.30) is related to the congestion toll but the monopoly facility, as the duopolists, charges more than just the pure un-internalized congestion of each carrier (which would have been $1 - s_0$, rather than $1 + s_0$). The second term is the mark-up from exploitation of market power that arises given the locational preferences of consumers. The third mark-up arises because of lack of facility competition: the monopoly internalizes the interrelation of demands.

From (6.17) and (6.29) it is not immediate, however, to conclude that monopoly facility charges are higher than duopoly charges, for two reasons. First, both (6.17) and (6.29) are actually a system of fixed points, since Q_0 and Q_1 depend on both P_0 and P_1 . Second, perhaps more fundamentally, capacities will likely differ in the two cases; prices and capacities are decided simultaneously and therefore, prices can be compared in two ways (see Spence, 1975; Basso, 2005):

- (i) Compare prices as if capacities were fixed,
- (ii) Compare actual prices, taking into consideration the difference in capacities.

The first one is useful because it may represent a short-term case, but it is also useful in performing the second comparison. In what follows we will, when feasible, perform both comparisons. For this, we first look at the capacity rules under monopoly. Taking the first-order conditions for capacities, and using (6.10), (6.13) and (6.29), we get:

$$K_{h} = \left(\frac{a(\alpha + \beta_{h})(N_{h} + 1)}{m_{h} N_{h}}\right)^{1/2} Q_{h}, \qquad h = 0,1$$
(6.31)

The monopoly capacity rule (6.31) is identical to the open-loop duopoly capacity rule (6.24). Obviously, since their pricing rules are different, the consumption levels, and hence actual capacities, will be different in the two cases. Delay times will be equal however,

given the assumption about linearity of the delay function. From (6.25), (6.28) and (6.31) we have (superscript *M* stands for monopoly):

$$D(Q_0^M, K_0^M) = a \frac{Q_0^M}{K_0^M} = \left(\frac{a m_0 N_0}{(\alpha + \beta_0)(N_0 + 1)}\right)^{1/2} = D(Q_0^{do}, K_0^{do}) < D(Q_0^{dc}, K_0^{dc})$$
(6.32)

Note that the sufficient condition for an interior solution in the monopoly case is the same as (6.26). DBVD found that the duopolists would provide longer delays than the monopolist. Here we find that the duopolists provide longer delays than a monopolist only if capacity decisions are made prior to price decisions (which is the case for DBVD). When the capacity and price decisions are made simultaneously, or when capacity investments are not observable prior to price decisions, the duopolists would provide the same service quality (congestion delay) as the monopolist.

Next, we compare the monopoly and duopoly prices for given capacities. Obtaining monopoly prices for given capacities involves solving system (6.29). This leads to:

$$P_0^M(\mathbf{K}) = \frac{2t + V - c_0}{2}$$
, $P_1^M(\mathbf{K}) = \frac{2t + V - c_1}{2}$ (6.33)

Thus, given the capacities, the monopoly prices are, somewhat surprisingly, actually independent of the capacities! This means that the monopolist would charge prices (6.33) independently of whether it can choose capacities or not (provided, of course, that it leads to an interior solution). A facility's price decreases with the marginal cost of carriers at that facility, but is independent of the marginal cost of carriers at the other facility. This is in contrast to the duopoly case, where a fall in carriers' marginal cost at one facility induces a fall of the other facility's charge. This distinction between duopoly and monopoly pricing might be used as an empirically testable hypothesis.

To compare the prices, recall that the duopoly prices are derived in (6.19). From (6.19) and (6.33) we can easily show that for given capacities, the duopoly prices are, as expected, smaller than the monopoly prices:

$$P_0^M(\mathbf{K}) - P_0^*(\mathbf{K}) = \frac{t[t(2t+V-c_0)+2g_0(2t+V-c_1)]}{2(4g_0g_1-t^2)} > 0$$
(6.34)

Moreover, the fact that the monopoly pricing rules do not depend on capacities allows us to further show that the actual duopoly prices will be smaller. Notice that $\mathbf{P}^*(\mathbf{K})$ in (6.19) denotes the duopoly pricing rules for given **K**. Letting \mathbf{K}^{do} denote the actual capacities in the open-loop duopoly, we have that $\mathbf{P}^*(\mathbf{K}^{do})$ are the actual prices in the open-loop duopoly. Hence,

$$P_0^{do} = P_0^*(\mathbf{K}^{do}) < P_0^M(\mathbf{K}^{do}) = P_0^M(\mathbf{K}^M) = P_0^M \Longrightarrow P_0^{do} < P_0^M$$
(6.35)

where the first inequality follows from the fact that the monopoly prices are higher than the duopoly prices for given capacities (equation (6.34)), and the last two equalities follow from the fact that the monopoly pricing rules do not depend on capacities. Similarly, we obtain that:

$$P_0^{dc} = P_0^*(\mathbf{K}^{dc}) < P_0^M(\mathbf{K}^{dc}) = P_0^M(\mathbf{K}^M) = P_0^M \Rightarrow P_0^{dc} < P_0^M$$
(6.36)

Hence, overall we have the following result:

Proposition 6.3: For facility h (h = 0,1) we have: (i) $P_h^{do} < P_h^{dc} < P_h^M$; and (ii) $D_h^{do} = D_h^M < D_h^{dc}$. Thus, an open-loop duopoly will have lower prices than a closed-loop duopoly, which in turn will have lower prices than a monopoly. As for the service levels (congestion delays), duopoly facilities provide lower service quality than a monopoly only if capacity decisions are made prior to the facility pricing decisions. If the decisions are simultaneous, the duopolists would provide the same service level as the monopolist.

6.4.2 The social optimum

. ..

The social optimum arises when a central planner chooses facility prices \mathbf{P} and capacities \mathbf{K} to maximize social welfare. Unlike the non-vertical setting of DBVD, we now have the surplus of three agents to consider (facilities, carriers, and final consumers) rather than just two agents (facilities and final consumers). With two facilities, we need to consider five surpluses in the social-welfare function:

$$SW(\mathbf{P}, \mathbf{K}) = CS + \Phi_1 + \Phi_2 + \pi^1 + \pi^2$$
(6.37)

where CS is consumer surplus, Φ_h is the aggregate (equilibrium) profit for carriers at facility h and π^h is the (equilibrium) profit of facility h. Recall that we have identified four types of consumers defined by three locations z of the linear city: z^l , \tilde{z} and z^r ; see equations (6.1) and (6.2) and Figure 6.1. With consumers being uniformly distributed with density one per unit of length, the consumers' surplus is thus given by:

$$CS = \int_{0}^{|z'|} [V - p_0(Q_0, Q_1) - \alpha D_0 - tz)] dz + \int_{0}^{\overline{z}} [V - p_0(Q_0, Q_1) - \alpha D_0 - tz)] dz + \int_{0}^{1-\overline{z}} [V - p_1(Q_0, Q_1) - \alpha D_1 - tz)] dz + \int_{0}^{z'-1} [V - p_1(Q_0, Q_1) - \alpha D_1 - tz)] dz$$

Note that Q_0 and Q_1 -which are given by (3)- do not depend on z, whereas z^l , z^r and \tilde{z} depend on Q_0 and Q_1 . Hence, we will obtain an expression dependent on Q_0 and Q_1 . Using (6.4) to replace p_0 and p_1 both in the integrands and in z^l , z^r and \tilde{z} , and solving the integrals we get:

$$CS = \frac{t}{2} \left(3Q_0^2 + 2Q_0Q_1 + 3Q_1^2 - 4 \right).$$
 (6.38)

It might seem that CS increases in t, and it is negative if there is no consumption. However, it is important to recall that both Q_0 and Q_1 are equilibrium values, in that they depend on the level of congestion and on the value of t. Indeed, an examination of (6.3) reveals that Q_0 and Q_1 will rise as t decreases, so the overall result is that CS actually falls as t increases, as expected. Also, recall that the maintained assumption has been that V is sufficiently large so that everyone in the [0,1] interval consumes. This implies that the minimum values of Q_0 and Q_1 for which the above CS expression is valid are when both are equal to 1 (each facility gets $\frac{1}{2}$ consumer from each side, left and right). Therefore, CS is never less than 2t.

Regarding the carriers' aggregate profit at a facility, it is straightforward, from (6.6), (6.4) and symmetry, to obtain:

$$\Phi_0(\mathbf{P}, \mathbf{K}) = (2t + V - c_0)Q_0 - Q_0P_0 - 3tQ_0^2 - tQ_0Q_1 - (\alpha + \beta_0)Q_0D_0$$
(6.39)

With these expressions for CS and Φ_h and the expressions for the facilities' profits given by (6.16), the welfare function (6.37) can be written as:

$$SW(\mathbf{P}, \mathbf{K}) = (2t + V - c_0)Q_0 + (2t + V - c_1)Q_1 - \frac{t}{2}(3Q_0^2 + 2Q_0Q_1 + 3Q_1^2) - 2t$$

- $(\alpha + \beta_0)Q_0D_0 - (\alpha + \beta_1)Q_1D_1 - m_0K_0 - m_1K_1$ (6.40)

Notice that SW above is not directly a function of P_0 and P_1 . Instead, it is a function of Q_0 and Q_1 and, through them, a function of prices.

At the social optimum, the prices and capacities are set such that social welfare is maximized. The first-order condition with respect to P_0 is,

$$\frac{dSW}{dP_0} = \frac{\partial SW}{\partial Q_0} \frac{\partial Q_0}{\partial P_0} + \frac{\partial SW}{\partial Q_1} \frac{\partial Q_1}{\partial P_0} + \frac{\partial SW}{\partial P_0} = 0$$

Calculating this –noticing that $\partial SW / \partial P_0 = 0$ – and using Ω^0 in (6.7) to obtain an expression for $(2t+V-c_0)$, (6.8) for the delay function and (6.11)-(6.12) for the derivatives of Q_0 with respect to prices, we get:

$$\left[P_{1}+3t\frac{Q_{1}}{N_{1}}-a(\alpha+\beta_{1})\frac{Q_{1}}{K_{1}}\left(\frac{N_{1}-1}{N_{1}}\right)\right]t=\left[P_{0}+3t\frac{Q_{0}}{N_{0}}-a(\alpha+\beta_{0})\frac{Q_{0}}{K_{0}}\left(\frac{N_{0}-1}{N_{0}}\right)\right]g_{0} \quad (6.41)$$

Similarly, the first-order condition with respect to P_1 leads to:

$$\left[P_{1} + 3t\frac{Q_{1}}{N_{1}} - a(\alpha + \beta_{1})\frac{Q_{1}}{K_{1}}\left(\frac{N_{1} - 1}{N_{1}}\right)\right]g_{1} = \left[P_{0} + 3t\frac{Q_{0}}{N_{0}} - a(\alpha + \beta_{0})\frac{Q_{0}}{K_{0}}\left(\frac{N_{0} - 1}{N_{0}}\right)\right]t \quad (6.42)$$

Since the terms in brackets on the LHS of (6.41) and (6.42) are the same, the terms in brackets on the right-hand side (RHS) of (6.41) and (6.42) are the same, and $g_0 > t$, $g_1 > t$ and $g_0 \neq g_1$, equations (6.41) and (6.42) hold only if each of the bracketed terms is zero. Using (6.8), this gives rise to the following social pricing rules:

$$P_0 = D_0(\alpha + \beta_0)(1 - s_0) - 3tQ_0s_0, \qquad P_1 = D_1(\alpha + \beta_1)(1 - s_1) - 3tQ_1s_1 \tag{6.43}$$

where $s_h = 1/N_h$ is a carrier's market share at facility *h*. This pricing rule for each facility is conceptually similar to the ones obtained by Pels and Verhoef (2004) for the case of two distant airports servicing two carriers, Basso (2005) for the case of two distant airports servicing *N* carriers, and Zhang and Zhang (2006) for the case of a single airport servicing *N* carriers. The socially optimal price at a facility consists of a *congestion term* –by which the facility charges to each carrier the un-internalized congestion it produces–, and a markdown, the *market-power term*, by which the facilities 'subsidize' the carriers so as to countervail the market power by monopoly or oligopoly carriers and induce the allocatively efficient output. One consequence of such subsidy is that the facilities may not recover their costs especially if the market-power term is large, even though we have constant returns to scale in the provision of capacity and the delay function is linear.¹⁷

Furthermore, notice that when there is a large number of carriers $(s_h \rightarrow 0)$, the carriers no longer have market power and hence the second terms on the RHS of price equations (6.43) vanish. Similarly, when there is a monopoly carrier at a facility $(s_h = 1)$ the monopolist perfectly internalizes congestion; consequently, there is no need to correct for congestion and so the first terms on the RHS of (6.43) vanish. Finally, (6.43) is easily comparable to the monopoly pricing equation (6.30). It can be also manipulated to yield:

$$P_0 = Q_0 g_0 \frac{N_0 - 1}{N_0 + 1} - 3t Q_0, \qquad P_1 = Q_1 g_1 \frac{N_1 - 1}{N_1 + 1} - 3t Q_1 \qquad (6.44)$$

which are more easily comparable to the duopoly pricing rules (6.17).

