ESSAYS ON AIR TRANSPORT AND PUBLIC POLICY

by

Chi-Lok Yuen

B.B.A., The Chinese University of Hong Kong, 1999
M. Phil., The Chinese University of Hong Kong, 2003
M.A., The University of British Columbia, 2004

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

DOCTOR OF PHILOSOPHY

in

THE FACULTY OF GRADUATE STUDIES

(Business Administration)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

December 2008

ABSTRACT

The rapid growth in air travel demand in the last few decades has led to two major public policy issues in the aviation industry. First, it has placed enormous pressure on the existing airport infrastructure. As a result, airlines, passengers and shippers are suffering from serious congestion delays at the facilities, implying a significant economic loss to the society. Second, given the increasing air travel demand, aviation greenhouse gas (GHG) emissions – a major contributor to man-made climate change – are growing at a rapid pace. Different public policies are proposed by policy makers and scholars in a number of different disciplines for the two issues. This thesis aims to investigate and evaluate some of the policies from different perspectives.

To deal with the airport congestion problem, the congestion pricing is considered to be one of the most feasible and simplest solutions, and economists also argue that it would be welfare-improving. However, the congestion pricing has not really been implemented at airports in practice. Chapter 2 considers the case of variable passenger time costs in the airport congestion pricing analysis, and examines its welfare-redistributive issues. This may help us to explain the unpopularity of the congestion pricing in practice. The Chapter also explores a case where the self-internalization – an important hypothesis suggested in the literature – may be incomplete. Chapter 3 investigates the effects of congestion pricing at a gateway on its hinterland's road tolls, road congestion and social welfare. The problem will become more practically relevant, as it is expected that gateway congestion pricing will be getting more popular in the near future.

Another possible solution to deal with the problem of capacity shortage at airports would be to utilize the existing facilities more efficiently. Chapter 4 measures the airport efficiency in China, and empirically investigates the factors, including competition and policy changes, affecting it. Chapter 5 considers another major public policy issue in the aviation industry – GHG emissions. The Chapter provides an analytical framework for examining the issue, and investigates the effects of unilateral GHG control measures on airline competition, market output, consumer benefits and world emissions.

TABLE OF CONTENTS

Appendixes

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGEMENTS

I would like to acknowledge and extend my heartfelt gratitude to the following persons who made the completion of this dissertation possible. First of all my special words of appreciation go to my supervisor, Professor Anming Zhang, for his innumerable comments, notes and suggestions on the drafts of my papers. He invested a lot of time and effort into discussing my work, as well as giving continuous support and encouragement throughout my time at UBC.

I am also genuinely grateful and obliged to the other members of my dissertation committee, Professors Tae Oum, David Gillen and Kevin Milligan, for their help, advice and support.

Many others have also helped me with important comments and suggestions about my work: I would like to thank Leonardo Basso, Xiaowen Fu, Heng Ju, and participants at OLOG division seminar, Canadian Economics Association Annual Meeting, and International Forum for Shipping, Ports and Airports. I would also like to thank my fellow students, Jing Shao, Ran Jin and Liping Liang for their friendship and support. In addition, I am tremendously thankful for the financial support I received from University Graduate Fellowship from UBC, Canadian Railway Fellowship in Transportation Studies, and Shelby L. Brumelle Memorial Graduate Scholarship.

Last but certainly not least, I am most deeply grateful to my family and friends for their love and support.

CO-AUTHORSHIP STATEMENT

The first drafts of the papers contained in Chapters 2, 3 and 4 of this thesis were written by me, but have subsequently benefited from major revisions by my advisor and co-author, Professor Anming Zhang, and Leonardo Basso for Chapter 3.

Chapter 1 Introduction

1.1 Research Topic and Content of the Thesis

The global air travel market has been growing significantly in the last few decades, and is expected to continue growing at a rapid pace in the coming years, especially in the Asia Pacific region. On one hand, the rapid growth in air travel implies that the aviation sector has had an increasingly important role in facilitating the flow of passengers and goods, thereby making it a key driving force for economic development and growth. On the other hand, the rapid traffic growth has also led to problems for the economy. First, as it has placed enormous pressures on facilities at major airports, airlines, passengers and shippers are suffering from delays at those airports, implying a huge economic loss to the society. Second, the rapid growth in air travel also leads to a huge increase in aviation greenhouse gas (GHG) emissions, which have been found to be a major contributor to man-made climate change. To deal with the two issues, different public policies are proposed or implemented in different countries. Due to the fact that the policies may have significant implications for the economy, they have received a great deal of attention from policy makers and scholars in a number of different disciplines. Some of the major policies on the issues are examined in four essays, which are Chapter 2 to Chapter 5, in this thesis.

The airlines' market liberalization, economic growth and globalization over the last several decades have significantly stimulated air passenger travel and cargo

movements, resulting in severe congestion at most of the major airports in the world. For example, US DOT (2006) estimated that aircraft delays cost passengers $9.4 billion. A recent study by the Oxford Economic Forecasting (2006) suggested that UK airport congestion cost (US) $3.2 billion in 2005. Numerous studies attempted to provide solutions for the airport congestion problem. Among the solutions, airport congestion pricing may be one of the most feasible and simplest solutions to deal with the problem. Since the late 1990s, the literature on airport congestion pricing has witnessed a revival (see Basso and Zhang, 2007a, for a comprehensive review). An important advancement of the general theory in the literature is to use the vertical-structure approach in the analysis since Brueckner (2002). In this approach, the airline market is formally modeled as oligopoly, which takes airport decisions, such as pricing and capacity investment, as given. In the paper, Brueckner argue that each airline at an airport only internalizes the congestion one of its flights imposing on its other flights but not that on other airline's flights (i.e., the so-called "self-internalization behavior"). As a result, the congestion is fully internalized under monopoly, and the congestion toll is unnecessary from the social efficiency point of view. For an oligopoly market, carriers only internalize the congestion they impose on themselves. The optimal congestion toll should be equal to the congestion cost imposed on other air carriers.

Following Brueckner (2002), Pels and Verhoef (2004) and Brueckner (2005) further considered the network issues in the analysis. In particular, Pels and Verhoef (2004) indicate that the optimal airport pricing would have two components: a positive congestion toll and a subsidy to correct the market-power distortion. As a result, the sign

of the optimal airport pricing is generally undetermined, while it tends to be positive when the airline market is more competitive. They also find that lack of cooperation among airport regulators in an aviation network may lead to welfare losses when compared to a no-tolling situation. On the other hand, Zhang and Zhang (2006) and Basso and Zhang (2007b) extend the literature by considering airport capacity as endogenous. Zhang and Zhang (2006) find that given the optimal airport congestion pricing, airline market structure has no impact on airport capacity and congestion for a welfare-maximizing airport that receives public subsidy. Basso and Zhang (2007b) investigate rivalry between congestion facilities (e.g., airports) and its effects on facility charges, capacities and congestion delays.

By using the vertical structure approach, Chapters 2 and 3 in this thesis look at two different issues in airport congestion pricing, which are less explored in the literature. Chapter 2, entitled "Airport Congestion Pricing and its Welfare Implications: The Case of Variable Passenger Time Costs," aims to examine two issues in the literature. First, while the self-internalization hypothesis – carriers may internalize some of the congestion they create, taking into account the delays they impose on themselves – is proposed in the literature (e.g., Brueckner, 2002), empirical results (e.g., Daniel, 1995; Brueckner, 2002; Mayer and Sinai, 2003; Morrison and Winston, 2007; Daniel and Harback, 2008) have not yet had consensus whether or not the self-internalization behavior exists. Second, although it is argued by economists that airport congestion pricing could be socially optimal, it has not really been implemented at airports worldwide.

The Chapter develops an analytical framework for airport congestion pricing by considering the variable passenger time costs in order to resolve the two issues. We find that, first, congestion delays will not be fully internalized at the carrier level under different market structures, including those airports dominated by a monopoly airline. Thus, congestion pricing is necessary regardless of the market structures. On the other hand, it suggests that the self-internalization may be incomplete, when the variable passenger time costs are taken into account in the analysis. Second, our results suggest that although the congestion pricing is welfare-improving, it may not be Pareto-improving. In particular, our analytical results and numerical simulation show that airlines and passengers are likely to be worse off if the congestion pricing is implemented. This may explain why congestion pricing appears unpopular in the real world. The result also highlights the importance of the welfare-redistributive issues in carrying out the congestion pricing schemes at airports in practice.

Chapter 3 is entitled "Effects of Gateway Congestion Pricing on Optimal Road Pricing and Hinterland." In practice, the gateway-hinterland intermodal transport system is very common. For example, cargos are shipped to the Port of New York (the gateway) and then are distributed to the hinterland mainly through the New Jersey highway system. Similarly, major gateways, such as Southern California, Antwerp, Rotterdam, Hamburg, Tokyo, Hong Kong, Shanghai and Singapore, serve their respective hinterlands by intermodal transport systems. As argued by Van Klink and Van den Berg (1998), those gateways are in a unique position to stimulate intermodal transport, and also use the intermodal systems to enlarge their hinterlands. In this type of system, cargo traffic

destined for (or originated from) an interior region – the hinterland – passes through the gateway. As final consumers of cargo care about "total cost" (sum of costs at gateway and hinterland), any operating/pricing decisions at one facility will have significant implications on another facility and region. We particularly consider how gateway congestion pricing will affect its hinterland. It is because more gateways are expected to adopt congestion pricing in the near future, which may have significant impacts on its hinterland.

This Chapter finds that if the gateway maximizes the joint profit of itself and its oligopoly carriers, its charge will rise, with part of this increase owing to congestion pricing. This increase in the gateway charge will lead to lower road tolls. Finally, while the change in road congestion is in general ambiguous, the hinterland's welfare will fall as a result of the increase in the gateway charge. In addition to the airport congestion pricing literature, this Chapter is also related to the rapidly growing literature on second-best road pricing in a transport network (see Small and Verhoef, 2007; Ubbels and Verhoef, 2008 for recent reviews). In terms of model structure and problem being examined, our analysis is close to three papers in this literature, namely, De Borger, et al. (2007), De Borger, et al. (2008) and Ubbels and Verhoef (2008). The major difference of this Chapter and those papers will be discussed in the following.

In addition to the congestion pricing discussed in Chapters 2 and 3, another possible solution to deal with the problem of capacity shortage at airports would be to utilize the existing facilities more efficiently. Assessment of airport efficiency has

become the focus of a large number of studies. Different methodologies have been used to measure the efficiency of airports in different regions around the world (see Oum, et al., 2003, for a comprehensive review). The airport efficiency problem is particularly important for China: In the past few decades, as rapid economic growth has significantly increased the demands for air services in China, it has placed enormous pressure on its airport infrastructure. A recent paper by Fung, et al. (2008) attempted to calculate the productivities for twenty-five major Chinese airports between 1995 and 2004. They find that over that period, airport efficiency was improving and the productivity among airports from different regions was converging. Using their data, Zhang and Yuen (2008) further investigate whether privatization through public listing improves airport performance.

Chapter 4, entitled "Effects of Competition and Policy Changes on Capacity Utilization Efficiency of Chinese Airports: An Empirical Investigation," estimates the Chinese airport efficiency and investigates the factors affecting it. To further the studies in the literature with an expanded dataset, Chapter 4 estimates both the efficiency level and its growth for twenty-five sample Chinese airports by using Data Envelopment Analysis. After controlling for hub status and other airports' characteristics, we find that: (i) publicly listed airports are significantly more efficient than non-listed airports; (ii) airports with more competition are more efficient than their counterparts; (iii) the airports' efficiency and the technical progress are positively correlated with the airport localization program, in which Chinese airports are transferred from the Chinese Aviation

Administration of China (CAAC) – a ministry-level organ of the Chinese central government – to its local counterparts.

Another important issue in the aviation industry is GHG emissions and its control measures. Public concerns over climate change have pressured both the aviation industry and regulators to mitigate aviation GHG emissions in hope of avoiding their adverse impacts on global climate. Thus, some governments consider unilaterally mandating some measures to control emissions from the industry. For example, one of the largest-scale unilateral GHG control measures would be the inclusion of the aviation sector in the EU emission trading system in 2012. Due to the potentially significant implications of the unilateral GHG control measures, the measures ought to be examined more rigorously. Although a growing number of empirical studies have been conducted to evaluate the unilateral control measure (e.g., Scheelhaase and Grimme, 2007; Benito, 2008; Hofer, et al., 2008; Forsyth, 2008; Forsyth and Ho, 2008; Scheelhaase, et al., 2008), there is lack of analytical work on the issue.

Chapter 5, entitled "Unilateral GHG Control Measure and Aviation Industry: A Theoretical Analysis," investigates the effects of unilateral GHG control measures on the aviation industry and total GHG emissions. The result suggests that if a country unilaterally takes actions to control aviation GHG emissions, the charges at its hub and spoke airports may increase. This increase in the airport charges will lead to a shift of domestic connecting flights between hub and spoke airports outside the country. It may also place its home airlines at a strategic disadvantage, as the home airlines usually

operate their hubs at the home country. Furthermore, the unilateral control measure will reduce GHG emissions in the country implementing the measure, while it may lead to an increase or a decrease in world emissions. It is because as mentioned, the domestic connecting flights may increase at another country, in which the airline network may be very inefficient in terms of GHG emissions.

1.2 References

Basso, L.J. and Zhang, A. (2007a) "An Interpretative Survey of Analytical Models of Airport Pricing", in: Lee, D. (Ed.), *Advances in Airline Economics*, Vol.2, Elsevier, 89-124.

Basso, L.J. and Zhang, A. (2007b) "Congestible Facility Rivalry in Vertical Structures", *Journal of Urban Economics*, 61(2), 218-237.

Benito, A. (2008) "Climate Change Impact as a Constraint to the Air Transport Growth: The European Case", Paper presented at World Conference on Transport Research, Berkeley, California.

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(7-8), 599-614.

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. and Harback, K.T. (2008) "(When) Do Hub Airlines Internalize their Self-Imposed Congestion Delays?", *Journal of Urban Economics*, 63, 583-612.

De Borger, B., Dunkerley, F. and Proost, S. (2007) "Strategic Investment and Pricing Decisions in a Congested Transport Corridor", *Journal of Urban Economics*, 62(2), 294-316.

De Borger, B., Proost, S. and Van Dender, K. (2008) "Private Port Pricing and Public Investment in Port and Hinterland Capacity", *Journal of Transport Economics and Policy*, 42(3), 527-561.

Forsyth, P. (2008) "Emissions Trading Schemes – Should International Aviation be Excluded?", *manuscript.*

Forsyth, P. and Ho, T.V. (2008) "Aviation's Indirect Carbon Emissions: Measurement and Implications", Paper for the German Aviation Research Society Workshop Aviation and Climate Change, Amsterdam, July.

Fung, M., Wan, K., Hui, Y.V. and Law, J. (2008) "Productivity Changes of Chinese Airports 1995–2004", *Transportation Research Part E*, 44(3), 521-542.

Hofer, C., Dresner, M. and Windle, R. (2008) "The Environmental Effects of Airline Carbon Emissions Taxation in the U.S.", Paper presented at Air Transport Research Society Conference, Athens.

Mayer, C. and Sinai, T. (2003) "Network Effects, Congestion Externalities, and Air Traffic Delays: Or Why Not All Delays are Evil", *American Economic Review*, 93, 1194-1215.

Morrison, S.A. and Winston, C. (2007) "Another Look at Airport Congestion Pricing", *American Economic Review*, 97(5), 1970-1977.

Oxford Economic Forecasting (2006) "The Economic Contribution of the Aviation Industry in the UK", Oxford Economic Forecasting, Oxford.

Pels, E. and Verhoef, E.T. (2004) "The Economics of Airport Congestion Pricing", *Journal of Urban Economics*, 55, 257-277.

Scheelhaase, J.D. and Grimme, W.G. (2007) "Emissions Trading for International Aviation – An Estimation of the Economic Impact on Selected European Airlines", *Journal of Air Transport Management*, 13, 253-263.

Scheelhaase, J.D., Grimme, W.G. and Schaefer, M. (2008) "The Inclusion of Aviation into the EU Emission Trading Scheme – Impacts on Competition between European and non-European Network Airlines", *manuscript*.

Small, K.A. and Verhoef, E.T. (2007) *The Economics of Urban Transportation*, Routledge, London.

Ubbels, B., and Verhoef, E.T. (2008) "Governmental Competition in Road Charging", *Regional Sceince and Urban Economics*, 38, 174-190.

US DOT (2006) "National Strategy to Reduce Congestion on America's Transportation Network", U.S. Department of Transportation, Washington, DC.

Van Klink, H.A. and Van den Berg, G.C. (1998) "Gateways and Intermodalism", *Journal of Transport Geography*, 6(1), 1-9.

Zhang, A. and Yuen, A. (2008), "Airport Policy and Performance in Mainland China and Hong Kong", in: Winston, C. and de Rus, G. (eds.), *Aviation Infrastructure Performance: A Study in Comparative Political Economy*, 159-192, Washington, D.C.: Brookings Institution Press.

Zhang, A. and Zhang, Y. (2006) "Airport Capacity and Congestion When Carriers have Market Power", *Journal of Urban Economics*, 60, 229-247.

Chapter 2 Airport Congestion Pricing and its Welfare Implications: The Case of Variable Passenger Time Costs[1]

2.1 Introduction

During the last several years, airlines and passengers have been suffering from congestion at 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 congestion by calling for the use of a price mechanism, under which landing fees are based on an aircraft's contribution to congestion (see Basso and Zhang, 2007a, for a comprehensive review). In particular, the intra-day utilization of airport facilities varies, and thus the congestion levels differ at different times of a day. As a result, "peak load" congestion pricing was proposed, which involves charging different landing fees at different times: during peak hours, flights are charged higher rates than during off-peak hours.[2] Congestion pricing is desirable from an efficiency point of view because social marginal cost varies depending on the time that the runway and slots are used, and so the peak-load pricing better reflects this varied

[1] A version of this Chapter is forthcoming as: Yuen, A. and Zhang, A. (2009) "Airport Congestion Pricing and its Welfare Implications: The Case of Variable Passenger Time Costs", *Pacific Economic Review.*

[2] The landing fee is the main component of airport pricing. At Hong Kong International Airport (HKIA), for instance, the airport charges consist of three components: landing, aircraft parking and terminal building charges (TBC). Landing fees are based on aircraft movements, while TBCs are passenger-based. In a "typical" case (a B747 with 418 seats, 66.8 percent load factor and 4-hour turnover), the landing fee accounts for 68.7 percent of the total charge, parking 6.6 percent and TBC 24.7 percent.

marginal cost than a flat fee structure.[3] A vast amount of literature has considered peak-load pricing at the airport level (e.g., Morrison, 1983; Morrison and Winston, 1989; Oum and Zhang, 1990; Daniel, 1995, 2001),[4] while Brueckner (2002, 2005), and Basso and Zhang (2008) further investigated peak-load pricing at both the airport and airline levels.

One important implication of our analysis is related to the debate in the literature on whether airlines internalize their self-imposed congestion. The self-internalization hypothesis is firstly examined in Brueckner (2002), which pointed out that, in making scheduling decisions under Cournot competition, carriers may internalize some of the congestion they create, taking into account the delays they impose on themselves. Further studies (e.g., Brueckner, 2005; Pels and Verhoef, 2004; Zhang and Zhang, 2006; Basso, 2008) have added insights to the literature from different perspectives. An important lesson from the studies is that congested airports should charge only the purely external congestion costs an airline impose on its rivals, thus an optimal congestion toll is decreasing in the airline's market share. To an extreme, a monopoly carrier will fully internalize all the congestion costs, and congestion pricing is unnecessary. However, whilst Brueckner (2002) and Mayer and Sinai (2003) found empirical evidence to support the self-internalization hypothesis, empirical results in Daniel (1995), and Daniel and Harback (2008) suggested that an atomistic model, in which carriers do not internalize any congestion, fitted the data better. In particular, Daniel (1995) argued that

[3] A less mentioned benefit of congestion pricing is that it may facilitate airlines' entry into routes from congested airports. Dresner, et al. (2002) found that among the airport barriers they analyzed, gate utilization during peak periods was the most significant deterrent to entry in the U.S. airline industry; Ciliberto and Williams (2007) measured the importance of operating barriers to entry, including limited access to airport facilities, as determinants of the hub premium.

[4] Although Arnott, et al. (1993) did not consider airports explicitly, their peak-load pricing model is applicable to congestible facilities in general.

Stackelberg/fringe competition would lead to the non-internalization result. Furthermore, Morrison and Winston (2007) found that the decline in aggregate net benefits from charging atomistic tolls instead of charging tolls which take self-internalization behavior into account is small. Brueckner and Van Dender (2007) further provided a unified analytical framework for the two arguments, which is based on whether airlines behave atomistically or not. Different from the literature, this Chapter explores a case where incomplete self-internalization may occur even if air carriers are non-atomistic.

The second objective in this Chapter is to explore the benefit distribution of airport congestion pricing. Although it is argued by economists that airport congestion pricing could be socially optimal, it has not really been implemented at airports worldwide, especially in the Asia-Pacific region. Due to their importance to local economies and their monopoly nature,[5] most airports are regulated by governments or outright publicly owned.[6] Airport operating decisions, including pricing, can be politically sensitive, and thus the welfare-redistributive issue plays an important role in the decision-making process.[7] In his simulation model, Daniel (2001) discussed the re-distributive issue and evaluated several price-and-rebate programs, some of which are self-financing and Pareto-improving. However an effective re-distributive mechanism is

[5] Fung, et al. (2006) found that the total benefit brought by the aviation industry was 7.02 percent of the Hong Kong's GDP at factor cost (at current prices) in 2003; The aviation industry in UK directly employed 186,000 people (full-time equivalents) in 2004, and helped to support over 520,000 jobs in total including those employed in its supply chain and in travel agents, and the jobs dependent on the spending of its employees (Oxford Economic Forecasting, 2006).

[6] World Bank (2003) noted that due to the "natural monopoly" characteristics of some airport services, governments developed regulatory policies for airport systems.

[7] For example, the Canadian Airport Authorities are required to give not less than 60 days advance public notice in local news media for all changes in user charges (excluding rent), together with an explanation for the noted increase.

usually not in place.[8] To shed some light on the issue, we discuss the re-distributive impact of congestion pricing by taking variable passenger time costs into account.[9]

In particular, this Chapter considers variable passenger time costs in airport congestion pricing analysis.[10] A common assumption made in the literature is that all passengers' value of travel time saving (VTTS) is identical. In other words, for a given number of flights, time costs due to congestion delays resulting from traveling in a congested period are the same for all travelers. This assumption may be unrealistic. For instance, business passengers' VTTS would be higher than that of their leisure counterparts. This observation is supported by empirical work which indicates that air travelers' VTTS depends, for example, on their travel purposes and wage rates. In particular, Hess, et al. (2007) found that business and leisure travelers' willingness-to-pay (WTP) for on-time performance are $10.39 and $7.02, respectively.[11] In addition, von Wartburg and Waters (2004) contain a detailed discussion on the relationship between VTTS and income. Their comprehensive survey of the relevant studies (e.g., MVA Consultancy, et al., 1987; AHGC, 1999; Gunn, et al., 1999; Wardman, 2001; Mackie, et

[8] In Japan, local residents are subsidized by the Narita Airport for aviation noise. Yet there is no similar re-distributive mechanism for congestion delays at the airport level in practice. In contrast, since 2003, the CAAC has required Chinese airlines to compensate their passengers for flight delays.

[9] Another reason argued in the literature to explain why the airport congestion pricing is unpopular in practice is the competitive issues due to the pricing. As argued in Brueckner (2008), the inverse association between a carrier's market size and the toll it pays would be controversial in practice, leading to potential political opposition from smaller carriers. In the paper, Brueckner also showed that a slot-distribution regime and the slot-auction regime may be able to reach the social optimum as the differentiated-toll regime does.

[10] Daniel (2001) considers the case where airlines are heterogeneous in time layover and queuing time values, while consumers' heterogeneity in time valuation is not taken into account. In addition, De Palma and Lindsey (2004) examined road congestion pricing with travelers' heterogeneity in wages, values of travel time, and the congestion characteristics of their vehicles. In the analysis, they considered atomistic road users. As a consequence, the congestion internalization by non-atomistic carriers – one of the major issues we want to address – was not examined.

[11] The WTP for on-time performance in the study is defined as the monetary amount which an individual is willing to pay for a 10-percent increase in the number on-time flights.

al., 2001) suggests that the income elasticity of VTTS is positive. As will be shown below, considering the heterogeneity in passenger time cost may have significant implications on airlines' self-internalization behavior, optimal airport congestion pricing and its welfare-redistributive issues.

The analytical model in this Chapter extends the framework in Brueckner (2002) by incorporating variable passenger time costs into the analysis, a case that was briefly discussed by Brueckner. We find that congestion costs may not be fully internalized under different market structures. First, as also discussed in the literature, each oligopoly airline at an airport only internalizes the congestion one of its flights imposing on its other flights but not that on other airlines' flights. Second, time costs of passengers with high travel benefits (for example, business travelers) are not fully reflected in the carriers' pricing, as the carriers are only concerned with marginal passengers, who are indifferent to travel in the peak and off-peak periods, in their profit maximization. Given this result, a social welfare-maximizing airport may need to impose a positive congestion toll to induce full internalization of congestion at the carrier level. In the welfare analysis, we find that if a positive congestion toll is imposed, airlines and some passengers who have low time valuation may be worse-off in the case of variable passenger time costs. This non-Pareto-improving result may explain why airport congestion pricing appears unpopular in practice.[12]

[12] It may be argued that a re-distributive mechanism may leave all groups better off. However, as discussed below, no such effective re-distributive mechanism is actually implemented in practice. On the other hand, congestion pricing may be opposed if there are doubts that compensation through the re-distributive mechanism will continue in perpetuity. In particular, Stiglitz (1998) identified this as a reason why airport congestion pricing has not been implemented in the U.S. It is noted that general aviation users

The Chapter is organized as follows. Section 2.2 discusses the airport congestion and delay in Asia Pacific. Section 2.3 sets up the model. Sections 2.4 and 2.5 consider carrier competition and airport pricing, respectively. Section 2.6 discusses the welfare implications of the congestion pricing on airport, air carriers and passengers. Numerical simulation is conducted in Section 2.7, and Section 2.8 concludes.

2.2 Airport Congestion and Delay in Asia Pacific

Although the problem of airport congestion and delay has been studied extensively in other parts of the world (e.g., Brueckner, 2002; Mayer and Sinai, 2003, for US; and Reynolds-Feighan and Button, 1999; Santos and Robin, 2007, for Europe), we are not aware of any published studies on the problem for the Asia-Pacific region. The aviation market in Asia Pacific has witnessed a substantial growth in the last few decades due to the rapid economic growth and air liberalization in the region. This fast-growing air transport demand has imposed great pressures on the existing airport capacity, and has led to congestion at major airports in the region. For instance, in China, passenger terminals operated, on average, 15 percent above their capacities, whereas cargo terminals and facilities could only handle 65 percent of potential demand (measured at existing prices) throughout most of the 1990s (Zhang and Chen, 2003). The congestion problem is particularly severe at some major Chinese airports. For example, due to

---

pay only nominal fees, and the introduction of congestion in combination with compensation would make the subsidy to general more transparent and politically untenable in the long run.

increasing congestion and delays, in 2007 the Civil Aviation Administration of China (CAAC) – the country's industry regulator – imposed daily ceilings of landings/taking-offs not exceeding 1,100 and 650 flights at Beijing Capital International Airport and Shanghai Pudong International Airport – the country's two busiest airports – respectively. In the southern China's Pearl River Delta, it is reported that 2,996 flights departing from Hong Kong were delayed in the first ten months of 2006, comparing with the 1,733 delayed flights in 2005, and 69.7 percent of flights from/to Macau were delayed in May 2006 (Zhang and Yuen, 2008).[13] For other Asian countries, the Thai government was required to reopen its old airport at Bangkok in March 2007 to ease air traffic congestion at the new airport which just began operation in 2006.

Furthermore, given the expected rapid growth in the air market in the Asia-Pacific region, the airport congestion problem may be likely to get worse in the future. On one hand, the rapid increase in air demand may be due to the fast-growing economy in the region; the Asian Development Bank forecasted that Asia would still register robust average GDP growth of 7.6 percent in 2008, despite the global credit crisis. On the other hand, the air liberalization in the region would further stimulate the rapid growth in its air market. For example, China has been moving towards a more liberal international policy regime since the early 2000s, which includes signing liberal air service agreements with U.S. and Hong Kong, and granting a significant number of 5th freedom traffic rights to foreign airlines (Zhang and Yuen, 2008). On the other hand, in December 2008, member

[13] It is noted that the flight delays at airports may also be due to, for example, weather, security, air traffic control and airlines' operation problems.

countries of the Association of Southeast Asian Nations (ASEAN) may sign the ASEAN Multilateral Agreement on the Full Liberalization of Air Freight Services and Air Services, which will lead to the full liberalization of freight and other air services among the countries, and seek to build a unified aviation market by 2015. Given the robust economic growth and air liberalization in the region, its air market is expected to experience a rapid growth; ICAO (2007) predicted that the market would grow at 5.8 percent per annum through 2025, and to become the largest aviation market in the world in terms of both passenger and freight traffic. As a consequence, more airports in the region may experience congestion, while the situation is expected to get worse at those currently congested airports, if no suitable measures are taken. For example, Figure 2.1 shows that while Hong Kong airport is currently operated under its capacity, the existing capacity would not be able to meet the demand in 2020. Thus there is an urgent need for policy makers in the region to look for solutions to relieve congestion delays at their airports.

[Figure 2.1a here]

[Figure 2.1b here]

Although airport expansion has been the major measure for dealing with the airport congestion problem in the Asia Pacific region,[14] it may not be sustainable in the future for several reasons. First, along with economic development, people are more aware of environmental issues (e.g., noise and air pollution concerns), and are also more organized in protecting their own interests.[15] Thus, airport expansion plans may face a strong opposition from local residents. For example, because of its local residents' opposition due to aviation noise and land expropriation, the expansion of Japan's Narita Airport has been limited. Second, congestion is not only experienced at airport facilities, but at airspace as well. For example, one of the major concerns for building the third runway at Hong Kong Airport is whether airspace congestion in the Pearl River Delta region can be relieved.[16] Finally, airport expansion usually involves a huge capital investment and is time-consuming,[17] whereas the demand in the future is highly uncertain.

---

[14] The CAAC commenced construction on 35 new airports in 2004; the Centre for Asia Pacific Aviation estimated that China would have 240 commercial airports by 2010 and the capital needed to meet that infrastructure requirement would top US$1.3 billion a year (Air Transport World, February 2006).

[15] In the environmental economics literature, the Environmental Kuznets Curve (EKC) hypothesis, which predicts pollution rises with income as long as income is relatively low and declines once income has exceeded a threshold level (e.g., per capita GDP US$8,000 found in Grossman and Krueger, 1995), is empirically supported in many studies (for example, Hettige, et al., 1992; Grossman and Krueger, 1993, 1995; Selden and Song, 1994). One of the explanations for the downward-sloping part of EKC is that when income increases, households choose more environmental quality through public action (McConnell, 1997). Note that per capita GDP of many Asian cities has been close to or above the threshold (for example, in 2006, the per capita GDP of Guangzhou and Shanghai were US$11,000 and US$7,200, respectively).

[16] Philip N.L. Chen, the Chief Executive of Cathay Pacific, argued that a third runway would not alleviate the insufficient time slots at Hong Kong Airport, if the problem of limited airspace would not be cleared up first (18th January 2007, South China Morning Post).

[17] Peter K.N. Lok, the former director of the Hong Kong Civil Aviation Department, estimated that the required investment of the proposed third runway at Hong Kong Airport could be up to HK$30 billion (US$3.8 billion) (16th January 2007, Ta Kung Pao).

As a result, airport expansion is not likely to be the only solution to ease airport congestion in the region, and other possible complementary solutions need to be explored. Here airport congestion pricing may be one of the most feasible and simplest solutions to deal with the problem.[18] Thus, although airport congestion pricing has not really been practiced in the region currently, airport congestion pricing analysis is practically relevant to its aviation market from a forward-looking perspective. It may also provide a better guidance for airports in the region to set an appropriate level of congestion tolls and help gain political support in the future.

2.3 The Model

We consider a three-stage model of airport and airline behavior, in which N air carriers are in service at a congested airport.[19] In the first stage, the airport decides airport charges for the carriers in the peak and off-peak periods. For simplicity, we assume that the off-peak period is never congested,[20] while the peak period is always congested.[21] In

[18] Another way to relieve airport congestion is by slot control. As argued by Czerny (2007), without uncertainty, the use of congestion pricing is equivalent to slot control. Thus the congestion pricing analysis, as in this Chapter, can also shed some light on the use of slot control in practice.

[19] For simplicity, this Chapter, as in Brueckner (2002), focuses on a single-airport case. As a consequence, the system-wide effects of congestion pricing in a network of airports (see Brueckner, 2005; Pels and Verhoef, 2004) and the competition between congested airports (see Basso and Zhang, 2007b) are not examined in the present analysis. In addition, the network benefits associated with the hub and spoke system, which may lead to flight delays (see Mayer and Sinai, 2003), are not considered in this Chapter.

[20] This could be the case that the off-peak traffic in equilibrium is small enough relative to airport capacity.

[21] This assumption rules out the possibility that a period shifting from the peak to off-peak, and vice versa in different scenarios. In order to make the analysis tractable, the assumption is necessary and has also been made in other peak-load pricing studies (e.g., Brueckner, 2002, 2005; Basso and Zhang, 2008). It will also be assumed in our numerical simulation in Section 2.8.

the second stage, each carrier chooses its output in terms of the number of flights in the two periods. Finally, consumers choose between three mutually exclusive alternatives, namely, no travel, travel in the peak or off-peak periods.

Consumers are represented by a continuum with index θ, which is uniformly distributed between 0 and Φ with a unit density. A consumer's utility depends on his/her travel decision and θ as follows:

$$V(\theta) = \begin{cases} x & \textit{if not travel} \\ x + b_o(\theta) & \textit{if travel in the off-peak period} \\ x + b_p(\theta) - H(\theta)t(n_p) & \textit{if travel in the peak period} \end{cases} \quad (2.1)$$

where x is consumption expenditure on other goods. The (gross) benefits from peak and off-peak travel are given by the functions $b_p(\theta)$ and $b_o(\theta)$, respectively. Here we assume $b_p^{'}, b_o^{'} > 0$, i.e., the (gross) travel benefits are higher for a high-θ passenger than a low-θ passenger. As such, θ may be interpreted as an index of the passenger's tendency to travel on business (as opposed to leisure travel), as travel is a crucial job requirement for business travelers. Furthermore, we assume that peak and off-peak travel would be vertically differentiated, such that all passengers will prefer traveling in the peak period as opposed to the off-peak period with identical delays and airfares. In other words, $b_p(\theta) > b_o(\theta)$ for all θ. This vertical-differentiation feature of air travel can arise if the peak period represents the day's more desirable travel time.

In addition, travelers in the peak period incur delay costs, $H(\theta)t(n_p)$, where $t(n_p)$ is the additional travel time (for example, in hours) due to congestion, which is increasing in n_p;[22] $H(\theta)$ is the VTTS for passenger θ.[23] We assume that a passenger's VTTS and gross travel benefit are correlated. In practice, passengers with higher gross travel benefits, such as business travelers, also have higher VTTS, as discussed in the introduction. Thus we further assume $H'(\cdot) > 0$. Note that if we consider constant time cost across passengers, $H(\cdot)$ is a constant and the model is reduced to that in Brueckner (2002). Moreover, since business travel must occur during the early and late peak hours to avoid disruption of the workday, peak-travel benefits should increase at a higher rate relative to off-peak benefits as θ increases, yielding $b_p'(\theta) - H'(\theta)t(n_p) > b_o'(\theta)$ for all θ. This is referred to as the single-crossing assumption. The assumption ensures that if a traveler with lower θ travels in the peak period, then the counterpart with higher θ must also travel in the peak period. It implies that, in the interior solution case, $\Phi > \theta^* > \underline{\theta} > 0$, where θ^* and $\underline{\theta}$ denote consumers who are indifferent to travel between peak and off-peak, and travel off-peak and not travel, respectively. Thus consumers with $\theta \in [\theta^*, \Phi]$ will travel in the peak period, consumers with $\theta \in [\underline{\theta}, \theta^*)$ will travel in the off-peak period, and consumers with $\theta \in [0, \underline{\theta})$ will not travel. To obtain the interior

[22] In this Chapter, we consider the airport capacity is exogenously given. As a result, the congestion delays only depend on the number of peak flights. The case of variable and endogenous capacity is examined in Zhang and Zhang (2006) and Basso (2008) in a congestion-pricing framework.

[23] The value of reliability in reducing congestion is also discussed in the literature. Small, et al. (2005) examined road users' values of travel time and reliability. On the other hand, for the cargo market, due to uncertain delays, firms are required to increase their inventories so as to prevent the shortage of inputs in production and goods to sell if delivery times are uncertain; Gausch and Kogan (2001) found that halving inventories could reduce unit production costs by 20 percent. Note that the value of reliability may be captured in $t(\cdot)$. It is because air passengers at a more congested airport (larger t) will be more vulnerable to uncertainty, as congestion usually increases at a very fast rate when the traffic volume is approaching to the capacity limit.

solution, we need to further assume that $b_p(\theta) > b_o(\theta)$ holds for high θ, and $b_p(\theta) - H(\theta)t(n_p) < b_o(\theta)$ holds for low θ in equilibrium, respectively.[24] In the following analysis, we will focus on the interior solution case, and thus the above conditions are assumed.

Given peak and off-peak airfares, f_p and f_o, respectively, consumers maximize their utility in (2.1) by choosing x and making a travel decision that is subject to the budget constraint $x + f_h \leq Y$, where f_h is the airfare in period $h \in (o, p)$ the consumer chooses, or $f_h = 0$ for those choosing not to travel. Y denotes consumers' income and is assumed to be uniform across consumers. Solving the consumer maximization problem and by the definition of $\underline{\theta}$ and θ^*, we have:

$$f_o = b_o(\underline{\theta}), \tag{2.2}$$

$$f_p - f_o = b_p(\theta^*) - H(\theta^*)t(n_p) - b_o(\theta^*). \tag{2.3}$$

Condition (2.2) suggests that in equilibrium, the off-peak fare is equal to the off-peak travel benefit for consumer $\underline{\theta}$, while condition (2.3) implies that the premium of the peak fare over the off-peak is equal to the difference between the peak and off-peak travel benefits for consumer θ^*.

[24] Note that the latter assumption is different from that in Brueckner (2002, 2005), in which passengers are not vertically differentiated (i.e., $b_p(\theta) < b_o(\theta)$ for low-θ passengers). More discussion about the condition for interior solutions in the vertically-differentiated case can be found in Basso and Zhang (2008).

2.4 Carriers' Competition

To solve the subgame perfect equilibrium, in this section we analyze the carriers' Cournot competition among N carriers in the second stage.[25] We denote n_p^k and n_o^k as the number of flights chosen by airline k in the peak and off-peak periods, respectively.[26] Thus, we have the number of total fights in period h, $n_h = \sum_k n_h^k$ for $h \in (o, p)$. These airlines are assumed to be symmetric such that they have a zero fixed cost, a constant operating cost per flight, c, and an additional operating cost per flight due to peak-period congestion, $g(n_p)$, with $g'(\cdot) > 0$. As in Brueckner (2002, 2005), all flights are assumed to use identical aircraft with fixed seat capacity, s, and a 100 percent load factor is assumed.

Given airport charges, p_p and p_o, in the peak and off-peak periods, respectively, each airline, k, chooses quantities to maximize its own profit:

$$\underset{n_p^k, n_o^k}{Max}\, \pi^k(n_p^k, n_o^k; n_p^{-k}, n_o^{-k}) = f_o\, s n_o^k + f_p\, s n_p^k - (n_o^k + n_p^k)c - n_p^k g(n_p) - p_o n_o^k - p_p n_p^k. \quad (2.4)$$

[25] Brander and Zhang (1990), and Oum, et al. (1993), for example, provided some empirical backing for Cournot behavior in the airline market.

[26] In the analysis, we consider the numbers of peak and off-peak flights, n_p and n_o, as real-valued. Note that they must be integers in practice. The integer constraint is ignored in this chapter as other studies in the literature (e.g., Brueckner, 2002, 2005; Pels and Verhoef, 2004; Zhang and Zhang, 2006).

