Airport Capacity Allocation with Network Airlines - Regulation of Congestion Externalities under Imperfect Competition, with Vertical Product Differentiation based on Network Density Effects

D I S S E R T A T I O N
of the University of St. Gallen,
School of Management,
Economics, Law, Social Sciences
and International Affairs
to obtain the title of
Doctor of Philosophy in Management

submitted by

Claudio Giovanni Noto

from

Ganterschwil (St. Gallen) and St. Gallen

Approved on the application of

Prof. Dr. Christian Laesser

and

Prof. Dr. Kuno Schedler

Dissertation no. 4483

epubli Berlin, 2016

Airport Capacity Allocation with Network Airlines - Regulation of Congestion Externalities under Imperfect Competition, with Vertical Product Differentiation based on Network Density Effects

D I S S E R T A T I O N
of the University of St. Gallen,
School of Management,
Economics, Law, Social Sciences
and International Affairs
to obtain the title of
Doctor of Philosophy in Management

submitted by

Claudio Giovanni Noto

from

Ganterschwil (St. Gallen) and St. Gallen

Approved on the application of

Prof. Dr. Christian Laesser

and

Prof. Dr. Kuno Schedler

Dissertation no. 4483

epubli Berlin, 2016

The University of St. Gallen, School of Management, Economics, Law, Social Sciences and International Affairs hereby consents to the printing of the present dissertation, without hereby expressing any opinion on the views herein expressed.

St. Gallen, November 2, 2015

The President:

Prof. Dr. Thomas Bieger

Acknowledgements

First, I would like to express my sincere gratitude to my supervisor, Prof. Dr. Christian Laesser, for his invaluable inputs and his unlimited confidence in this endeavor, and both my co-supervisors, Prof. Dr. Kuno Schedler and Prof. Dr. Rico Maggi, for providing me with their economic expertise. Moreover, I am indebted to Prof. Dr. Armin Schmutzler for his critical review of the model design and, in particular, to Achim Czerny for his encouragement and his thorough feedbacks at my two GARS Workshops.

Furthermore, I would like to extend very warm and special thanks to Peter Forsyth for showing interest in my work; his opinion on a late draft of this thesis and his general thoughts on the subject have been tremendously contributing to my progress. In addition, a large piece of cake is reserved for Barbara Heck for her accurate and expeditious professional editing, thereby boosting both my pace and morale during final approach.

Also, I owe deep appreciation to Hedi for coaching the boarding process of my train of thoughts back then and for something about something ever since, my brother Flavio for motivating this goal and for urgent proofreading missions, Johannes Binswanger for his constructive comments at an early stage, Tino and Florian for contributing their friendship, Peppi, Lupo and the fellows of Fighter Sqn 19 for keeping up the spirit, all family members for their understanding, my friends for staying in touch, and, finally, Binz for the seminal insight that its ultimately just for the drop-down menu. Fly like hell.

I dedicate this work to my dear parents for always maintaining their unconditional trust in each and every one of my undertakings, and to Co for her patience and her passionate support for all my ideas, even if they occasionally interfere with our common plans.

St. Gallen, December 2015 - cn.

Executive Summary

Excess demand for runway capacity at major airports is still causing congestion externalities and hence significant delay costs for airlines and passengers. Therefore, recent studies criticize the current administrative allocation of airport slots as both inefficient and inequitable, and propose alternatives such as a market allocation of slots or a congestion pricing solution. In perfect competition settings, theoretical models show allocation efficiency for both instruments; however, when market power prevails over the congestion externality, also adverse welfare effects may arise.

The literature suggests that modern airline competition involves network structures as a dominant strategy, aiming at achieving competitive advantages based on product differentiation in order to increase market power. The corresponding network effects supposedly induce two opposing effects for passengers: additional indirect travel benefits from network density and higher mark-ups on flight fares, thus constituting the dilemma of hub concentration. However, most airport capacity allocation models consider the flights of competing airlines as perfect substitutes, so that the flight fares are determined by total industry output and do not account for product differentiation based on asymmetric network structures.

Consequently, this study first provides a modified theoretical model to investigate the network hub of a dominant airline, featuring an asymmetric oligopoly with vertical product differentiation based on network density effects. Subsequently, a partial equilibrium analysis qualitatively evaluates the inefficiencies arising in the unregulated market and the welfare impact of an airport quota allocation, a secondary trading scheme, and a congestion pricing solution. In addition, a computer simulation illustrates the particular model characteristics and quantifies the inefficiencies by numeric results.

The setting with congestion externalities, market power and network density benefits reveals that all three instruments may improve allocation efficiency but that the potential aggravation of the output inefficiency and the network undersize also contains a welfare caveat. The network undersize arises because the airline faces concave network returns, while social welfare monotonously increases with the network size, so that the network benefits accentuate the market power distortion. The only first-best capacity allocation arises from individual quotas in conjunction with a use obligation, which in effect closely resemble the current administrative quota allocation in practice. This result challenges the recent criticism and the alternative propositions from the literature. Nonetheless, further research might investigate whether different regulation policies from other network industries might prove even more suitable in a hub-airport context.

Zusammenfassung

Nachfragebedingte Verkehrsprobleme an grossen Flughäfen verursachen aufgrund von Externalitäten nach wie vor ungerechtfertigt hohe Verspätungskosten für Passagiere und Fluggesellschaften. Die administrative Allokation von Flughafenzugangsrechten wird daher in vielen Studien als ineffizient und auch ungerecht kritisiert. Theoretische Modelle zeigen, dass sowohl eine freie Marktallokation als auch eine Spitzenbelastungssteuer in kompetitiven Märkten gleichsam effizient wären; falls allerdings Marktmacht vorhanden ist, treten auch negative Wohlfahrtseffekte auf.

Im Wettbewerb zwischen Fluggesellschaften gelten Netzwerkstrukturen als dominante Strategien, welche mittels Produktdifferenzierung komparative Vorteile erschaffen können. Aus systemischer Sicht begründen sie allerdings das Dilemma der Marktkonzentration, da die Netzwerkeffekte einerseits zusätzliche Vorteile für die Passagiere darstellen, aber auch die Marktmacht und damit die Flugpreise erhöhen. Die meisten Modelle zur Allokation von Flughafenkapazitäten betrachten aber die Flüge konkurrierender Fluggesellschaften als perfekte Substitute, und bilden daher keine Produktheterogenitäten und Asymmetrien zwischen den Firmen ab.

Diese Arbeit modifiziert daher ein theoretisches Flughafenmodell, um eine Wettbewerbssituation mit vertikaler Produktdifferenzierung darzustellen, welche auf asymmetrischen Netzwerkeffekten beruht. Eine Teilgleichgewichtsanalyse analysiert dann qualitativ die Ineffizienzen im unregulierten Markt und die daraus folgenden Wohlfahrtseffekte sowohl einer administrativen und einer marktbasierten Allokation von Zugangsrechten, als auch einer Lenkungsabgabe. Zudem veranschaulicht eine Computersimulation quantitativ die Modellspezifika, Ineffizienzen und Regulierungseffekte mittels numerischer Gleichgewichtswerte und Sensitivitätsanalysen.

Im vorliegenden Modell können zwar alle drei Instrumente die ökonomische Allokationseffizienz im Gesamtsystem verbessern. Da aber die Netzwerkprämie konkav zur Netzwerkgrösse verläuft, während die Gesamtwohlfahrt monoton ansteigt, bleibt das Netzwerk aus sozialer Sicht unterentwickelt. Die Netzwerkeffekte verstärken daher den negativen Einfluss der Marktmachtverzerrung, sodass in allen drei Fällen auch Wohlfahrtsminderungen auftreten können. Die einzige theoretisch effiziente Lösung stellen individuell bemessene Zugangsrechte in Verbindung mit einer Nutzungsverpflichtung dar. Da diese im Endeffekt stark der gegenwärtigen Slotallokation aus der Praxis gleichen, relativiert dieses Resultat die Kritik aus der Literatur. Zukünftige Arbeiten könnten dennoch untersuchen, ob sich auch Regulierungsschemata aus anderen Netzwerkindustrien auf Flughäfen übertragen liessen.

Contents

List of Tables

List of Figures

Introduction, Background and Literature

"Economic regulation introduces its own distortions, and at the end of the day there is a trade-off to be made between imperfect competition and imperfect regulation." (Starkie, 2008b, p.135)

1 Introduction

At major airports, excess demand for runway capacity has been causing congestion, leading to flight delays and to significant costs for both airlines and passengers (e.g. Cook, 2007b, p.97). Because congestion partly represents an externality, regulation has been introduced to efficiently allocate scarce airport capacity. Recent literature, however, has criticized the current administrative capacity allocation scheme as being both inefficient and inequitable (cf.,e.g., Matthews and Menaz, 2008, p.24).

Subsequently, a considerable number of studies have proposed and investigated alternative instruments to this classical economic allocation problem. As predicted by economic theory, their optimality has been shown under perfect competition. This also applies to imperfect competition settings that assume perfectly elastic demand, as indicated in Section 4.2. In the presence of market power, where prices are endogenous functions of demand, however, the airport capacity allocation does not only affect the congestion externality but also the market power distortion that arises in this market. Consequently, the impact of capacity allocation on market power induces second-order effects on efficiency that may offset or even overcompensate the positive welfare impact of regulation.

In this context, the dual market distortion by market power and the congestion externality is known as the dual distortion. Based on theoretical equilibrium models featuring imperfect competition with finite demand elasticity, a few studies have shown the potentially adverse welfare effects from allocation instruments in conjunction with this dual distortion. This indicates that both current as well as proposed alternative allocation schemes may not succeed in replicating the socially optimal market structure.

Economic theory describes the airline industry structure as one of product heterogeneities, where competition among airlines is based on differentiation strategies that allow successful players to implement competitive advantages (see e.g. Gillen and Morrison, 2008, pp.178

and O'Connell, 2006, p.54). These differentiation strategies translate into higher markups which, in turn, permit these airlines to outperform their competitors (Holloway, 2002, p.23-24). In this respect, it is widely accepted that modern large airlines have adopted hub-and-spoke network structures, which allow them to differentiate their product against both other networking and non-networking competitors (O'Connell 2006, p.60, Zhang et al. 2011, p.803 or Oum et al. 2012, p.432). As a result, the market structure at current large network hubs should justifiably be characterized as an asymmetric oligopoly rather than as a perfect competition setting or a symmetric oligopoly with flights as homogenous goods.

In a network, product quality is driven by network density, which allows the connectivity within the network to increase. The connectivity may manifest in, e.g., higher flight frequencies or a wider destination choice (cf. e.g. Joppien, 2003; Gödeking, 2010). This higher product quality in terms of connectivity thus is likely to induce additional passenger benefits and thus to improve an airline's profitability (see Laffont et al., 1998, p.4-5 in Section 3.1.3). One might hence argue that network airlines take advantage of asymmetric network density effects in order to compete against other non-networking airlines based on vertical differentiation by product quality. As economic theory suggests, in turn, product heterogeneity increases market power. However, the higher product quality associated with the corresponding indirect network benefits and advantages also increases the travel benefits for the networking airlines' passengers. As a consequence, customers might be willing to afford the corresponding hub premium arising from these indirect travel benefits. In this case, these network services would constitute a net benefit for the customers, despite the higher market power that is potentially associated with them (cf. e.g. Starkie 1998, p.114 and 2008b, pp.171). This would allow the networking airline's market share to increase along with market power while at the same time increasing allocation efficiency. This controversial implication has been denoted by Starkie (2008a, p.194) as the "dilemma of hub concentration."

The assessment of any allocation instrument concerned with congestion externalities at major airports can hence only be performed if the natural market structure arising from the above asymmetric density benefits is known. Consequently, Langner (1996, pp.15) urged the need to reflect network effects in the discussion about airport capacity allocation. Also Brueckner (2002a) suggests that airline asymmetries other than cost-side differences have to be accounted for in future allocation efficiency assessments. Moreover, Aguirregabiria and Ho (2010, p.1) note that the effect of network effects has thus far been neglected in airline competition models. The above dilemma indicates that the presence of network effects opposing the traditional market power distortion substantially complicates the corresponding welfare analysis (Berry, 1990, p.394).

However, most recent studies apply theoretical models where flights are homogenous products. This means that there is only one market price, which is determined by total industry output and is identical across all firms (Vives, 2001, p.94). However, as Berry (1990, p.394) mentions, this kind of "traditional market power" only yields profits "by restricting output and driving up prices", while it would be the "socially ambiguous nature of much of product differentiation in this industry" which would make welfare analysis "particularly difficult". In-line with the above notion about the market structure of the airline industry, he concludes that "both simple cost-reducing and naive market power stories are inappropriate for the airline industry" (idem). It may thus be deemed questionable whether the current market structure of major airports has adequately been reflected in the recent airport capacity allocation models in the literature.

Moreover, in most recent studies the formalizations of either airline asymmetries or network effects that may arise in a network-hub market structure have rarely been provided. More generally, this also concerns demand-side heterogeneities that allow the airlines to endogenously differentiate their prices, although these heterogeneities have already been identified by Berry (1990, p.398) as representing a different kind of market power than the traditional output inefficiency and hence constituting a major driver of the airlines' airport demand. The above reasoning hence suggests that the question about efficient airport capacity allocation at major network hubs of dominant airlines remains as of yet unanswered.

Therefore, this study investigates airport capacity allocation at a stylized network-hub airport that reflects an airline duopoly with a networking airline and a non-networking competitor. For this purpose, it first presents a theoretical partial equilibrium model based on Brueckner's (2002a) symmetric airport model. While that model considers flights as homogenous goods and does not include network effects, this model has been modified to account for vertical product differentiation based on network density benefits. The duopoly consists of a non-networking airline that is supposed to provide direct flight benefits from point-to-point transportation only and a networking airline that is supposed to offer additional indirect benefits from network density to its customers. These benefits presumably arise from its optimized network structure at the network hub. As a result, the networking airline can endogenously differentiate its product from the leisure airline, which yields both a higher market share as well as a higher profitability. This ultimately leads to a demand-side asymmetry in imperfect airline competition, which enables "product differentiation to affect both costs and demand" (Berry, 1990, p.394). Subsequently, the unconstrained market equilibrium reveals the natural market structure at this network hub and its inefficiencies as compared to the theoretical social optimum. In response to these inefficiencies, this investigation considers the welfare effect of the three most prominent capacity allocation instruments: a quota solution, a secondary quota trading scheme, and congestion pricing. The first of the three instruments has already been established in practice, whereas the latter two are known from economic theory and are suggested by recent contributions in the literature.

In light of the inefficiencies in an unregulated market, the analysis firstly assesses an allocation of individual quotas that are tailored according to the efficient market shares of either airline. In addition, it evaluates the current administrative airport quota scheme. These quotas are known in practice as the airport slots. For modeling purposes, their allocation is replicated by a symmetric arbitrary constraint that concerns both airlines. Secondly, the analysis assesses both the trading potential and the welfare effects of a secondary trading market for these airport slots. This concerns both the individual quotas and the arbitrary constraints. Moreover, strategic airline behavior is introduced in two distinct forms: Either airline may hold on to its unused number of quotas instead of selling them, or it may buy some slots that have previously been utilized by its competitor in order to hoard them unused. The former type of strategic behavior has been discussed in the literature as the babysitting of slots, while the latter type will be referred to as strategic slot hoarding. The opportunities for strategic airline behavior thus may or may not arise, depending on the application of a set of trading rules. These trading rules are hence shown to have a crucial impact on the welfare effect of secondary trading. Lastly, the analysis considers the particularities that arise from a congestion pricing scheme in the presence of network effects.

On the one hand, the results from the theoretical model partly replicate the theoretical results from recent studies. On the other hand, they yield some novel insights considering the natural market structure based on the asymmetric network effects across both airlines. Most importantly, the asymmetric network effects dictate that the social optimum simultaneously involves an optimal individual output for either airline. As a central result, this study shows that neither of the alternative allocation schemes under investigation replicates the social optimum. At the same time, they all contain a welfare caveat based on the overall output inefficiency and the underprovision of network services in equilibrium. These insights lead to important considerations concerning the above instruments, which in certain aspects differ from the previously established results.

As the results also show, the model generally yields conditions under which second-best efficiency is reached or welfare caveats arise rather than concise results. This owes to the fact that in analogy to Brueckner (2002a), the model analysis relies on generic functions. In order to illustrate the resulting ambiguities, the study, in addition, provides a numeric simulation that is based on a specification and parametrization of the generic model. The specified functions draw on common notions from the literature and the basic functions are linear. The subsequent numeric parameters are arbitrarily chosen, yet they aim at providing a stylized representation of the generic model. This yields numeric solutions for all endogenous variables that ultimately enable the study to quantitatively evaluate the different values of passenger utility, airline profits and welfare both in equilibrium as well as in the social optimum. The illustrative value of these results is further enhanced by means of a sensitivity analysis for different parameter sets. In this respect, the simulation provides additional findings that could not have been inferred from the formal analysis of the generic model. However, the parametric specification and the parameter choice limit the generalizability of these results.

From an instrumental perspective, this study thus provides an important contribution to the airport capacity allocation discussion: It introduces an airline asymmetry based on network density effects into an airport model so that flights become imperfect substitutes. Their degree of differentiation is endogenous in the density of the business airline's network. The benefits that arise from this network density thus represent vertical product differentiation based on product quality. In practice, these indirect network benefits are supposedly based on, e.g., flight frequency, connectivity and schedule delays within the network. They hence reflect additional utility from travel options and flexibility for the customers. In this model, for simplicity the network density is directly approximated by the business airline's peak-period flight volume. The indirect utility arising from these benefits is implemented according to Belleflame and Peitz's (2010) general foundations for utility from network goods.

Subsequently, the model formalizes both a theoretically optimal quota scheme as well as the administrative grandfathering quota scheme from practice. In addition, it implements the two most prominent alternative allocation instruments proposed in the recent literature: a congestion pricing scheme and a secondary trading option for the administratively allocated airport quotas. These two alternative allocation instruments have already been investigated in the recent airport capacity models both under perfect competition and with traditional market power. However, as described above, the theoretical models applied in these investigations neither include the network aspect of supply nor the implications of heterogenous demand arising from product differentiation. Hence, they reflect neither demand-side asymmetries nor asymmetric network effects as applied in this study. This also concerns the administrative quota scheme from practice. In addition, an individual quota scheme has rarely been investigated in this context. However, in view of airline asymmetries, individual quotas seem more appropriate than an anonymous symmetric quota scheme because they account for the efficient asymmetric market shares instead of just evenly reducing supply across the competitors. This reasoning is underlined by the fact that in an asymmetric model with market power, a symmetric grandfathering quota allocation that constrains both airlines cannot be endogenously justified based on efficiency.

As already mentioned, the results obtained in the analysis challenge some of the results from previous work in the literature. Most importantly, the analysis shows that the current market structure with large dominant airlines plausibly points toward inefficiencies with both a market-based capacity allocation as well as a tax-based scheme. This lets us suppose that the welfare results arising from perfect competition settings and from traditional market power models are not directly applicable in a network hub environment. Consequently, these results suggest the need to find and investigate suitable adaptions of regulation from other industry sectors concerned with dominant, asymmetric networks. If such regulation enabled the internalization of externalities and at the same time compensated for the market power distortion, it might dismantle the natural but inefficient association of market dominance and network density benefits while letting the customer value flourish.

There are, however, two important limitations to the analysis of the generic model: First, the general output inefficiency, which significantly affects allocation efficiency and supports the welfare caveat of the allocation instruments, is based on the assumption of downward sloping demand arising from traditional market power. Although the airline market may be deemed as (at least partly) incontestable, some recent capacity allocation studies prefer the application of market power with inelastic demand. Consequently, they find neither the same inefficiencies nor the same welfare caveats as this model with endogenous demand. Second, the airline asymmetry is based on the notion that network structures and services arise from incumbency advantages of major airlines. The airline asymmetry is therefore completely exogenous to the model. As a consequence, the competing airline cannot provide indirect utility by definition, although in practice new entrant competitors might engage in network configurations as well. This rigid market structure may arguably be criticized as oversimplifying the current industry structure from practice. Nevertheless, the model serves as an illustrative stylized setting to investigate the impact of an asymmetric airline network on allocation efficiency that is motivated by reflections on the prevailing market structure. Consequently, this analysis provides yet another perspective on the characteristics of airport capacity allocation. Its particular results from the inclusion of the network density effects may thus be considered along with those of the preceding studies from literature that feature perfect competition and traditional market power settings.

The remainder of this study is divided into five chapters referred to as the Background and Literature, Parts I to III, and the Conclusion. These chapters are organized as follows: The first chapter presents the background and the corresponding economic theory on the airport capacity allocation problem from practice and provides an overview of the most relevant recent studies on the topic from the literature. Part I first introduces the generic model of the network hub airport. Subsequently, it computes both the equilibrium and the socially optimal output allocation at the airport in order to reveal the inefficiencies that arise in an unregulated market solution. The results are then discussed along with the model assumptions that led to these results. Part II formally implements the three above-mentioned allocation instruments and investigates their impact on allocation efficiency. Part III provides the simulation based on a parametric form of the generic model. This simulation is used to graphically illustrate the model's main properties and to obtain quantitative results concerning the above inefficiencies. The last chapter presents a summary of all relevant results and draws conclusions about the particular findings of this study. In addition, it considers the main limitations of the model with respect to both conceptual and technical aspects and offers perspectives and directions for further research.

The segmentation into Parts I to III is chosen to emphasize the three unique contributions that this study makes to the recent discussion in the literature: the instrumental contribution of the asymmetric airline model, the subsequent investigation of airport capacity allocation in light of network density benefits, and the quantitative illustration of the generic results by means of the parametric model simulation. All tables and figures are own illustrations; the graphs in Part III were generated based on the simulation software. Lastly, note that throughout the whole study the masculine form is used for simplicity; however, this convention always implies the inclusion of the feminine form.

2 Background

This section first defines the airport capacity problem and the scope of this study. Thereafter, it briefly reviews the background of the airport capacity regulation as a classical economic allocation problem. Subsequently, it describes the current allocation scheme from practice and recent criticism from the literature together with an overview of the proposed alternatives. The foundations of those allocation instruments are presented in more detail in Section 3. The main insight from this section is that the air traffic infrastructure may be characterized as a mixed public good that is concerned with congestion externalities. Consequently, the regulation of the airport capacity problem is justified by general welfare considerations.

2.1 Airport Capacity Problem

2.1.1 Definitions and Scope of this Study

An airport is typically separated into two distinct parts: The airside and the landside. The airside represents all movement areas that are used for aircraft flight operations. These areas are also referred to as to the airport facilities. The landside consists of all passenger and ground transport interfaces that are necessary to connect the passengers to the air operations. These two parts are both functionally and spatially separated by the passenger terminals (see Moosecker, 2010, p.8 and Karakus, 2009, p.15).

Together with the so-called en-route airspace, the airport facilities represent the air traffic infrastructure. The en-route airspace represents the spatial dimensions within which the aircraft proceed from airport to airport. They include all arrival and departure routes, the airways and all other kinds of air traffic areas off the ground (Majumdar, 2007, p.67). These two parts of the air traffic infrastructure differ not only in function but also in the regulations associated with them. Therefore, this study only considers the capacity problems concerning the facilities required for aircraft operations at airports but includes neither en-route airspace problems nor land-side constraints.

2.1.2 Airport Congestion as an Externality

The limiting factor within the airport facilities generally is the runway capacity (Bauer, 2008, p.153). The runway capacity describes the number of airplane take-offs and landings that can be performed on a runway within a defined period of time. If the demand for aircraft movements exceeds this runway capacity and demand is not regulated, the airport facilities suffer from overuse, which leads to airport congestion. Airport congestion hence arises from the rival consumption of operational airport capacity as an exhaustible resource.

Airport congestion, in turn, leads to potential and actual flight delays. These flight delays cause congestion costs to the airlines and time costs to the passengers. These delay costs are "accepted to be real, large but poorly understood quantitatively" (Cook, 2007b, p. 97). If the delay costs are not fully accounted for in the decision rationales of all participants at a congested airport, they constitute external effects (cf., e.g., Feess-Dörr, 1995). These externalities from congestion constitute a market failure. As a consequence, the corresponding resource allocation is economically inefficient (Mas-Colell et al., 1995, p. 350).

From the perspective of the International Air Transport Association (IATA), the problem of airport congestion should be countered by an augmentation of airport capacities in the first place. This refers both to expansion plans for airport facilities as well as to the optimization of capacity utilization. These supply-side measures are strictly preferred over the regulation of existing capacities, which should only be performed when "all possibilities of developing the limiting components of airports have been exhausted" (IATA, 2010, p.1).

Both in the literature and practice, however, there has been a broad consensus that capacity expansions sufficient to cover both the current and the expected excess demand was not a realistic outlook for the near future. This concerns at least Europe and the USA as the world's two main traffic regions, where the motivation for large infrastructure projects seems to be severely limited both due to the physical as well as the political presuppositions. In this respect, the latest growth study from Eurocontrol reports that the airport capacity increases in Europe planned as of 2007 would amount to 41% by the year 2030, while the respective air traffic growth in terms of flight volume is forecast to reach a magnitude of 70% to 120% (EUROCONTROL, 2008, p.1).

On the one hand, these severe limitations only leave room for optimization strategies in order to increase capacities on the supply side. Such strategies aim at increasing airport capacities and optimizing air traffic flows given the current amount of airport infrastructure, some of which have already been implemented in practice.[1] On the other hand, the current and expected shortages foster the case for the investigation of the allocation of excess demand to a fixed limited amount of airport capacity. Although distinct capacity allocation schemes have been implemented at major airports worldwide, recent contributions in the literature have severely criticized the corresponding allocation efficiency. This mainly concerns the

[1] See Section 2.5.2, which refers to the systems of collaborative airport decision and demand management as examples.

administrative allocation of airport slots that is applied in many world regions. In line with a number of other current studies, this work is therefore devoted to the problem of airport capacity allocation.

2.2 Typology

Airport capacity may generally be perceived as a public good. However, economic theory defines a pure public good as inexhaustible and non-rival in consumption. Moreover, a public good “may or may not be” exclusive (Mas-Colell et al., 1995, p.359-360). In this respect it contrasts with a private good, which is exclusive by definition. A runway, however, can only be used by one aircraft at a time which makes airport capacity a naturally limited, exhaustible resource. Moreover, the use of airport facilities is exclusive concerning both the physical access properties as well as the permits required and restrictions imposed for airlines to operate.[2] In this respect, airport capacity seems to be more suitably characterized as a private good.

However, Kost (2003, pp.98) states that on legal grounds, the air traffic infrastructure exhibits a large degree of public accessibility as long as the sovereign character of infrastructure provision and of the public laws concerning private ownership conditions prevail. In this respect, the air traffic infrastructure is subject to fundamental rights to be granted by public authorities such as the commercial freedom of action and equal treatment (Kost, 2003, p.110 and p.116). Moreover, rivalry in consumption only occurs when the traffic volume actually exceeds the nominal capacity of the facilities (cf., e.g., Baumol, 1975, p. 27). Below that level, the use of infrastructure remains non-rival. As a result, the air traffic infrastructure may be considered as an exclusive public good. This typology, however, remains unsatisfactory with regard to the potential rivalry in consumption.

The relevance of these subtleties in the classification of goods is lucidly articulated by Buchanan (1965, p.2): “While it is evident that some goods and services may be reasonably classified as purely private, even in the extreme sense, it is clear that few, if any, goods satisfy the conditions of extreme collectiveness.” Accordingly, the cases where consumption involves “some publicness” but only “within a finite range” are exactly the interesting ones (idem). Given the above properties, airport capacity is therefore classified as a mixed public good within the context of this study. The public property rests on the notion of the free market access which is legally provisioned, so that the infrastructure at least theoretically is

[2] In practice, numerous operational and safety regulations need to be fulfilled. Moreover, airlines need to obtain air traffic rights “within a complex web of bilateral air service agreements“ of their sovereign home states (Doganis, 2002, pp.31-34).

accessible by any airline despite its associated regulation.[3] The resource limitation depending on the intensity of utilization is hence characterized as an occasional or potential rivalry in consumption. The mixed property thus refers to this limitation.

Given the above characteristics, the mixed public air traffic infrastructure may be suggested to correspond to a so-called club good, as presented by Buchanan (1965), whose general theory of clubs exactly aims at the specification of goods within the above ambiguity between the private and the public. Buchanan's (1965) club refers to an "ownership-membership arrangement" where users group together to finance private goods or services. Rivalry then occurs through congestion within the club and decreases the valuation of the club good (idem, p.7). In other words, the utility that each individual receives from the consumption of the good or service is functionally related to the number of participants within the club. As opposed to the case of a public good, however, the cost from consumption rivalry is included in the cost-benefit consideration of each user and depends on the club size (i.e., the number of members). The "central question in a theory of clubs" is the determination of the optimal "membership margin," which refers to the "the size of the most desirable cost and consumption sharing arrangement". For any given facility size there will thus be an optimal size of the club (idem, p.2 and p.8).

The club specification hence corresponds to the above typology of airport capacity allocation but for one important distinction: Within the theory of the club, congestion is already internalized and thus affects the ex-ante determination of the optimal club size. In the airport case, however, congestion is partly externalized in the unconstrained equilibrium. This signifies that congestion internalization is not driven by the participants but needs to be imposed by the authorities. If this regulation, however, is characterized as an integral part of the system, the airport capacity becomes a genuine club good in Buchanan's sense. Note that this consideration is interesting in its own right but does not add any further implications to the setup of this study.

2.3 Regulation Policy

As briefly mentioned in Section 2.2, the source of inefficiency within the airport capacity allocation consists in the consumption rivalry: This rivalry leads to airport congestion and thus to potential and actual flight delays. These delays cause costs to the airlines and to the

[3] In the European Union, any airline holding a European air carrier license is generally allowed to operate (Van Reeven, 2005, p.711). Although this only applies if the carrier is able to obtain all permission requirements stipulated by the associated regulation, from a competition policy perspective the market may be characterized as freely accessible.

passengers. If these costs are not accounted for in the decision rationales of the participants, they constitute negative external effects. External effects can hence be defined as exogenous variables in the utility and profit functions of individuals and firms. That signifies that these variables cannot be influenced by the individuals and firms in their output decisions (cf. e.g. Feess-Dörr, 1995).[4] As a consequence, the externalities constitute a market failure and the optimality criteria for allocation efficiency are not met (cf. again Mas-Colell et al., 1995, p. 350). From an economic welfare perspective, these inefficiencies thus account for the need of regulation.