¹⁷ Brueckner (2002), Basso (2005) and Zhang and Zhang (2006) have obtained the same result under different model settings. All these studies have explicitly considered imperfect competition in the carriers' market. When such imperfect competition is absent, earlier studies (Morrison, 1983; Zhang and Zhang, 1997; and De Borger and Van Dender, 2006) have shown that under the constant returns to scale in the provision of capacity and a linear delay function, the optimal pricing and optimal provision of capacity lead to exact cost recovery for a congestible facility (e.g., airport). The issue of budget adequacy is further discussed by Basso (2005) in the context two distant airports, but the conclusions there apply to this competing-facilities case as well: two part-tariffs, or cost-recovery two part tariffs if the carriers do not make enough profits, may resolve the problem. If fixed fees are not feasible for some reason, the less efficient alternative of Ramsey-Boiteux prices is called for.

To derive the socially optimal capacities, it is useful to point out that the pricing rules (6.43) are obtained as if we were maximizing directly in terms of (Q_0, Q_1) rather than (P_0, P_1) , because the pricing rules are in fact derived from $\partial SW / \partial Q_h = 0$. Hence:

$$\frac{dSW}{dK_h} = \frac{\partial SW}{\partial Q_0} \frac{\partial Q_0}{\partial K_h} + \frac{\partial SW}{\partial Q_1} \frac{\partial Q_1}{\partial K_h} + \frac{\partial SW}{\partial K_h} = \frac{\partial SW}{\partial K_h} = 0$$

From $\partial SW / \partial K_h = 0$ and (6.40), it follows immediately that the social capacity rules are given by:

$$K_{h} = \left(\frac{a(\alpha + \beta_{h})}{m_{h}}\right)^{1/2} Q_{h}, \qquad h = 0,1$$
(6.45)

The capacity rules (6.45) can be compared to the monopoly capacity rules (6.31), giving rise to:

Proposition 6.4: Conditional on facility charges, the monopoly capacity rules are the same as the socially-optimal capacity rules if and only if $N_h \rightarrow \infty$, i.e., the downstream carriers' markets are perfectly competitive.

Proposition 6.4 shows that when the downstream markets are imperfectly competitive, the monopoly capacity rules will be different from the social capacity rules. This is in contrast to what was found by DBVD in their analysis without an intermediate, carriers' market; they found that the monopoly and socially optimal capacity rules were identical. Their result had in fact a precedent in Oum, et al. (2004), who analyzed price and capacity decisions by a single congestible airport. Since they did not formally derive the airport's demand from the equilibrium of the airlines' market, their setting is actually quite close to DBVD's, with the exception that DBVD had two facilities with interdependent demands. Our result here shows that, when one takes into consideration that the congestible facilities may be upstream providers of an input, as in the case of airports, seaports or telecommunication networks, the monopoly capacity rules will coincide with the socially optimal rules only for the case of atomistic carriers (see also Basso, 2005; Zhang and Zhang, 2006).

Both Oum et al. and DBVD correctly pointed out that, since the pricing rules are different, consumption levels will be different, and hence actual capacities will be different. However, taking advantage of the assumption of a linear congestion function, DBVD showed that, in their case where the facilities interact directly with final consumers, the monopolist provides exactly the same service quality –i.e., the same level of delays– as welfare-maximizing facilities. In our case, from (6.45) and (6.32) it follows that:¹⁸

¹⁸ Note that the sufficient condition for an interior solution at the social optimum remains the same as (26), that is, $m_h \le a(\alpha + \beta_h)$.

$$D(Q_{h}^{W}, K_{h}^{W}) = a \frac{Q_{h}^{W}}{K_{h}^{W}} = \left(\frac{a m_{h}}{\alpha + \beta_{h}}\right)^{1/2} > \left(\frac{a m_{h} N_{h}}{(\alpha + \beta_{h})(N_{h} + 1)}\right)^{1/2} = D(Q_{h}^{M}, K_{h}^{M})$$
(6.46)

where W stands for welfare maximization. Inequality (6.46) thus leads to:

Proposition 6.5: Monopoly pricing and capacity choices result in a higher level of service quality (shorter congestion delays) than the social optimum.

Proposition 6.5 shows that when the facilities are input providers, the monopolist would no longer provide the same level of service quality as welfare-maximizing facilities. However, an important issue is: if the congestion levels were equal, does that mean that the monopolist is providing the *socially optimal* level of service quality? The answer is no. As it has been pointed out, capacities and prices are decided jointly so they cannot be analyzed separately. And since service quality is a result of both the level of demands induced and the capacities chosen, it cannot be looked at on its own; just as with capacity, delays must be analyzed together with the pricing rules. Hence, in the first-best case, the service quality obtained is optimal given the W pricing but it would no longer be optimal if the pricing rule is different. In other words, both the capacities and the delay levels are most likely *not a second best*: If we were to look for optimal capacities subject to the monopoly pricing, different capacity rules, and consequently different congestion levels would follow (see Basso, 2005).

This argument may seem a little bit difficult to see in the present case, given that the central planner and the monopolist will have different capacity rules to start with. Nevertheless, we can make it more transparent by analyzing a fourth type of facility ownership: the monopolist vertically integrates with the carriers at the facilities. As we shall show in the next Section, in this case, this *hyper-monopolist* will have exactly the same service level as the central planner, just as in the case of DBVD. However, this will not imply that the facilities provide the socially optimal quality level in a second-best sense.

6.5 Vertical Integration

Consider now a vertical-integration case, in which a single owner of the facilities tries to maximize the sum of the facilities' profit and the carriers' profit. This case is relevant in the real world; for example, in the case of airports, it has been often argued that strategic collaboration between airports and airlines would solve the incentive and coordination problems regarding capacity and pricing in the vertical structure (see, e.g., Beesley, 1999; Forsyth, 1997; Starkie, 2001; for a more complete list see Basso, 2005).

Thus, the problem of the hyper-monopolist is:

$$\max_{\mathbf{P},\mathbf{K}}\sum_{h}(\Phi_{h}+\pi^{h}).$$

The objective function is easily obtainable as:

$$\sum_{h} (\Phi_{h} + \pi^{h}) = (2t + V - c_{0})Q_{0} + (2t + V - c_{1})Q_{1} - t(3Q_{0}^{2} + 2Q_{0}Q_{1} + 3Q_{1}^{2}) - (\alpha + \beta_{0})Q_{0}D_{0} - (\alpha + \beta_{1})Q_{1}D_{1} - m_{0}K_{0} - m_{1}K_{1}$$
(6.47)

Notice that just like the welfare function (6.40), this objective function depends indirectly on P_0 and P_1 . The pricing rules are obtained in the same fashion as in the social-optimum case, and are given by:¹⁹

$$P_0 = D_0(\alpha + \beta_0)(1 - s_0) + 3tQ_0(1 - s_0) + Q_1t$$
(6.48)

This pricing equation is conceptually similar to the ones obtained by Basso (2005) for the case of two distant airports that vertically integrate with the airlines, and by Basso and Zhang (2006) for the case of peak-period pricing by a vertically integrated airport. It shows that the price consists of a *congestion toll term* –by which the facility charges each carrier for the un-internalized congestion it produces– and a mark-up. The congestion-toll term is represented by the first term on the RHS of (6.48), which depends on the number of carriers at the facility. The toll is greatest for atomistic carriers ($s_0 \rightarrow 0$) and is zero for a monopoly carrier ($s_0 = 1$). By the mark-up, the facility increases the carriers' marginal costs, thereby inducing the cartel level of output and maximizing the carriers' joint profit. In particular, the mark-up takes into account both the competition between the carriers within a facility (represented by the other facility and its carriers (third term). Note that equation (6.48) is easily comparable to (6.43) for the central-planner case and (6.30) for the monopoly case.²⁰

Like the pricing rule, the capacity rule can also be obtained in the same fashion as we obtained the social capacity rule, that is, by realizing that the pricing rules are in effect obtained from $\partial \left(\sum_{h} \Phi_{h} + \pi^{h} \right) / \partial Q_{h} = 0$. It turns out that the hyper-monopolist's capacity rule is the same as the central planner rule (6.45); as a consequence, the level of service quality (delays) is, from (6.46), also the same:

$$D(Q_0^{VI}, K_0^{VI}) = a \frac{Q_0^{VI}}{K_0^{VI}} = \left(\frac{a m_0}{\alpha + \beta_0}\right)^{1/2} = D(Q_0^W, K_0^W)$$
(6.50)

where superscript VI stands for the vertical-integration case. Does this mean that the hypermonopolist would have no incentive to distort quality and would be providing the optimal level of delays? As we argued earlier, the level of delays in (6.50) is optimal given the social optimal pricing. If a constrained central planner is forced to price as the hyper-

¹⁹ To save notations, in this Section various expressions will be written for facility 0 only, rather than for both facilities 0 and 1.

²⁰ It can be further modified to obtain an equivalent expression, $P_0 = Q_0 g_0 (N_0 - 1) / (N_0 + 1) + tQ_1$, which can be more easily compared to the duopoly pricing rule (17), or the alternative social pricing rule (44).
monopolist does, it would choose a different capacity rule and thereby induce a different – in fact, probably superior– service level. We formally state the result in Proposition 6.6:

Proposition 6.6: When the monopolist vertically integrates with the carriers at the facilities, it would provide the same congestion level as the central planner. Nevertheless, the hyper-monopoly service level is not socially optimal in a second-best sense. In effect, in the fully ex-ante symmetric case, it is too low with respect to the second best.

<u>Proof</u>: The first part of the Proposition has been indicated by (6.50). To prove the second and third parts, we consider the fully ex-ante symmetric case, namely, $N_0 = N_1 = N$, $\beta_0 = \beta_1 = \beta$ and $m_0 = m_1 = m$. Let $\tilde{S}W(\mathbf{K}) \equiv SW(\mathbf{P}^{VI}(\mathbf{K}))$ be a second-best social welfare function, where $\mathbf{P}^{VI}(\mathbf{K})$ represents the hyper-monopolist pricing rule. Hence we have:

$$\frac{d\tilde{S}W}{dK_0} = 0 \Leftrightarrow \frac{dSW}{dK_0}\Big|_{\mathbf{P}^{W}(\mathbf{K})}$$

Calculating $\frac{dSW}{dK_0} = \frac{\partial SW}{\partial Q_0} \frac{\partial Q_0}{\partial K_0} + \frac{\partial SW}{\partial Q_1} \frac{\partial Q_1}{\partial K_0} + \frac{\partial SW}{\partial K_0}$ from (6.40), we get:

$$\begin{bmatrix} P_0 + 3t \frac{Q_0}{N_0} - a(\alpha + \beta_0) \frac{Q_0}{K_0} \left(\frac{N_0 - 1}{N_0} \right) \end{bmatrix} \frac{\partial Q_0}{\partial K_0} + \begin{bmatrix} P_1 + 3t \frac{Q_1}{N_1} - a(\alpha + \beta_1) \frac{Q_1}{K_1} \left(\frac{N_1 - 1}{N_1} \right) \end{bmatrix} \frac{\partial Q_1}{\partial K_0} + a(\alpha + \beta_0) \left(\frac{Q_0}{K_0} \right)^2 - m_0 = 0$$

Evaluating this at $\mathbf{P}^{VI}(\mathbf{K})$, which is given by (6.48) an its counterpart for facility 1, yields:

$$\frac{dSW}{dK_0}\Big|_{\mathbf{P}^{V_1}(\mathbf{K})} = \left[3tQ_0 + tQ_1\right]\frac{\partial Q_0}{\partial K_0} + \left[3tQ_1 + tQ_0\right]\frac{\partial Q_1}{\partial K_0} + a(\alpha + \beta_0)\left(\frac{Q_0}{K_0}\right)^2 - m_0 = 0. \quad (6.51)$$

From (6.13) and (6.14) it follows that:

$$\frac{\partial Q_0}{\partial K_0} = \frac{g_1 (N_0 + 1) a(\alpha + \beta_0) Q_0}{N_0 (g_0 g_1 - t^2) {K_0}^2}, \qquad \frac{\partial Q_0}{\partial K_0} = -\frac{t (N_0 + 1) a(\alpha + \beta_0) Q_0}{N_0 (g_0 g_1 - t^2) {K_0}^2}.$$

Replacing this in (6.51) and then looking into the symmetric-capacities solution for the fully symmetric case, we then obtain:

$$a(\alpha+\beta)\frac{Q^2}{K^2}\left[\frac{(N+1)4t}{N(g+t)}+1\right]-m=0.$$

Since $\Psi > 0$, we can conclude that:

$$D(\tilde{Q}^{W}, \tilde{K}^{W}) = a \frac{\tilde{Q}^{W}}{\tilde{K}^{W}} = \left(\frac{a m}{(\alpha + \beta)(1 + \Psi)}\right)^{1/2} < D(Q^{W}, K^{W}) = D(Q^{VI}, K^{VI}) \qquad \blacksquare$$

Hence, the service level provided by the hyper-monopolist is not optimal: at the very least, in the fully symmetric case, a constrained central planner who is forced to use the same pricing rule as the hyper-monopolist would choose a higher service level. The congestion delays of the hyper-monopolist are not second best. The intuition is simple: given that prices will be higher, the central planner will compensate consumers and carriers by providing a higher quality service. Overall, this shows that, just as with capacity, congestion delays cannot be analyzed separately from pricing.