Aggregating all flights by all the carriers, we have $\Phi - \theta^* = s\sum n_p^k$ and $\underline{\theta} = \Phi - s\sum(n_p^k + n_o^k)$. By conditions (2.2) and (2.3), the peak and off-peak fares can be rewritten as:

$$f_o = b_o[\Phi - s\sum(n_p^k + n_o^k)], \tag{2.5}$$

$$f_p = b_p(\Phi - s\sum n_p^k) - H(\Phi - s\sum n_p^k)t(n_p) - b_o(\Phi - s\sum n_p^k) + b_o[\Phi - s\sum(n_p^k + n_o^k)]. \tag{2.6}$$

By substituting (2.5) and (2.6) into (2.4), and solving the maximization problem, the Cournot equilibrium is characterized by the first-order conditions (assuming the second-order conditions),

$$\Omega_o = \frac{\partial \pi^k}{\partial n_o^k} = b_o(\underline{\theta}) - b_o^{'}(\underline{\theta})(n_o^k + n_p^k)s - \frac{p_o}{s} - \frac{c}{s} = 0, \qquad \forall k, \tag{2.7}$$

$$\begin{aligned}\Omega_p = \frac{\partial \pi^k}{\partial n_p^k} &= [b_p(\theta^*) - H(\theta^*)t(n_p) - b_o(\theta^*)] \\ &- [n_p^k H(\theta^*)t^{'}(n_p) - (n_p^k / n_p)(\Phi - \theta^*)H^{'}(\theta^*)t(n_p)] - \frac{1}{s}[g(n_p) + n_p^k g^{'}(n_p)] \\ &- (n_p^k / n_p)(\Phi - \theta^*)[b_p^{'}(\theta^*) - b_o^{'}(\theta^*)] - \frac{(p_p - p_o)}{s} = 0, \qquad \forall k.\end{aligned} \tag{2.8}$$

We further assume that a carrier's marginal profit declines when another carrier's output rises:

$$\frac{\partial^2 \pi^k}{\partial n_h^k \partial n_h^j} < 0, \qquad j \neq k \text{ and } h \in (o, p). \tag{2.9}$$

Conditions (2.9) imply the carriers' outputs are "strategic substitutes" (Bulow, et al., 1985) and ensure that various comparative static properties are "well defined" (see Dixit, 1986).

By conditions (2.7) and (2.8), imposing symmetry and adding up, we obtain:

$$b_o(\underline{\theta}) = \frac{c}{s} + \frac{p_o}{s} + \frac{(\Phi - \underline{\theta})}{N} b_o^{'}(\underline{\theta}), \tag{2.10}$$

$$\begin{aligned} & b_p(\theta^*) - H(\theta^*)t(n_p) - b_o(\theta^*) \\ & \quad = \frac{1}{s} g(n_p) + \frac{(p_p - p_o)}{s} + \frac{(\Phi - \theta^*)}{N}[\frac{1}{s} H(\theta^*)t^{'}(n_p) - H^{'}(\theta^*)t(n_p)] \\ & \quad + \frac{1}{N}\frac{n_p}{s} g^{'}(n_p) + \frac{(\Phi - \theta^*)}{N}[b_p^{'}(\theta^*) - b_o^{'}(\theta^*)]. \end{aligned} \tag{2.11}$$

By (2.2), the LHS of (2.10) is the off-peak airfare. The first two terms on the RHS represent a carrier's average cost per seat – the sum of the operating cost and airport charge *per seat* – in the off-peak period. The last term is a mark-up owing to market power in the Cournot oligopoly. As N increases, the market power mark-up diminishes.

By (2.3), the LHS of (2.11) is the premium in the peak fare over the off-peak. The first two terms on the RHS represent a carrier's additional average cost per seat in the peak period over the off-peak, which is the sum of congestion costs incurred by the

carrier and the difference between the peak and off-peak airport charges *per seat*. The third term is the marginal change in delay costs incurred by the *marginal* peak traveler with $\theta = \theta^*$ due to a marginal increase in the number of peak travelers. On the other hand, it also implies a marginal decrease in the peak fare charged by the airlines due to the change in congestion delays (see equation (2.3)). As the decrease in the fare will affect the revenue from all peak travelers, the term is multiplied by $(\Phi - \theta^*)$. The term is also multiplied by *1/N*, which suggests that each carrier only takes its own revenue change due to the decrease in the peak fare into account. An important catch here is that the terms $H(\theta)$ for $\theta \in (\theta^*, \Phi]$ do not appear in equation (2.11), which suggests that the marginal delay costs for those passengers with $\theta \in (\theta^*, \Phi]$ are not taken into account by the pricing of the air carriers. On the other hand, note that the difference of the variable passenger time cost case considered in this Chapter with the constant time cost case is that in our case, the marginal delay cost imposed on travelers increases in θ. Thus the marginal delay costs for other peak travelers with $\theta \in (\theta^*, \Phi]$ will be higher than that for the marginal traveler. The fourth term suggests that each carrier only charges for the congestion affecting its own operating cost in the peak period. For a monopoly carrier, $N = 1$, the marginal congestion cost on its flights is fully internalized (charged). On the other extreme, for atomistic (perfectly competitive) carriers, $N \to \infty$, the congestion cost is totally ignored by each carrier. For $1 < N < \infty$, partial internalization is obtained. Again, the fifth term is a mark-up owing to market power in the Cournot oligopoly.

The comparative static results concerning the carriers' aggregate output with respect to p_p and p_o are now derived. Totally differentiating (2.7) and (2.8) with respect to airport charges p_p and p_o, and by the second-order conditions and (2.9), we obtain:

$$\frac{\partial n_p}{\partial p_p} < 0, \frac{\partial n_o}{\partial p_p} > 0, \tag{2.12}$$

$$\frac{\partial n_p}{\partial p_o} > 0, \frac{\partial n_o}{\partial p_o} < 0. \tag{2.13}$$

Inequalities (2.12) and (2.13) suggest that the number of flights in a period is decreasing in airport charge of the period, while it increases in that of the other period.[27] The comparative static results will be used in the welfare analysis in Section 2.6.

2.5 Airport Pricing

In the airport pricing stage (stage 1) the airport chooses peak and off-peak charges, taking the subsequent carrier and consumer behavior into account to maximize the social welfare – the sum of airport and airlines profits, and passengers' travel benefits.[28] To obtain the optimal charges, we first derive the optimal allocation of passenger, and then

[27] Similar results are found in Basso and Zhang (2008).

[28] We assume that the consumer surplus is the sum of all passengers' net benefits. Note that some public policies may also aim to re-distribute the surplus among consumers. For example, a government may place more weight on poor people than their rich counterparts in its policy-making decision. To take this into account, we may need to consider a more general welfare function, which is usually used in the public economics literature (e.g., Cooter and Helpman, 1974).

discuss what airport charges can induce such optimal allocation. Thus the airport solves the following maximization problem:[29]

$$\underset{\underline{\theta},\theta^*}{Max}\, SW = \int_{\underline{\theta}}^{\theta^*} b_o(\theta)d\theta + \int_{\theta^*}^{\Phi}[b_p(\theta) - H(\theta)t(n_p)]d\theta - (n_o + n_p)c - n_p g(n_p). \tag{2.14}$$

The first-order conditions yield:

$$b_o(\underline{\theta}) - c/s = 0, \tag{2.15}$$

$$[b_p(\theta^*) - H(\theta^*)t(n_p) - b_o(\theta^*)] - \frac{1}{s}t'(n_p)\int_{\theta^*}^{\Phi} H(\theta)d\theta - \frac{1}{s}[g(n_p) + n_p g'(n_p)] = 0. \tag{2.16}$$

Equation (2.15) suggests that the travel benefit of traveler $\underline{\theta}$ in social optimum is equal to the flight operating cost *per seat*. This result is equivalent to the one obtained by Brueckner (2002), which is because in the optimal allocation between travelers and non-travelers, we only need to consider the operating cost and off-peak travel benefits, which do not involve the peak benefit function upon which variable passenger time costs are considered – the major difference between the model of this Chapter and that in Brueckner.

The terms in the first bracket on the LHS of (2.16) represent the difference of the marginal traveler's travel benefits between the peak and off-peak periods. The second

[29] For simplicity, the airport's operating, maintenance, administration and other costs are normalized to zero.

term captures the congestion effect due to the marginal traveler: to serve this passenger, the number of peak flights should be increased by *1/s*. This term thus represents the total additional time costs imposed on the peak passengers when shifting the marginal traveler to the peak period. The last term is the additional carriers' operating cost of shifting the traveler from the off-peak period to the peak. Thus, equation (2.16) suggests that the marginal costs and benefits of switching the marginal traveler from the off-peak period to the peak are equal in the social optimum.

Next, we turn to the discussion of the optimal airport pricing. As noted above, the optimal airport charges are set such that the optimal allocation of passengers can be achieved. Thus, conditions (2.15) and (2.16) are satisfied. By comparing conditions (2.10) and (2.11), and conditions (2.15) and (2.16), the optimal airport pricing can be obtained as follow:

$$p_o = -\frac{s(\Phi - \underline{\theta})}{N} b_o^{'}(\underline{\theta}), \tag{2.17}$$

$$\begin{aligned} p_p - p_o = {} & t^{'}(n_p)\int_{\theta^*}^{\Phi} H(\theta)d\theta - \frac{(\Phi - \theta^*)}{N}[H(\theta^*)t^{'}(n_p) - sH^{'}(\theta^*)t(n_p)] \\ & + \frac{N-1}{N} n_p g^{'}(n_p) - \frac{s(\Phi - \theta^*)}{N}[b_p^{'}(\theta^*) - b_o^{'}(\theta^*)]. \end{aligned} \tag{2.18}$$

Equation (2.17) suggests that there is a subsidy for the off-peak fare to correct the "market distortion" due to carriers' market power. Specifically, for atomistic carriers, there is no market distortion at the carrier level and so the subsidy is unnecessary in the

social optimal airport charge. Note that although a subsidy to off-peak flights is welfare-improving, it may still be politically infeasible for the government to provide subsidies to the industry.[30] Furthermore, this may leave the airport in a position of financial deficit, if the optimal airport charge is negative.[31]

Equation (2.18) suggests the premium paid by carriers using the airport in the peak period. The third term on the RHS of (2.18) is included to induce the carriers to fully internalize congestion costs imposed on all flights – as discussed in Section 2.3, each carrier only charges for the congestion affecting its own flights but not those operated by others. This term is first identified in Brueckner (2002). However, as also discussed in the literature (e.g., Pels and Verhoef, 2004; Zhang and Zhang, 2006; Basso and Zhang, 2008), charging only the positive congestion toll though would not be the first best because, in this case, subsidies would be needed to correct for downstream carriers' market power. Thus, the last term in (2.18), which is positive by the single crossing condition, is included so as to correct the "market distortion" due to carriers' market power. As a result, the peak premium in (2.18) may be positive or negative, depending on the magnitude of the congestion-toll component and the last term in (2.18). More specifically, for small N, the premium will be negative, as the market-power effect

[30] For instance, Section VI. B, Paragraph 12 of the Use of Airport Revenue Policy states: "DOT policy forbids direct subsidy of air carrier operations. Direct subsidies are considered to be payments of airport funds to carriers for air service."

[31] In practice, the airport charge is likely to include a component to cover operating, maintenance and administration costs (recall that to simplify the analysis, those costs are normalized to zero as mentioned in footnote 27), and is therefore the optimal airport charge is likely to be positive. On the other hand, to make the airport financially self-supporting, a budget constraint may be imposed in the analysis. As shown in Zhang and Zhang (1997), the airport charge will then have a mark-up over the social marginal cost, which depends on the elasticity of demand.

dominates; for large N, the premium will be positive, as the congestion-externality effect dominates. This result is also discussed in Basso and Zhang (2008).

In (2.18), our major focus is the first and second terms on the RHS, which are the difference between marginal congestion costs (incurred by peak travelers) and the congestion costs internalized at the carrier level. If the first term is larger than the second, it implies that the carriers do not fully internalize the congestion costs. By comparing the two terms, we have the following proposition:

Proposition 2.1 The congestion costs will not be fully internalized at the carrier level under different market structures, and thus the congestion pricing is always necessary.

Proof:

$$t'(n_p)\int_{\theta^*}^{\Phi} H(\theta)d\theta - \frac{(\Phi-\theta^*)}{N}[H(\theta^*)t'(n_p) - sH'(\theta^*)t(n_p)]$$

$$> t'(n_p)[\int_{\theta^*}^{\Phi} H(\theta)d\theta - (\Phi-\theta^*)H(\theta^*)] + \frac{s(\Phi-\theta^*)}{N}H'(\theta^*)t(n_p)$$

$$> \frac{s(\Phi-\theta^*)}{N}H'(\theta^*)t(n_p) > 0 \qquad Q.E.D.$$

Proposition 2.1 suggests that there is no market structure that induces full internalization of congestion at the carrier level. This result is different from the literature (e.g., Brueckner, 2002), which suggests that a congestion toll is unnecessary for an airport served by a monopoly carrier, in the case of congestion being fully internalized by the

carrier. The intuition is as follows: To maximize its profit, the monopolist will only consider the marginal traveler's tradeoff in choosing the peak and off-peak periods. Different from a social welfare-maximizing airport, the monopoly carrier will have no incentive to take into account the costs and benefits of inframarginal travelers with $\theta > \theta^*$, who have higher time valuations than the marginal counterparts in our case of variable passenger time costs.[32] Thus the congestion cost is not fully internalized by the monopoly carrier. Note that the first three terms on the RHS of (2.18) can be considered as a "pure" congestion toll, which is positive, and are included so as to induce full internalization of congestion at the carrier level.

The result in Proposition 2.1 is also discussed in Brueckner (2002) over the effects of non-separability between time cost and travel benefit. In this regard, the proposition can be considered as providing formal proof of Brueckner's result with specific functional forms. Moreover, this can provide a better framework to examine the underlying factors affecting the extent of congestion internalization and peak allocation. As mentioned, the uninternalized congestion costs can be represented by the difference between the first two terms on the RHS of (2.18). The fact that the extent of congestion internalization decreases in N has already been discussed in the literature. Thus we will now focus on other factors in the following discussion by considering the monopoly case. In such a case, the uninternalized congestion costs (i.e., the difference between the first two terms) are:

[32] It is noted that this bias will be weaker in practice as airlines can usually differentiate service within flights, for example, according to business class and different categories of economy class. To an extreme, the bias will vanish if perfect price discrimination is feasible.

$$t'(n_p)[\int_{\theta^*}^{\Phi}[H(\theta)-H(\theta^*)]d\theta+(\Phi-\theta^*)H'(\theta^*)st(n_p). \tag{2.19}$$

The first term of (2.19) is positive and increases in the difference between $H(\theta^*)$ and $H(\theta)$ for $\theta \in (\theta^*, \Phi]$. Thus, if the difference between the time valuations of inframarginal and marginal passengers increases, the uninternalized congestion cost increases. The second term of (2.19) is also positive and suggests that the larger the $H'(\theta^*)$, the greater the amount of uninternalized congestion.

2.6 Welfare Analysis

As argued in Section 2.5, the imposition of congestion pricing induces the full internalization of congestion delays at the carrier level. Thus social welfare will be improved. Yet the welfare implications on different groups are uncertain. The major concern here is how the revenue from congestion pricing is re-distributed. In this section, we will consider the case that the revenue will be regarded as a source of airport revenue and will not be re-distributed to air carriers and travelers. This is because as noted in the introduction, an effective re-distributive mechanism is usually not in place. At the same time, as discussed above, a direct subsidy to air carriers and travelers is less observed in practice. As a result, we will focus on the pure congestion toll in the peak period (the first three terms on the RHS of (2.18)) as in Brueckner (2002, 2005). The peak charge is thus

positive as discussed above. In the following, we will examine the welfare implications of the imposition of the congestion toll on the airport, air carriers and travelers.

First, the airport will be better-off as the airport profit is increased from zero to positive after the imposition of the congestion pricing. Second, we consider the change in consumer utility. After the imposition of the positive congestion toll, some passengers previously traveling in the peak period will choose to travel in the off-peak period (by inequalities (2.13)). For θ^* and $\tilde{\theta}^*$ denoting the marginal passenger between the peak and off-peak periods before and after the imposition of the congestion pricing, respectively, we have $\tilde{\theta}^* > \theta^*$. Thus we compare three groups of travelers who may be affected by the imposition of the congestion toll: (i) travelers θ_1 who travel in the peak period before and after the congestion toll is imposed (i.e., $\tilde{\theta}^* < \theta_1 \leq \Phi$); (ii) travelers θ_2 who change from traveling in the peak period to the off-peak (i.e., $\theta^* \leq \theta_2 \leq \tilde{\theta}^*$); (iii) travelers θ_3 who travel in the off-peak period before and after the congestion toll is imposed (i.e., $\underline{\theta} \leq \theta_3 < \theta^*$).

Equation (2.10) suggests that the off-peak fare will not change after the imposition of the congestion toll. On the other hand, the off-peak travel benefits remain unchanged, even though the number of off-peak flights increases. Recall that the off-peak travel benefit is independent of n_o. As a result, θ_3's utility will not change with the

imposition of the congestion toll. Second, we consider the change in θ_2 's utility. Comparing θ_2 's utility before and after the imposition of the congestion toll, we have:

$$U_2^B = b_p(\theta_2) - H(\theta_2)t(n_p^B) - f_p^B > b_o(\theta_2) - f_o^B = b_o(\theta_2) - f_o^A = U_2^A, \tag{2.20}$$

where the superscripts "*B*" and "*A*" represent variables in equilibrium before and after the imposition of the congestion toll, respectively. In words, (2.20) suggests that before the imposition of the congestion toll, as θ_2 choose to travel in the peak period instead of the off-peak, their utility for traveling in the peak period is higher than that in the off-peak. Furthermore, after the imposition of the congestion toll, their off-peak travel benefits and the off-peak fare remains changed. Thus we can conclude that θ_2 will be worse-off after the imposition of the congestion toll.

Furthermore, by (2.20) and the fact that θ_2 prefer to travel in the off-peak period than the peak after the imposition of the congestion toll, we have:

$$H(\theta_2)[t(n_p^B) - t(n_p^A)] + (f_p^B - f_p^A) < 0. \tag{2.21}$$

As there are fewer passengers in the peak period after the imposition of the congestion pricing (i.e., $n_p^B > n_p^A$), we have $t(n_p^B) > t(n_p^A)$. Hence, the first term on the LHS of (2.21) is positive and represents the time cost saving for θ_2 after the imposition of the congestion toll, if they traveled in the peak period. On the other hand, as there are fewer

people who will travel in the peak period, the peak fare will increase after the imposition of the congestion toll as $\partial f_p / \partial \theta^* > 0$ (by the single crossing condition and (2.3)). Thus the second term on the LHS of (2.21) is negative. Inequality (2.21) suggests that after the imposition of the congestion pricing, the time cost saving for θ_2 is not enough to compensate for the increase in peak fare if they traveled in the peak period.

For θ_1, they will travel in the peak period before and after the imposition of the congestion toll. The difference in θ_1's utility before and after the imposition of the congestion toll is:

$$U_1^A - U_1^B = H(\theta_1)[t(n_p^B) - t(n_p^A)] + (f_p^B - f_p^A). \tag{2.22}$$

Again, the first term on the RHS of (2.22) is positive and represents the time cost saving for θ_1 after the imposition of the congestion toll. The second term on the RHS of (2.22) is negative, as argued, and represents the difference of the peak fares before and after the imposition of the congestion pricing. It shows that the change in θ_1's utility can be positive or negative, which depends on their time valuation, as it determines the magnitude of the time cost saving after the imposition of the congestion toll. For travelers who have high time valuation, their change in utility can be positive due to the substantial benefits from the time cost saving. However, in the Brueckner (2002) case, where $H(\theta_1) = H(\theta_2)$, (2.21) and (2.22) imply $U_1^A - U_1^B < 0$, which suggests that θ_1 will always be worse-off after the imposition of the congestion toll. In other words, if we

assume that the value of time for all passengers is the same (i.e., $H(\cdot)$ is constant), no passengers will be better-off if the positive congestion toll is imposed. The above discussion leads to the following proposition:

Proposition 2.2 After the imposition of the congestion pricing, (i) the original off-peak travelers will have no change in utility, (ii) travelers shifting from the peak period to the off-peak will be worse-off; and (iii) travelers remaining in the peak period may be better-off or worse-off depending on their time valuation.

It is uncertain whether carriers are better-off or worse-off after the imposition of the congestion pricing. On one hand, the imposition of the congestion toll will increase the average cost per flight at a given congestion level. On the other hand, the imposition of the congestion toll not only decreases the number of peak passengers of a carrier, but also its rivals. This will lower the congestion cost experienced by the carrier. Thus, the imposition of the congestion toll may solve the coordination problem among air carriers facing the congestion externality. As a result, air carriers may be better-off or worse-off after the imposition of the congestion toll. The impact of the congestion toll on carriers will be further examined in the following section. To summarize the discussion in this section, it is noted that after the imposition of congestion pricing, airlines and some passengers who have low time valuation may be worse-off in the case of variable passenger time costs. This non-Pareto-improving result may explain why the airport congestion pricing appears unpopular in practice.

2.7 Numerical Simulation

This section conducts a simulation analysis to further illustrate the analytical results of our model. Here we consider symmetric airlines, and use specific functional forms and related parameters described as follows. First, we consider the case that the (gross) benefit functions, $t(n_p)$ and $H(\theta)$ are linear. Thus,

$$b_o(\theta) = \alpha_o + \gamma_o \theta,$$

$$B_p(\theta, n_p) = b_p(\theta) - H(\theta)t(n_p) = \alpha_p + \gamma_p \theta - (v + \phi\theta)(tn_p),$$

where $\alpha_o, \gamma_o, \alpha_p, \gamma_p, t, v, \phi \geq 0$, $\alpha_p \geq \alpha_o$ and $\gamma_p \geq \gamma_o$. Moreover, to illustrate the importance of the variable passenger time cost in the airport congestion pricing analysis, we consider different slopes of $H(\theta)$, namely $\phi = 0$, 0.0001, 0.0002, 0.0003 and 0.0004. For comparison among different ϕs, we assume $v = 1 - \phi(\Phi/2)$, which suggests that for different ϕs, the time valuations (i.e., $H(\theta)$) of the median consumers with $\theta = \Phi/2$ are all equal to 1. The $H(\theta)$ for different ϕs is shown in Figure 2.2.

[Figure 2.2 here]

Furthermore, we consider carriers facing constant marginal operating cost function and constant marginal delay cost. Thus, the total cost of carrier *k* is:

$$(n_o^k + n_p^k)c + n_p^k \varphi n_p,$$

where $c, \varphi \geq 0$.

Table 2.1 shows the parameterization for our simulation exercise. Further we assume that $\Phi = 5{,}000$, which implies that the market size is 5,000. Note that it is not the purpose of this Chapter to describe a real-life aviation network accurately. Thus our choice of parameters does not correspond to any actual cases.[33] Yet we calculated the equilibrium price elasticities of demand and compared them with estimates from the literature to confirm the plausibility of the parameterization. In particular, we found that the calculated price elasticities in equilibrium in the simulation results are between -0.52 and -1.29, which are consistent with the existing empirical literature.[34] At last, we will also conduct a sensitivity analysis for some crucial parameters to check the robustness of our results.

[Table 2.1 here]

Table 2.2 presents the simulation results for passenger allocation under different market structures and ϕs. The results are consistent with our analytical results. First, $\underline{\theta}$ (the allocation between travelers and non-travelers), as expected, remains the same for

[33] Note that, in practice, a take-off-and-landing (TOL) operation takes about two minutes. The cost in travel time delay imposed by an additional TOL operation on a typical traveler might be of the order of dollars. Thus, for setting $t = 100$, the units of the measurement would be cents.

[34] The overall mean price elasticity of the air passenger market estimated by Brons, et al. (2002) is -1.146.

each market structure among cases with different ϕ s. As off-peak passengers do not experience any congestion in our model, the allocation will only depend on the operating costs, off-peak travel benefits and market power. Second, the socially optimal θ^* (the allocation between peak and off-peak travelers) increases in ϕ, which suggests that there are fewer people who will travel in the peak period when the value of time is more sensitive to the change in θ (i.e., larger $H^{'}(\cdot)$). This is because the total additional time cost to peak passengers when shifting the marginal traveler to the peak period is larger, when $H^{'}(\cdot)$ is larger, as discussed in Section 2.5. Third, the θ^* s are larger under monopoly and smaller under an oligopoly (and perfect competition) than the social optimum, which indicates underuse of the peak period under monopoly, while it is overused under an oligopoly (and perfect competition).

[Table 2.2 here]

Figure 2.3 plots the congestion toll (i.e., the uninternalized congestion costs) against the market structure variable N, the number of airlines at the airport. This is consistent with our analytical results. First, under monopoly while no congestion toll is required for $\phi = 0$ (as discussed in Brueckner, 2002), a positive congestion toll should be imposed in the variable passenger time cost cases. This result is consistent with our Proposition 2.1, which suggests that, under monopoly, the carrier does not fully internalize the congestion cost as a social welfare-maximizing airport does. Thus, a positive congestion toll should be imposed in the variable passenger time cost cases. The

figure also supports the argument made by Brueckner (and this Chapter) that congestion is negatively correlated with airport concentration. Furthermore, the marginal increase in the optimal congestion toll falls as N rises. That is, although airport congestion is worsened owing to more serious externalities as the number of carriers rises, the marginal external effect falls. In particular, Figure 2.3 shows that the level of congestion toll remains almost unchanged for $N > 5$ suggesting, therefore, that for the practical purpose of congestion pricing we might just treat $N > 5$ as atomistic airlines and apply the results of the road-pricing literature – charging the toll as in perfect competition (see for example, Small, 1992 and Small and Verhoef, 2007). Moreover, as $H^{'}(\cdot)$ is larger, the congestion toll is larger under different market structures. Intuitively, the uninternalized congestion costs are larger, if $H^{'}(\cdot)$ is larger. Thus, a larger congestion toll is required to achieve the social optimum.

[Figure 2.3 here]

Table 2.3 shows the welfare change after the imposition of the congestion pricing. It suggests that the congestion toll is not necessarily welfare-improving. It is because the congestion toll decreases the number of peak travelers, which will further reinforce the distortion due to downstream carrier market-power markups. However, it is noted that they are all welfare-improving in different cases under perfect competition. It is because no marker-power distortion exists in such cases. As a result, the social welfare optimum can be achieved by imposing the congestion toll under perfect competition.

[Table 2.3 here]

In Section 2.6, we discuss the welfare change across different groups after the imposition of the congestion pricing. Table 2.4 presents the change in consumer surplus, producer surplus and airport revenue after the imposition of the congestion toll in the duopoly case. It shows that the consumer and producer surplus will decrease, which implies that both carriers and consumers will be worse-off after the imposition of the congestion toll. In particular, the consumer and producer surplus in the peak period will decrease, while those in the off-peak period increase. Finally, the airport will be better-off as it receives positive revenue from the congestion pricing. Nevertheless as both air carriers and most consumers are worse-off after the imposition of the congestion toll in our simulation results, the implementation of the pricing may then face strong opposition from them. This may explain why the congestion toll is still unpopular in reality, although it has strong support from economists based on efficiency arguments. On the other hand, Table 2.4 also shows that with a larger $H^{'}(\cdot)$, the larger the social gain from imposing the congestion toll. This is because the uninternalized congestion increases in ϕ. This result is consistent with Daniel (2001) which found that the benefits from congestion pricing increase with the degree in traveler heterogeneity in trip-timing preferences. The result implies that if we ignore the heterogeneity in passenger time cost, we may underestimate the welfare gains from the implementation of the congestion pricing. Similar results are also found for different market structures.

[Table 2.4 here]

Note that any carriers' welfare loss due to the congestion pricing may be easier to be offset by a re-distributive program between the airport and air carriers, or through vertical integration, while those solutions are more difficult to compensating the loss in the consumer market. Thus the welfare impact on consumers is particularly important in our welfare-redistributive analysis. Hence we also take a closer look at the consumer side by investigating the change in utility across different consumer groups after the imposition of the congestion toll. Table 2.5 presents the result with $\phi = 0.0004$. The result is consistent with Proposition 2.2. First, the travelers shifting from the peak to the off-peak period are worse-off. Second, the original off-peak travelers have no change in utility. Third, the travelers staying in the peak period may be better-off or worse-off depending on their time valuations. In particular, those having very high time valuations may be better-off as they have significant gains from the time cost saving after the imposition of the congestion toll. Nonetheless, in our simulation results, most travelers will not be better-off after the imposition of the congestion toll. Furthermore, for $\phi = 0$, 0.0001 and 0.0002, all passengers will not be better-off after the imposition of the congestion toll.

[Table 2.5 here]

In order to achieve the social optimum, as also suggested by Pels and Verhoef (2004), both congestion and market-power distortion should be taken into account to determine optimal airport pricing. Thus, we also examine the welfare change after the

imposition of the optimal airport pricing given in equations (2.18) and (2.19). Table 2.6 shows the calculated optimal airport pricing in our simulation exercise. Under monopoly, a negative optimal charge should be imposed (i.e., subsidy). This suggests that the airport is underused due to downstream carriers' market power. Under oligopoly competition, a positive optimal charge should be imposed, which suggests that the airport is overused due to the congestion externalities in carrier competition. Moreover, the optimal airport pricing is less than the congestion toll. This is because market-power mark-up alleviates the congestion at the airport. Finally, the congestion toll is equivalent to the optimal airport charge under perfect competition as market power vanishes. Table 2.7 shows the change in welfare if the optimal airport pricing is imposed. Under monopoly, the welfare is improved by almost 30 percent. The gains decrease initially as N rises and falls as N approaches infinity.

[Table 2.6 here]

[Table 2.7 here]

To check the robustness of the simulation results, we also conducted sensitivity analysis by varying the parameters in the time cost and travel benefit functions, while all of the remaining parameters are kept as above. In particular, we conduct the sensitivity analysis, in which (i) t is increased from 100 to 200; (ii) γ_p is increased from 1.2 to 1.5. We found that these changes yielded similar results as those in the benchmark case.

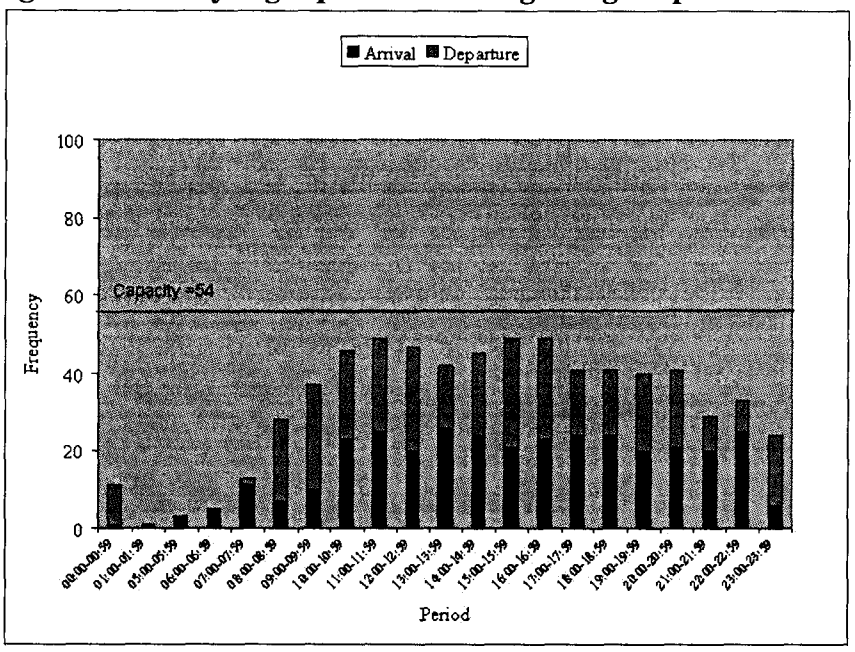
2.8 Concluding Remarks

This Chapter examined the airport congestion-pricing problem that arises when passengers are heterogeneous in time valuation, and investigated welfare-redistributive issues after the imposition of the congestion pricing scheme. In practice, the value of time is different among passengers, depending on their characteristics, such as income and travel purposes. Although this fact is supported by extensive empirical work, it is less discussed in the analytical work in the airport congestion literature. The present analysis is a first attempt to formally study the implications of variable passenger time costs in the airport congestion pricing analysis.

In the case of variable passenger time costs, we found that under different market structures at the downstream carrier level, from monopoly to perfect competition, congestion delays will not be fully internalized by the carriers. This is because each carrier is only concerned with the delays affecting its marginal peak traveler, thus the delay costs incurred by other passengers in the peak period are not fully reflected in the airline's pricing. Furthermore, our results suggested that although the implementation of the congestion pricing may be welfare-improving, it may not be Pareto-improving. As the airport industry is considered politically-sensitive, any airport policy, including a change in pricing, is difficult to implement without support from its users, namely airlines and passengers. This may explain why congestion pricing appears unpopular in practice.

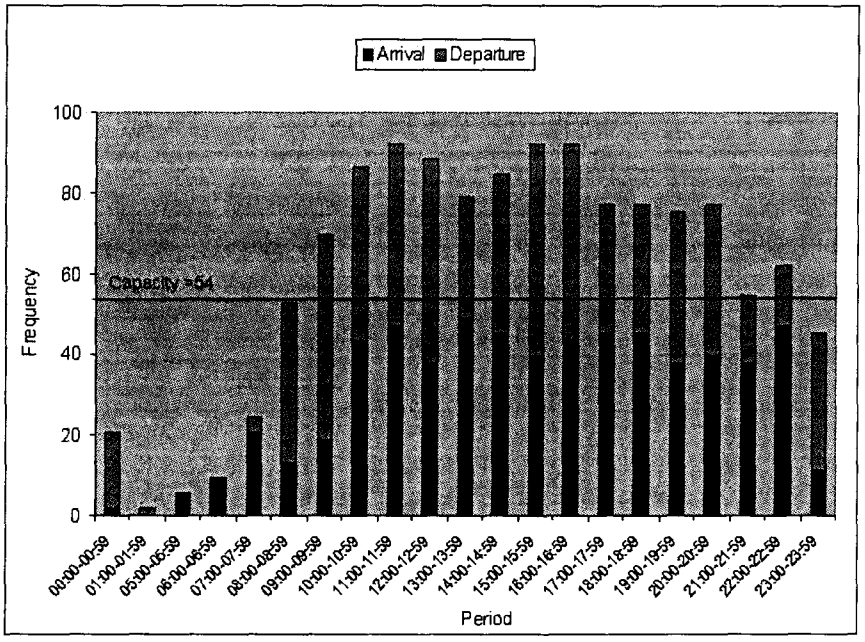
A valuable extension to this Chapter is to consider an asymmetric airline case. This Chapter found that passengers with high time valuation may be better-off when congestion pricing is implemented. In practice, there is market segmentation in the airline market: legacy carriers usually serve a higher portion of business travelers, while low-cost carriers serve a higher portion of leisure travelers (O'Connell and Williams, 2005). As a result, the implementation of congestion pricing may have different implications on legacy and low-cost carriers and their respective customers. In particular, the introduction of congestion pricing may induce an increase in demand for legacy carrier services due to the time cost savings in traveling. Considering the cost side, legacy carriers operate more flights during the congested period for hubbing, as an example, and thus operating cost savings due to a reduction in congestion in the period will be more significant for them than others. Using data from the Minneapolis-St. Paul Airport, Daniel (2001) found that Northwest enjoyed the greatest gain due to the implementation of congestion pricing as compared to regional counterparts. Thus, the decision of an airport to implement congestion pricing may also need to take into account the potential competitive issues in the downstream carrier market.

Figure 2.1a Daily flight pattern at Hong Kong Airport in 2007



Note: Flight schedule on 8th November (Thursday) 2007; the runway capacity is 54 flights/hour.
Source: Hong Kong International Airport website

Figure 2.1b (Estimated) Daily flight pattern at Hong Kong Airport in 2020



Note: Assuming (i) the average annual passenger growth is 5 percent (HKIA, "Master Plan 2020"); (ii) the flight increase in each period is proportional to the passenger growth; and (iii) the current capacity is kept unchanged.