The airport capacity allocation problem constitutes a classical economic allocation problem, well-known from welfare theory (cf. e.g. Baumol, 1965). Its principal aspects can be drawn from environmental economics: According to the opportunity cost principle, an allocation is efficient if it realizes the highest possible social welfare from a constrained resource. For this purpose, a multitude of economic allocation instruments exists. These will be referred to in Section 3.2.1. In general, optimality requires that the marginal utility of all individuals equals the marginal total costs (cf. e.g. Baumol, 1975; Baumol and Oates, 1979; Feess-Dörr, 1995 and Button and Verhoef 1998). As a consequence, the externalities need to be included in the participants' decision rationales. This is referred to as the internalization of external effects. The allocation thus becomes efficient because all market distortions causing losses of economic rent are eliminated. As a result, the sum of consumer rents and firms' profits is maximized (cf., e.g., Baumol, 1965, pp.24). In the airport context, it is important to note that allocation efficiency does not require the congestion costs to completely vanish. Rather, they need to be optimally balanced against the benefits that arise from all airport operations (cf. again Baumol, 1965, pp.24).

As Savage (2006, p.350) notes, the fact that a large share of economic activity was devoted to transportation made it a "major focus or government regulatory activity". Accordingly, Wolf (2004, pp.201) states that airport infrastructure is typically subordinated to national competition laws due to its monopolistic character. This is valid regardless of the increasingly privatized provision of airport services including complete deregulation (idem, pp.205). However, as Baumol (1965, pp. 20 and 29) stresses, government intervention is not justified by the existence of externalities alone. Rather, the imposition of regulation needs to be based on the case where an unregulated market process fails to achieve externality internalization on its own: For example, the presence of institutional or legal barriers that prevent internalization requires external intervention in order to restore allocation efficiency.

Following the definitions by Mas-Colell et al. (1995, p.364), congestion costs can be specified

[4] The original definition of Feess-Dörr (1995) only involves individuals.

as rivalrous, multilateral externalities. Multilaterality signifies that multiple participants are concerned with congestion, where rivality means that the externality could be completely absorbed by some participants and thus that the externality is exhaustible on its own. For illustration purposes, suppose that delays from congestion may be plausibly assumed to concern more than one single flight. As a consequence, they can thus be compensated by a cancellation or delay of flight movements of one single or a few participants only. Additional flights at a constrained airport should thus be allowed until the marginal cost of capacity use equals social marginal cost (Matthews and Menaz, 2008, p.24-25). The above classification is transferred to the airport case based on Baumol's (1975, p. 19) example concerning road traffic congestion, and its properties conceptually correspond to the general typology of goods (see Section 2.2). The rivalrous multilateral externalities hence contrast with inexhaustible externalities such as noise emissions or environmental pollution (cf. Baumol, 1975, pp.19).

Matthews and Menaz (2008, p.24) indicate that two diverse problems arise from limited airport capacities: scarcity and congestion.[5] Congestion refers to the state where demand exceeds supply and creates an overuse of the airport facilities. Scarcity denotes the effect that some participants may not gain market access at all because the resource has been quantitatively constrained. This distinction is based on a quota solution that restricts airport activity by means of airport slots. With regard to the above reasoning, however, this study proposes a slightly more generalized interpretation of the airport capacity problem: As scarcity is a problem that arises with a limited resource, it simply illustrates the property of rivalry in consumption. This rivalry characterizes both a private and a mixed public good, while it is dissociated from a pure public good. Congestion thus arguably constitutes the consequence of a resource scarcity that lacks exclusion, such as in the mixed public good case. This resource scarcity becomes manifest in excess demand and in congestion. As a result, scarcity rather denotes the capacity shortage itself rather than its consequence, which means that congestion simply follows from resource scarcity. Under the condition of non-exclusiveness, these two terms may therefore be stated to indicate a consequential relationship rather than two separate problems.

Put differently, the above definition of scarcity ex ante presupposes a quota solution that might yield exclusion for some participants. A broader view on the capacity allocation problem that abstracts from a particular regulation scheme may be obtained if congestion is interpreted as simply denoting the welfare effect of resource scarcity. The questions to answer are hence the following two: Firstly, to what extent does the congestion damage affect allocation efficiency? Secondly, which regulative measures for restricting congestion best account for an increase in social welfare given the natural market structure?

[5] This notion corresponds to the view in NERA's (2004, pp.50) report to the European Commission.

As already delineated, an efficient allocation instrument exactly balances excess demand and limited capacity to an equilibrium where the consumer and the producer rent net of congestion costs are maximized. In this respect, note that with a private good, the pricing process would match demand and supply to yield an optimal output level such as in the club good case (see above), whereas with an exogenous capacity constraint, some participants will simply be excluded from the consumption of a good or access to a facility. Justifying such an exclusion in a market solution would require that the excluded participants' valuation of capacity is lower than the cost that they have to afford for its use. Note, however, that such a market solution need not necessarily correspond to allocation efficiency if other market distortions are present as well. This problem is further explored in Section 3.3.2.

Feess-Dörr (1995, p.19) state that regulation needs not only to precisely meet its goal. Rather, from a policy point of view, its implementation requires political and social feasibility, while from an economic perspective, it needs to be cost efficient, exhibit minimal transaction costs and account for its own macroeconomic and sectoral impact. In this respect, Starkie (2008b, p.135) points out that regulation is concerned with competition because an external intervention itself alters the market structure. As a consequence, regulation itself may further distort the resource allocation and affect rent distribution. This implicit risk requires a trade-off between imperfect competition and imperfect regulation. As Dixit and Stiglitz (1996, p.188) surmise, the three innate problems accompanying a market consist of external affects, distributional equity, and the natural market structure (with regard to size effects). As a consequence, they note that the "basic issue" of production consists of the question whether market solutions lead to the "socially optimal quantities" of goods given "prices, output, market entry and exit". This study's investigation of allocation efficiency corresponds exactly to the above basic question about the natural market structure and its effect on prices and outputs with regard to the social optimum. Ultimately, Savage (2006) defines three distinct issues that regulation in transport is concerned with: product quality, safety, and competition. This model is concerned with the impact of regulation policy on allocation efficiency by considering its effect on the market structure, but abstracts from the focal points of product quality and safety regulations.

2.4 Current Allocation Schemes

Europe and the United States of America represent the two "largest and most complex air traffic management systems in the world" (Donohue and Zellweger, 2001, p.7). From a systemic perspective, these two areas crucially differ in respect to their regulation policies concerning airport capacity allocation: As both Liang et al. (2001, p.18) and Gillen (2008, p.52) summarize, this major difference consists in the notion that European regulators generally seem to prefer an administrative coordination scheme based on airport quotas, whereas the US policymakers seem to rather favor market approaches.[6] Most other world regions concerned with capacity regulation follow the European approach, such as Canada, Australia and South America.[7]

2.4.1 Two different Paradigms

Access to congested European airports is restricted by means of airport quotas. For this purpose, the major airports suffering from congestion are defined as so-called coordinated airports, where the efficient number of aircraft movements per unit of time is determined. A corresponding number of quotas is defined that reflects this operational airport capacity and is subsequently allocated to the distinct participating airlines at an airport (see e.g. Ulrich, 2008, pp.9). In this respect, the airport quotas are usually referred to as airport slots. One airport slot includes the right to use "the full range of airport infrastructure necessary to operate an air service at a coordinated airport on a specific date and time" (Tanner, 2007, p.99).[8]

In the European Union these airport slots constitute legally established access rights. These access rights are based on the recommendation by the International Air Transport Organization (IATA) known as the Worldwide Slot Guidelines (WSG; currently IATA, 2014).[9] They have been incorporated in the European Slot Allocation Regulation based on EU regulation (EEC) No.95/93 and (EC) No. 793/2004 (see again Ulrich, 2008, pp.9 and Bauer, 2008, p.152), and are implemented in practice by the European Airport Coordinators Association's (EUACA) EU Slot Guidelines Nr. 1 to 3 (EUACA, 2013, 2014a,b). Because a neutral coordinator supervises this quota allocation, this scheme is usually referred to as the administrative allocation of airport slots in the literature. Note that these airport slots must not be confused with the tactical "slots" imposed by air traffic control for operational flow management in case of capacity disturbances during daily operations (see Section 2.4.3).

In contrast to the European case, Whalen et al. (2007, pp.7) report that in the US region, historically, only four airports have been coordinated by administrative slots, whereas at

[6] In both regions, airport operators charge landing-fees for commercial aircraft in order to account for the use of infrastructure and facilities. These user fees are not considered to be of a regulatory nature within the scope of this study.

[7] In the meantime, internationally leading hubs have also developed in the Asian-Pacific region (De Wit et al., 2009, pp.639). From the literature, however, little is known about regulation in these systems.

[8] An airport slot (or 'slot') is defined as a "permission given by a coordinator for a planned operation to use the full range of airport infrastructure necessary to arrive or depart (...) on a specific date and time," (IATA, 2014, p.16).

[9] Previously, this document was known as the Worldwide Scheduling Guidelines (see, e.g., IATA, 2010).

present only two of them remain. However, in some cases conferences were also held by the regulator in order to "secure voluntary cutbacks from all operating airlines" at congested airports (cf. idem, p.9). Otherwise, capacity is generally available on a first-come first-served basis (Madas and Zografos, 2010, p.275). In that case, excess demand becomes manifest in aircraft queuing for take-off and landing. This regional difference in regulation policy also concerns other topics of US airport regulation: As Graham (2004, p.63) points out, there are major distinctions with respect to ownership, financing, airline relationships as well as facility and service provision rules, which may also have an impact on how airports are operated. Liang et al. (2001, p.18) conclude that the European system generates a more uniform demand throughout the day, whereas the utilization of US airports points toward more pronounced peak-period patterns. On an overall level, however, they find that the situation at airports "working close to their capacity limits" are similar in both regions and hence do not consider the effects of the different regulation policies to be fundamentally distinct.

2.4.2 The current administrative Allocation Scheme

The European administrative airport slot allocation takes place in semi-annual strategic coordination conferences where all airports and airlines concerned bargain over the slot allocation under supervision of a neutral coordinator. The initial quota allocation for this bargaining is performed by the coordinator according to the airlines' operational requests and is free of charge. The airlines' requests are prioritized according to the so-called grandfathering rights of the participating airlines. These grandfathering rights reflect the principle that air routes operated regularly during one period are granted airport access at first priority during the next period. As a result, established connections and schedule changes have priority over the expansion plans of the airlines (see Ulrich, 2008, pp.10 and De Wit and Burghouwt, 2008, p.150).

Within this process, a marginal share of quotas remains unallocated as it is reserved for market entry of new participants and for expanding airlines. This quota reserve constitutes the so-called slot pool. After the final quota allocation, a monitoring process enforces the return of insufficiently utilized quotas. In this respect, a use-it-or-lose-it policy is implemented by means of the so-called 80-20-Rule, which dictates that slots must be operated on four out of five sequential traffic days. If an airport slot is not properly operated, it has to be handed back for a subsequent reallocation according to other airlines' requests or for the augmentation of the slot pool (cf. Ulrich, 2008, pp.17).

In practice, this administrative allocation process is described as "extremely successful" (Ulrich, 2008, p.10) and is argued to ensure transparency, fairness and non-discrimination (Bauer, 2008, p.152). Moreover, Ulrich (2008, p.19) claims a number of crucial advantages such as a nearly worldwide applicability and acceptance of the system, the well-defined "priority rules valid for everybody," the low cost of the system itself, and a relative flexibility to respond to special situations despite the regulated environment. Also Bauer (2008, p.152) finds that the administrative allocation policy has been "very successful in maintaining (...) coherence and stability in the international air transport system" and argues that the planning reliability for airlines has been "vital" for their decisions on major capital investments and for their schedule adjustments according to passenger demand. Moreover, he concludes that the schedule continuity reduces search and transaction costs for passengers while providing a "high certainty about flight movements" to both airport and air traffic control authorities (idem, p.155). As a consequence, the grandfather rights as the "foundation of the current slot allocation" are judged as an important principle allowing for the efficient use of airport capacity while at the same time ensuring fairness (again Bauer, 2008, p.152 and p.155). As a main requirement, however, Ulrich (2008, p.13) points out that the slot coordinator has to strictly adhere to the principles of neutrality, transparency, and non-discrimination.

In contrast to the above reasoning, however, recent literature has raised severe criticism of the above administrative quota allocation scheme: The grandfathering privilege of established carriers does not allow the allocation of slots according to the willingness to pay of the airlines and does not reflect the real social value of the airport slots, so that the airport capacity allocation is both inefficient and inequitable (cf. Matthews and Menaz, 2008, pp.24 or De Wit and Burghouwt, 2008, pp.148). Moreover, the prioritization of established airlines implies entry barriers for business rivals and reduces the degree of competition in the market (Daniel, 2009b, p.22). These "unequal chances" (Kilian, 2008, p.255) from the above market barriers thus compromise the general European infrastructure policy goals (cf. Van Reeven, 2005, p.711).[10]

In addition to the above efficiency and distributional concerns, Button (2005, p.55) mentions further potential inefficiencies based on the system design: The administrator might deviate from the target function that stipulates social efficiency either for comprehensible goals such as environmental policy or for disguised political reasons such as the provision of rents to voters and other stakeholders. In this respect, one might imagine that the public perception of social welfare is erroneously diverted from maximum economic rent to minimum aircraft noise or other emissions, operating hour restrictions or minimum flight delays. In such cases, the regulator might simply be interested in reducing total output (Verhoef, 2010, p.326). Moreover, Kost (2003, p.108) raises legal doubts concerning the warranty of commercial

[10] Van Reeven (2005, p.711) points out that the EU infrastructure policy requires air traffic constraints to be non-discriminating and externalities to be borne according to the user-pays-principle.

freedom of action and equal opportunities which need to be granted based on current general economic and competition policies. This criticism climaxes in Daniel's (2009b) radical notion that the administrative allocation scheme may be "accurately, if uncharitably, characterized as industry regulation of entry and exit explicitly designed to preserve incumbent airlines' private quasi-property rights" in the use of major public airports.

From an economic point of view, both the quota determination and allocation might be efficient as long as perfect information prevails and the market is competitive (see Section 3.2.1). The administrative allocation from practice, however, might suggest that complete information may not be available to the neutral coordinator. Consequently, first-best allocation efficiency might be difficult to reach in an administrative grandfathering quota scheme. On this account, the current European legislation is considering an extension toward the secondary trading of airport slots (see EC, 2011). This policy extension draws, among others, on final report SDG (2011), which assessed the efficiency potential for alternative allocation schemes based on the expertise of different policy experts and airport coordinators.

2.4.3 Airport Slots vs. Calculated Take-Off Times (CTOT)

Note that there is a risk of confusion concerning the use of the term "slot" in the aviation context because this term is used in two distinct meanings. The official connotation of a slot refers to the airport quotas as discussed above, which constitute the airport access rights as defined by the corresponding legal documents. At the same time, however, the term slot is also utilized as a proxy for the calculated take-off times (CTOT), which were introduced in conjunction with the tactical air traffic flow management (ATFM) system in Europe.

The ATFM system is designed to optimize queuing and holding times for aircraft in the case of unforeseen disturbances within the daily airport operations that decrease the declared capacity of airport facilities or certain regions of the en-route airspace. For this purpose, it provides an individual calculated take-off time (CTOT) for each aircraft concerned with a capacity downgrade, which accounts for the beginning of an aircraft's trajectory in space as a function of time in order to maintain an "orderly and expeditious flow of traffic" (Tanner, 2007, p.35). Due to the tactical character of this system, it does not represent a strategic regulation policy but rather a tool for demand management given temporary shortages of supply. Despite their distinct nature, these CTOTs are commonly also referred to as "slots" (idem:60). By contrast, the airport slots investigated in this study constitute a strategic instrument for airport capacity regulation in the long run. Because the properties and managing processes of the two instruments are genuinely distinct (Cook, 2007b, p.97), they must not be confused: In official terms, slots denote strategic airport quotas and not tactical CTOTs.

2.4.4 En-Route Airspace

The en-route airspace in general has to date not been regulated on a strategic level (see e.g. Daniel, 1995; Odoni, 2001).[11] On this account, Daniel (1995) explains that the allocation of airspace demand reflects a dynamic short-term problem based on the characteristics of the flexible instantaneous operational flight planning.[12] In addition, Liang et al. (2001, p.38) mention that European airspace is excessively fragmented: They calculate that European core airspace has nearly twice as many airspace sectors than the comparable core airspace in the US. As each sector is controlled by its national authority, the number of European air traffic control centers would be only half of the American ones. They attribute this fragmentation to the existence of the manifold national states and boundaries (cf. idem, p.18). As a consequence, they suggest that fragmentation and the influence of a large number of different authorities might make common policy-making difficult.[13]

Concerning the supply side, EUROCONTROL does not expect the en-route airspace to become a critical capacity constraint (EUROCONTROL, 2008, p.1). However, other authors suspect present as well as future congestion problems within the en-route airspace system (cf., e.g., Majumdar, 2007, p.65). Moreover, as Oster and Strong (2007, p.2) calculate, the user fees for airspace use depend on the mileage traveled through each sector, but do not mirror the operating costs of air traffic control - let alone congestion and the opportunity costs for foregone capacity. These arguments indicate the need to further assess the topic of strategic en-route regulation. However, this subject exceeds the scope of this study and is not further addressed.

2.5 Recent Criticism

This section briefly reviews the criticism from recent literature on the current administrative quota allocation scheme. The arguments are attributed to five different views: an economic

[11] In this respect, the tactical air traffic flow management (ATFM) is not considered as a strategic regulation mechanism (see Section 2.4.3).

[12] See Tanner (2007) for comprehensive details on the principles and practice of operational flight planning. Operational flight planning starts seven days prior to a flight event while the definitive routing depends, e.g., on weather, winds and operational constraints and is not determined until shortly prior to a flight.

[13] This problem was addressed by the Single European Sky (SES) project as of 1999, where the first operational improvements were projected to rise as of 2012 (see Van Houtte, 2007 and EUROCONTROL, 2009). A major reform implementation has not yet been reported.

efficiency perspective, a view on technical and transactional efficiency, the political debate on distributional effects and legal considerations about policy accordance. In this respect, economic efficiency is reached if a limited resource achieves the maximum amount of economic rent for both passengers and airlines. It implies a socially optimal allocation of capacity. Technical efficiency refers to the utilization degree of allocated capacity and may therefore also be considered as effectiveness. The transactional efficiency concerns the costs accruing from the regulation system itself. The distributional debate addresses the distinct costs and benefits that accrue to each stakeholder from resource allocation. Ultimately, the legal aspects cover the accordance of the quota allocation scheme with current regulation policy.

2.5.1 Economic Efficiency

In the economic context, efficiency is commonly measured by social welfare as the sum of consumer and producer surplus (see Section 3.2.2). Allocation efficiency based on airport quotas thus requires both the determination of the optimal output quantity as well as the socially optimal allocation of the constrained resources to the participants (Forsyth and Niemeier, 2008, pp.65).

The main argument against economic efficiency under the administrative scheme concerns the social value of the quota allocation: It claims that the airport access rights are not allocated to those airlines that can make the "most valuable and beneficial use" of them (De Wit and Burghouwt, 2008, p.152). Moreover, the quota scheme still induces airport delays while the carriers do not recognize the externalities which they cause (Matthews and Menaz, 2008, p.26). As a consequence, the scarce airport capacity is suspected to be available at a price below its marginal social costs. On this account, Madas and Zografos (2010, p.281) also claim an insufficient "mismatch management" of capacity versus demand. This claim is illustrated by the fact that, e.g., 9% of airlines' demand for airport capacity could not be allocated while at the same time an 8.7% of allocated slots were returned early within one season.[14]

The second issue with economic efficiency concerns free market access: Because the grandfathering rights are conceived to create barriers of entry, the established airlines are suspected of "generating scarcity rents" (Gillen and Morrison, 2008, p.174) based on limited competition (Gillen and Morrison, 2008, p.174). According to Matthews and Menaz (2008, p.27), these rents would still motivate the major airlines to hold on to their allocated capacity, even if that capacity were not efficiently used. This inefficient slot usage occurs despite (or even

[14] These numbers reportedly concern the European 2002 summer season.

due to) the provisions to avoid improper slot hoarding by the airlines for the purpose of restricting competition in terms of the 80-20-Rule. The improper utilization points toward slot preservation to keep out newcomers and to secure rents by means of "baby-sitting" these slots with flights that are not worthwhile otherwise, such as uneconomic routes with low load factors, with small aircraft, or even lending them to alliance partners (De Wit and Burghouwt, 2008, p.152). From the above critical perspective, the quota allocation can be considered to guarantee the established airlines a degree of market concentration which allows them to both obtain scarcity rents while at the same time to profit from economic cost savings.

As Gillen and Morrison (2008, pp.175) point out, however, the evaluation of allocation efficiency against market concentration is not unambiguous: Firstly, large airlines could be realizing cost savings due to economies of scope such as lower average operating costs than two or more airlines. In that case, a forced allocation of airport capacity to other competitors would result in an efficiency loss on the production side. Moreover, if the total amount of airport quotas constrained overall flight volume below the monopoly output of a single airline, a quota reallocation to new entrant competitors would not lead to an increase in outputs and in consumer surplus despite the higher number of competitors. Also in this case, lowering the barriers of entry and increasing competition would improve neither productive efficiency nor consumer welfare. As a consequence, the question of optimal regulation concerns the naturally occurring market structure in the first place. This question is further explored in Section 3.1 from a theoretical perspective, while the natural market structure in this model is revealed by the equilibrium computation in Section 6.5.

2.5.2 Technical Efficiency

The technical efficiency is dissociated from the economic efficiency in the way that it measures the number of utilized versus the number of allocated quotas rather than the social value of the allocation itself (Bauer, 2008, p.152-154).[15] The core instrument of the administrative regime governing technical efficiency is the slot monitoring process (see Section 2.4.2 above). It ensures that the slots allocated are either used or returned in a timely manner and only cause a minimized opportunity cost of forgone capacity. After an extensive review, Bauer (2008, p.169) concludes that technical efficiency of the administrative regime is satisfactory. Also in this respect, however, Madas and Zografos (2010, p.281) claim empirical evidence for their insufficient "misuse management": 15% of the initially allocated slots were not operated

[15] As Bauer (2008) mentions, the technical efficiency of slot utilization may also include other dimensions such as aircraft weights or passenger numbers. In this respect he argues that the ratio allocated to used slots was most meaningful, however, because most other metrics could not be properly quantified or administered.

during the summer period of 2002 in Europe. Based on this number, they estimate revenue losses of 20 million euros caused by late cancellations of unused slots at congested airports.

In addition, Daniel (1995) expresses qualified objections concerning the quotas' effectiveness in actually reducing congestion. The argument is that slots are "too long for precise peak spreading", which signifies that they cannot prevent traffic from peaking within the single time windows as defined by each slot. Consequently, the number and length of slot windows will critically affect the congestion levels and efficiency gains achieved.[16] As a result, Daniel (1995) concludes that a first-best solution required one slot window per operation, which in practice could prove to be difficult to implement. In the meantime, however, extensive research has been undertaken on this subject. This research has led to various practical implementations of collaborative decision management systems (referred to as airport CDM; see Modrego et al., 2009) and of arrival and departure management systems (referred to as AMAN/DMAN systems; see Deau et al., 2009). Such systems aim at optimizing the tactical queuing in airport demand with regard to actual aircraft movements and operational capacity variations at an airport.

The above contradicting perspectives are not further resolved in this study as they concern the design and implementation technicalities of the quota scheme. Nevertheless, they may also be conceived of as supporting the above complaints about economic efficiency and thus the rationale of this study. As a shortcut solution to the congestion argument, the remainder of this study will assume that the instrument of airport quotas effectively controls for flight delays.

2.5.3 Transactional Efficiency

Transactional efficiency considers whether the regulation scheme itself only causes minimum cost (Feess-Dörr, 1995, p.19). This type of efficiency represents an implementation issue rather than an economic problem concerning the resource allocation itself. Some authors stress that complicated regulation schemes would be difficult and expensive to implement and would later cause excessive costs to the users, which might even overturn the efficiency gains of the regulation scheme itself.

Although specific data on transaction costs for the current administrative scheme are not available in the literature, Ulrich (2008, p.19) notes that direct costs from slot allocation are estimated to be less than EUR 2.50 per flight movement. He concludes that any commercial

[16] The current time-frame for an aircraft movement within an airport slot is fifteen minutes (-5/+10).

system for slot allocation would be much more expensive. Without further investigation, one may conclude that the transactional efficiency of the current system remains within plausible dimensions.

2.5.4 Distributional Effects

The distribution of rents to the participants ultimately constitutes a political aspect of the airport capacity allocation debate Button (2005, p.47): In contrast to economic efficiency, it is not concerned with the overall maximization of social value but with the distribution of the gains and losses arising from the capacity regulation across the participants.

According to Kilian (2008, pp. 255), the main point for criticism involves the unequal chances arising both from privileging the established airlines and the rising barriers to market entry for new entrants. In this respect, he attributes little help to the slot-pool because the latter contains commercially invaluable slots only. These slots are supposedly located in the extreme off-peak periods where flight demand is difficult to accommodate, and therefore also referred to as "moonlight slots" (see also Matthews and Menaz, 2008, p.27). As mentioned in Section 2.4.2 above, the distributional criticism culminates in Daniel's (2009b, p.22) notion that the "quasi property rights in public airports" for established carriers are protected by the current administrative quota scheme. In a similar but less extreme manner, Starkie (1998, p.112) notes that the grandfathering rights might lead to an inequitable distribution of profits to established carriers. According to Gillen and Morrison (2008, pp.174), this inequality also exists if the scarcity rents from protected market access only depend on the number of flights but not on the number of competitors.

These distributional concerns are not explicitly considered in this study as the investigation focuses on the evaluation of allocation efficiency. As explained below in Section 3.2.3, these two issues are distinct problems that can and need to be solved separately: Once efficiency is achieved in the resource allocation, any required distribution of rents can be achieved by means of transfers - at least in theory. Put differently, this signifies that an allocation with a higher market concentration and lower competition could be a relative deterioration for passengers but might nevertheless account for an overall welfare gain and should be undertaken. Nonetheless, due to the importance of the distributional effects, further research might focus on this problem. In this respect, a stakeholder analysis would help to judge the political feasibility of each beneficial reform option and to catalyze the consideration of potentially required wealth transfers.

2.5.5 Legal and Policy Accordance

From an infrastructure policy perspective, the reflection of all externalities and infrastructure costs by taxes and charges is a declared European Union policy goal in transportation as a whole. In this respect, both the polluter pays principle as well as some “clear fiscal incentives” should help to reduce congestion in all modes of transport, and an “appropriate charging” of all participants should improve the usage of scarce infrastructure capacity (Van Reeven, 2005, p.722). From a competition point of view, the sector-specific transport policy prescribes that the European airline market be open to all licensed EU air carriers. As a consequence, only two cases exist where restrictions on air traffic are explicitly allowed: Firstly, on traffic within closed airport systems; and secondly, in the case of serious congestion or environmental problems (Van Reeven, 2005, p.711).

As a result, the restrictions on market access and air traffic arising from the European slot allocation regulation appear to implement a goal regarding overall congestion and are justified within the above sector-specific regulation policy. However, both the competition and the user-pays aspects are less evident: Airport slots are free of charge while according to Starkie (1998, p.112), the normal landing fees at overcrowded airports cover neither infrastructure costs nor congestion damages.[17] Whether such a level of airport charges is in accordance with the general infrastructure policy, however, at least seems to be arguable.

In addition to the above concerns, the grandfathering allocation also invokes major concerns regarding the general fundamental rights with public infrastructure: Kost (2003, pp.98) postulates that the public authorities need to grant the public rights of commercial freedom of action and of equal treatment in every sector regardless of whether the provision of services or infrastructure is public or private. On the one hand, he assesses the grandfathering quota scheme as a systematic rather than an arbitrary approach able to provide stability within the public service in air transportation. Nevertheless, he finds that this approach causes serious violations of both of the above public rights that needed to be granted on the grounds of public law (idem, p.110 and 116).

From a legal and policy point of view, the above arguments lead to the conclusion that the current air traffic infrastructure regulation scheme based on grandfathering rights is generally acceptable. However, like any costless policy, it challenges the user-pays principle regarding infrastructure use and congestion. Moreover, from a legal perspective this instrument points toward difficulties in granting the fundamental public rights of freedom of action and equal treatment to all participants.

[17] According to the estimates by Oster and Strong (2007, p.2), the same applies for the en-route airspace.

2.5.6 Summary

In sum, the above discussion confirms the notion that the main criticism concerns the "highly controversial principle of grandfather rights" as a foundation for a socially optimal allocation of airport access rights (Matthews and Menaz, 2008, p.26). As claimed by Gillen and Morrison (2008, p.181), however, an optimal resource allocation requires the knowledge of the natural market structure. Therefore, this study will abstract from the above technical and legal issues and focus on allocation efficiency by providing a model that appropriately reflects a market structure with asymmetric network density effects.

2.6 Alternative Instruments

2.6.1 Fundamental Concepts

The fundamental propositions for market-based externality regulation in economic theory have been developed throughout the last century. These ideas arose with the concept of the PIGOUVIAN tax, based on Pigou's (1924) seminal contribution. They led to the road and congestion pricing literature, which was established by the fundamental work of Vickrey (1969) and Levine (1969). The airport congestion pricing scheme was developed from the road pricing idea and reflects the crucial difference that airport users - as opposed to road users - are considered as non-atomistic (cf., e.g., Button, 2008, p.578).

2.6.2 Recent Propositions

In the recent capacity allocation discussion, the two most prominent alternative instruments are known as congestion pricing and secondary trading: As mentioned above, a congestion pricing scheme aims at charging the marginal social costs from congestion externalities to the users. This concept exactly represents a PIGOUVIAN tax that internalizes the negative external effects. By contrast, a secondary trading scheme replicates a market place for the trading of airport slots among the participants. This scheme requires an initial allocation of airport quotas. Subsequently, it introduces opportunity costs for slot holding and thus explicitly commercializes the social value of scarce airport infrastructure. As a consequence, trading according to these values should occur and lead to an efficient market allocation of capacities (cf. e.g. Brueckner, 2009a, pp.682). In addition, such a free market is expected to lower barriers of entry and increase competition to the benefit of the consumer (Gillen and Morrison, 2008, p.174). The current administrative quota allocation and the two proposed alternative instruments above reflect the three fundamental economic instruments for the regulation of externalities, which are explained in Section 3.2.1.

However, Borenstein (1988, p.382) noted that there was “no theoretical justification” that the airlines’ profit-maximizing rationales correspond to the socially efficient surplus generation within market approaches for resource allocation, especially in the presence of output inefficiencies related to market power. These welfare caveats in conjunction with market power are delineated in detail in Section 3.3.2.

2.6.3 Other Schemes

In addition to the above two propositions, some studies have suggested slot auctions as an alternative scheme for the initial allocation such as, e.g., Rassenti et al. (1982) and Gale (1994). Similarly to the secondary trading idea, an auction should yield a capacity allocation according to its market value. However, because the current grandfathering quota scheme has been profoundly established in the industry, a transition from the current allocation to a novel initial allocation within the established market shares is expected to yield implementation problems. At the same time, a secondary trading market based on the current allocation scheme might ultimately yield the same or an even higher degree of allocation efficiency as a quota auction (cf. Sentance, 2003, p.56). Therefore, the importance of this idea has diminished in the recent literature, so that an initial quota auctioning scheme is not further discussed in this study.