6.6 Concluding Remarks

Our main objectives in writing this paper are to contribute to the understanding of rivalry between congestible facilities -such as airports, seaports and telecommunication networksand to explicitly incorporate downstream carriers -both their decision-making and market structure -- into the analysis of facility competition. In our vertical facility-carrier-consumer structure with imperfectly competitive output (carriers) markets, we found that the duopolists' equilibrium prices increase with both the consumers' value of time and the carriers' cost sensitivity to congestion delays; entrance of a new carrier to any of the facilities depresses the prices charged by both facilities; and lower marginal cost of the carriers at one facility will induce a higher facility price at that facility but a lower facility price at the other facility. In terms of service level, we found that the duopoly facilities provide longer congestion delays than a monopolist only if capacity decisions are made prior to the facility pricing decisions. When the capacity and pricing decisions are made simultaneously, or when capacity investments are not observable prior to the pricing decisions, the duopolists would provide the same level of service quality (delays) as the monopolist. Furthermore, a monopolist would provide a higher level of service than the central planner. Our analysis showed that when the monopolist vertically integrates with the carriers at the facilities, it would provide the same congestion level as the central planner. Nevertheless, the monopoly service level is not socially optimal in a second-best sense. In effect, in the fully ex-ante symmetric case, it is too low with respect to the second best. Finally, despite the fact that we assume constant returns to scale and a linear delay function, the optimal pricing and capacity will probably not lead to cost recovery, owing to market power at the carriers' level.

Our analysis has also pointed to (at least) two areas for future research. First, it would be desirable to be able to rank the alternatives we have discussed in this paper, in terms of social welfare. For example, we know that, in terms of consumer surplus, the open-loop duopoly dominates both the closed-loop duopoly and the monopoly, because it has both lower prices and shorter delays. But it is not clear whether the closed-loop duopoly is superior to monopoly (because of its lower prices but higher delays as compared to monopoly), or under which conditions it will be. The same goes for any comparisons with the vertical integration case. To compare the final outcomes for consumer surplus and social welfare under different ownerships and market structures is extremely hard analytically; and so it would probably need to be undertaken numerically.

A second, perhaps more interesting, issue that could be further studied is how the prices at each facility compare to one another in the different cases. For example, it is easy to see that the price difference in the monopoly case will always be (even if capacities are fixed) equal to half the difference of carriers' marginal costs. In the duopoly case, however, the difference will depend on many other parameters of the output (carriers) market, such as the number of firms. If the only source of asymmetry is carriers' marginal cost, the difference between prices in the duopoly case would be smaller than in the monopoly case. However, it is reasonable to ask: might it happen that with enough asymmetry the monopoly and duopoly price differentials have opposite signs? Again, it seems to us that such an analysis would be feasible only numerically.

6.7 References

- Abiru, M. (1988) Vertical Integration, Variable Proportions and Successive Oligopolies, Journal of Industrial Economics, 36, 315-325.
- Adler, N. and Berechman, J. (2001) Measuring airport quality from the airlines' viewpoint: An application of data envelopment analysis, *Transportation Policy*, 3, 171-181.
- Baake, P. and Mitusch, K. (2004) Competition with congestible Networks, *Discussion Paper* 402, DIW, Berlin.
- Basso, L.J. (2005) Airport ownership: Effects on pricing and capacity, *Working Paper*, Sauder School of Business, The University of British Columbia (<u>Http://Ssrn.Com/Abstract=849584</u>).
- Basso, L.J. (2006) On input markets surplus and its relation to the downstream market game, *Working Paper*, Sauder School of Business, The University of British Columbia.
- Basso, L.J. and Zhang, A. (2006) Sequential peak-load pricing in a vertical setting: The case of airports and airlines, *Working Paper*, Sauder School of Business, The University of British Columbia.
- Beesley, M.E. (1999) Airport regulation, in M.E. Beesley (ed.), *Regulating Utilities: a New Era?*: Institute of Economic Affairs, London.
- Bradley, M.A. (1998) Behavioural models or airport choice and air route choice, in J. de D. Ortuzar, D. Hensher and S. Jara-Diaz (eds.), *Travel Behaviour Research: Updating* the State of Play, Elsevier, Amsterdam, pp. 141-159.
- Braid, R.M. (1986) Duopoly pricing on congested facilities, *Working Paper No.322*, Department of Economics, University of Columbia.
- Brander, J.A. and Zhang, A. (1990) Market conduct in the airline industry, *Rand Journal of Economics*, 21, 567-583.
- Brander, J.A. and Zhang, A. (1993) Dynamic behaviour in the airline industry, International Journal of Industrial Organization, 11, 407-435.
- Brueckner, J.K. (2002) Airport congestion when carriers have market power, American Economic Review, 92(5), 1357-1375.
- Brueckner, J.K. (2005) Internalization of airport congestion: A network analysis, International Journal of Industrial Organization, 23(7-8), 599-614.
- De Borger, B. and Van Dender, K. (2006) Prices, capacities and service levels in a congestible Bertrand duopoly, *Journal of Urban Economics*, forthcoming.

- De Neufville, R. (1995) Management of multi-airport systems, Journal of Air Transport Management, 2, 99-110.
- De Palma, A. and Leruth, L. (1989) Congestion and game in capacity: A duopoly analysis in the presence of network externalities, *Annales D'Economie et de Statistique*, 15/16, 389-407.
- Dixit, A. (1986) Comparative statics for oligopoly, International Economic Review, 27, 107-122.
- Gillen, D.W. and Lall, A. (2004) Competitive advantage of low-cost carriers: some implications for airports, *Journal of Air Transport Management*, 10, 41-50.
- Gillen, D.W. and W.G. Morrison (2003) Bundling, integration and the delivered price of air travel: Are low cost carriers full service competitors?, *Journal of Air transport Management*, 9, 15-23.
- Forsyth, P. (1997) Price regulation of airports: Principles with Australian applications, *Transportation Research Part E-Logistics and Transportation Review*, 33(4), 297-309.
- Fournier, G.M., Hartmann, M.E. and Zuehlke, T. (2006) Airport substitution by travellers: why do we have to drive to fly? in D. Lee (ed.), *Advances in Airline Economics*, Vol. 2, forthcoming, Elsevier.
- Fudenberg, D. and Tirole, J. (1984) The fat-cat effect, the puppy-dog ploy, and the lean and hungry look, *American Economic Review*, 74(2), 361-366.
- Hart, O. and Tirole, J. (1990) Vertical integration and Market Foreclosure, *Brookings* Papers on Economic Activity. Microeconomics, 1990, pp. 205-286.
- Ishii, J., K. Van Dender and S. Jun (2005) Air travel choices in multi-airport markets, *Working Paper*, Department of Economics, University of California at Irvine.
- Lee, E. and Staelin, R. (1997) Vertical Strategic Interaction: Implications for Channel Pricing Strategy, *Marketing Science*, 2, 161-190.
- Mason, K.J. (2000) The propensity of business travellers to use low cost airlines, *Journal of Transport Geography*, 8, 107-119.
- Morrison, S.A. (1983) Estimation of long-run prices and investment levels for airport runways, *Research in Transportation Economics*, 1, 103-130.
- Ndoh, N.N., Pitfield, D.E. and Caves, R.E. (1990) Air transportation passenger route choice: A nested multinomial logit analysis, in: M.M. Fischer, P. Nijkamp and Y.Y. Papageorgiou (eds.), *Spatial Choices and Process*, North-Holland, Amsterdam, pp. 349-365.

- Oum, T., Zhang, A. and Zhang, Y. (2004) Alternative forms of economic regulation and their efficiency implications for airports, *Journal of Transport Economics and Policy*, 38, 217-246.
- Pels, E., Nijkamp, P. and Rietveld, P. (1997) Substitution and complementarity in aviation: Airports vs. airlines, *Transportation Research E*, 33, 275-286.
- Pels, E., Nijkamp, P. and Rietveld, P. (2000) Airport and airline competition for passengers departing from a large metropolitan area, *Journal of Urban Economics*, 48, 29-45.
- Pels, E., Nijkamp, P. and Rietveld, P. (2001) Airport and airline choice in a multiple airport region: An empirical analysis for the San Francisco Bay area, *Regional Studies*, 35(1), 1-9.
- Pels, E., Nijkamp, P. and Rietveld, P. (2003) Access to a competition between airports: A case study for the San Francisco Bay Area, *Transportation Research A*, 37(1), 71-83.
- Pels, E. and Verhoef, E.T. (2004) The economics of airport congestion pricing, *Journal of Urban Economics*, 55(2), 257-277.
- Salop, S. and Scheffman, D. (1987) Cost-Raising Strategies, Journal of Industrial Economics, 36(1), 19-34.
- Spence, A.M. (1975) Monopoly, quality, and regulation, *Bell Journal of Economics*, 6(2), 417-429.
- Starkie, D. (2001) Reforming UK airport regulation, Journal of Transport Economics and Policy, 35, 119-135.
- Van Dender, K. (2005) Duopoly prices under congested access, *Journal of Regional Science*, 45(2), 343-362.
- Young, A.R. (1991) Vertical Structure and Nash Equilibrium: A note, *Journal of Industrial Economics*, 39(6). 717-722.
- Zhang, A. and Zhang, Y. (1997) Concession revenue and optimal airport pricing, Transportation Research Part E-Logistics and Transportation Review, 33(4), 287-296.
- Zhang, A. and Zhang, Y. (2006) Airport capacity and congestion when carriers have market power, *Journal of Urban Economics*, forthcoming.

7 CONCLUSIONS

7.1 Summary of the Thesis

This thesis has researched the issue of airport pricing and privatization through five papers: a literature review of the subject, four theoretical papers that looked at different aspects of airport pricing and capacity decisions under different types of ownership, and one paper that looked at the general relation between input and output pricing.

In the first paper, in Chapter 2, we presented an up-to date review of theoretical papers of airport pricing, highlighting the main results in the literature. We claimed that the models in the literature can be grouped into two broad *approaches*: the *traditional approach*, which has used a classical partial equilibrium model where the demand for the airport depends on airport charges and on congestion costs of both passengers and airlines, and the airline market is not formally modeled. And the *vertical structure approach* which instead recognizes that airports provide an input for the airline market –which is modeled as a rather simple oligopoly– and that it is the equilibrium of this downstream market which determines the airport's demand: the demand for airports is therefore a derived demand.

We showed that the questions examined with the two approaches have not perfectly overlapped. The *traditional approach* has been wider in scope, having been used to analyze issues such as optimal capacity investments, the effect of concession revenues, privatization and so on. The *vertical structure* approach on the other hand, has mainly focused on calculating the additional toll that airlines should be charged to attain maximization of social welfare. But the two approaches have not only examined different questions, but also grew somewhat disconnected, which raises the questions of transferability of results. In the third paper, in Chapter 4, I showed that abstracting from the airline market, as is done in the *traditional approach* is a reasonable approximation only when airlines behave competitively, but it is not when airlines have market power. In the latter case, the derived demand for the airport would not be dependent only on its *full-price*, as it is assumed. As a result, the integration of the airport demand with respect to the full price, which is said to capture *consumer surplus*, would not adequately capture the surpluses of passengers and airlines because market power and congestion effects preclude it.

This result however, was an *ad-hoc* answer to a broader question: how "consumers' surplus" measures coming from areas under input demand curves capture the effects of direct purchasers and downstream final consumers? This question pertains to the economic analysis of input markets in general, and because of that, it is central to the analysis of airport pricing as well. The second paper in Chapter 3 looked at answering the question in the more general possible way, within the context of differentiable oligopoly games. It was explained there that the most general results available in the literature regarding the relation between input and output markets surplus, hinge on a number of strong simplifying assumptions. I argued that those assumptions can be relaxed by linking the input markets surplus question to results from another stream of literature, which characterizes a function that firms in oligopoly collectively, yet unintentionally, maximize. I show that the input

markets surplus change measure (obtained by integration under the input demands derived from the equilibrium of a downstream oligopoly game) is equal to the change in a function for which critical points coincide with the equilibria of the downstream game. In particular, if the downstream game is *potential*, the input market surplus is shown to be equal to the change in the exact potential function. The proposition proved synthesizes and generalizes the established results on the relation between input and final market surplus measures, providing guidance to policy analysts who seek to infer the total welfare effects of input market price changes from information on the input market demands only. In particular, it allows us to assess how much information about the downstream market is captured in the derived demand for inputs even when there are externalities in production downstream, which is precisely the case of airlines given the congestion problem at airports.