Figure 2.2 Time valuation for different ϕs

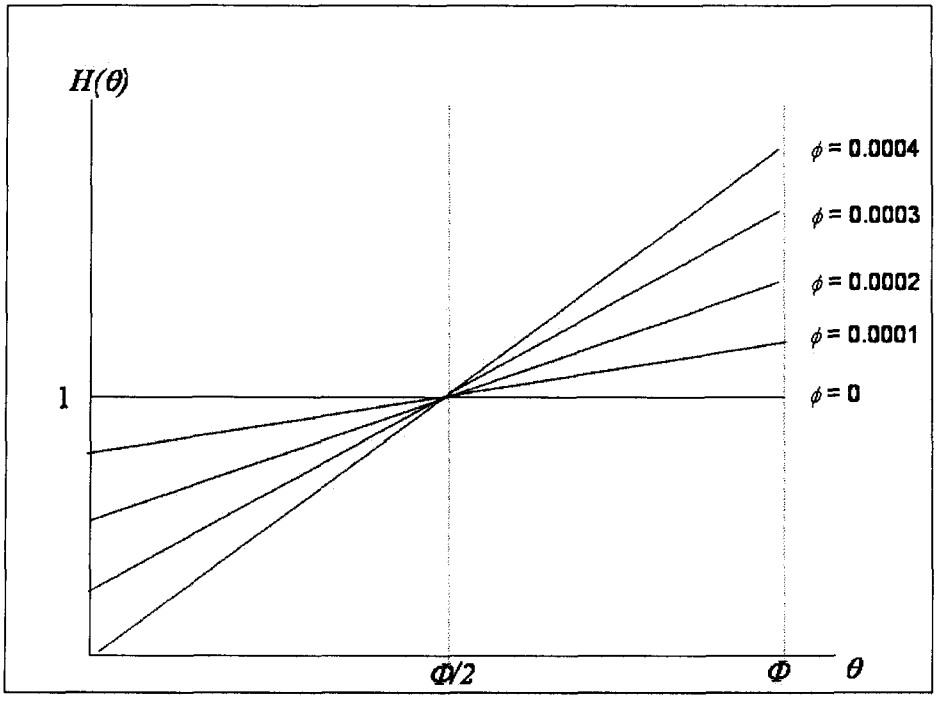


Figure 2.3 The congestion toll under different market structures

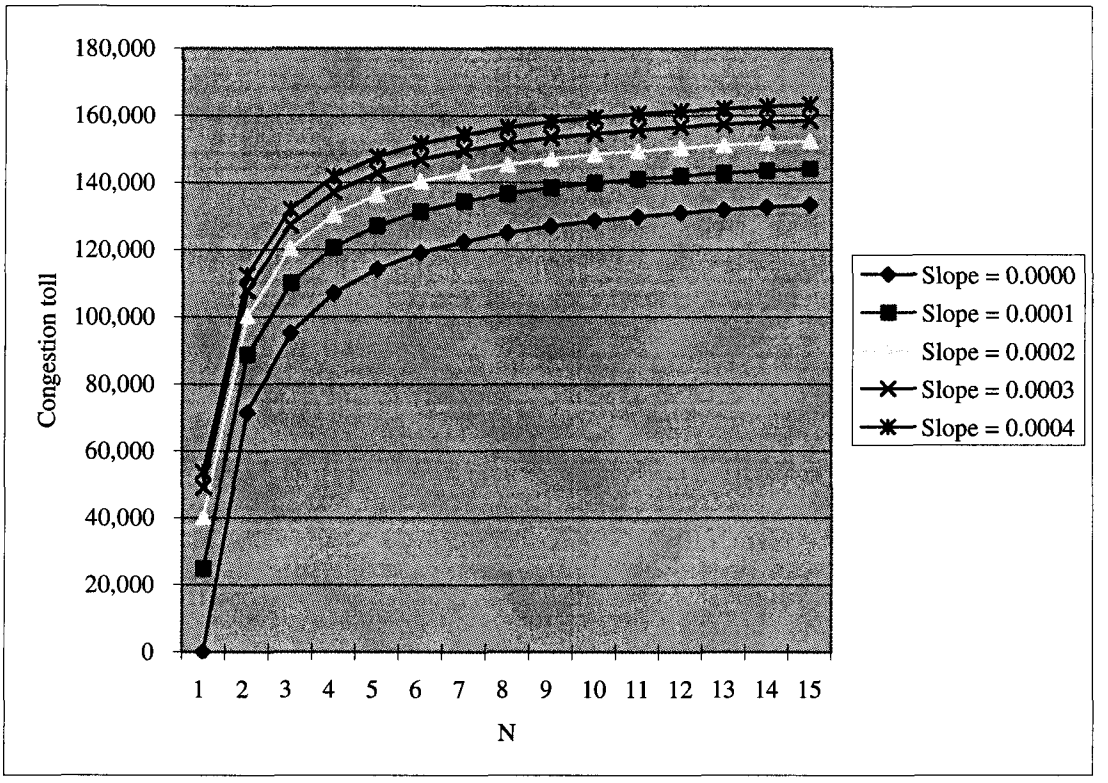


Table 2.1 Parameter values for the numerical simulation

Benefit functions		Airline cost function	
α_o	0	c	100,000
γ_o	0.8	s	200
α_p	0	φ	4,000
γ_p	1.2		
t	100		
v	1		

Table 2.2 The allocation of passengers

	Social Optimum		$N=1$		$N=2$		$N=3$		$N \to \infty$	
ϕ	$\underline{\theta}$	θ^*	$\underline{\theta}$	θ^*	$\underline{\theta}$	θ^*	$\underline{\theta}$	θ^*	$\underline{\theta}$	θ^*
0.0000	625	3,572	2,813	3,889	2,083	3,519	1,719	3,334	625	2,778
0.0001	625	3,713	2,813	3,943	2,083	3,566	1,719	3,376	625	2,817
0.0002	625	3,842	2,813	4,000	2,083	3,620	1,719	3,427	625	2,865
0.0003	625	3,956	2,813	4,060	2,083	3,682	1,719	3,487	625	2,925
0.0004	625	4,055	2,813	4,120	2,083	3,750	1,719	3,557	625	3,001

Table 2.3 Welfare with congestion toll (in millions)

	$N=1$		$N=2$		$N=3$		$N \to \infty$	
ϕ	Welfare	% change	Welfare	% change	Welfare	% change	Welfare	% change
0.0000	7.099	0.00%	8.203	-0.36%	8.588	0.26%	9.084	5.11%
0.0001	6.935	-0.40%	8.038	-0.16%	8.422	0.79%	8.917	6.79%
0.0002	6.801	-0.58%	7.900	0.16%	8.283	1.42%	8.776	8.59%
0.0003	6.692	-0.62%	7.787	0.53%	8.168	2.07%	8.658	10.33%
0.0004	6.602	-0.72%	7.692	0.86%	8.071	2.62%	8.559	11.78%

Table 2.4 Welfare change with congestion toll ($N = 2$)

ϕ	Off-peak Consumer Surplus	Peak Consumer Surplus	Off-peak Producer Surplus	Peak Producer Surplus	Airport revenue	Total
0.0000	331,387	-472,750	307,824	-630,038	434,469	-29,108
0.0001	457,051	-589,590	402,270	-765,922	483,075	-13,116
0.0002	555,998	-659,640	466,400	-841,338	491,274	12,694
0.0003	624,190	-693,470	501,380	-869,516	478,184	40,767
0.0004	664,400	-379,520	513,040	-1,187,245	455,062	65,737

Table 2.5 Passengers' utility change ($\phi = 0.0004$)

	Utility change for different θs										
θ	0	500	1,000	1,500	2,000	2,500	3,000	3,500	4,000	4,500	5,000
$N = 1$	0	0	0	0	0	0	0	0	0	-26	-12
$N = 2$	0	0	0	0	0	0	0	0	-50	-175	5
$N = 3$	0	0	0	0	0	0	0	0	-50	-26	34
$N \to \infty$	0	0	0	0	0	0	0	0	0	93	199

Table 2.6 Optimal airport pricings (in thousands)

	$N=1$		$N=2$		$N=3$		$N \to \infty$	
ϕ	Off-peak	Peak	Off-peak	Peak	Off-peak	Peak	Off-peak	Peak
0.0000	-700	-114	-350	14	-233	57	0	142
0.0001	-700	-78	-350	37	-233	75	0	152
0.0002	-700	-52	-350	53	-233	89	0	160
0.0003	-700	-34	-350	65	-233	99	0	166
0.0004	-700	-22	-350	74	-233	106	0	171

Table 2.7 Welfare change with optimal airport pricing

ϕ	$N=1$	$N=2$	$N=3$	$N \to \infty$
0.0000	27.97%	10.35%	6.05%	5.11%
0.0001	28.06%	10.76%	6.71%	6.79%
0.0002	28.28%	11.26%	7.45%	8.59%
0.0003	28.58%	11.77%	8.20%	10.33%
0.0004	28.90%	12.23%	8.82%	11.78%

2.9 References

AHCG (1999) "The Value of Travel Time on UK Roads", Accent Marketing & Research and Hague Consulting Group, The Hague, Netherlands.

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. (2008) "Airport Deregulation: Effects on Pricing and Capacity", *International Journal of Industrial Organization*, 26, 1015-1031.

Basso, L.J. and Zhang, A. (2007a) "An Interpretative Survey of Analytical Models of Airport Pricing", in: Lee, D. (Ed.), *Advances in Airline Economics*, Vol.2, Elsevier, 80-124.

Basso, L.J. and Zhang, A. (2007b) "Congestible Facility Rivalry in Vertical Structures", *Journal of Urban Economics*, 61(2), 218-237.

Basso, L.J. and Zhang, A. (2008) "Sequential Peak-load Pricing: The Case of Airports and Airlines", *Canadian Journal of Economics*, 41(3), 1087-1119.

Borins, S.F. (1978) "Pricing and Investment in a Transportation Network: The Case of Toronto Airport", *Canadian Journal of Economics*, 11, 680-700.

Brons, M., Pels, E., Nijkamp, P. and Rietveld, P. (2002) "Price Elasticities of Demand for Passenger Air Travel: A Meta-analysis", *Journal of Air Transport Management*, 8, 165-175.

Brander, J.A. and Zhang, A. (1990) "Market Conduct in the Airline Industry: An Empirical Investigation", *RAND Journal of Economics*, 21, 567–583.

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(7-8), 599-614.

Brueckner, J.K. (2008) "Price vs. Quantity-Based Approaches to Airport Congestion Management", *manuscript.*

Brueckner, J.K. and Van Dender, K. (2007) "Atomistic Congestion Tolls at Concentrated Airports? Seeking a Unified View in the Internalization Debate", *manuscript.*

Bulow, J.I., Geanakoplos, J.D. and Klemperer, P.D. (1985) "Multimarket Oligopoly: Strategic Substitutes and Complements", *Journal of Political Economy*, 93, 488-511.

Carlin, A. and Park, R.E. (1970) "Marginal Cost Pricing of Airport Runway Capacity", *American Economic Review*, 60, 310-319.

Ciliberto, F. and Williams, J.W. (2007) "Limited Access to Airport Facilities and Market Power in the Airline Industry", Working paper, Department of Economics, University of Virginia.

Cooter, R. and Helpman, E. (1974) "Optimal Income Taxation For Transfer Payments Under Different Social Welfare Criteria", *Quarterly Journal of Economics*, 88(4), 656-670.

Czerny, A. (2007) "Congestion Management under Uncertainty in a Two-Airport System", *manuscript.*

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.

Daniel, J.I. and Harback, K.T. (2008) "(When) Do Hub Airlines Internalize their Self-Imposed Congestion Delays?", *Journal of Urban Economics*, 63, 583-612.

De Palma, A. and Lindsey, R. (2004) "Congestion Pricing with Heterogeneous Travelers: A General-Equilibrium Welfare Analysis", *Networks and Spatial Economics*, 4, 135-160.

Dixit, A. (1986) "Comparative Static for Oligopoly", *International Economic Review*, 27, 107-122.

Dresner, M., Windle, R. and Yao, Y. (2002) "Airport Barriers to Entry in the U.S.", *Journal of Transport Economics and Policy*, 36, 389-405.

Fung, M., Law, J. and Ng, L. (2006) "Economic Contribution to Hong Kong of the Aviation Sector: A Value-Added Approach", *Chinese Economy*, 39(6), 19-38.

Gausch, J.L. and Kogan, J. (2001) "Inventories in Developing Countries: Levels and Determinants, a Red Flag on Competitiveness and Growth", World Bank Working Paper, No. 2552.

Grossman, G. and Krueger, A. (1993) "Environmental Impacts of a North American Free Trade Agreement", in: Garber, P. (Ed.), The Mexico-U.S. Free Trade Agreement, Cambridge, M.A.: MIT Press.

Grossman, G. and Krueger, A. (1995) "Economic Growth and the Environment", *Quarterly Journal of Economics*, 110(2), 353-377.

Gunn, H., Tuinenga, J.G., Cheung, Y.H.F. and Kleijn, H.J. (1999) "Value of Dutch Travel Time Savings in 1997", in Meersman, H., Van de Voorde, E. and Winkelmans, W. (eds.), *Proceedings of the 8th World Conference on Transport Research*, 3, Pergamon: Amsterdam.

Hess, S., Adler, T. and Polak, J.W. (2007) "Modelling Airport and Airline Choice Behaviour with the Use of Stated Preference Survey Data", *Transportation Research Part E*, 43, 221-233.

Hettige, H., Lucas, R.E.B. and Wheeler, D. (1992) "The Toxic Intensity of Industrial Production: Global Patterns, Trends and Trade Policy", *American Economic Review*, 82, 478-481.

ICAO (2007) "Performance Indicators," *ICAO Journal*, 62(2), 5.

Levine, M.E. (1969) "Landing Fees and the Airport Congestion Problem", *Journal of Law and Economics*, 12, 79-108.

Mackie, P.J., Wardman, M., Fowkes, A.S., Whelan, G., Nellthorp, J. and Bates, J. (2001) "Values of Travel Time Savings in the UK: A Report on the Evidence", ITS Working Paper No. 567, Institute of Transport Studies, University of Leeds.

Mayer, C. and Sinai, T. (2003) "Network Effects, Congestion Externalities, and Air Traffic Delays: Or Why Not All Delays are Evil", *American Economic Review*, 93, 1194-1215.

McConnell, K.E. (1997) "Income and the Demand for Environmental Quality", *Environment and Development Economics*, 2, 383-399.

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. and Winston, C. (1989) "Enhancing the Performance of the Deregulated Air Transportation System", *Brookings Papers on Economic Activity, Microeconomics*, Special Issue, 61-123.

Morrison, S.A. and Winston, C. (2007) "Another Look at Airport Congestion Pricing", *American Economic Review*, 97(5), 1970-1977.

MVA Consultancy, Institute of Transport Studies, University of Leeds, and Transport Studies Unit, University of Oxford (1987) "The Value of Travel Time Savings", *Policy Journals*, Newbury, Berks, U.K.

O'Connell, J.F. and Williams, G. (2005) "Passengers' Perceptions of Low Cost Airlines and Full Service Carriers: A Case Study Involving Ryanair, Aer Lingus, Air Asia and Malaysia Airlines", *Journal of Air Transport Management*, 11(4), 259-272.

Oum, T.H., 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.H. and Zhang, Y. (1990) "Congestion Tolls, Lumpy Investment and Cost Recovery", *Journal of Public Economics*, 43, 353-374.

Oxford Economic Forecasting (2006) "The Economic Contribution of the Aviation Industry in the UK", Oxford Economic Forecasting, Oxford.

Pels, E. and Verhoef, E.T. (2004) "The Economics of Airport Congestion Pricing", *Journal of Urban Economics*, 55, 257-277.

Reynolds-Feighan, A.J. and Button, K.J. (1999) "An Assessment of the Capacity and Congestion Levels at European airports - A New Highway Pricing and Investment Policy", *Journal of Air Transport Management*, 5(3), 113-134.

Santos, G. and Robin, M. (2007) "Delays at European Airports: Causes and Solutions", Paper presented at the 11th World Conference on Transport Research, Berkeley.

Selden, T.M. and Song, D. (1994) "Environmental Quality and Development: Is There a Kuznets Curve for Air Pollution Emissions?", *Journal of Environmental Economics and Management*, 27, 147-162.

Small, K.A., Winston, C. and Yan, J. (2005) "Uncovering the Distribution of Motorists' Preferences for Travel Time and Reliability", *Econometrica*, 73(4), 1367-1382.

Stiglitz, J.E. (1998) "The Private Uses of Public Interests: Incentives and Institutions", *Journal of Economic Perspectives*, 12(2), 3-22.

von Wartburg, M. and Waters II, W.G. (2004) "Congestion Externalities and the Value of Travel Time Savings", Chapter 2, Towards Estimating the Social and Environmental Costs of Transportation in Canada, A Research Report prepared for Transport Canada, 54-140.

Wardman, M. (2001) "A Review of British Evidence on Time and Service Quality Valuations", *Transportation Research Part E*, 37, 107-128.

World Bank (2003) "Market Power: Airports – Vertical Integration between Airports and Airlines", *Public Journal*, No. 250.

Zhang, A. and Chen, H. (2003) "Evolution of China's Air Transport Development and Policy towards International Liberalization", *Transportation Journal*, 42, 31-49.

Zhang, A. and Yuen, A. (2008) "Airport Policy and Performance in Mainland China and Hong Kong", in: Winston, C. and de Rus, G. (eds.), *Aviation Infrastructure Performance: A Study in Comparative Political Economy*, Washington, D.C.: Brookings Institution Press, 159-192.

Zhang, A. and Zhang, Y. (1997) "Concession Revenue and Optimal Airport Pricing", *Transportation Research Part E*, 33(4), 287-296.

Zhang, A. and Zhang, Y. (2006) "Airport Capacity and Congestion When Carriers have Market Power", *Journal of Urban Economics*, 60, 229-247.

Chapter 3 Effects of Gateway Congestion Pricing on Optimal Road Pricing and Hinterland[35]

3.1 Introduction

The gateway-hinterland intermodal transport system is very common. In this type of system, a gateway (seaport or airport) serves its hinterland: traffic destined for (or originated from) an interior region – the hinterland – passes through the gateway. In particular, goods from the rest of the world are first brought to a gateway, and then are trucked to the hinterland via the latter's road system. For example, cargos are shipped to the Port of New York (the gateway) and then are distributed to the hinterland mainly through the New Jersey highway system. Similarly, major gateways, such as Southern California, Antwerp, Rotterdam, Hamburg, Tokyo, Hong Kong, Shanghai and Singapore, serve their respective hinterlands by intermodal transport systems. As argued by Van Klink and Van den Berg (1998), those gateways are in a unique position to stimulate intermodal transport, and also use the intermodal systems to enlarge their hinterlands.[36]

An obvious characteristic of the intermodal system is that both the gateway and its hinterland's road system are prone to congestion. But while road congestion pricing has been implemented in several places and has been gaining momentum, gateway

[35] A version of this Chapter is published as: Yuen, A., Basso, L. and Zhang, A. (2008) "Effects of Gateway Congestion Pricing on Optimal Road Pricing and Hinterland", *Journal of Transport Economics and Policy*, 42(3), 495–526.

[36] A number of studies in the literature (for example, Notteboom, 1997; Kreukels and Wever, 1998) further argued that hinterland access is also important for the competitiveness of a gateway.

congestion pricing is relatively rare. For example, despite the call for its implementation by economists for many years (for example, Levine, 1969; Carlin and Park, 1970), airport congestion pricing 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.[37] However, it is expected that more gateways will adopt congestion pricing in the near future, for (at least) two reasons. First, the economic growth and globalization over the last several decades have significantly stimulated cargo movements, resulting in severe congestion at most of the major gateways in the world; economic loss due to such congestion is huge.[38] Thus there is an urgency to look for solutions, which include congestion pricing, to relieve gateway congestion.[39] Second, there has been an acceleration in worldwide private investment in container terminals during the 1990s: in 1991 the public sector handled 42 percent of container port throughput, while in 2003 its share fell to about 22 percent (Midoro, et al., 2005). Also, a number of major airports in Europe, Australia, New Zealand and Asia were recently privatized, or are in the process of being privatized. According to several authors (for example, Poole, 1990; Gillen, 1994; Vasigh and Haririan, 1996) private transport facilities such as airports are more likely to adopt congestion pricing for their services.

[37] In the case of seaports, it is noted that congestion pricing has been adopted at some ports such as the Ports of Los Angeles and Long Beach.

[38] For example, the twelve shipping lines, which accounted for about 70 percent of trans-Pacific trade, estimated that their costs in 2005 would rise 10 percent because of port congestion (Bloomberg News, December 4, 2005). In the case of airports, a recent study by the Oxford Economic Forecasting (2006) suggested that UK airport congestion cost (US)$3.2 billion in 2005; further, US DOT (2006) has estimated that aircraft delays cost passengers $9.4 billion.

[39] The literature on airport congestion pricing has witnessed a revival (for example, Morrison and Winston, 1990; Daniel, 1995, 2001; Brueckner, 2002, 2005; Pels and Verhoef, 2004; Zhang and Zhang, 2006; Basso, 2007; Basso and Zhang, 2007a; Brueckner and Van Dender, 2007), which may provide a better guidance for gateways to set an appropriate level of congestion tolls and help build political support.

Given the expectation that gateway congestion pricing will become more popular, it is important to investigate how such a change in gateway pricing will affect its hinterland. This is a practically relevant question because the traffic flows that are coming from gateways are certainly stressing their hinterlands' road system. For example, truck traffic in the province of British Columbia is expected to increase by 50 percent between now and 2021, generated primarily by the Vancouver International Airport, the Deltaport and Fraser Surrey Docks,[40] while according to US GAO (2003), from 1993 through 2001 truck traffic on urban highways in the U.S. increased more than twice as much as passenger traffic, meaning that freight traffic is contributing to worsening congestion at a faster rate than passenger traffic. This indeed translates into substantial social costs; for example Berechman (2007) finds that the additional highway traffic due to a 6.4 percent container throughput expansion at the Port of New York will induce annual costs that range from $0.66 billion to $1.62 billion, while California highways would need up to $30 billion over the next 20 years to maintain and handle the influx of cargo at coastal ports (Park, 2005).

In this Chapter we investigate the effect of congestion pricing implemented at a gateway on its hinterland's optimal road tolls, road congestion and social welfare. Thus, we attempt to add insights into the existing road pricing literature in which the gateway-hinterland pricing linkage is less explored. Table 3.1 lists the recent studies in the literature of road pricing and network interactions (also see Small, 1992 and Small and Verhoef, 2007 for literature reviews). Our main point is that, as the gateway and

[40] The Gateway Program website (www.th.gov.bc.ca/gateway/).

hinterland transport facilities are parts of an intermodal transport system, their pricing decisions, even if taken independently, affect each other's performance: thus, externality and strategic issues may arise in such pricing. In particular, an increase in gateway charge increases the costs incurred by cargo transshipped at the gateway to its hinterland, thereby reducing the cargo demand. This will then have implications for the hinterland's road usage and congestion level, and for the optimal road tolls.

[Table 3.1 here]

More specifically, we consider a transport system consisting of two congestible facilities, namely, a gateway and a road, each of which belongs to an independent region (the gateway region or the hinterland region). The gateway is served by its oligopoly carriers that carry both the "local" and "transit" traffic. The local traffic refers to cargo that is destined for the gateway region, whereas the transit traffic refers to cargo that is destined for the hinterland region. Thus, the transit traffic uses both the gateway and the road whilst the local traffic uses only the gateway. In addition to the transit cargo traffic, the road in the hinterland has its own local users (local commuters) as well. The road charges for congestion externalities, but it may or may not price discriminate between the two types of traffic. We compare the case in which the gateway implements congestion pricing, with the no-congestion-pricing base case in which the gateway's charge is just equal to its operating cost.

We examine how the demands derived from the equilibria of carrier and consumer markets – the cargo and local road user markets – respond to changes in facility charges. We find that, for a given gateway price, if the road price discriminates between the local and transit cargo traffic, a particular type of traffic decreases in its own toll while increasing in the toll on the other type of traffic.[41] With uniform pricing however, if the carriers' value of travel time savings is much larger than that of the local users, an increase in the (single) road toll may reduce the local traffic while increasing the transit cargo traffic. We also show that if the gateway maximizes the joint profit of itself and its oligopoly carriers, its charge will rise, with part of this increase owing to congestion pricing. In response to the increase in gateway charge, the hinterland will reduce its (optimal) road tolls independently of whether it is able to price discriminate between the local and transit cargo traffic. In addition, while the change in the road's congestion level is in general ambiguous, the hinterland's welfare will fall as a result of the increase in the gateway charge. Our analysis then suggests that a move by the gateway towards congestion pricing would negatively affect the hinterland. As a consequence, if the gateway and the road were to belong to the same jurisdiction, this negative impact would be taken into account in the optimal gateway congestion pricing.

This Chapter is related to the rapidly growing literature on second-best road pricing in a transport network (see Small and Verhoef, 2007 and Ubbels and Verhoef, 2008 for recent reviews). In terms of model structure and problem being examined, our

[41] If both tolls rise, then traffic levels may fall or rise, depending on the relative magnitudes of toll increases imposed on the two types of traffic. For example, one possible case is that a small increase in the transit toll accompanied by a large increase in the local toll would lead to an increase in transit traffic.

analysis is close to three papers in this literature, namely, De Borger, et al. (2007), De Borger, et al. (2008), and Ubbels and Verhoef (2008). Like this Chapter, these papers consider a serial network with multiple congestible facilities, where the facilities are linked by through-traffic and where the facility operators of the different regions strategically choose their pricing and capacity investment decisions. The major differences between this Chapter and the three papers are: (i) these studies have considered the facilities as service providers to final consumers and therefore users are assumed to be atomistic. But whilst roads may provide services directly to final consumers (drivers), gateways (seaports, airports) are input providers that reach final consumers (shippers) only through carriers and, therefore, these facilities are in an intermediate market and not in the final market. Hence, we allow the carriers to possess market power in the output market, meaning that their behavior is not atomistic.[42] (ii) De Borger, et al. (2007) consider a symmetric set-up, while our analysis is concerned with two asymmetric regions, gateway and hinterland, with different objectives and carrier market structures. (iii) Unlike De Borger, et al. (2008), the present Chapter considers endogenous road tolls at the hinterland and so the congestion at the hinterland has been taken into account in the examination of the impact of gateway pricing. And (iv), we consider both freight and passenger transportation, which induces a mix of facility users

[42] The U.S. Postal Service, FedEx Express and UPS accounted for 89 percent of the domestic air shipments in the U.S. in the first half of 2005 (Colography Group Inc., "Domestic Air Cargo Trends Report 2005"). In sea shipping, the top 25 container carriers controlled 84 percent of the world's slot capacity at the beginning of 2007 (Paul, 2007). Furthermore, many major ports/port terminals either are "dedicated" to one or two particular carriers, or are dominated by a few major carriers.

that introduces heterogeneity in time valuations. As will be seen below, this heterogeneity plays an important role in some of our results.[43]

The Chapter is organized as follows: Section 3.2 sets up the model and analyzes the consumer problem. Section 3.3 examines carrier behavior and derives important comparative static results that will be used in the subsequent analysis. Section 3.4 investigates the effects of gateway congestion pricing on the optimal road tolls, road congestion and the hinterland's social welfare. Section 3.5 contains concluding remarks.

3.2 The Model

Consider a gateway situated in region G, and an adjacent road situated in its hinterland, region H. N homogenous cargo carriers, belonging to region G, bring Y_G units of cargo into G via the gateway, while they use both the gateway and the road to bring X cargo units into H.[44] For example, Y_G and X may represent the number of containers (for example, twenty-foot equivalent unit, or TEU). Both the gateway and the road are congestible but, while the gateway is used only by the carriers, the road is also used by

[43] This Chapter is also related to Zhang (2007) who examines the impact of road pricing and capacity investment in a gateway-hinterland system. Similar to the above three studies, Zhang assumes atomistic carriers. While all of the four studies have examined capacity decisions, the capacities are taken as given in this Chapter so that we could focus on the gateway-hinterland interaction with non-atomistic carriers.

[44] In reality, the cargos processed at a gateway may be shipped out also by rail or inland waterways. The modal split differs greatly among gateways depending on the geographical situation and existing infrastructure. For example, over 50 percent of Rotterdam's containers are shipped out by truck, while this percentage is 75 percent for the New York/New Jersey ports. Here, we consider only the road case for the following reasons: (i) substitution to other modes is generally infeasible in the short term given the existing transport network, geographical limitation and product characteristics. For example, in North America, short-haul (<500km) freight transport is generally done by trucking, whereas long-haul (>500km) is mainly by rail. (ii) Probably, only the road would have enough local traffic to make the problem interesting. And (iii) we could assume that substitution to other modes is implicitly captured by the price elasticity of demand.

region H's local traffic, which we assume to be private cars and denote by Y_H. Figure 3.1 shows the basic network structure.

[Figure 3.1 here]

The stylized problem we want to analyze is the following: Given that the gateway charge and road tolls are set non-cooperatively by the respective regions, what would be the effects on the road tolls (and region H's social welfare) of the gateway moving towards congestion pricing? For this, we consider a base case in which the gateway does not implement congestion pricing but the road does, and an alternate case in which the gateway also implements congestion pricing. We model the problem as a three-stage game. In the first stage, both the gateway and the road decide their charges to users. In the second stage, the N carriers compete in Cournot fashion, deciding on their quantities both for consumption at region G and for delivery to region H.[45] In the third stage, demands for cargo and the local road traffic are realized. We assume, in this Chapter, that capacities of the gateway and the road are fixed.[46]

We shall examine the subgame perfect equilibrium of this game, starting with the stage-3 analysis. Here, we assume that the demands for cargo in both regions come from competitive markets which would either resell the goods to final consumers, or use them

[45] Cournot competition between air carriers has been a common assumption in the theoretical airport pricing literature, which in fact has some empirical backing. See Basso and Zhang (2007b) for a review of the airport pricing literature.

[46] Variable capacities at either the gateway or road (or both) are considered as an important, yet difficult, extension to the analysis presented here. We discuss the issue further in the concluding remarks.

to produce final goods to be sold (this assumption will simplify analysis in Section 3.4). The inverse demands for cargo units in regions G and H are given by:

$$P_G = \alpha_G - \beta_G Y_G, \qquad P_X = \alpha_X - \beta_X X. \tag{3.1}$$

P_G and P_X are the prices that the carriers charge to cargo consumers in each case, and parameters α and β are positive.[47] Note that, on one hand, these prices need not be the same since in serving the market in region H, a *delivery* service is provided as well.[48] On the other hand, consumers of cargo services do not care about congestion costs; that is, they only care about the final price and the cargo actually being delivered. As is to be seen below, the carriers do have extra costs induced by congestion and those costs will indeed be reflected in their prices.

Local traffic on the road, Y_H, depends linearly on a "full price" ρ_H, with the inverse demand function being:

$$\rho_H = \alpha_H - \beta_H Y_H. \tag{3.2}$$

The full price is the sum of the road toll and congestion cost. And, although the contribution to congestion of one unit of X (for example, a truck with a container) and one unit of Y_H (for example, a private car) are probably not the same, here we simply let

[47] Because this Chapter is somewhat intensive in notation, we provide a glossary of variables in Appendix A.

[48] Here we assume, for simplicity, that consumers in region G do not require any delivery service from the carriers.

the equivalence factor be one, so total traffic on the road is $Y_H + X$.[49] Thus, assuming that (average) congestion delay (for example, in hours) is a linear function of the total-traffic-to-capacity ratio, we can write the full price as:

$$\rho_H = \tau_H + \gamma_H \frac{Y_H + X}{K_H}, \tag{3.3}$$

where τ_H is the road toll imposed on local road users, and $\gamma_H (> 0)$ is the delay cost parameter for local road users which is defined as the product of their value of travel time savings (VTTS) and the (constant) slope of the congestion delay function. Here, we assume that the slope of the congestion delay function is the same for all road users.[50] K_H is the road's capacity. From (3.2) and (3.3) we can write the road local traffic as:

$$Y_H = z_H(\tau_H, X). \tag{3.4}$$

Applying the implicit function theorem to (3.2) and (3.3), we obtain the following useful partial derivatives:

$$\frac{\partial z_H}{\partial \tau_H} = \frac{-1}{\beta_H + (\gamma_H / K_H)} < 0, \qquad \frac{\partial z_H}{\partial X} = -\frac{\gamma_H / K_H}{\beta_H + (\gamma_H / K_H)} < 0,$$

[49] In his study on the social costs of additional traffic in the congested New Jersey highway system due to the expansion of the Port of New York, Berechman (2007) suggests that a "private car equivalence" ratio of 3 or 4 for each freight truck movement is to be used.

[50] This is a standard and reasonable assumption. As pointed out by an anonymous referee, in principle, the slope could differ for cars and trucks due to differences in preferred travel speeds, maneuverability and difficulties in passing.

$$1+\frac{\partial z_H}{\partial X}=\frac{\beta_H}{\beta_H+(\gamma_H/K_H)}>0. \tag{3.5}$$

Inequalities (3.5) show that (i) an increase in the local road toll will, as expected, reduce the local road traffic (the first inequality); (ii) an (exogenous) increase in transit traffic will decrease the local road traffic (the second inequality); (iii) an (exogenous) increase in transit traffic will, while reducing the local road traffic, increase the overall road traffic (the third inequality).

Having analyzed the demands, we now specify firms' costs. The carriers are symmetric in the sense that they have the same cost function, with carrier i's cost function being given as:

$$\begin{aligned} C^i(Y_G^i, X^i; \mathbf{Y}_G^{-i}, \mathbf{X}^{-i}, Y_H) = (Y_G^i + X^i)\left(c_G + t_G + \gamma_G \frac{Y_G + X}{K_G}\right) \\ + X^i\left(c_H + t_H + \gamma_X \frac{Y_H + X}{K_H}\right) \end{aligned} \tag{3.6}$$

where $Y_G = \sum_i Y_G^i$, $X = \sum_i X^i$ and K_G denotes the gateway's capacity. The first term on the right-hand side (RHS) represents the cost of bringing the cargo into region G. Here, the unit cost has three components: (i) a fixed unit operating cost c_G; (ii) the gateway charge t_G; and (iii) the congestion cost, where γ_G is the delay cost parameter – the product of the VTTS and the slope of the congestion delay function – for carriers at the gateway. The second term on the RHS represents the cost to the carrier of delivering the cargo to region H, which includes a fixed unit operating cost c_H, road toll t_H, and delay costs due

to road congestion, where γ_X is the delay cost parameter for carriers at the road. We assume hereafter that $\gamma_X > \gamma_H$, since the cost of an hour of delay for a carrier's truck is obviously larger than that for a car.[51] Note that while the slope of the congestion delay function is assumed to be the same for all facility users, a larger VTTS implies a larger γ.

The cost function in (3.6) makes apparent some important aspects of our modeling. First, both the gateway charge and the road toll are levied per unit of cargo. Note, however, that under a "fixed proportions" assumption – the number of cargo units per plane/ship is fixed and the same across carriers – the gateway charge per unit of cargo and the gateway charge per vehicle are equivalent. Second, the gateway is unable to price discriminate between the cargo remaining in region G and the cargo delivered to region H. This is, we believe, a sensible assumption as carriers would have both types of cargo in each vehicle. Third, in principle, the road would be able to price discriminate between the local traffic and the cargo traffic coming from region G. This is feasible in certain situations and may in fact be an important tool for the road authority or the region H's government. For instance, the government can implement cordon pricing – charging a fee to traffic entering its region. In such a case, only the transit traffic is required to pay for the fee. In what follows, we will nevertheless explore both the price-discrimination case and the "uniform pricing" case, in which the road charges a uniform toll to the two types

[51] US DOT (2003) recommends a value of travel time between $10.60/hour and $21.20/hour (in 2000 US$) for cars and light trucks, depending on the purpose of the trip (personal, commuting or "on-the-clock" travel). On the other hand, a study by Levinson and Smalkoski (2003) find a mean value of $49.42/hour for commercial vehicle operators in Minnesota, whereas De Jong (2000) finds a range of values from $36/hour to $48/hour. The costs of travel delays used by Transport Canada (2006) range from CAD$24.71/hour to CAD$31.35/hour (in 2002 CAD$) for work and work-related trips, while from CAD$7.63/hour to CAD$9.67/hour for non-work-related trips.

of traffic.[52] Note that the uniform-pricing case does not necessarily mean trucks and cars are charged exactly the same tolls; instead it should be understood as the existence of some *a priori* rule which says that cars and trucks are charged according to some fixed ratio (for example, a truck pays three times what a car pays). To save on notation – but without losing insight – we will set the ratio to one in the uniform-pricing case, that is, $\bar{t}_H = t_H = \tau_H$.

3.3 Carriers' Competition

After solving for the consumer problem, in this section we analyze the carriers' subgame in the second stage. Carrier i's optimization problem is:

$$\underset{Y_G^i, X^i}{Max}\ \pi^i(Y_G^i, X^i; \mathbf{Y}_G^{-i}, \mathbf{X}^{-i}) = (\alpha_G - \beta_G Y_G)Y_G^i + (\alpha_X - \beta_X X)X^i - C^i(Y_G^i, X^i; \mathbf{Y}_G^{-i}, \mathbf{X}^{-i}, Y_H).$$

Notice that, because this stage occurs prior to the decisions of local road users, the Y_H term in total cost specification (3.6) is actually seen by carriers as functions of X – see equations (3.4) and (3.5). The Cournot-Nash equilibrium of the carriers' subgame is characterized by the first-order conditions (assuming the second-order conditions and other regularity conditions hold). Deriving these and imposing symmetry, it is easy to obtain the following equations:

[52] Another possible case, which is examined by De Borger, et al. (2005) and De Borger, et al. (2007), is that the road imposes a toll on the local users but not the transit traffic. De Borger, et al. (2005) argue that the case resembles the current situation in many countries, where fuel taxes are the main tolling instrument. However, as the fuel taxes are not an effective policy tool for targeting congestion (Lindsey, 2007), this case is not discussed in our congestion-pricing analysis.

$$\Omega_{Y_G} = \frac{\partial \pi^i}{\partial Y_G^i} = (\alpha_G - \beta_G Y_G - c_G - t_G - \gamma_G \frac{Y_G + X}{K_G}) - \frac{\gamma_G}{K_G}\frac{Y_G + X}{N} - \beta_G \frac{Y_G}{N} = 0, \tag{3.7}$$

$$\begin{aligned}\Omega_X = \frac{\partial \pi^i}{\partial X^i} &= (\alpha_X - \beta_X X - c_G - c_H - t_G - t_H - \gamma_G \frac{Y_G + X}{K_G} - \gamma_X \frac{Y_H + X}{K_H}) \\ &- \frac{\gamma_G}{K_G}\frac{Y_G + X}{N} - \frac{\gamma_X}{K_H}\frac{X}{N}(1 + \frac{\partial z_H}{\partial X}) - \beta_X \frac{X}{N} = 0.\end{aligned} \tag{3.8}$$

Equations (3.4), (3.7) and (3.8) implicitly define the gateway and road demands as functions of the facility charges:

$$Y_G = Y_G(t_G, t_H, \tau_H), \qquad X = X(t_G, t_H, \tau_H), \qquad Y_H = Y_H(t_G, t_H, \tau_H). \tag{3.9}$$

These are the demands that the gateway and the road take into account when deciding on their charges, and are the *derived demands* from the equilibria of stages 2 and 3 of the game.[53] The shapes of these demands can be obtained through straightforward yet tedious comparative statics, and are contained in the following lemmas (their proofs are provided in Appendix B). The first lemma is as follows:

Lemma 3.1 The traffic responses to changes in the gateway charge are as follows:

$$\frac{\partial Y_G}{\partial t_G} < 0, \qquad \frac{\partial X}{\partial t_G} < 0, \qquad \frac{\partial Y_H}{\partial t_G} > 0, \qquad \frac{\partial (Y_H + X)}{\partial t_G} < 0. \tag{3.10}$$

[53] Hence, these demands are different than the ones in Section 3.2; for example, Y_H here is different from $Y_H \equiv z_H$ in equation (3.4). To save on notations though, we do not write them differently.

The first two terms show that, as expected, if the gateway charge increases, then Y_G, X and hence total traffic at the gateway $Y_G + X$ will decrease. Because in particular the amount of cargo delivered to region H will decrease (second inequality), congestion on the road – and hence the full price perceived by the local road traffic – will diminish, inducing a rise in local road traffic, Y_H (third inequality). Overall though, total traffic on the road falls (last inequality), which also implies that an increase in the gateway charge, *ceteris paribus*, reduces congestion at both the gateway and the road.

Responses of the gateway and road demands to changes in the road tolls depend on whether the road price discriminates between the two types of traffic. The price-discrimination results are given in Lemma 3.2:

Lemma 3.2 With price discrimination at the road, the traffic responses to changes in the road tolls are given as:

$$\frac{\partial X}{\partial t_H} < 0, \quad \frac{\partial Y_G}{\partial t_H} > 0, \quad \frac{\partial (Y_G + X)}{\partial t_H} < 0, \quad \frac{\partial Y_H}{\partial t_H} > 0, \quad \frac{\partial (Y_H + X)}{\partial t_H} < 0. \tag{3.11}$$

$$\frac{\partial Y_H}{\partial \tau_H} < 0, \quad \frac{\partial X}{\partial \tau_H} > 0, \quad \frac{\partial (Y_H + X)}{\partial \tau_H} < 0, \quad \frac{\partial Y_G}{\partial \tau_H} < 0, \quad \frac{\partial (Y_G + X)}{\partial \tau_H} > 0. \tag{3.12}$$

The inequalities in (3.11) show what happens when the road toll levied on transit cargo increases: Because the unit cost of transit cargo rises, the carriers will reduce the

amount of transit cargo they bring into region H (first inequality) to favor the cargo that stays in region G (second inequality). Overall, however, total traffic at the gateway will fall as a result of the larger road toll (third inequality). Next, since transit traffic on the road falls, road congestion will also fall, thereby boosting up local road traffic (fourth inequality). Total traffic on the road will still be smaller though (fifth inequality), which shows that an increase in the cargo road toll reduces total traffic and congestion at both the road and the gateway.

The inequalities in (3.12) show, on the other hand, what happens when the road toll levied on local users increases. The obvious decrease in the local traffic (first inequality) induces an increase in the volume of cargo on the road (second inequality) due to falling congestion. Overall, however, total traffic on the road will be diminished (third inequality). Next, since the carriers increase the amount of cargo that goes to region H, gateway congestion will rise and consequently, the carriers will reduce the amount of cargo that remains in region G (fourth inequality). Still, total traffic at the gateway will rise (fifth inequality), indicating that an increase in the local traffic toll reduces congestion on the road *but increases* congestion at the gateway. Different from the comparative statics with respect to t_G and t_H, in this case an increase in τ_H increases congestion at one facility but reduces congestion at the other facility. This happens because τ_H affects both the local and transit traffic at the gateway only indirectly through the change in congestion, but not through the change in price directly.

Now consider the case in which the road uses *uniform pricing*, that is, $\bar{t}_H = t_H = \tau_H$. In Appendix B we show:

Lemma 3.3 With the uniform pricing at the road, the traffic responses to changes in the single road toll are given as:

1. $\dfrac{\partial Y_H}{\partial \bar{t}_H} < 0$ and $\dfrac{\partial (Y_H + X)}{\partial \bar{t}_H} < 0$; (3.13)

2. If $\gamma_X - \gamma_H < \beta_H K_H$, then

$$\frac{\partial X}{\partial \bar{t}_H} < 0, \quad \frac{\partial Y_G}{\partial \bar{t}_H} > 0, \quad \frac{\partial (Y_G + X)}{\partial \bar{t}_H} < 0; \tag{3.14a}$$

If $\gamma_X - \gamma_H \geq \beta_H K_H$, then

$$\frac{\partial X}{\partial \bar{t}_H} \geq 0, \quad \frac{\partial Y_G}{\partial \bar{t}_H} \leq 0, \quad \frac{\partial (Y_G + X)}{\partial \bar{t}_H} \geq 0. \tag{3.14b}$$

Lemma 3.3 indicates that an increase in $\bar{t}_H$ reduces both the local traffic and total traffic on the road. However, the responses of transit traffic X, local gateway traffic Y_G and total gateway traffic $Y_G + X$ to changes in the (single) road toll depend on the relative sizes of carriers' and local road users' values of travel time savings. Recall that a larger value of travel time savings implies a larger γ. From (3.14a) and (3.14b) one can see the existence of a "cut-off rule": if the difference between the values of travel time savings for cargo and local road users is small, an increase in the toll reduces X on the

road. But if the difference is large, then X will *increase*.[54] This happens because a large difference in the values of time makes the unit cost of X *decrease* if $\bar{t}_H$ increases, inducing an increase in the volume of X.[55] The intuition is as follows: If $\bar{t}_H$ increases, the smaller the γ_H, the larger the decrease in local road traffic Y_H because the users care, relatively speaking, more about the toll than congestion – it can formally be seen from (3.5) that $\partial|(\partial z_H / \partial \tau_H)| / \partial \gamma_H < 0$. If, in addition, γ_X is large, the decrease in Y_H leads to strong congestion cost savings for carriers, which may in fact dominate the increase in the road toll itself, leading to a reduction in the unit cost of X.[56] The other two inequalities in (3.14a) and (3.14b) are then expectable: If the unit cost of X increases (decreases), the carriers will reduce (increase) their output level, while increasing (decreasing) the volume of Y_G. The latter effect however is dominated by the change in X.