3 Theoretical Foundations

This section first explores the natural market structure that supposedly arises from asymmetric airlines at a large network hub airport. Thereafter, it briefly presents the theoretical background concerning the resource allocation problem in light of congestion externalities and with regard to the dual market distortion, which arises if congestion occurs in conjunction with market power. In addition, it provides a short introduction to the two commonly used oligopoly models and a distinction between strategic competition and strategic behavior.

Its main implications are that the asymmetric network market structure is most suitably reflected by an oligopoly model with product differentiation and COURNOT quantity competition, that both the current as well as the proposed alternative airport capacity allocation schemes are based on the three basic economic instruments for the regulation of congestion externalities, and that the dual distortion induces welfare caveats to these allocation instruments that do not arise with externalities alone.

3.1 Market Structure

As the efficiency of regulation crucially depends on the natural market structure, the network structures currently observed in practice raise the question of whether the airline market at major airports is one of a natural monopoly (cf. Gillen and Morrison, 2008, p.175). This section therefore briefly explores the general features of competitive advantages related to the market structure, and justifies the asymmetric duopoly assumed in this model.

3.1.1 Economies of Scope and Scale

In the literature, economies of scale and scope are generally defined as cost-side advantages. Economies of scale prevail if average costs decrease when output quantity is increased. They hence occur mainly associated with fixed costs and can differ over the range of output: At low levels of production, decreasing average costs point toward economies of scale. However, at high levels of output, average costs may increase again, e.g., due to inefficiencies in the firm size. The initial economies of scale thus revert to actual diseconomies. This relationship depicts the u-shape of the standard average cost curve. With linear average costs, marginal costs are constant and there are no economies of scale. By contrast, economies of scope arise if any output quantity of a good can be produced at the lowest total cost by one single firm rather than by a number of firms. Also, the degree and sign of this scope effect may vary with the level of output. A natural monopoly hence only exists with declining average costs or with small scale diseconomies where sharing the output between two firms is costlier than single-firm delivery. In contrast, a competitive market evolves when the scale advantages of a single firm fall short of total market demand. In that case, obtaining production at minimum cost requires a large number of firms (Train, 1992, p.5-8).

Natural monopolies emerge from either of these two sources. The existence of scale economies is dependent on the relation of average costs relative to market demand, i.e., on whether this advantage exists over a “sufficient range of output relative to demand” (Train, 1992, pp.5). The same is the case for the economies of scope because both effects depend on the overall cost situation: If the cost curves exhibit subadditivity, then any good is produced at a smaller cost by one single firm than by multiple firms. Consequently, a natural monopoly arises regardless of how the output is divided across the firms (idem, p.5). The question of the natural market structure hence is one of average costs and total market demand.

An application from the airline case is the question of whether it is more efficient to satisfy demand on a route by one airline rather than by two or more competing flights on the same route. As in the case of the natural market structure, the answer depends on total demand: If both flights clear the market at sustainable prices, there are no economies of scope. If only half of the capacity can be accommodated, the single firm delivery would be more efficient. This means that the seat load factor determines whether there are potential economies of scope. If it is achieves one-hundred percent, no advantage is gained by the single firm delivery. The succeeding question is whether it is efficient to replace two small, fully-booked aircraft with one large yet also fully-booked one. This question concerns the economies of scale: As stated above, if marginal costs are constant, nothing is to be gained by the use of a larger aircraft.

On the one hand, Doganis (2002, p.120) states the general rule that “the larger an aircraft the lower its direct operating costs per unit of output”. This would signify that direct operating cost increases would not be proportionate to aircraft size and would exhibit economies of scale. Also Harback (2005, p.2-3) argues that hub-and-spoke networks exploit economies of scale in aircraft size. On the other hand, Jäggi (2000, pp.61) reports that there is a consensus in the literature in the sense that the realization of economies of scale in the airline industry is negligible. In other words, raising output based on aircraft size would be unlikely to yield cost advantages on the production side. The discussion in the literature hence seems to remain controversial. In order to focus on the network density effects, cost-side advantages in terms of economies of scale and scope are not considered in this model (also see Section 10.4.2).

3.1.2 Economies of Network Density

In contrast to the economies of scope and scale that accrue on the cost side only, network benefits and benefits from the "vertical integration of regional feed" are recognized as "revenue related economies". These so-called "network synergies" are attributed to the network "density and general market presence" of an airline (Jäggi, 2000, pp.61). As Gillen and Morrison (2008, p.177) illustrate, e.g., adding one destination in a network increases the number of connections in the dimension of $2n$, where n is the number of airports already served. The corresponding benefits that arise to passengers of a network airline as compared to the customers of a non-networking airline may be described as the differentiation of identical products based on associated services or values. This concept was introduced by Friedman (1983, pp.50) in general oligopoly theory and corresponds to Wöckner's (2011) notion of product differentiation by product image.

Indirect benefits from network goods have been described and formalized in economic network theory (see, e.g., Belleflame and Peitz, 2010, pp.550). According to their nature, they correspond to the prominent concept of customer value from the business administration literature. The latter describes the overall value of a product to the customer that not only includes utility from consumption but also from certain attributive values of a product (Woodruff, 1997, p.142). While product attributes within the utility function were introduced by Dixit and Stiglitz (1977) and Krugman (1980) in the monopolistic competition literature, the integration of customer value into the concept of economic utility was introduced by Wichers's (1996) theory of individual behavior. Customer value as a competitive advantage against competitors was proposed by Woodruff (1997) and is now a standard argument in marketing theory (cf., e.g., Shankar and Carpenter, 2012). The economic foundation for this competitive advantage is simply vertical product differentiation (see e.g. Wöckner, 2011, pp.15).

In this respect, Joppien (2003, p.124) describes how an airline's market share in terms of passenger volume represented an S-curve relative to the market concentration at an airport. This relationship indicates that passengers prefer to fly with the carrier that has the higher market presence. Wei and Hansen (2005, p.325) empirically confirm this "s-curve effect of service frequency on an airlines market share." The fact that indirect travel benefits in the S-curve relation may either be abstractly based on market presence or concisely measured by service frequency ultimately illustrates the distinction between direct and indirect flight benefits. In this respect, real additional services and travel options arising from network density should hence further justify the concept of indirect utility based on network density benefits. Consequently, this indirect utility may be expected to induce an additional willingness to pay on the customer side, which can be commercialized by the networking airline. As a consequence, Gillen and Morrison (2008, p.177) indicate that allocation efficiency could therefore even increase along with market concentration. This reasoning, finally, corresponds to the dilemma of airport concentration, as referred to by Starkie (2008a).

Daniel (1995, p.357) confirms that complex effects may arise in a network based on interactions between connecting flights. As Czerny (2006, p.4) notes, these interactions based on "the network character of the industry" yields that the airline's demand for airport capacity normally is complementary. Aguirregabiria and Ho (2010, p.1-2) argue that the airlines' profit functions across routes are interdependent due to complementary demand. This interdependency is referred to as supermodularity. The latter implies that negative profits might be taken on some routes if they can generate profits on other routes. Also, Joppien (2003, p.121ff.) states that market power is distinct across routes in a heterogeneous market structure, so that airlines can apply cross-subsidies or cash-transfers to maximize their profits. Different capacity valuations and complex network adjustment patterns may therefore arise from network airlines concerned with network benefits and complementary demand.

In this study, the above concept of network density economies is simplified to the point where the latter constitute additional indirect utility for customers that arises from the network structure. This indirect utility depicts the benefits from potential travel options that passengers gain from, e.g., a wider destination choice or a higher flight frequency. In this respect, the network effects are referred to as density benefits. Although these benefits depend on network width and depth, the latter dimensions of the network are unified in the abstract notion of the network density, which itself is approximated by the networking airline's flight volume. This specification is formally delineated in Section 5.2.2.

3.1.3 Network Density Benefits as Vertical Product Differentiation

In most recent studies concerning airport capacity allocation, flights are considered as perfect substitutes. Hence, they account for transportation between two points in space only and may not differ in price. The equilibrium with all firms remaining in the market thus yields one uniform market price for all firms, regardless of whether firms enjoy market power or not (also see Section 3.4). Based on the above argument concerning the demand-side network economies, this study introduces the idea of imperfect substitution of flights on identical routes or markets. This product heterogeneity should reflect the differences in customer value between flights of airlines with distinct business models. As already stated, such distinctions may be justified by Friedman's (1983, pp.50) identical product differentiation, which suggests that identical products from different sellers can be distinguished based on side advantages such as a better repair service or more generous return policies. In the same sense but more abstractly, Wöckner (2011, pp.15) provides product differentiation by image as an argument.

The customer value between two distinct flights may thus differ as a function of the additional services provided. With regard to networking airlines, one might think of advantages such as higher flight frequencies and lower schedule delays based on optimized network planning, which yield more travel options and flexibility at a higher connectivity and at shorter travel times. E.g., Laffont et al. (1998, p.4-5) note potential gains for airlines stemming from passenger benefits based on “interconnections”. Their notion of interconnection exactly denotes the idea of the above connectivity, where passengers may swap across flights in case of travel plan changes, missed connections or, simply, because they “enjoy the convenience of increased departure time variety” (idem, p.4). In this respect, they also mention that these connectivity benefits are related to an airline’s market share.

Flight-related ground services are not explicitly included, but travelers may also expect a high quality of associated services such as, e.g., re-booking assistance and options in cases of misconnections or altered travel plans. These network features are ultimately left unspecified in this model but are summarized as the network density benefits. They account for vertical product differentiation. A dissociation of vertical and horizontal product differentiation is provided in Section 3.1.4 below.

As Gillen and Morrison (2008, p. 178) point out, air traffic networks are likely to exhibit a natural market structure of oligopoly and product differentiation. In such a market structure, a change in market power affects prices, outputs and both airline profits as well as consumer welfare. Essentially, an oligopoly setting with product heterogeneities preserves the notion that competing airlines may increase profits based both on their degree of market power as well as their ability to differentiate their product. Allowing flights of different airlines on an identical route to become imperfect substitutes thus yields endogenous market power in quantity competition with price differentiation. If, in turn, product differentiation is presumably based on network benefits that arise from the density of an airline network, an airline’s market concentration at its network hub is hence a crucial factor for its profitability. Ultimately, product differentiation based on network effects provides an additional rationale for strategic airline competition and behavior concerning the airport access rights as key network assets.

3.1.4 Horizontal and Vertical Product Differentiation

Economic theory distinguishes horizontal from vertical product differentiation: Horizontal differentiation refers to variations in product attributes with regard to heterogenous consumer tastes. In this case, at uniform prices, all goods encounter demand and are delivered. Horizontal dissociation hence does not only depend on objectively determined product characteristics but also on customers preferences. By contrast, vertical heterogeneities account for differences in product quality. Demand for different qualities, in turn, is only justified based on consumer income or wealth. As a consequence, multiple goods of distinct qualities must be offered at distinct prices in order to yield a positive demand for all goods: If prices were identical, only the highest quality good would be sold (cf. Wöckner, 2011, pp.15 and pp.93).

The motivation for firms to differentiate their products is the following: With product differentiation, product substitutability is reduced and competition on the market decreases. Therefore, under both above forms of product differentiation, the firms' profits increase along with a higher degree of product heterogeneity (cf. Wöckner, 2011, p.46 or Belleflame and Peitz, 2010, p.51). In this respect, product differentiation yields substitute relationships across differentiated products in terms of finite cross-demand elasticities, which formally reflect the strategic concerns of market power and market presence across firms (cf. Friedman, 1983, p. 52).[18]

3.1.5 Implications

From the perspective of this study, the potential scale economies are too small to justify a natural monopoly in the airline market. Neither, the network character of the industry supports the flight market to be supplied by a large number of price-taking firms. Rather, a small number of strategically interdependent firms seems to adequately reflect the stylized, asymmetric market structure at a network hub. An oligopoly exactly reflects such a strategic interdependence between firms: namely, neither monopolists nor perfect competitors face "strategic concerns" regarding other firms. Their only difference in modeling is thus the form of the demand curve: Monopolists face a downward sloping demand curve while perfect competitors cope with horizontal demand because they do not have an "effect on market

[18] Vives (2001, p.144) denotes substitute goods by a positive cross-demand elasticity of price $\partial D_i/\partial p_j \geq 0$ for $j \neq i$, and hence by a price-elasticity of demand $\partial P_i/\partial q_j \leq 0$ equal to or below zero based on the inverted demand function.

demand and supply functions". In oligopoly, demand is generally downward sloping so that prices are endogenous functions of output (Friedman, 1983, p.6-8).[19]

Ultimately, also (Gillen and Morrison, 2008, p.178) opine that an oligopoly seems to precisely reflect the airline market structure. The market is therefore assumed to reflect an airline oligopoly with product differentiation, where the natural market structure depends on the importance of the network benefits. As described above, these indirect benefits depend on network density and constitute vertical product differentiation based on product quality. The airlines' strategic concerns thus directly depend on the degree of cross-demand elasticities. This property contrasts to a competitive industry, where the interdependency between the firms vanishes because the firms have to accept the market prices as given (Friedman, 1983, p.9).

Specifically, Train (1992, p.7) shows an illustrative case for a natural duopoly*:* If economies of scale are exhibited over a small range of output only and demand is high, a single firm would face much higher average costs than two firms dividing the market. This reasoning should exactly justify the duopoly between a networking and a non-networking airline modeled in this study: Based on the S-curve above, the network density benefits may be assumed to exhibit increasing returns for a small network size but decreasing returns for a large network size. This relationship precisely corresponds to the above argument of scale economies over a limited range of output. Therefore, it represents the theoretical foundation for the airline duopoly considered in this study.

Holloway (2002, p.23-24) notes that the prevailing market structure may yield competitive advantages in competition. In this respect, Oum et al. (2012, p.432) show that a network structure can be a dominant strategy in oligopoly competition. This result may be explained by the notion that networks provide "superior connectivity and wider market coverage" (O'Connell, 2006, p.60) so that hub connectivity supposedly yields a "tremendous commercial impact" (Gödeking, 2010, p.21). This impact may be related to the passenger benefits arising from connectivity, as denoted by Laffont et al. (1998) in Section 3.1.3. In the same sense, Brueckner (2002b, p.10-11) argues that density economies constitute the basic rationale for the prevailing hub-and-spoke structures in the airline market. As a consequence, one might think of network airlines to differentiate their products against competitors based on network density effects as similarly expressed by Starkie (2008a, p.197).

The competitive network advantage on the production side is supposed to be based on network density, and because network synergies are recognized as revenue related economies,

[19] Technically, this argument omits monopolistic competition, which might also qualify for interdependent effects. However, the latter is characterized by many small firms, where each firm offers a distinct variant of a product (Friedman, 1983, p.7). This also seems not to reflect the current airline market.

airport access rights may justifiably be denoted as “key business assets” for the network airlines (Jäggi, 2000, pp.61 and 271). Consequently, Aguirregabiria and Ho (2010, p.1) show that network density effects may constitute heterogeneities that translate a higher market concentration into higher markups: They empirically find differences in prices and hub sizes across airlines when profit functions account for the complementarity of flight services. The above arguments thus strongly support that a theoretical model should capture the market structure of a large network hub by taking into account product differentiation based on network density effects.

3.2 Resource Allocation with Externalities

3.2.1 Economic Allocation Instruments

In economic theory, there are three general economic instruments to compensate for market distortions caused by external effects: quotas, taxation, and decentralized bargaining. Quotas reduce externalities to a socially optimal level by restricting the activity from which they arise. Taxation aims at adjusting the costs of an activity so that they include the external effects caused by that activity. Finally, decentralized bargaining may internalize all externalities based on free negotiations. If markets are perfectly competitive and if perfect information is available, all of the above three instruments are equally efficient (Mas-Colell et al., 1995, pp.351 and pp.356). This implies that market power must not prevail, and that all costs and benefits must be known to the regulator and to the market participants.

As already pointed out in Section 2.6, both the actual and the recently proposed alternative instruments for airport capacity allocation are founded on the three general economic instruments. Technically, the following prerequisites must be met for these instruments to become efficient: An optimal quota regulation requires a correct determination of the number of quotas according to the socially optimal level of the activity concerned. In addition, the quotas need to be allocated appropriately across the participants so that both the efficient overall output as well as the optimal market shares are replicated (cf.,e.g., Forsyth and Niemeier, 2008). For this purpose, the regulator needs to be able to exactly quantify the efficient level of the corresponding activity. Moreover, the optimal quota allocation to the market participants must be feasible and sustainable.

A congestion tax needs to equal the marginal costs of the externality. Therefore, the exact monetary equivalent of that externality must be known to the tax setting agency in order to impose the correct amount of taxes to each participant. In this respect, note that agents can also be subsidized for not undertaking an activity rather than being taxed while performing it. This signifies that the net tax may become negative. From a welfare perspective, a subsidy with transfers thus may exactly replicate an efficient tax solution (cf. Mas-Colell et al., 1995, p.376).

Finally, according to the Coase theorem, a decentralized bargaining solution is supposed to yield an efficient internalization of the externality: Coase (1960) states that individuals always find an efficient allocation in free negotiations if the property rights of the externality are unambiguously distributed among the participants. More precisely, the property rights need to be institutionally defined and enforceable by law, and the external effect must be measurable and quantifiable. In this case, the bargaining process allows complete internalization and pricing processes (cf. Baumol, 1975, p. 20). In analogy to a classical market situation that is efficient under perfect competition (cf. Mas-Colell et al., 1995, p. 359), an efficient decentralized bargaining solution could thus be achieved based on a market place for the externality.[20] A secondary trading scheme for airport quotas could hence be argued to replicate such a market place.

In theory, both airport quotas as well as a correctly determined congestion tax thus are equally efficient as long as the participants' costs and benefits are known to the regulator. The fact that this perfect information might be difficult to obtain lets us already suspect that allocation efficiency might be difficult to achieve both with administrative quotas as well as with congestion pricing. Also in a competitive market solution, incomplete information among participants may preclude an optimally efficient solution (Mas-Colell et al., 1995, p. 368ff.). In this respect, Matthews and Menaz (2008, pp.25) point out that the participants may prefer not to disclose their private information when facing competition in the allocation of their airport slots, which Jäggi (2000, p. 197) suitably classifies as the airlines' strategic "key business assets." In the same sense, strategic airline behavior in terms of entry deterrence within a trading process may constitute a market failure that precludes an efficient market outcome (see Section 3.5.2). Moreover, Rietveld and Verhoef (1998, p.359) stress that the internalization of externalities is only possible with completely allocated property rights. As Baumol (1975, p. 20) mentions, however, the allocation of property rights is often prevented by institutional obstacles in practice.[21] These arguments indicate that the efficiency of all three above economic instruments may be degraded by imperfect information, non-institutionalized property rights, and by the presence of market distortions other than

[20] In this respect, Mas-Colell et al. (cf. 1995, p. 359) argue that perfect competition was a legitimate assumption specifically in the case of multilateral externalities.

[21] In addition, the expectation of incomplete information may increase the desire for central coordination (Baumol, 1965, p. 201). This tendency increases the difficulty to implement regulation policy based on a decentralized market solution.

externalities - such as market power and strategic behavior.

Based on the above arguments, this study assumes perfect information both for the regulator as well as for the participants. This allows the welfare investigation to focus on the resource allocation problem rather than on instrument design and implementation issues. As a consequence, both the number of quotas and the size of the congestion tax are correctly determined by assumption. Otherwise, incomplete or imperfect information could preclude an efficient solution as early as the conceptual stage with all three above allocation instruments (cf. Mas-Colell et al., 1995, pp.368): One might argue that both the determination of the efficient number of quotas is difficult and the monetary equivalent of the externalities for optimal taxation is hard to obtain, while the market participants are unlikely to reveal all their preferences in a decentralized bargaining solution. Although these problems may arguably occur in practice, this study only considers the allocation problem of efficiently designed regulation schemes.

3.2.2 Efficiency Measure: Economic Rent

Economic efficiency is generally measured in terms of overall social welfare (cf. Mas-Colell et al., 1995, p.326). Social welfare, in turn, is commonly quantified by the Marshallian aggregate surplus, which denotes the sum of consumer surplus and firms' profits (Vives, 2001, p.101).[22] However, economic efficiency does not depend on the distribution of wealth across all participants. As discussed in the next section, this distributional issue is generally independent of the problem of allocation efficiency (also see Blaug and Lloyd, 2010).

The consumer surplus corresponds to each consumer's willingness to pay minus the price for the consumption of the good concerned. Geometrically, it denotes the distance between the demand curve and the market price. The producer surplus equals the firms' net gains arising from the turnover of all goods sold minus their production costs. Social welfare thus implicitly accounts for damages from externalities because the latter affect consumer utility (Matthews and Menaz, 2008, p.26). By contrast, if firms yield excess profits based on mark-ups over marginal costs, the excess profits themselves do not alter welfare. They only affect the distribution of the economic rent between consumers and producers. The market structure giving rise to these excess profits, however, causes a net welfare effect because it reduces outputs below their optimal levels. This market distortion based on output quantity

[22] Vives (2001, p.83) mentions that the Marshallian concept only approximated the actual social welfare because the latter is based on the Hicksian consumer surplus. He formalizes the approximation error to a percentage of $1/\sqrt{n}$ for a price change in one good, where n is the number of goods in the economy (idem, pp.89). This error may hence be considered as negligible (also see Ng, 2010 for a comprehensive review of this problem).

is referred to as the deadweight loss. It arises from downward sloping demand, which yields that a small output diminution increases the price of all remaining sales. Consequently, a firm enjoying market power will charge a price higher than marginal costs while serving a lower number of customers as compared to marginal cost pricing under perfect competition (cf. Mas-Colell et al., 1995, p.385).

3.2.3 Allocation Efficiency and Distributional Equity

In any allocation problem, the issue of economic efficiency is independent of the (possibly controversial) political issue of wealth distribution across participants (cf. Mas-Colell et al., 1995, p. 307). In this context, the wealth distribution determines the accrual of all costs, benefits, and economic rents to the participants. This independence has two important implications: Firstly, once an efficient allocation is achieved any distributional outcome can be reached by an appropriate wealth transfer. Secondly, multiple efficient allocations are all equivalent from an economic perspective. Which solution is chosen will therefore be determined by political goals (see Baumol and Oates, 1979, pp.176). However, as Baumol and Oates (1979, p. 174) also note, in the “world of political reality” the distributional questions may be more prominent than the economic efficiency issues themselves. As a consequence, the actual feasibility of regulation policy is reported to be less dependent on the overall welfare gain for society than on the distribution of wealth across the participants (Rietveld and Verhoef, 1998, p. 288).

In line with most previous work in the field, this study is only concerned with the issue of economic allocation efficiency and will abstract from political and social feasibility issues. However, as the distributional aspects seem vastly important in the political process for a potential implementation of any allocation scheme, further research might provide a thorough distributional analysis concerning all primary and secondary effects of current and alternative airport capacity allocation.

3.3 Dual Distortion

In oligopoly markets with congestion externalities, market power introduces an additional distortion to the resource allocation. As already indicated in Section 3.2.1, the resulting dual distortion may alter the efficiency results of all above allocation instruments, and may even yield ambiguous or adverse welfare effects from regulation.

3.3.1 Market Power and Congestion Externalities

From an economic perspective, the two market distortions in terms of congestion externalities and market power may be described as follows: The congestion externality generally increases outputs above the social optimum because the congestion costs are only partly accounted for in the cost-benefit rationale of the market participants. By contrast, market power causes the market size to decrease from the social optimum because the profit-maximizing prices are higher and output is lower than under perfect competition. The excess profits for the firms induce a deadweight loss (see Section 3.2.2 above). Moreover, the degree of market power determines the degree to which the congestion costs are internalized, which means that the quantity distortion from higher market power is additionally increased by a higher internalization of congestion (cf. Brueckner, 2002a, p.1359). As a result, the impact of market power and of the congestion externality on overall output are opposed to each other. Hence, in an unregulated oligopoly with externalities and market power, the resulting market distortion may be referred to as the dual distortion.

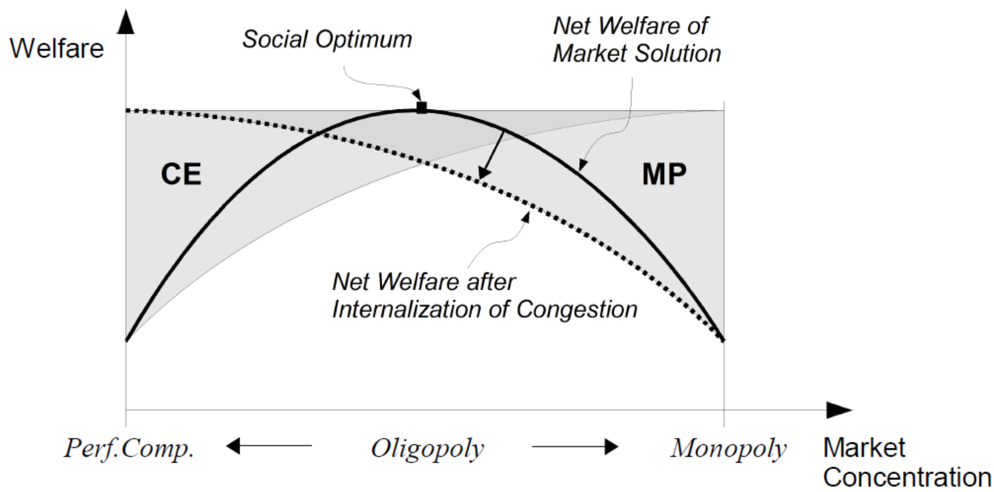

Fig. 3.1: Foundations - Dual Distortion

This dual distortion is depicted in Figure 3.1. The graph represents the net welfare level that can be reached as a function of market concentration.[23] Market concentration is determined by the number of firms in equilibrium: A high number of firms signifies a low market concentration whereas a low number of firms induces a high market concentration. As the importance of market power increases with a decreasing number of firms it is proportionate to market concentration. Both overall output and the congestion externality, by contrast,

[23] This own illustration is based on the properties of Brueckner's results in his (2002a) homogenous products model. The functional forms are stylized and for illustration purposes only.

are inversely related to market concentration because the internalization of an externality decreases along with the market share of a firm.

Net welfare of the unconstrained equilibrium as a function of market concentration (i.e., the number of firms in the market) is depicted by the solid concave line. In this respect, an increasing market concentration signifies that the market power distortion (MP) increases while the congestion externality (CE) decreases. The extreme to the right is a non-price-discriminating monopoly which fully internalizes congestion but implies a large deadweight loss. The other extreme is a perfectly competitive market as shown on the left-hand side: In contrast to a monopoly, perfect competition is not concerned with market power but suffers from excessive congestion because congestion costs are not internalized at all. In the center, the welfare maximum by definition denotes the social optimum. This optimum may occur based on two distinct causalities: Either, a social planner may allocate outputs so as to fully internalize congestion and at the same time remove the deadweight loss that is imposed by the market power distortion; alternatively, this allocation may naturally occur in equilibrium if the two opposing output distortions cancel each other out by coincidence.[24] In all other cases, the resulting market structure will lead to a trade-off between the two distortions at different levels of overall welfare.

As a consequence, regulation policy that aims at internalizing congestion may only yield a welfare improvement as long as its negative output effect does not overturn the efficiency gain of the initial regulation goal itself. In this case, the removal of the congestion externality generally represents a second-best solution to the problem of the dual market distortion. However, if the output effect of the congestion internalization increases the market power distortion by an amount that overcompensates the previous efficiency gain then regulation deteriorates allocation efficiency on an overall level. The dual distortion arising from the two above market failures thus illustrates that the internalization of the congestion externality may either increase allocation efficiency in a second-best manner or induce an adverse welfare effect, depending on the initial size and direction of the dual distortion. The respective outcomes are shown by the dashed line in the graph in above Figure 3.1, which denotes overall welfare when only the market power distortion is present.

This result signifies that an optimal output allocation scheme needs to internalize the congestion externality while at the same time accounting for its own output effect on the market power distortion. The presence of asymmetric network density effects which affect the con-

[24] The third occurrence of a socially optimal allocation is a monopoly with perfect price discrimination, which both fully internalizes congestion and removes the market power distortion and thus the deadweight loss at the same time (cf. Brueckner, 2002a). As justified in Section 23.5, however, perfect price discrimination is not considered in this study.

sumers' utility from flights and depend on the market presence of the networking firms, in turn, should be expected to further complicate the above relationship.

3.3.2 Welfare Caveat of Allocation Instruments

The welfare caveat of the output quantity distortion from market power applies to all three allocation schemes considered in this study: In a quota scheme, the quota allocation may decrease overall output below its socially efficient level if the number quotas is not correctly determined and at the same time the unregulated flight volumes are significantly reduced by the market power distortion. In this case, the relative efficiency gain from reduced congestion may be overturned by the higher output inefficiency (in terms of the deadweight loss). However, even if the number of slots is correctly chosen, an asymmetric market structure may yield that the constrained output becomes inefficiently low even though the dual distortion initially would lead to an excessive overall output. This case is explicitly encountered in the later analysis (see Section 11.4).

In a congestion pricing scheme, allocation efficiency may deteriorate if output is already distorted by market power. Because taxation of the externality decreases output, the congestion tax works in the same direction as the market power distortion. Consequently, the output reduction caused by the congestion tax might reduce output farther away from the social optimum. This result arises if the market power distortion already was large in relation to the congestion externality so that the initial output resulting from the dual distortion had been close to the social optimum. Although the actual output shift may be relatively small in absolute terms because the congestion externality is not important, congestion pricing hence has a negative welfare impact in this case. Note that the initial output may be higher or lower than the efficient output for this welfare caveat to arise. The crucial argument is that offsetting the congestion externality moves output farther away from the optimum than the initial distortion.

In the contrasting case where the congestion externality is important in relation to the market power distortion, however, the welfare effect of the tax is strictly positive: Although the internalizing effect of the tax yields a large output effect, the remaining market power distortion is small. Consequently, the output decreasing effect of the tax shifts the equilibrium closer to the social optimum. This yields a beneficial welfare result, even though overall output becomes inefficiently low because it now undershoots the social optimum. The ambiguity of congestion pricing under market power hence depends on the size of the market power effect relative to the congestion externality: The tax yields a positive welfare result as long as the absolute value of the output deviation after regulation remains smaller than the initial distortion beforehand (cf. Brueckner, 2002a, p.1368).