In the third paper, in Chapter 4, I used a model of vertical relations between airports and airlines (in order to adequately set-up the central planner benchmark case), to examine both analytically and numerically, how ownership affects airports' prices and capacities. This paper was motivated by the fact that it has been argued that privatized airports would charge more efficient congestion prices and would be more responsive to market incentives for capacity expansions. Furthermore, the privatized airports would not need to be regulated since price elasticities are low, so allocative inefficiencies would be small, and collaboration between airlines and airports, or airlines countervailing power, would solve the problem of airports' market power. Results showed a somewhat unattractive picture for unregulated privatization when compared to first- and second-best benchmark cases though: (i) private airports would be too small in terms of both, traffic and capacity and, despite the fact that they may be less congested, they would induce important deadweight losses; (ii) the arguments that airlines countervailing power or increased cooperation between airlines and airports may make regulation unnecessary seem to be overstated; and (iii) things may deteriorate further if privatization is done on an airport by airport basis rather than in a system. Also, I showed that two features of air travel demand that have not been incorporated previously in the literature -demand differentiation and schedule delay cost- play important roles on airports' preferences regarding the number of airlines using the airport.

The model of Chapter 4 and most airport pricing models are usually models of congestion pricing and not peak-load pricing in that, even if there is more than one period, the demands between periods are not interdependent. Hence, the only way to fight excess usage is to dampen the demand. When periods are interdependent, pricing can be used not only to dampen the demand but to re-distribute consumers through the periods, 'flattening' the demand curve. Proponents of privatization have argued that private airports would use more efficient peak-load pricing schemes. The fourth paper, in Chapter 5 investigated airport peak-load pricing (PLP), analyzing both the price level and price structure with vertically differentiated peak and off-peak travel. One of the main features of the model is that both players, airport and airlines, may use peak-load pricing, giving rise, potentially, to sequential peak-load pricing. We found that, compared to the *public* (benchmark) airport, a private, profit-maximizing airport would charge both higher peak and off-peak runway prices, as well as a higher peak/off-peak price differential. As a consequence, airport privatization would lead to both fewer total air passengers and fewer passengers using the

premium peak hours for their travel, both of which reduce social welfare. And although those passengers who still use the peak period would benefit from less congestion delays, overall it would not be economically efficient to have such a lower level of peak congestion. Moreover, the analysis showed that whilst private airports will always use peak-load pricing a public airport may, somewhat surprisingly, actually charge a peak price that is lower than the off-peak price, as a mean to *correct* the peak-load pricing used by airlines. A private airport strategically collaborating with the airlines would have fewer incentives to use peak-load pricing than a pure private airport.

Now, the model of Chapter 4 was sort of the worse-case scenario for unregulated privatization, because the airports do not really face any type of competition. Hence, in the final paper, in Chapter 6, we investigated rivalry between congestible input providers, i.e. airports or seaports serving the same region, and its effects on facility charges, capacities and congestion delays. We found that the duopolists' equilibrium prices would increase with both the consumers' value of time and the carriers' cost sensitivity to congestion delays; entrance of a new carrier to *any* of the facilities would depress the prices charged by *both* facilities; and lower marginal cost of the carriers at one facility. In terms of service level, we found that the duopoly facilities would provide longer congestion delays than a monopolist only if capacity decisions were made prior to the facility pricing decisions. When the capacity and pricing decisions are made simultaneously, the duopolists would provide the same level of service quality (delays) as the monopolist. Furthermore, monopoly pricing and capacity choices would result in a higher level of service quality (shorter delays) than the social optimum.

7.2 Future work

There are many areas of future research that can be identified. They range from those that seem to be straightforward extensions of models already established, to some more fundamental questions, related to actual policy, but which would probably require to step a little further away from the models already known in the literature. Starting from the former, vertical structure models should be used to reassess those issues that have been examined only with the traditional approach. Among others, the effects of concession revenues on budget adequacy, the efficiency of alternative regulation mechanisms, or the importance of indivisibilities in capacity expansion.

A second line of future research would have to deal with the pricing of airport networks (both private and public cases), and airlines' choice of route structure. Most of the papers in the literature –and particularly the papers in this thesis– look at airports in isolation or at round trip travel between two airports at most. However, real airport networks are more complex than that; and although some good progress has been made on this subject (as discussed in Chapter 2), there is still work to do because in two out of three papers there is no route structure decisions from the part of airlines, and it would be through route structure decisions that airports may actually compete: they would be competing for connecting passengers. The third paper on the other hand, although having route structure

decisions, does not consider congestion, capacity decisions or airline competition. Hence, further work in the pricing of airport networks –including effects of privatization and regulation mechanisms– should be pursued. Note also that, in most cases, all the airports have been assumed to be national.

But perhaps the most important aspects of future research have to do with actual policies and, to date, important unanswered questions. Let me mention a few that I think are important. First, it is seldom true that airports are priced as in a system, and it is seldom true that airport managers have access to all the information that they would need to do what is best. Hence, how should public airports be priced when they are not in a system, and when information is incomplete? And given this, what are the costs and gains of privatization? And what would be a good and feasible regulation mechanism? Second, most of the papers in the literature –and particularly the papers in this thesis– have assumed that public airports would maximize social welfare. Because of this, one obtains that the airport may need to subsidy certain city-pair markets to restore allocative efficiency lost by market power at the airline level. But, is it reasonable to ask this from an airport manager? After all, the airline industry went through a large process of deregulation and it would seem here that regulation would be being re-imposed through airport pricing. What is then a reasonable objective function for a public airport? Charging only the uninternalized congestion may not be the best idea if market power effects are too strong. Could it be better than an airport seeks to maximize throughput subject to budget adequacy? Or to minimize delays subject to a throughput constraint and budget adequacy? Third and lastly, in this thesis I have discussed mainly about pricing approaches. But a different answer has been attempted in European airports, namely, the more administrative approach of slot allocation. What would be the optimal allocation mechanism? And how would that mechanism compare to the best congestion pricing alternative?

APPENDICES

APPENDIX A

Appendix for Chapter 3: On input market surplus and its relation to the downstream market game

A.1) Conjectural variations game with many inputs Proof that $H_1(\mathbf{q}, \mathbf{w}) = \sum_i \pi^i(\mathbf{q}, \mathbf{w}) + ((n-1-\nu)/n)CS(Q)$ fulfills condition (3.2).

Recall that $CS(Q) = \int_{0}^{Q} P(Z) dZ - QP(Q)$ and $\sum_{i} \pi^{i}(\mathbf{q}, \mathbf{w}) = QP(Q) - \sum_{i} C(q_{i}, \mathbf{w})$. Replacing these two in H_{1} leads to:

$$H_1(\mathbf{q}, \mathbf{w}) = \frac{1+\nu}{n} QP(Q) + \left(\frac{n-1-\nu}{n}\right) \int_0^Q P(Z) dZ - \sum_i C(q_i, \mathbf{w})$$

And then, differentiating with respect to q_i we get:

$$\frac{\partial H_1(\mathbf{q}, \mathbf{w})}{\partial q_j} = P(Q) + \frac{1+\nu}{n} Q P'(Q) - C_q(q_j, \mathbf{w})$$
(a.1)

Next, it is easy to see that

$$\frac{\partial \pi^{j}(\mathbf{q}, \mathbf{w})}{\partial q_{j}} = P(Q) + (1 + \nu)q_{j}P'(Q) - C_{q}(q_{j}, \mathbf{w})$$
(a.2)

And since the Nash Equilibrium of the oligopoly game is symmetric, from (a.1) and (a.2) we get

$$\frac{\partial H_1(\mathbf{q}^*, \mathbf{w})}{\partial q_j} = P(Q^*) + \beta Q^* P'(Q^*) - C_q(q_j^*, \mathbf{w}) = \frac{\partial \pi^j(\mathbf{q}^*, \mathbf{w})}{\partial q_j}$$

A.2) Externalities (congestion) in production game:
Proof that
$$H_2(\mathbf{q}, w) = \sum_i \pi^i (\mathbf{q}, w) + ((n-1)/n)CS(Q) + (n-1) \int_0^Q C_Q(Z/n, Z, w) dZ$$

fulfills condition (3.2).

Recalling that $CS(Q) = \int_0^Q P(Z) dZ - QP(Q)$ and $\sum_i \pi^i(\mathbf{q}, \mathbf{w}) = QP(Q) - \sum_i C(q_i, Q, \mathbf{w})$, H_2 can be re-written as:

$$H_{2}(\mathbf{q}, w) = QP(Q) - \sum_{i} C(q_{i}, Q, w) + \frac{n-1}{n} \left[\int_{0}^{Q} P(Z) dZ - QP(Q) \right] + (n-1) \int_{0}^{Q} C_{Q}\left(\frac{Z}{n}, Z, w\right) dZ$$

which leads to

$$H_{2}(\mathbf{q},w) = \frac{1}{n}QP(Q) - \sum_{i}C(q_{i},Q,w) + \frac{n-1}{n} \left[\int_{0}^{Q}P(Z)dZ\right] + (n-1)\int_{0}^{Q}C_{Q}\left(\frac{Z}{n},Z,w\right)dZ$$

Differentiating with respect to q_j and evaluating at the Nash equilibrium, which is symmetric, we get:

$$\frac{\partial H_2(\mathbf{q}^*, w)}{\partial q_j} = \left\{ \frac{1}{n} \left[QP'(Q) + P(Q) \right] - \sum_i C_q \left(\frac{Q}{n}, Q, w \right) \frac{1}{n} - \sum_i C_Q \left(\frac{Q}{n}, Q, w \right) \right\} \\ + \frac{n-1}{n} P(Q) + (n-1)C_Q \left(\frac{Q}{n}, Q, w \right) \right\} \frac{\partial Q}{\partial q_j}$$

From where it is direct that:

$$\frac{\partial H_2(\mathbf{q}^*, w)}{\partial q_j} = P(Q) + (Q/n)P'(Q) - C_q(Q/n, Q, w) - C_Q(Q/n, Q, w) = \frac{\partial \pi^j(\mathbf{q}^*, w)}{\partial q_j} \qquad \blacksquare$$

A.3) Differentiated demands game Proof that: $p^{j}(\mathbf{q}) = (\partial h(\mathbf{q}) / \partial q_{j})(1/\alpha_{j})$ leads to a potential function given by $F(\mathbf{q}, \mathbf{w}) = \sum_{k=1}^{2^{n}} \left[H_{k}^{\dagger}(\prod_{i \in \Gamma_{k}} q_{i}^{\alpha_{i}}) \prod_{i \in \Gamma_{k}} q_{i}^{\alpha_{i}} \right] - \sum_{j} C^{j}(q_{j}, \mathbf{w}).$

To prove this, I will use Lemma 3.1; I will show that firm *i*'s profit can be written as the sum of two functions: the potential function and a function that depends only on \mathbf{q}_{-i}

(statement 3). We have $h(\mathbf{q}) = \sum_{k=1}^{2^n} H_k (\prod_{i \in \Gamma_k} q_i^{\alpha_i})$. Let $K_j = \{k/j \in \Gamma_k\}$; then

$$\frac{\partial h(\mathbf{q})}{\partial q_j} = \sum_{k \in K_j} \left[H_k^{\dagger} (\prod_{i \in \Gamma_k} q_i^{\alpha_i}) \frac{\alpha_j}{q_j} \prod_{i \in \Gamma_k} q_i^{\alpha_i} \right]$$

Firm *j*'s revenues are then:

$$R^{j}(\mathbf{q}) = p^{j}(\mathbf{q})q_{j} = \frac{\partial h(\mathbf{q})}{\partial q_{j}} \frac{q_{j}}{\alpha_{j}} = \sum_{k \in K_{j}} \left[H_{k}^{'}(\prod_{i \in \Gamma_{k}} q_{i}^{\alpha_{i}}) \prod_{i \in \Gamma_{k}} q_{i}^{\alpha_{i}} \right]$$