Finally, equations (3.7) and (3.8) also allow us to obtain the carriers' (subgame) equilibrium pricing rules. From (3.7) we obtain the pricing rule for cargo that stays in region G:

---

[54] Glazer (1981) and Niskanen (1987) derive similar results for two traffic types and a continuum of traffic types, respectively. We thank an anonymous referee for pointing out this to us.

[55] It is a well-known property of symmetric Cournot games that the equilibrium output is a decreasing function of marginal cost.

[56] To formally obtain the cut-off rule, note from (3.6) that the unit cost of X is given by $c_X = c_G + t_G + [\gamma_G (Y_G + X) / K_G] + c_H + \bar{t}_H + [\gamma_X (Y_H + X) / K_H]$. As we want to see how the change in $\bar{t}_H$ affects carriers' equilibrium outputs, we hold X and Y_G constant here. Hence, if $\bar{t}_H$ increases, the unit cost changes by $1 + (\gamma_X / K_H)(\partial z_H / \partial \tau_H) = (\beta_H K_H + \gamma_H - \gamma_X) / (\beta_H K_H + \gamma_H)$, where the equality follows from (3.5). It is then easy to check that the unit cost change is negative if and only if $\gamma_X - \gamma_H \geq \beta_H K_H$.

$$P_G = c_G + t_G + \gamma_G \frac{Y_G + X}{K_G} + \frac{\gamma_G}{N} \frac{Y_G + X}{K_G} + \frac{1}{N} \beta_G Y_G . \tag{3.15}$$

This pricing rule is known by now (see, for example, Brueckner, 2002; Basso and Zhang, 2007a). The first three terms on the RHS of (3.15) represent a carrier's average cost, which includes congestion cost at the gateway. The fourth term is a mark-up that represents internalization of congestion by a carrier: A marginal increase in its output would induce an increase in congestion, but the carrier only charges for the congestion it imposes on its own traffic. The fifth term is a mark-up owing to market power in the Cournot oligopoly. As N increases, both the market power and congestion internalization mark-ups diminish.

On the other hand, from (3.8) we can obtain the pricing rule for cargo that is delivered to region H:

$$\begin{aligned} P_X = & (c_G + c_H + t_G + t_H + \gamma_G \frac{Y_G + X}{K_G} + \gamma_X \frac{Y_H + X}{K_H}) \\ & + \frac{\gamma_G}{N} \frac{Y_G + X}{K_G} + \frac{\gamma_X}{N} \frac{X[1 + (\partial z_H / \partial X)]}{K_H} + \frac{1}{N} \beta_X X . \end{aligned} \tag{3.16}$$

The pricing rule (3.16) is slightly more complicated as this type of cargo uses both the gateway and the road. The first bracketed term on the RHS represents average costs which, this time, include congestion costs at both the gateway and the road. The second and third terms represent internalization of congestion at the gateway and the road

respectively. The one catch in the case of the road is that, while still internalizing only the congestion they impose on themselves, carriers realize that an increase in their cargo volume would induce a response from local road users. The last (fourth) term is, again, a market power mark-up.

3.4 Effects of Gateway Congestion Pricing

In the facility pricing stage (stage 1) regions G and H simultaneously choose charges at their facilities, taking the subsequent carrier and consumer behavior into account.[57] As indicated earlier, our principal objective in the present Chapter is to examine the effect of gateway congestion pricing on optimal road tolls. To focus on this objective, we shall consider a joint-profit maximizing case, in which the gateway maximizes the sum of the gateway's and carriers' profits. As is to be shown below, in this case the gateway charge will in general rise and will include a congestion pricing component. This will then make the linkage between gateway congestion pricing and road tolls to be demonstrated as clearly as possible.

In addition to this modeling consideration, focusing on the joint-profit maximizing case is relevant in practice. On one hand, vertical integration between private ports and shipping lines is quite common. Ocean carriers have entered the port industry via the development of dedicated terminals at major load centers (see, for example,

[57] Note that governments, consumers and firms outside the system (i.e., Regions G and H) are not taken into account in the analysis.

Brennan, 2002; Notteboom, 2004).[58] In the case of airports, on the other hand, it has been often argued that more strategic collaboration between airlines and airports (or greater airlines' countervailing power) may improve efficiency of privatized airports and solve congestion problems, by allowing a better alignment of incentives (see, for example, Beesley, 1999; Condie, 2000; Forsyth, 1997; Starkie, 2001; Productivity Commission, 2002; Civil Aviation Authority UK, 2004). This joint-profit maximizing case has been analytically examined by Basso (2008), who notes that, with sufficient carriers' symmetry, a simple two-part tariff pricing mechanism may be enough for the outcome of joint-profit maximization to arise. Finally, although potential gateway competition is not considered in the present analysis, in practice carriers may choose among gateways in the same region based on their pricings and congestion levels. Facing such competition and to retain or attract traffic, a particular gateway operator may also be required to take its carriers' interests into account when making its pricing decision, thereby leading to the pricing behavior close to joint-profit maximization.[59]

3.4.1 Gateway pricing

We first consider the gateway's pricing decision. Conditional on region H's pricing decisions, the gateway chooses a charge to maximize the joint profit of the gateway and carriers:

[58] In 1991, the shipping lines controlled 11 percent of world container port handling; but by 2001, this share had risen to 19 percent (Slack and Fremont, 2005).

[59] Nevertheless, although focusing on the joint-profit maximizing case, we will also discuss, below, gateway pricing for the cases with the gateway being pure public or pure private.

$$\underset{t_G}{Max} \qquad \Pi + (t_G - C_G)(Y_G + X), \tag{3.17}$$

where $\Pi = \sum_i \pi^i$ is the aggregate carriers' profit, and C_G denotes the (constant) unit operating cost for the gateway to provide service to the carriers. The first-order condition yields:

$$\begin{aligned}&[t_G^* + C_G - \frac{N-1}{N}(\gamma_G \frac{Y_G + X}{K_G} + Y_G \beta_G)]\frac{\partial Y_G}{\partial t_G} \\ &\quad + \{t_G^* + C_G - \frac{N-1}{N}[\gamma_G \frac{Y_G + X}{K_G} + \gamma_X (1 + \frac{\partial z_H}{\partial X})\frac{X}{K_H} + X\beta_X]\}\frac{\partial X}{\partial t_G} = 0.\end{aligned}$$

Rearranging the first-order condition, we obtain:

$$\begin{aligned}t_G^* = C_G &+ \frac{N-1}{N}\gamma_G \frac{Y_G + X}{K_G} + \frac{N-1}{N}\gamma_X (1 + \frac{\partial z_H}{\partial X})\frac{X}{K_H}\frac{\partial X}{\partial t_G} / (\frac{\partial Y_G}{\partial t_G} + \frac{\partial X}{\partial t_G}) \\ &+ \frac{N-1}{N}[Y_G \beta_G \frac{\partial Y_G}{\partial t_G} / (\frac{\partial Y_G}{\partial t_G} + \frac{\partial X}{\partial t_G}) + X\beta_X \frac{\partial X}{\partial t_G} / (\frac{\partial Y_G}{\partial t_G} + \frac{\partial X}{\partial t_G})].\end{aligned} \tag{3.18}$$

According to the gateway pricing rule (3.18), the gateway charge is equal to the gateway's operating cost (the first term) plus three additional terms.[60] The second and third terms on the RHS are included so as to induce the carriers to fully internalize congestion costs at both the gateway and the road – as discussed in Section 3.3, at both

[60] Note that in the case that the gateway price discriminates cargos based on their destinations, the cargos destined for the gateway and hinterland will be charged by $C_G + [(N-1)/N][\gamma_G (Y_G + X)/K_G + Y_G \beta_G)$ and $C_G + [(N-1)/N]\{\gamma_G (Y_G + X)/K_G + \gamma_X [1 + (\partial z_H / \partial X)](X / K_H) + X\beta_X\}$, respectively.

facilities, a carrier only charges for the congestion it imposes on its own traffic. Thus, joint-profit maximization will indeed induce congestion pricing.[61] But the fourth term further increases the price. This term is put in place to countervail the business-stealing effect – a carrier does not take into account profits lost by its competitors when it expands its output. In other words, as a first mover of the game, the gateway will induce a cartel outcome at the carriers' level. Similar results are found in Basso (2008) and Basso and Zhang (2007a) in the setting of a single gateway region (that is, region H and its congestible road were not considered). Finally, note that the gateway does not take into account the external effects of its pricing on the hinterland, as it can be seen from the maximization problem (3.17). This "tax competition issue" – the local tax policy disregards its effects on the other region's revenue – has been discussed in the literature (for example, De Borger, et al., 2007, 2008).

Since it is easy to verify that if $N > 1$ then $t_G^* > C_G$ – see (3.5) and Lemma 3.1 – we can conclude that:[62]

Proposition 3.1 For an oligopoly carrier market, a gateway pursuing joint-profit maximization will induce an increase in the gateway charge beyond its operating cost, part of which corresponds to congestion tolls.

[61] Note that we assume both gateway and road are congested. Yet if the gateway's capacity is in excess, no congestion pricing will be implemented. As a result, our question about its effects on its hinterland will not arise.

[62] Due to the complexity of the problem, this Chapter considers linear function of demands and delay costs, as in other studies in the literature (for example, Pels and Verhoef, 2004; Basso and Zhang, 2007a; De Borger, et al., 2007; Ubbels and Verhoef, 2007). Nevertheless, it should be noted that the analytical results in this section may not hold if more general functional forms are considered.

Before analyzing how the new gateway pricing scheme affects the hinterland's road tolls and welfare, it is important to point out two things. First, for $N = 1$, the gateway charge in (3.18) is equal to the operating cost. As Basso (2008) explains, this is so because, on one hand, a monopoly carrier will fully internalize the congestion costs (Brueckner, 2002), and on the other hand, there is no need to correct for the business-stealing effect. Therefore, for the special case of $N = 1$, the gateway charge will remain unchanged in the joint-profit maximization and, consequently, there will be no impact on the hinterland's road tolls and welfare.

Second, a public gateway maximizing the gateway's social welfare – the sum of gateway and carriers' profits plus the consumer surplus of its cargo market, Y_G – would impose a congestion toll with the same functional form as the joint-profit maximizing gateway does, i.e., the second and third terms on the RHS of (3.18), but it should be also noted that the levels of the toll would be different due to different equilibrium outputs in the two cases. Charging only these tolls though would not be the first best because, in this case, subsidies would be needed to correct for downstream market power (Pels and Verhoef, 2004). While in general the balance between congestion tolls and optimal subsidies may be positive or negative, congestion tolls are, as shown, always positive for an oligopoly carrier market.[63] On the other hand, for a pure private gateway, which only maximizes its own profit, its pricing rule can be shown to be equal to that given in (3.18) plus a positive mark-up term if N is finite (see, for example, Basso, 2008). This result

[63] In particular, for a monopoly, the congestion toll components vanish and thus the optimal gateway charge is below marginal cost and perhaps even negative, while for a perfectly competitive market, the optimal subsidies are zero and thus the charge is above marginal cost.

implies "double marginalization" – a well-known result in vertical control literature since Spengler (1950). Understandably, the mark-up term will vanish when the downstream carrier market is perfectly competitive ($N \to \infty$); in that case, its pricing rule is equivalent to (3.18).

3.4.2 Effects on road tolls, congestion and welfare

To examine the effects of the gateway congestion pricing on its hinterland, we first derive the optimal road tolls for both the price-discrimination and uniform-pricing cases.[64] In the price-discrimination case, the hinterland chooses road tolls to maximize its social welfare conditional on region G's pricing decision (superscript *PD* denoting "price discrimination"):

$$\underset{t_H,\tau_H}{Max}\, SW_H^{PD} = [\int_0^{Y_H}(\alpha_H - \beta_H y)dy - \rho_H Y_H] + [\int_0^{X}(\alpha_X - \beta_X x)dx - P_X X] + (Y_H \tau_H + X t_H), \tag{3.19}$$

where the first bracketed term in (3.19) is consumer surplus of the local road user market, and the third bracketed term represents road profit. On the other hand, the benefit to region H that arises from the transit cargo is the sum of the profits of firms that demanded the cargo, and final consumer surplus. Because we assumed in Section 3.2 that the

[64] As pointed out by an anonymous referee, the road-pricing scenario envisaged here might be motivated more from a forward-looking than a current perspective. While toll roads are prevalent around the world, only a few roads have tolls that are designed to price congestion. Nevertheless, road pricing has been gaining support, and there is a reasonable prospect that toll roads generally – and congestion pricing specifically – will become more widespread in the next few years (Lindsey, 2007).

demands for cargo in both regions come from competitive markets, the benefit to region H can be obtained as the area under the demand for transit cargo, as expressed in the second bracketed term in (3.19).[65] Finally, because carriers were assumed to belong to region G, their profits do not enter into region H's objective function.

The first-order conditions yield:

$$(\tau_H^* - \gamma_H \frac{Y_H}{K_H})\frac{\partial Y_H}{\partial \tau_H} + (t_H^* - \gamma_H \frac{Y_H}{K_H} + X\beta_X)\frac{\partial X}{\partial \tau_H} = 0,$$

$$(\tau_H^* - \gamma_H \frac{Y_H}{K_H})\frac{\partial Y_H}{\partial t_H} + (t_H^* - \gamma_H \frac{Y_H}{K_H} + X\beta_X)\frac{\partial X}{\partial t_H} + X = 0.$$

By Cramer's rule and some further algebraic manipulations, we obtain the road-toll rules for local and transit traffic:

$$\tau_H^* = \gamma_H \frac{Y_H}{K_H} + \gamma_X \frac{X}{K_H}, \tag{3.20}$$

$$t_H^* = \gamma_H \frac{Y_H}{K_H} - \gamma_X \frac{X}{K_H}\frac{\partial Y_H}{\partial t_H} / \frac{\partial X}{\partial t_H} - X / \frac{\partial X}{\partial t_H} - X\beta_X. \tag{3.21}$$

Equation (3.20) shows a well-known result: the toll for local users will be equal to the marginal congestion cost – due to a marginal increase in Y_H – imposed on all road users

[65] If either cargo demand is from an oligopoly market, then the area under the cargo demand curve would fail to correctly capture the sum of firms' profit and final consumer surplus; see details in Basso (2005).

(local and transit cargo). Note that due to the difference in time valuations between local users and carriers, the composite of local and cargo traffic at the road will affect the level of τ_H^*. For instance, for a given total road traffic, more cargo traffic implies a higher optimal road toll since $\gamma_X > \gamma_H$. The charge on transit cargo though, is more complicated. The first term in (3.21) represents the marginal congestion costs – due to a marginal increase in X – imposed on local road users. But because carriers' output decisions happen before the decisions of local users, carriers know that an increase in their output will be followed by a decrease in local road traffic, which will save them on congestion costs. These cost savings, which happen at the expense of local traffic, are charged by the road authority in the second term of (3.21). The third term, related to the elasticity of transit cargo demand, is obviously a mark-up that arises from revenue maximization from the part of the road authority. And the fourth term – a mark-down – flows from the fact that cargo traffic induces benefits for the region, and thus carriers are "subsidized" as a way to countervail their output contraction that stems from their monopolistic behavior (because of the induced cartel).

Equations (3.20) and (3.21) define the reaction functions of the optimal discriminating road tolls with respect to any gateway charge t_G. When the gateway charge increases, following joint-profit maximization, there are two opposing effects on the optimal local road toll τ_H^*. On one hand, the increase in t_G will reduce the transit traffic at the road ($\partial X / \partial t_G < 0$ by Lemma 3.1), thereby decreasing the marginal congestion cost imposed on transit cargo. On the other hand, the increase in t_G will

increase the local road traffic ($\partial Y_H / \partial t_G > 0$ by Lemma 3.1), thereby increasing the marginal congestion cost imposed on local road users. But since the former effect is direct while the latter is indirect, the former dominates the latter, and hence the optimal road toll on local traffic will fall after the gateway implements its new pricing scheme. As for the optimal cargo road toll t_H^*, the increase in gateway charge will again, by Lemma 3.1, have a positive impact on the first three terms in (3.21) but a negative impact on the last term. As the latter dominates the former, the cargo road toll will fall after the gateway implements its new pricing scheme. All these results can be proved formally (see Appendix B), and hence we collect them into the following proposition:

Proposition 3.2 If the road can price discriminate between the local and transit cargo traffic, the road's best response to an increase in gateway charge is to decrease the road tolls:

$$\frac{\partial \tau_H^*}{\partial t_G} < 0, \qquad \frac{\partial t_H^*}{\partial t_G} < 0. \tag{3.22}$$

Thus, if the carrier market is oligopolistic and the gateway maximizes the joint profit of itself and its carriers, the hinterland will reduce the road tolls.

If, on the other hand, we impose uniform pricing at the road and then solve problem (3.19), we can derive the (single) road-toll rule for the uniform-pricing case:

$$\bar{t}_H^* = \gamma_H \frac{Y_H}{K_H} - X(1+\beta_X \frac{\partial X}{\partial \bar{t}_H})/(\frac{\partial Y_H}{\partial \bar{t}_H} + \frac{\partial X}{\partial \bar{t}_H}). \tag{3.23}$$

Intuitively, since in the uniform-pricing case region H can only charge a single road toll to both local users and cargos, there will be a trade-off between the welfare considerations of the two markets in the optimal road toll. This can be seen analytically by noticing that (3.23) can be rewritten as:

$$\bar{t}_H^* = \tau_H^* \frac{\partial Y_H}{\partial \bar{t}_H}/(\frac{\partial Y_H}{\partial \bar{t}_H} + \frac{\partial X}{\partial \bar{t}_H}) + t_H^* \frac{\partial X}{\partial \bar{t}_H}/(\frac{\partial Y_H}{\partial \bar{t}_H} + \frac{\partial X}{\partial \bar{t}_H}), \tag{3.24}$$

where τ_H^* and t_H^* are given in (3.20) and (3.21), respectively. Thus, (3.24) shows that $\bar{t}_H^*$ is a weighted average of τ_H^* and t_H^*. The relative importance of a market in the toll depends on its responsiveness to the toll change, as given by the magnitudes of $\partial Y_H / \partial \bar{t}_H$ and $\partial X / \partial \bar{t}_H$. The interesting catch here though, is that the weight that multiplies t_H^* may be negative if the values of travel time savings are too different, because $\partial X / \partial \bar{t}_H$ would be positive in that case (Lemma 3.3). This may lead to the conjecture that, if the road uses uniform pricing and the difference in values of travel time savings is large, the optimal (single) toll may actually increase with an increase in gateway pricing. This, however, is not the case, as shown in Appendix B. Hence:

***Proposition* 3.3** If the road uses a uniform toll, the road's best response to an increase in gateway charge is to decrease the (single) road toll:

$$\frac{\partial \bar{t}_H^*}{\partial t_G} < 0. \tag{3.25}$$

Thus, if the carrier market is oligopolistic, and the gateway maximizes the joint profit of itself and its carriers, the hinterland will reduce the (single) road toll.

As the implementation of congestion pricing by the gateway will reduce road tolls, a natural question is: What will be its effect on road congestion? It turns out that this effect is, in general, undetermined. To illustrate this point, consider the price-discrimination case.[66] With the road pricing rules given in (3.20) and (3.21), the demands at the road can be written as:

$$X = X(t_G, t_H^*(t_G), \tau_H^*(t_G)), \qquad Y_H = Y_H(t_G, t_H^*(t_G), \tau_H^*(t_G)) \tag{3.26}$$

Totally differentiating (3.26) with respect to t_G, we have:

$$\frac{d(Y_H + X)}{dt_G} = \frac{\partial(Y_H + X)}{\partial t_G} + \frac{\partial(Y_H + X)}{\partial t_H}\frac{\partial t_H^*}{\partial t_G} + \frac{\partial(Y_H + X)}{\partial \tau_H}\frac{\partial \tau_H^*}{\partial t_G} \tag{3.27}$$

[66] The uniform-pricing case can be analyzed similarly, and the result is available upon request.

One can see, from (3.27), that a change in gateway charge has two kinds of impact. The *direct* impact, captured by the first term on the RHS, is negative by Lemma 1, suggesting that an increase in gateway charge reduces total road traffic and hence road congestion. On the other hand, the *indirect* impact through the change in t_H^* and τ_H^*, captured by the second and third terms on the RHS, is positive by (3.22) and Lemma 3.2. As a consequence, an increase in gateway charge will increase total traffic at the road and hence road congestion. To examine the net effect, we substitute previous results into (3.27) and simplify the resulting expressions, obtaining that:

$\frac{d(Y_H + X)}{dt_G} < 0$ if and only if

$$\frac{\partial t_H^*}{\partial t_G} + [\frac{\gamma_X}{K_H} - \frac{N+1}{N}\Delta/(\beta_G + \frac{\gamma_G}{K_G})(1 + \frac{\partial z_H}{\partial X})]\frac{\partial z_H}{\partial \tau_H}\frac{\partial \tau_H^*}{\partial t_G} > -\beta_G/(\beta_G + \frac{\gamma_G}{K_G}) \tag{3.28}$$

where $\Delta \equiv (\beta_G + \frac{\gamma_G}{K_G})[\beta_X + \frac{\gamma_X}{K_H}(1 + \frac{\partial z_H}{\partial X})] + \beta_G \frac{\gamma_G}{K_G} > 0$. This result implies that if $\partial t_H^* / \partial t_G$ and $\partial \tau_H^* / \partial t_G$ are small in magnitude, then $d(X + Y_H)/dt_G < 0$ and hence an increase in gateway charge reduces congestion at the road.

Whilst the implementation of congestion pricing by the gateway has, in general, an undetermined effect on road congestion, the impact of gateway congestion pricing on the hinterland's social welfare is negative. Essentially, an increase in gateway charge

reduces region H's welfare. Since in our setting the implementation of congestion pricing by the gateway raises its charge, the hinterland's welfare will fall as a result. The result is given in Proposition 3.4 (the proof is given in Appendix B):

Proposition 3.4 If the carrier market is oligopolistic, and the gateway maximizes the joint profit of itself and its carriers, the hinterland's social welfare will fall in both the price-discrimination case and the uniform-pricing case.

Although the impact of gateway congestion pricing on the hinterland's aggregate welfare is determined by Proposition 3.4, the respective welfare implications on road profit, and consumer surplus of the local road user and cargo markets are less clear. Consider the full costs incurred by the local road users and cargo consumers. On one hand, the full costs may fall as the road tolls decrease after the imposition of gateway congestion pricing (by Propositions 3.2 and 3.3). On the other hand, the impact on congestion costs is generally undetermined as discussed above. As a result, the overall effects on the full costs and thus the respective welfare are generally undetermined. Similarly, the impact on road profit is also generally undetermined, given the decrease in road tolls and the impact on total traffic being undetermined.

3.5 Concluding Remarks

Our main objective in this Chapter is to investigate the effects of congestion pricing implemented at a gateway (port or airport) on its hinterland's optimal road tolls,

road congestion and social welfare. We have examined the effects of facility charges on equilibrium outputs (derived demands). We found that, with the uniform pricing at the road, the effect of the single road toll on the changes in the local and transit cargo traffic, and hence the congestion level at the gateway, depends on the values of travel time savings for the transit cargo and local road users. Further, our analysis showed that when the gateway maximizes the joint profit of itself and its carriers, the gateway charge will rise after congestion pricing is implemented at the gateway. This increase in gateway charge will lead to lower road tolls and lower social welfare for the hinterland. However, the road congestion level may or may not go up with the gateway congestion pricing.

The Chapter has also raised a number of other issues and avenues for future research. First, the modal split for gateway traffic to its hinterland is an interesting and practically relevant issue. In extending the present analysis to such a case, competition between transport modes at the hinterland needs to be taken into account. Second, this Chapter considers a pure serial transport network, and so the competition between transport facilities (gateways or roads, or both) is not considered. Facility competition is examined by Basso and Zhang (2007a) and De Borger, et al. (2008). Yet the former does not consider the gateway-hinterland problem, whereas the latter does not consider the oligopoly market structure at the gateway. Finally, as discussed earlier, capacity decisions in a transport network have been analyzed in several recent studies, but oligopoly market structure at the gateway was not considered. Considering both the capacity investment and oligopoly market structure introduces substantial complexity to the modeling; as a

consequence, deriving general analytical results appears to be rather difficult. Thus numerical simulations may be required to obtain further insights.

Figure 3.1 Basic network structure

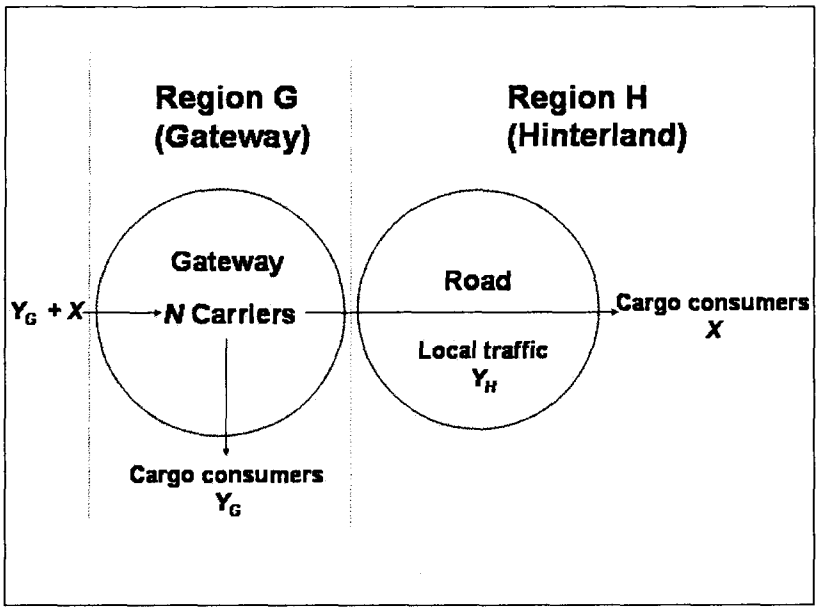


Table 3.1 Recent studies on road pricing and network interactions

Authors	Year	Parallel (P)/ Serial (S)/ with transit(T)	Pricing (P)/ Capacity (C)	Public control(G)/ Private control(P)/ Free toll(F)	Single facility operator (S)/ Duopoly (D)/ More than two (M)
Verhoef, et al.	1996	P	P	G	S
Liu and McDonald	1998	P	P	GPF	S
Liu and McDonald	1999	P	P	GF	S
De Palma and Lindsey	2000	P	P	GPF	D
Levinson	2000	S	P	G	M
Small and Yan	2001	P	P	GPF	S
Verhoef	2002	Generalized	P	G	S
Engel, et al.	2004	P	P	P	M
Verhoef and Rouwendal	2004	PS	PC	GF	S
Verhoef and Small	2004	PS	P	GPF	S
De Borger, et al.	2005	P(T)	P	G	D
Van Dender	2005	P	P	P	D
De Borger and Van Dender	2006	P	PC	P	D
Proost and Sen	2006	No explicit network	P	G	D
De Borger, et al.	2007	S(T)	PC	G	D
Verhoef	2007	PS	PC	GPF	S
Small and Verhoef (Ch. 6)	2007	PS	P	P	M
Ubbels and Verhoef	2008	S	PC	G	D
Zhang	2007	S	PC	GP	D
De Borger, et al.	2008	PS	PC	P	D
This Chapter	2008	S	P	GP	D

Source: The authors and Ubbels and Verhoef (2008).

3.6 References

Basso, L.J. (2005) "On Input Markets Surplus and its Relation to the Downstream Market Game", Working paper, Sauder School of Business, University of British Columbia.

Basso, L.J. (2008) "Airport Deregulation: Effects on Pricing and Capacity", *International Journal of Industrial Organization*, 26(4), 1015-1031.

Basso, L.J. and Zhang, A. (2007a) "Congestible Facility Rivalry in Vertical Structures", *Journal of Urban Economics*, 61(2), 218-237.

Basso, L.J. and Zhang, A. (2007b) "An Interpretative Survey of Analytical Models of Airport Pricing", in: Lee, D. (Ed.), *Advances in Airline Economics*, Vol.2, Elsevier, 89-124.

Beesley, M.E. (1999) "Airport Regulation", in Beesley, M.E. (ed.) *Regulating Utilities: A New Era?* Institute of Economic Affairs, London.

Berechman, Y. (2007) "The Social Costs of Global Gateway Cities: The Case of the Port of New York", Paper presented at the International Conference on Gateways and Corridors, Vancouver, B.C.

Brennan, J.R. (2002) "Brave New World", *Containerisation International*, 35, 39-41.

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, 599-614.

Brueckner, J.K. and Van Dender, K. (2007) "Atomistic Congestion Tolls at Concentrated Airports? Seeking a Unified View in the Internalization Debate", *manuscript.*

Carlin, A. and Park, R. (1970) "Marginal Cost Pricing of Airport Runway Capacity", *American Economic Review*, 60, 310-319.

Civil Aviation Authority (2004) *Airport Regulation: Looking to the Future, Learning from the Past*, CAA, London.

Condie, S. (2000) "Whither Airport Regulation?", in: Bradshaw, B. and Smith, H.L. (eds.), *Privatization and Deregulation of Transport*, St. Martin's Press, Inc, New York, 364-393.

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.

De Borger, B., Dunkerley, F. and Proost, S. (2007) "Strategic Investment and Pricing Decisions in a Congested Transport Corridor", *Journal of Urban Economics*, 62(2), 294-316.

De Borger, B., Proost, S. and Van Dender, K. (2005) "Congestion and Tax Competition on a Parallel Network", *European Economic Review*, 49(8), 2013-2040.

De Borger, B., Proost, S. and Van Dender, K. (2008) "Private Port Pricing and Public Investment in Port and Hinterland Capacity", *Journal of Transport Economics and Policy*, 42(3), 527-561.

De Borger, B. and Van Dender, K. (2006) "Prices, Capacities and Service Levels in a Congestible Bertrand Duopoly", *Journal of Urban Economics*, 60, 264-283.

De Jong, G. (2000) "Value of Freight Travel-Time Savings", in Hensher, D.A. and Button, K.J. (eds.) *Handbook of Transport Modelling*, Elsevier.

De Palma, A. and Lindsey, R. (2000) "Private Roads: Competition under Various Ownership Regimes", *Annals of Regional Science*, 34, 13-35.

Engel, E., Fisher, R. and Galetovic, A. (2004) "Toll Competition among Congested Roads", *Topics in Economic Analysis and Policy*, 4, Article 4.

Forsyth, P. (1997) "Price Regulation of Airports: Principles with Australian Applications", *Transportation Research Part E*, 33, 297-309.

Gillen, D.W. (1994) "The Market and Aviation Infrastructure: Pricing, Productivity and Privatization", *Transportation Research Circular*, No. 425 (App), 82-86.

Glazer, A. (1981) "Congestion Tolls and Consumer Welfare", *Public Finance*, 36(1), 77-83.

Kreukels, T. and Wever, E. (1998) "North Sea Ports in Competition", in Kreukels, T. and Wever, E. (eds.), *North Sea Ports in Transition: Changing Tides*, Royal Van Gorcum.

Levine, M.E. (1969) "Landing Fees and the Airport Congestion Problem", *Journal of Law and Economics*, 12, 79–108.

Levinson, D. (2000) "Revenue Choice on a Serial Network", *Journal of Transport Economics and Policy*, 34, 69-98.

Levinson, D. and Smalkoski, B. (2003) "Value of Time for Commercial Vehicle Operators in Minnesota", TRB International Symposium on Road Pricing, University of Minnesota.

Lindsey, R. (2007) "Congestion Relief: Assessing the Case for Road Tolls in Canada", C.D. Howe Institute Commentary, No.248, C.D. Howe Institute, Toronto.

Liu, L.N. and McDonald, J.F. (1998) "Efficient Congestion Tolls in the Presence of Unpriced Congestion: A Peak and Off-peak Simulation Model", *Journal of Urban Economics*, 44, 352–366.

Liu, L.N. and McDonald, J.F. (1999) "Economic Efficiency of Second-best Congestion Pricing Schemes in Urban Highway Systems", *Transportation Research*, 33B, 157-188.

Midoro, R., Musso, E. and Parola, F. (2005) "Maritime Liner Shipping and the Stevedoring Industry: Market Structure and Competition Strategies", *Maritime Policy and Management*, 32(2), 89-106.

Morrison, S.A., and Winston, C. (1990) "The Dynamics of Airline Pricing and Competition", *American Economic Review*, 80(2), 389-393.

Niskanen, E. (1987) "Congestion Tolls and Consumer Welfare", *Transportation Research Part B*, 21B(2), 171-174.

Notteboom, T.E. (1997) "Concentration and the Load Center Development in the European Container Port System", *Journal of Transport Geography*, 5, 99-115.

Notteboom, T.E. (2004) "Container Shipping and Ports: An Overview", *Review of Network Economics*, 3(2), 86-106.

Oxford Economic Forecasting (2006) "The Economic Contribution of the Aviation Industry in the UK", Oxford Economic Forecasting, Oxford.

Park, C. (2005) "Who Will Pay to Build Bigger Highways?", *Business Press*, San Bernardino, California.

Paul, J. (2007) "Wave of Consolidation – Ever Bigger Container Ships", *Business Lines*, Chennai, India.

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.

Productivity Commission (2002) *Price Regulation of Airport Services*, Report No. 19, AusInfo Canberra, Australia.

Proost, S. and Sen, A. (2006) "Urban Transport Pricing Reform with Two Levels of Government: A Case Study of Brussels", *Transport Policy*, 13, 127-139.

Slack, B. and Fremont, A. (2005) "Transformation of Port Terminal Operations: From the Local to the Global", *Transport Reviews*, 25(1), 117-130.

Small, K.A. (1992) *Urban Transportation Economics*, Harwood Academic Publishers, Chur, Switzerland.

Small, K.A. and Verhoef, E.T. (2007) *The Economics of Urban Transportation*, Routledge, London.

Small, K.A. and Yan, J. (2001) "The Value of 'Value Pricing' of Roads: Second-Best Pricing and Product Differentiation", *Journal of Urban Economics*, 49, 310-336.

Spengler, J. (1950) "Vertical Integration and Anti-trust Policy", *Journal of Political Economy*, 58, 347-352.

Starkie, D. (2001) "Reforming UK Airport Regulation", *Journal of Transport Economics and Policy*, 35, 119-135.

Transport Canada (2006) *The Cost of Urban Congestion in Canada*, Transport Canada, Environmental Affairs, Ottawa.

Ubbels, B., and Verhoef, E.T. (2008) "Governmental Competition in Road Charging", *Regional Science and Urban Economics*, 38, 174-190.

US DOT (2003) *Revised Departmental Guidance: Valuation of Travel Time in Economic Analysis*, U.S. Department of Transportation, Washington, DC.

US DOT (2006) "National Strategy to Reduce Congestion on America's Transportation Network", U.S. Department of Transportation, Washington, DC.

US GAO (2003) *Freight Transportation: Strategies Needed to Address Planning and Financing Limitations*, U.S. General Accounting Office, Washington, DC.

Van Dender, K. (2005) "Duopoly Prices under Congested Access", *Journal of Regional Science*, 45, 343-362.

Van Klink, H.A. and Van den Berg, G.C. (1998) "Gateways and Intermodalism", *Journal of Transport Geography*, 6(1), 1-9.

Vasigh, B. and Haririan, M. (1996) "Airport Privatization: A Simple Welfare Analysis", *Journal of Economics and Finance*, 20, 89-95.

Verhoef, E.T. (2002) "Second-best Congestion Pricing in General Static Transportation Networks with Elastic Demands", *Regional Science and Urban Economics*, 32, 281-310.

Verhoef, E.T. (2007) "Second-best Road Pricing through Highway Franchising", *Journal of Urban Economics*, 62, 337-361.

Verhoef, E.T., Nijkamp, P. and Rietveld, P. (1996) "Second-best Congestion Pricing: The Case of an Untolled Alternative", *Journal of Urban Economics*, 40(3), 279-302.

Verhoef, E.T. and Rouwendal, J. (2004) "Pricing, Capacity Choice and Financing in Transportation Networks", *Journal of Regional Science*, 44, 405-435.

Verhoef, E.T. and Small, K.A. (2004) "Product Differentiation on Roads: Constrained Congestion Pricing with Heterogeneous Users", *Journal of Transport Economics and Policy*, 38, 127-156.

Zhang, A. (2007) "Congestion Pricing and Capacity Investment in a Gateway-Intermodal System", Paper presented at the International Conference on Gateways and Corridors, Vancouver, B.C.

Zhang, A. and Zhang, Y. (2006) "Airport Capacity and Congestion When Carriers have Market Power", *Journal of Urban Economics*, 60, 229-247.

Chapter 4 Effects of Competition and Policy Changes on Capacity Utilization Efficiency of Chinese Airports: An Empirical Investigation[67]

4.1 Introduction

In the past few decades, rapid economic growth has significantly increased the demands for air services in China; between 1980 and 2005, the number of air passengers and cargo tonnage grew at an average rate of 16.8 percent and 18.2 percent per annum, respectively. This increasing air demand has placed enormous pressure on China's airport infrastructure. The situation is expected to get worse, as air travel is forecasted to grow at the still fast rate of 7.4 percent per year for the Chinese market over the next twenty years.[68] Thus, in addition to the infrastructure investment, there is an urgent need for Chinese airports to utilize their existing capacities more efficiently in order to relieve the pressure. Furthermore, as the liberalization of the airline industry continues, more foreign airlines will be allowed to operate in China, and will have increasing freedom to choose where they base their gateways in China. This would also put pressure on Chinese airports to further improve their own efficiency,[69] as the airlines want to locate at efficient airports in order both to reduce their operating costs and to improve the quality

[67] A version of this Chapter is forthcoming as: Yuen, A. and Zhang, A. (2009) "Effects of Competition and Policy Changes on Chinese Airport Productivity: An Empirical Investigation", *Journal of Air Transport Management.*

[68] A forecast made by Boeing (http://www.boeing.com/companyoffices/aboutus/boechina.html).

[69] In this Chapter, we consider the "efficiency" as the efficiency of physical capital utilization. In airport operation, airports may use other inputs, for example, personnel and outsourcing activities, to produce their outputs. The efficiency of utilizing those inputs is not considered in this Chapter.

of their service. Thus an empirical investigation of factors affecting Chinese airport efficiency has become important.

Assessment of airport efficiency has become the focus of a large number of studies. Different methodologies have been used to measure the efficiency of airports in different regions around the world (see Oum, et al., 2003, for a comprehensive review). Due to a lack of data, however, it is difficult to assess airport efficiency in China. A recent paper by Fung, et al. (2008) attempted to calculate the productivities for twenty-five major Chinese airports between 1995 and 2004. They found that over that period, airport efficiency was improving and the productivity among airports from different regions was converging. Using their data Zhang and Yuen (2008) further investigated whether privatization through public listing improves airport performance. Although they found a positive and statistically significant relationship between Chinese airport efficiency and public listing, a large portion of the variance in efficiency and its growth are still left unexplained by their regression models. Furthermore, their panel dataset does not capture the effects of policy changes on Chinese airport efficiency after 2004, during which several important industry reform measures have taken place.

This Chapter investigates the effects of China's competition and aviation policy reform (for example, the airport localization program in which Chinese airports have been transferred from the central government to its local counterparts, and listing airports on stock markets) on the efficiency of Chinese airports. Our sample data consist of a panel of twenty-five major airports for the period from 1995 to 2006. This new dataset

may provide a better basis for investigating the effectiveness of recent policy changes on improving the Chinese airport efficiency. In particular, we use Data Envelopment Analysis to compute efficiency scores for each airport. We then run regressions to examine the effects of the competition and aviation policy reform on the efficiency scores by controlling a set of airport characteristics and event variables.