With secondary trading, market power might constitute a caveat to efficiency in conjunction with strategic behavior, network benefits, and scarcity rents: Network airlines might be willing to afford higher prices for access rights than their smaller competitors, effectuating entry deterrence to increase market concentration (Matthews and Menaz, 2008, p.36). This kind of entry deterrence might be additionally motivated if dominant airlines maintain "hub networks based on network density benefits" (Starkie, 2008a, pp.193), where network density is a competitive advantage that allows carriers to obtain scarcity rents referred to as so-called hub premiums (see Starkie, 1998, p.114 or Starkie, 2008b, pp.171). The competition for airport access may hence include strategic airline behavior as the required "strategic interdependence" is typically reflected by an oligopolistic market structure (cf. Wöckner, 2011, p.3, Carlton and Perloff, 2005, p.350 and cf. Friedman, 1983, p.8). Depending on the profits at stake, this strategic airport demand might either suppress any slot trading at positive prices or increase the competitive position of a dominant airline against its inferior competitors at the expense of social welfare. Secondary trading is thus suspected to preserve incumbent advantages and subsequently preclude a more efficient allocation.

Regarding the above efficiency concern, however, also note that the passengers should be willing to pay a corresponding "hub-premium" if they appreciate the indirect density benefits that arise from network structures (Berry, 1990, p.394). This, in turn, would signify that a low market concentration at a hub airport would not only imply a high degree of market power for the network airline but also higher passenger utility based on these network density benefits. Because these two properties have a counteracting effect on welfare, they were introduced as the dilemma of airport concentration by Starkie (2008a, p.194). The conjunction of network benefits and market power thus introduces an additional ambiguity in the potential impact of a secondary slot trading market solution.

3.3.3 Internalization Debate

In the oligopoly context, the airlines take account of the congestion damage that they impose on their own flight operations (as above see, e.g., Brueckner, 2002a, p.1359). As Daniel (1995) points out, the computation of the congestion tax needs to reflect this internalized portion of congestion; otherwise, the price for flight operations in a congested period exceeds the actual externality. While Daniel (2009a) argues that supply and demand would still lead to flight delays under inferior toll levels, Daniel and Harback (2009) stress that erroneously imposing non-internalizing fees on internalizing airlines may result in worse welfare results than no congestion pricing at all.

As a consequence, the internalization implication is essential for the design of an efficient congestion pricing scheme. In this respect, a number of studies has empirically questioned whether airlines really internalize the flight delays that they cause on their own flight operations: Both Brueckner (2002b) and Mayer and Sinai (2003) find weak empirical evidence for internalization when comparing excess versus minimum travel times between major airports in the US. In a similar sense, Santos and Robin (2010) find a negative empirical correlation between airport concentration and flight delays in European airport data. They validate this correlation as evidence for the partial internalization of each airline's own effect on airport congestion. In contrast, both Harback (2005) and Rupp (2009) report not to find empirical evidence for internalization in real US flight data.[25] Also, Daniel and Harback (2005) and (2008), who consider traffic peaking at single airports, reject the internalization hypothesis. Brueckner and Dender (2008) attempt to unify these contrasting views by arguing that internalization is only important when outputs of different airlines are imperfect substitutes.

The above controversy is referred to as the internalization debate. Considering the above arguments, one may conclude that this debate remains undecided on empirical grounds. At the same time, economic theory suggests that the degree of internalization depends on the market structure: Under perfect competition with inelastic demand, congestion is fully externalized, whereas a monopolist internalizes his flight delays entirely (see, e.g., Brueckner, 2002a). Being a theoretical contribution, this study thus maintains the assumption of the partial internalization of congestion that arises from market power in the COURNOT oligopoly setting.

3.4 Oligopoly Models

The two common oligopoly settings in economic theory are the COURNOT and the BERTRAND models. This section briefly overviews both models and explains why competition with homogeneous goods is not suitable to study market power with network density effects. It concludes that the model in this study should reflect imperfect competition with heterogeneous goods. Note that the notation in this section follows the cited contributions from the literature and thus may deviate from the notation used later in this study's model.

[25] The flight delays within the previous studies have been measured based on excessive travel times between airport. However, both Harback (2005) and Rupp (2009) note that an adequate flight delay measure requires to compare actual versus scheduled arrival times.

3.4.1 Cournot Competition with homogeneous Goods

In oligopoly theory, quantity competition according to COURNOT is typical for a situation with a small number of producers. Firms set their production quantities as a reaction function of each competitor's output and depending on each firm's cost functions. As a consequence, the output quantities may differ between firms. Nevertheless, the joint output determines one single market price according to the inverse demand functions and allows for excess profits (see Vives, 2001, p.93ff.).

Formally, Vives (2001, pp.93) denotes the firms' reaction functions as $R_i(q_1, q_2, ..., q_j)$. These reaction functions determine the firms' individual outputs $q_i(R_i)$ as best replies to the production quantities of their competitors. The best reply functions are typically decreasing and yield a unique NASH-equilibrium (idem, p.96). Consumers are passive with an aggregate inverse demand function $D(p)$, which states the demand for goods as a function of prices. The subsequent price function of outputs $P(Q)$ is an aggregate function of individual quantities $Q = \sum(q_1, q_2, ..., q_i)$, so that total industry output determines one uniform market price identical across all firms. Consequently, the firms' profits are denoted as $\pi_i = P(Q) \cdot q_i - c_i(q_i)$, where the output quantity is assumed to determine the market prices based on inverse demand. Profit maximization with regard to output thus yields the firms' reaction functions and ultimately the individual equilibrium outputs. Individual outputs $q_i \neq q_j$ with $i \neq j$ may differ across firms, so that also profits may be distinct. If aggregate output exceeds a certain value, firms may even take losses, hence $\pi_i \gtreqless 0$ (idem, p.94).[26]

Friedman (1983, p.47ff.) generally characterizes COURNOT competition to reflect firms that have to decide on their production plans in advance, so that the output level would be the firm's strategic variable. Prices, in turn, would later be adjusted according to market conditions "in order to match sales to production". Thus, airline competition with its long-term capital investments and resource planning decisions might be well justified to correspond to this kind of competition.

3.4.2 Bertrand Competition with homogeneous Goods

Under BERTRAND price competition, firms select prices independently. However, with homogeneous products, demand is satisfied by the producer with the lowest price only (Vives, 2001, p.117). Hence, if the marginal costs differ across firms, the more efficient firm serves

[26] In addition, Vives (2001, p.94) denotes the assumptions for strictly downward sloping functions and describes the presuppositions for comparative statics and for firm entry and exit (idem, pp.101 and pp.107).

the entire market. This firm will either set a monopoly price or a price near marginal costs of the less efficient firm, whichever is lower (idem, p.123). If the marginal costs are identical, by contrast, the market will become perfectly competitive and symmetric while profits vanish (Mas-Colell et al., 1995, p.387ff.). Therefore, the BERTRAND model reflects downward sloping demand but does not allow firms to yield mark ups on their competitive prices.

In a BERTRAND duopoly only one unique NASH equilibrium exists where either firm has a positive output: Competition will lead both firms to set their prices $p_i = p_j = c$ equal to uniform marginal costs, so that profits are $\pi = 0$ for both firms (Proposition 12.C.1 in Mas-Colell et al., 1995, p.388). As discussed above, this only holds when marginal costs are constant and identical across firms (Vives, 2001, p.123). In the context of homogeneous goods, therefore, price competition according to a BERTRAND oligopoly does not allow the study of market power at all - except for the corner solution of a single remaining firm that represents a monopoly producer. Moreover, by assumption the firms must supply demand, even if it is higher than the optimal competitive supply at the market price (cf. Vives, 2001, p.117ff.). This property does not suitably reflect a market that is characterized by quantity constraints.[27]

3.4.3 Heterogeneous Goods

Heterogeneous goods yield that prices may differ across the firms both in COURNOT and in BERTRAND settings: In the COURNOT case, the firm's profit function becomes $\pi_i = p_i(q) \cdot q_i - c_i(q_i)$, where p_i denotes the individual price for firm i's good while q denotes the vector of all firms' individual output quantities q_i. Under BERTRAND price competition, the modification is similar but for the fact that firms choose the prices which, in turn, yield the output quantities so that turnover in the profit function can be written as $p_i \cdot q_i(p)$ and p is the price vector of all goods (Vives, 2001, p.149).

Consequently, introducing product differentiation also renders market power to the BERTRAND model (see Mas-Colell et al. 1995, p.395, Vives 2001, p.143ff. or Friedman 1983, p.50ff.). Nevertheless, the BERTRAND case remains less competitive because the firms assume that competitors do not follow their own price adjustments (Vives, 2001, p.155). In comparison, in a COURNOT setting, the degree of competition is reduced, because each firm expects its competitors to reduce their prices along with its own price cuts. The heterogeneous COURNOT

[27] Vives (2001, p.118) assumes plausibility if there are "large costs to turn customers away", as in regulated industries or under consumer protection laws. However, especially with increasing marginal costs, firms may not want to produce up to demand. This case is referred to as the EDGEWORTH problem and is not considered in this study (idem, p.143).

duopoly case thus seems appropriate for a framework that investigates the potentially anti-competitive effects of regulation on welfare.

3.4.4 Implications

As Friedman (1983, p.47) points out, the "outstanding" advantage of the Bertrand model is that the firm is setting the price - as "who else should?" He notes, however, that in reality firms should be imagined to have differentiated products so as to choose both prices and quantities at the same time. In this respect, a firm that simultaneously performs its price and output decisions essentially reflects a Bertrand price chooser. The contrasting case concerns a firm that has to plan its production far in advance. It may be assumed to face high costs of inventory and frequent alterations in price because it aims at matching its sales volume to its preceding production choice. This kind of firms is better represented by Cournot competition, where prices are determined by the market conditions (Friedman, 1983, p.48).

Accordingly, Mas-Colell et al. (1995, p.395) offer an illustrative interpretation of the Cournot model where the latter could be seen as a "proxy for second-stage price competition." This signifies that the quantity choices may be seen as driving quantity competition in the long run, whereas the subsequent price adjustments would reflect price competition in the short run depending on the previously chosen output quantities. Therefore, under the assumption of homogeneous goods and market power, the airline industry is suitably reflected by Cournot competition due to its long-term capacity and resource planning. By contrast, Bertrand price competition with homogeneous goods does not allow prices to divert from marginal costs and implies that output and price decisions are performed simultaneously.

The Cournot oligopoly model with homogeneous good hence represents market power in a way that firms can "set" the market price by choosing their production quantities and hence may obtain positive profits. Subsequently, they can raise their profit margins by increasing their own share on aggregate output (Vives, 2001, p.100).[28] Because products are identical, however, the market price still only depends on the total output and is hence the same for all firms. By definition, thus, firms cannot differentiate their own prices further away from marginal costs in order to obtain higher premiums on the revenue side (Vives, 2001, p.94). Consequently, the Cournot oligopoly model with homogeneous goods represents market power only in the "traditional" sense as characterized by Berry (1990).

[28] Note that also decreasing production costs would increase profits. This study, however, is concerned with the revenue side and therefore abstracts from cost side considerations.

Also with homogeneous goods, firms have a motivation to exercise strategic behavior in order to lower the competitors' outputs. They can do so, e.g., by deterrence of entry or by supporting and exploiting constraints that affect the other firms, such as incumbent-friendly regulation. However, as they do not face the opportunity to gain a competitive advantage with a differentiation strategy that translates into higher price markups, strategic concerns should be limited to increasing the share of traditional market power profits without expanding overall output. In a constrained environment, however, the potential of higher markups might substantially increase the motivation to strategically enhance market dominance by constraining the competitors' resources. Aguirregabiria and Ho (2010, p.1) also support the notion that airline competition is not likely to be of identical prices and provide empirical price differences across airlines when profit functions account for complementarity. As both Cournot and Bertrand competition only allow for one uniform market price across all firms, however, the homogenous setup does not suitably reflect competitive advantages that arise from product differentiation. Therefore, the strategic competition based on differentiated prices and markups is based on product differentiation in this model. As a consequence, heterogeneous Cournot competition is deemed to most closely reflect the presumed market structure revealed in Section 3.1.

Most network studies in the recent literature assume oligopoly markets with homogeneous flights: Both Brueckner (2002b) and Hong and Harker (1992) assume quantity competition in a Cournot duopoly either with two symmetric airlines or with randomly different cost functions between airlines. Consequently, both only account for traditional market power. Czerny (2010) assumes market power between multiple airlines but with inelastic demand. In his case, however, the subject matter of analysis is congestion at the airport level with total flight volume only. A detailed literature review is provided in Section 4.2.

3.4.5 Ambiguities from Market Power

Vives (2001, p.113) notes a "sometimes counter-intuitive relationship" between market concentration and welfare in Cournot oligopoly models because the latter two properties can actually be proportionate rather than opposed to each other. In the Cournot model, a higher market concentration means a lower number of firms, which yields lower competition and higher overall profits (cf. Mas-Colell et al., 1995, pp.393). If welfare is measured by the sum of consumer surplus and firm profits, however, a higher market concentration only implies lower welfare in the case of identical firms with constant marginal costs. Under asymmetric costs or with economies of scale, by contrast, welfare will rise along with market concentration because the firms with a higher productive efficiency can increase their market shares (Vives, 2001, p.101). The same may be true for passenger benefits that increase with the market share of a firm, as illustrated by the dilemma of airport concentration in Section 3.3.2.

This study's network density benefits as a source of product differentiation are expected to exactly reflect this dilemma, which implies that the impact of market concentration and competition on welfare may remain ambiguous in the current model. As Dixit (1996, p.96) confirms, oligopoly models with product differentiation usually do not yield predictive generalized outcomes. He concludes that each case was therefore "best left to be treated *sui generis*". This notion further motivates the study at hand.

3.5 Strategic Competition and Strategic Behavior

The dissociation of strategic competition and strategic behavior constitutes a general difficulty in the determination of government regulation (Carlton and Perloff, 2005, p.378). One viable explanation in the context of this study is therefore provided in the following.

3.5.1 Definitions

Strategic competition may be stipulated as the production decisions of interdependent firms: Firms that take into account the actions and reactions of competitors are generally said to operate strategically. Based on this broad definition, strategic interactions occur in all market forms but for two exceptions: They are absent under perfect competition due to the price-taking property and in monopoly markets due to the natural lack of competitors (Wöckner, 2011, p.3). This coincides with Friedman's (1983, p. 8) notion that the "strategic interdependence" between firms is conceptually reflected in the market form of an oligopoly, and with the concept of product differentiation as one example of strategic quantity competition (Wöckner, 2011, pp.15). An oligopoly setting thus exactly provides the interdependency required for strategic competition.

In contrast to strategic competition, Carlton and Perloff (2005, p.350) define strategic behavior as actions by firms that "influence the market environment" in order to increase profits by reducing competition from "actual and potential rivals." Strategic behavior may thus include more complications than "simply setting prices or quantities" (idem). In this respect, two kinds of strategic behavior exist: The cooperative form aims at raising total industry profits based on collusion across firms. The non-cooperative form aims at enhancing a firm's profits "at the expense of its rivals" and is "more likely to occur in industries with small numbers of buyers and sellers" (OECD, 2015, p.81). Non-cooperative strategic actions include price predation, limit pricing, investments to lower production cost, and raising the rivals' costs (Carlton and Perloff, 2005, pp.351, 367 and 371). Ultimately, most of these strategies aim at entry deterrence (see e.g. Wöckner, 2011, p.36). To succeed with a non-cooperative strategy, Carlton and Perloff (2005, p.352) argue that a competitor must have an advantage over his rivals.[29]

Since this study focuses on competition between firms, only non-cooperative strategic behavior is considered. Cooperative actions such as collusion or cartelization, which serve for the benefit of the whole industry, are ruled out by assumption. However, they also represent an interesting subject for investigation and constitute a topic for further research.

3.5.2 Strategic Airline Behavior

While price predation and limit pricing are not interesting within the scope of this work, both the concepts of investments to lower production cost and entry deterrence exactly capture the airlines' strategic behavior as proposed by this study: First, dominant airlines might engage in bidding for airport access rights in a market-based system to partly or completely deny access by outbidding its rivals in entry deterrence. Second, as a modification of the "lower production cost" argument from a revenue point of view, airlines might invest in airport capacity to increase market concentration and lower competition for the sake of enhancing their own profit margin.[30] In this context, competition for airport capacity might be considered as an investment to increase network premiums (Belleflame and Peitz, 2010) and may even be worthwhile if costly.

In addition, a network airline's rationale for entry deterrence is supported by the supermodularity (i.e., demand complementarity) of its profit function (Aguirregabiria and Ho, 2010, p.8), which implies that network airlines need entire networks of access rights (Langner, 1996, p.15ff.): Established networks are reported to make entry and expansion for competitors difficult, not only based on airport capacity constraints but also on the "exclusionary conduct" of incumbents (Starkie, 2008b, p.25). The OECD (1988, pp.75-76) stated that airport access restrictions had become "a crucial issue for competition" and proposed that regulation had

[29] Established firms supporting government regulation that implies barriers for market entry may thus also be perceived to behave in non-cooperative strategic manner (cf. Carlton and Perloff, 2005, pp.372). This may be the case, e.g., with grandfathering rights.

[30] This reasoning projects the "investments to lower production costs" argument on the revenue side.

to “prevent unfair or exclusionary practices by incumbent airlines with market power.”[31]

Recent studies argue that strategic behavior emerged in the current administrative capacity allocation from an inferior price for capacity use (see e.g. Matthews and Menaz, 2008; De Wit and Burghouwt, 2008). Competition issues are claimed to occur based on “the implied barriers to entry given the allocation and scarcity of take-off or landing slots”. These barriers provided “a substantial wealth transfer to grandfathering airlines” (Gillen and Morrison, 2008, p.173). Specifically, airlines are suspected to “strategically ’hoard’ slots (...) by continuing to use them with the sole purpose of preventing entry” or by having them “’babysat’ by non-competing airlines or smaller aircraft”. Keeping this capacity occupied would allow the hoarding airlines to “still charge monopoly prices” (Gillen and Morrison, 2008, p.176). As Starkie (2008b, p.25) puts it, new entrants find difficulties in competing against established networks of dominant airlines due to their organizational structure, sunk costs, and due to preclusion of entry “on a scale necessary for effective competition” through airport constraints. He also reports barriers to be “often reinforced by the exclusionary conduct of the incumbents.” According to Aguirregabiria and Ho (2010, p.2), entry deterrence is based on the demand complementarity of airlines’ profits in networks. This complementarity implies that “even negative profits might be taken on some routes, if these can generate profits on other routes”. Thus, airlines can simply use “networks as strategy to deter the entry of competitors” (idem, p.1). The above arguments are consistent with general oligopoly theory, where barriers of entry are conceived as “asymmetries that favor the incumbent and allow him to earn a rent” (Vives, 2001, p. 205).

Hence, strategic behavior is suspected to occur exactly as stated above: by entry deterrence and the exploitation of market power for the sake of pricing advantages. This is not surprising, as airport access rights are “key business assets” for airlines that constitute resource heterogeneities in competition (Jäggi, 2000, p. 197). In other words, they represent comparative advantages for competitors at airports and constitute a corresponding strategic position. In the context of this study, the network structure of a dominant airline accounts for such a competitive advantage. Economically speaking, a network allows for vertical product differentiation based on network density effects (Belleflame and Peitz, 2010). Product heterogeneity, in turn, is acknowledged to reduce competition in the market (Wöckner, 2011, p.14). If dominant network airlines compete against smaller point-to-point airlines, this heterogeneity might be considered as a comparative advantage and hence account for market distortions based on strategic behavior.

[31] In addition, Belleflame and Peitz (2010) distinguish between direct and indirect effects of strategic investments on the competitors profits, and conclude that there can be a negative effect on an entrant (i.e., on a competitor) of an incumbent’s investment. The indirect effect is to account for the impact of an ex-post change to an already chosen ex-ante-decision on the competitors profit, and is denoted as strategic effect.

3.5.3 Implications

The above discussion shows the difficulty of a clear-cut dissociation between competitive and strategic behavior: It does not seem evident whether striving for network advantages should be considered as strategic behavior or simply as strategic competition. One might argue, however, that the typology of these effects is not actually important for the choice of regulation policy. The question is rather about how the degree of competition in the market affects the firms and customers: While too strong a regulation policy might deter firms from beneficial price competition, underregulation might lead to excessive market power and anti-competitive strategic actions (Carlton and Perloff, 2005, p.378).

These considerations also point to a general pitfall for regulation: As Gillen and Morrison (2008, p.179) note, regulation policy in practice seems to focus on competition itself rather than on the effects of competition on the stakeholders. Competition authorities would not maximize economic efficiency but rather, would encourage competition. In this sense, also Starkie (2008b, p.23) argues that the aim of the European market deregulation was to impose competition on the national airlines "in their duopoly markets". Consequently, this study focuses on the investigation of the impact of market concentration on customer value rather than on competition. As Carlton and Perloff (2005, p.378) point out, the effects of strategic behavior on social welfare may generally be ambiguous. In line with Dixit's (1996) statement that the outcome of oligopolies is not generally predictable, they conclude that the "welfare implications of strategic behavior" need to be considered "case by case." This notion strongly supports the motivation of this study.

4 Previous Studies

This section presents work from the literature that is considered both relevant and fundamental for the specification of the model and the subsequent investigation of the capacity allocation schemes.

4.1 Overview

The airport capacity allocation problem can be described as an application of regulation theory to transport economics, where the specific case of airport congestion arises from the air traffic management (ATM) literature. This classification attempt is illustrated in Figure 4.1.

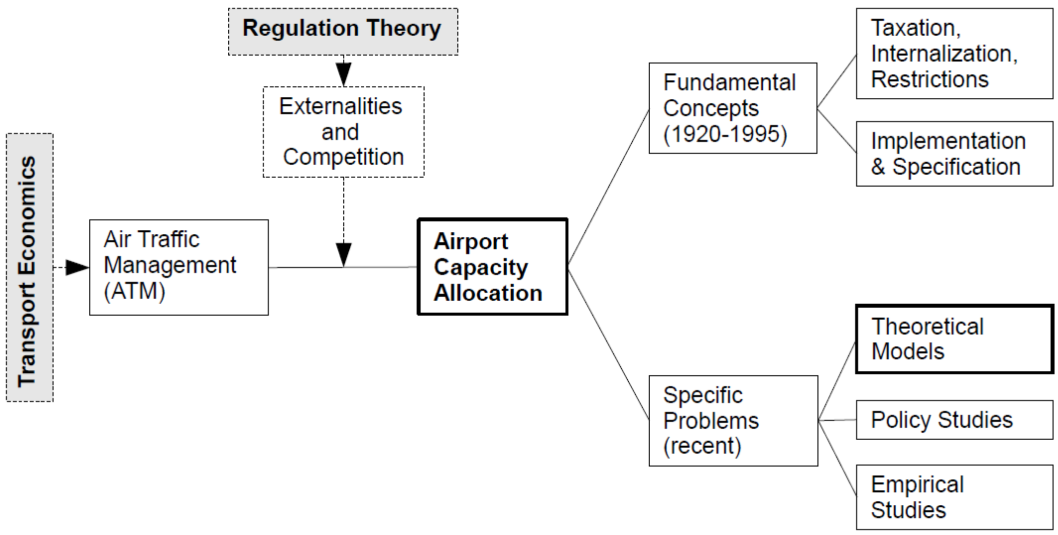

Fig. 4.1: Airport Capacity Allocation in the Literature

The stream of relevant literature can thus be subdivided into two parts: On the one hand, it consists of the fundamental concepts revolving around general capacity allocation issues. These concepts include the airport case but have arisen within the broader context of transport economics and, in particular, from the road pricing literature. On the other hand, it contains the specific problems in respect to airport capacity allocation that have appeared rather recently. While the former general contributions from the literature were outlined in Section 2.6, this section aims at reviewing the latter substream specifically devoted to the airport capacity allocation problem.

For this purpose, the relevant recent work is further categorized into three distinct groups: theoretical models, policy studies, and empirical studies. The first group consists of those studies that base their results on their own theoretical models. This group constitutes the central part of the review as it reflects the previous models applied in the airport capacity allocation problem and is thus essential to this study in terms of context, content and methodology. Note that the studies included in the first group are mostly concerned with the three instruments of secondary trading, congestion pricing and slot auctions while only the former two instruments are considered in this study. Therefore, the review is focused on those two instruments but for the sake of completeness also presents some results concerning slot auctions in an additional subsection. The second set of studies consists of policy papers focusing on the sector-specific evaluation and implementation of the above allocation instruments. Unlike the first group, these studies are not based on theoretical airport models but apply distinct evaluation methods dedicated to investigating and assessing the suitability of different allocation instruments in practice. Finally, the third group represents a small number of studies that provide rare empirical evidence on the welfare effects of congestion pricing and on the historical experience with secondary trading markets. Because the contributions to this group are not numerous the following review combines them together with the studies of the second group into one single subsection.

Next to the above studies, an additional subsection referred to as the model extensions reviews two further topics: the formalization of network density benefits in recent models and the fundamental contributions on vertical differentiation models. These two topics represent the essential foundations for the extension of the theoretical model in this study because they formally warrant those innovations. As they are not directly associated with the current stream of relevant literature they are not shown in Figure 4.1. The last subsection completes the review by describing the gap in the literature that this study aims to exploit.

This survey makes no claim to be complete. Therefore, I also refer to Zhang and Czerny (2012), who offer a comprehensive overview of recent research on the topic, including the analytic foundations where applicable.

4.2 Theoretical Models

The recent theoretical contributions provide partial equilibrium models to qualitatively assess the allocation efficiency of a single or several different allocation schemes. The full set of alternative instruments considered includes slot auctions, slot sales, secondary trading of airport slots, and congestion pricing. This section only presents the results on congestion pricing and secondary trading.[32]

[32] Slot auctions are reviewed in the subsequent section whereas slot sales are usually deemed inefficient and thus are not further addressed.

As expected, some studies replicate the efficiency of these alternative schemes when markets are perfectly competitive. However, others also reveal and investigate the inefficiencies and the technical problems that may arise under imperfect competition with market power. Most studies adopt an economic perspective while a few also consider distributional aspects. This study contributes to the above field of theoretical studies by investigating multiple allocation instruments in a single theoretical airport model.

The two previous studies that most closely relate to this work are Brueckner (2002a) and Verhoef (2010): First and foremost, Brueckner (2002a) considers congestion pricing under different market structures ranging from a monopoly to perfect competition. It is his model that is directly adopted and modified for the purpose of this study and therefore represents both its essential theoretical foundation as well as the main reference for the design and comparison of the congestion pricing scheme. Verhoef (2010), in turn, provides the only study to account for market power while simultaneously comparing both the secondary trading and the congestion pricing instruments. Although his model is distinct from Brueckner (2002a), it provides the basis for the formal framework of the secondary trading scheme and represents a benchmark for the corresponding evaluation.

Further studies of close interest include the basic consideration of a market-based resource allocation by Hong and Harker (1992) as well as the seminal contributions on airport congestion pricing from Daniel (1995) and (2001) as well as their application to the design of an optimal quota system in Sieg (2010) and Daniel (2014). Moreover, they comprise other theoretical models to investigate the welfare effects of congestion pricing such as Brueckner (2002b), Barbot (2005), Zhang and Zhang (2006) and Brueckner and Dender (2008) as well as of a slot reallocation as in Barbot (2004). Finally, this subset is completed by those studies comparing multiple instruments within a single framework with or without market power. They include Brueckner (2009a), Basso and Zhang (2010), and Czerny (2010).

In the following, first the contributions of Brueckner (2002a) and Verhoef (2010) are reviewed due to their central role for this study. The remainder of the above studies are clustered according to their respective topics and then presented in chronological order.

4.2.1 Essential Foundations

Brueckner (2002a) provides the basic single-airport model fundamental for this study. He investigates a congestion tax as an instrument against congestion in a symmetric setting with flights as homogenous products. The main feature of his model is that it considers a variety of market structures and hence accounts for a varying degree of market power: both a perfectly price-discriminating and a non-price-discriminating monopoly, a Cournot and a Bertrand oligopoly with a variable number of firms, and perfect competition.

The central implication of this investigation is that a congestion tax that internalizes congestion can have an adverse welfare effect if the airlines' outputs are already suffering from the output distortion induced by market power. This illustrates the dual distortion, where the congestion externality yields excessive outputs while the market power distortion reduces the outputs. As a result, congestion pricing adversely affects welfare if the total market distortion after imposition of the tax exceeds the initial dual distortion. This generally is the case when market power is important relative to the congestion externality. In this respect, economic theory suggest that the congestion externality increases and market power decreases with the number of firms in the market (Mas-Colell et al., 1995, pp.393). Accordingly, Brueckner (2002a, p.1370) finds that the internalization of congestion is an increasing function of the degree of market power. Therefore, monopolists are presumed to fully internalize congestion whereas in a perfect competition market the congestion costs are assumed to be completely external. As a consequence, Brueckner (2002a) suggests that a congestion tax may only play a role in airport capacity regulation when market power is not important so that the congestion externality prevails. In addition, he finds that full price discrimination in a monopoly is efficient because in addition to the complete internalization of congestion, it does not induce a deadweight loss. This result draws on general economic theory (see Mas-Colell et al., 1995, p.387).

Brueckner (2002a) thus provides a fundamental framework for the investigation of congestion pricing with various degrees of market power and consequently establishes the ambiguity of congestion pricing, which relates to the size of the residual market power effect against the congestion externality. However, this analysis is limited to congestion pricing and abstains from the investigation of other alternative instruments. Moreover, the market consists of symmetric airlines and hence abstracts from product differentiation. Therefore, he mentions the need for the "development of a convincing asymmetric model" that reflects other differences than the cost structure because "a planner would not allow high-cost firms to operate at the social optimum" (idem, p.1368).

Verhoef (2010) models a single congested airport with two airlines in Cournot competition. The two airlines are asymmetric based on heterogeneous costs, which yields a cost-efficient and an inefficient airline. This heterogeneity is referred to as "cost regularity", which refers to the notion of a typical market split between low-cost and full-service airlines (also see Section 10.2.1 on this distinction). While maintaining the market power assumption, Verhoef (2010) investigates the secondary trading for airport quotas as well as a congestion pricing scheme.

Firstly, the study provides a microfoundation for a potential inefficiency of a secondary-trading market with market power that exactly refers to the dilemma of airport concentration: Due to the cost regularity, the productive efficiency in equilibrium can only increase along with an increase in market concentration. The welfare effect of a higher market share of the cost-efficient airline hence remains undetermined because it also implies a higher degree of market power. The cost asymmetry, in turn, only yields two possible outcomes from secondary quota trading: Either the less efficient airline is bought out of the market, so that the more efficient airline becomes a monopolist, or aggregate airport demand is too low to permit slot trading at positive prices. The monopoly yields an ambiguous welfare result and may even become less efficient than no regulation, whereas allocation efficiency is not affected if trading does not take place. Put differently, the welfare effect of secondary trading is indeterminate because it depends on the degree of market power incurred against the higher productive efficiency reached and the congestion externality internalized. Similarly to the result in Brueckner (2002a), the dominance of market power tends to be beneficial if the congestion is large. In the opposing case, the unconstrained market equilibrium is more efficient because it avoids overpricing from the potential monopoly.