Denoting by K_j^c the complement set of K_j , we can re-write this as

$$R^{j}(\mathbf{q}) = \underbrace{\sum_{k=1}^{2^{n}} \left[H_{k}^{'}(\prod_{i \in \Gamma_{k}} q_{i}^{\alpha_{i}}) \prod_{i \in \Gamma_{k}} q_{i}^{\alpha_{i}} \right]}_{R(\mathbf{q})} - \underbrace{\sum_{k \in K_{j}^{C}} \left[H_{k}^{'}(\prod_{i \in \Gamma_{k}} q_{i}^{\alpha_{i}}) \prod_{i \in \Gamma_{k}} q_{i}^{\alpha_{i}} \right]}_{-r^{j}(\mathbf{q}_{-j})}$$

which shows that firm *i* profits can be written as $\pi^i(\mathbf{q}, \mathbf{w}) = R(\mathbf{q}) + r^i(\mathbf{q}_{-i}) - C^i(q_i, \mathbf{w})$, or

$$\pi^{i}(\mathbf{q}) = \underbrace{R(\mathbf{q}) - \sum_{j} C^{j}(q_{j}, \mathbf{w})}_{F(\mathbf{q}, \mathbf{w})} + \underbrace{r^{i}(\mathbf{q}_{-i}) + \sum_{j \neq i} C^{j}(q_{j}, \mathbf{w})}_{\Theta^{i}(\mathbf{q}_{-i}, \mathbf{w})}$$

Therefore, by Lemma 3.1, the potential function exists and is given by $F(\mathbf{q}, \mathbf{w}) = R(\mathbf{q}) - \sum_{j} C^{j}(q_{j}, \mathbf{w})$

A.4) Differentiated demands game Proof of equation (3.6), that is:

$$F_{1}(\mathbf{q},\mathbf{w}) = \sum_{j=1}^{n} \left[(a - bq_{j})q_{j} - C^{j}(q_{j},\mathbf{w}) \right] - e \sum_{k=1}^{n-1} \sum_{j=k+1}^{n} q_{k}q_{j} \equiv \sum_{j=1}^{n} \pi^{j}(\mathbf{q},\mathbf{w}) + e \sum_{k=2}^{n} \sum_{j=1}^{k-1} q_{k}q_{j}$$

We have:

$$F_{1}(\mathbf{q}, \mathbf{w}) = \sum_{j=1}^{n} \left[(a - bq_{j})q_{j} - C^{j}(q_{j}, \mathbf{w}) \right] - e \sum_{k=1}^{n-1} \sum_{j=k+1}^{n} q_{k}q_{j}$$

$$\Rightarrow \quad F_{1}(\mathbf{q}, \mathbf{w}) = \sum_{j=1}^{n} \left[\left(a - bq_{j} - \sum_{k \neq j} eq_{k} + \sum_{k \neq j} eq_{k} \right) q_{j} - C^{j}(q_{j}, \mathbf{w}) \right] - e \sum_{k=1}^{n-1} \sum_{j=k+1}^{n} q_{k}q_{j}$$

$$\Rightarrow \quad F_{1}(\mathbf{q}, \mathbf{w}) = \sum_{j=1}^{n} \left[p^{j}(\mathbf{q})q_{j} - C^{j}(q_{j}, \mathbf{w}) \right] + \sum_{j=1}^{n} \sum_{k \neq j} eq_{k}q_{j} - e \sum_{k=1}^{n-1} \sum_{j=k+1}^{n} q_{k}q_{j}$$

And, therefore:

$$F_{1}(\mathbf{q}, \mathbf{w}) = \sum_{j=1}^{n} \pi^{j}(\mathbf{q}, \mathbf{w}) + e \sum_{k=2}^{n} \sum_{j=1}^{k-1} q_{k} q_{j}$$

A.5) Differentiated demands game: calculation of consumer surplus

Proof that
$$CS = \int_{\mathbf{p}}^{a} \sum_{i} q_{i}(\mathbf{p}) dp_{i} = (b/2) \sum_{k} q_{k}^{2} + e \sum_{k=2}^{n} \sum_{j=1}^{k-1} q_{k} q_{j}$$

First, inverting the system of indirect demands lead to $q^{j}(\mathbf{p}) = A - Bp_{j} + \sum_{i \neq j}^{n} Ep_{i}$, where

$$a = \frac{A}{B - (n-1)E}, \ b = \frac{B - (n-2)E}{(B - (n-1)E)(B + E)} \text{ and } e = \frac{E}{(B - (n-1)E)(B + E)}$$

Since $\frac{\partial q_i}{\partial p_j} = \frac{\partial q_j}{\partial p_i}$, the solution of the line integral is path-independent. We take a linear integration path as follows: $\Psi_i(\sigma) = p_i + \sigma(a - p_i)$, $\sigma \in [0,1]$, $\forall i \in [1..N]$. Thus $\Psi_i(\sigma = 0) = p_i$ and $\Psi_i(\sigma = 1) = a$. Further, $d\Psi_i(\sigma) = (a - p_i)d\sigma$. Changing variables, we can re-write

$$CS = \int_{\mathbf{p}}^{d} \sum_{i} q_{i}(\mathbf{p}) dp_{i} = \int_{0}^{d} \sum_{i} q_{i}(\Psi(\sigma)) \cdot (a - p_{i}) d\sigma$$

from where we get $CS = \sum_{i=1}^{n} (a - p_i) \int_{0}^{1} q_i (\Psi(\sigma)) d\sigma = \sum_{i=1}^{n} \left(bq_i + \sum_{j \neq i} eq_j \right) \int_{0}^{1} q_i (\Psi(\sigma)) d\sigma$.

Next, we calculate *I*. Replacing $q_i(\Psi)$ we get $I = \int_{0}^{1} \left(a - b\Psi_i(\sigma) + \sum_{j \neq i} e\Psi_j(\sigma)\right)$. Replacing Ψ_i and Ψ_i and reordering we obtain

$$I = \int_{0}^{1} q_i d\sigma + \int_{0}^{1} \left(-aB + Bp_i + a\sum_{j \neq i} E - \sum_{j \neq i} Ep_j \right) \sigma d\sigma$$
$$I = q_i + \int_{0}^{1} \left(-q_i + A - aB + a(n-1)E \right) \sigma d\sigma$$
$$I = \frac{q_i}{2} + \frac{A - a(B - (n-1)E)}{2}$$

The second term on the right hand side of I is zero because a(B - (n-1)E) = A. Therefore, replacing I in CS we obtain

$$CS = \sum_{i=1}^{n} \left(bq_i + \sum_{j \neq i} eq_j \right) \frac{q_i}{2} = \frac{b}{2} \sum_i q_i^2 + \frac{e}{2} \sum_i \sum_{j \neq i} q_i q_j = \frac{b}{2} \sum_k q_k^2 + e \sum_{k=2}^{n} \sum_{j=1}^{k-1} q_k q_j \qquad \blacksquare$$

APPENDIX B

Appendix for Chapter 4: Airport Ownership, effects on pricing and capacity

B.1) Existence, unicity and stability of Cournot-Nash equilibria in the airline market

First-order and second-order derivatives of airline *i*'s profit function (equation 4.7) are:

$$\phi_{i}^{i} = \left(AS - 2BS^{2}Q_{i} - ES^{2}\sum_{j \neq i}Q_{j} - c - \sum_{h=1,2}P_{h}\right) - S(g(Q_{i}) + Q_{i}g'(Q_{i})) - (\alpha S + \beta)\sum_{h=1,2}(D^{h} + Q_{i}D_{Q}^{h})$$
(b.1)

$$\phi_{ij}^{i} = -ES^{2} - (\alpha S + \beta) \sum_{h=1,2} \left(D_{Q}^{h} + Q_{i} D_{QQ}^{h} \right)$$
(b.2)

$$\phi_{ii}^{i} = -2BS^{2} - S(2g'(Q_{i}) + Q_{i}g''(Q_{i})) - (\alpha S + \beta) \sum_{h=1,2} (2D_{Q}^{h} + Q_{i}D_{QQ}^{h})$$
(b.3)

where subscripts in ϕ denotes partial derivatives and D_0^h denotes the derivative of the delay function with respect to Q, evaluated at Q and K_h . From (b.2) it can be seen that $\phi_{ij}^i < 0$, because D_Q and D_{QQ} are positive, implying that the game is not supermodular (hence disabling this approach to existence, uniqueness and stability). $\phi_i^i = 0$, $\forall i$, are the necessary conditions for Nash equilibria. As for the sufficient conditions, both the first and third terms on the right hand side of ϕ_{ii}^{i} are negative; the sign of the second term is not obvious though, because while $g'(\cdot) < 0$, the sign of $g''(\cdot)$ is unclear. Under assumptions (a) and (b) regarding schedule delay, is easy to verify that g">0 but, further $2g'(Q_i) + Q_ig''(Q_i) = 0$. Thus ϕ_{ii}^i is negative and the existence of Nash equilibria is guaranteed (as long as the solution is interior which is assumed for now). To prove uniqueness, first note that best reply correspondences, $\Psi_i(\mathbf{Q}_{-i})$, defined by $\phi_i^i(\Psi_i(\mathbf{Q}_{-i}), \mathbf{Q}_{-i}) = 0$, are actually continuously differentiable functions of the sum of quantities of other firms, that is, $\Psi_i(\mathbf{Q}_{-i}) = \Psi_i(\sum_{i \neq i} Q_i)$. Next, the slope of each best reply function is given by the ratio between $-\partial \phi_i^i / \partial \sum_{i \neq i} Q_j$ and ϕ_{ii}^i , but it is easy to check that $-\partial \phi_i^i / \partial \sum_{i \neq i} Q_j \equiv \phi_{ij}^i$. From (9) and (10) then, it follows that the slope of each best-reply function is greater than -1, implying there is a unique Cournot-Nash equilibrium, which is symmetric.¹

¹ See theorem 2.8 in Vives, (1999). Additionally, for a different approach to existence it would have been enough to note that, since all best reply functions are continuous and strictly decreasing, the best reply mapping $\Psi = (\Psi_1, ..., \Psi_N)$ has at least one fixed point; see theorem 2.7 in Vives (1999).

As for Cournot (or *tatonnement*) stability, a sufficient condition is that the best reply mapping is a contraction: $\phi_{ii}^i + \sum_{i \neq i} |\phi_{ij}^i| < 0$. In this case this is:

$$(\alpha S + \beta) \sum_{h=1,2} \left((N-3) D_Q^h + (N-2) N^{-1} Q D_{QQ}^h \right) - \left(2B - (N-1)E \right) S^2 < 0$$

where symmetry was imposed $(Q_i = Q_j = Q/N \quad \forall i, j)$. Evidently, this condition holds for N very small.

Reference: Vives, X. (1999) Oligopoly Pricing: old ideas and new tools, MIT Press.

B.2) Free entry long run equilibrium in the airline market

The free entry long run equilibrium is obtained when $\phi^i = 0 \quad \forall i \text{ or, equivalently, when the revenue per flight, } S \cdot t^i(Q_i, \mathbf{Q}_{-i})$, equals average cost. Using equation (4.7), with free entry

$$AS - (B + (N-1)E)S^{2}\frac{Q}{N} - Sg(\frac{Q}{N}) - c - \sum_{h} P_{h} - (\alpha S + \beta)\sum_{h} D(Q, K_{h}) = 0$$
 (b.4)

Equations (4.8) and (b.4) together determine the free entry equilibrium $Q(P_h, K_h)$ and $\overline{N}(P_h, K_h)$. To see this equilibrium graphically, first note that under (a) and (b), the marginal revenue of each firm, $MR_i(Q_i, \mathbf{Q}_{-i})$, is decreasing in Q_i (recall that $\phi_{ii}^i < 0$ and that airline's cost are convex). Further, it intersects the inverse demand function for flights, $S \cdot t^i(Q_i, \mathbf{Q}_{-i})$, which is first increasing and then decreasing (see discussion in Section 4.2.1), at its maximum.² Next, both marginal and average cost functions are convex and increasing, the former being larger than the latter. Therefore, the free entry equilibrium is as in Figure b.1.

² Proof: Revenues are $S \cdot t^{i}Q_{i}$, therefore $MR_{i}(Q_{i}, \mathbf{Q}_{-i}) = S(Q_{i} \cdot t_{i}^{i} + t_{i})$. Imposing $MR_{i}=S t^{i}$, we obtain that $S \cdot Q_{i} \cdot t_{i}^{i} = 0$. Since we are ruling out $Q_{i} = 0$, marginal revenue and inverse demand intersect when $t_{i}^{i} = 0$.