Our empirical results reveal that airport localization has a strong impact on airport efficiency; the efficiency of the localized airports is significantly higher than that of their counterparts. Furthermore, there is statistically significant evidence suggesting that airports with more competition are more efficient than their counterparts. There is also strong evidence that publicly listed airports are significantly more efficient than non-listed airports. We do not find, however, any statistically significant correlation between Chinese airport efficiency and two specific policy changes at the airline level, namely the signing "open skies" agreements, and airline mergers arranged by the China's State Council in 2003. Finally, we use the Malmquist index method to investigate the effects of the competition and aviation policy reform on *changes* in the efficiency of Chinese airports. We find that efficiency growth and its component, technical efficiency, do not have a statistically significant relationship with the airport localization program, competition intensity or stock market listing. However, technical progress – the other component of efficiency growth – is positively, and statistically significantly, correlated with the airport localization program dummy.

This Chapter is organized as follows. Section 4.2 describes recent policy changes in the Chinese aviation market and their potential effects on airport efficiency. Sections 4.3 and 4.4 discuss, respectively, the methodology and data employed by the Chapter. Section 4.5 reports the empirical results on airport efficiency levels, and Section 4.6 examines the changes in airport efficiency. Section 4.7 contains concluding remarks.

4.2 Recent Policy Changes and Airport Efficiency

As part of the general economic reform, the reform of the aviation industry in China began in the late 1970s (see Zhang, 1998; Zhang and Chen, 2003; and Zhang and Yuen, 2008, for reviews). The "Report on Civil Aviation Reform Measures", which was passed by the State Council in January 1987, stated that the long-term goal of the industry reform was to separate the Civil Aviation Administration of China (CAAC) as the regulator from direct involvement in airline and airport operations. This goal would be achieved through the airport localization program, in which airports are turned over to local governments. As a pilot program of the airport localization program, operation of the Xiamen Airport and Shanghai Hongqiao International Airport (including all fixed and working capital and all personnel) was transferred to their municipal governments in 1988 and 1993, respectively. The CAAC, however, was still heavily involved in the late 1980s and 1990s. The localization program regained momentum in the early 2000s and was completed by 2003, when the CAAC transferred ownership and control of all its remaining airports, except the Beijing and Tibet Airports, to their respective local governments.

The airport localization program, on one hand, increased the initiatives for local and private investment in airport capacity expansion. On the other hand, airport efficiency was expected to improve after the implementation of the localization program. As pointed out by Zhang and Yuen (2008), as opposed to the "soft budget" approach taken by the CAAC, the localization program made the airports more financially accountable and consequently improves their efficiency.[70] Furthermore, as the efficiency of airports has significant implications for local economies, local governments may have greater incentives to improve their airport efficiency than would the CAAC.

The second recent policy change that may affect Chinese airport efficiency is allowing Chinese airports to be listed on stock markets. Although attracting private funds was one rationale for airport listing, the principal objective was to improve airport efficiency (Zhang and Yuen, 2008). Since the initial public offering (IPO) of Xiamen Gaoqi International Airport, six Chinese airport companies have been listed on stock exchanges in Hong Kong, Shanghai, and Shenzhen. In the literature, there are a number of studies empirically examining the performance of Chinese listed companies. Sun and Tong (2003) found that there was an improvement in state-owned enterprises' earnings ability, real sales and workers' productivity, but not in profit returns or leverage after listing. Wang (2005), on the other hand, found a sharp decline in post-issue operating performance of IPO firms. Zhang and Yuen (2008) investigated the effect of listing on Chinese airport efficiency, and found that the listed airports had higher efficiency scores

[70] As part of the localization program, the central government began to phase out its subsidization of airports in 2006.

than did unlisted airports, while the correlation between efficiency growth and listing was statistically insignificant.

The policy changes in the *airline market* may also affect Chinese airport efficiency. One of the prominent changes there is the opening of the market to foreign airlines gradually. For the past five years, China has been moving toward a more liberal international policy regime, which has significantly increased the number of international connections for China's airports (Zhang and Yuen, 2008). The bilateral "open skies" agreements may increase the passenger and cargo traffic at airports. Given a fixed amount of input, this will imply an efficiency improvement. This will, however, also increase the percentage of international traffic at the airports; and Oum, et al. (2003) found that the airports with heavy reliance on international passenger traffic had lower gross total factor productivities (TFPs) than average airports.

Another major change in the airline market is the consolidation of the airlines in China. In October 2002, under the State Council's arrangement, the three megacarriers – Air China, China Southern Airlines and China Eastern Airlines – took over fourteen relatively minor carriers (most of which were under the CAAC control). This substantial change in the market structure of the downstream carriers may have significant implications for Chinese airport efficiency. For example, after the mergers, the bargaining power of the three new airlines groups was likely strengthened in their negotiation with airports. Thus, they might be able to impose more pressure on airports

for further improvement of their productivities, leading to a possible reduction of airport charges.

4.3 Methodology

To investigate the effect of competition and policy changes on Chinese airport efficiency, we use a two-stage procedure. See, for example, Ali and Flinn (1989) and Kalirajan (1990) for an application of the two-stage analysis. In the first stage, we calculate the productive efficiency from 1995 to 2006 for each airport. In the second stage, we run regressions to examine the effects of competition and policy changes on the productive efficiency of airports, while controlling for a set of independent variables.

In particular, in the first stage, we need to calculate the productive efficiency of airports, which is reflected by the relationship between the outputs the airport produces and the inputs the airport uses in a given period of time. Empirical applications of the efficiency measurement are feasible by a non-parametric technique known as Data Envelopment Analysis (DEA).[71] A DEA model gives an efficiency score for each airport in each year. For the output-oriented model, the efficiency score has a value between zero and one. Airports with an efficiency score of unity are located on the frontier in the sense that their outputs cannot be further expanded without a corresponding increase in input. Airports with an efficiency score below one are inefficient. The DEA model defines the

[71] Useful references on DEA include Farrell (1957), Banker, et al. (1984, 1989), Charnes, et al. (1978, 1981), Seiford and Thrall (1990), and Lovell (1993).

efficiency score of any airport as the ratio of the airport's output to that can be produced for an airport on the efficient frontier with the same level of input.

The DEA approach is widely used in measuring the performance of airports. A number of advantages to estimate airport efficiency are discussed in the literature: (i) it does not require any assumption concerning either the technology or the behaviors of actors (for example, cost minimization) (Pels, et al., 2001); (ii) it can be done without some detailed operating information (such as input costs); (iii) it can handle multiple outputs.[72] Gillen and Lall (1997) applied DEA to assess terminal and airside operations of twenty-one top US airports from 1989 to 1993. This is followed by a number of papers using DEA to evaluate performance of airports in different countries (for example, Parker, 1999; Sarkis, 2000; Vasigh and Hamzaee, 2000; Chin and Siong, 2001; Martin and Roman, 2001; Pels, et al., 2001; Abott and Wu, 2002). In order to measure Chinese airport efficiency, Fung, et al. (2008) applied DEA to assess productive efficiency for twenty-five major Chinese airports over the 1995-2004 period. They focused primarily on whether airport efficiency was improving over the time period and whether efficiency among the airports from different regions was converging, and found positive answers to both questions.

[72] The DEA also has some well-known weaknesses, especially its estimates' sensitivity to outliers. Instead, other methodologies, e.g., stochastic frontier analysis (SFA), are also used for measuring airport efficiency. However, input prices are necessary for the SFA, while the information is not available for Chinese airports.

In the second stage, in order to investigate the influence of competition and aviation policy reform in China on the efficiency of Chinese airports, we consider the efficiency scores as a function of:

- the airport localization program;
- competition intensity;
- public listing;
- airport characteristics; and
- other important events in the Chinese aviation market.

In particular, we run the following ordinary least squares (OLS) regression to examine the effect of the airport localization program (*LOCAL*), competition intensity (*d*), and public listing (*LIST*) on the productive efficiency of airports while controlling for airport characteristics and other factors:

$$\theta = \alpha + \beta_1 LOCAL + \beta_2 d + \beta_3 LIST + \beta_4 (LOCAL)\mathrm{X}(LIST) + \sum_{i=1}^{16} \beta_{5i} X_i + \varepsilon, \qquad (4.1)$$

where θ is the DEA efficiency scores; *LOCAL* is a dummy for the airport under airport localization program; *d* is a proxy for the competition intensity. In (4.1), we also take into account the interaction between the localization program and listing. *X* is a set of controlling variables, including a dummy variable for airline mergers in 2003, hub dummy, congestion dummy, dummies for tourist city and coastal city, city population,

provincial per-capita gross domestic product (GDP), and input and output indexes. Additionally, we also include the "Guangzhou New Airport" dummy and "Shanghai Pudong" dummy to capture the effects of opening new airports in Guangzhou and Shanghai in 2004 and 1999, respectively. The next section will further discuss how we construct the variables.

It is notable that as the efficiency scores have an upper bound of one, there may be a truncation bias in the OLS regression model. Thus, as in Gillen and Lall (1997), we also run the Tobit regression model (Tobin, 1958). Our Tobit model can be represented as follows:

$$\hat{\theta} = \beta_1 LOCAL + \beta_2 d + \beta_3 LIST + \beta_4 (LOCAL) X (LIST) + \sum_{i=1}^{16} \beta_{5i} X_i + \varepsilon,$$

$$\theta = \begin{cases} 1 & if \ \hat{\theta} \geq 1 \\ \hat{\theta} & if \ \hat{\theta} < 1 \end{cases} \tag{4.2}$$

where the variables in the Tobit regression model are the same as that in our OLS regression model.

4.4. Sample Airports and Variable Construction

4.4.1 Sample airports

We consider a panel dataset for twenty-five major Chinese airports in the period between 1995 and 2006. Our sample includes airports in China that represent different ownerships, operational characteristics, and regional locations (as shown in Table 4.1). The data are compiled from various sources, including Statistical Data on Civil Aviation of China, Chinese Statistics Yearbooks (various years) and company annual reports of listing airports.

[Table 4.1 here]

To measure the efficiency by DEA (i.e., the efficiency scores), one must first identify outputs that an airport produces and the inputs it uses in producing those outputs. The summary statistics for the inputs and outputs that we consider for DEA is given in Table 4.2. In particular, on the input side, we consider two physical capital input measures: the number of runways and terminal size. In our sample, the Beijing and Guangzhou Airports had two runways in 2006, while others had one each. The terminal size ranged from 2,500 square meters (the Hailar Airport) to 336,000 square meters (the Beijing Airport). It is noted that operational costs, the number of employees, the number of gates, and other soft-cost inputs are also considered as inputs in some other DEA airport efficiency studies (for example, Gillen and Lall, 1997; Sarkis, 2000; Pels, et al.,

2001). The data for these input measures are not available for most Chinese airports, however.

[Table 4.2 here]

On the output side, we consider passenger volume, air cargo volume, and the number of aircraft movement as outputs of airports in our DEA. The number of passengers served is the most commonly used output for measuring airport efficiency. In China, most airports primarily serve passenger traffic; during the study period (i.e., 1995-2006), the growth of the average passenger volume of our sample airports increased by 210.4 percent. We also consider cargo traffic as another output, as its importance has increased over the years (Zhang, et al., 2004). For instance, the average cargo traffic increased substantially, by 308.2 percent, in the study period. The third output is aircraft movement, which also grew rapidly during the study period (i.e., by 205.5 percent). Several other measures, including operating revenues and general aviation, have been considered in the literature (for example, Sarkis, 2000). Another important output for airport performance measurement is non-aeronautical revenue (Zhang and Zhang, 1997; Oum, et al., 2003). In addition to their aviation business, airports also derive revenue from concessions, retailing, advertising and other services. The proportion of non-aeronautical revenue to the total revenue varies significantly among airports around the world. For example, in 2003, non-aeronautical revenue accounted for only 27 percent of New York John F. Kennedy International Airport's total revenue, while it constituted 68 percent of total revenue at the Brisbane Airport (Oum, et al., 2003). In China, the portion

of non-aeronautical revenue for airports is still relatively low. For example, it only contributed 27.3 percent to the total revenue of the Beijing Airport in 2006. Nevertheless, the non-aeronautical business is becoming a more important source of revenue for Chinese airports (for example, in 2006, 47.3 percent of total revenue at the Guangzhou Airport was from its non-aeronautical business). Although we cannot include the non-aeronautical revenue as an output measure in our DEA due to the data limitation, it would be worthwhile to include the output measure in the future studies on Chinese airport performance if the data are available.

4.4.2 Variable construction

In the second stage, we run the regression models in (4.1) and (4.2) to examine the effects of the airport localization program, competition intensity, and public listing on the airport efficiency, while controlling for airport characteristics and other factors. The independent variable is the efficiency score θ, obtained from DEA in the first stage. Three explanatory variables in our regression models are the airport localization dummy (*LOCAL*), the listing dummy (*LIST*), and competition intensity (d). The dummy variable for airport localization (*LOCAL*) is one for airports in a particular year if under the local government ownership and control, and zero otherwise. An exception is Beijing Capital International Airport, which is excluded from the CAAC's localization program. In constructing the airport localization dummy variable, we consider the Beijing Airport to be one for all years. Since the CAAC has direct control over the airport, potential agency problems between CAAC and local governments would not arise for the Beijing Airport.

The listing dummy (*LIST*) is one for airports in a particular year if the airport is listed on stock exchanges, and zero otherwise. Table 4.3 shows the airports in our sample that are listed on different stock exchanges in different years. In addition, we consider a variable for interaction between listing and the localization program. After the localization program is implemented, local governments are the major shareholders of their airports. If the airports are public listed, the local governments may have more incentive to improve the airports' performance, as it may imply an increase in their stock prices.

[Table 4.3 here]

In order to investigate how competition affects airport efficiency, we need to construct a proxy for the intensity of airport competition. The most commonly used proxy for market competition is the Herfindahl index (HHI). For example, Zhang, et al. (2001) quantified the effect of market competition (and also ownership) on the productive efficiency and efficiency growth of Chinese industrial firms by using HHI as a proxy for market competition. The construction for HHI for our analysis is less feasible. In practice, airports serve many different markets (for example, long- and short-distance markets, transshipment markets, and different origin-destination markets) through air carriers. Thus to construct HHI for each airport, it requires very comprehensive market data in the whole study period, which are not available for the Chinese aviation market. Consequently, we instead use another proxy for the intensity of airport competition: (log) distance of a particular airport with another nearest airport in our sample airports, or the

Hong Kong and Incheon Airports, which may be competing with our sample airports.[73] Thus, we consider an airport to be facing more competition if there is another one in proximity. For example, the distance between the Shenzhen and Hong Kong Airports is only about thirty-two kilometers, indicating that the Shenzhen Airport is considered to be facing strong competition from the Hong Kong Airport in our analysis.

To control for airport characteristics and other potential factors to explain the Chinese airport efficiency, we further consider sixteen controlling variables in our regression models as follows.

Airport Characteristics

We consider twelve variables that capture the major airport characteristics of our sample airports:

- *Hub Status*: We control for the size and location advantages possessed by hub airports, and thus the expected sign of the effect of hub status on efficiency is positive. Specifically, we have the hub dummy for the international and regional hubs. As in Zhang and Yuen (2008), the Beijing, Shanghai and Guangzhou Airports are considered as the three international hubs, while the Chengdu,

[73] In addition to the logarithm of distance, we also used distance as a proxy for competition intensity. Our major results remain unchanged. On the other hand, a more accurate measure of accessibility than either distance or the logarithm of distance could be derived by identifying actual travel distances or travel times by rail or road. But these values are not available for our sample airports.

Kunming, Shenyang, Shenzhen, Urumqi, and Xian Airports are the six regional hubs in our regression models.

- *Congestion*: We consider a dummy for congested airports. Oum, et al. (2003) found empirical evidence to support the conjecture that airports with severe capacity shortage may be forced to find ways to be more efficient. Thus, the expected sign of the effect of congestion on efficiency is positive. While there are no flight delay data for Chinese airports, we consider the following index for congestion:

$$C = \frac{Average\ daily\ aircraft\ movement}{Runway\ Capacity} \tag{4.3}$$

The average daily aircraft movement in (4.3) is the total aircraft movement in year t divided by 365. We assume that the capacity for a runway is 35 aircraft movements per hour and an airport operates 12 hours per day. Thus the runway capacity in (4.3) is the number of runway in year t times 420. Note that for an airport in which the average daily aircraft movement is less than runway capacity (i.e., $C < 1$), it may still be subject to congestion in peak periods. In our empirical analysis, we consider the congestion dummy is equal to one if $C > 0.8$, and zero otherwise. The threshold for congestion dummy is relatively arbitrary. Thus we will check the

robustness of our empirical results by considering other thresholds (e.g., 0.7 and 0.9), and see whether our results are sensitive to the threshold selection.[74]

- *Local Economy*: The provincial per-capita GDP of the province an airport located is used as a proxy for the local economy.

- *Coastal City*: We consider a dummy for airports located at coastal cities. Since 1978, the economy at the coastal provinces in China has been growing at a much faster rate than its inland counterparts; Jones, et al. (2003) found that annual growth rates of coastal cities were, on average, three percentage points higher than they were in non-coastal cities. The difference in economic development between the coastal provinces and non-coastal provinces may explain the difference in enterprise productivity in China. For example, Fleisher and Chen (1997) showed that the TFP was roughly twice as high in the coastal provinces as they were in non-coastal provinces; they demonstrated that investment in higher education and foreign direct investment helped explain this productivity gap. Thus, the expected sign of the effect of the costal city dummy on efficiency is positive. As did Fleisher and Chen, we also consider Beijing to be a coastal city.

- *Tourist City*: A tourist city dummy is introduced to reflect particular airports with a high percentage of leisure passengers. In our analysis, we consider the Sanya and Kunming Airports are located at tourist cities.

[74] Although some function of the ratio of movements to capacity can be used to construct a continuous variable for congestion, there is no priori reason to choose a particular functional form for that.

- *Population*: The population of the city where the airport is located is used as a proxy for the market size served by the airport.

- *Demand and Supply Shocks*: Airports are a capital-intensive industry and the capital investment in runways and terminals are largely indivisible (see for example, Oum and Zhang, 1990; and Zhang and Zhang, 2003). This characteristic can affect the role that the yearly efficiency scores play in measuring airport performance. In particular, large and lumpy investments can deteriorate the efficiency measurement because they constitute a large increase in input. Similarly, demand shocks might significantly alter the efficiency scores through a large output expansion. To control for these potentially external shocks, we consider the following input and output indexes:

$$InputI^t(x^t) = \frac{x^t}{x^{1995}}, \qquad OutputI^t(y^t) = \frac{y^t}{y^{1995}} \tag{4.4}$$

where x^t (y^t) is the input (output) level at time t, with $InputI^{1995}(x^{1995}) = 1$ and $OutputI^{1995}(y^{1995}) = 1$. The input variables are runway and terminal area indexes, whereas the output variables are passenger, cargo and aircraft movement indexes.

We expect that the signs of the effects of input indexes on efficiency are negative, while those of output indexes are positive.[75]

Event Variables

In addition to airport characteristics, we also control for four important events in the Chinese aviation market that may have potentially affected the efficiency of our sample airports during the study period:

- *Airline Mergers (AMERGER)*: As discussed in Section 4.2, the airline mergers in China may have affected airport efficiency; we consider a dummy for the airline-mergers event.

- *Open Skies Agreements (OPENS)*: In our empirical analysis, we particularly focus on two major air service agreements: the liberal air service agreements signed between China, and the U.S. and Hong Kong in July and September 2004, respectively. It is because the two regions have close economic relationships with China; the U.S. and Hong Kong were the largest and third largest trading partners of China in 2005, respectively. As international hub airports are more likely to be affected by the open skies agreements, we consider a dummy (*OPENS*) for the

[75] The input and output indexes in (4.4) look similar in construction to DEA efficiency score itself, which may raise concerns about the correlation between the indexes and the error term in (4.1). Thus, we also run regression of the residuals from our OLS model given in (4.1) on the indexes, and found no statistically significant relationships among them.

international hubs in 2005 and 2006.[76]

- *Guangzhou New Airport Dummy*: A new Guangzhou Baiyun Airport was opened in August 2004. Compared to the old airport, the terminal increased 3.8 times in size; additionally, the new airport operates one more runway than did the old airport. This may suggest a structural change in its operation. We capture this potential effect on its efficiency by introducing a dummy for the new airport, and expect that the sign of the effect of the dummy on efficiency is negative.

- *Shanghai Pudong Dummy*: In 1999, Shanghai Pudong International Airport was opened and replaced Shanghai Hongqiao International Airport (our sample airport) as Shanghai's international airport, and took over all international flights, including regional flights to Hong Kong and Macau. Consequently, in 2000, the passenger traffic of Hongqiao Airport was decreased by 15.4 percent, while that of our sample airports was increased by 6.4 percent. To control for the demand shock, we add a dummy to capture this structural change in demand at Hongqiao Airport. We expect that the sign of the effect of the dummy on efficiency is negative.

[76] To check the robustness of our results, we also use a dummy for all airports that were affected by the open skies agreements and find that our main results remained unchanged.

4.5 Empirical Results

In this section, we discuss the results for our two-stage analysis. First, we report the DEA efficiency scores for our twenty-five sample Chinese airports in the period between 1995 and 2006. In addition to this, we will investigate the influence of competition and aviation policy reform in China on the efficiency of Chinese airports through a discussion of the results from different regression models.

4.5.1 Efficiency scores

The results of the DEA efficiency scores of the twenty-five sample airports in the period between 1995 and 2006 are shown in Table 4.4. The scores of the airport in the first nine years (i.e., 1995 to 2004) are also reported in Fung, et al. (2008). Yet, it should be noted that different from Fung, et al., we consider the number of runway as one of the inputs measures rather than the runway length. Furthermore, we adjust for the input changes due to the opening of the new Guangzhou Baiyun Airport in 2004.[77] With the new dataset, we find that the pooled average efficiency score of the period in 2004-2006 increased by 6 percent, compared to that in the period of 1995-2003.

Additionally, the Beijing, Shanghai, Shenzhen and Chongqing Airports were considered to be efficient, while the Kashi, Hailar, Hohhot, Lanzhou and Harbin Airports

[77] In Fung, et al. (2008), the terminal size and runway length of the old airport are used as the inputs of the Guangzhou Airport for calculating the efficiency scores for the airports in that year, while we adjust the input on the pro-rata basis.

were the least efficient airports in the last three consecutive years (2004-2006). We also observe that among our sample airports during that period, Guangzhou Baiyun Airport recorded the greatest drop of efficiency scores, from 1.000 in 2004 to 0.617 in 2006. The reason for this drop in efficiency is most likely because the inputs (i.e., the number of runway and terminal size) of the airport were increased significantly after the opening of the new airport. As discussed above, the capital investments in airports are largely indivisible. This drop in efficiency score, therefore, may not necessarily indicate that the airport become less efficient. This potential problem will be controlled for in our regression analysis.

4.5.2 Regression results

In the second stage, we run the OLS model given in (4.1) and the Tobit model given in (4.2) to examine the effects of the airport localization program, competition intensity, and public listing on the productive efficiency of airports while controlling for airport characteristics and other factors. The regression results are presented in Table 4.5. The adjusted R-square of the OLS regression model is 0.71, which means that our model can explain 71 percent of the variation of the airport efficiency in our panel dataset. Thus there is a significant improvement in explanatory power of our OLS model, comparing with that in Zhang and Yuen (2008), in which the adjusted R-square is from 0.29 to 0.41.

[Table 4.5 here]

As expected, we find that the relationship between airport efficiency and the (log) distance with the nearest airport (a proxy for airport competition) is statistically correlated in the OLS model and the Tobit model. The negative sign of the coefficients suggests that airports with more competition are more efficient than their counterparts. On the other hand, both the OLS model and the Tobit model show that the efficiency scores are positively correlated with the airport localization program dummy and that the coefficients are statistically significantly different from zero. The result suggests that the airport localization program effectively improved Chinese airport efficiency. It also provides an empirical evidence to support the argument made in Zhang and Yuen (2008) that, by making the airports more financially accountable through the localization program, the airports would improve their efficiency.

The efficiency scores are positively correlated with public listing and the coefficients are statistically significant. This result is consistent with Zhang and Yuen (2008). It should be cautioned, however, that the correlation result we find in the regression does not imply causation between the two variables. As also noted in Zhang and Yuen, this listing effect on efficiency level does not necessarily imply that privatization through public listing improves airport performance; the positive listing effect might arise because the Chinese government chose the most efficient airports to be listed (see more discussion about the endogenity issues in Appendix C). As a result, it would be useful for us to further look at whether listing can effectively improve the efficiency growth of Chinese listed airports in the next section. We also find that the coefficient for the interaction between the airport localization program and listing is

statistically significant in the Tobit model. The positive sign of the coefficient suggests that after the localization program was implemented, listed airports are more efficient than their counterparts. However, this positive correlation is found to be statistically insignificant in the OLS model.

For the airport characteristics variables, the impact of hub status – whether airports are international or regional hubs – on the efficiency score is positive and statistically significant. The results may suggest that hub airports possess the size and location advantages. On the other hand, the relationship between efficiency and congested dummy is positive and statistically significant.[78] This result is consistent with Oum, et al. (2003). We also find that the airport location and market size may explain the productive efficiency of our sample airports; the efficiency score is positively correlated with the coastal city dummy and city population. A negative relationship between the efficiency score and provincial per-capita GDP, however, is found, and is statistically significant. It would probably be better to use the per-capita GDP at the *city* level (rather than at the *provincial* level) but unfortunately, the complete data were not available for the sample airports. Finally, the results suggest that the efficiency score is negatively correlated with the terminal area index and that the coefficient is statistically significant, while the coefficient of the aircraft movement index is positive, and statistically significant.

[78] It is argued that airports with capacity shortages may be forced to operate more efficiently. However, causality could also run in the opposite direction. For example, an airport with inefficient air traffic control or poor performance at turning flights around may be prone to congestion delays. As a result, endogenity issues may arise. Given this, we ran regression models excluding the congestion variable, and found that the empirical results related to the policy dummies and competition intensity – the major focus of this Chapter – remain unchanged.

We did not find any statistically significant correlation between Chinese airport efficiency and the four event variables we controlled for, namely the signing open-skies agreements, airline mergers and the opening of new Guangzhou and Shanghai Pudong Airports. This may be because, first of all, the effects have already been captured by other control variables including the input and output indexes. For instance, the significant change in the number of runway and terminal size after the opening of new Guangzhou Airport have already been reflected in the runway and terminal area indexes. Second, it is also possible that the period of our panel dataset is not long enough to fully reflect the effects of the events. For instance, the Sino-US air service agreement signed in July 2004 will double the number of airlines that can fly between the countries and will permit a nearly five-fold increase in Sino-US air services over the next six years from 2004. Since our dataset, however, only covered the first two years of the agreement, it may not fully reflect the impact of the agreement on Chinese airports and their productivities.

4.6 Growth in Efficiency

Having examined levels of efficiency for the sample airports, we will now turn to the changes in levels of efficiency. This examination is useful in that, if the low level of efficiency at some airports is due to their low starting point, a faster growth rate in efficiency after the policy changes could reduce and eliminate the gap. For instance, Fung, et al. (2008) found that efficiency among the airports from different regions was converging. The efficiency growth is measured by the Malmquist index:

$$M_O^{t+1}(x^t, y^t, x^{t+1}, y^{t+1}) = \frac{D_O^{t+1}(x^{t+1}, y^{t+1})}{D_O^t(x^t, y^t)} \times \left[\frac{D_O^t(x^{t+1}, y^{t+1})}{D_O^{t+1}(x^{t+1}, y^{t+1})} \times \frac{D_O^t(x^t, y^t)}{D_O^{t+1}(x^t, y^t)} \right]^{\frac{1}{2}}, \qquad (4.5)$$

where D_O is an output distance function of airport O. The distance function is the inverse of the output-oriented efficiency score, which we calculated in Section 4.5. The superscripts on D_O indicate the time periods within which the efficiency scores are calculated. The superscripts on x and y indicate the time periods of the data used in the calculation of the efficiency scores. As a measure of the overall efficiency change, a Malmquist index M_O^{t+1} greater (less) than unity indicates that the overall efficiency of airport O has increased (declined) from period t to period $t+1$.

Note that equation (4.5) also represents a decomposition of efficiency change from period t to period $t+1$. The ratio outside the bracket on the right-hand side measures the change in "technical efficiency (*EFFECH*)" of airport O from period t to period $t+1$. Greater (smaller) than unity implies that the technical efficiency has improved (declined) in reference to the production frontier from period t to period $t+1$. The bracketed term represents the geometric mean of the shift in production frontier. When the value of this term is greater (less) than unity, it implies that the "technology (*TECH*)" of the industry has progressed (regressed) from period t to period $t+1$.

Table 4.6 presents the descriptive statistics for efficiency growth rates and its two components. We find that Chinese airport efficiency grew at a faster rate in the period between 2004 and 2006 than it did prior to 2004; the mean of the Malmquist index in the period is 1.199, while the mean in the period between 1996 and 2003 is only 1.048. The increase in efficiency growth rates is mainly due to the improvement in the technological change of the industry, as the mean of technical progress increased from 1.009 to 1.133 in the period between 2004 and 2006. The results suggest that the overall Chinese airport efficiency has improved, which might be attributed to the recent policy changes in the Chinese aviation market.

[Table 4.6 here]

We run the OLS regression model to investigate the effect of competition and policy reform on the growth of efficiency and its two components. All dependent variables are the same as in Section 4.5, except the input and output indexes for demand and supply shocks. In this section, we will consider indexes for the growth rates of the input and output variables. Table 4.7 presents the regression results of efficiency growth rates. The adjusted R-square of the OLS regression is 0.76, implying our model can explain 76 percent of the data variations of the airport efficiency growth.

[Table 4.7 here]

The OLS regression results show that the efficiency growth does not have a statistically significant relationship with the airport localization program dummy, the (log) distance with the nearest airport (a proxy for airport competition) or public listing. For the airport characteristics variables, we find that the efficiency growth is positively correlated with the coastal city dummy but the coefficient is statistically insignificant. Furthermore, the coefficient of terminal area growth index is negative and statistically significant. At the same time, the efficiency growth is positively correlated with the cargo growth and aircraft movement growth indexes and the coefficients are statistically significant. For the event variables, we find that the relationship between the efficiency growth and Guangzhou new airport dummy is negative and statistically significant, while we do not find any statistically significant correlation between the Chinese airport efficiency growth and other three events which we controlled for.

The regression results of technical efficiency (*EFFECH*) in Table 4.7 show that we do not find any statistically significant correlation between the technical efficiency improvement and the airport localization program dummy, the (log) distance with the nearest airport (a proxy for airport competition), or public listing. We, however, find that the technical efficiency improvement is negatively correlated with the terminal area growth index and the coefficient is statistically significant. At the same time, the coefficient of the aircraft movement growth index is positive and statistically significant. Finally, as for the event variables, we do not find any statistical correlation between the technical efficiency improvement and the four events we controlled for.

The OLS results in Table 4.7 show that the technical progress (*TECH*) is positively correlated with the airport localization dummy, and statistically significant. This result suggests that the airport localization program has a positive impact on the technical progress of the industry. This may be due to the fact that, besides the Shanghai and Xiamen Airports, airport controls were transferred from the central government to the provincial or municipal governments almost at the same time. As a result, after the completion of the airport localization program in 2003, the efficiency of most Chinese airports has been improved. Thus, we observe technical progress in the industry in the period after 2003. However, we do not find any statistically significant correlation between technical progress and the (log) distance with the nearest airport or public listing. For the airport characteristics variables, the coefficient of the terminal area growth index is negative and statistically significant. At the same time, the coefficient of the aircraft movement growth index is positive and statistically significant. Again, for the event variables, we do not find any statistical correlation between the technical progress and the four events we controlled for.

4.7 Concluding Remarks

The main purpose of this Chapter is to assess the effects of China's competition and aviation policy reform on both the level and growth of its airports efficiency. In particular, the newly available panel dataset enabled us to investigate the effect of recent policy changes, including airport listing and the airport localization program, on airport efficiency. We found that there is strong evidence that publicly listed airports are

significantly more efficient than non-listed airports. This result, however, requires special attention, as it does not necessarily imply that public listing can cause an improvement of airport efficiency (as also discussed in Zhang and Yuen, 2008). Furthermore, by using a larger panel dataset, we found results similar to Zhang and Yuen, that the correlation between listing and efficiency growth is statistically insignificant.

Our empirical results further supported the argument made by Zhang and Yuen (2008) that, as airports would be financially more accountable after the airport localization program, the program could result in an improvement in the efficiency of airports. There is also some evidence suggesting that airports with more competition are more efficient than their counterparts. Finally, the policy changes at the airline level, namely the airline mergers and signing open-skies agreements, do not seem to have a statistically significant correlation with the Chinese airport efficiency.

Due to data limitation, we only considered two inputs and three outputs of Chinese airports when estimating the efficiency scores by DEA. In practice, different airports are very different in their operating characteristics and in their services provided. Exclusion of some inputs and outputs may yield biases in measuring airport efficiency. For example, if the so-called soft-cost input, which is measured by all expenses not directly related to capital and personnel, is not included, the efficiency measurement may favor the airports that outsourced most of their services (Oum, et al., 2003). On the other hand, as discussed in Section 4.3, non-aeronautical businesses are getting more important for Chinese airports. Exclusion of non-aeronautical revenue as an output may bias

efficiency against the airports that generate more revenues from non-aeronautical businesses. Further research may need to consider extending our analysis by using a larger set of inputs and outputs in measuring the airports' efficiency.

Table 4.1 Sample airports and operating characteristics in 2006

	Region	Number of runway	Terminal size (square meter)	Passenger volume (person)	Cargo volume (ton)	Aircraft movement (flight)
Beijing Capital International	Northern	2	336,000	48,748,298	1,148,259	378,888
Changsha Huanghua	Central and South	1	34,000	6,592,602	60,094	71,139
Chengdu Shuangliu International	Southwest	1	92,000	16,280,225	285,127	155,484
Chongqing Jiangbei International	Southwest	1	13,321	8,050,007	116,512	88,929
Dalian Zhoushuizi International	Northeast	1	33,000	6,351,089	105,838	56,374
Guangzhou Baiyun International	Central and South	2	320,000	26,222,037	626,598	232,404
Hailar Dongshan	Northern	1	2,490	150,561	606	2,584
Harbin Taiping International	Northeast	1	67,000	3,643,232	41,540	33,863
Hefei Luogang	Eastern	1	13,835	1,851,464	20,284	24,000
Hohhot Baita	Northern	1	19,339	1,509,643	8,883	21,468
Jinan Yaoqiang	Eastern	1	11,000	3,696,305	36,269	41,901
Kashi	Northwest	1	7,931	444,332	1,294	4,017
Kunming Wujiaba International	Southwest	1	76,700	14,443,607	215,389	135,573
Lanzhou Zhongchuan	Northwest	1	27,000	1,861,148	13,253	21,902
Nanning Wuxu	Central and South	1	25,885	2,244,234	26,668	27,248
Qingdao Liuting	Eastern	1	15,650	6,791,240	97,835	72,008
Sanya Fenghuang	Central and South	1	18,230	3,905,956	23,827	32,850
Shanghai Hongqiao International	Eastern	1	82,000	19,336,517	323,518	177,626
Shenyang Taoxian International	Northeast	1	83,538	5,343,566	85,590	48,931
Shenzhen Baoan International	Central and South	1	111,026	18,356,069	549,879	169,493
Taiyuan Wusu	Northern	1	25,800	2,843,482	26,250	38,356
Tianjin Binhai International	Northern	1	25,500	2,766,504	95,820	54,948
Urumqi Diwopu International	Northwest	1	68,359	5,136,028	72,592	51,602
Xiamen Gaoqi International	Eastern	1	127,000	7,501,004	169,574	77,355
Xian Xianyang International	Northwest	1	25,700	9,368,958	93,349	99,315

Source: Aviation Policy and Research Center (APRC), Chinese University of Hong Kong.

Table 4.2 Summary statistics of inputs and outputs for DEA

Year	Average number of runway	Average terminal size (square meter)	Average passenger volume (person)	Average cargo volume (ton)	Average aircraft movement (flight)
1995	1.04	19,236	2,878,981	41,600	27,734
1996	1.04	20,178	3,113,395	50,940	31,136
1997	1.04	30,997	3,197,762	60,778	33,464
1998	1.04	30,997	3,284,180	69,987	37,822
1999	1.04	41,298	3,497,035	88,635	39,735
2000	1.04	50,847	3,723,238	93,164	41,004
2001	1.04	56,633	4,146,812	97,686	46,187
2002	1.04	57,063	4,687,961	108,805	51,031
2003	1.04	57,063	4,618,482	105,234	51,695
2004	1.05	57,063	6,416,740	121,874	65,398
2005	1.08	66,492	7,654,047	140,545	74,711
2006	1.08	66,492	8,937,524	169,794	84,730

Table 4.3 Listing airports

Airport	Listing year	Stock exchange	State share in 2003 (%)	State share in 2006 (%)
Beijing Capital International	2000	Hong Kong	65	53.8
Guangzhou Baiyun International	2003	Shanghai	60	52.4
Shanghai Hongqiao International	1998	Shanghai	63	57.6
Shenzhen Baoan International	1998	Shenzhen	64	54.7
Xiamen Gaoqi International	1996	Shanghai	75	68.0

Source: Zhang and Yuen (2008)

Table 4.4 Efficiency scores, θ

	1995	1996	1997	1998	1999	2000
Beijing Capital International	0.843	0.816	1.000	1.000	0.634	0.848
Guangzhou Baiyun International	1.000	1.000	0.991	0.936	0.937	1.000
Shanghai Hongqiao International	1.000	1.000	1.000	1.000	1.000	1.000
Shenzhen Baoan International	0.594	0.613	0.745	0.728	0.709	0.559
Kunming Wujiaba International	0.590	0.724	0.902	0.951	1.000	0.535
Chengdu Shuangliu International	0.862	0.827	1.000	1.000	1.000	0.441
Xian Xianyang International	0.631	0.574	0.609	0.629	0.552	0.945
Xiamen Gaoqi International	1.000	1.000	0.332	0.317	0.309	0.335
Chongqing Jiangbei International	0.655	0.589	0.743	0.821	0.775	1.000
Dalian Zhoushuizi International	0.522	0.383	0.361	0.404	0.363	0.529
Qingdao Liuting	0.361	0.429	0.727	0.755	0.749	0.833
Shenyang Taoxian International	0.392	0.380	0.595	0.502	0.436	0.735
Changsha Huanghua	0.588	0.594	1.000	1.000	1.000	0.566
Harbin Taiping International	0.382	0.394	0.145	0.148	0.142	0.160
Urumqi Diwopu International	0.313	0.200	0.312	0.328	0.309	0.412
Jinan Yaoqiang	0.187	0.340	0.529	0.461	0.403	0.618
Sanya Fenghuang	0.130	0.123	0.223	0.406	0.250	0.231
Tianjin Binhai International	0.440	0.797	1.000	0.609	0.370	0.440
Nanning Wuxu	0.319	0.348	0.184	0.206	0.171	0.246
Hefei Luogang	0.462	0.467	0.463	0.341	0.212	0.313
Lanzhou Zhongchuan	0.495	0.223	0.693	0.550	0.556	0.843
Taiyuan Wusu	0.086	0.072	0.091	0.167	0.172	0.260
Hohhot Baita	0.270	0.145	0.098	0.096	0.134	0.143
Kashi	0.051	0.048	0.087	0.092	0.071	0.129
Hailar Dongshan	0.070	0.085	0.174	0.130	0.122	0.177
Average Efficiency	0.490	0.487	0.560	0.543	0.495	0.532
Standard deviation	0.291	0.304	0.341	0.322	0.321	0.296

Table 4.4 Efficiency scores, θ (Cont'd)

	2001	2002	2003	2004	2005	2006
Beijing Capital International	0.874	0.848	0.822	1.000	1.000	1.000
Guangzhou Baiyun International	1.000	1.000	1.000	1.000	0.668	0.617
Shanghai Hongqiao International	0.995	0.897	0.784	1.000	1.000	1.000
Shenzhen Baoan International	0.640	0.722	0.840	1.000	1.000	1.000
Kunming Wujiaba International	0.536	0.582	0.591	0.693	0.698	0.794
Chengdu Shuangliu International	0.484	0.527	0.584	0.770	0.781	0.869
Xian Xianyang International	1.000	1.000	0.846	0.968	1.000	0.947
Xiamen Gaoqi International	0.323	0.358	0.386	0.418	0.404	0.421
Chongqing Jiangbei International	1.000	1.000	1.000	1.000	1.000	1.000
Dalian Zhoushuizi International	0.513	0.488	0.471	0.609	0.581	0.578
Qingdao Liuting	0.887	0.818	0.770	0.865	0.839	0.805
Shenyang Taoxian International	0.207	0.222	0.237	0.287	0.256	0.275
Changsha Huanghua	0.538	0.540	0.579	0.607	0.578	0.615
Harbin Taiping International	0.169	0.175	0.182	0.213	0.210	0.216
Urumqi Diwopu International	0.128	0.162	0.272	0.354	0.324	0.322
Jinan Yaoqiang	0.727	0.647	0.601	0.612	0.597	0.571
Sanya Fenghuang	0.284	0.325	0.337	0.426	0.416	0.441
Tianjin Binhai International	0.393	0.298	0.352	0.559	0.565	0.562
Nanning Wuxu	0.253	0.234	0.236	0.251	0.261	0.259
Hefei Luogang	0.327	0.276	0.215	0.233	0.255	0.268
Lanzhou Zhongchuan	0.186	0.175	0.188	0.219	0.206	0.205
Taiyuan Wusu	0.357	0.265	0.278	0.270	0.347	0.365
Hohhot Baita	0.180	0.149	0.136	0.180	0.197	0.222
Kashi	0.068	0.071	0.085	0.106	0.099	0.093
Hailar Dongshan	0.209	0.172	0.124	0.115	0.154	0.155
Average Efficiency	0.491	0.478	0.477	0.550	0.537	0.544
Standard deviation	0.314	0.308	0.291	0.324	0.309	0.307

Table 4.5 OLS and Tobit regression results on efficiency scores

	OLS Model		Tobit Model	
	Coefficient	t-value	Coefficient	t-value
(Log) distance	-0.162***	-3.77	-0.147***	-2.98
Airport localization dummy	0.104***	2.81	0.106**	2.56
Listing dummy	0.160***	1.80	0.189*	1.94
Airport localization x listing	0.148	1.42	0.200*	1.68
International hub	0.297***	5.39	0.323***	5.03
Regional hub	0.268***	9.26	0.258***	8.16
Congestion dummy	0.117**	2.08	0.234***	3.26
Tourism city dummy	-0.022	-0.49	-0.016	-0.33
Costal city dummy	0.222***	6.19	0.215***	5.52
Runway index	-0.147	-0.59	-1.199	-0.28
Terminal area index	-0.026***	-6.79	-0.030***	-6.86
Passenger handled index	0.011	0.89	0.010	0.71
Cargo handled index	-0.008**	-2.09	-0.008*	-1.89
Air movements index	0.040***	3.09	0.039***	2.78
City population (10000)	0.000***	11.41	0.000***	10.72
Per capita GDP	-0.000***	-5.22	-0.000***	-4.49
Open skies dummy	0.041	0.38	0.569	0.25
Airline merger dummy	-0.020	-0.55	-0.020	-0.49
Guangzhou dummy	-0.286	-1.57	0.128	0.04
Shanghai dummy	0.044	0.44	-0.122	-0.9
Constant	0.873***	3.00	1.884	0.44

Number of observation: 300

Adjusted R-square for OLS model: 0.7127

Log of the likelihood function for Tobit model: 40.8409

* The coefficient is significant at the 90% level.