Subsequently, Verhoef (2010) specifies the above ambiguity by the consideration of a service obligation. By doing so, he finds that the ambiguity only arises in the absence of a use obligation, so that the efficient airline could strategically hoard its slots after trading in order to achieve its monopoly output. By contrast, if a use obligation were imposed, the efficient airline would expand its output to the socially optimal volume and thus, allocation efficiency would be reached. Although the ambiguity of the single firm market provision can be turned into a net positive effect, however, Verhoef also mentions that a service obligation for a monopolist who can freely set prices would be a “rather theoretical construct” (p.327).

Moreover, Verhoef (2010) computes an optimal congestion tax that includes the second-order effects from market power, which are the source of the inefficiency in Brueckner (2002a). As a result, he finds that the optimal tax for the efficient airline is negative: It constitutes a subsidy for the provision of an optimal output, which represents a compensation for the opportunity costs from abstaining from the output quantity distortion normally inferred by the market power effect. In this respect, he indirectly confirms the welfare caveat arising from market power. Although the allocation becomes efficient, Verhoef (2010) indicates that the tax for the inefficient airline would be sufficiently high as to yield a complete market exclusion. Consequently, however, the tax incidence for the airport authorities would remain zero, so that such a subsidy could not be financed without a loss for the government. These results generally correspond to Brueckner's (2002a) expectations for heterogeneity based on costs (see above).

In summary, both above schemes yield a corner solution where the least efficient airline is driven out of the market. In the case of congestion pricing, this solution is efficient but yields a financing problem. With secondary trading, the welfare result of the exclusive market access is ambiguous but could be turned into a net positive effect by means of a service obligation for the efficient airline. Verhoef (2010) thus presents an airline asymmetry as called for by Brueckner (2002a) and replicates the ambiguity of both above allocation schemes arising from market power. In this respect, his framework, which considers both congestion pricing and secondary trading, is conceptually close to the setting in this study. However, his model is based on cost-side differences only as it "deliberately refrains" from modeling the "counteracting force of frequency benefits" (idem, p.321).

4.2.2 Slot (Re-)Allocation

Hong and Harker (1992) provide a symmetric discrete choice model of a triangular network structure and assume quantity competition in a Cournot duopoly. The airlines are asymmetric based on randomly different cost functions. The model assumes perfect information in terms of all passenger demand and airline cost functions. By estimating the commercial value of the airport access rights based on each airline's respective valuation of market access, the authors find a higher overall allocation efficiency based on higher industry profits if the slots are allocated according to their commercial value as compared to an arbitrary allocation. Because the airlines have different cost functions, these profits are unevenly distributed.

While this result provides a fundamental basis for the subsequent studies, however, Hong and Harker (1992) only compute the airlines' willingness to pay for airport access but do not further investigate any concrete capacity allocation scheme. As a consequence, any strategic considerations that could affect a corresponding trading solution remain out of scope, although the airlines compete in oligopoly. Moreover, the analysis does not account for airport congestion, so that the externality problem and its welfare caveat in conjunction with market power are lacking in the analysis.

From a network perspective, the model reflects transfer passengers on the triangular routes but incorporates neither an actual hub- and spoke network structure that accounts for flight frequency nor any other kinds of network density benefits. Therefore, the authors themselves stress the importance of a better passenger demand estimation and a more realistic network-airline scheduling instead of the stylized patterns and indicate that both specifications might change the above result. In this respect, also the randomly distinct cost functions cannot be considered as a real differentiation of the airlines' business models in response to demand heterogeneity.

Barbot (2004) In contrast to the preceding study, Barbot (2004) finds a welfare-decrease after a market reallocation of airport access rights. This slot reallocation is investigated in a single airport model with a peak- and an off-peak period and multiple airlines, where it is based on a competitive market-based allocation process. As a result of this reallocation, Barbot (2004) reveals that the airlines engage in higher price differentiation, which ultimately lowers welfare. This price differentiation occurs between peak-period and off-peak-period flights.

Although the author states that this result is warranted by the empirical data on price differentiation, she indicates that the welfare result depends on the criteria used for the assessment and might be reversed under some instances. Nevertheless, the result is consistent with economic theory, which suggests that firms may aim to increase their competitive position based on product differentiation. As a consequence, market power rises and welfare is reduced. In this respect, the study is also concerned with the dilemma of hub concentration in a similar sense as this model. In her case, the welfare result from hub concentration is thus adverse.

Sieg (2010) considers the assignment of property rights in the airport slot allocation in a single airport, single airline setting under information asymmetry. While the airport is a profit-maximizing private entity, consumer demand is stochastic and only known to the airline but not to the airport operator. The study then compares two distinct types of slot allocations: a scheme where the quotas are sold to the airlines and a scheme where the quotas are allocated administratively. In this respect, the author designates the former commercial scheme as implying full property rights in the sense of an unconditional ownership, whereas the latter administrative scheme reflects the current grandfathering allocation scheme from practice, which does not formally imply ownership but still infers "quasi-property rights" to the established airline.

As a result, he finds that the airport prefers the grandfathering allocation over slot sales because the use-it-or-lose-it rule implies a higher degree of slot utilization as compared to unconditional ownership. Correspondingly, the higher utilization increases the airport's profits arising from airport charges for aircraft movements. By contrast, when travel demand is low, the higher output causes airline profits to fall and thus allocation efficiency to decrease. As a consequence, Sieg (2010, p.34) stipulates an adverse welfare effect both from the use-it-or-lose-it rule and from airports that are authorized to assign the property rights of airport slots because they would engage in renting the slots as a substitute for aircraft movement fees rather than in selling the quotas to the participating airlines.[33]

Daniel (2014) does not actually evaluate the efficiency of a slot allocation scheme but rather provides a technical contribution concerning the effectiveness of airport quotas to control for congestion. His suggestion arises from the reported problem that the operating windows associated with each particular slot would be excessive. As a consequence, the slots would inhibit the precise "peak spreading" of flights during periods of intense traffic, so that congestion would occur despite the allocation of airport quota (idem, p.19). He thus concludes that the airport delay problems could not be avoided under the current slot system.[34]

By applying a stochastic pricing model to the queuing problem arising with the quota that was originally designed to compute optimal congestion pricing tolls, he thus computes large efficiency gains in favor of a system with quotas that concisely define narrow time intervals for the corresponding aircraft operations. For the allocation of these slots, he generally proposes a slot auction to improve allocation efficiency against the current administrative scheme. However, Daniel also concedes that both markets and auctions would be difficult to implement given the very detailed nature of these precise quotas. Consequently, he suggests that airports should rather turn towards the application of a congestion pricing scheme. This indicates the practical difficulties of complex allocation schemes that are theoretically efficient.

4.2.3 Congestion Pricing only

Daniel (1995) first of all presents a perfect competition model of a single airport and shows significant welfare improvements for an airport congestion pricing scheme as compared to administratively allocated airport slots, which is mainly based on "intertemporal traffic adjustments" (idem, p.366). In particular, he provides a sophisticated stochastic bottleneck model that differs from the usual partial equilibrium models used in the other theoretical

[33] Note that this result contradicts with the usual notion about low output inefficiencies, where an extrinsic increase of outputs decreases the firm's profits but increases social welfare. This contrast also arises in the later discussion of secondary trading in Section (14.2.1).

[34] Daniel (2009a) precedes this study and thus yields the same results.

studies. In this respect, Daniel's (1995) contribution includes a detailed methodical section on how to compute the external costs at the airport under various conditions in a stylized hub-and-spoke network schedule situation.

His distributional consideration concludes that large airlines generally profit from a toll system whereas small airlines are expelled from the valuable peak-periods at congested airports. Moreover, the efficiency result is put into perspective by indicating that, unlike the theoretical basics of congestion pricing, the monetary value of the congestion externality would not be well known yet. In addition, Daniel (1995, p.357) stresses that the interactions within airline networks must be expected to substantially complicate the effects from airport charges on airport demand, which also concerns the allocation problem in a system of multiple network airports (idem, p.366).

Daniel (2001) extends the above model and computationally simulates the effect of a congestion pricing scheme on the equilibrium. The simulation is calibrated based on data from a major US airport (Minneapolis St. Paul). As a result, the model reveals welfare gains for most stakeholders such as major airlines, passengers and the airport while it finds losses for smaller airlines and private air traffic. While this outcome may have been expected based on the congestion pricing rule, the quantitative evaluation also reveals that the overall welfare improvement from congestion pricing even allows the financing of airport programs, yielding welfare gains for all stakeholders.

Brueckner (2002b) validates Brueckner's (2002a) results in a stylized network with four airports in a Cournot duopoly. For this purpose, it computes the efficient amount for a congestion tax while also considering the market price as an inverse function of demand. The idea of market power is justified by the notion that the market values are specific for each route, so that the competition level does not need to be equal across the entire network. More precisely, some routes feature perfect competition while others are monopoly markets (idem, p.10). This virtually replicates the idea of competition with cross-subsidies within a network as captured by Aguirregabiria and Ho's (2010) particular airline profit function for hub-and-spoke networks (see Section 4.4.1 below).

In contrast to Brueckner (2002a), the network model reveals that the internalization of congestion in a network is not a function of market power but of the respective airlines' market share at an airport. The disclosure of this result owes to the network structure, which enables the model to depict multiple markets with different degrees of market power simultaneously. Unlike in the single airport setting, therefore, the effect of market power can be considered while leaving the market power distortion aside. Although this analysis is also limited to the symmetric airline perspective and to congestion pricing, it yields a valuable contribution for future studies of airport capacity allocation across an airport network. However, Brueckner (2002b) also admittedly abstracts from a microfoundation of network effects such as "economies of traffic density", connecting passengers, flight frequencies and the typical "traffic patterns at hub airports" (idem, pp. 4 and 10).

Barbot (2005) is the only setting mentioned here to also provide a vertical differentiation framework. As common in the vertical differentiation literature, the model represents a quality choice game with two stages where the airlines first choose their product qualities and thereafter decide on their outputs. As a result, there is a low-quality and a high-quality firm. The quality distinction between the two firms, however, is not further specified.[35]

In this setting, Barbot finds that congestion charges decrease welfare if compared against the usual weight-based operating fees at airports: Although the airport and the low-quality airline enjoy a higher profitability under the tax-based scheme, the forgone profits of the high-quality airline overturn these gains, so that an overall welfare loss results. The author points out that this outcome contrasts with the previous results on congestion pricing and attributes this difference to both her particular model characteristics and the peculiar airport case. Because this setting distinguishes itself from the traditional market power models applying homogenous Cournot quantity competition, it provides an important contribution for future reference.

Zhang and Zhang (2006) consider airport congestion charges from a perspective that is distinct from most other studies: They investigate an airline oligopoly with market power where the welfare effect of the airport's regulation policy is the subject of study. Their rationale reflects that the congestion internalization of airlines with market power will lead to lower congestion taxes, so that the airport may no longer be able to finance its capacity investments based on this tax. Although the capacity allocation turns out to be efficient, from a broader view the system is not sustainable: Namely, as a result, Zhang and Zhang (2006) find that a public airport that maximizes social welfare will require government subsidies for its capital investments. Under other regimes such as airport privatization or budget constraints, by contrast, the airport will engage in capacity overinvestment. As a consequence, the marginal congestion costs fall short of the marginal capacity cost. Although this study is concerned with the efficiency of public infrastructure rather than air traffic operations, it

[35] Also see Section 4.4.2 on the vertical differentiation literature.

provides an important contribution with regard to the effect of capacity allocation schemes on airport financing.

Brueckner and Dender (2008) also borrow Brueckner's (2002a) model while enhancing it to reflect an asymmetric airline structure. The asymmetry stems from the implementation of a Stackelberg leadership with one dominant and several atomistic firms both in a Cournot oligopoly and in perfect competition. The purpose of the model is to present the optimal taxes for a congestion pricing scheme in various asymmetric market types with a varying degree of market power. As a result, the model confirms the findings of both Daniel (1995) and Brueckner (2002a), as the congestion tolls include compensatory elements for the different degrees of market power for both internalizing and non-internalizing airlines.

Although the Stackelberg setting considers asymmetric airlines, the asymmetry results from the assumption about the specific market form but not from a corresponding micro-foundation of the demand structure. As a consequence, the equilibrium is only asymmetric in flight volumes but not in prices or in the degree of market power. The model hence reflects neither endogenous product differentiation nor network effects and thus, with respect to the model properties, may not be considered as a predecessor of this study.

4.2.4 Secondary Trading and Congestion Pricing

Brueckner (2009a) modifies the original Brueckner (2002a) setup by assuming that the single airport is always congested. This signifies that the peak period by definition serves the entire market, which in a vertical setup would correspond to full market coverage. In further contrast to the (2002a) model the airport is used by two airlines that serve two completely separate markets. Additionally, the market structure is also a Cournot duopoly but market power is neutralized by perfectly elastic demand. Consequently, the airlines cannot yield economic rents, so that the output distortion based on market power completely vanishes. Moreover, the airline costs are assumed to be convex in order to reflect diseconomies of scale.

As a result, Brueckner (2009a) finds that both slot sales at uniform prices and a first-come, first-served capacity allocation were inefficient due to the negligence of congestion internalization and thus too high traffic volumes. In correspondence to the horizontal demand specification, however, both congestion pricing as well as secondary trading (and also slot auctions) are found to be fully efficient.[36]

[36] Brueckner (2008) precedes this study and replicates the same results.

Note that this airline asymmetry allows the airlines to charge different prices in the two markets. However, because the markets are separated and the model thus abstracts from cost and demand side heterogeneities, this setting does not actually reflect competition with differentiated products in an asymmetric market structure.

Basso and Zhang (2010) use Brueckner's (2009a) model with asymmetric airlines based on market separation and additionally introduce the airport authority as a profit-maximizing stakeholder. They suggest that this modification is both necessary to „generate sensible results“ with horizontal demand and also plausible in the case where airlines are concerned with congestion.

The investigation of both secondary trading and congestion pricing consequently proves efficiency for both allocation schemes in accordance with Brueckner (2009a). Considering the airport also as a stakeholder, however, they also find that the solutions yield distinct revenues for the airport. As a consequence, the efficiency equivalence of the above scheme vanishes when the airport chooses the scheme that maximizes its profits. In this case, the airport would choose to auction the airport slots, which, in turn, would yield a higher traffic volume and higher congestion. This rationale is similar to that in Sieg (2010), where conditional property rights in airport slots increase the overall traffic volume and thus are beneficial for the airport but detrimental for the airlines and social welfare.[37]

Czerny (2010) investigates the welfare benefits of congestion pricing to a market-based slot allocation in terms of a slot auction both for a single airport as well as in an airport network. The market is perfectly competitive, so that airlines do not enjoy market power and thus the analysis may focus on the congestion externality and the network property of the market. However, arguing that resource management under uncertainty were extensively analyzed in the field of environmental economics but that these results were of limited use for the management of runway capacities due to the interdependent demand between different airports, he applies stochastic shocks both to airport demand and to marginal congestion costs in order to account for uncertainty regarding the social benefits and costs of airport operations (idem, p.372).

In the investigation, the model gradually extends the single airport setting to a symmetric network of two congested and one uncongested airport, and eventually applies non-linearity

[37] Note that Sieg (2010) contrasts with this case because his airport prefers a grandfathering allocation over selling unconditional property rights. The welfare outcome, however, is ultimately the same (see Section 4.2.2 above).

and uncertainty to congestion costs. The main finding is that congestion pricing is superior to a slot allocation in all cases considered.[38] A more general rule expresses that an airport slot auction yields a higher allocation efficiency in the airport network whereas a congestion pricing scheme is more beneficial when congestion costs are affected by uncertainty.[39] Although the model investigates an airport network, however, it abstracts from specific network effects in similar manner to Brueckner (2002b).

4.2.5 Slot Auctions

Slot auctions have been proposed, e.g., by Rassenti et al. (1982), Gale (1994) or Button (2007) and (2008) while criticism thereof has been expressed by Borenstein (1988), Jones et al. (1993), Sentance (2003) and Daniel (2014), among others.

Rassenti et al. (1982, p.402) consider that economic efficiency requires distributing the airport access rights according to the airlines' airport demand, so that the willingness to pay ultimately determines the airport capacity allocation. Hence, they present a combinatorial auction system for the allocation of slot packages that are useful for schedule optimization. Gale (1994, p.25) shows a slot auction in a symmetric duopoly and finds that the allocation is asymmetric but not monopolistic. More generally, Hong and Harker (1992, p.321) find that an endogenous allocation of access rights would distribute the airport slots to the airlines according to their commercial value, so that the total profits of the airline industry would increase. In this respect, they address the corresponding "economic advantage" of a slot auction. Also Button (2007) and (2008) argue that an appropriate kind of a slot auction scheme would optimally allocate the scarce airport capacity and thus "significantly" ameliorate the efficiency of resource allocation (Button, 2008, p.578).

However, Borenstein (1988, p.358) stresses that a "competitive market allocation" by auctioning, selling or re-sales of airport slots would not generally need to ensure allocation efficiency in the presence of output quantity distortion based on market power. He reports that the relation between firms' profits and total surplus would depend on the actual market structure, so that there might be an "extreme divergence between the private and social value" from operations even in the absence of cooperation among firms. In this respect, the auctioning of

[38] In the (rare) deterministic cases of a single airport with non-linear congestion costs and a network with linear congestion costs and steep marginal costs, the slot system may supersede the congestion pricing scheme.

[39] Very similar settings are found in Czerny (2006) and (2007), which explore uncertain airport demand and stochastic quadratic marginal external cost and thus precede Czerny (2010). While Czerny (2007) shows the same results, Czerny (2006), however, finds that independent demand and perfect information yield optimal welfare results for both instruments. In addition, Czerny (2006) reveals that slot auctions are favored over congestion pricing when airport demand is complementary whereas the opposite holds for monopolistic isolated airports.

airport slots were prone to large inefficiencies by excessive outputs on frequent connections and an underprovision of flight services in oligopoly markets (idem, p.375). Moreover, Jones et al. (1993) investigate different kinds of slot auctions. They find that a simple auction of airport slots as single entities could never possibly yield the chance for airlines to build stable and useful timetables. For a combinatorial auction concerning entire packages of slots, they note that no market would emerge due to the valuation diversity of the distinct slots within one package. The conclusion made is that the demand externality caused "simple market solutions" to be insufficient for capacity allocation at major airports, so that a coordinator would still remain essential (idem, p.48). Despite this efficient result for a commercial quota allocation, also Hong and Harker (1992, p.321) and, more recently, Daniel (2014, p.24) address the complex problem of implementing the complementary and interdependent airport demand in a practicable auctioning mechanism.

In addition, Sentance (2003) compares an initial slot auction to the current administrative regime and a secondary trading solution. His result is that slot auctions in the first place would compromise the current administrative quota allocation for the sole purpose of accounting for environmental concerns and thus see no "compelling case for developing slot auctions". They hence conclude that slot auctions would be associated with few benefits but large costs, so that a secondary trading solution would yield even better efficiency result while representing a far more practicable solution (idem, p.53 &56).

Lastly, some of the studies about secondary trading and congestion pricing schemes considered in previous Section 4.2.4 also include slot auctions. The results are similar to those above. For perfectly competitive markets, slot auctions yield efficiency gains that are equivalent to the two alternative instruments previously mentioned: e.g., Basso and Zhang (2010) find that in equivalence to secondary trading and congestion pricing, slot auctions could lead to the social optimum. Moreover, they show that an auction can even turn out to be superior if the airport has market power but its profitability is not excessive. Also Brueckner (2009a) replicates this efficiency result for all three above schemes. Unfortunately, the market power studies do not reveal evidence on an initial quota allocation based on auctions. However, the corresponding welfare caveat was already pointed out by Borenstein (1988), so that the inefficiencies concerning secondary trading and congestion pricing from the dual distortion as presented in the economic foundations in Section 3.3.2 should also be generally applicable to this type of allocation scheme as well.

4.3 Policy Studies and Empirical Evidence

The following policy papers focus on the analysis of advantages, disadvantages, and potential threats of the different allocation instruments on descriptive grounds, i.e., abstracting from proper models. Apart from the usual efficiency concerns, they also consider distributional aspects regarding the potential implementation of alternative schemes from a policy perspective. In addition, a number of contributions provides case studies for specific regions or airports.

The general concerns relating to the current administrative quota allocation are comprehensively illustrated by Matthews and Menaz (2003) and (2008), to mention only a few but important studies. NERA (2004), NEXTOR (2004), Whalen et al. (2007) and Berardino (2009) provide evaluations of alternative allocation instruments aimed at European and US policymakers. Odoni (2001) illuminates the distributional aspects of congestion pricing, along with a number of issues concerning the implementation into a practicable and equitable system. Gruyer and Lenoir (2003) describe a combinatorial auction suitable for the initial allocation and an appropriate secondary trading scheme for a subsequent market reallocation of airport slots.

In their comprehensive reviews, Madas and Zografos (2006), (2008), and (2010) present different sophisticated evaluation schemes for the assessment of optimal policies, thereby providing a methodological contribution. By applying these methods to the knowledge base from recent literature, they assess and illustrate the advantages and disadvantages of different alternative instruments for distinct generic airport settings and stakeholders.

From a practical perspective, a multitude of single case studies investigates the perspectives for secondary trading at different specific airports. These include as Mehndiratta and Kiefer (2003) for San Francisco International Airport, Tether and Metcalfe (2003) for London Heathrow, MacDonald (2006) for smaller US community airports, De Wit and Burghouwt (2007) for Amsterdam Schiphol and Ball et al. (2007) for New York La Guardia, which was one of the few slot constrained airports in the US region at that time.

Despite the vast amount of policy studies, however, Fukui (2010) presents empirical evidence from actual implementations of market approaches and shows that the potential for slot trading has been limited as of to date. Based on expert interviews with three national airport coordinators from Europe, Noto and Laesser (2013) therefore review both the current administrative scheme as well as the propositions for alternative allocation instruments and assess their perspectives in practice. By doing so, they recapitulate the economic efficiency arguments from the capacity allocation discussion while also indicating the advantages of the administrative quota allocation, which arise from the coordinators' experience and expertise concerning the actual heterogeneous, asymmetric market structure.

Nevertheless, Santos and Robin (2010) find that the flight delays at European airports are still significant. Moreover, Bel and Fageda (2009) consider the effects of flight delays on the airlines' network choices and their consequential impact on hub concentration. They find that hub concentration positively correlates with flight delays. In addition, Bel and Fageda (2010) investigate the relationship between airport privatization and airport regulation. Based on empirical data, they find that a more detailed regulation is more likely to arise under private ownership conditions than at public airports. The above findings hence indicate that the topic of scarce airport capacity allocation continues requiring a thorough discussion and further research.

4.4 Foundations for Model Extension

4.4.1 Network Effects and Density Benefits

Network density benefits in the airline context constitute the central topic of this study. Economic theory has described and formalized such network effects in a general sense, which applies both to cost-side network economies as well as to the utility from network goods (see, e.g., Belleflame and Peitz, 2010). Moreover, in the applied literature, the airlines' struggle for network advantages and structures is an established result (see Section 3.1.2). Surprisingly, however, the benefits from network density have rarely been captured in the theoretical models in the airport capacity allocation context. In this respect, e.g., Brueckner and Zhang (2001), Aguirregabiria and Ho (2010), Czerny (2010) and Fageda and Flores-Fillol (2013) have formalized network effects on utility and welfare but either do not reflect airline asymmetries or do not consider airport capacity allocation. Nevertheless, these studies provide valuable formal support for the justification of this study's model design.

Brueckner and Zhang (2001) explicitly model passenger benefits from flight frequency within a network structure. In that way, their setting associates to the model at hand. Their investigation, however, considers a single monopolistic airline's problem of network type and design in a flight-fare versus frequency context but does not engage in the capacity allocation discussion. Nevertheless, Brueckner and Zhang (2001) provide important support concerning the framework of this study by presenting a proposition for the formalization of network density benefits.

By contrast, Czerny (2010) assesses capacity allocation issues by introducing indirect flight benefits based on airline flight frequency (for a summary, see Section 4.2). These benefits, however, only depend on the overall flight volume at an airport and are not airline specific. This means that Czerny (2010) introduces actual network density benefits in the airport capacity discussion but does not consider an asymmetric airline structure that includes a comparative network advantage. Consequently, the model provides a solid justification for network density based on flight volume as indirect utility but does not replicate a heterogeneous market structure such as the one suggested by this study for a central hub airport with a dominant airline.

Aguirregabiria and Ho (2010) provide a network airline profit function that reflects the complementarity of market entry and exit decisions across different routes to investigate a route-structure optimization problem from the perspective of a single airline. They provide a so-called supermodular profit function that enables the model to "incorporate the entry deterrence motive" for networks (idem, p.2). This work also provides some most valuable, formalized theoretical support for the premise underlying the model at hand but ultimately does not investigate the problem of airport capacity allocation itself.

Fageda and Flores-Fillol (2013) consider airline network structures and congestion. They find that a hub-and-spoke network structure could adversely affect allocation efficiency based on airport congestion. More precisely, their rationale suggests that congestion costs cause the network airlines to reduce flight frequency at their hubs and to return to a more direct-type route structure. However, also in their model the network is also supposed to benefit both the customers as well as the airlines. As a consequence, the decreasing hub dominance of networks decreases welfare. Although this study also abstracts from the consideration of actual capacity allocation, its essential contribution to the study at hand consists in the passengers' utility function that explicitly values flight frequency. Consequently, the airlines can commercialize their network advantages. Fageda and Flores-Fillol (2013) thus provide great support for the network density benefits idea. Their explicit consideration of flight frequency, however, makes us also aware of the very basic specification of network density applied in the model at hand.[40]

Overall, the above studies all provide valuable foundations for the essential network density benefit argument. However, the frameworks from Brueckner and Zhang (2001), Aguirregabiria and Ho (2010) and Fageda and Flores-Fillol (2013) consider optimization problems for single airlines and thus abstract from the actual airport capacity allocation. By contrast, Czerny (2010) discusses the airport capacity allocation in light of network density benefits but

[40] In this respect, see the corresponding model limitation in Section 23.4.

his symmetric market structure does not account for an (asymmetric) comparative advantage of a network airline against its competitors.

Finally, it is worth mentioning that both Brueckner (2002b) and Hong and Harker (1992) extend their investigations to airport networks. As mentioned in Section 4.2, however, both admittedly abstract from actual network effects and thus only consider the direct utility from flights. Consequently, their studies include neither cost-side advantages nor indirect passenger benefits from network operations.

4.4.2 Vertical Product Differentiation

The vertical product differentiation literature mainly draws on the seminal contributions of Shaked and Sutton (1982) and Gabszewicz and Thisse (1986). While Shaked and Sutton (1982) show that competitors will vertically differentiate their products in order to yield positive profits in price competition, Gabszewicz and Thisse (1986) are concerned with the stability of equilibria by comparing quality competition in the vertical differentiation model to price competition under horizontal differentiation based on consumer taste (due to Hotelling, 1929).

Most studies commonly apply Tirole's (1988) standard setting, which draws on the above foundations and is also used in a modified form for the parametric specification of the generic model in this study (see Section 15.2).[41] This model involves two firms that compete in a two-stage game where they first choose their product qualities and then resolve their production quantities. Because these decisions are interdependent in a similar manner to the reaction functions in the Cournot model, which determine the firm's contingent output choices, the model is solved through backward induction.

In Tirole (1988), the only Nash equilibrium consists of the corner solution where one firm chooses the lowest possible quality while its competitor chooses the highest possible quality. This result ultimately follows from the assumption that the market is covered. Choi and Shin (1992) extend this result by showing that the equilibrium becomes an interior yet still fixed solution if the market is assumed not to be covered.[42] In turn, Wauthy (1996) shows that the degree of market coverage depends on the quality choices of the firms in equilibrium and thus is endogenous to the model if market coverage is not determined by a specific assumption.

Recent contributions relying on analytic solutions include Ecchia and Lambertini (2006) and Lambertini (2006). The former provide a model with convex quality costs where either

[41] By contrast, Brueckner's (2002a) model is based on the horizontal specification.

[42] In Choi and Shin (1992) the quality levels amount to 2/7 and 4/7 of the highest possible quality.

the chosen quality levels are fixed proportions of maximum quality and costs and yield an interior solution, or there is no duopoly equilibrium because the low-quality firm exits the market. The existence of the equilibrium depends on the absolute magnitude of the highest achievable quality level. By contrast, the latter consider quality improvements based on capital investments and thus abstract from variable costs. Similarly, he also finds either a monopoly or a duopoly with complete or full market coverage depending on the cost of capital in relation to the population wealth. His specification hence provides the foundations for the model calibration in Section 15.3.2.

Finally, Motta (1993) provides a general framework used to investigate both price and quantity competition in the second-stage while controlling either for fixed or variable quality costs in the first stage. However, due to the increasing computational complexity, the model yields polynomial terms for the equilibrium conditions and hence relies on numeric solutions. Because the results can be generalized, this setting constitutes an important contribution for the development of a vertical differentiation model with quality choice as a strategic variable and externalities that ultimately yield convex costs and thus require computational methods (see the corresponding proposition in Section 23.4).

4.5 Conclusion: Gap in Literature

As this literature review shows, both the efficiency of the proposed alternative instruments under perfect competition as well as the welfare caveats in conjunction with market power have been addressed by recent theoretical studies: In perfect competition settings, Daniel (1995) and (2001) show large efficiency gains for congestion pricing whereas, e.g., Brueckner (2008) and (2009a) as well as Basso and Zhang (2010) find first-best allocation efficiency for both secondary trading and congestion pricing schemes in imperfect competition with inelastic demand, as long as strategic airline behavior is ruled out. By contrast, mainly Brueckner (2002a) and Verhoef (2010) show the welfare caveats arising from both above allocation schemes based on market power in conjunction with congestion externalities and, if applicable, the associated strategic hoarding of airport quotas.

However, Langner (1996, pp.15) already criticizes that alternative allocation instruments do not account for the "network characteristics of flight services". This notion is strongly supported by Aguirregabiria and Ho (2010, p.1), who argue that the effect of airline networks in the context of entry deterrence has to date been neglected. In the same sense, also Brueckner (2002a) suggests investigating airline asymmetries for their impact on congestion pricing and efficiency. Most prominently, he points to the requirement of a "richer framework where passenger valuation of flight frequency is explicitly considered". In doing so, he judges "cost differences across firms" as not being a "useful source of asymmetry" because high-cost firms would not operate at the social optimum (idem, p.1368). This hence underlines that the issues of competitive advantages and product differentiation should not be neglected in the investigation of optimal airport capacity allocation at large network-hub airports because they may both yield or increase market power.