Figure b.1: Free-entry long-run equilibrium

 \overline{N} is given by Q/\widetilde{Q}_i , where \widetilde{Q}_i is determined by the profit maximization first-order condition marginal revenue equals marginal cost and the zero profit condition average cost equals revenue per flight. At this point average cost and $S \cdot t^i$ are tangent.³

B.3) Derivation of equation (4.11):

$$PS(P_{h}, K_{h}, N) = \int_{\Theta(P, K, N)}^{A} \sum_{i}^{N} q_{i}(\Theta) d\Theta_{i} = \frac{(B + (N - 1)E)S^{2}Q(P_{h}, K_{h}, N)^{2}}{2N}.$$

Since $\partial q_i / \partial \theta_j = \partial q_j / \partial \theta_i$ the line integral has a solution that is path-independent. We take a linear integration path as follows:

$$\Theta_i(\sigma) = \theta_i(P_h, K_h, N) + \sigma(A - \theta_i(P_h, K_h, N)) \quad , \quad \sigma \in [0,1] \quad , \quad \forall i \in [1..N]$$
(b.5)

Thus $\Theta_i(\sigma = 0) = \theta_i(P_h, K_h, N)$ and $\Theta_i(\sigma = 1) = A$. Further, $d\Theta_i(\sigma) = (A - \theta_i(P_h, K_h, N))d\sigma$. Changing variables, we can rewrite

$$PS(P_h, K_h, N) = \int_{\Theta(P, K, N)}^{A} \sum_{i}^{N} q_i(\Theta) d\theta_i = \int_{0}^{1} \sum_{i}^{N} q_i(\Theta(\sigma)) \cdot (A - \theta_i(P_h, K_h, N)) d\sigma$$

³ Proof: The first-order condition is $S(Q_i \cdot t_i^i + t_i) = C_i^i$, while the zero profit condition is $S \cdot t_i = C^i / Q_i$. Together the imply that $S \cdot Q_i \cdot t_i^i + C^i / Q_i = C_i^i$ and therefore $S \cdot t_i^i = (Q_i \cdot t_i^i - C^i) / Q_i^2$. Thus, at \tilde{Q}_i , $AvC^i = S \cdot t^i$ and they have the same slope: they are tangent. which is equivalent to

$$PS(P_h, K_h, N) = \sum_{i}^{N} \underbrace{\left(A - \theta_i(P_h, K_h, N)\right)}_{i} \underbrace{\int_{0}^{1} q_i(\Theta(\sigma)) d\sigma}_{ii}$$
(b.6)

Let us first restate *I* differently. Recalling that $\theta_i(P_h, K_h, N) = A - Bq_i - \sum_{j \neq i} Eq_j$, that $q_i = SQ_i$ and that in equilibrium $Q_i = Q/N$ $\forall i$, we get $I = \frac{SQ(P_h, K_h, N)}{N} (B + (N - 1)E)$. Next, we calculate *II* in (b.6). Replacing $q_i(\Theta)$ we get $II = \int_{0}^{1} \left(a - b\Theta_i(\sigma) + \sum_{j \neq i} e\Theta_j(\sigma)\right)$. Next, replacing Θ_i and Θ_j with (b.5) and reordering we obtain

$$II = \int_{0}^{1} q_{i}(P_{h}, K_{h}, N) d\sigma + \int_{0}^{1} \left(-Ab + b\theta_{i} + A\sum_{j \neq i} e - \sum_{j \neq i} e\theta_{j} \right) \sigma d\sigma$$

$$II = q_{i}(P_{h}, K_{h}, N) + \int_{0}^{1} \left(-q_{i}(P_{h}, K_{h}, N) + a - Ab + A(N - 1)e \right) \sigma d\sigma$$

$$II = \frac{SQ(P_{h}, K_{h}, N)}{2N} + \frac{a - A(b - (N - 1)e)}{2}$$

In the first term on the right hand side we imposed symmetry; the second term is zero because $A = \frac{a}{b - (N-1)e}$. Therefore, replacing *I* and *II* in (b.6), we obtain

$$PS(P_h, K_h, N) = \sum_{i}^{N} \left[\frac{SQ(P_h, K_h, N)}{N} \left(B + (N-1)E \right) \cdot \frac{SQ(P_h, K_h, N)}{2N} \right]$$

and, obviously, nothing inside the brackets depend on *i* anymore so we obtain

$$CS(P_{h}, K_{h}, N) = \frac{(B + (N - 1)E)S^{2}Q(P_{h}, K_{h}, N)^{2}}{2N}$$

B.4) Derivation of Equation (4.17):

$$\frac{d\Phi}{d\rho} = -Q(P, K_h, N) - \frac{(N-1)ES^2Q}{N} \frac{\partial Q}{\partial P} - (\alpha S + \beta) \frac{Q}{N} \sum_h D_Q^h \frac{\partial Q}{\partial \rho}.$$

Equation (4.16) shows that total profits in the airline market may be written as

$$\Phi(Q,\rho) = QS\left[A - \frac{QS}{N}(B + (N-1)E)\right] - Q[c+\rho]$$

Where ρ is given by (4.13). Straightforward calculation of $\frac{d\Phi}{d\rho} = \frac{\partial\Phi}{\partial Q} \cdot \frac{\partial Q}{\partial \rho} + \frac{\partial\Phi}{\partial \rho}$, leads to

$$\frac{d\Phi}{d\rho} = \left[AS - c - \rho - 2\frac{QS^2}{N}(B + (N-1)E)\right]\frac{\partial Q}{\partial \rho} - Q$$

On the other hand, equation (4.14) is:

$$QS^{2}\left(\frac{2B}{N}+\frac{(N-1)}{N}E\right)+\rho+c-AS+(\alpha S+\beta)\frac{Q}{N}\sum_{h}D_{Q}^{h}=0$$

which leads to: $AS - c - \rho = (\alpha S + \beta) \frac{Q}{N} \sum D_Q^h + \frac{QS^2}{N} (2B + (N-1)E)$. Replacing this in $\frac{d\Phi}{d\rho}$, gives us the desired result.

B.5) *Proof that at the optimum,* $K_1 = K_2 = K$

We prove this for the SPA case. The proofs for the other cases are analogous.

 $\frac{\partial P}{\partial K_1} = -(\alpha S + \beta) \left(\frac{Q}{N} D_{QK_1}^1 + D_{K_1}^1 \right) \text{ does not depend on } K_2, \text{ therefore, } Q \cdot \partial P / \partial K_1 \text{ is a}$ function of Q and K_1 only. Also, $\partial P / \partial K_1$ is decreasing, goes to infinity when $K_1 \rightarrow Q$ and to zero when $K_1 \rightarrow \infty$. Hence, there is only one $K_1(Q,r,N) > Q$ that satisfies (4.21). By symmetry of P with respect to K_1 and K_2 , the same goes for $K_2(Q,r,N)$. Replacing these in (4.20) one obtains the optimal Q and then optimal $K_1 = K_2 = K$.

B.6) Showing that dQ/dN and dK/dN cannot be signed a priori

We show this for the SPA case. The other cases are analogous. Differentiating both (4.20) and (4.21) with respect to N and solving for dQ^{SPA}/dN and dK^{SPA}/dN :

$$\frac{dQ^{SPA}}{dN} = \frac{\pi_{QN}\pi_{KK} - \pi_{QK}\pi_{KN}}{\pi_{QQ}\pi_{KK} - \pi_{QK}} \qquad \qquad \frac{dK^{SPA}}{dN} = \frac{\pi_{KN}\pi_{QQ} - \pi_{QK}\pi_{QN}}{\pi_{QQ}\pi_{KK} - \pi_{QK}}$$

where, for second-order conditions to hold, π_{QQ} , π_{KK} and the denominators must be negative; also $\pi_{QK} = P_{QK}Q + P_K > 0$, $\pi_{QN} = P_{QN}Q + P_N > 0$ and $\pi_{KN} = P_{KN}Q < 0$ -see equations in (4.10). Therefore, the signs cannot be determined a priori and, as a consequence, it cannot be known now how P^{SPA} change with N

B.7) *Proof of Proposition* 4.1

From (4.22), we can write

$$SW(Q, K_h; N) = \pi + \frac{(B + (N - 1)E)S^2Q^2}{2N} + QS \left[A - \frac{QS}{N} (B + (N - 1)E) - g\left(\frac{Q}{N}\right) - \alpha \sum D^h \right] - Q \left[c + P + \beta \sum D^h\right]$$

We differentiate this with respect to Q and evaluate the resulting expression at $Q^{SPA}(K)$, the optimal private quantity for the given K, which makes the term $\partial \pi / \partial Q$ nil. Using equation

(4.8) to replace
$$AS - c - S\left(g'(\frac{Q}{N})\frac{Q}{N} + g(\frac{Q}{N})\right)$$
, we obtain:

$$\frac{\partial SW}{\partial Q}\Big|_{Q^{SPA}(K)} = -P_Q Q + \frac{BS^2 Q}{N} - \frac{(N-1)}{N}(\alpha S + \beta)Q\sum_h D_Q^h\Big|_{Q^{SPA}(K)}$$

Replacing $P_Q = (\alpha S + \beta) \sum_h (\frac{N-1}{N} D_Q^h + \frac{Q}{N} D_{QQ}^h) + \frac{S^2 (2B + (N-1)E)}{N}$, we finally get

$$\frac{\partial SW}{\partial Q}\Big|_{Q^{SPA}(K)} = \frac{Q}{N}(\alpha S + \beta)\sum_{h}(2D_{Q}^{h} + QD_{QQ}^{h}) + \frac{S^{2}Q(3B + (N-1)E)}{N}\Big|_{Q^{SPA}(K)} > 0$$

which shows that the SPA induces fewer flights. The equivalence follows from the decreasing monotonicity of P with respect to Q.

B.8) Proof of Proposition 4.2

Expression (4.22) for SW leads to $\partial SW / \partial K_1 |_{K^{SPA}(Q)} = -Q(\alpha S + \beta)D_{K_1} - QP_{K_1}$. Replacing

$$\frac{\partial P}{\partial K_1} = -(\alpha S + \beta) \left(\frac{Q}{N} D_{QK_1}^1 + D_{K_1}^1 \right), \text{ it is finally obtained that}$$

$$\partial SW / \partial K_1 |_{K^{SPA}(Q)} = Q^2 (\alpha S + \beta) D_{QK_1} / N < 0.$$

B.9) Proof of Proposition 4.5

Let $Q^{W}(K)$ and $Q^{JP}(K)$ be the W and JP quantity rules respectively. Consider $(\pi + \Phi)(Q^{JP}(K)) \equiv (\pi + \Phi)(K)$ and differentiate this with respect to K (could be either K_1 or K_2). We get $\frac{d(\pi + \Phi)(K)}{dK} = \frac{\partial(\pi + \Phi)Q^{JP}(K)}{\partial Q}\frac{\partial Q^{JP}(K)}{\partial K} + \frac{\partial(\pi + \Phi)}{\partial K} = \frac{\partial(\pi + \Phi)}{\partial K}$ by using the first-order conditions. Using (4.27) we obtain $\frac{d(\pi + \Phi)(K)}{dK} = -Q^{JP}(K)(\alpha S + \beta)D_Q(Q^{JP}(K), K) - r$. However, we

also know that $r = -Q^{W}(K)(\alpha S + \beta)D_{Q}(Q^{W}(K), K)$ from (4.24) so we get

$$\frac{d(\pi + \Phi)(K)}{dK}\Big|_{Q^{TPT}(K^{W}), K^{W}} = (\alpha S + \beta) \Big[Q^{W}(K^{W}) D_{Q} \Big(Q^{W}(K^{W}), K^{W} \Big) - Q^{JP}(K^{W}) D_{Q} \Big(Q^{JP}(K^{W}), K^{W} \Big) \Big]$$
(b.7)

Since $Q^{IP}(K^W) < Q^W(K^W)$ by Proposition 4.3 and $D_Q < 0$ and is decreasing in Q, (b.7) is negative. Therefore $K^W > K^{JP}$ and thus $Q^{JP}(K^{JP}) < Q^{JP}(K^W) < Q^W(Q^W)$ which implies that $Q^{JP} < Q^W$, $K^{JP} < K^W$.

B.10) Proof of Proposition 4.6

Consider the following second best social welfare function: $\tilde{S}W(K) \equiv SW(Q^{JP}(K))$ differentiate it and evaluate it at K^{JP} . We get:

$$\frac{d\widetilde{S}W}{dK}\Big|_{K^{JP}} = \frac{\partial SW}{\partial Q}\Big|_{Q^{JP}(K^{JP}),K^{JP}} \cdot \frac{\partial Q^{JP}(K)}{\partial K}\Big|_{K^{JP}} + \frac{\partial SW}{\partial K}\Big|_{Q^{JP}(K^{JP}),K^{JP}}$$
(b.8)

We are interested in the sign of (b.8). If it is positive, then second best SW capacities are larger than the JP ones. The third term in the right hand side is zero because

 $\frac{\partial SW}{\partial K}\Big|_{Q^{I^{p}}(K^{I^{p}}),K^{I^{p}}} = \frac{\partial(\pi + \Phi)}{\partial K}\Big|_{Q^{I^{p}}(K^{I^{p}}),K^{I^{p}}} = 0 \text{ (the first equality is because they have the same capacity rule, the second from first-order condition). The first term is positive by Proposition 4.3. The second term is also positive because <math>\partial Q^{I^{p}}(K)/\partial K = -(\pi + \Phi)_{QK}/(\pi + \Phi)_{QQ}$, but $(\pi + \Phi)_{QK} > 0$ and $(\pi + \Phi)_{QQ} < 0$. Therefore, (A.4) is positive and JP capacities are below 2nd best social welfare capacities.