** The coefficient is significant at the 95% level.

*** The coefficient is significant at the 99% level.

Table 4.6 Descriptive statistics of efficiency growth rates

	Malmquist Index		Technical efficiency (*EFFECH*)		Technological change (*TECH*)	
	Mean	Standard Deviation	Mean	Standard Deviation	Mean	Standard Deviation
1996	1.139	0.354	1.009	0.306	1.128	0.034
1997	0.964	0.351	1.297	0.583	0.765	0.135
1998	1.142	0.286	1.028	0.273	1.115	0.062
1999	1.070	0.221	0.913	0.171	1.173	0.096
2000	0.916	0.251	1.219	0.368	0.764	0.110
2001	1.030	0.321	0.954	0.299	1.083	0.049
2002	1.089	0.126	0.979	0.126	1.115	0.052
2003	1.054	0.164	1.028	0.178	1.029	0.071
2004	1.266	0.152	1.168	0.145	1.087	0.085
2005	1.162	0.135	1.001	0.123	1.161	0.040
2006	1.172	0.067	1.018	0.056	1.152	0.058

Table 4.7 OLS regression results on Malmquist Index and it components

	Malmquist Index		EFFECH		TECH	
	Coefficient	t-value	Coefficient	t-value	Coefficient	t-value
(Log) distance	-0.008	-0.25	0.038	0.64	-0.011	-0.29
Airport localization dummy	0.012	0.42	-0.072	-1.39	0.081**	2.40
Listing dummy	-0.063	-0.93	-0.039	-0.32	-0.021	-0.27
Airport localization x listing	0.121	1.57	0.121	0.88	-0.025	-0.28
International hub	-0.012	-0.29	-0.028	-0.38	-0.025	-0.51
Regional hub	-0.001	-0.07	-0.003	-0.08	-0.011	-0.42
Congestion dummy	0.006	0.16	0.030	0.41	-0.001	-0.01
Tourism city dummy	0.017	0.52	-0.024	-0.41	0.012	0.32
Costal city dummy	0.033	1.19	0.027	0.55	0.002	0.05
Runway growth index	-0.141	-0.90	-0.381	-1.36	0.159	0.87
Terminal area growth index	-0.078***	-13.52	-0.070***	-6.78	-0.026***	-3.91
Passenger growth index	0.108	1.50	-0.143	-1.10	0.119	1.41
Cargo growth index	0.106***	5.87	0.040	1.23	0.044**	2.11
Air movement growth index	0.845***	17.26	0.905***	10.30	-0.031	-0.54
City population (10000)	0.000	0.89	0.000	-0.19	0.000	0.31
Per capita GDP	-0.000	-0.48	0.000	-0.24	0.000	0.18
Open skies dummy	0.034	0.46	-0.088	-0.68	0.129	1.52
Airline merger dummy	0.024	0.88	-0.015	-0.32	0.013	0.41
Guangzhou dummy	-0.247**	-2.18	-0.009	-0.04	-0.121	-0.91
Shanghai dummy	-0.065	-0.89	-0.029	-0.22	-0.041	-0.48
Constant	0.961***	10.36	0.903***	5.43	1.031***	9.51
Adjusted R-square	0.7623		0.4351		0.1874	
Number of Observation	275		275		275	

** The coefficient is significant at the 95% level.

*** The coefficient is significant at the 99% level.

4.8 References

Abbott, M. and Wu, S. (2002) "Total Factor Productivity and Efficiency of Australian Airports", *Australian Economic Review*, 35, 244-260.

Ali, M. and Flinn, J.C. (1989) "Profit Efficiency among Basmati Rice Producers in Pakistan Punjab", *American Journal of Agricultural Economics*, 71(2), 303-310.

Banker, R.D., Charnes, A. and Cooper, W.W. (1984) "Some Models for Estimating Technical and Scale Inefficiencies in Data Envelopment Analysis", *Management Science*, 30, 1078-1092.

Banker, R.D., Charnes, A., Cooper, W.W., Swarts, J. and Thomas, D.A. (1989). "An Introduction to Data Envelopment Analysis with some of its Models and their Uses", in: Chan, J.L. (Ed.), *Research in Government and Nonprofit Accounting*, 5, 125-164, Greenwich, CT: JAI Press.

Charnes, A., Cooper, W.W. and Rhodes, E.L. (1978) "Measuring the Efficiency of Decision-making Units", *European Journal of Operation Research*, 2, 429-444.

Charnes, A., Cooper, W.W. and Rhodes, E.L. (1981) "Evaluating Program and Managerial Efficiency: An Application of Data Envelopment Analysis to Program Follow Through", *Management Science*, 27, 668-697.

Chin, A.T.H. and Siong, L.E. (2001) "Airport Performance: A Comparative Study between Changi Airport and Airports in the New York–New Jersey Metropolitan Area", Paper presented at World Conference on Transport Research, Seoul.

Farrell, M.J. (1957) "The Measurement of Productive Efficiency", *Journal of the Royal Statistics Society, Series A: General*, 120, 253-281.

Fleisher, B.M. and Chen, J. (1997) "The Coast–noncoast Income Gap, Productivity, and Regional Economic Policy in China", *Journal of Comparative Economics*, 25, 220-236.

Fung, M., Wan, K., Hui, Y.V. and Law, J. (2008) "Productivity Changes of Chinese Airports 1995–2004", *Transportation Research Part E*, 44(3), 521-542.

Gillen, D. and Lall, A. (1997) "Developing Measures of Airport Productivity and Performance: An Application of Data Envelopment Analysis", *Transportation Research Part E*, 33(4), 261-273.

Jones, D.C., Li, C. and Owen, A.L. (2003) "Growth and Regional Inequality in China during the Reform Era", *China Economic Review*, 14, 186-200.

Kalirajan, K.P. (1990) "On Measuring Economic Efficiency", *Journal of Applied Econometrics*, 5 (1), 75-85.

Lovell, C.A.K. (1993) "Production Frontiers and Productive Efficiency", in Fried, H.O., Lovell, C.A.K. and Schmidt, S.S. (Eds.), *The Measurement of Productive Efficiency: Techniques and Applications*, 3-67, New York: Oxford University Press.

Martin, J. and Roman, C. (2001) "An Application of DEA to Measure the Efficiency of Spanish Airports prior to Privatization", *Journal of Air Transport Management*, 7, 149-157.

Oum, T.H., Yu, C. and Fu, X. (2003) "A Comparative Analysis of Productivity Performance of the World's Major Airports: Summary Report of the ATRS Global Airport Benchmarking Research Report – 2002", *Journal of Air Transport Management*, 9, 285-297.

Oum, T.H. and Zhang, Y. (1990) "Airport Pricing: Congestion Tolls, Lumpy Investment, and Cost Recovery", *Journal of Public Economics*, 43, 353-374.

Parker, D. (1999) "The Performance of BAA before and after Privatization", *Journal of Transport Economics and Policy*, 33, 133-145.

Pels, E., Nijkamp, P. and Rietveld, P. (2001) "Relative Efficiency of European Airports", *Transport Policy*, 8(3), 183-192.

Sarkis, J. (2000) "An Analysis of the Operational Efficiency of Major Airports in the United States", *Journal of Operations Management*, 18, 335-351.

Seiford, L.M. and Thrall, R.M. (1990) "Recent Developments in DEA: the Mathematical Programming Approach to Frontier Analysis", *Journal of Econometrics*, 46, 7-38.

Sun, Q. and Tong, W.H.S. (2003) "China Shares Issue Privatization: The Extent of its Success", *Journal of Financial Economics*, 70, 183-222.

Tobin, J. (1958) "Estimating the Relationship for Limited Dependent Variable", *Econometrica*, 26, 24-36.

Vasigh, B. and Hamzaee, R.G. (2000) "Airport Efficiency: An Empirical Analysis of the US Commercial Airports", Paper presented at the Fourth Annual Conference of the Air Transport Research Group, Amsterdam.

Wang, C. (2005) "Ownership of Operating Performance of Chinese IPOs", *Journal of Banking and Finance*, 29, 1835-1856.

Zhang, A. (1998) "Industrial Reform and Air Transport Development in China", *Journal of Air Transport Management*, 4, 155-164.

Zhang, A. and Chen, H. (2003) "Evolution of China's Air Transport Development and Policy towards International Liberalization", *Transportation Journal*, 42, 31-49.

Zhang, A., Hui, G., Hui, Y.V., Leung, L. and Cheung, W. (2004), *Air Cargo Logistics Industry in China and Hong Kong*, 2004, London: Ashgate.

Zhang, A. and Yuen, A. (2008), "Airport Policy and Performance in Mainland China and Hong Kong", in: Winston, C. and de Rus, G. (eds.), *Aviation Infrastructure Performance: A Study in Comparative Political Economy*, 159-192, Washington, D.C.: Brookings Institution Press.

Zhang, A. and Zhang, Y. (1997), "Concession Revenue and Optimal Airport Pricing", *Transportation Research Part 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.

Zhang, A., Zhang, Y. and Zhao, R. (2001) "Impact of Ownership and Competition on the Productivity of Chinese Enterprises", *Journal of Comparative Economics*, 29, 327-346.

Zhang, A., Zhang, Y. and Zhao, R. (2002) "Profitability and Productivity of Chinese Industrial Firms: Measurement and Ownership Implications", *China Economic Review*, 13, 65-88.

Chapter 5 Unilateral GHG Control Measure and Aviation Industry: A Theoretical Analysis

5.1 Introduction

Public concerns over climate change have pressured both the aviation industry and regulators to mitigate aviation greenhouse gas (GHG) emissions in hope of avoiding their adverse impacts on global climate. As a result, while a number of airlines set up their own voluntary schemes to reduce their GHG emissions,[79] some governments also consider unilaterally mandating the implementation of some measures to control emissions from the industry. Among these control measures, the one of the largest scale would be the inclusion of the aviation sector in the European Emission Trading System (EU ETS) in 2012.[80] This kind of unilateral action has been strongly opposed by the industry and various countries, based partly on economic concerns, such as competitive issues.[81] For example, Air France argued that the EU ETS would give a competitive edge to those airlines with a hub outside of Europe.[82] It is because connecting flights between

[79] Currently, a number of airlines, including Air Canada, American Airlines, Cathay Pacific, Continental Airlines, Delta Airlines, Northwest Airlines and Qantas, have introduced their carbon offset programs for their customers. More discussion about voluntary emission control measures for aviation can be found in ICAO (2007).

[80] In July 2008, the European Parliament passed the second reading vote of a directive which includes GHG emissions from flights to, from and within the EU in the EU ETS from 2012. Under the new legislation, all airlines will be covered (irrespective of nationality). They will be able to sell surplus allowances if they reduce their emissions and will need to buy additional allowances if their emissions grow.

[81] For example, in the ICAO's meeting in July 2007, the US representative objected to the inclusion of international flights in the EU ETS without their consent, and pointed out that it runs contrary to EU Member State international legal obligations under the Chicago Convention on International Civil Aviation, and under numerous bilateral air services agreements, including those with the U.S.

[82] Reuters, June 9, 2008.

hub and spoke airports in the EU are subject to related GHG charges resulting from the EU GHG emission control measure, while those flights outside the EU are not.

Due to the potentially significant impacts of unilateral GHG control measures on GHG emissions, national welfare and the aviation industry, the economic arguments for and against the unilateral GHG control measures ought to be examined more rigorously. Although there is a growing empirical literature (e.g., Arthur Anderson, 2001; CE Delft, 2005, 2007; Scheelhaase and Grimme, 2007; Scheelhaase, et al., 2008) examining the effects of GHG control mechanisms in the aviation industry, discussion of the issue so far lacks much relevant theoretical work. Thus this Chapter aims to provide an analytical framework for investigating GHG emission control measures in the aviation industry. In particular, we will first examine how the unilateral measures affect the competition of airlines with different networks. For example, to serve the same market, airlines may face different GHG charges if their hub locations differ (as argued by Air France mentioned above). Thus the measures may give some airlines strategic advantages in the competition. The empirical results in Scheelhaase, et al. (2008) showed that network carriers based outside the EU are likely to gain a significant competitive advantage, compared to EU network carriers, if the aviation industry is included in the EU ETS. Our second objective is to look at how the unilateral GHG emission control measures affect world emissions and its distribution among countries. In particular, the unilateral

measures may affect airlines' pricing and network utilization, which, in turn, could affect world emissions.[83]

In the literature, a few theoretical works have been published on pollution control measures on the aviation industry; for example, Brueckner and Girvin (2008) explored the impacts of airport noise regulation on flight frequency and airfares. One of the major differences between this Chapter and theirs is that we consider asymmetric airlines. In our analysis, two airlines operate flights between two countries, and belong to each country, respectively. Under such duopoly competition, the two airlines use different networks to serve the market. As shown below, the introduction of asymmetry in the analysis is necessary for us to examine the competitive issues and the change in world emissions with the unilateral action to control GHG emissions – the main objectives of this Chapter. However, due to the complexity after considering the asymmetric case, we cannot take into account flight frequency in our model as the studies in the literature. Yet an important extension to this Chapter could examine the impacts of GHG control measures on air service quality, such as flight frequency: more frequent flights will reduce scheduled flight delays (Douglas and Miller, 1974). To derive the airlines' demand, we consider an infinite linear city model, where a consumer's utility depends on their brand preference. A number of studies in the literature (e.g., Brueckner and Zhang, 2001; Brueckner, 2004; Brueckner and Flores-Fillol, 2007) consider similar spatial

[83] From a different perspective, Hoel (1991) showed that the unilateral reductions of GHG emissions may affect the outcome of international negotiations about reduced emissions, which may imply higher total emissions than if no action is taken.

models to examine airlines' network choice between fully-connected (FC) and hub-and-spoke (HS) networks.

In the analysis, we compare the case in which one of the countries unilaterally imposes GHG charges on both airlines with the no-action base case in which both countries have no action regarding GHG emissions. We find that if a country price discriminates between domestic and international flights, the GHG charges may be positive or negative, while they will be positive when its home airline's profit is not taken into account in the country's airport pricings. On the other hand, under uniform pricing, when a country moves toward GHG pricing, the airport charge will increase. Note that the adverse impacts of the positive GHG charges on the home airline may be larger relative to their foreign counterpart, as the home airline will operate connecting flights between the hub and spoke airports within its home country, where the GHG pricing is implemented. Thus, such an initiative may raise competitive concerns from those airlines. We also show that the positive GHG charges will reduce the total output in the airline market, while domestic flights connecting hub and spoke airports may be shifted from the home country to the foreign counterpart. Nevertheless, the consumer surplus of the air travel market will decrease. Finally, we find that world emissions may increase when a country takes the GHG control actions unilaterally. This is due to the fact that the domestic connecting flights may be shifted to a country with a less efficient network. This is a very interesting result, with potentially important implications for governments deciding to implement unilateral GHG control measures.

The Chapter is organized as follows: Section 5.2 discusses the relationship between global climate change and the aviation industry, including GHG control measures implemented in the industry in practice. Section 5.3 sets up the model, and Section 5.4 examines airline behavior. Section 5.5 derives the unilateral GHG pricing, and Section 5.6 investigates its effects on airline competition, market output, consumer welfare and GHG emissions. Section 5.7 concludes.

5.2 Climate Change and Aviation

An increasing amount of scientific evidence indicates that the global climate change is predominantly a result of increases in GHGs caused by human activities (see Stern, 2006, for a comprehensive review). Transportation, in particular, is one of the most important sources of GHG emissions. As with other modes of transportation, air transport produces GHGs during operation. The aviation GHG of most concern is carbon dioxide (CO_2), which has long residence time in the atmosphere. Other aircraft emissions affecting climate include water vapor (H_2O), nitrous oxide (NO_X), sulphate (SO_4) and soot particles. Currently, the aviation industry contributes 3 percent of the total man-made contribution to climate change (IPCC, 2007). Given the fact that the direct contribution of the industry to the global gross domestic product (GDP) is only about 1 percent, the industry share of the man-made contribution to climate change is thus significant. Furthermore, air travel grew significantly in the past decade and will continue growing at a rapid pace in the coming years. This rapid growth in air traffic may lead to a significant increase in GHG emissions from the sector, even with improvements in

aircraft fuel efficiency.[84] IPCC (2007) found that CO_2 emissions from global aviation increased by a factor of about 1.5 from 1990 to 2000, and predicted that the emissions will continue to grow by around 3-4 percent each year, unless additional measures are taken; Olsthoorn (2001) estimated that between 1995 and 2050, aviation emissions of CO_2 may increase by a factor of 3-6.

Given its global character, it is sensible for countries to act collectively to control GHG emissions. As of May 2008, more than 180 countries have signed and ratified the Kyoto Protocol,[85] an international agreement that sets binding targets for countries and regions to reduce GHG emissions. Yet, while the Protocol accounts for domestic aviation emissions, emissions from international aviation are not included.[86] It may be because allocating international aviation emissions to specific countries may be difficult. Instead, the Protocol calls for countries to pursue limitation or reduction of GHG emissions from international flights through the International Civil Aviation Organization (ICAO).[87] Currently, the ICAO focuses mainly on technological and operational measures in GHG emission control, whereas the progress of market-based measures through the organization is still limited.[88] Instead, a number of airlines set up their own voluntary schemes, in which passengers can pay extra to offset their emissions. The revenues from the passengers are then spent on various offset programs. Hodgkinson, et al. (2007)

[84] IPCC (2007) forecasted that technology developments might offer only a 40-50 percent improvement in fuel efficiency over 1997 level by 2050, while air traffic growth is estimated to be around 5 percent each year.

[85] The participating countries now contribute to about 64 percent of global GHG emissions in total.

[86] Annex I Parties of the Protocol.

[87] Article 2.2 of the Protocol.

[88] In particular, the ICAO has decided not to work towards a new global legal market-based instrument under the organization in 2004.

argued that the schemes do serve a useful purpose in enabling passengers who are concerned about their emissions to make an essentially carbon neutral flight. On the other hand, British Airways voluntarily participated in the UK ETS to reduce their emissions of CO_2 equivalent to below set targets. In return, it received an incentive payment from the UK Government. However, the effectiveness of those voluntary measures is still questionable.

As a result, some governments are also considering unilaterally mandating some measures to control emissions from the aviation industry. A growing number of empirical studies have been conducted to evaluate the proposed measures. Scheelhaase and Grimme (2007) provided an empirical estimation of the impacts of the inclusion of the aviation sector in the EU ETS on operating costs and transport demand for full service and low cost airlines, and found that the financial impact on low cost carriers and regional airlines is likely to be significantly greater than for network carriers. Furthermore, Scheelhaase, et al. (2008) found that EU network carriers are likely to encounter competitive disadvantages compared to airlines from non-EU countries on long-haul flights. Benito (2008) estimated the aviation CO_2 emissions reduction with the proposed EU ETS by taking into account the substitution between air travel and high-speed railways in the EU. In addition to the EU ETS mentioned above, the UK government levies an "Air Passenger Duty," which is designed to reduce CO_2, on all airline tickets sold in the country. However, IATA (2006) opposed the levy because the climate benefits due to the reduction in emissions is estimated to be £53 million only, while the GDP losses would be £400 million plus losses to air travelers. Hofer, et al.

(2008) investigated the effects of an air travel carbon emissions tax on total travel-related carbon emissions in the U.S. They highlighted the fact that emissions taxes on air travel may cause potentially significant air-to-automobile diversion effects, which may substantially reduce the environmental benefits of air travel carbon emissions taxes. Forsyth (2008), and Forsyth and Ho (2008) discussed the Australian ETS by taking direct emissions from aviation and indirect emissions from other industries, such as tourism, into account.

5.3 The Model

This analysis considers a model with an airline in country H (i.e., airline H) and an airline in country F (i.e., airline F) serving a same origin-destination (OD) market. As depicted in Figure 5.1, a simple air transport network is considered, as this is likely the simplest structure allowing us to address the main questions concerning us. To simplify the analysis, we focus on the OD market between airports H and F (i.e., the spoke airports in countries H and F, respectively). To serve the market, each airline will bring their passengers from the origin to its hub, which is located at its respective country (i.e., airports A and B for airlines H and F, respectively), and then from the hub to the destination. In practice, it is very common for airlines operating HS networks in both their domestic and international markets after market deregulation in the aviation industry (see Zhang, et al., 2008, for a review). One example relevant for our analysis is the air travel market between Vancouver and Munich. In this market, Air Canada will take up passengers from Vancouver and fly to Toronto, which is Air Canada's hub for its

transatlantic market, then from Toronto to Munich. On the other hand, Lufthansa serves the market by connecting at Frankfurt. The assumption that airlines choose (or are only allowed to use) airports at their home country as their hubs can be justified for at least two reasons. First, although the international aviation market has been moving toward greater liberalization, cabotage is usually prohibited, and thus airlines are restricted to operate flights within the domestic borders of another country.[89] Second, most major airports around the world are suffering from congestion. Home carriers usually dominate those airports, implying that foreign carriers with limited access to the airport facilities may have difficulties operating their own flights at those airports.[90]

[Figure 5.1 here]

In the model, aircrafts emit GHGs during the flight, where the emission intensity (i.e., emissions per flight between two airports) depends on the route they fly. On the other hand, for simplicity, we will abstract from other possible OD markets in the network (e.g., H to B, A to H) in our analysis.[91] The stylized problem that we want to analyze is the following: when one of the countries adopts (unilateral) GHG pricing, what are its effects on airlines' operating decisions, competition, market output, consumer benefits and world emissions? We model the problem as a two-stage game. In the first

[89] A few exceptions include the creation of single aviation markets in the EU, and between Australia and New Zealand, in which cabotage rights are granted to airlines in the markets.

[90] Ciliberto and Williams (2007) measured the importance of operating barriers to entry, including limited access to airport facilities, as determinants of the hub premium; they found that exclusive access to and dominance of gates at the market endpoint airports are key determinants of the hub premium.

[91] Under this simplification, our model will not take the advantages of HS network into account, including those due to demand and cost complementarities, which have been widely discussed in the literature (e.g., Oum, et al., 1995; Hendricks, et al., 1995, 1999; Brueckner and Zhang, 2001; Brueckner, 2004; Brueckner and Flores-Fillol, 2007).

stage, countries determine the charges, with part of them owing to GHG pricing, at their respective hub and spoke airports. In the second stage, the two airlines compete in Cournot fashion.[92]

To derive the demand faced by the two airlines, we specify and solve the consumers' problem as follows. We consider an infinite linear city, where potential consumers are distributed uniformly with a density of one consumer per unit of length. Basso and Zhang (2007a) use a similar spatial model to investigate the rivalry between congestible facilities. Airlines F and H are located at 0 and 1, respectively, and the locations of the airlines are exogenous. If a consumer located at z chooses to travel by F, then she derives a net benefit (utility):

$$U_F = V - p_F - 4\tau|z|, \tag{5.1}$$

where V is the gross travel benefit; p_F is the (equilibrium) airline F's airfare; 4τ is a parameter capturing consumers' transportation cost (i.e., the intensity of consumers' brand preference to airlines: consumers located closer to an airline's location will enjoy a higher benefit than others if they fly by that airline, given the same airfare), and τ is assumed to be positive.[93] Similarly, if the consumer chooses to travel by H, then she derives a net benefit (utility):

[92] Cournot competition between air carriers has been a common assumption in the theoretical airport pricing literature, which in fact has some empirical backing (e.g., Brander and Zhang, 1990; Oum, et al., 1993). See Basso and Zhang (2007b) for a review of the airport pricing literature.

[93] This transportation cost is multiplied by 4 in order to simplify most of the equations in the Chapter (see, e.g., equations (5.6)).

$$U_H = V - p_H - 4\tau|1-z|, \tag{5.2}$$

where p_H is the (equilibrium) airline H's airfare.

Assuming that everyone with $z \in [0,1]$ travels and both airlines capture some of the travelers,[94] then the consumer with $\tilde{z}$ will be indifferent between the two airlines when (also see Figure 5.2):

$$V - p_F - 4\tau\tilde{z} = V - p_H - 4\tau(1-\tilde{z}),$$

or when

$$\tilde{z} = \frac{1}{2} + \frac{p_H - p_F}{8\tau}. \tag{5.3}$$

[Figure 5.2 here]

Condition (5.3) suggests that, when the airline H's (airline F's) fare decreases, more (less) passengers will choose airline H rather than airline F. Furthermore, airlines F and H serve consumers with $z < 0$ and $z > 1$, respectively. The cutoff points z^l and z^r for the

[94] To have everyone with $z \in [0,1]$ travel, we assume $2V \geq p_H + p_F + 4\tau$, and to have both airlines capture some of the travelers, we assume $|p_H - p_F| < 4\tau$. Both of which are our maintained assumptions in the following analysis.

customers between travel with airlines F and H, respectively, and not travel by air (which utility is normalized to zero) can be derived as follows:[95]

$$z^l = -\frac{V - p_F}{4\tau}, \qquad z^r = 1 + \frac{V - p_H}{4\tau}. \tag{5.4}$$

By conditions (5.3) and (5.4), we can derive the demand system as follows:

$$q_F = \frac{V + 2\tau}{4\tau} + \frac{p_H - 3p_F}{8\tau},$$

$$q_H = \frac{V + 2\tau}{4\tau} + \frac{p_F - 3p_H}{8\tau}. \tag{5.5}$$

For simplicity, q_i is measured by the number of flights. This measurement is equivalent to the number of passengers if each flight has an equal number of passengers, which holds when all flights use identical aircraft and have the same load factor. Note from (5.5) that an airline loses traffic when its fare rises, while it gains traffic when another airline's fare rises. In other words, the service provided by an airline is a substitute to another. From the demand system (5.5), we can obtain the following inverse demand functions faced by the two airlines:

[95] The "no travel by air" option may also include traveling by other modes of transportation, such as automobile and high-speed railways. In such cases, we may need to further consider the consumer benefits and emissions from other modes of transportation in the following analysis. This kind of diversion effects are empirically examined in Benito (2008) and Hofer, et al. (2008).

$$p_F(q_H, q_F) = 2\tau + V - 3\tau q_F - \tau q_H,$$

$$p_H(q_F, q_H) = 2\tau + V - 3\tau q_H - \tau q_F. \tag{5.6}$$

The inverse demand system (5.6) will be used to solve the Cournot-Nash equilibrium in Section 5.4.

5.4 Airline Competition

We shall examine the subgame perfect equilibrium of the two-stage game, starting with the analysis of the airlines' subgame in the second stage. Having characterized the demands in Section 5.3, we now specify airlines' cost. The carriers are symmetric in the sense that they have the same cost function. The airline i's total cost function is given as:

$$C_i(q_i) = (c + t_i^S + t_i^U + t_{-i}^S)q_i, \tag{5.7}$$

where c is the (constant) marginal cost per flight.[96] Recall that an airline will fly from a spoke airport to another spoke airport through the hub located at its home country. As a result, airline i will incur per-flight charges at the home country's spoke and hub airports, t_i^S and t_i^U, respectively, and a per-flight charge at the foreign country's spoke airport, t_{-i}^S. The airport charges will be determined in Section 5.5. In our analysis, we assume that the

[96] Here, we assume that the airlines are operating under constant returns to scale. By using different cost functions, Kumbhakar (1990, 1992) and Hansen, et al. (2001) empirically found that airlines operate under increasing returns to scale, while Caves, et al. (1984) and Chua, et al. (2005) found constant returns to scale.

per-flight aircraft emissions in each route are exogenous.[97] Furthermore, the GHG taxes considered in Section 5.5 depend only on airlines' output. Thus, there is no incentive for airlines to improve their efficiency in GHG emissions. As a result, the GHG abatement cost is not included in (5.7). Yet an interesting extension to the present analysis would be considering incentive mechanisms for airlines to improve their emission intensity.

In the second stage, airline i chooses the number of flights to maximize its own profit. Thus its optimization problem is:

$$\underset{q_i}{Max}\ \pi_i(q_i; q_{-i}) = p_i q_i - C_i(q_i) \qquad (for\ i = \text{F, H}) \quad (5.8)$$

$$= (2\tau + V - 3\tau q_i - \tau q_{-i})q_i - (c + t_i^S + t_i^U + t_{-i}^S)q_i .$$

The first-order condition for the maximization problem is:

$$\frac{\partial \pi_i}{\partial q_i} = (2\tau + V - 6\tau q_i - \tau q_{-i}) - (c + t_i^S + t_i^U + t_{-i}^S) = 0, \qquad (for\ i = \text{F, H}) \quad (5.9)$$

where $\partial^2 \pi_i / \partial q_i^2 = -6\tau < 0$. Thus, the second-order condition holds.

[97] In practice, some airlines have made some progress in GHG emission control. For example, Air Canada reduced the emission intensity of its flight operations by 28 percent since 1990 by investing in energy efficient new aircrafts, including Embraer regional jets and Boeing 777s.

Solving equations (5.9), the airline i's equilibrium output is:

$$q_i = \frac{2\tau + V - c - t_i^S - t_{-i}^S}{7} - \frac{6t_i^U - t_{-i}^U}{35}. \qquad (for\ i = \text{F, H}) \quad (5.10)$$

Conditions (5.10) suggest that airline i's output decreases in the charges at the home country's airports (both the hub and spoke airports) and the foreign country's spoke airport, while it increases in the foreign country's hub airport charge. Furthermore, the total market output, $Q = q_i + q_{-i} = [2(2\tau + V - c - t_i^S - t_{-i}^S) - (t_i^U + t_{-i}^U)]/7$, decreases in the hub and spoke airport charges at both home and foreign countries.

Substituting (5.10) into (5.6), the equilibrium prices are:

$$p_H(q_F, q_H) = (2\tau + V)(1 - \frac{4}{7}\tau) + \frac{\tau}{35}[20(c + t_H^S + t_F^S) + 17t_H^U + 3t_F^U], \qquad (5.11)$$

$$p_F(q_H, q_F) = (2\tau + V)(1 - \frac{4}{7}\tau) + \frac{\tau}{35}[20(c + t_H^S + t_F^S) + 17t_F^U + 3t_S^U]. \qquad (5.12)$$

5.5 Country Stage

To examine the effects of the unilateral GHG emission control measure, we first derive the optimal airport charges, with part of them owing to GHG pricing. In this Chapter, we focus on the unilateral action from a country to control GHG emissions. Thus, in the country stage, we consider country H chooses charges t_H^U and t_H^S at its hub

and spoke airports, respectively, taking the subsequent carriers behavior into account. The airport charges at country F are assumed to be exogenous.[98] Country H will maximize the following welfare function:

$$\underset{t_H^S, t_H^U}{Max}\, SW_H = \alpha CS + \beta \pi_H + TR_H - \gamma E_H \,, \tag{5.13}$$

where CS is total consumer surplus of the air travel market. α and β, which are positive, capture the importance (relative to tax revenue) of consumer surplus and home airline's profit, respectively, in country H's policy-making decision.[99] TR_H is country H's tax revenue, which is the sum of tax revenues from its spoke and hub airports (i.e., $t_H^S(q_H + q_F)$ and $t_H^U q_H$, respectively). Recall airline H will use both the hub and spoke airports at country H, while airline F will only use the spoke airport at the country (see Figure 5.1). E_H is the GHG emissions related to country H, the positive γ is the country H's GHG (per-unit) abatement cost or the cost to meet a GHG control target, for example, those set under the Kyoto Protocol.[100] Thus, the last term in (5.13) is the total cost on GHG emission reduction in order to meet a particular emission target. Note that the last

[98] If both countries' airport pricings are treated as endogenous, we need to consider the strategic behaviors between the two countries (see Chapter 3 of this thesis, Small and Verhoef, 2007, and Ubbels and Verhoef, 2008, for recent reviews of the studies on the strategic behaviors of transport facilities).

[99] The magnitudes of α and β depend on a number of factors, including the percentage of foreign ownership stake in the home airline, the proportion of consumers from country H relative to the whole air travel market, costs of taxation from other sectors, etc. For example, in case of France, which supports national "champion" such as Air France, β tends to be large.

[100] For example, a country may offset its GHG emissions through some project-based transactions, in which the buyer purchases emission credits from a project that can verifiably demonstrate GHG emission reductions compared with what would have happened otherwise. The most notable examples of such activities are under the Clean Development Mechanism (CDM) and the Joint Implementation (JI) mechanisms of the Kyoto Protocol. In 2006, the CDM and JI grew sharply to a value of about US$5 billion (World Bank, 2007).

term in (5.13) may also be interpreted as the benefits of emission reduction to the country and its consumers. In particular, for $\gamma = 0$, country H does not account for any GHG emission costs to the country. Thus, the pricing derived from the maximization problem in (5.13) will not include any GHG component – the base case considered in this Chapter. In the following, we will compare the equilibrium in the base case with that in the case of $\gamma > 0$.

We further specify the GHG emissions related to country H as follows:

$$E_H = e_H q_H + \delta e_K (q_H + q_F), \tag{5.14}$$

where e_H and e_K are the (per-flight) emissions of domestic travel in country H (i.e., between A and H as shown in Figure 5.1) and international travel (i.e., between A and F or between B and H), respectively. In practice, the emissions depend on a number of factors, including aircraft type, load factor and flight distance. In our present analysis, those factors are treated as given, thus the per-flight emissions are exogenous. Yet it is an important extension for this Chapter to consider the factors endogenously in the analysis. More details about this extension will be discussed in the concluding remarks of this Chapter. On the other hand, the international emissions are discounted by a factor $\delta \in [0,1]$ in counting the emissions related to country H. The different weight of domestic and international emissions in counting the emissions related to country H may be due to different treatments between the two kinds of flights in the existing GHG control

mechanisms. For example, as mentioned, emissions from domestic flights are included in the Kyoto Protocol, while those from international flights are not.

An important assumption here is that country H will not take the emissions of domestic flights of another country into account. Note that although GHG emissions may have global effects, GHG emission control measures considered by most countries in practice aim to limit the emissions within its own territory, rather than global emissions. For example, under the EU ETS, the EU Member States are required to limit their emissions below *national* emission caps. On the other hand, in evaluating the effectiveness of the inclusion of the aviation sector in the EU ETS, the Full Assessment Report (EC, 2006) submitted by the European Commission only took into account the emissions from flights to, from and within the EU, but not from the flights outside the EU. Similarly, CE Delft (2005), a report commissioned by the European Commission, also evaluated different GHG control options by comparing the emissions from the flights within the European airspace.

In the following, we will explore both the price discrimination case and the "uniform pricing" case, in which country H charges a uniform (per-unit) charge to flights at its hub and spoke airports. Note that the uniform pricing case does not necessarily mean flights at the hub and spoke airports are charged exactly the same taxes. Instead, it should be understood as the existence of some a *priori* rule which says that flights at the hub and spoke airports are charged according to some fixed ratio (for example, the hub airport charges three times what the spoke airport charges). To save on notation – but

without losing insight – we set the ratio to one in the uniform pricing case, that is, $t_H^S = t_H^U = \bar{t}_H$. In practice, the uniform pricing case in our analysis may be more relevant. For example, the proposal for including the aviation sector in the EU ETS requires domestic and international flights to participate in the same ETS. As a result, the flights are required to buy emission permits from the ETS at a same price. Nevertheless, both the price discrimination case and the uniform pricing case will be examined in the following.

In the price discrimination case, country H chooses airport taxes at its hub and spoke airports to maximize its social welfare given in (5.13). The first-order conditions with respect to the charges give rise to the welfare-maximizing pricing rules. We obtain (the derivation is straightforward but long, and is hence given in Appendix D):

$$t_H^U = \{\frac{5[14-\beta+\tau(\beta-4\alpha)]}{\Omega_1} e_H + \frac{30(2-\tau)\beta\delta}{\Omega_1} e_K\}\gamma + K_1, \tag{5.15}$$

$$t_H^S = \{\frac{6\beta+17\alpha\tau-6\beta\tau}{\Omega_1} e_H + \frac{[70-65\beta+(29\beta+14\alpha)\tau]\delta}{\Omega_1} e_K\}\gamma + K_2, \tag{5.16}$$

where $\Omega_1 = -(\tau-1)^2\beta^2 + 2(34\tau+3\alpha\tau-70)\beta - 4(2\alpha\tau+3)\alpha\tau + 140 > 0$ by the second-order conditions. K_1 and K_2 are constants.

In the base case, country H does not take the GHG emission cost into account, i.e., $\gamma = 0$. Thus, equations (5.15) and (5.16) suggest that in the base case, the equilibrium charges at the hub and spoke airports at country H are equal to K_1 and K_2, respectively.

In particular, K_1 and K_2 may be considered as trade taxes (or subsidies) under imperfect international competition, which have been widely discussed in the strategic trade policy literature (e.g., see Brander, 1995, for a comprehensive review). When country H moves toward (unilateral) GHG pricing (i.e., $\gamma > 0$), the bracketed terms in (5.15) and (5.16) (multiplied by a positive γ) are those additional components to the base case. An interesting catch here is that, once GHG emissions are taken into account, the airport charges may increase or decrease. This result seems contradictory to the traditional wisdom that a positive charge should be imposed to internalize the negative externality (e.g., pollution or GHG emissions in our case). To understand why the signs of the additional terms are undetermined, we make note of three points. First, an increase in airport charges may decrease the airlines' output, thereby reducing GHG emissions. Second, an increase in t_H^U, which is only imposed on airline H, may shift customers to airline F. Thus domestic connecting flights in country H decrease, while those in country F increase. Recall that only the emissions from the former are counted in country H's emissions, but not that from the latter. As a result, the emissions related to country H may decrease. Third, a decrease in airport charges may confer a strategic advantage to its home airline. Note that the home airline will use both the hub and spoke airports, while its foreign counterpart will only use the spoke airport. Thus, the decrease in the airport charges will benefit the home airline more than its foreign counterpart. While the first two have positive effects on the airport charges after GHG emissions are taken into account, the last effect implies that negative GHG charges are possible.