In this respect, a number of studies have addressed the need to consider the comparative advantage arising from network density benefits at a network hub airport in the debate on airport capacity allocation. Most importantly, Starkie's (2008a) "dilemma of airport concentration" arising from the correlation of market concentration and passenger benefits indicates that this topic has been discussed intensively. However, the above literature review shows that the asymmetric market structure of a hub airport with a dominant, networking airline has not yet been formalized in the recent theoretical airport capacity models investigating allocation efficiency. As a consequence, the study at hand differs from recent theoretical work in the literature and, most notably, from its invaluable fundamentals consisting of Brueckner (2002a) and Verhoef (2010). It constitutes an independent, unique contribution featuring the innovation of an exogenous airline asymmetry based on network density effects.

Part I. Foundation: Generic Model

„Both simple cost-reducing and naive market power stories are inappropriate for the airline industry." (Berry, 1990, p.394)

5 Generic Model

This section presents the generic model that is used for the investigation of the three different capacity allocation schemes in Part II. For this purpose, first the causalities underlying the model design are explained. Thereafter, the formal setting is presented in mathematical terms. Lastly, the two main assumptions for the later analysis are shown.

5.1 Causal Relations

The causalities implemented in this model can be shown in a *circular flow diagram* (cf. e.g. Backhouse and Giraud, 2010). This flow diagram depicts the relevant economic interactions between passengers and airlines and is shown in Figure 5.1.[43] The respective causal relations draw on common notions from the literature and on this study's proposition concerning the impact of network density benefits on supply and demand.

5.1.1 Demand and Supply

The basic relations between supply and demand within the model are standard in theoretical partial equilibrium models. They are depicted in solid lines and are described as follows:

The fundamental driver of the airlines' airport capacity demand is the market for air transportation. In this market, the *customer value function* determines the passengers' air travel demand as a function of the travel price, the related travel benefits and travel costs. Subsequently, the matching of market demand and supply determines outputs (i.e. the number of flights). Assuming an oligopoly market structure implies that the airlines decide on their profit maximizing outputs based on inverted travel demand. These demand functions thus

[43] Backhouse and Giraud (2010) describe some generic forms of circular flow diagrams and state that the latter have been traditionally used in economics to illustrate the economic relations between households and firms (i.e., income and production streams).

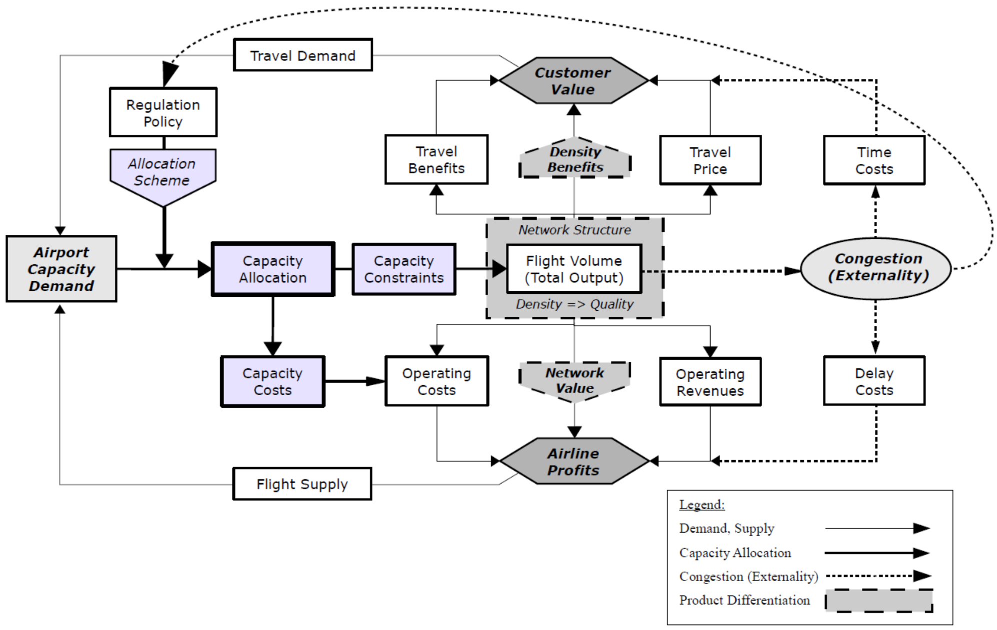

Fig. 5.1: Causal Relations in the Generic Model

yield air travel demand as an endogenous function of flight fares and ultimately allow airlines to determine equilibrium prices along with their outputs.

In particular, the corresponding passenger utility from air travel includes two distinct types of benefits: On the one hand, the direct travel benefits denote direct utility arising from transportation itself. This type of benefits occurs in most recent theoretical models (such as in Brueckner, 2002b or Czerny, 2010). On the other hand, however, there are also indirect benefits from an airline's network density. As further explained below in Section 5.2, these indirect benefits abstractly reflect utility that arises from product quality that arises from a connective network as compared to a multitude of single, non-coordinated point-to-point flights. One may think of benefits such as higher travel flexibility, lower schedule delays or a wider travel choice as they normally arise from an optimized network structure.

5.1.2 Network Density Benefits

The model's unique contribution is shown within the **dashed, gray boxes** and the corresponding **dashed lines:** Network density benefits as a source of product quality, which affects airline profits and customer value. These network effects are envisioned as providing additional, indirect travel benefits for passengers arising from the networking characteristics of flight services and their associated services. One may think, e.g. of a wide destination choice, high flight frequencies on a route, or schedule optimization and thus low schedule delays. In conjunction with flight re-booking possibilities and the associated ground services that assist in case of changes in travel plans or irregularities, these properties arguably provide customer value to passengers, in addition to the mere transportation services of air travel. And as these network density effects enter the passengers' utility function, they also have an impact on the network airline's profit maximizing rationale and thus on airport capacity demand.

In this model, however, the above network characteristics are not explicitly modeled. More specifically, they are supposed to be implicitly based on the density of an airline's network. Consequently, they are abstractly reflected as indirect travel benefits that are based on an airline's network density. In sum, this network density again abstractly constitutes a product quality that can endogenously be chosen by the airline - at least to a certain extent. Subsequently, this variable product quality affects demand and flight fares: It enables the network airline to collect a network premium based on the corresponding customer value.

Generally, the economic literature usually relates network economies to cost-side benefits of network operations. These are presumed to accrue, e.g., from economies of scale based on the concentration of passenger streams and the subsequent application of larger aircraft (see Section 3.1.1). In contrast to this general connotation, however, it is important to note that the network effects in this study relate to passenger benefits based on network services rather than on the cost-side economies of network operations: As will be specified in Section 5.2, the aircraft size is taken as invariable, so that only the number of all flights operated in a network is considered. Consequently, the network effects reflect increasing returns for passengers on the demand side, rather than scale economies of network density for airlines on the supply side.

5.1.3 Product Differentiation

Product differentiation arises in that the opportunity to create network structures is reserved for one airline only. This means that the other airline is limited to offering point-to-point services. As a consequence, network benefits can only be created by the networking airline. Correspondingly, the commercial network value also accrues only to that airline. The non-networking competitor is bound to remain with direct benefits from air transportation only. Put differently, the network effects provide the foundation for product differentiation based on an exogenous airline asymmetry. As a consequence, flights become heterogeneous goods so that airlines can also differentiate their prices and thus enhance their profits. Ultimately, this heterogeneity generally decreases competition in the market, because the network airline can increase its flight fares versus the uniform market price that would prevail if flights were homogeneous goods.

This distinction is not directly visible in the graph in Figure 5.1 because the latter only depicts the relations of the networking airline. Nonetheless, it can easily be retraced by leaving out the dotted gray fields and the associated relations for the non-networking competitor. This shows that network benefits are absent both in the non-networking competitors profits as well as for its customers. Consequently, the network effects affect both demand and supply, and hence translate into an asymmetric airport capacity demand and into the corresponding airport capacity allocation. Ultimately, the airline asymmetry will hence affect the first-order conditions for allocation efficiency. It thus has an impact on the welfare effects of the different capacity allocation instruments.

5.1.4 Congestion

Congestion arises from the peak-period output of the airlines and denotes the excessive usage of airport capacity. As a consequence, airport demand exceeds supply so that flights operated during the peak period experience flight delays. These delays, in turn, cause congestion costs to both airlines. Moreover, passengers are directly affected by the time costs caused by the flight delays arising from airport congestion. Hence, in addition to production costs and travel benefits, the above profit and utility maximizing rationales of airlines and passengers are also affected by congestion. The effects of congestion are represented with **dotted lines** in Figure 5.1.

Most importantly, airport congestion provides the actual rationale for airport capacity regulation. This is explained as follows: On the one hand, each airline takes account of the impact of congestion on its own operations.[44] On the other hand, however, every airline also contributes to the congestion affecting its competitor. Therefore, at least part of the airport congestion constitutes a negative externality and thus represents a market distortion. As a consequence, resource allocation efficiency is diminished. This, in turn, justifies applying airport capacity allocation schemes as a regulation policy in order to improve the inefficiencies arising from external congestion (see Section 2.3).

[44] The question whether airlines really internalize the proportion of delay imposed on themselves is discussed in Section 3.3.3 based on recent contributions from the literature.

5.1.5 Airport Capacity Allocation

The impact of airport capacity allocation regulation on airlines and passengers is displayed in **bold lines**: From an airline perspective, under capacity regulation firms cannot directly choose their output volume. Instead, they manifest their airport capacity demand as an endogenous variable from their profit maximization. Subsequently, airport capacity is allocated to the airlines in order to maximize the overall social welfare in view of the inefficiencies caused by external congestion. Hence, depending on the respective allocation scheme and the respective capacity utilization rules, the airlines' production decision may be constrained by two distinct factors: Either by means of a resource allocation limitation, or by production costs resulting from a corresponding airport capacity allocation scheme. These capacity costs directly enter the airlines' profit functions and hence affect the flight volume decision. Finally, with flight fares as an inverse function of demand the airport capacity allocation changes flight fares and customer value.

In the model, the allocation scheme is exogenously determined by regulation policy. This allows the evaluation of the allocation process under different allocation instruments. In this respect, the regulation policy determines whether there are capacity costs according to an exogenous pricing function (as with taxation), a capacity constraint (as with airport quotas), or a market-based capacity pricing among participants (as with secondary trading of airport quotas). Whether an airline's airport demand affects its competitors' capacity costs and output depends on the allocation scheme put into effect. Therefore, the allocation scheme is shown as a moderator variable in Figure 5.1. In the absence of regulation (that is, under a first-come, first-served policy), capacity demand directly determines the allocation, and capacity costs (unlike congestion costs!) equal zero. Consequently, regulation also affects the degree of competition in the market.

Subsequently, it is important to recall that the network density effects exhibit an ambiguity from a welfare perspective: On the one hand, passengers are faced with the higher travel prices that arise from the airline's commercialization of the network value. This occurs because the corresponding product differentiation increases market power. On the other hand, the network airline's customers profit from the additional indirect travel benefits based on network density. As already outlined in Section 3.3.2, this complication was introduced by Starkie (2008a) as the dilemma of hub concentration. As a result, both the natural market structure and thus the impact of increasing or decreasing the networking airline's market share against its non-networking competitor on allocation efficiency initially remain unclear. Unfortunately, these causalities cannot be replicated in the graph from Figure 5.1 above. Nevertheless, they straightforwardly illustrate that the demand-side network density effects substantially complicate the welfare effect of airport capacity allocation.

5.2 Formal Setting

Formally, the generic model is based on Brueckner's (2002a) airport model for symmetric airlines under different market types. Subsequently, it is modified to implement the above causal relations arising from product differentiation based on network density effects. The model reflects an asymmetric airline duopoly with COURNOT quantity competition based on vertical product differentiation. The basic notation draws from Brueckner (2002a), but is altered to account for the specific characteristics where necessary. The central assumptions and specifications taken are discussed in Detail in Section 10.

5.2.1 Airport and Network Structure

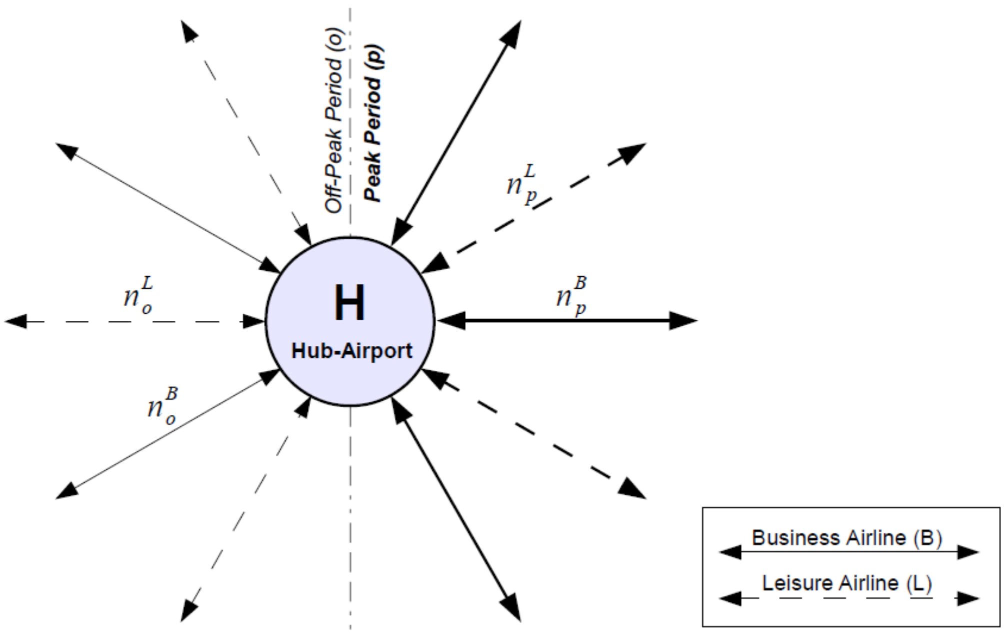

Fig. 5.2: Network Structure in the Generic Model

The model reflects a partly congested hub airport with two periods that connects to several uncongested destinations. By assumption, peak-period demand exceeds airport capacity and leads to congestion, which, in turn, causes time costs for passengers and congestion costs for the airlines. The off-peak period is free of flight delays. This setting is replicated from Brueckner (2002a).

Subsequently, Brueckner's (2002a) original model is modified to reflect two stylized airlines in a duopoly with vertical product differentiation. Above all, the airline asymmetry thus requires the individualization of flight volumes across airlines B and L as

$$N_p = n_p^B + n_p^L \quad \text{and} \quad N_o = n_o^B + n_o^L,$$

where uppercase letters are aggregates, and subscripts o, p denote the off-peak and the peak period. The corresponding network structure is depicted in Figure 5.2. The typical characteristics of the two asymmetric airlines are the following:

Business airline B reflects a dominant network airline with the model airport as its hub. Its network structure is assumed to constitute an incumbent advantage that provides *indirect network density benefits*. The indirect density benefits are assumed to abstractly reflect additional utility from travel flexibility, choice options, and connectivity within the network. The degree of these indirect travel benefits, in turn, is associated with a higher product quality. Product quality is presumed to abstractly reflect an advantageous design of the network schedule, which may, e.g., be manifested in a high flight frequency, attractive departure and arrival times throughout the day, and low schedule delays. For simplicity, however, product quality is not explicitly modeled, but approximated by means of the network density. The network density, in turn, directly corresponds to the business airline's peak-period output.

By contrast, the leisure airline L reflects a *point-to-point* carrier that only offers simple direct flight services and thus direct flight benefits only. It may be considered as a new entrant competitor or even as residual supply from multiple independent, non-networking airlines. As a consequence, the business airline enjoys an exogenous competitive advantage over the leisure airline. As explained above, this advantage is based on the opportunity to provide additional indirect travel benefits for the passengers from an interconnected network structure. The heterogeneous demand system is therefore modeled as follows.

5.2.2 Passenger Utility

Individuals maximize gross utility

$$U(\theta, x) \equiv u(\theta, x) + \eta$$

which is quasi-linear in travel utility $u(\theta, x)$ and residual consumption η of a numeraire good.

In terms of travel, individuals can choose between the following options: The first option is to travel during the off-peak period. Off-peak travel utility is independent of the airline choice because the respective costs and benefits are assumed to be identical. As a second option, passengers can travel during the peak period. In that case, they have to choose between the business or the leisure airline because the flights of the two airlines assumedly differ in quality based on their associated indirect travel benefits. Lastly, individuals can choose not to travel at all. Apart from these possibilities, individuals are assumed not to have outside opportunities for travel.[45]

Travel utility from these options can thus be formalized as a discrete-choice demand system

$$u(\theta, x) \equiv \begin{cases} b_o(\theta) & \text{for } x_0 = 1, \\ b_p(\theta) - t(N_p) & \text{for } x_p^L = 1, \\ b_p(\theta) - t(N_p) + d(\theta, n_p^B) & \text{for } x_p^B = 1, \end{cases} \tag{1}$$

where utility consists of direct flight benefits from air transportation $b_o(\theta)$ and $b_p(\theta)$ during the off-peak and peak period, passenger time costs $t(N_p)$ and indirect network density benefits $d(\theta, n_p^B)$ from peak-business flights. Additive separability and the distinct utilities of the three travel options already indicate the vertical nature of product differentiation across flights.

Within these functions, parameter $\theta \in [\Theta - 1, \Theta]$ denotes each individual's preference for peak-period travel.[46] It is assumed to be uniformly distributed with unit density $f(\theta) = 1$. Choice vector $x = (x_o, x_p^L, x_p^B)$ denotes each individual's discrete travel choice as a binary variable $x_o, x_p^L, x_p^B \in \{0, 1\}$ for off-peak-period travel and peak-period travel with the leisure or the business airline, respectively. Condition

$$x_o \cdot x_p^L = x_o \cdot x_p^B = x_p^L \cdot x_p^B = 0$$

ensures that only one option can be chosen from this set. Finally, $[\Theta - 1, \Theta]$ with $\Theta > 1$ represents the population range that corresponds to the unit density of the consumer taste. Its scale variability allows a later calibration of the model for either full or partial market coverage (see Sections 5.3.2 and 8.4).[47]

Following Lambertini (2006, p.165), the consumer taste for quality is formally defined as a

[45] This assumption may be justified by unavailability of High-Speed Rail (HSR) or by disproportionately high time costs for alternative travel modes. Also, it serves to confine the scope of this study.

[46] Tirole (1988, p.296) denotes θ as consumer taste for quality, whereas Lambertini (2006, p.161-162) defines θ as the marginal willingness to pay for quality. The former nomenclature equals the case of horizontal product differentiation (such as in Brueckner, 2002a), while the latter corresponds to the formal definition of θ under vertical product differentiation (see below).

[47] The general formalization of this demand system follows Ben-Akiva and Lerman (1985, pp.43), while the notation for flight benefits and time costs are adopted from Brueckner (2002a), and the specification for vertical differentiation draws on Tirole (1988, p.296) and Lambertini (2006, p.161-162).

marginal willingness to pay

$$\theta \equiv \frac{\alpha}{\partial u/\partial y}, \tag{2}$$

where $\partial u/\partial y > 0$ is the marginal utility of income and $\alpha > 0$ is any positive parameter. With utility assumed concave in income, the second-order derivative of income utility is given by $\partial^2 u/\partial y^2 \leq 0$. Assuming that incomes strictly increase from left to right within the interval $[\Theta - 1, \Theta]$ implies that the marginal willingness to pay θ is also positive and monotonously increases. The above derivation shows that θ formalizes the demand heterogeneity that is necessary to account for vertical product differentiation because it ultimately accounts for income differences across the population. This specification is standard in the vertical differentiation literature but deviates from Brueckner's (2002a) horizontal differentiation model (cf. e.g. Lambertini, 2006, p.163).[48]

The direct flight benefits from both periods are symmetric across airlines. However, it is assumed that direct flight benefits from peak-period travel $b_p(\theta)$ are higher than travel benefits $b_o(\theta)$ from an off-peak period flight, at least for a substantial portion of travelers on the right-hand side of the population range. This means that peak-period travel is more desirable than off-peak period travel for most individuals. The travel decision thus directly corresponds to a product quality choice and represents vertical product differentiation between peak and off-peak period flights. Formally, this assumption can be stated as $b_p(\theta) > b_o(\theta)$ for $\theta \in [\Theta_0, \Theta]$ where Θ_0 is arbitrarily small but satisfies $\Theta - 1 \leq \Theta_0 < \Theta$.[49]

Moreover, direct flight benefits are assumed to strictly increase with θ, which is formally provided by conditions $b_o(\theta)'$, $b_p(\theta)' > 0$. Every traveler within $\theta \in [\Theta - 1, \Theta]$ hence experiences a different individual value of the direct travel benefits, depending on his willingness to pay for air travel. This assumption implies that direct travel benefits increase with higher income, regardless of the travel period. In other words, travel utility generally increases with individual wealth. In addition, it is presumed that peak-period benefits increase more steeply than off-peak benefits with the individuals' marginal willingness to pay. This is assured by

[48] Defining the preference for peak-period travel as the willingness to pay allows the model to indirectly control for income, which is required to yield heterogeneous demand with vertical product differentiation based on quality. By contrast, in Brueckner's (2002a) model with *horizontal* product differentiation, population heterogeneity is defined based on *consumer taste* $\theta \in [0, 1]$. I am deeply indebted to Maria del Pilar Socorro Quevedo for a corresponding comment on an earlier specification of this model at the GARS Workshop 2014.

[49] Note that standard vertical differentiation models generally apply a monotonous quality order (cf. e.g. Tirole, 1988, p.296 or Lambertini, 2006, p.163). If $\Theta_0 > \Theta - 1$ the direct flight benefit functions cross over at the left-hand side of $[\Theta - 1, \Theta]$ and gross product qualities are actually reversed. However, according to utility (1) the travel decision also involves delay costs. This means that net benefits from air travel may hence also induce a quality reversion for low values of θ. Therefore, it is of little importance whether Θ_0 is actually larger or equal to $\Theta - 1$ and hence whether crossing occurs between $b_p(\theta)$ and $b_o(\theta)$ or not. This contrasts with Brueckner's (2002a) horizontal differentiation specification, where $b_o(\theta)$ and $b_p(\theta)$ in addition need to intersect at an "intermediate value" (idem, p.1361). This yields *single crossing*, which ensures that passengers are properly ordered across off-peak and peak travel according to their preferences.

assumption $b_p(\theta)' > b_o(\theta)'$. This condition yields that with higher income, peak-period travel becomes increasingly worthwhile as compared to off-peak-period travel.

The above two assumptions are directly adopted from Brueckner (2002a) but also correspond to the standard properties of a vertically differentiated duopoly (see Tirole, 1988, p.296 or Lambertini, 2006, p.163). They imply that peak-period travel becomes the more desirable for passengers, the less the flight fare matters to them. As a consequence, the marginal willingness to pay may also be referred to as a preference for peak-period travel. This preference, in turn, may justifiably be argued to inversely represent the price-sensitivity for air travel of the individuals. A brief discussion of these assumptions is provided in Section 10.5.1.

Function $t(N_p)$ denotes passenger time costs, which are also symmetric across airlines but only occur in the congested peak period. Therefore, they directly depend on total peak traffic volume. In correspondence with Brueckner (2002a), time costs are assumed non-decreasing, convex and homogenous across individuals with $t' > 0$ and $t'' > 0$. For simplicity, time costs are identically distributed across the individuals and hence do not depend on peak-travel preference θ.

Finally, $d(\theta, n_p^B)$ denotes the network density benefits for the passengers. As already delineated, these benefits abstractly denote product quality based on indirect additional travel utility arising from services and benefits associated with network structures. These benefits are not specifically modeled but are abstractly assumed to directly depend on network density. Because network services are deemed important mainly for business travelers, the density benefits are also a function of peak-flight preference. This specification is motivated in correspondence with the above arguments about the direct flight benefits.

For simplicity, network density itself is directly approximated by the business airline's peak-period output. On the one hand, this specification serves as a most simple generic specification of network density. On the other hand, this implies that off-peak-period flights do not generate network benefits, because they presumably do not form a part of the airline's peak-period network structure. Put differently, the business airline needs to provide its network during the peak-period travel times in order to provide indirect travel benefits. Otherwise, the related network services and their additional utility are not applicable.

Formally, the network density benefits function $d(\theta, n_p^B)$ is generally specified as a concave function with $d(\Theta-1, 1) = d(\Theta, 0) = 0$. This means that the density benefits are zero at both extremes to the left and to the right. This is justified by the assumption that individuals $\theta = \Theta - 1$ at the lower end of the consumer continuum do not appreciate network density.

Correspondingly, passengers $\theta = \Theta$ to the right-hand side of the population range highly appreciate network density, but $n_p^B = 0$ at the same time dictates that the network size is zero. In between these extremes, the network density benefits are strictly positive. This is achieved by presumptions $\frac{\partial d(\theta, n_p^B)}{\partial \theta} > 0$ and $\frac{\partial d(\theta, n_p^B)}{\partial n_p^B} > 0$, which define that the network density benefits are monotonously increasing both with peak-period preference and with network density.

Moreover, network density benefits are assumed to increase with increasing network density for every network user (i.e. peak-period passenger of the business airline). Therefore, the density benefits function is supposed to be twice continuously differentiable, with cross-partial derivatives $\frac{\partial^2 d(\theta, n_p^B)}{\partial \theta \partial n_p^B} > 0$ and $\frac{\partial^2 d(\theta, n_p^B)}{\partial n_p^B \partial \theta} > 0$ strictly positive. Concavity arises from the fact that θ and n_p^B are directly but inversely related to each other. This will become clear in the equilibrium computation in Section 6.3. Both this general specification as well as the additive separability of the network density benefits in the utility function are based on Belleflame and Peitz's (2010, p.554) generic proposition for utility from network goods.[50] The network density benefits function constitutes the central innovation to Brueckner's (2002a) model, as well as the key contribution of this study to the capacity allocation debate.

Lastly, note that parameter Θ can also referred to as the *population wealth*. Consequently, θ may also express each individual's *propensity to consume*. As explained above, the latter may be claimed to inversely reflect passengers' price sensitivity for air travel. Overall, the above demand system reflects network density benefits as vertical product differentiation. In that respect, it induces an exogenous demand asymmetry based on the perceived quality of heterogeneous flight benefits across periods and airlines.

5.2.3 Airline Profits

Finally, the airlines are profit maximizers. With f_o and f_p^i denoting flight fares, the profit function can be written as

$$\Pi^i[n^i(n^j), n^j(n^i)] = n_o^i \cdot f_o(n^i, n^j) + n_p^i \cdot f_p^i(n^i, n^j) - c \cdot (n_o^i + n_p^i) - n_p^i \cdot g(n_p^i + n_p^j). \qquad (3)$$

The pairs $n^i(n^j), n^j(n^i)$ denote each airline's output vectors $n^i = (n_o^i, n_p^i)$ and $n^j = (n_o^j, n_p^j)$. The above notation already makes clear that in oligopoly the firms are interdependent, as each

[50] Formally, Belleflame and Peitz (2010, p.554) propose a structure $U_{ij} = a_i + f_i(n_j)$ for consumer i. Utility hence consist of direct benefits a_i, and of indirect benefits $f_i(n_j)$ that depend on the number of network users n_j, where $f(0) = 0$ and $f(n_j)' > 0$. This model's specification is hence in-line with literature, with the number of network users replaced by the number of flights.

airline's output ultimately is determined as a reaction function of its competitor's outputs (see 6.1 below). The profit function denotes each airline's net gains as the total turnover from all output, minus operating and congestion costs.

The notion of turnover as a simple multiplicative term illustrates two important properties of the model: First, endogenous pricing is assumed to be discretionary, so that price discrimination is ruled out. Second, both the number of seats per aircraft and the seat load factor are set to unity. This means that the number of passengers directly corresponds to output in terms of flight volume. This simplification is without loss of generality and again follows Brueckner (2002a), so that the seat load factor assumption assures market clearing in equilibrium.

The two airlines' specific cost functions are derived from a generic total cost function

$$C_i(n_o^i, n_p^i, n_p^j) = c_i(n_o^i, n_p^i) + G(n_p^i, n_p^j),$$

which distinguishes between additively separable operating costs and congestion costs. The operating cost function is further specified as $c_i(n_o^i, n_p^i) = c\cdot(n_o^i + n_p^i)$, which denotes constant marginal operating costs. Congestion costs, in turn, presumably depend on overall peak-period output and on an airline's own peak-period flight share. They are hence specified as $G_i(n_p^i, n_p^j) = n_p^i \cdot g(n_p^i + n_p^j)$, where $g(n_p^i + n_p^j)$ denotes congestion as a function of aggregate peak-period output. Both cost functions are assumed to be monotonously increasing and continuously differentiable in both arguments.

5.3 Key Assumptions

Before turning to the results, this section presents two crucial assumptions for the analysis of the model's equilibrium and social optimum and the subsequent efficiency investigation of the different allocation schemes.

5.3.1 Partial Equilibrium Perspective

Above all, the following analysis adopts a partial-equilibrium perspective. For this purpose, income is assumed large enough so that the flight market represents a small share of each consumers' overall budget only. Under this premise, income effects are captured by the single composite numeraire good only. This means that the quasi-linear preferences (1) let us abstract from substitution effects from all other sectors in the economy (cf. Mas-Colell et al., 1995, p.50 and pp.316). As a consequence, this yields a demand system as a function of flight fares only when income is held constant (Jehle and Reny, 2011, p.50).[51] The advantage of this approach is two-fold: First, it reduces the scope of the analysis to the airline sector and thus avoids the complexities of a general equilibrium perspective. Second and as a direct consequence, the welfare analysis as the central part of the investigation of the allocation schemes becomes much more convenient.

For the efficiency analysis, recall that social welfare corresponds to the MARSHALLIAN aggregate surplus as defined in the social optimum computation from Section 7.1. This means that welfare reflects the sum of consumer surplus and net airline profits (cf. e.g. Mas-Colell et al., 1995, p.326). The consumer surplus is generally defined as utility from consumption minus the consumption costs. This consumer surplus hence reflects the net utility from consumption and may be referred to as indirect utility (Vives, 2001, p.76). From a partial equilibrium perspective net (or indirect) utility can hence directly be used to evaluate the individuals' consumption choices. As already described above, the latter thus become function of flight fares only. Consequently, no further conversions or computations are necessary for the welfare analysis: The net utility can directly be used to quantify the costs and benefits of the passengers under the distinct allocation schemes. Put differently, the consumer surplus becomes an „appropriate measure of welfare change" because it „corresponds directly to the indirect utility function" (Vives, 2001, p.77).

This simplification follows most theoretical contributions in the field such as Brueckner (2002a) and (2002b), Zhang and Zhang (2006), Czerny (2010) or Verhoef (2010). For a justification of this approach also see the discussion in Section 10.6.

5.3.2 Partial Market Coverage

In addition to the partial equilibrium perspective, the model also assumes partial market coverage. This means that in equilibrium some individuals choose not to travel.[52]

The importance of this assumption relates to two different reasons: On the one hand, the degree of congestion is a function of total peak-period traffic and hence is endogenous to the model. However, congestion can only vary with the different peak-period outputs under

[51] Formally, this is shown in that utility becomes $u(f) = u[x(f)]$, and thus is not affected by cross-price effects (Vives, 2001, p.76).