B.11) Analysis of the closed-loop game when airports are private and independent

Airports first choose capacities (simultaneously) and then prices. Over or underinvestment in capacity will be par rapport to the open-loop. From the profit functions in (4.34), and

noting that $\frac{\partial Q^h(P_1, P_2)}{\partial P_1 \partial P_2} = \frac{\partial^2 Q(P_1 + P_2)}{\partial P^2}$, is then easy to obtain that

$$\frac{\partial^2 \pi^h}{\partial P_h \partial P_k} = (P_h - C') \frac{\partial^2 Q}{\partial P^2} + \frac{\partial Q}{\partial P} - \left(\frac{\partial^2 Q}{\partial P^2}\right)^2 C''$$
(b.9)

$$\frac{\partial^2 \pi^h}{\partial P_h^2} = (P_h - C') \frac{\partial^2 Q}{\partial P^2} + 2 \frac{\partial Q}{\partial P} - \left(\frac{\partial^2 Q}{\partial P^2}\right)^2 C''$$
(b.10)

(b.10) being negative is a necessary condition for existence in the open-loop case. If this is true, then (b.9) is negative as well, but also $\partial^2 \pi^h / \partial P_h^2 + \left| \partial^2 \pi^h / \partial P_h \partial P_k \right| < 0$. Hence, the best reply mapping is a contraction and therefore there is a unique, symmetric and stable Nash equilibrium in the second stage, which is denoted by $\hat{P}_h(K_1, K_2)$. To know whether capacities are going to be smaller or larger than in the open-loop game, we look at the first stage: $\frac{d\pi^h}{dK_h} = \frac{\partial \pi^h}{\partial K_h} + \frac{\partial \pi^h}{\partial P_h} \frac{\partial \hat{P}_h}{\partial K_h} + \frac{\partial \pi^h}{\partial P_k} \frac{\partial \hat{P}_k}{\partial K_h}$. Evaluating this at the open-loop capacity makes the first term on the right hand side vanish. The second term is zero by the envelope theorem. Thus

$$\frac{d\pi^{h}}{dK_{h}}\Big|_{K_{h}^{0-L}} = \frac{\partial\pi^{h}}{\partial P_{k}}\frac{\partial\hat{P}_{k}}{\partial K_{h}} = \frac{\partial\pi^{k}}{\partial P_{h}}\frac{\partial\hat{P}_{h}}{\partial K_{h}}\frac{\partial\hat{P}_{h}}{\partial\hat{P}_{h}}$$
(b.11)

where the symmetry of the problem was used. It is easy to check that the first derivative on the right hand side is negative; the second is positive: $\partial \hat{P}_h / \partial K_h = -(\partial^2 \pi^h / \partial P_h \partial K_h)(\partial^2 \pi^h / \partial P_h^2)^{-1}$. Investment makes an airport tough then, in that $\frac{\partial \pi^k}{\partial P_h} \frac{\partial \hat{P}_h}{\partial K_h} < 0$. The third derivative is negative because prices are strategic substitutes (b.9 is negative). Hence (b.11) is positive, which shows that closed-loop capacities are

(6.9 is negative). Hence (6.11) is positive, which shows that closed-loop capacities are larger than open-loop capacities: airports over invest in capacity following *top-dog* strategies. This directly leads to higher prices $(\partial \hat{P}_h / \partial K_h < 0)$ but the effect on traffic cannot be signed.

B.12) Derivation of equations (4.37) and (4.38)

If airports individually use two part tariffs, the program they face in the open loop is

$$\max_{P_h, K_h, T_h} Q_h(P_1 + P_2, K_1, K_2) P_h - C(Q_h) - K_h r + T_h N \quad \text{st} \quad T_1 + T_2 \le \phi^1(P_1 + P_2, K_1, K_2)$$
(b.12)

where $\phi^1 = \phi^i$, i=1,...,N, is the profit of airline 1 given P_1 , P_2 and K_1 , K_2 (i.e. downstream equilibrium profit). It is easy to check that in equilibrium the constraints must be binding; if not, airports have incentives to increase their fees. Noting that $\partial \phi^1 / \partial P_h = \partial \phi^1 / \partial P$ and recalling that $N\phi^1(P_1 + P_2, K_1, K_2) = \Phi$, we get the following first-order conditions

$$\frac{\partial Q}{\partial P}P_h + Q - C'\frac{\partial Q}{\partial P} + \frac{\partial \Phi}{\partial P} = 0$$
 (b.13)

$$\frac{\partial Q}{\partial K_{h}}P_{h} - C'\frac{\partial Q}{\partial K_{h}} + \frac{\partial \Phi}{\partial K_{h}} = 0$$
 (b.14)

$$T_1 + T_2 = \phi^1(P_1 + P_2, K_1, K_2)$$
 (b.15)

It is easy to see that (b.13) and (b.14) are identical to the first order-conditions of the setting in which each airport, in the open-loop, maximizes own profit plus airlines profits (the collaboration idea). Hence, pricing and capacity rules will be identical (as in the SPA case). Next

$$\frac{\partial \Phi}{\partial P} = \left[AS - P - c - S \left(\frac{Q}{N} g' \left(\frac{Q}{N} \right) + g \left(\frac{Q}{N} \right) \right) \right] \frac{\partial Q}{\partial P} + \left[-(\alpha S + \beta) \sum D_{Q}^{h} - 2 \frac{QS^{2}}{N} (B + (N - 1)E) \right] \frac{\partial Q}{\partial P}$$

From $\Omega(Q, P, K_h, N) = 0$ in (4.8), it can be seen that the first term in brackets on the right hand side is equal to $(\alpha S + \beta) \sum (QD_Q^h / N + D^h) + QS^2 (2B + (N-1)E) / N$. Replacing this, simplifying and then replacing the resulting $\partial \Phi / \partial P$ back into (b.13), we get

$$P_{h} = 2C' + ((N-1)/N)(\alpha S + \beta)Q\sum_{h} D_{Q}^{h} + ((N-1)/N)ES^{2}Q$$

Imposing symmetry and adding P_1 and P_2 , (4.37) is obtained. For capacities, we have

$$\frac{\partial \Phi}{\partial K_{h}} = \left[AS - P - c - S\left(\frac{Q}{N}g'\left(\frac{Q}{N}\right) + g\left(\frac{Q}{N}\right)\right) - (\alpha S + \beta)\sum(QD_{Q}^{h} + D^{h}) - 2\frac{QS^{2}}{N}(B + (N - 1)E)\right] \frac{\partial Q}{\partial K_{h}} - (\alpha S + \beta)QD_{K}^{h}$$

replacing this in (b.14) and using the first-order condition on P get us equation (4.38).

ι.

APPENDIX C

Appendix for Chapter 5: Sequential Peak-Load Pricing in a Vertical Setting, the Case of Airports and Airlines

C.1) Conditions for an interior allocation of consumers

(i) If $(P_p - P_o)/S < \overline{\theta}(B_p - B_o)$, then the peak period is used, that is $\theta^* < \overline{\theta}$. (ii) If $\underline{\theta}B_o < (c + P_o)/S$, then some consumers will not fly, that is $\theta^f > \underline{\theta}$. (iii) If $\overline{\theta}$ is large enough, then the off-peak period is used, that is $\theta^* > \theta^f$.

Proof:

First, equivalent conditions for interior allocations, but in terms of Q_p and Q_o are: The peak is used: $\theta^* < \overline{\theta} \Leftrightarrow (\overline{\theta} - \theta^*)/S > 0 \Leftrightarrow Q_o > 0$ Some consumers do not fly: $\theta^f > \underline{\theta} \Leftrightarrow (\overline{\theta} - \theta^f)/S < (\overline{\theta} - \underline{\theta})/S \Leftrightarrow Q_o + Q_p < (\overline{\theta} - \underline{\theta})/S$ The off-peak is used: $\theta^* > \theta^f \Leftrightarrow (\theta^* - \theta^f)/S > 0 \Leftrightarrow Q_o > 0$

With this, the proofs of each part are:

- (i) Note that $(-\Omega^{p} + \Omega^{o})$ in (5.12) is strictly increasing in Q_{p} , and $(-\Omega^{p} + \Omega^{o})_{Q_{p} \to \infty} > 0$. Also, $(-\Omega^{p} + \Omega^{o})_{Q_{p}=o} = (P_{p} - P_{o}) - \overline{\theta}S(B_{p} - B_{o})$. Hence, if $P_{p} - P_{o} < \overline{\theta}S(B_{p} - B_{o})$, then $(-\Omega^{p} + \Omega^{o})_{Q_{p}=o} < 0$ and $Q_{p} > 0$.
- (ii) From $\Omega^{o}=0$ in (5.10) we get that $(Q_{o} + Q_{p})B_{o}S^{2}(N+1)/N = B_{o}\overline{\theta}S c P_{o}$, This imply that $Q_{o} + Q_{p} < (B_{o}\overline{\theta}S - c - P_{o})/(B_{o}S^{2})$. Hence, a sufficient condition for $Q_{o} + Q_{p} < \frac{\overline{\theta} - \theta}{S}$ is: $(B_{o}\overline{\theta}S - c - P_{o})/(B_{o}S^{2}) < \overline{\theta} - \theta/S$, which leads to $\theta B_{o} < (c + P_{o})/S$.

(iii) From $\Omega^o = 0$ we know that $Q_o + Q_p = \frac{(B_o \theta S - c - P_o)N}{B_o S^2 (N+1)}$. Hence, $Q_o > 0$ is equivalent to $Q_p < \frac{(B_o \overline{\theta} S - c - P_o)N}{B_o S^2 (N+1)} \equiv \widetilde{Q}_p$. In order to ensure $Q_p < \widetilde{Q}_p$, we need that $(-\Omega^p + \Omega^o)_{Q_p = \widetilde{Q}_p} > 0$ (see proof of part i). Straightforward algebra gives us

$$(-\Omega^{p} + \Omega^{o})(\tilde{Q}_{p}) = (\alpha S + \beta) \left[D(\tilde{Q}_{p}) + \frac{\tilde{Q}_{p}}{N} D'(\tilde{Q}_{p}) + \frac{(P_{p} - P_{o})}{(\alpha S + \beta)} - \frac{(B_{p} - B_{o})(c + P_{o})}{B_{o}(\alpha S + \beta)} \right],$$

so that a sufficient condition for $(-\Omega^{p} + \Omega^{o})_{\tilde{Q}_{p}} > 0$ is

$$D(\tilde{Q}_p) > \frac{(B_p - B_o)(c + P_o)}{B_o(\alpha S + \beta)} - \frac{(P_p - P_o)}{(\alpha S + \beta)}.$$
 And since $\partial \tilde{Q}_p / \partial \overline{\theta} > 0$, the condition is

always fulfilled for θ large enough.

Part (i) says that the peak period is used if the airport price differential between peak and off-peak is not too large. Specifically, the per-passenger airport price differential has to be smaller than the incremental benefit, for the highest consumer type, of changing from the off-peak to the peak. Clearly, when the airport does not practice PLP, the peak is always used. Part (ii) says that if $\underline{\theta}$ is low enough, then some consumers will not fly. In particular, the lowest consumer type must have a willingness to pay for off-peak travel that is smaller than the airlines' per-passenger marginal cost for an off-peak flight. Finally, part (iii) implies that Brueckner (2002, 2005)'s *single crossing property*, which imposes that $B_p(\theta) < B_o(\theta)$ for small θ values, is not needed to have a non-empty off-peak, and that a smaller airport price differential between peak and off-peak increases the likelihood of the off-peak been used. The lower bound for $\overline{\theta}$ cannot be made explicit because of the non-linearity of the delay function. For a linear delay function $D(Q_p, K) = \partial Q_p / K$, the lower

bound on
$$\overline{\theta}$$
 is given by $\widetilde{\theta} = \frac{SK(N+1)}{\delta B_o(\alpha S + \beta)N} ((B_p - B_o)(c + P_o) - B_o(P_p - P_o)) + c + P_o$,

while a lower bound not depending on N, would be $2\theta N/(N+1)$.