To see how the airline H's profit in (5.13) affects the airport charges after the unilateral GHG taxes are imposed, we consider a particular case that country H does not take airline H's profit into account (i.e., $\beta = 0$). As a result, the last effect discussed above vanishes. In this case, equations (5.15) and (5.16) reduce to:

$$t_H^U = \frac{5e_H}{2(\alpha\tau + 5)}\gamma + K_4 > 0, \tag{5.17}$$

$$t_H^S = \frac{1}{(7 - 2\alpha\tau)}[\frac{17\alpha\tau}{(\alpha\tau + 5)}e_H + 14\delta e_K]\gamma + K_3 > 0, \tag{5.18}$$

where the denominators are positive by the second-order conditions. K_3 and K_4 are constants. Equations (5.17) and (5.18) suggest that, when country H does not take airline H's profit into account, both the airport charges at the hub and spoke airports will increase, after country H implements GHG pricing. Note that the GHG taxes are increasing in γ. In other words, as the GHG abatement cost increases, country H will charge higher taxes on flights. Equations (5.17) and (5.18) show that the charge at the spoke airport increases in both e_H and e_K, while the charge at the hub airport increases only in e_H. This observation implies that the two charges serve different roles in the GHG control. Recall that only airline H will be subject to t_H^U. Thus, an increase in t_H^U will decrease the output by airline H, thereby reducing the emissions from the domestic flights in country H. As a result, if the domestic flights in country H are less efficient in GHG emission (a large e_H), a higher t_H^U is necessary. Note that an increase in t_H^U will also increase the output by airline F, thereby increasing the emissions from the domestic

flights in country F, though country H does not take this into account. On the other hand, an increase in t_H^S will reduce the output of both airlines, thereby reducing both the emissions from the domestic flights in country H and international flights operated by the two airlines. Thus, t_H^S increases in both e_H and e_K. The above discussion leads to:

Proposition 5.1 In the price discrimination case, when a country moves toward GHG pricing, its airport charges may increase or decrease, while the charges will increase if the country does not take its airline's profit into account.

We next consider the uniform pricing case, in which country H chooses a single airport tax at the hub and spoke airports (i.e., $t_H^S = t_H^U = \bar{t}_H$) to maximize its social welfare given in (5.13). The first-order condition with respect to the charge gives rise to the welfare-maximizing pricing rules. We obtain (see the derivation in Appendix D):

$$\bar{t}_H = \frac{35}{\Omega_2}(11e_H + 15\delta e_K)\gamma + K_5, \tag{5.19}$$

where $\Omega_2 = -401\alpha\tau + 814\beta\tau - 1540\beta + 1820 > 0$ by the second-order condition. K_5 is a constant.

Equation (5.19) suggests that when country H moves toward GHG pricing (i.e., from $\gamma = 0$ to $\gamma > 0$), the (single) airport charge will increase. The airport charge is a weighted average of e_H and e_K. The weight of e_K depends on δ, the discount factor of

the international emissions in counting country H's GHG emissions. Note that for $\delta = 1$, the weight of e_H is less than that of e_K in the airport charge given in (5.19). This result may be due to the existence of a trade-off in country H's airport pricing between domestic emission reduction in country H and airline H's profit. In particular, a measure to reduce the domestic emissions, which are only from the flights operated by airline H, will increase airline H's operating cost, but have no effect on that for airline F. As a result, this will confer a strategic advantage to airline F. Yet this tradeoff does not exist in a measure to reduce the international emissions, as it affects both airlines' operating cost by the same amount. As a result, for $\delta = 1$, country H may have more incentives to reduce the international emissions than domestic ones. Finally, it is important to note that, with uniform pricing, the effects of the GHG measure on airline H will be larger than that on airline F. This is due to the fact that airline H will be subject to the airport charge at both the hub and spoke airports in country H, while airline F will only incur such charge at the spoke airport but not at the hub airport in country H. As equation (5.19) suggests that the airport charge will increase after the imposition of GHG pricing, the pricing may confer a strategic advantage to airline F in the competition. As a result, the GHG pricing may raise competitive concerns. This leads to the following proposition:

Proposition 5.2 In the uniform pricing case, when a country moves toward GHG pricing, its airport charge will increase, which will place its airline at a strategic disadvantage.

Before analyzing how the unilateral GHG pricing affects the airline competition and total emissions, it is important to note that the results in this section may have

significant implications for empirical studies on the issue. We show that the implementation of the GHG pricing may have different impacts on home and foreign airlines. As a result, the pricing may alter the strategic relationship between the home and foreign airlines. If this effect is not taken into account in the empirical studies on the issue, the estimation in those studies may underestimate the effects of the GHG pricing on the welfare of the country implementing the measures and its airlines' profits. On the other hand, the empirical studies may overestimate the effects on consumers, as it does not account for their option to choose an airline less affected by the GHG control measures.

5.6 Effects of Unilateral GHG Pricing

In this section, we further investigate the effects of unilateral GHG pricing examined in the previous section on market outputs, consumer surplus and GHG emissions. We focus on two particular cases: the price discrimination case in which airline H's profit is not taken into account in country H's airport pricings (i.e., $\beta = 0$) and the uniform pricing case. Note that as the change of airport charges in the general price discrimination case is undetermined as shown in the previous section, its effects on market outputs and emissions are also generally undetermined.

5.6.1 Market outputs and consumer surplus

For the price discrimination case with $\beta = 0$, substituting equations (5.17) and (5.18) into (5.10) gives the following expressions for the changes in market outputs under unilateral GHG pricing (superscripts A and B for variables after and before the imposition of the unilateral GHG pricing, respectively):

$$q_H^A - q_H^B = -\frac{\gamma[(12-\alpha\tau)e_H + 2(5+\alpha\tau)\delta e_K]}{4(\alpha\tau+5)(7-2\alpha\tau)} < 0, \tag{5.20}$$

$$q_F^A - q_F^B = -\frac{\gamma[(2-3\alpha\tau)e_H + 2(5+\alpha\tau)\delta e_K]}{4(\alpha\tau+5)(7-2\alpha\tau)} < or > 0, \tag{5.21}$$

$$Q^A - Q^B = -\frac{\gamma(e_H + 2\delta e_K)}{2(7-2\alpha\tau)} < 0, \tag{5.22}$$

where the denominators are positive by the second-order conditions for country H's welfare maximization problem.

Equation (5.20) suggests that airline H's output will decrease after the imposition of unilateral GHG pricing. After the policy change, both airline H and F's (equilibrium) fares will increase, while the increase for the former will be more than that for the latter (see equations (5.11) and (5.12)). Given this, we can break down the decrease in airline H's output into two parts (also see Figure 5.3): (i) an increase in airport charges will increase the fare charged by airline H, and lead to a contraction of airline H's market at the region where consumers located at $z > 1$ (i.e., z^r decreases); (ii) the airline H's

market share in the market for consumers with z between 0 and 1 will decrease (i.e., $\tilde{z}$ increases). As mentioned, an increase in airport charges will have a stronger (positive) effect on airline H's fare than that on airline F's fare. Thus, more consumers in the region of $0 \leq z \leq 1$ will choose airline F over airline H. Equation (5.21) suggests that the airline F's output may increase or decrease. The result is undetermined because airline F's market at the region where consumers located at $z < 0$ will contract (i.e., $|z^l|$ decreases), while its market share in the market for $0 \leq z \leq 1$ will increase.

[Figure 5.3 here]

Nevertheless, equation (5.22) suggests that the total market output will decrease after the imposition of the unilateral GHG pricing. Note that as the market size in the region of $0 \leq z \leq 1$ does not change, the decrease in the total market output is due to the decrease in the market at the two extremes (i.e., $z < 0$ and $z > 1$). On the other hand, we can conclude that consumer surplus of the air travel market will decrease after the imposition of the unilateral GHG pricing. It is because, on one hand, fewer consumers will choose to travel by air (i.e., Q decreases). On the other hand, passengers traveling on both airlines H and F will face higher fares. This result suggests that it is important to consider the adverse impact on consumer benefits, when a country plans to implement GHG control measures.

Proposition 5.3 In the price discrimination case, when a country moves toward GHG pricing without taking its airline's profit into account, its airlines' and total market outputs will decrease, while that of the foreign airline may increase or decrease. Consumer surplus will decrease.

For the uniform pricing case, by substituting equation (5.19) into (5.10), the changes in market outputs under the unilateral GHG pricing are given below:

$$q_H^A - q_H^B = -\frac{11(11e_H + 15\delta e_K)\gamma}{\Omega_2} < 0, \tag{5.23}$$

$$q_F^A - q_F^B = -\frac{4(11e_H + 15\delta e_K)\gamma}{\Omega_2} < 0, \tag{5.24}$$

$$Q^A - Q^B = -\frac{15(11e_H + 15\delta e_K)\gamma}{\Omega_2} < 0. \tag{5.25}$$

Equations (5.23) to (5.25) suggest that the total market output, and airline H and F's outputs will decrease after country H implements GHG pricing. Comparing (5.23) and (5.24), we can see that the decrease in airline H's output will be greater than that for airline F's output. This is due to fact that airline H will be charged twice (at the hub and spoke airports in country H) by the expression given in (5.18), while airline F will be charged once (at the spoke airport in country H). Thus, the airline H's fare will increase more than that for airline F's fare. As a result, airline H's market share in the market for $0 \le z \le 1$ will decrease (i.e., $\tilde{z}$ increases). On the other hand, as in the price discrimination case, the decrease in total market output in (5.25) is due to the decrease in the market at the two extremes (i.e., $z < 0$ and $z > 1$). For the consumer surplus, since

airfares increase and output decreases, consumers will be worse off after the imposition of unilateral GHG pricing.

Proposition 5.4 In the uniform pricing case, when a country moves toward GHG pricing, both its home and foreign airlines' outputs will decrease, and consumer surplus will decrease.

5.6.2 GHG emission

In this section, we examine how the unilateral GHG control measure will affect world emissions. This section may be related to the pollution haven hypothesis discussed in the environmental economics literature (see, e.g., Copeland and Taylor, 1994), which suggests that industries that are highly (local) pollution-intensive may migrate from high-income countries to its low-income counterparts, causing world pollution to increase.[101] Yet note that the effect of GHG emissions is global: its impact is independent of where in the world the emissions occur. In the following, we will explore a possibility that the aviation GHG emissions may be shifted from a country to another, which may lead to an increase in world emissions. As in the previous section, we focus particularly on the price discrimination case, in which airline H's profit is not taken into account in country H's airport pricings (i.e., $\beta = 0$) and on the uniform pricing case. The world emissions are the sum of the emissions from airlines H and F, thus:

[101] Empirical studies of the pollution haven hypothesis can be found in, for example, Grossman and Krueger (1991), Xu (2000), Eskeland and Harrison (2003), Kahn (2003), Levinson and Taylor (2004), and Ederington, et al. (2005).

$$E = q_H(e_H + e_K) + q_F(e_F + e_K), \tag{5.26}$$

where e_F is the (per-unit) emissions due to the connecting flights operated by airline F between the hub and spoke airports in country F (i.e., the route between B and F as shown in Figure 5.1). Thus, the first term on the RHS of (5.26) is the total emissions from airline H's flights, and the second term is those from airline F's flights. Note that domestic flights in a country with a lower e_i are more efficient in terms of GHG emission than that in another country. Thus, the difference between e_H and e_F may capture the difference of factors affecting emission intensity of domestic routes in countries H and F, including aircraft fuel efficiency, route structure and load factor.

For the price discrimination case with $\beta = 0$, by substituting equations (5.20) and (5.21) into (5.26), and rearranging the equation gives:

$$\begin{aligned} & E^A - E^B \\ & = -\frac{\gamma[(12-\alpha\tau)e_H + 2(5+\alpha\tau)\delta e_K]}{4(\alpha\tau+5)(7-2\alpha\tau)}e_H - \frac{\gamma(e_H + 2\delta e_K)}{2(7-2\alpha\tau)}e_K - \frac{\gamma[(2-3\alpha\tau)e_H + 2(5+\alpha\tau)\delta e_K]}{4(\alpha\tau+5)(7-2\alpha\tau)}e_F \end{aligned} \tag{5.27}$$

Equation (5.27) suggests that the total GHG emissions may increase or decrease after the imposition of unilateral GHG pricing. It is a very interesting result. To deal with the GHG problem, many environmental groups advocate that given the absence of an

international agreement, a country ought to take unilateral actions to reduce its environmentally harmful emissions, as they at least contribute in the right direction (Hoel, 1991). Yet here we show that a country's unilateral action may possibly worsen the problem rather than improve it. To understand the intuition behind the result, we need to further examine the factors affecting the sign of world emission change. From equation (5.27), it can be shown that:

$$E^A - E^B > 0$$

$$\textit{if and only if} \quad \text{(i) } e_F > \frac{(12-\alpha\tau)e_H^2 + (5+\alpha\tau)[2e_H(\delta+1)+4e_K\gamma]e_K}{(3\alpha\tau-2)e_H - 2(5+\alpha\tau)\delta e_K},$$

$$\text{and} \quad \text{(ii) } (2-3\alpha\tau)e_H + 2(5+\alpha\tau)\delta e_K < 0. \tag{5.28}$$

Condition (5.28) suggests that total GHG emissions will increase after the imposition of the unilateral GHG pricing, if and only if the (per-flight) emissions for the domestic flights in country F (i.e., e_F) is larger than a threshold (i.e., (i) in (5.28) holds) and the domestic flights in country F increases after country H implements the GHG pricing (it would be the case when (ii) in (5.28) holds, see equation (5.21)). In Proposition 5.3, we find that domestic flights in country H and total output will decrease after imposing the GHG pricing, while the domestic flights in country F may increase. Thus, the emissions from domestic flights in country H and international flights will decrease, while that in country F may increase. These two opposing effects on total emissions depend on the emission intensity of the three routes. When e_F is very large, the latter effect may dominate. In other words, the unilateral GHG pricing may result in a shift of domestic

flights from one country to another, which may have a less efficient domestic network. As a result, the unilateral GHG pricing may increase world emissions. Although this Chapter does not explicitly consider global cooperation in control measures, the potential adverse impact of the unilateral GHG control action on world emissions control may provide an argument to support international cooperation on the issue.

Proposition 5.5 In the price discrimination case, when a country moves toward GHG pricing, total emissions may increase or decrease, depending on aircrafts' emission efficiency in that country and its foreign counterparts.

Proposition 5.5 may have significant implications for implementing GHG control measures unilaterally, like the inclusion of the aviation industry in the EU ETS. As EU airlines operating flights between EU airports may face a stricter regulation on GHG emissions (i.e., need to buy permits for their emissions), they may be required to charge a higher fare to their customers. As a result, more consumers may choose non-EU airlines, and thereby increasing hubbing activities outside the EU. For example, after the inclusion of the aviation sector in the EU ETS, it is expected that more domestic connecting flights in the EU for hubbing will be shifted to the U.S., as US airlines, which usually operate their hubs in the U.S., may gain a larger market share of the transatlantic market (Scheelhaase, et al. 2008). On the other hand, Table 5.1 shows the emission intensity of domestic flights in some selected routes in the U.S. and the EU. It suggests that the US domestic flights are less efficient than that in the EU in terms of GHG emissions. Thus, to evaluate the effectiveness of the inclusion of the aviation industry in the EU ETS, it is

important to take the change in GHG emissions outside the EU into account, as the unilateral GHG control measure may shift flights from a more efficient network (e.g., in the EU) to a less efficient one (e.g., in the U.S.).[102] However, the change in GHG emissions outside the EU due to the EU ETS is usually ignored in the present empirical studies (e.g., CE Delft, 2005; EC, 2006).

[Table 5.1 here]

For the uniform pricing case, by substituting equations (5.23) and (5.24) into (5.26), the change in total GHG emissions with the unilateral GHG pricing is given below:

$$E^A - E^B = -\frac{\gamma}{\Omega_2}(11e_H + 15e_K)(15e_K + 11e_H + 4e_F) < 0. \tag{5.29}$$

Equation (5.29) suggests that the total GHG emissions will decrease if a country unilaterally imposes GHG pricing. As shown in Proposition 5.4, domestic flights in countries H and F as well as international flights will decrease after imposing unilateral GHG pricing. Thus, the reduction in GHG emissions by (5.29) is the sum of the emissions reduction from the three routes. As shown in (5.29), the reduction is increasing in the (per-unit) emissions in the three routes (i.e., e_H, e_F and e_K).

[102] Note that for the U.S. and the EU moving towards more liberalizing air market, EU airlines may allow operating their hubs in the U.S. Given a stricter regulation on GHG emissions in the EU, EU airlines may operate more flights through their U.S. hubs.

Proposition 5.6 In the uniform pricing case, when a country moves toward GHG pricing, total GHG emissions will decrease.

5.7 Concluding Remarks

The main objective of this Chapter is to investigate the effects of unilateral GHG pricing implemented at a country's airports on the airline market competition and world emissions. We explored both the price discrimination case and the uniform pricing case, in which a country imposes a single charge at both hub and spoke airports. In the price discrimination case, we found that the unilateral GHG pricing may be positive or negative. This is due to the fact that positive charges are imposed to internalize the negative externality, while the country may subsidize airlines, which may confer a strategic advantage to its home airlines. Yet if the home airline's profit is not taken into account or in the uniform pricing case, the airport charges will increase when the country unilaterally implements GHG pricing. In such cases, a strategic advantage is conferred to foreign airlines. It is because the effects of such (positive) GHG charges have a less adverse impact on the foreign airlines than on the home airlines. Furthermore, the increase in airport charges due to the unilateral GHG pricing leads to a reduction in total market output, and the home airlines' market share will decrease. Consumer surplus of the air travel market will decrease. Finally, in the price discrimination case, total GHG emissions may increase after the unilateral GHG pricing is imposed. This may be due to the fact that domestic flights connecting hub and spoke airports may be shifted from a more efficient flight network to a less efficient one.

The Chapter has also raised a number of other issues and avenues for future research. First, our analysis assumes that the airlines' network structure is fixed. Given the airline industry moves toward more liberalization, airlines will have more flexibility to choose different networks for their flight operations. From a forward-looking perspective, it might be important to take the response of airlines network choice into account, when we examine the effects of GHG control measures on the airline industry. For example, note that the emission intensity is increasing in flight distance and the number of taking-offs and landings. As a result, given GHG pricing, airlines may be less likely to use HS networks than FC networks, since the HS networks usually involve longer flying distances and more taking-offs and landings, which would incur higher costs due to the GHG pricing. On the other hand, the unilateral GHG control measure may also lead to inefficient airlines network choice: For example, Albers, et al. (2008) argue that airlines may add an intermediate stop outside of the EU on long international flights, so that the EU emissions charge only applies to the final short leg. As a result, the flying distance and the number of taking-offs and landings may increase, thereby increasing world emissions (see, e.g., Morrell and Lu, 2007, for more discussion about the relationship between airlines network choice and environmental costs).

Second, we take the emission technology of aircrafts to be exogenous. In the long run, airlines may control its aircraft emissions by changing, for example, its speed of aircraft replacement, aircraft size and load factor. Thus, it is also practically relevant to take the airlines' decisions which affect emission technology into account, when we

examine the effects of GHG control measures. Finally, this Chapter considers a GHG pricing in a single market – the aviation market. Several empirical studies have been conducted to examine the side effects of the pricing on other industries, for example, tourism (Forsyth, 2008; Forsyth and Ho, 2008), automobile (Hofer, et al., 2008) and high-speed railways (Benito, 2008). Thus, it is interesting to take other related industries into account in our theoretical analysis.

Finally, this Chapter only considered GHG emissions, while other externalities due to air transportation, such as congestion discussed in Chapters 2 and 3, noise and local air pollution, are omitted in the analysis. Note that contemporary estimates (e.g., Quinet and Vickerman, 2004; Schipper, 2004) indicate that the environmental costs of noise and air pollution are at least as large as the costs of GHG emissions. The presence of these other externalities would not be a concern if they were independent of GHG emissions. However, they are likely to be interdependent to a degree. Congestion probably contributes to air pollution and GHG emissions through greater fuel consumption as aircraft fly longer routes, queue on taxiways, and circle airports waiting to land. For similar reasons, congestion contributes to noise. More generally, efforts to reduce one externality are likely to affect the magnitude and spatial distribution of other externalities. It would be interesting to extend this Chapter to consider the interdependence of GHG emissions and other externalities into account in the analysis.

Figure 5.1 Network structure

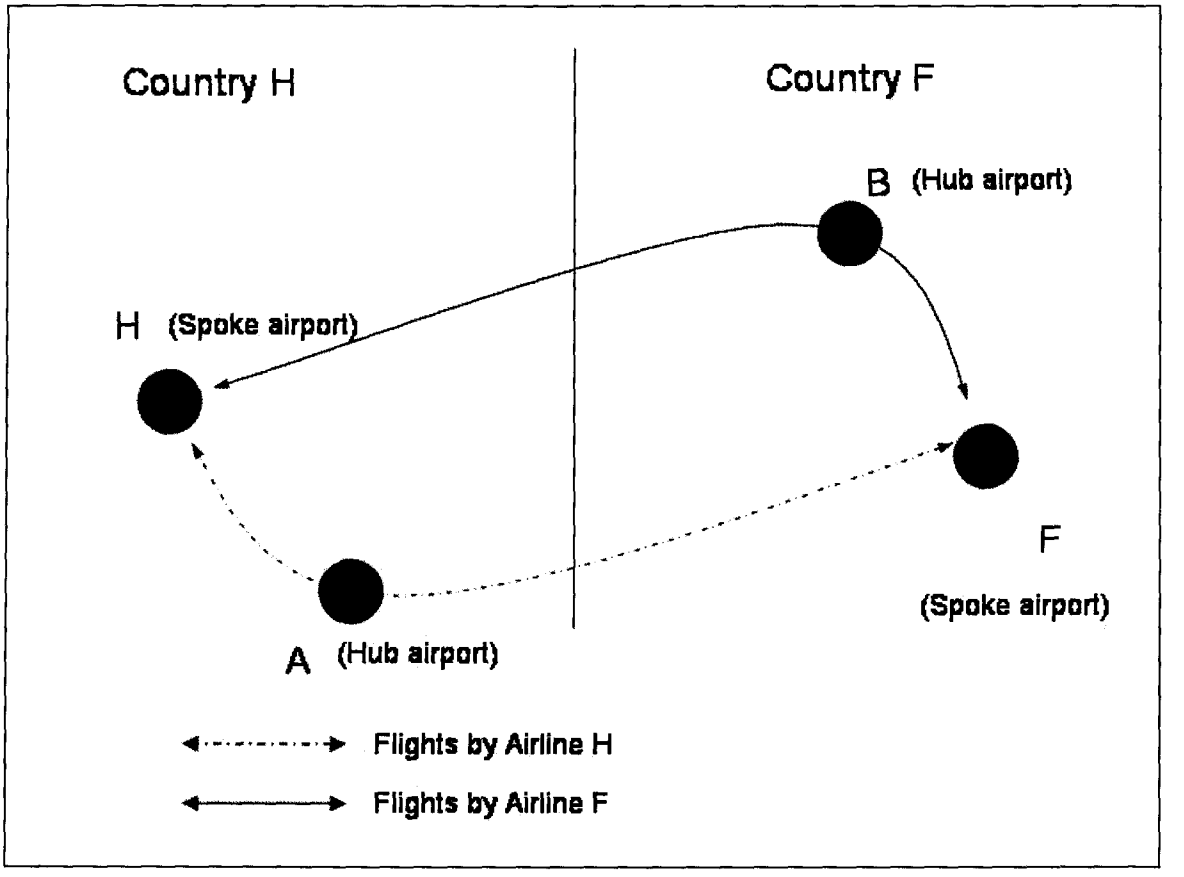


Figure 5.2 Consumers' distribution and airlines' catchment areas

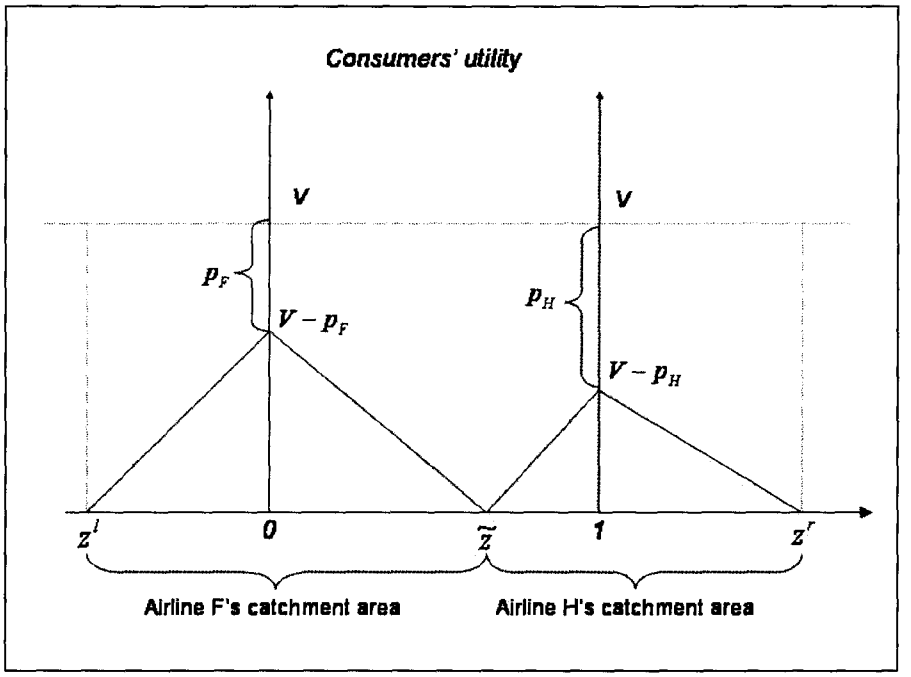


Figure 5.3 The change in market size and its share after GHG pricing

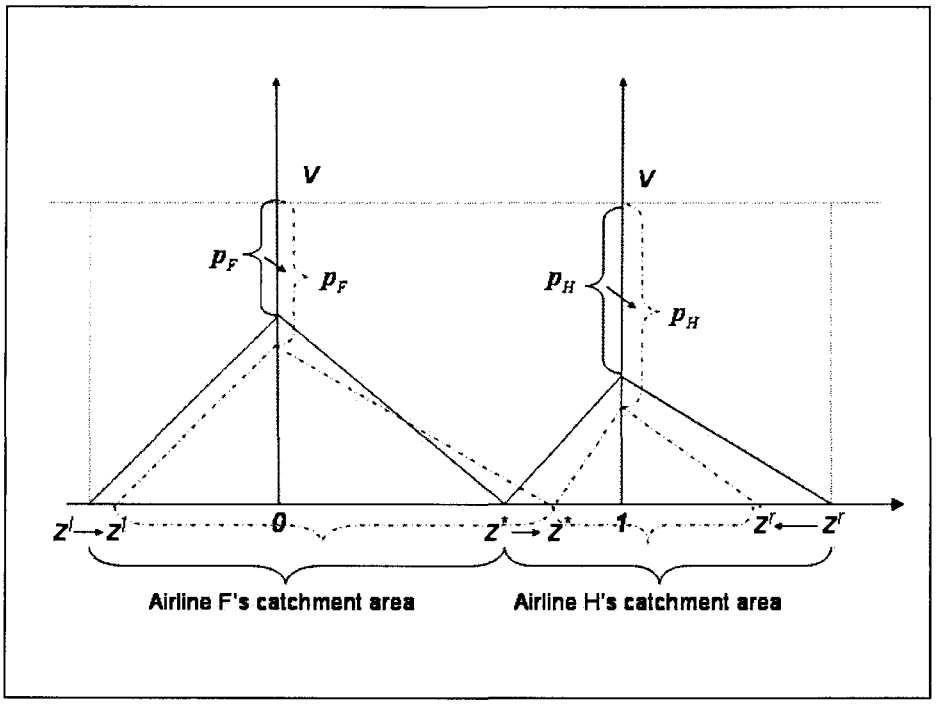


Table 5.1a CO_2 Emissions of selected routes in US domestic markets[a]

From New York to	Per-passenger CO_2 emission (kg)*	Distance (km)	Emission per distance (kg/km)
Columbus	130.3	774	0.168
Detroit	131.4	816	0.161
Louisville	154.6	1061	0.146
Chicago	157.6	1186	0.133
New Orleans	197.4	1889	0.105
Denver	206.7	2607	0.079
Average	163.0	1389	0.132

Table 5.1b CO_2 Emissions of selected routes in EU domestic markets

From London to	Per-passenger CO_2 emission (kg)*	Distance (km)	Emission per distance (kg/km)
Berlin	97.0	946	0.103
Milan	114.2	978	0.117
Barcelona	110.7	1146	0.097
Rome	134.8	1442	0.093
Lisbon	141.7	1564	0.091
Athens	217.6	2242	0.097
Average	136.0	1386	0.099

*Economy class

Source: ICAO Carbon Emissions Calculator

[a.] It is shown from the two tables that emission per km decline strongly in the U.S., but not in the EU. Note that the emissions per passenger tend to fall with distance because a significant portion of flight emissions occur during taking-off and landing, and also because larger aircraft are used for long-distance flight which produce fewer emissions per passenger (e.g., Schipper, 2004).

5.8 References

Albers, S., Buhne, J. and Peters, H. (2008) "Will the EU-ETS Instigate Airline Network Reconfigurations?" *Journal of Air Transport Management, forthcoming.*

Arthur Andersen (2001) Emissions Trading for Aviation, a study commissioned by the International Air Transport Association, London.

Basso, L.J. and Zhang, A. (2007a) "Congestible Facility Rivalry in Vertical Structures", *Journal of Urban Economics*, 61, 218-237.

Basso, L.J. and Zhang, A. (2007b) "An Interpretative Survey of Analytical Models of Airport Pricing", in: Lee, D. (Ed.), *Advances in Airline Economics*, Vol.2, Elsevier, 89-124.

Benito, A. (2008) "Climate Change Impact as a Constraint to the Air Transport Growth: The European Case", Paper presented at World Conference on Transport Research, Berkeley, California.

Brander, J.A. (1995) "Strategic Trade Policy", in: Grossman, G. and Rogoff, K. (Eds.), *Handbook of International Economics*, Vol. III, Elsevier Science, Amsterdam.

Brander, J.A. and Zhang, A. (1990) "Market Conduct in the Airline Industry: An Empirical Investigation", *RAND Journal of Economics*, 21, 567–583.

Brueckner, J.K. (2004) "Network Structure and Airline Scheduling", *Journal of Industrial Economics*, 52, 291-312.

Brueckner, J.K. and Flores-Fillol, R. (2007) "Airline Schedule Competition", *Review of Industrial Organization*, 30, 161-177.

Brueckner, J.K. and Girvin, R. (2008) "Airport Noise Regulation, Airline Service Quality, and Social Welfare", *Transportation Research Part B*, 42, 19-37.

Brueckner, J.K. and Zhang, Y. (2001) "A Model of Scheduling in Airline Networks: How a Hub-and-Spoke System Affects Flight Frequency, Fares and Welfare", *Journal of Transport Economics and Policy*, 35, 195-222.

Caves, D.W., Christensen, L.R. and Tretheway, M.W. (1984) "Economies of Density versus Economics of Scale: Why Trunk and Local Service Airline Costs Differ", *RAND Journal of Economics*, 15(4), 471–489.

CE Delft (2005) Giving Wings to Emission Trading – Inclusion of Aviation under the European Emission Trading System (ETS): Design and Impacts, study commissioned by the European Commission, Delft.

CE Delft (2007) Allocation of Allowances for Aviation in the EU ETS: The Impact on the Profitability of the Aviation Sector under High Levels of Auctioning. Final report, Delft.

Chua, C.L., Kew, H. and Yong, J. (2005) "Airlines Code-Share Alliances and Costs: Imposing Concavity on Translog Cost Function Estimation", *Review of Industrial Organization*, 26(4), 461-487.

Ciliberto, F. and Williams, J.W. (2007) "Limited Access to Airport Facilities and Market Power in the Airline Industry", Working paper, Department of Economics, University of Virginia.

Copeland, B.R. and Taylor, M.S. (1994) "North-South Trade and the Environment", *Quarterly Journal of Economics*, 109, 755–787.

Douglas, G.W. and Miller, J.C. (1974) "Quality Competition, Industry Equilibrium, and Efficiency in the Price-constrained Airline Market", *American Economic Review*, 64 (4), 657-669.

EC (2006) Full Assessment Report, accompanying document to the Proposal for a Directive of the European Parliament and of the Council, European Commission.

Ederington, J., Levinson, A. and Minier, J. (2005) "Footloose and Pollution-free", *Review of Economics and Statistics*, 87, 92-100.

Eskeland, G.S. and Harrison, A.E. (2003) "Moving To Greener Pastures? Multinationals and The Pollution Haven Hypothesis", *Journal of Development Economics*, 70, 1-23.

Forsyth, P. (2008) "Emissions Trading Schemes – Should International Aviation be Excluded?", *manuscript*.

Forsyth, P. and Ho, T.V. (2008) "Aviation's Indirect Carbon Emissions: Measurement and Implications", Paper for the German Aviation Research Society Workshop Aviation and Climate Change, Amsterdam, July.

Grossman, G.M. and Krueger, A.B. (1991) "Environmental Impacts of a North American Free Trade Agreement", NBER Working Paper No. 3914, National Bureau of Economic Research, Cambridge, Massachusetts.

Hansen, M.M., Gillen, D., Djafarian-Tehrani, R. (2001) "Aviation Infrastructure Performance and Airline Cost: A Statistical Cost Estimation Approach", *Transportation Research Part E*, 37, 1-23.

Hendricks, K., Piccione, M. and Tan, G. (1995) "The Economics of Hubs: The Case of Monopoly", *Review of Economic Studies*, 62, 83-99.

Hendricks, K., Piccione, M. and Tan, G. (1999) "Equilibria in Networks", *Econometrica*, 67, 1407-1434.

Hodgkinson, D., Coram, A. and Garner, R. (2007) Strategies for Airlines and Aircraft Emissions and Climate Change: Sustainable, Long-Term Solutions, Hodgkinson Group, Working Paper No 2, June.

Hoel, M. (1991) "Global Environmental Problems: The Effects of Unilateral Actions Taken by One Country", *Journal of Environmental Economics and Management*, 20, 55-70.

Hofer, C., Dresner, M. and Windle, R. (2008) "The Environmental Effects of Airline Carbon Emissions Taxation in the U.S.", Paper presented at Air Transport Research Society Conference, Athens.

IATA (2006) Economics Briefing Impact of the Rise in UK Air Passenger Duty, International Air Transport Association.

ICAO (2007) Report on Voluntary Emissions Trading for Aviation, Preliminary Edition, International Civil Aviation Organization.

IPCC (2007) Fourth Assessment Report, Technical Summary, Intergovernmental Panel on Climate Change.

Kahn, M.E. (2003) "The Geography of US Pollution Intensive Trade: Evidence from 1958 to 1994", *Regional Science and Urban Economics*, 33, 383-400.

Kumbhakar, S.C. (1990) "A Reexamination of Returns to Scale, Density and Technical Progress in U.S. Airlines", *Southern Economic Journal*, 57, 428-442.

Kumbhakar, S.C. (1992) "Allocative Distortions, Technical Progress, and Input Demand in U.S. Airlines: 1970–1984", *International Economic Review*, 33, 723-737.

Levinson, A. and Taylor, M.S. (2004) "Unmasking the Pollution Haven Effect", NBER Working Paper No.10629, National Bureau of Economic Research, Cambridge, Massachusetts.

Morrell, P. and Lu, C. (2007) "The Environmental Cost Implication of Hub-hub versus Hub By-pass Flight Networks", *Transportation Research Part D*, 12, 143-157.

Olsthoorn, X. (2001) "Carbon Dioxide Emissions from International Aviation 1905-2050", *Journal of Air Transport Management*, 7, 87-93.

Oum, T.H., 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.H., Zhang, A. and Zhang, Y. (1995) "Airline Network Rivalry", *Canadian Journal of Economics*, 28, 836-857.

Quinet, E. and Vickerman, R. (2004), *Principles of Transport Economics*, Cheltenham and Northampton, Mass: Edward Elgar.

Scheelhaase, J.D. and Grimme, W.G. (2007) "Emissions Trading for International Aviation – An Estimation of the Economic Impact on Selected European Airlines", *Journal of Air Transport Management*, 13, 253-263.

Scheelhaase, J.D., Grimme, W.G. and Schaefer, M. (2008) "The Inclusion of Aviation into the EU Emission Trading Scheme – Impacts on Competition between European and non-European Network Airlines", *manuscript*.

Schipper, Y. (2004) "Environmental Costs in European Aviation", *Transport Policy*, 11, 141-154.

Small, K.A. and Verhoef, E.T. (2007) The Economics of Urban Transportation, Routledge, London.

Stern, N. (2006) Stern Review Report on the Economics of Climate Change, a report to the Prime Minister and the Chancellor of the Exchequer on the Economics of Climate Change, United Kingdom.

Ubbels, B., and Verhoef, E.T. (2008) "Governmental Competition in Road Charging", *Regional Science and Urban Economics*, 38, 174-190.

World Bank (2007) State and the Trends of Carbon Market 2007, The World Bank Institute, Washington, D.C.

Xu, X. (2000) "International Trade and Environmental Regulation: Time Series Evidence and Cross Section Test", *Environmental and Resource Economics*, 17, 233-257.

Zhang, A., Zhang, Y. and Clougherty, J.A. (2008) "Competition and Regulation in Air Transport", *Handbook of Transport Economics*, Edward Edgar, forthcoming.

Chapter 6 Conclusions

6.1 Thesis Summary and Future Research

This thesis has examined several public policies on the problems of capacity shortage at airports and aviation GHG emission controls through four Chapters. In particular, Chapters 2, 3 and 5 addressed several issues of major policy interest in air passenger and freight transport: growing congestion delays at airports and on intermodal transport links, GHG emissions and exercise of market power by airports and airlines. Each of these issues arises from a type of market failure. The three Chapters analyzed how these failures can be rectified by imposing usage charges or taxes. They also considered the welfare-distributive effects of the policies – which is a critical determining factor whether the policies are likely to be adopted. Along with the analytical works in the three Chapters, Chapter 4 examined the factors affecting efficiency of airport capacity utilization. In the following paragraphs, I will provide a brief summary of findings presented in the previous Chapters and identify some future research areas to conclude this thesis.

Chapter 2 examined the airport congestion-pricing problem that arises when passengers are heterogeneous in time valuation, and investigated welfare-redistributive issues after the imposition of airport congestion pricing. In the case of variable passenger time costs, we found that the airlines' self-internalization behavior is incomplete in the sense that each airline may not fully internalize the congestion costs it imposed on its

peak-period passengers. On the other hand, our results suggested that although the implementation of the congestion pricing may be welfare-improving, it may not be Pareto-improving. This may explain why congestion pricing appears unpopular in practice. Chapter 3 investigated the effects of congestion pricing implemented at a gateway on its hinterland's optimal road tolls, road congestion and social welfare. Our analysis showed that when the gateway maximizes the joint profit of itself and its carriers, the gateway charge will rise after congestion pricing is implemented at the gateway. This increase in gateway charge will lead to lower road tolls and lower social welfare for the hinterland.

Chapter 4 assessed the effects of China's competition and aviation policy reform on both the level and growth of its airports efficiency. We found empirical evidence that (i) publicly listed airports are significantly more efficient than non-listed airports; (ii) the airport localization program could result in an improvement in the efficiency of airports; (iii) airports with more competition are more efficient than their counterparts. Chapter 5 investigated the effects of unilateral GHG pricing implemented at a country's airports on the airline market competition and world emissions. We found that the unilateral GHG pricing may confer a strategic advantage to foreign airlines. Furthermore, the increase in airport charges due to the unilateral GHG pricing leads to a reduction in total market output, while home airlines' market share will decrease. Finally, world emissions may increase or decrease after the unilateral GHG pricing is imposed.