[52] The terminology of market coverage varies between different contributions in the field: Ecchia and Lambertini (2006) refer to this situation as to a partially covered market while both Wauthy (1996) and Motta (1993) call this as an uncovered market. I will stick to the former definition because I find it more intuitive.

the distinct capacity allocation schemes if not all travelers are allocated to the peak period; otherwise, the entire market is served with peak-period flights and congestion always remains at its maximum value and thus becomes invariable. Therefore, the investigation of airport congestion becomes uninteresting when the market is fully covered. As a consequence, the consideration of endogenously variable congestion requires the model airport to be partly congested only. This property is achieved by assuming partial market coverage.

On the other hand, partial market coverage is important from a technical point of view: In this COURNOT oligopoly model, the demand elasticity is finite, so that the endogenous flight fares are determined by both airlines' outputs. The formal specification of the demand functions thus differs depending on whether the market is fully or only partly served: When full market coverage is observed, the demand specification for partial market coverage becomes invalid and needs to be revised (cf. e.g. Lambertini, 2006, p.64 or Ecchia and Lambertini, 2006, p.86). As a result, the discontinuity of the discrete-choice demand system also justifies considering partial market coverage only.[53]

The literature on vertical differentiation has established that market coverage in a vertical equilibrium depends on population heterogeneity and on the degree of product differentiation (see e.g. Wauthy, 1996, p.348). In order to achieve partial market coverage, the following two conditions are therefore adopted: First, it is assumed that at the left-hand end of the population range some individuals do not travel. Partial market coverage thus requires that the left-most traveler's willingness to pay is larger than the lower bound of population interval $\theta \in [\Theta - 1, \Theta]$. This condition is implemented by restricting Θ to an intermediate magnitude in absolute terms (cf. e.g. Ecchia and Lambertini, 2006, p.86). For this purpose, it would actually be sufficient to only restrict the peak-period market coverage because the off-peak-period outputs do not affect congestion. However, the lower off-peak-period limit only depends on exogenous parameters whereas market separation between peak and off-peak is endogenous. The former case is therefore much easier to formalize. Because it is also without loss of generality, partial market coverage based on overall output is preferred and implemented. Still, this distinction is only important for the calibration of population range Θ in the specified model but not for the analytic investigation of the generic model (see Section 8.4).

Second, the utility definition from equation (1) leads us to suspect that the network density benefits $d(\theta, n_p^B)$ might induce a corner solution where the business airline's peak-period flights satisfy all demand. This would mean that the business airline's network preempted

[53] In addition, partial market coverage avoids that the inverse demand functions might not be found in COURNOT competition, because with full market coverage total demand is no longer a function of prices (see Motta, 1993, p.116).

the entire market because all individuals could afford a peak-business flight and were best-off with that solution. Yet, this case neither yields any results worth investigating. Therefore, network density benefits are assumed to be limited in magnitude. In addition to the boundary on Θ from the above first assumption, this requires a condition on the size of the density benefits relative to the direct flight benefits and to congestion. This condition ensures that both airlines operate in the peak period and that the off-peak period is served at all.

As is common in the literature, the conditions both on Θ as well as on $d(\theta, n_p^B)$ are explored ex-post based on the first-order conditions of the equilibrium (cf. e.g. Choi and Shin, 1992; Ecchia and Lambertini, 2006). As θ denotes the consumers' marginal willingness to pay for a flight indicates that partial market coverage requires an appropriate specification of the considered population range by means of Θ (cf. e.g. Ecchia and Lambertini, 2006, p.86). In turn, the fact that the degree of market coverage depends on individuals' wealth shows that market coverage ultimately represents an income effect. The appropriate choice of maximum consumer wealth Θ thus ensures that the left-most individuals on the θ-scale abstract from consumption. This guarantees a non-degenerate, partial market coverage in the vertical differentiation model.

Formally, partial market coverage is achieved simply by calibrating parameter Θ in the equilibrium first-order conditions in such way that overall output volume remains smaller than unity. Other than in horizontal differentiation models, the location of the population range on the horizontal scale of wealth remains variable but its size remains unity at all times. For the numeric solution of a specified model, the appropriate calibration of population range parameter Θ needs to be evaluated iteratively. This procedure is described in greater detail in Section 15.3.2 of the simulation in Part III. For the analytic investigation of the generic model, it is sufficient to presume that Θ has a magnitude that yields partial market coverage.

6 Equilibrium

This section develops the unregulated market equilibrium based on the airlines' profit maximization and inverted demand resulting from the individuals' utility maximization.

6.1 Profit Maximization

Under COURNOT competition, airlines maximize their profits given that flight fares are endogenous functions of demand. The profit maximization problem is hence given by

$$\max_{n_o^i, n_p^i} \Pi^i \left[f_o(n^i, n^j), f_p^i(n^i, n^j) \right] \quad \text{s.t.} \quad n_o^i, n_p^i \geq 0, \tag{4}$$

where the constraint states that outputs cannot be negative. The fact that the demand functions depend on both airlines' output vectors stipulates that optimal outputs in oligopoly are ultimately given by reaction functions, which denote each airline's optimal output choice given its competitor's output (see Section 3.4.1).

Due to the non-negativity condition, constrained maximization applies. Optimality for the firms thus requires the simultaneous solution of

$$\text{(eq.i): } {}^{\partial \Pi^i}/_{\partial n_k^i} \leq 0 \quad \text{and} \quad \text{(eq.ii): } n_k^i \cdot {}^{\partial \Pi^i}/_{\partial n_k^i} = 0 \tag{5}$$

for each airline $i \in \{B, L\}$ and both periods $k \in \{o, p\}$, with all $n_k^i > 0$ (cf. Jehle and Reny, 2011, pp.591). The endogenous flight fares are found by the inversion of the demand functions from utility maximization. Derivation of these flight fares is provided at the end of this section.

The constrained maximization according to conditions (5) is performed as follows: Equality condition (eq.ii) is evaluated first because inequality in condition (eq.i) applies depending on the result of (eq.ii): Namely, if there is a non-zero solution $n_k^i > 0$ to (eq.ii) it automatically makes condition (eq.i) hold with equality. It is easily inferred from (eq.ii) that this applies when n_k^i balances the gradient of the maximized function to zero, so that ${}^{\partial \Pi^i}/_{\partial n_k^i} = 0$. In this case, an interior optimum has been found. If, in addition, strict concavity can be shown, this optimum represents a unique global maximum.

However, if there is no strictly positive solution to (eq.ii), the only result that complies with the equality is $n_k^i = 0$. Obviously, this happens when ${}^{\partial \Pi^i}/_{\partial n_k^i} < 0$, so that any non-zero solution would violate the non-negativity constraint. Consequently, this corner solution can no longer fulfill (eq.i) with equality. The inequality therefore indicates that the corner solution is governing. This case arises when the global optimum of (4) is outside the valid parameter range and thus is void. The corner solution therefore denotes a local extreme point. It constitutes a local maximum when it complies with concavity (cf. Jehle and Reny, 2011, pp. 570 and pp.591).

6.2 Utility Maximization

The utility maximization problem is generally denoted as

$$\max_{\eta,x} U(\theta,x) \quad s.t. \quad \eta + x \cdot f \leq I,$$

where personal income I needs to be sufficient for consumption of the numeraire good η and the flight chosen. The partial equilibrium perspective, however, allows us to reduce the above problem to

$$\max_{x} \left[u(\theta,x) - x \cdot f \right],$$

which simply denotes the maximization of flight utility net of flight fares (cf. e.g. Vives, 2001, p.76-77).[54] All consumption choices that solve the above problem can now be subsumed by a maximum value function, which states flight demand as a function of flight fares (cf. Jehle and Reny, 2011, p.28 and p.50). I refer to $cv(\theta, f)$ as the customer value function, which corresponds to net or indirect utility and therefore can directly be used in the welfare analysis (cf. Vives, 2001, p.76). Due to the discrete-choice setting, however, utility (1) is not continuously differentiable. Therefore, the demand functions need to be determined by evaluating the indifference conditions between each adjacent option of choice (cf. Ben-Akiva and Lerman, 1985, p.44).[55]

6.3 Inverted Demand

Evidently, the first individual $\theta \in [\Theta - 1, \Theta]$ from the left will travel when the travel fare exceeds the travel benefit, so that a positive net utility from this flight arises. Further to the right, passengers will switch to peak-period travel with the leisure airline as soon as it

[54] Formally, the budget constraint can be re-arranged and substituted for the numeraire good in the utility function because it is always binding with equality. As income is held constant for every individual, with quasi-linear preferences it does not affect first-order conditions (cf. Mas-Colell et al., 1995, p.318).

[55] This contrasts with models applying a single, continuous and strictly quasi-concave utility function, where differentiation yields the demand function as a unique solution to the consumer's utility maximization problem (Jehle and Reny, 2011, p.28).

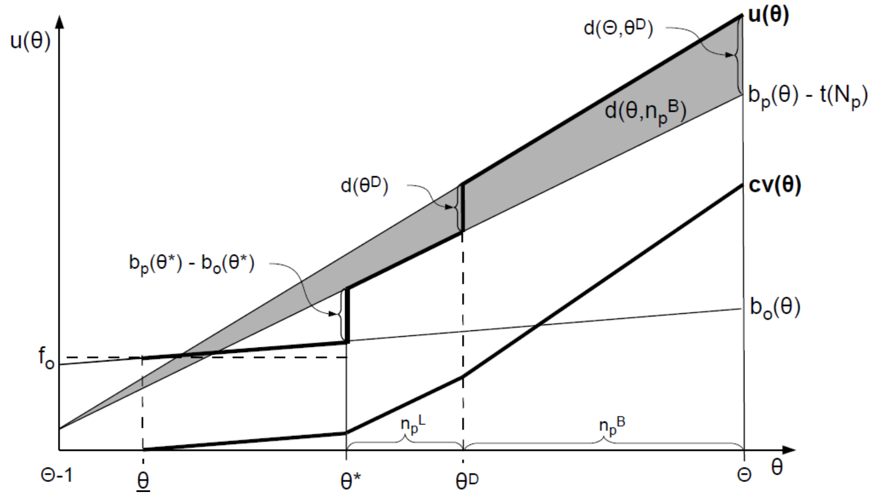

Fig. 6.1: Travel Benefits and the Characteristic Values of θ

becomes more valuable, despite the congestion incurred. Lastly, the higher-θ individuals will revert to peak-business travel as soon as the additional density benefits overcompensate the higher peak-business flight fare. Customer value as a maximum value function that describes all solutions to the above problem can hence be written as

$$cv(\theta, f) \equiv \max \left[0,\ b_o(\theta) - f_o,\ b_p(\theta) - b_o(\theta) - t(N_p) - f_p^L + f_o,\ d(\theta, n_p^B) - \left(f_p^B - f_p^L\right)\right]. \quad (6)$$

To compute inverse demand, the critical consumers that separate demand according to (6) are denoted by their characteristic peak-travel preference: From left to right, the first passenger traveling at all is denoted as $\underline{\theta}$. Next, the passenger switching from off-peak travel to peak-period travel with the leisure airline is referred to as θ^*. Lastly, the passenger that is indifferent about peak travel with the leisure airline and a peak flight from the business airline is denoted θ^D. By assumption, indifferent consumers choose the next travel option to their right. Figure 6.1 illustrates the customer value function and its corresponding characteristic values of peak-travel preference. The notation again draws on Brueckner (2002a).

The indifference conditions for the critical passengers can be extracted from (6) into an equation system. After cross-substitution, inversion and re-arrangement the latter reads

$$\begin{aligned} f_o(\underline{\theta}) &= b_o(\underline{\theta}), \\ f_p^L(\theta^*) &= b_p(\theta^*) - b_o(\theta^*) - t(N_p) + f_o(\underline{\theta}), \\ f_p^B(\theta^D) &= d(\theta^D) + f_p^L. \end{aligned} \quad (7)$$

For brevity, the fare difference $f_p^L - f_o$ between peak-period leisure and off-peak-period flights is referred to as the peak premium. Similarly, the markup between leisure and business flights

$f_p^B - f_p^L$ in the peak period is denoted as the density premium. And lastly, $f_p^B - f_o$ is the peak-density premium.

At this point, note that individuals $\theta \in [\underline{\theta}, \theta^*[$ are off-peak passengers, while individuals $\theta \in [\theta^*, \theta^D[$ are peak-leisure travelers and individuals $\theta \in [\theta^D, \Theta]$ are peak-business travelers. With both the seat load factor and aircraft size at unity, the characteristic values relate to the airlines' output variables as

$$n_p^B = \Theta - \theta^D, \;\; n_p^L = \theta^D - \theta^* \text{ and } \;\; n_o^B + n_o^L = \theta^* - \underline{\theta}.$$

These equivalences can be re-arranged to show that $\underline{\theta} = \Theta - n_o^B - n_o^L - n_p^B - n_p^L, \quad \theta^* = \Theta - n_p^B - n_p^L$ and $\theta^D = \Theta - n_p^B$. Using these equivalences in equation (7) finally yields the inverted demand functions, which denote the flight fares as a function of output.

6.4 First-Order Conditions

The off-peak derivatives are symmetric across airlines because flights are homogeneous in the off-peak period. Condition (eq.i) from above is therefore also symmetric and reads

$$b_o(\underline{\theta}) - \left[n_o^i + n_p^i\right] \cdot b_o'(\underline{\theta}) \geq c. \tag{8}$$

The marginal benefits term $b_o'(\underline{\theta})$ reflects "traditional" market power. Its effect is best illustrated by re-arranging it to the right-hand side in equation (8), so that the strict monotonicity of $b_o' > 0$ and condition $f_o(\underline{\theta}) = b_o(\underline{\theta})$ from equation (7) make clear that the marginal benefits term has a net positive impact on the flight fare by raising it above marginal costs (also cf. Brueckner, 2002a, pp.1363). Due to the inverse relationship of flight volumes and the critical θ's, however, the direction of a change in the marginal benefits term with changing output initially remains ambiguous. This indicates the existence of a unique solution.

The peak-period derivatives, by contrast, are heterogeneous across airlines. In an interior solution, the off-peak-period conditions (8) can be used with equality and thus simplify the peak-period terms to

$$B^* - TG^* \;\; \geq \;\; n_p^L \cdot \left[B^{*'} + TG^{*'}\right], \tag{9}$$

$$B^* - TG^* + d(\theta^D) \;\; \geq \;\; n_p^B \cdot \left[B^{*'} + TG^{*'} + d'(\theta^D)\right]. \tag{10}$$

The shortcuts $B^* \equiv b_p(\theta^*) - b_o(\theta^*)$ and $TG^* \equiv t(N_p) + g(N_p)$ subsume the relative peak-flight benefits of the critical peak-period passenger as well as overall time and congestion costs in short notation. Correspondingly, $B^{*'}$ and $TG^{*'}$ denote their first derivatives with regard to overall peak-period output.

Observe that marginal congestion costs are included in each airline's profit maximizing rationale only to the extent of each airline's own share of output. This means that each competitor's delay costs are not accounted for, which results in the externality of congestion in equilibrium. Therefore, the marginal terms $B^{*'}$ and $TG^{*'}$ represent the distortions of equilibrium outputs by market power and congestion externalities. Obviously, those terms are evaluated at θ^* as this value denotes total peak-period output $N_p = 1 - \theta^*$ and determines the overall level of congestion.

The inequalities in the above conditions are interpreted as follows: First, recall that conditions (8) to (10) denote an interdependent equation system, which means that the equations need to be solved simultaneously. The first-order conditions thus constitute the airlines' best-response functions to each competitor's output, representing a simultaneous NASH equilibrium as is standard in COURNOT Oligopoly competition (cf. e.g. Jehle and Reny, 2011, p.173).

Starting with the peak period conditions, either there is any valid output combination $n_p^L, n_p^B > 0$ with $n_p^L + n_p^B = N_p = 1 - \theta^*$ that makes (9) and (10) hold with equality (as explained in Section 6.1). This solution represents an interior optimum. Or, in contrast, if any negative output $n_p^L, n_p^B < 0$ is required to fulfill both equations, this negative output violates the non-negativity constraint from (4). Consequently, the respective output reverts to zero. The respective condition is still satisfied as it holds with the inequality governing. This case represents a corner solution that replaces the global maximum by a local maximum, as the former is not within the proper range of definition. Once peak-period outputs are determined, off-peak-period outputs $n_o^L, n_o^B \geq 0$ are similarly resolved based on symmetric off-peak-period conditions (8).

Both the interior optimum and a potential corner solution are discussed next. Note that this computation contrasts with Brueckner's (2002a) problem, where symmetry and additive separability yield uni-variate first-order conditions that can be solved independently.

6.5 Interior Optimum

In the interior optimum, (8) holds with equality for both airlines. Symmetry thus implies that

$$n_o^B + n_p^B = n_o^L + n_p^L. \tag{11}$$

This signifies that each airline has to offset its output changes across periods, so that the overall market is divided equally across the two airlines. The model hence does not allow for market exit. Total output is thus determined by marginal costs and the functional form of off-peak-period benefits according to (8).

In the peak period, the equilibrium conditions also hold with equality if the interior solution is valid. These conditions illustrate the impact of the dual distortion: Consider first the right-hand side of equation (9) and recall that marginal direct benefits $B^{*'}$ strictly increase with θ^*. Marginal delay costs $TG^{*'}$, by contrast, strictly decrease with θ^* because they monotonously increase with $N_p = 1 - \theta^*$. Consequently, the sign of the term within brackets is generally ambiguous. Notice, however, that for any positive output n_p^L, the difference between actual direct flight benefits and delay costs $B^* - TG^*$ must have the same sign as the bracket on the right-hand side. Moreover, observe that based on the above monotonicity, there is single crossing between B^* and TG^*. As a result, the sign of $B^* - TG^*$ reveals whether output is relatively low or relatively high, where the output deviation ultimately denotes the size and the sign of the dual distortion.

If congestion costs are important relative to market power then $TG^{*'}$ is large in relation to $B^{*'}$. In that case, the right-hand side of (9) becomes negative and thus also the difference $B^* - TG^*$. When delay costs are important because $TG^{*'}$ is high, function TG^* is relatively steep compared to benefits B^*. The single crossing between the two functions tends to lie toward the right-hand side of population spectrum $[\Theta - 1, \Theta]$. Although the equilibrium value is to the left of the crossing point where $B^* - TG^* < 0$, this indicates that θ^* comes to rest at a relatively high value. As a consequence, the peak-period output is relatively low when $TG^{*'}$ is large in relation to $B^{*'}$.

By contrast, flight benefits that are important relative to congestion shift the single-crossing point toward the left where the sum of $B^{*'} + TG^{*'}$ is positive. Consequently, the equilibrium value for θ^* is located to the right of the crossing point. Nevertheless, the crossing point lies farther to the left in the θ-continuum so that peak-period output is relatively large. The sign of $B^{*'} + TG^{*'}$ hence determines the sign of $B^* - TG^*$ and demonstrates the two counteracting, single effects of the dual distortion. This is illustrated in Figure 6.2 below.

The airline asymmetry can be shown by comparison of equilibrium conditions (9) and (10): Recall first that the two sides of (9) can have either sign but that the sign needs to be equal on both sides for the condition to be binding with equality. Correspondingly, the two equilibrium conditions differ in that the business airline is additionally concerned with the actual network benefit $d(\theta^D)$ at the left-hand side and with the marginal network benefit $n_p^B \cdot d'(\theta^D)$ on the right-hand side. These two terms hence formally denote the airline asymmetry.

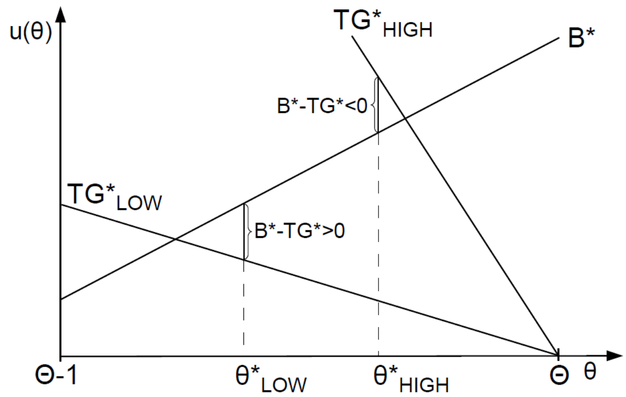

Fig. 6.2: Peak-Period Output and Dual Distortion

The impact of this difference is formally revealed as follows: First, the network benefit $d(\theta^D) > 0$ is always positive, which means that the left-hand side of (10) is also larger than the left-hand side of (9). In order for equilibrium condition (10) to be balanced, the peak-period output of the business airline therefore generally needs to be higher than the peak-period output of the leisure airline. There is, however, a complication in this simple relationship: The sign of the marginal network benefit $d'(\theta^D)$, which is generally ambiguous. As a consequence, the above airline asymmetry is more pronounced when $d'(\theta^D) < 0$. By contrast, a positive sign of $d'(\theta^D)$ decreases the business airline's peak-period output so that the airline asymmetry could become reversed.

However, the division of (10) by n_p^B enables us to show that a reversal of the asymmetry does not occur: The latter yields $d(\theta^D)/n_p^B$ on the left-hand side and $d'(\theta^D)$ on the right-hand side. These two terms represent the average and the marginal network density benefit for traveler θ^D, respectively. Recalling that the network value directly enters the endogenous peak-business flight fare f_p^B in (7), it can consequently be argued that the average network benefit must exceed the marginal network benefit in equilibrium. Otherwise, reducing flight volume by one unit would raise the network benefit and thus the peak-density premium by more than its own average at the margin. In addition, such an output contraction would reduce congestion so that the overall profits of the business airline would increase. This implies that

$$d(\theta^D) > n_p^B \cdot d'(\theta^D) \tag{12}$$

in equilibrium and hence that $n_p^B > n_p^L$ in the interior optimum. The network benefit $d(\theta^D)$ will subsequently be referred to as the network value and is further investigated in Section 8.2.

Lastly, note that the left-hand side of (10) equals $f_p^B - f_o$ from (7). The business airline's density premium includes both the traditional mark-up as well as the marginal network benefit, and with $d(\theta^D) > 0$ it follows that $f_p^B > f_p^L > f_o$. With both a higher flight fare and a higher output in the peak period, business airline profits are higher than the profits of the leisure airline. Obviously, the leisure airline's higher off-peak-period market share cannot compensate the lower off-peak-period flight fares. The density benefits thus increase both the peak-period market share and the profitability of the business airline. Since the second-order conditions show that both profit functions are strictly concave, the interior solution is a unique global maximum (see Appendix B).

6.6 Corner Solution

In contrast to the interior optimum, the simultaneous solution of peak-period conditions (eq.ii) might also yield a negative off-peak business output $n_o^B < 0$. This may arise if the optimization constraints are not considered because a negative off-peak business output would increase the business airline's peak-period mark-up while keeping total peak-period output and thus congestion and market power constant. In this case, the non-negativity constraint requires the business airline to revert to corner solution $n_o^B = 0$. As a consequence, the business airline's off-peak condition (8) no longer holds with equality so that the symmetry condition (11) is altered to

$$n_p^B \geq n_o^L + n_p^L. \tag{13}$$

As (10) shows, this corner solution may arise if the network value is large because the network density benefits are relatively important. In contrast to the interior solution, it allows overall output to become asymmetric across airlines. As a consequence, the business airline's superior profitability over the leisure airline further increases. In that case, the business airline's off-peak condition (8) holds with inequality only and cannot be used to simplify the business airline's peak-period derivative. Condition (10) thus becomes invalid as the peak-business first-order condition must be altered to $B^* - TG^* + d(\theta^D) = n_p^B \cdot \left[B^{*'} + TG^{*'} + d'(\theta^D)\right] + \left(n_p^B - n_o^L - n_p^L\right) \cdot b_o'$, where the additional term on the right-hand side compensates for the excessive network value.

The occurrence of this corner solution is checked ex-post by computation of the interior optimum from (8), (9) and (10), and subsequent evaluation of asymmetry condition (13). If the latter holds, the corner solution becomes governing so that the equilibrium needs to be re-computed with $n_o^B = 0$. Due to the exogenous asymmetry of the network density benefits, other corner solutions are unlikely to occur and therefore remain unexplored.

6.7 Second-Order Conditions

The interior optimum represents a unique global maximum if it can be shown to be strictly concave. The concavity of a multidimensional function, in turn, is assured if its Hessian matrix is negative definite (cf. Jehle and Reny, 2011, p.559 and pp.570). If the corner solution can be shown to be concave it only represents a local maximum, as it hinges on the non-negativity constraint. In addition, the concavity formally proves that both extrema do represent maxima and not minima.

The derivation in Appendix B.2 shows negative definiteness for both the interior and the corner solutions.[56] Therefore, the above interior solution denotes the global maximum of the airlines' profits whereas the corner solutions denote their local maxima where the non-negativity constraint is binding.

[56] In this respect, note that negative definiteness represents a sufficient but not a necessary condition for concavity (Mas-Colell et al., 1995, p.933). This signifies that both above solutions may denote maxima independent of the value of their Hessians.

7 Social Optimum

In contrast to the equilibrium, this Section determines the hypothetical capacity allocation needed to create maximum social welfare, based on the social welfare function.

7.1 Determination

Welfare is defined as the MARSHALLIAN aggregate surplus, consisting of consumer surplus and net airline profits (cf. e.g. Mas-Colell et al., 1995, p.326). In partial equilibrium, the customer value function directly corresponds to indirect utility and is equivalent to consumer surplus (cf. Vives, 2001, p.77). It can hence directly be used in the welfare function, so that the social optimum is defined in analogy to Brueckner (2002a) as

$$\max_{\underline{\theta},\theta^*,\theta^D} W(\theta) = \int_{\underline{\theta}}^{\Theta} cv(\theta, f)d\theta + \Pi^L + \Pi^B \qquad \text{s.t.} \quad \left\{\underline{\theta}, \theta^*, \theta^D\right\} \in [\Theta - 1, \Theta] \tag{14}$$

The boundary condition technically ensures that the critical θ's are within the valid population range. In addition, condition

$$\underline{\theta} \leq \theta^* \leq \theta^D$$

ensures that the critical θ's are allocated in an ordered manner from left to right within range $[\Theta - 1, \Theta]$, so that no leap-frogging occurs. Otherwise, the optimum might yield negative individual outputs for some periods or airlines. The above ordering rule hence simply corresponds to the non-negativity constraint from equilibrium problem (4) but is related to the critical θ's.[57]

Constrained optimization invokes similar necessary conditions as in the equilibrium. However, note that the boundary conditions are now two-sided so that

$$\text{(so.i)} \ {}^{\partial W}/_{\partial\theta} \leq 0, \quad \text{and} \quad \text{(so.iia)} \ [\theta - (\Theta - 1)] \cdot {}^{\partial W}/_{\partial\theta} = 0 \qquad \text{or} \quad \text{(so.iib)} \ (\Theta - \theta) \cdot {}^{\partial W}/_{\partial\theta} = 0 \tag{15}$$

for all $\theta \in \left\{\underline{\theta}, \theta^*, \theta^D\right\}$. Condition (so.ii) is now dual and applies depending on which boundary value is restricting:

[57] By contrast, in equilibrium the individual outputs are subject to non-negativity constraints before they are translated into the critical θ's. In the market solution, hence, the ordering condition is implicitly ensured based on the non-negativity constraints.

Either any $\theta \in [\Theta - 1, \Theta]$ can solve (8) with equality and thus represents an interior optimum, or, there is no such θ within $[\Theta - 1, \Theta]$ and hence there is no interior solution. In that case, the boundary conditions determine the corner solution: If the candidate solution for any characteristic θ is below the lower bound $\Theta - 1$, condition (so.iia) becomes governing, and the lower corner solution $\theta = \Theta - 1$ applies. By contrast, if the candidate θ is higher than upper bound Θ then condition (so.iib) is involved. In that case, the resulting corner solution is $\theta = \Theta$.

For the derivation of the optimum, the welfare function is continuously differentiable. In contrast to the equilibrium, the discrete-choice boundary conditions are not required and the social optimum can directly be determined by standard analysis. Notice, however, that both the direct flight benefits and congestion in the peak period depend on overall peak-period output: Depending on whether there is an interior or a corner solution in $\theta^* = 1 - N_p$, it is either θ^D or θ^* that determines overall peak-period output.

In order to ensure that the respective first-order conditions of the peak-period θ's each take account of their potential effect on congestion and direct flight benefits, the welfare function needs to be enhanced by a minimum value function concerning θ^* and θ^D. Substitution of airline profits from (3) and customer value from (6) into (14) yields this computable welfare function as

$$W(\theta) = \int_{\underline{\theta}}^{\min\left[\theta^*, \theta^D\right]} b_o(\theta) d\theta + \int_{\min\left[\theta^*, \theta^D\right]}^{\Theta} b_p(\theta) d\theta + \int_{\theta^D}^{\Theta} d(\theta, n_p^B) d\theta - \left(1 - \min\left[\theta^*, \theta^D\right]\right) \cdot \left[t(N_p) + g(N_p)\right] - (1 - \underline{\theta}) \cdot c. \quad (16)$$

The minimum value function $\min\left[\theta^*, \theta^D\right]$ ascertains that each peak-period first-order condition takes into account its own impact on congestion and direct flight benefits. Note that the delay cost functions are not concerned with the integral operator because they do not depend on peak-travel preference.

Formally, the minimum conditions are treated as follows: First, the optimization of the welfare function treats the critical θ's as independent variables. This means that these are evaluated directly and independently from a global point of view. Consequently, the welfare function is differentiated with respect to the critical θ's, assuming in each case that the respective variable is governing the minimum value function. After derivation of the first-order conditions, each unconstrained solution is checked for compliance with the boundary conditions $\left\{\underline{\theta}, \theta^*, \theta^D\right\} \in [\Theta - 1, \Theta]$ and ordering rule $\underline{\theta} \leq \theta^* \leq \theta^D$. If violations occur, the unconstrained solutions are invalid and are replaced by the respective corner solutions.

Otherwise, the results represent interior solutions for the social optimum. Consequently, the social optimum is either a global or a local maximum, but the solution set is fully valid and well-ordered across the population range.

Note that the social optimum computation differs from the equilibrium: In the latter, the airlines compete interdependently in NASH fashion. Hence, the first-order conditions represent simultaneous, mutual reaction functions that need to be solved as an equation system. In the social optimum, the three strategic variables are solved independently and checked against the boundary conditions in order to distinguish the interior and the corner solutions.