C.2) Proof of Proposition 5.2

(i) Differentiating (5.12) with respect to N we get: $\frac{\partial Q_p}{\partial N} = -\frac{\partial (-\Omega^p + \Omega^o)/\partial N}{\partial (-\Omega^p + \Omega^o)/\partial Q_p}.$ This

leads to

$$\frac{\partial Q_p}{\partial N} = \frac{\frac{Q_p S^2 (B_p - B_o)}{N^2} + \frac{(\alpha S + \beta)Q_p D_Q (Q_p)}{N^2}}{\frac{S^2 (N+1)(B_p - B_o)}{N} + (\alpha S + \beta) \left(\frac{N+1}{N} D'(Q_p) + \frac{Q_p}{N} D''(Q_p)\right)}$$

which can be written as

$$\frac{\partial Q_p}{\partial N} = \frac{Q_p}{N(N+1)} \left(\frac{S^2(B_p - B_o) + (\alpha S + \beta)D_Q(Q_p)}{S^2(B_p - B_o) + (\alpha S + \beta)D'(Q_p) + \frac{(\alpha S + \beta)}{N+1}Q_pD''(Q_p)} \right)$$

from where $0 < \partial Q_p / \partial N < Q_p / (N(N+1))$ follows.

(ii) From $\Omega^o = 0$ in (5.10) we know that $Q_o + Q_p = \frac{(B_o \theta S - c - P_o)N}{B_o S^2 (N+1)}$, from where it is

direct that
$$\frac{\partial (Q_o + Q_p)}{\partial N} = \frac{(B_o \overline{\partial} S - c - P_o)}{B_o S^2 (N+1)^2} = \frac{Q_o + Q_p}{N(N+1)} > 0$$

(iii) From parts (i) and (ii) we know that $\partial Q_o / \partial N > Q_o / (N(N+1))$. If the off-peak is used for all N, then $Q_o > 0$ and therefore $\partial Q_o / \partial N > 0$.

C.3) Proof of Proposition 5.5

To prove parts (i) and (ii), it is useful to first state the following Lemma:

Lemma C.1: If two prices P_1 and P_2 are given by the fixed points $P_1 = f(Q(P_1))$ and $P_2 = g(Q(P_2))$ respectively, where f is continuously differentiable in $(Q(P_1); Q(P_2)), Q$ is continuously differentiable in $(P_1; P_2), f(Q(P_2)) > g(Q(P_2))$, and either $Q(\cdot)$ is non-increasing and $f(\cdot)$ is non-decreasing, or $Q(\cdot)$ is non-decreasing and $f(\cdot)$ is non-increasing, then $P_1 > P_2$.

<u>*Proof*</u>: We prove this by contradiction. Suppose that $P_1 \leq P_2$. Denote $\tilde{P}_2 = f(Q(P_2))$. Applying the mean-value theorem to $P_1 = f(Q(P_1))$ and $\tilde{P}_2 = f(Q(P_2))$ yields:

$$P_1 - \widetilde{P}_2 = f'(\overline{Q})(Q(P_1) - Q(P_2))$$

where \overline{Q} is some point between $Q(P_1)$ and $Q(P_2)$. Further applying the mean-value theorem to $Q(P_1)$ and $Q(P_2)$, the above equation becomes:

$$P_1 - \widetilde{P}_2 = f'(\overline{Q})(Q(P_1) - Q(P_2)) = f'(\overline{Q})Q'(\overline{P})(P_1 - P_2) \ge 0$$

where the inequality arises because $f'(\overline{Q})Q'(\overline{P}) \leq 0$ and the assumption that $P_1 \leq P_2$. Thus, $P_1 \geq \widetilde{P}_2$. But since by assumption $f(Q(P_2)) > g(Q(P_2))$ or, equivalently, $\widetilde{P}_2 > P_2$, we obtain $P_1 > P_2$, thus resulting in a contradiction.

Now, we can prove Proposition 5.5.

(i) That $P_o^W < P_o^{\pi}$ follows from writing the pricing rules (5.34) and (5.29) as $P_o^{\pi} = f_o(Q_o + Q_p)$ and $P_o^W = g_o(Q_o + Q_p)$. Since total traffic $Q_0 + Q_p$ is, by (5.15), downward sloping in P_0 , $f_o(\cdot)$ is increasing and $f_o(\cdot) > g_o(\cdot)$, then $P_o^W < P_o^{\pi}$ by Lemma C.1.

Next, taking derivative of (5.34) and (5.29) with respect to N gives us:

$$\frac{dP_o^W}{dN} = -S^2 B_o \left[-\frac{Q_o + Q_p}{N^2} + \frac{\partial (Q_o + Q_p)}{\partial N} \frac{1}{N} \right] \text{ and}$$
$$\frac{dP_o^\pi}{dN} = S^2 B_o \left[-\frac{Q_o + Q_p}{N^2} + \frac{\partial (Q_o + Q_p)}{\partial N} \left(\frac{N+1}{N} \right) \right].$$

But, from Proposition 5.2.2 we know that $\partial (Q_o + Q_p) / \partial N = (Q_o + Q_p) / (N(N+1))$. Therefore $\frac{dP_o^W}{dN} = \frac{S^2 B_o (Q_o + Q_p)}{N(N+1)} > 0$ and $dP_o^{\pi} / dN = 0$.

(ii) That $\Delta P_{p-o}^{W} < \Delta P_{p-o}^{\pi}$ follows from writing the pricing rules (5.35) and (5.30) as $\Delta P_{p-o}^{\pi} = f_{p-o}(Q_p)$ and $\Delta P_{p-o}^{W} = g_{p-o}(Q_p)$. Since peak traffic Q_p is, by (5.16), downward sloping in ΔP_{p-o} , $f_{p-o}(\cdot)$ is non-decreasing as long as the (unsigned) term $D^{\prime\prime\prime}(\cdot)$ is non-negative (or if is negative, its magnitude is not too large) and $f_{p-o}(\cdot) > g_{g-o}(\cdot)$, then $\Delta P_{p-o}^{W} < \Delta P_{p-o}^{\pi}$ by Lemma C.1.

 $\frac{d\Delta P_{p-o}^{"}}{dN} > 0$ follows from differentiation of (5.35) with respect to N. We get:

$$\frac{d\Delta P_{p-o}^{W}}{dN} = (\alpha S + \beta) \left[\frac{Q_p D'(Q_p)}{N^2} + \frac{\partial Q_p}{\partial N} \frac{(N-1)}{N} D'(Q_p) + Q_p D''(Q_p) \frac{(N-1)}{N} \right]$$
$$+ S^2 (B_p - B_o) \left[\frac{Q_p}{N^2} - \frac{\partial Q_p}{\partial N} \frac{1}{N} \right]$$

Since $\partial Q_p / \partial N > 0$ by Proposition 5.2.1, the first term in the RHS is positive. Since $\partial Q_p / \partial N < Q_p / (N(N+1))$ by Proposition 5.2.1, $\frac{Q_p}{N^2} > \frac{\partial Q_p}{\partial N} \frac{1}{N}$ making the second term in the RHS positive as well. Therefore, $d\Delta P_{p-o}^W / dN > 0$.

 $\frac{d\Delta P_{p-o}^{*}}{dN} = 0$ if the delay function is linear, follows from differentiating (5.30) and then imposing $D''(Q_p) = 0$, $D'''(Q_p) = 0$. We get:

$$\frac{d\Delta P_{p-o}^{\pi}}{dN} = \left((\alpha S + \beta)D'(Q_p) + (B_p - B_o)S^2\right) \left[-\frac{Q_p}{N^2} + \frac{\partial Q_p}{\partial N}\frac{(N+1)}{N}\right], \text{ where } D'(Q_p) \text{ is a}$$

constant. With a linear delay function Proposition 5.2.1 changes to $\partial Q_p / \partial N = Q_p / (N(N+1))$, making $d\Delta P_{p-o}^{\pi} / dN = 0$. If we consider $D''(Q_p) > 0$, the sign is then undetermined and depends on the values of, for example, B_p , B_o and $D'''(Q_p)$.

(iii) $Q_p^W > Q_p^{\pi}$ flows from part (ii), and the comparative statics in (5.16) or Proposition 5.1.3. $\frac{dQ_p^{\pi}}{dN} > \frac{dQ_p^W}{dN} = 0$ follows from replacing ΔP_{p-o}^{π} in the sub-game equilibrium equation (5.12); we get:

$$\left(-\Omega^{p} + \Omega^{o}\right) (\Delta P_{p-o}^{\pi}) = \frac{2Q_{p}^{\pi}(B_{p} - B_{o})S^{2}(N+1)}{N} - \overline{\theta}S(B_{p} - B_{o}) + (\alpha S + \beta) \left(D(Q_{p}^{\pi}) + \frac{N+2}{N}Q_{p}^{\pi}D'(Q_{p}^{\pi}) + (Q_{p}^{\pi})^{2}D''(Q_{p}^{\pi})\right) = 0$$

As in the proof of Proposition 5.2.1, use this to calculate

$$\frac{dQ_p^{\pi}}{dN} = -\frac{\partial(-\Omega^p + \Omega^o)(\Delta P_{p-o}^{\pi})/\partial N}{\partial(-\Omega^p + \Omega^o)(\Delta P_{p-o}^{\pi})/\partial Q_p}$$

and to prove that $0 < dQ_p^{\pi} / dN < Q_p^{\pi} / (N(N+1))$, which shows that peak traffic increases with N.

Similarly, replacing ΔP_{p-o}^{W} in the sub-game equilibrium equation (5.12), we get:

$$(-\Omega^{p} + \Omega^{o})(\Delta P_{p-o}^{W}) = Q_{p}^{W}(B_{p} - B_{o})S^{2} - \overline{\theta}S(B_{p} - B_{o}) + (\alpha S + \beta)(D(Q_{p}^{W}) - Q_{p}^{W}D'(Q_{p}^{W})) = 0$$

Which does not depend on N, hence $dQ_p^W / dN = 0$.

(iv) $Q_t^w > Q_t^\pi$ flows from part (i) and the comparative statics in (5.15) or Proposition 5.1.1. To prove $\frac{dQ_t^\pi}{dN} > \frac{dQ_t^W}{dN} = 0$, use the sub-game equilibrium equation (5.10) to prove $\frac{dQ_o^\pi}{dN} > \frac{dQ_o^W}{dN} = 0$ in an analogous way as in part (iii). The result then follows directly from this and part (iii).

(v) Direct from part (iii) and $D'(Q_p) > 0$ (equation 5.4).

(vi)Consider the SW function in (5.33). We can rewrite it in terms of total traffic, Q_t , and peak-traffic, by replacing $Q_o = Q_t - Q_p$. This gives us:

$$SW(Q_t, Q_p) = \overline{\theta}S(B_pQ_p + B_o(Q_t - Q_p)) - c(Q_t) - C(Q_t) - Kr$$

- $\frac{S^2}{2}(B_o(Q_t - Q_p)^2 + 2B_o(Q_t - Q_p)Q_p + B_pQ_p^2) - (\alpha S + \beta)Q_pD(Q_p)$

Now, SW is globally concave in (Q_t, Q_p) because

$$\frac{\partial^2 SW}{\partial Q_t^2} = -B_o S^2 < 0, \quad \frac{\partial^2 SW}{\partial Q_p^2} = -(B_p - B_o) S^2 - (\alpha S + \beta) \left(2D'(Q_p) + Q_p D''(Q_p) \right) < 0$$

and

$$\frac{\partial^2 SW}{\partial Q_o^2} \cdot \frac{\partial^2 SW}{\partial Q_p^2} - \left(\frac{\partial^2 SW}{\partial Q_o \partial Q_p}\right)^2 = B_o S^2 \left[(B_p - B_o) S^2 + (\alpha S + \beta) \left(2D'(Q_p) + Q_p D''(Q_p) \right) \right] > 0$$

Since (Q_t^W, Q_p^W) maximizes SW, and from parts (iii) and (iv), $(Q_t^W, Q_p^W) > (Q_t^{\pi}, Q_p^{\pi})$, then $SW^{W} > SW^{\pi}$. Finally, since $(Q_{t}^{\pi}, Q_{p}^{\pi})$ increases with N, SW^{π} increases with N, while SW^w does not change with N because (Q_t^w, Q_p^w) does not. 100

C.4) Parameter values for numerical simulation

θ

Bp

Bo

0.1

0.13

The values on Table c.1 were used to simulate the model and obtain the dead weight-losses presented in the paper. Most of the values were used in the simulations of chapter 4, and were actually drawn from results from the literature. See Table 4.1 and footnote 25 in Chapter 4. The capacity level, which in this Chapter is exogenous and in Chapter 4 was endogenous, is close to the capacity JFK airport had in 2000 (56 flights per hour).

Airlines Airports Demand 40 61,000 S 60 Κ 50 α $\overline{ heta}$ β 3000 1,000 Ν varies С 2,000

с

36,000

5

Lp

Table c.1: Parameter values for the numerical simulation