There are many areas of future research that can be identified from the previous Chapters. First, a valuable extension to Chapter 2 is to consider an asymmetric airline case. Daniel (2001) empirically found that Northwest enjoyed the greatest gain due to the implementation of congestion pricing as compared to regional counterparts. It may be because in the air travel market, legacy carriers usually serve a higher portion of business travelers, who are expected to benefit more from congestion pricing schemes as discussed in Chapter 2, while low cost carriers serve a higher portion of leisure travelers (O'Connell and Williams, 2005). Thus, the potential competitive issues in the downstream carrier market may need to be taken into consideration in regards to implementing congestion pricing. In Chapter 3, we considered a serial transport network, in which cargos to the gateway will only use the road facility at the hinterland. In extending the present analysis, it is practically relevant to consider the inland modal split for gateway traffic to its hinterland. Second, it is also interesting to examine the competition between transport facilities (gateways or roads, or both) (e.g., Basso and Zhang, 2007; De Borger, et al., 2008).

Chapter 4 examined efficiency of international and local airports in China, in which due to data limitation, we only considered the airports' efficiency of utilizing physical capital. In other words, only physical infrastructures, such as runways and terminals, are considered as inputs in measuring airport efficiency. To measure the overall efficiency, further research may need to consider extending our analysis by using a larger set of inputs and outputs. On the other hand, it is important to extend our analysis to compare the efficiency of airports in the Northeast Asia region. It is because open-

skies agreements among China, South Korea and Japan are likely to be reached in the near future. A better understanding about their airports, especially those local airports, may shed some light on the open-skies agreement negotiation.

Chapter 5 examined GHG control measures on the aviation sector, which is an issue being practically important, but discussion of the issue so far lacks much relevant theoretical work. The Chapter has raised a number of other issues and avenues for future research. First, this Chapter considered GHG pricing in a single market – the aviation market. Several empirical studies have been conducted to examine the side effects of the pricing on other industries, for example, tourism (Forsyth, 2008; Forsyth and Ho, 2008), automobile (Hofer, et al., 2008) and high-speed railways (Benito, 2008). Thus, it is interesting to take other related industries into account in our theoretical analysis. Second, we only considered airline pricings as endogenous. Yet in response to GHG control measures implemented by governments, airlines may also maximize their profits by adjusting other operating decisions, such as load factors, aircraft size, network, etc. Thus, it is also practically relevant to consider those decisions in a theoretical analysis. This may shed some light on the discussion about the mechanisms chosen to control the aviation GHG emissions.

Finally, to rectify the market failures discussed in Chapter 2, 3 and 5, this Chapter particularly focused on a price mechanism (e.g., usage charges or taxes). However, other policy tools are also implemented or discussed in practice. The choice of the price-based and quantity-based policy tools in the air transport market is related to the discussion in

Wetizman (1974). The paper, which is referring to pollution management (amongst others), showed that under uncertainty about benefits and costs of pollution, the expected welfare depends on the choice between prices (e.g., pollution taxes) or quantities (e.g., emission standards) as instruments. The analytical work on comparing different policy instruments in dealing with airport congestion is still limited. A few examples include Czerny (2007) which evaluates the performance of slot constraints and congestion pricing in alleviating airport congestion. The paper found that due to the network character of the airline industry, the demand for airport capacities normally is complementary, which favors the use of slot constraints compared to congestion pricing from a social efficiency point of view. In contrast, for monopolistic airports, prices as instruments constitute a dominant choice. On the other hand, Brueckner (2008) found that a uniform optimal airport charge may lead to the result that flight volumes tend to be too low for large carriers and too high for small carriers. But quantity-based regimes, where the airport authority allocates a fixed number of slots via free distribution or an auction, lead carriers to treat total flight volume (and thus congestion) as fixed, and this difference generates an efficient outcome as long as the number of slots is optimally chosen.

On the other hand, both quantity-based and price-based mechanisms have also been implemented in practice in controlling GHG emissions – the major issue examined in Chapter 5. The price-based mechanism (e.g., emission taxes) to control GHG emissions has been discussed in the analytical model in the Chapter. For the quantity-based measures, a country committed to limiting its GHG emissions could implement a quota system. Under such a scheme, airlines must obtain quotas for their GHG emissions.

The government may sell the emission quotas to airlines, or allocate the quotas to them based on, for example, grandfathering rule. A similar approach with the quota system to control GHG emissions is the so-called "cap-and-trade" scheme. Under the scheme, a government imposes a cap on each airline's GHG emission level.[103] An example of the cap-and-trade mechanism is the inclusion of the aviation sector in the EU emission trading system (EU ETS). Given this, if the emissions from the airline are more than the cap, it is required to purchase permits from other sectors or from the project based Kyoto instruments "Joint Implementation" and "Clean Development Mechanism"; if the emissions are less than the cap, the airline can sell its permits to other participating firms. While the quantity-based GHG control mechanisms are not analytical examined in this thesis, it would be practically important to examine those mechanisms in future works.

[103] For example, under the EU ETS, the total number of allowances allocated to an airline will be calculated on the basis of the average total emissions reported for the years 2004-2006 by that airline.

6.2 References

Basso, L.J. and Zhang, A. (2007) "Congestible Facility Rivalry in Vertical Structures", *Journal of Urban Economics*, 61(2), 218-237.

Benito, A. (2008) "Climate Change Impact as a Constraint to the Air Transport Growth: The European Case", Paper presented at World Conference on Transport Research, Berkeley, California.

Brueckner, J.K. (2008) "Price vs. Quantity-Based Approaches to Airport Congestion Management", *manuscript.*

Czerny, A. (2007) "Congestion Management under Uncertainty in a Two-Airport System", *manuscript.*

Daniel, J.I. (2001) "Distributional Consequences of Airport Congestion Pricing," *Journal of Urban Economics*, 50, 230-258.

De Borger, B., Proost, S. and Van Dender, K. (2008) "Private Port Pricing and Public Investment in Port and Hinterland Capacity", *Journal of Transport Economics and Policy*, 42(3), 527-561.

Forsyth, P. (2008) "Emissions Trading Schemes – Should International Aviation be Excluded?", *manuscript.*

Forsyth, P. and Ho, T.V. (2008) "Aviation's Indirect Carbon Emissions: Measurement and Implications", Paper for the German Aviation Research Society Workshop Aviation and Climate Change, Amsterdam, July.

Hofer, C., Dresner, M. and Windle, R. (2008) "The Environmental Effects of Airline Carbon Emissions Taxation in the U.S.", Paper presented at Air Transport Research Society Conference, Athens.

O'Connell, J.F. and G. Williams (2005) "Passengers' Perceptions of Low Cost Airlines and Full Service Carriers: A Case Study Involving Ryanair, Aer Lingus, Air Asia and Malaysia Airlines," *Journal of Air Transport Management*, 11(4), 259-272.

Weitzman, M.L. (1974) "Prices vs. Quantities", *Review of Economic Studies*, 41(4), 477-491.

Appendix A Glossary of variables for Chapter 3

Y_G : units of cargo consumed in region G

X : units of cargo consumed in region H

Y_H : region H's local traffic

α : intercepts of the indirect demands

β : (absolute) slopes of the indirect demands

K_G : the gateway's capacity

K_H : the road's capacity

C_G : the gateway's unit operating cost of providing service to carriers

$\bar{t}_H$: the road toll imposed on local road users in the uniform-pricing case

Local road users:

ρ_H : the full price perceived by local road users

τ_H : the road toll imposed on local road users in the price-discrimination case

γ_H : the delay cost parameter for local road users

Carriers:

N : number of carriers

P_G : price imposed on cargo consumers in region G

P_X : price imposed on cargo consumers in region H

c_G : the fixed unit operating cost at the gateway

c_H : the fixed unit operating cost at the road

t_G : the gateway charge

t_H : the road toll imposed on cargo in the price-discrimination case

γ_G : the delay cost parameter for carriers at the gateway

γ_X : the delay cost parameter for carriers at the road

Appendix B Proofs in Chapter 3

As we will extensively use the following expression in our proofs, we first define:

$$\Delta \equiv (\beta_G + \frac{\gamma_G}{K_G})[\beta_X + \frac{\gamma_X}{K_H}(1+\frac{\partial z_H}{\partial X})] + \beta_G \frac{\gamma_G}{K_G} > 0.$$

Proof of Lemma 3.1: Totally differentiating Ω_{Y_G} and Ω_X given in (3.7) and (3.8) with respect to t_G, and solving the resulting equations, we obtain:

$$\frac{\partial Y_G}{\partial t_G} = -\frac{N}{N+1}\frac{1}{\Delta}[\beta_X + \frac{\gamma_X}{K_H}(1+\frac{\partial z_H}{\partial X})] < 0, \qquad \frac{\partial X}{\partial t_G} = -\frac{N}{N+1}\frac{\beta_G}{\Delta} < 0.$$

Using these inequalities, we have:

$$\frac{\partial Y_H}{\partial t_G} = \frac{\partial z_H}{\partial X}\frac{\partial X}{\partial t_G} > 0, \qquad \frac{\partial (X+Y_H)}{\partial t_G} = (1+\frac{\partial z_H}{\partial X})\frac{\partial X}{\partial t_G} = -\frac{N}{N+1}\frac{\beta_G}{\Delta}(1+\frac{\partial z_H}{\partial X}) < 0.$$

Q.E.D.

Proof of Lemma 3.2: Totally differentiating Ω_{Y_G} and Ω_X given in (3.7) and (3.8) with respect to t_H and solving the resulting equations, we obtain:

$$\frac{\partial Y_G}{\partial t_H} = \frac{N}{N+1}\frac{1}{\Delta}\frac{\gamma_G}{K_G} > 0, \qquad \frac{\partial X}{\partial t_H} = -\frac{N}{N+1}\frac{1}{\Delta}(\beta_G + \frac{\gamma_G}{K_G}) < 0.$$

Using these results, we have:

$$\frac{\partial Y_H}{\partial t_H} = \frac{\partial z_H}{\partial X}\frac{\partial X}{\partial t_H} = -\frac{N}{N+1}\frac{1}{\Delta}(\beta_G + \frac{\gamma_G}{K_G})\frac{\partial z_H}{\partial X} > 0, \qquad \frac{\partial (X+Y_G)}{\partial t_H} = -\frac{N}{N+1}\frac{\beta_G}{\Delta} < 0,$$

$$\frac{\partial (X+Y_H)}{\partial t_H} = (1+\frac{\partial z_H}{\partial X})\frac{\partial X}{\partial t_H} = -\frac{N}{N+1}\frac{1}{\Delta}(\beta_G + \frac{\gamma_G}{K_G})(1+\frac{\partial z_H}{\partial X}) < 0.$$

Next, totally differentiating Ω_{Y_G} and Ω_X with respect to τ_H yields:

$$\frac{\partial Y_G}{\partial \tau_H} = \frac{N}{N+1}\frac{1}{\Delta}\frac{\gamma_G}{K_G}\frac{\gamma_X}{K_H}\frac{\partial z_H}{\partial \tau_H} < 0, \qquad \frac{\partial X}{\partial \tau_H} = -\frac{N}{N+1}\frac{1}{\Delta}(\beta_G + \frac{\gamma_G}{K_G})\frac{\gamma_X}{K_H}\frac{\partial z_H}{\partial \tau_H} > 0.$$

Using these results, we have:

$$\frac{\partial Y_H}{\partial \tau_H} = \frac{\partial z_H}{\partial \tau_H} + \frac{\partial z_H}{\partial X}\frac{\partial X}{\partial \tau_H} < 0, \qquad \frac{\partial Y_G}{\partial \tau_H} + \frac{\partial X}{\partial \tau_H} = -\frac{N}{N+1}\frac{1}{\Delta}\beta_G\frac{\gamma_X}{K_H}\frac{\partial z_H}{\partial \tau_H} > 0,$$

$$\frac{\partial (X+Y_H)}{\partial \tau_H} = \frac{\partial z_H}{\partial \tau_H} - \frac{N}{N+1}\frac{1}{\Delta}(\beta_G + \frac{\gamma_G}{K_G})\frac{\gamma_X}{K_H}\frac{\partial z_H}{\partial \tau_H}(1+\frac{\partial z_H}{\partial X}) < 0.$$

For the last inequality, we note that although the first term on the RHS is negative and the second term is positive, it can be shown, by substituting Δ and after some tedious manipulations, that the overall effect is negative. *Q.E.D.*

Proof of Lemma 3.3: First prove Part 2. Totally differentiating Ω_{Y_G} and Ω_X given in (3.7) and (3.8) with respect to $\bar{t}_H$ and solving the resulting equations, we obtain:

$$\frac{\partial Y_G}{\partial \bar{t}_H} = \frac{N}{N+1}\frac{1}{\Delta}\frac{\gamma_G}{K_G}(1+\frac{\gamma_X}{K_H}\frac{\partial z_H}{\partial \bar{t}_H}), \qquad \frac{\partial X}{\partial \bar{t}_H} = -\frac{N}{N+1}\frac{1}{\Delta}(\beta_G + \frac{\gamma_G}{K_G})(1+\frac{\gamma_X}{K_H}\frac{\partial z_H}{\partial \bar{t}_H}),$$

$$\frac{\partial (X+Y_G)}{\partial \bar{t}_H} = -\frac{N}{N+1}\frac{1}{\Delta}\beta_G(1+\frac{\gamma_X}{K_H}\frac{\partial z_H}{\partial \bar{t}_H}).$$

By (3.5), we have:

$$1+\frac{\gamma_X}{K_H}\frac{\partial z_H}{\partial \bar{t}_H} = \frac{\beta_H K_H - (\gamma_X - \gamma_H)}{\beta_H K_H + \gamma_H}. \tag{A.3.1}$$

With this identity we can show:

$$1+\frac{\gamma_X}{K_H}\frac{\partial z_H}{\partial \bar{t}_H} > 0 \quad \Leftrightarrow \quad \gamma_X - \gamma_H < \beta_H K_H$$

Part 2 of Lemma 3.3 then follows from this relationship as well as the previous results.

To show Part 1 of the lemma, we first show, using (A.3.1), that $\partial(Y_H + X)/\partial \bar{t}_H < 0$:

$$\frac{\partial(Y_H + X)}{\partial \bar{t}_H} = \frac{\partial z_H}{\partial \bar{t}_H} + \left(1 + \frac{\partial z_H}{\partial X}\right)\frac{\partial X}{\partial \bar{t}_H} = \frac{\partial z_H}{\partial \bar{t}_H} - \beta_H \frac{\partial z_H}{\partial \bar{t}_H}\frac{\partial X}{\partial \bar{t}_H}$$

$$= \frac{\partial z_H}{\partial \bar{t}_H}[1 + \beta_H \frac{N}{N+1} \frac{1 + \frac{\gamma_X}{K_H}\frac{\partial z_H}{\partial \bar{t}_H}}{\beta_X + \frac{\gamma_X}{K_H}(1 + \frac{\partial z_H}{\partial X}) + \beta_G \frac{\gamma_G}{K_G} / (\beta_G + \frac{\gamma_G}{K_G})}]$$

$$= \frac{\partial z_H}{\partial \bar{t}_H}[\beta_X + \beta_G \frac{\gamma_G}{K_G} / (\beta_G + \frac{\gamma_G}{K_G}) + \frac{\gamma_X}{K_H}(1 + \frac{\partial z_H}{\partial X}) + \beta_H \frac{N}{N+1}(1 + \frac{\gamma_X}{K_H}\frac{\partial z_H}{\partial \bar{t}_H})]$$

$$= \frac{\partial z_H}{\partial \bar{t}_H}[\beta_X + \beta_G \frac{\gamma_G}{K_G} / (\beta_G + \frac{\gamma_G}{K_G}) + \frac{\gamma_X}{K_H} \cdot \frac{\beta_H K_H}{\beta_H K_H + \gamma_H} + \beta_H \cdot \frac{N}{N+1} \cdot \frac{\beta_H K_H - (\gamma_X - \gamma_H)}{\beta_H K_H + \gamma_H}]$$

$$= \frac{\partial z_H}{\partial \bar{t}_H}[\beta_X + \beta_G \frac{\gamma_G}{K_G} / (\beta_G + \frac{\gamma_G}{K_G}) + \frac{\beta_H \gamma_X (N+1) + N\beta_H (\beta_H K_H - (\gamma_X - \gamma_H))}{(N+1)(\beta_H K_H + \gamma_H)}]$$

$$= \frac{\partial z_H}{\partial \bar{t}_H}[\beta_X + \beta_G \frac{\gamma_G}{K_G} / (\beta_G + \frac{\gamma_G}{K_G}) + \frac{\beta_H (\gamma_X + N\beta_H K_H + N\gamma_H)}{(N+1)(\beta_H K_H + \gamma_H)}] < 0$$

We then show, also using (A.3.1), that a sufficient condition for $\partial Y_H / \partial \bar{t}_H < 0$ is $\gamma_X \geq \gamma_H$:

$$\frac{\partial Y_H}{\partial \bar{t}_H} = \frac{\partial z_H}{\partial \bar{t}_H} + \frac{\partial z_H}{\partial X}\frac{\partial X}{\partial \bar{t}_H} = \frac{\partial z_H}{\partial \bar{t}_H} + \frac{\gamma_H}{K_H}\frac{\partial z_H}{\partial \bar{t}_H}\frac{\partial X}{\partial \bar{t}_H}$$

$$= \frac{\partial z_H}{\partial \bar{t}_H}[1 - \frac{\gamma_H}{K_H} \frac{N}{N+1} \frac{1 + \frac{\gamma_X}{K_H}\frac{\partial z_H}{\partial \bar{t}_H}}{\beta_X + \frac{\gamma_X}{K_H}(1 + \frac{\partial z_H}{\partial X}) + \beta_G \frac{\gamma_G}{K_G} / (\beta_G + \frac{\gamma_G}{K_G})}]$$

$$= \frac{\partial z_H}{\partial \bar{t}_H}[\beta_X + \beta_G \frac{\gamma_G}{K_G} / (\beta_G + \frac{\gamma_G}{K_G}) + \frac{\gamma_X}{K_H}(1 + \frac{\partial z_H}{\partial X}) - \frac{\gamma_H}{K_H}\frac{N}{N+1}(1 + \frac{\gamma_X}{K_H}\frac{\partial z_H}{\partial \bar{t}_H})]$$

$$= \frac{\partial z_H}{\partial \bar{t}_H}[\beta_X + \beta_G \frac{\gamma_G}{K_G} / (\beta_G + \frac{\gamma_G}{K_G}) + \frac{\gamma_X}{K_H} \cdot \frac{\beta_H K_H}{\beta_H K_H + \gamma_H} - \frac{\gamma_H}{K_H} \cdot \frac{N}{N+1} \cdot \frac{\beta_H K_H - (\gamma_X - \gamma_H)}{\beta_H K_H + \gamma_H}]$$

$$= \frac{\partial z_H}{\partial \bar{t}_H}[\beta_X + \beta_G \frac{\gamma_G}{K_G} / (\beta_G + \frac{\gamma_G}{K_G}) + \frac{\beta_H K_H [\gamma_X (N+1) - \gamma_H N] + \gamma_H N(\gamma_X - \gamma_H)}{K_H (N+1)(\beta_H K_H + \gamma_H)}]$$

where the first term is negative whereas the second term is positive if $\gamma_X \geq \gamma_H$, which is satisfied (recall that $\gamma_X > \gamma_H$ is our maintained assumption). *Q.E.D.*

Proof of Proposition 3.2: Differentiating (3.20) with respect to t_G, we obtain:

$$\frac{\partial \tau_H^*}{\partial t_G} = \frac{\partial Y_H}{\partial t_G}\frac{\gamma_H}{K_H} + \frac{\partial X}{\partial t_G}\frac{\gamma_X}{K_H} = \frac{\partial z_H}{\partial X}\frac{\partial X}{\partial t_G}\frac{\gamma_H}{K_H} + \frac{\partial X}{\partial t_G}\frac{\gamma_X}{K_H}$$

$$< \frac{\partial z_H}{\partial X}\frac{\partial X}{\partial t_G}\frac{\gamma_X}{K_H} + \frac{\partial X}{\partial t_G}\frac{\gamma_X}{K_H} \qquad \text{(recall } \gamma_X > \gamma_H \text{ is our maintained assumption)}$$

$$= (1 + \frac{\partial z_H}{\partial X})\frac{\partial X}{\partial t_G}\frac{\gamma_X}{K_H} < 0.$$

Differentiating (3.21) with respect to t_G, we obtain:

$$\frac{\partial t_H^*}{\partial t_G} = \frac{\gamma_H - \gamma_X}{K_H}\frac{\partial Y_H}{\partial t_G} - \frac{\partial X}{\partial t_G}\beta_X - \frac{\partial X}{\partial t_G}/\frac{\partial X}{\partial t_H} < -\frac{\partial X}{\partial t_G}\beta_X - \frac{\partial X}{\partial t_G}/\frac{\partial X}{\partial t_H}$$

$$= \frac{N}{N+1}\frac{\beta_G}{\Delta}\beta_X - \beta_G/(\beta_G + \frac{\gamma_G}{K_G}) = \beta_G[\frac{N}{N+1}\frac{\beta_X}{\Delta} - 1/(\beta_G + \frac{\gamma_G}{K_G})]$$

$$= [\beta_G/(\beta_G + \frac{\gamma_G}{K_G})]\{\frac{N}{N+1}\beta_X/[\beta_X + \frac{\gamma_X}{K_H}(1 + \frac{\partial z_H}{\partial X}) + \beta_G\frac{\gamma_G}{K_G}/(\beta_G + \frac{\gamma_G}{K_G})] - 1\} < 0.$$

Proof of Proposition 3.3: Rewrite the road toll (3.23) as:

$$\bar{t}_H^* = \frac{(Y_H + X)\frac{\gamma_H}{K_H}\frac{\partial Y_H}{\partial \bar{t}_H}}{\frac{\partial Y_H}{\partial \bar{t}_H} + \frac{\partial X}{\partial \bar{t}_H}} + \frac{[Y_H\frac{\gamma_H}{K_H} - X\beta_X - X(1/\frac{\partial X}{\partial \bar{t}_H} + \frac{\gamma_H}{K_H}\frac{\partial Y_H}{\partial \bar{t}_H}/\frac{\partial X}{\partial \bar{t}_H})]\frac{\partial X}{\partial \bar{t}_H}}{\frac{\partial Y_H}{\partial \bar{t}_H} + \frac{\partial X}{\partial \bar{t}_H}}$$

Differentiating the above expression with respect to t_G, we obtain:

$$\frac{\partial \bar{t}_H^*}{\partial t_G} = \frac{(\frac{\partial Y_H}{\partial t_G} + \frac{\partial X}{\partial t_G})\frac{\gamma_H}{K_H}\frac{\partial Y_H}{\partial \bar{t}_H}}{\frac{\partial Y_H}{\partial \bar{t}_H} + \frac{\partial X}{\partial \bar{t}_H}} + \frac{[\frac{\partial Y_H}{\partial t_G}\frac{\gamma_H}{K_H} - \frac{\partial X}{\partial t_G}\beta_X - \frac{\partial X}{\partial t_G}(1/\frac{\partial X}{\partial \bar{t}_H} + \frac{\gamma_H}{K_H}\frac{\partial Y_H}{\partial \bar{t}_H}/\frac{\partial X}{\partial \bar{t}_H})]\frac{\partial X}{\partial \bar{t}_H}}{\frac{\partial Y_H}{\partial \bar{t}_H} + \frac{\partial X}{\partial \bar{t}_H}}$$

For $\gamma_X > \gamma_H$, the first term is negative by Lemmas 3.1 and 3.3.

Next, we show the second term is negative:

1) If $\gamma_X - \gamma_H < \beta_H K_H$,

$$\frac{\partial Y_H}{\partial t_G}\frac{\gamma_H}{K_H} - \frac{\partial X}{\partial t_G}\beta_X - \frac{\partial X}{\partial t_G}[1/\frac{\partial X}{\partial \bar{t}_H} + \frac{\gamma_H}{K_H}(\frac{\partial z_H}{\partial \bar{t}_H} + \frac{\partial z_H}{\partial X}\frac{\partial X}{\partial \bar{t}_H})/\frac{\partial X}{\partial \bar{t}_H}]$$

$$= -\frac{\partial X}{\partial t_G}\beta_X - (\frac{\partial X}{\partial t_G} + \frac{\partial Y_H}{\partial t_G})/\frac{\partial X}{\partial \bar{t}_H} = -\frac{\partial X}{\partial t_G}[\beta_X + (1+\frac{\partial z_H}{\partial X})/\frac{\partial X}{\partial \bar{t}_H}]$$

$$= -\frac{\partial X}{\partial t_G}[\beta_X - [\Delta\frac{N+1}{N}/(\beta_G + \frac{\gamma_G}{K_G})]\frac{1+\partial z_H/\partial X}{1+\frac{\gamma_X}{K_H}\frac{\partial z_H}{\partial \bar{t}_H}}] < -\frac{\partial X}{\partial t_G}[\beta_X - [\Delta\frac{N+1}{N}/(\beta_G + \frac{\gamma_G}{K_G})]\frac{1+\partial z_H/\partial X}{1+\frac{\gamma_H}{K_H}\frac{\partial z_H}{\partial \bar{t}_H}}]$$

$$< -\frac{\partial X}{\partial t_G}[\beta_X - \Delta\frac{N+1}{N}/(\beta_G + \frac{\gamma_G}{K_G})] < -\frac{\partial X}{\partial t_G}[\beta_X - \frac{N+1}{N}[\beta_X + \frac{\gamma_X}{K_H}(1+\frac{\partial z_H}{\partial X}) + \beta_G\frac{\gamma_G}{K_G}/(\beta_G + \frac{\gamma_G}{K_G})]$$

$$< -\frac{\partial X}{\partial t_G}\{-\frac{1}{N}\beta_X - \frac{N+1}{N}[\frac{\gamma_X}{K_H}(1+\frac{\partial z_H}{\partial X}) + \beta_G\frac{\gamma_G}{K_G}/(\beta_G + \frac{\gamma_G}{K_G}) < 0.$$

This, together with Lemma 3.3, implies the second term is negative.

2) If $\gamma_X - \gamma_H \geq \beta_H K_H$,

$$\frac{\partial Y_H}{\partial t_G}\frac{\gamma_H}{K_H} - \frac{\partial X}{\partial t_G}\beta_X - \frac{\partial X}{\partial t_G}[1/\frac{\partial X}{\partial \bar{t}_H} + \frac{\gamma_H}{K_H}(\frac{\partial z_H}{\partial \bar{t}_H} + \frac{\partial z_H}{\partial X}\frac{\partial X}{\partial \bar{t}_H})/\frac{\partial X}{\partial \bar{t}_H}]$$

$$= \frac{\partial Y_H}{\partial t_G}\frac{\gamma_H}{K_H} - \frac{\partial X}{\partial t_G}\beta_X - \frac{\partial X}{\partial t_G}/\frac{\partial X}{\partial \bar{t}_H} + \frac{\partial z_H}{\partial X}\frac{\partial X}{\partial t_G}\frac{\gamma_H}{K_H} - \frac{\partial X}{\partial t_G}\frac{\gamma_H}{K_H}\frac{\partial z_H}{\partial \bar{t}_H}/\frac{\partial X}{\partial \bar{t}_H}$$

$$= -\frac{\partial X}{\partial t_G}\beta_X - \frac{\partial(Y_H + X)}{\partial t_G}/\frac{\partial X}{\partial \bar{t}_H} > 0.$$

This, together with Lemma 3.3, implies the second term is negative. *Q.E.D.*

Proof of Proposition 3.3: Consider the price discrimination case. Totally differentiating (3.19) with respect to t_G yields:

$$\frac{dSW_H^{PD}(t_G, t_H^*(t_G), \tau_H^*(t_G))}{dt_G} = \frac{\partial SW_H^{PD}}{\partial t_G} + \frac{\partial SW_H^{PD}}{\partial t_H}\frac{\partial t_H^*}{\partial t_G} + \frac{\partial SW_H^{PD}}{\partial \tau_H}\frac{\partial \tau_H^*}{\partial t_G} = \frac{\partial SW_H^{PD}}{\partial t_G}$$

where the second equality follows by the first-order conditions of (3.19). Further,

$$\frac{\partial SW_H^{PD}}{\partial t_G} = -Y_H \frac{\gamma_H}{K_H}(\frac{\partial Y_H}{\partial t_G} + \frac{\partial X}{\partial t_G}) + \tau_H^* \frac{\partial Y_H}{\partial t_G} + (X\beta_X + t_H^*)\frac{\partial X}{\partial t_G}$$

$$= -Y_H \frac{\gamma_H}{K_H}(1 + \frac{\partial z_H}{\partial X})\frac{\partial X}{\partial t_G} + \tau_H^* \frac{\partial z_H}{\partial X}\frac{\partial X}{\partial t_G} + (X\beta_X + t_H^*)\frac{\partial X}{\partial t_G}$$

$$= [-Y_H \frac{\gamma_H}{K_H}(1 + \frac{\partial z_H}{\partial X}) + (Y_H \frac{\gamma_H}{K_H} + X \frac{\gamma_X}{K_H})\frac{\partial z_H}{\partial X}]\frac{\partial X}{\partial t_G} + [Y_H \frac{\gamma_H}{K_H} - X(1/\frac{\partial X}{\partial t_H} + \frac{\gamma_X}{K_H}\frac{\partial z_H}{\partial X})]\frac{\partial X}{\partial t_G}$$

$$= -X \frac{\partial X}{\partial t_G} / \frac{\partial X}{\partial t_H} < 0.$$

The proof for the uniform-pricing case is similar. Applying these inequalities and Proposition 3.1, then results in Proposition 3.3. *Q.E.D.*

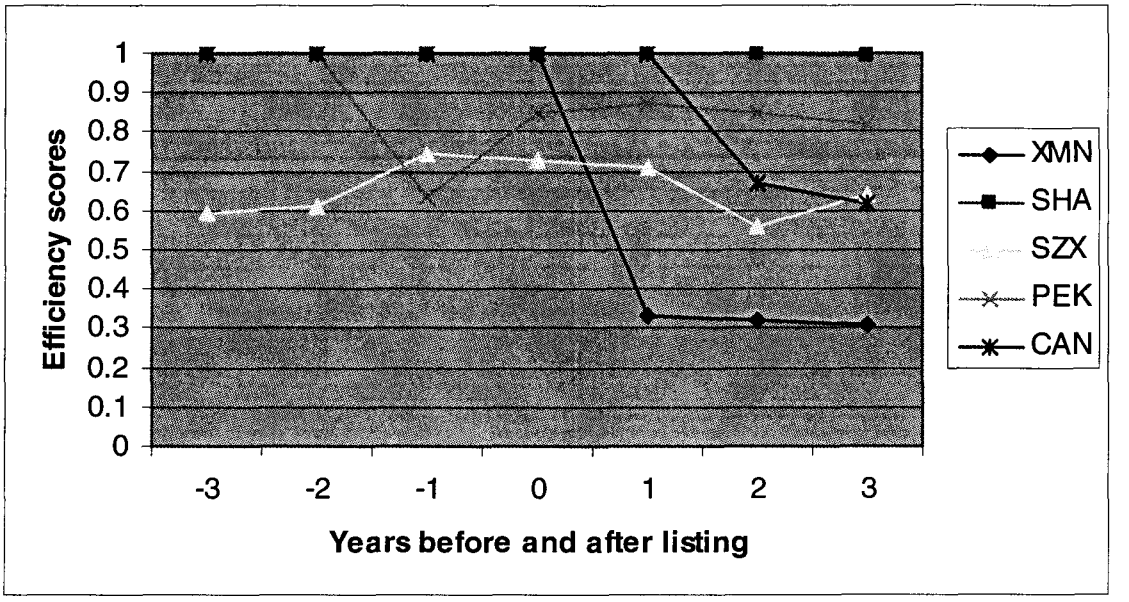
Appendix C Endogenity issues in listing dummy (Chapter 4)

The endogenity issue in our regression models may arise due to the listing mechanism in China. Before the revision of the Securities Law in China in October 2005, Chinese companies were required to be profitable in three consecutive years prior to initial public offering. According to China's National Audit Office, nine of the twelve major airports and thirty seven of the thirty eight secondary airports in China were not profitable in 2001 (Li, 2002). This suggests that only several Chinese airports were able to fulfill the listing requirement. If so, as also argued in Zhang and Yuen (2008), the listing requirement might serve as a device to screen out under-performing Chinese airports for listing. As a result, if high productivity (i.e., the dependent variable in our regression models) will lead to high profits, then an endogenity problem may exist in our regression models. Yet in China, high efficiency of state owned enterprises (SOE), such as airports, may not necessarily imply high profitability (see, e.g., Zhang, et al., 2002), thus whether the endogenity exists is uncertain.

Note that studies in the literature examining the effects of listing or privatization on the profitability and efficiency of SOEs may be potentially subject to similar endogenity problems. The general approach in the literature (e.g., Megginson, et al., 1994; Dewenter and Malatesta, 2001) to deal with the problem is to compare the firms' performance before and after listing, rather than comparing listed or privatized firms with their counterparts. Figure AC.1 shows the efficiency for each sample listed Chinese airport, both for the three years before its listing and the three years after its listing (if the data are available). It seems that there is a drop in efficiency after listing. However, we are unable to draw any conclusion about the relationship between listing and efficiency here, as we do not control for other changes before and after listing in the comparison. Given the fact that only six Chinese airports are currently listed on stock markets, this rather limited data points preclude a more rigorous before/after study as in the literature.

[Figure AC.1 here]

Figure AC.1 Efficiency change before and after public listing



Notes: XMN: Xiamen Gaoqi International Airport; SHA: Shanghai Hongqiao International Airport; SZX: Shenzhen Baoan International Airport; PEK: Beijing Capital International Airport; CAN: Guangzhou Baiyun International Airport.

References

Dewenter, K.L. and Malatesta, P.H. (2001) "State-Owned and Privately Owned Firms: An Empirical Analysis of Profitability, Leverage, and Labor Intensity", *American Economic Review*, 91(1), 320-334.

Li, Jinhua (2002) The Report Submitted to the National Standing Committee of the National People's Congress by Li Jinhua, Auditor General of the National Audit Office of China, 2002, Beijing.

Megginson, W.L., Nash, R.C. and Van Randenborgh, M. (1994) "The Financial and Operating Performance of Newly Privatized Firms: An International Empirical Analysis", *Journal of Finance*, 49(2), 403-452.

Zhang, A. and Yuen, A. (2008), "Airport Policy and Performance in Mainland China and Hong Kong", in: Winston, C. and de Rus, G. (eds.), *Aviation Infrastructure Performance: A Study in Comparative Political Economy*, pp. 159-192, Washington, D.C.: Brookings Institution Press.

Zhang, A., Zhang, Y. and Zhao, R. (2002) "Profitability and Productivity of Chinese Industrial Firms: Measurement and Ownership Implications," *China Economic Review*, 13, 65-88.

Appendix D Proofs in Chapter 5

Derivation of GHG pricing for the price discrimination case

To solve the welfare maximization problem in (5.13), we first derive consumer surplus CS:

$$CS = \int_0^{|z_l|}[V - p_H(q_H, q_F) - 4\tau z]dz + \int_0^{\tilde{z}}[V - p_H(q_H, q_F) - 4\tau z]dz$$
$$+ \int_0^{1-\tilde{z}}[V - p_F(q_H, q_F) - 4\tau z]dz + \int_0^{z_F}[V - p_F(q_H, q_F) - 4\tau z]dz$$

Using (5.4), (5.5) and (5.6), we can solve the integrals to obtain:

$$CS = \tau(q_F^2 + 6q_F q_H + q_H^2 - 4)/2. \tag{A.5.1}$$

Substitute the inverse demand system (5.6) into airline H's profit function, we have:

$$\pi_H = (2\tau + V - c)q_H - (t_H^S + t_H^U + t_F^S)q_H - 3\tau q_H^2 - \tau q_H q_F. \tag{A.5.2}$$

With (A.5.1) and (A.5.2), the welfare maximization problem (13) can be written as:

$$\begin{aligned} SW_H &= \alpha[\tau(q_F^2 + 6q_F q_H + q_H^2 - 4)/2] + \beta[(2\tau + V - c)q_H \\ &\quad - t_F^S q_H - 3\tau q_H^2 - \tau q_H q_F + t_H^S q_F - \alpha[e_H q_H + \delta e_K(q_H + q_F)] \end{aligned} \tag{A.5.3}$$

The first-order conditions for the maximization problem with respect to t_H^S and t_H^U are,

$$\frac{dSW_H}{dt_H^S} = \frac{\partial SW_H}{\partial q_H}\frac{\partial q_H}{\partial t_H^S} + \frac{\partial SW_H}{\partial q_F}\frac{\partial q_F}{\partial t_H^S} + \frac{\partial SW_H}{\partial t_H^S} = 0, \tag{A.5.4}$$

$$\frac{dSW_H}{dt_H^U} = \frac{\partial SW_H}{\partial q_H}\frac{\partial q_H}{\partial t_H^U} + \frac{\partial SW_H}{\partial q_F}\frac{\partial q_F}{\partial t_H^U} + \frac{\partial SW_H}{\partial t_H^U} = 0. \tag{A.5.5}$$

and assuming the second-order conditions are satisfied.

By solving conditions (A.5.4) and (A.5.5),

$$t_H^U = \{\frac{5[14-\beta+\tau(\beta-4\alpha)]}{\Omega_1}e_H + \frac{30(2-\tau)\beta\delta}{\Omega_1}e_K\}\gamma + K_1,$$

$$t_H^S = \{\frac{6\beta+17\alpha\tau-6\beta\tau}{\Omega_1}e_H + \frac{[70-65\beta+(29\beta+14\alpha)\tau]\delta}{\Omega_1}e_K\}\gamma + K_2.$$

where $\Omega_1 = -(\tau-1)^2\beta^2 + 2(34\tau+3\alpha\tau-70)\beta - 4(2\alpha\tau+3)\alpha\tau + 140 > 0$ by the second-order conditions; and

$$\begin{aligned}
K_1 = \frac{1}{\Omega_1}\{&[(\tau-1)^2\beta^2 + [69-3\tau(2\alpha+11)]\beta + \alpha\tau(8\alpha\tau+9) - 70]t_F^U \\
&+[(\tau-1)^2\beta^2 - [75+3\tau(2\alpha-13)]\beta + \alpha\tau(8\alpha\tau+26) - 70]t_F^S \\
&-[2\tau^3 + (V-c-4)\tau^2 + 2(c-V+1)\tau + V - c]\beta^2 \\
&-[-2(6\alpha+39)\tau^2 + 3[(2\alpha+13)(c-V)+50]\tau + 75(V-c)]\beta \\
&-16\alpha^2\tau^3 - 4\alpha\tau^2[13+2\alpha(V-c)] + (70-26\alpha\tau)(V-c) + 140\tau\}
\end{aligned}$$

$$\begin{aligned}
K_2 = \frac{1}{\Omega_1}\{&-[(\tau-1)^2\beta^2 - [75-3\tau(2\alpha+11)]\beta + 8\alpha\tau(\alpha\tau-1) - 70]t_F^U - 30\beta(\tau-2)t_F^S \\
&-30\beta[2(V-c) - (4-V+c)\tau + 2\tau^2]\}
\end{aligned}$$

Derivation of GHG pricing for the uniform pricing case

The first-order condition for the maximization problem with respect to $\bar{t}_H$ is,

$$\frac{dSW_H}{d\bar{t}_H} = \frac{\partial SW_H}{\partial q_H}\frac{\partial q_H}{\partial \bar{t}_H} + \frac{\partial SW_H}{\partial q_F}\frac{\partial q_F}{\partial \bar{t}_H} + \frac{\partial SW_H}{\partial \bar{t}_H} = 0, \tag{A.5.6}$$

and assuming the second-order condition is satisfied.

By solving conditions (A.5.6),

$$\bar{t}_H = \frac{35}{\Omega_2}(11e_H + 15\delta e_K)\gamma + K_5,$$

where $\Omega_2 = -401\alpha\tau + 814\beta\tau - 1540\beta + 1820 > 0$ by the second-order condition; and

$$K_5 = \frac{1}{\Omega_2}\{[(655 - 401\tau)\beta + 499\alpha\tau - 655]\bar{t}_F + [(810\tau - 1470)\tau + (405\tau - 735)(V - c)]\beta \\ - 25[24\alpha\tau^2 + 6[2\alpha(V - c) - 7]\tau - 21(V - c)]\}.$$