7.2 Off-Peak-Period Output Condition

The optimal threshold for off-peak travel is determined by (so.i) as

$$b_o(\underline{\theta}) \geq c. \tag{17}$$

The inequality is interpreted as explained above: If an interior solution exists, the lower bound of travelers is given by $b_o(\underline{\theta}) = c$. This means that the left-most traveler $\underline{\theta}$ should be chosen so as to equate his marginal benefit $b_o(\underline{\theta})$ from off-peak travel with marginal cost c. Due to the monotonicity of the direct travel benefits functions, all passengers to the right of that critical passenger would enjoy benefits $b_o(\theta) > c$ in excess of marginal costs. Consequently, they would all travel. By contrast, passengers to the left of $\underline{\theta}$ could not sufficiently compensate their travel expenses with their benefits, because they could only yield $b_o(\theta) < c$. Accordingly, they would not travel. This interior solution thus arises when the leftmost individual's direct flight benefit is smaller than marginal costs.

By contrast, if the travel benefit of the left-most individual exceeds marginal costs, this individual always has a positive net utility from travel. This means that $b_o(\theta) > c$ for all passengers across the whole population range. Consequently, the corner solution arises where all individuals travel. In this case, condition (so.i) evidently cannot hold with equality as the required lower bound would technically be located beyond the lower limit of the population range. Correspondingly, condition (so.iia) becomes governing and reverts the invalid candidate interior solution to $\underline{\theta} = \Theta - 1$. From this it follows that $b_o(\Theta - 1) > c$ so that (17) holds with inequality only.

The above considerations show that a model specification with $b_o(\Theta - 1) < c$ yields an interior solution $\underline{\theta} > \Theta - 1$, where some individuals do not travel. By contrast, assuming $b_o(\Theta - 1) > c$ yields the corner solution, where the overall passenger number spreads across the entire population range. In this respect, note that the partial market coverage assumption from the equilibrium applies accordingly but may yield a different result for market coverage in the social optimum: Outputs may be higher or lower, depending on the size and direction of the dual distortion. Moreover, these conditions demonstrate that assumption $c < b_o(\Theta)$ ensures that the (rather theoretical but degenerate) overshoot $\underline{\theta} > \Theta$ at the right-hand side of the population range will not occur.

Lastly, observe that equation (17) does not explicitly determine the allocation of off-peak-period outputs across the two airlines in the social optimum. As flights are homogenous products in the off-peak period, however, this allocation has no impact on welfare. For the quantitative analysis in Part III the off-peak-period output will be presumed to fully accrue to the leisure airline based on distributional grounds.

7.3 Peak-Period Output Condition

Condition (so.i) for overall peak-period output reads

$$B^* - TG^* \geq \left[n_p^L + n_p^B\right] \cdot TG^{*'}. \tag{18}$$

Comparison of (18) to equilibrium peak-leisure output from (9) shows that there are two important differences across the two terms. Firstly, the marginal flight benefits term has vanished. As this term enters condition (9) on the right-hand side and is positive, the resulting overall peak-period output in (18) must be overall higher when other things remain equal. This means that the market power distortion is removed from equation (18), and thus that output is no longer depressed inefficiently. Secondly, observe that in contrast to (9), the entire share of the peak-period output is included on the right-hand side of (18). Recalling that $TG^{*'}$ denotes marginal congestion and time costs, this signifies that the full amount of delay costs is accounted for in the social optimum. In other words, in (18) congestion is fully internalized, which again decreases output. Whether in that case the final flight volume is above or below the equilibrium value from (9) ultimately depends on which of the two distortions was previously dominant.

Ultimately, peak-period condition (18) demonstrates that in the social optimum, the net benefit from the left-most peak-period flight must compensate the marginal social costs of this flight. If marginal congestion and time costs are lower than the net benefit evaluated at $\theta^* = 1 - N_p$, then peak-period output needs to expand. By contrast, if the marginal delay costs exceed the first peak-period passenger's net benefit, then peak-period output needs to contract.[58] Moreover, because B^* is strictly increasing and TG^* and $N_p \cdot TG^{*'}$ are strictly decreasing with θ^*, there is single crossing between the left-hand and the right-hand side of (18). This means that algebraically, there is a unique positive solution that enables condition (18) to hold with equality.

Lastly, it remains to be checked whether this solution complies with the above ordering and boundary conditions. If the ordering condition is infringed because $\theta^* > \theta^D$, then the solution needs to revert to $\theta^* = \theta^D$ in order to avoid a degenerate optimum with a negative peak-leisure output. In that case, a corner solution arises where the leisure airline does not serve the peak period at all.

7.4 Optimal Network Size Condition

Differentiation of welfare condition (16) for θ^D yields the condition (so.i) for peak-period flights of the business airline as

$$B(\theta^D) - TG(\theta^D) + d(\theta^D) \geq \left(\Theta - \theta^D\right) \cdot TG^{D'} + \int_{\theta^D}^{\Theta} \frac{\partial d(\theta, n_p^B)}{\partial n_p^B} d\theta. \tag{19}$$

This condition is similar to condition (19) above but exhibits two crucial distinctions:[59] On the one hand, delay costs and flight benefits are now evaluated at θ^D rather than at θ^*. As explained above, this ensures that all effects from peak-business output variations are captured accordingly. On the other hand, condition (19) now includes the network density benefits. More precisely, the left-hand side contains the network value. The integral on the right-hand side, in turn, accounts for the marginal network density benefit from a network expansion that accrues to all network users. In the same fashion as above, condition (19) dictates that the business airline's peak-period output and hence the network size are optimized when the overall net benefits equal all delay costs plus the marginal network density benefits.

Observe that the overall network density benefits are strictly increasing with peak-business output because the partial derivative on the right-hand side of (19) is monotonously increasing in n_p^B by definition and because a corresponding decrease in $\theta^D = \Theta - n_p^B$ shifts the lower bound of the integral to the left. Strict concavity of $d(\theta^D)$ thus implies single crossing

[58] This rule is well known from economic theory, as it generally applies to any cost-benefit problem.

[59] Formally, the integral remains after derivation because network density benefits are multivariate in both the critical $\theta's$ and in peak-travel preference. For interpretive purposes, the partial derivation operator has been changed to ∂n_p^B according to ${}^{\partial d(\theta, 1-\theta^D)}/_{\partial\theta^D} = {}^{\partial d(\theta, n_p^B)}/_{\partial(1-n_p^B)} = -{}^{\partial d(\theta, n_p^B)}/_{\partial n_p^B}$ with the corresponding change of sign.

between both sides of (19), so that a unique, strictly positive solution $\theta^D > 0$ exists that fulfills condition (so.ii) and thus makes (19) hold with equality. Once again, if this solution complies with the population range, it denotes the globally optimal network size. If not, it is replaced by one of the respective corner solutions. The relevant corner solutions and their respective constraints are analyzed in the following subsection.

7.5 Corner Solutions

For consideration of the corner solutions, recall first that according to the above ordering rule θ^D is dominant both against θ^* and $\underline{\theta}$. This means that toward the left-hand side of the consumer continuum, θ^D is only bounded by $\Theta - 1$. Equivalently, the upper bound for θ^D toward the right-hand side is Θ. Moreover, θ^D is uniquely determined by equation (19). Consequently, the corner solutions for θ^* and $\underline{\theta}$ need to be evaluated hierarchically, with the solution of θ^D as a precondition.

Therefore let us first consider the potential corner solutions for θ^D by supposing that the network density benefits are important relative to direct flight benefits and congestion costs. In this case, the extent of the network benefits for all passengers on the right-hand side of (19) increasingly disbalances the left-hand side of (19). If equivalence across the two sides is no longer met, θ^D needs to be limited to $\theta^D = 1 - \Theta$, which yields an optimal network size of $\Theta - \theta^D = 1$. The network hence extends to the entire passenger continuum and dominates the entire market, although condition (17) suggests an interior solution for $\underline{\theta}$.[60]

By contrast, imagine the opposite case where congestion costs are prohibitively high but flight and density benefits very weak so that serving the peak-period is not worthwhile at all. In that case, the planner may want to technically allocate a negative share of output to the peak period. As a consequence, based on the non-negativity constraint $\theta^D \leq \Theta$ from the above boundary conditions, both airlines' peak-period outputs would remain at zero so that $\theta^D = \Theta$. Note, however, that both these corner solutions are not likely to occur as they require an important disbalance between delay costs and the benefits from flights and network density. Therefore, they are of no particular value for the investigation and are hence not further discussed.[61] For the remainder of this part, an interior solution for the optimal network size is thus assumed. Moreover, an interior solution for θ^D is also required in order to comply with the partial market coverage assumption.

[60] In Brueckner's (2002a, p.1362) horizontal specification, intermediate crossing ensures $b_o(\underline{\theta}) > b_p(\underline{\theta})$ and hence warrants an interior solution $\underline{\theta} < \theta^*$. In this case, this problem is omitted.

[61] For the purpose of illustration, the sensitivity of the simulation in Part III will illustrate a fully dominant network.

As for the corner solutions of θ^*, let us also assume that congestion costs are important enough that the peak period cannot dominate the entire market. This corresponds to the partial market coverage assumption for peak-period flights from Section 5.3.2. The relevant corner solution for the leisure airline thereafter concerns the question of whether the leisure airline has a positive peak-period output or whether the peak period is exclusively served by the business airline. To answer this question, compare the dominant peak-business output from (19) with the optimal peak-leisure output from (18): Due to single crossing of functions B and TG it becomes clear that the smaller difference $B-TG$ on the left-hand side of the two equations yields a lower θ and hence a governing interior solution for the corresponding first-order condition. Unfortunately, however, this relationship cannot be explored on analytical grounds in the generic model because the relative sizes of the density benefits terms in (18) are unknown.

Nevertheless, a brief consideration of the inequality in equation (18) permits us to stipulate a simple rule concerning the applicability of the corner solution for the leisure airline: Suppose that peak-business output (as independently determined from (19)) were large enough to fulfill inequality (18) with $n_p^L = 0$. Any positive $n_p^L > 0$ would also comply with the inequality but would require that (18) holds with equality according to condition (so.ii). The equation hence becomes imbalanced when corner solution $n_p^L = 0$ is governing. A rearrangement of (18) shows that the business airline needs to exclusively serve the peak period if

$$n_p^B \geq \frac{B^* - TG^*}{TG^{*\prime}}.$$

with $n_p^L = 0$ so that $\theta^* = \theta^D$. In the opposite case, n_p^B cannot make (18) hold with inequality. As a consequence, a positive $n_p^L > 0$ would be required to reach the optimal overall peak output. This, in turn, would balance equation (18) to equality, denoting an interior solution where the peak period is served by both airlines so that $\theta^* < \theta^D$. Use of the above condition in (18) and re-arrangement hence yields optimum peak-leisure output as an implicit function

$$n_p^L(n_p^B) = \max\left[0, \frac{B^* - TG^*}{TG^{*\prime}} - n_p^B\right]. \tag{20}$$

Formally, the social optimum is derived by first solving equation (19), which independently yields the optimal peak-business flight volume. If $n_p^L = 0$ subsequently fulfills condition (18), the corner solution becomes governing. By contrast, if the inequality is not fulfilled, then condition (18) holds with equality and can be solved for a positive $n_p^L > 0$. Finally, substitution of peak-flight volumes into (17) yields the overall off-peak-period output.

Concerning the off-peak period outputs, a fully dominant network would nullify all other individual outputs so that $n_o^L = n_o^B = 0$. As stated above, however, this case is ruled out by assumption. Partial market coverage in the peak period thus generates positive off-peak-period outputs the social optimum, which yields $n_o^L = n_o^B > 0$ due to symmetry of (17). According to (7) the off-peak-period flight fare remains at marginal cost.

7.6 Second-Order Conditions

In analogy to the equilibrium, Appendix B.1 shows the welfare function to be strictly concave for a valid solution set $\theta^{OPT} = (\underline{\theta}, \theta^*, \theta^D)$ that complies with the above boundary conditions $\underline{\theta} < \theta^* \leq \theta^D \in [\Theta - 1, \Theta]$. In accordance with the above reasoning, this valid solution is assumed to be an interior solution with partial market coverage $\underline{\theta} > \Theta - 1$ and a non-negative network size $\theta^D < \Theta$. This interior solution is independent of the outcome of peak-leisure output rule (20), so that $\theta^* \leq \theta^D$ is sufficient. The valid solution above hence represents a unique global maximum and, therefore, constitutes the social optimum.

8 Model Properties

This section extends the model analysis by revealing the three most important characteristics of the model at hand: The monotonicity of the network density benefits from a passenger or social perspective, the concavity of the network value from an airline view, and the conjectural variations of both airlines' outputs following an asymmetric capacity constraint. In addition, the last subsection explains the model calibration for partial market coverage.

8.1 Monotonicity of Network Density Benefits

Firstly, recall that the network density benefits have been defined as monotonously increasing both with peak-period preference as well as with network density (see Section 5.2.2). Moreover, utility equation (1) shows that peak-period flights of the business airline generate the same direct flight benefits at the same amount of congestion as the peak-leisure flights but in addition provide density benefits.

From a welfare perspective, this has the following implication: On the downside, any increase of peak-period output generally increases congestion. On the upside, the increasing flight volume generates peak-period travel benefits for a higher number of travelers and decreases the market power distortion. Whether a peak-period output expansion is worthwhile therefore depends on the optimal ratio of flight benefits and congestion. This ratio is indicated by social optimum condition (18), as pointed out above. However, along with the optimal ratio, the respective market shares also have an impact on welfare: As peak-period condition (20) shows, the leisure airline must only have a non-zero peak-period output if the optimal network size has been reached. However, the above observation about the definition of the network density benefits reveals that replacing peak-period output flights of the leisure airline with flights of the business airline monotonously increases network size and thus gross utility, while congestion remains constant. This effect is henceforth referred to as the monotonicity of the network density benefits.

Subsequently, given any optimal amount of peak-period output, the monotonicity of the network density benefits implies that net utility and thus social welfare are always higher when this output is exclusively provided by the networking airline. The above reasoning hence dictates that corner solution $n_p^L = 0$ must always prevail in the social optimum so that $n_p^L > 0$ can never occur. This model property could not be inferred from social optimum condition (20) and therefore further narrows down the above generic result from Section 7.5. It ultimately dictates that the interior solution from condition (20) will never take place. As a consequence, the leisure airline is completely expelled from the peak period, and the critical thetas simplify to $\theta^* = \theta^D$.

8.2 Concavity of Network Value

As the social optimum computation has shown, the network density benefits are monotonously increasing in network size, so that the business airline's output unambiguously increases utility for all passengers and hence strictly has a net positive effect on overall welfare. By contrast, the derivation of the equilibrium can be used to show that from an airline perspective, the network density benefits function becomes a one-dimensional function of peak-period output. This one-dimensional function denotes the value of the business airline's network, rather than the latter's concise benefits for each passenger; therefore, it is referred to as the network value. The equilibrium derivation shows that it is the network value that is relevant for the airline's profit maximization, rather than the network density benefits function. As the network value can be shown to be a concave function of peak-period output, the resulting model property is referred to as the concavity of the network value.

Formally, the concavity of the network value can be derived as follows: First, according to inverse demand function (7), the business airline can choose its profit maximizing network size n_p^B with regard to the propensity to consume θ. This output choice determines the corresponding peak-business flight fare. However, because price discrimination is not available to the business airline, only the marginal willingness to pay of its left-most passenger on the θ-scale is relevant for the airline's profit maximization. The critical consumer θ^D thus denotes the equivalence between peak-period travel with the leisure and the business airline. As a consequence, based on equivalence $\theta^D = \Theta - n_p^B$, the relevant peak-travel preference directly relates to the business airline's output choice (see 6.1). Therefore, in the business airline's optimal peak-period output decision, the two-dimensional network density benefits function from (1) reduces to a one-dimensional function

$$d(\theta^D) \equiv d(\theta, n_p^B)|_{\theta=\theta^D}.$$

This function formalizes that the business airline can only commercialize its network based on the network density benefit of critical passenger θ^D. The term $d(\theta^D) = d(\Theta - n_p^B)$ in equilibrium condition (10) subsequently reflects the commercial value of the business airline's network and hence is referred to as the network value. Ultimately, the network value function hence shows that the business airline's only variable to control for its commercial network density benefits is the network size.

Because the density benefits for passengers increase both with peak-travel preference θ and network density n_p^B, a higher network density is counter-balanced by a decreasing willingness to pay of the critical passenger. This, in turn, implies that the network value is a concave function of the business airline's peak-flight volume. This concavity generically exhibits $d(\Theta - 1, 1) = d(\Theta, 0) = 0$ (see 5.2.2), which means that network density benefits are zero at either side of the passenger continuum: To the left because the left-most individual within the population range does not value network density, and to the right because network size at $\theta^D = \Theta$ equals zero. Thus, monotonicity and differentiability of $d(\theta, n_p^B)$ imply $d'(\Theta - 1) > 0$ as well as $d'(\Theta) < 0$, and concavity yields $d(\theta^D) > 0$ for $\theta^D \in]\Theta - 1, \Theta[$ with $d(\theta^D)'' < 0$.

The network value function is illustrated as a linear example in Figure 8.1. Recalling that the network density benefits $d(\theta, n_p^B)$ are assumed to increase both with the consumer's willingness to pay θ and with network size n_p^B, the density benefits function is decomposed into two linear functions for comprehensiveness: Function $d(\theta, \overline{n}_p^B)$ denotes the potential density benefit for all passengers who may travel with the business airline, given a fixed network size $\overline{n}_p^B$. This function monotonously increases from left to right, because network density benefits strictly increase with a higher willingness to pay θ. By contrast, function $d(\overline{\theta}, n_p^B)$ depicts the density benefit for an arbitrary passenger $\overline{\theta}$ with increasing network density (where network density is approximated by peak-flight volume n_p^B, as defined above in Section 5.2.2). This function increases from right to left, because network density $n_p^B = 1 - \theta^D$ inversely relates to the willingness to pay of the critical consumer who constitutes the left-most peak-period passenger of the business airline.

Combining the two inversely related functions again and substituting equivalence $n_p^B = 1 - \theta^D$ yields the network value function $d(\theta^D)$. As already mentioned, profit maximization implies choosing a network size $n_p^B = 1 - \theta^D$ with respect to the willingness to pay of critical consumer θ^D. Based on the above equivalence, this choice is simultaneous (i.e. one-dimensional). From the inverse relationship of the two sub-functions it thus follows that the network value function is concave.

The density benefit $d(\theta^D)$ of critical consumer θ^D determines the peak-density premium, while the density benefit for every other peak-period passenger follows $d(\theta, \overline{n}_p^B)$ and hence is higher than $d(\theta^D)$. However, this additional customer value cannot be commercialized by the business airline because price discrimination is not available by assumption: Increasing the network size increases network density but shifts the critical consumer to the left. By contrast, decreasing the network size increases the critical consumer's willingness to pay, but decreases network density. These properties thus illustrate the concavity of the business airline's network value.

Finally, note that the profit maximizing network size emerges where the marginal benefit from a network expansion equals its marginal costs. As equilibrium condition (10) shows, however, this rationale does not only include the network value but also time costs, congestion costs and market power. Therefore, as shown in Figure 8.1, the optimal network size does not generally correspond to the maximum network value.

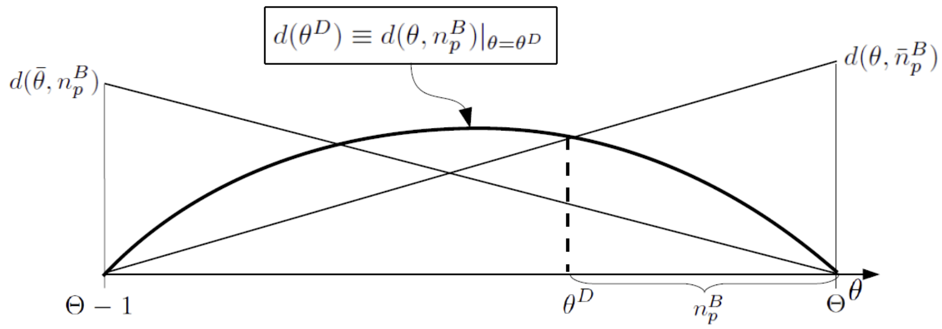

Fig. 8.1: Network Value Function

Under the above conditions, $d'(\theta) \lesseqgtr 0$ can take either sign. This means that also the sign of $d'(\theta^D)$ in equilibrium condition (10) is ambiguous. For clarity, let us first restate this term as $d'(\theta^D) = d'(\Theta - n_p^B) = -d'(n_p^B)$. Function $d'(n_p^B)$ can thus justifiably be referred to as the marginal network value. The advantage of this re-statement is that the direction of a change in network value (i.e., the marginal network value) directly corresponds to changes in network size and does not have to be inverted by using $\theta^D = \Theta - n_p^B$.

With this simplification, the following can be said about the equilibrium value of $d'(\theta^D)$: As stated above, concavity causes $d'(\theta^D) > 0$ at the far left-hand side of passenger continuum $[\Theta - 1, \Theta]$. Due to the above sign change, the marginal network value is negative by definition. This means that any network expansion decreases the network value when network size is already high. A decreasing network value, in turn, reduces the peak-density premium $f_p^B - f_p^L$. In addition, time and congestion costs strictly increase with increasing overall peak-period output. In general, expansion of an already large network therefore does not seem to be worthwhile.[62] Notwithstanding this adjustment, a large network size may be profit maximizing if the additional profits from the output expansion overcompensated the

[62] Note that the total effect of a network expansion on the full peak-period flight fare f_p^B of the business airline (and not just the premium) remains ambiguous: As the endogenous flight fare equations from inverse demand in (7) show, it also depends on the resulting output adjustment of the leisure airline and the peak-leisure airfare. The total effect of output expansion on the flight fare could hence be negative and not too large, or even positive. Moreover, the same output compensation also affects congestion and time costs: Because peak-leisure output generally inversely reacts to peak-business output, an increase in delay costs would be disproportionately lower than the network expansion itself.

lower density premium and the higher delay costs. Evidently, this could only be the case if network density benefits were important relative to time and congestion costs.

The opposite occurs at the right-hand end of population spectrum $]\Theta - 1, \Theta[$, where by definition $d'(\theta^D) < 0$. Again, due to the above sign change, the marginal network value is positive in this section. Corresponding to the above, this means that increasing output also increases the peak-density premium. As long as the leisure airline partly compensates any peak-period output expansion of the business airline, the latter can at the same time increase both its flight fare and its output within this side of the density benefits function. A network expansion must hence be profitable, unless time and congestion costs are dominant over the network density benefits and erode all additional benefits. Put differently, this means that a low peak-period equilibrium output of the business airline can only occur if delay costs are important as opposed to the density benefits function.

In sum, the above reasoning suggests that the equilibrium network size is large and thus the marginal network value is negative if the network density benefits function is important relative to the delay cost functions. By contrast, if time and congestion costs are more prominent than the density benefits the network size is likely to remain small.

8.3 Conjectural Variation: Endogenous Output Adjustments

As previously stated, in a duopoly model with endogenous demand, generally each airline's output is a reaction to its competitor's output (see Section 3.4.1). In other words, the equilibrium first-order conditions determine each firm's output change as a function of its competitor's output change in order to yield maximum profits. Note, however, that there is a fundamental difference between an unconstrained steady state and an equilibrium arising from an exogenous shock that asymmetrically constrains one competitor's output.

Formally, the firm's mutual output adjustments are explicitly determined by the slope of their reaction functions. In economic theory, the corresponding output adjustment path has been defined as the conjectural variation (see Giocoli, 2010, p.138 and p.140).[63] In this model's notation, the latter can be written as

$$n_p^i(n_p^j)' = {}^{\partial n_p^i}\!/\!_{\partial n_p^j} \equiv -\frac{{}^{\partial^2 \Pi^i}\!/\!_{\partial n_p^i \partial n_p^j}}{{}^{\partial^2 \Pi_i}\!/\!_{\left(\partial n_p^i\right)^2}}. \tag{21}$$

[63] Formally, Giocoli (2010) denotes the slope $n_p^i(n_p^j)'$ of the reaction function as $R_i'(n_j)$ and defines the conjectural variation as $\nu_{ij} = {}^{\partial n_i}\!/\!_{\partial n_j}$.

These variations denote the adjustment paths of both firms' outputs as functions of their competitor's choices until a steady state is reached where both firm's outputs simultaneously are mutual best responses and hence no further output variations occur. This steady state constitutes a Nash equilibrium, where neither firm can further increase its profit by deviating from the currently chosen output (cf. e.g. Eichenberger, 2010, p.90). Giocoli (2010, p.140) mentions that the Cournot equilibrium represents the special case of both firms' output choices where $\partial n_p^i / \partial n_p^j = 0$.

However, suppose that one airline changes its output as an answer to an exogenous shock to its profit function. In this model, such a shock may occur through the imposition of asymmetric quotas or through the introduction of a secondary trading opportunity following an asymmetric quota allocation. In such a case, the unconstrained firm adjusts its output according to its conjectural variation as long as it is not directly concerned with a constraint on its own and thus is free in choosing its output volume. The unconstrained firm's conjectural variation at its pre-shock equilibrium but with the competitor's constrained output is hence no longer zero.

According to (21), the conjectural variation describes the marginal output change that will take place subsequent to an exogenous shock. The corresponding output change of an unrestricted airline i is captured by the term

$$\mathrm{d}n_p^i = \partial n_p^i / \partial n_p^j \cdot \mathrm{d}n_p^j, \tag{22}$$

where $\mathrm{d}n_p^j$ denotes the output change of competitor j based on the exogenous shock, and $\partial n_p^i / \partial n_p^j$ corresponds to the conjectural variation as defined in equation 21. Although this output compensation is caused by an exogenous shock, it ultimately occurs endogenously within the first-order condition of airline i. Therefore, it will henceforth be referred to as the *endogenous output adjustment*.

Finally, note that any constrained airline cannot further adjust its own output as it has been constrained by definition. Therefore, the endogenous output adjustment according to the conjectural variation only concerns an unconstrained firm after introduction of an exogenous shock that alters its competitors output. It occurs until the conjectural variation again equals zero and thus a new equilibrium is reached.

This effect will come into play in the analysis of the allocation schemes in Part II whenever a quantity constraint is imposed that is binding for one airline only. The transmission path of the endogenous output adjustment is explored in more detail by evaluation of the specific first-order conditions within the quota case in Section 11.

8.4 Partial Market Coverage

The condition for partial market coverage is defined ex-post by definition of the model parameters in terms of the equilibrium first-order conditions (see 5.3.2). This subsection provides the corresponding formal evaluation, which has been adopted from the fundamental vertical differentiation models of Lambertini (2006, p.164) and Choi and Shin (1992, p.231).

Partial market coverage requires that the left-most individual within the defined range $\theta \in [\Theta - 1, \Theta]$ have a non-positive utility from a flight, so that he does not travel. In terms of model notation, this presumption can be formally denoted as $\underline{\theta} > \Theta - N$. Now, from the substitution of inverse demand equations (7) into the utility definitions (1), it is known that $f_o = b_o(\underline{\theta})$ in equilibrium. Using equivalence $\underline{\theta} = \Theta - N$, the above presumption can hence be written as requiring

$$f_o > b_o(\Theta - N).$$

In a specified model, the equilibrium values could now be used to explicitly find the critical size of population wealth parameter Θ that would yield partial market coverage. In the basic model, however, this condition unfortunately cannot be evaluated further because the generic functions do not allow the above implicit condition to be made explicit. The inequality lets us at least conclude that the critical Θ is a maximum rather than a minimum. This means that the population's propensity to consume must not be exceedingly large in order to yield partial market coverage.

Alternatively, the above condition can also simply be stipulated as

$$N(\Theta) < 1. \tag{23}$$

In an explicit model, the sum of all equilibrium outputs can hence be expressed as a function of Θ and by calibration of the latter be set to any positive value less than unity. In the generic model, however, total output $N(\Theta)$ is again an implicit function and thus cannot be resolved explicitly. Nevertheless, the implicit determination by equations (8) to (10) allows us to constitute the following relationships: For partial market coverage Θ can higher if marginal costs c are higher, delay costs $T(\cdot)$ and $G(\cdot)$ become more important, and marginal benefits $b'(\cdot)$ and $d(\cdot)'$ are steeper. By contrast Θ has to be lower when the benefit functions $b(\cdot)$ and $d(\cdot)$ are more important.

9 Discussion

The following discussion first descriptively reviews the model results of both the equilibrium and the social optimum from Sections 6 and 7. Subsequently, it reveals the inefficiencies that arise from the two market distortions in the unconstrained equilibrium by comparison of those two outcomes. Finally, the last two subsections briefly recap the three characteristic model properties arising from those results, as shown in Section 8, and stress the dissociation of this study from Brueckner's (2002a) original work.

9.1 Equilibrium

In the unconstrained COURNOT equilibrium, airlines maximize their profits given the endogenous demand functions of flight fares. This means that they take into account the effect of their output choice on market prices. The airlines' output choices are then determined by mutual reaction functions that determine optimal flight volumes as a best response to each competitor's output.

The above determination of the unconstrained equilibrium shows that the individual market shares of both airlines depend on the importance of the network density benefits relative to the dual distortion from congestion and market power: If density benefits are moderate relative to flight benefits and delay costs, the equilibrium is an interior solution with positive outputs for both airlines in both periods. In this case, the market is dominated by the business airline both in terms of peak-period output and flight fares and therefore also in terms of profits. Overall output, however, remains symmetric across airlines, which means that any output change in one period is offset by the respective airline in the other period.

By contrast, if network density benefits are relatively important, a corner solution arises where the business airline's peak-period output rises above total output of the leisure airline: Although the business airline's off-peak-period output is restricted to zero, total outputs are no longer symmetric. This means that the corner solution extends the asymmetry between the two airlines in terms of flight fares, outputs and profits from the peak-period to the entire market. As mentioned above, however, a degenerate equilibrium with extreme density benefits and thus market preemption by the business airline is ruled out by assumption.

It is worth pointing out that the business airline's network size decision has a market-power effect: it determines the network density premium that can be earned as a mark-up against the peak-period flights of the leisure airline. However, this mark-up follows the concavity of the network value and hence does not necessarily work in the same direction as traditional market power, which yields a premium based on an output reduction. As a consequence, the density premium constitutes a market power effect that is based on product differentiation and on the endogenous choice of product heterogeneity in terms of the airline's network density. Correspondingly, the network quality choice of the networking airline is based on demand and exactly represents the kind of demand-side endogenous market power that was originally proposed by Berry (1990).

9.2 Social Optimum

The social optimum is generally characterized as follows: Overall output is determined according to the relative importance of delay costs and flight benefits. In this respect, the optimum conditions require that the output-increasing congestion externality and the output-decreasing market power distortion cancel each other out, so that the dual distortion becomes absent. Note that, in general, the optimal amount of congestion costs is therefore not zero.

Subsequently, the optimal market shares are the following: The optimal peak-business output depends on the importance of the network density benefits against delay costs and direct flight benefits. When the network effects are not important, this yields a small network, which may even allow an interior solution where both airlines participate in the peak period. By contrast, a corner solution where only the business airline provides peak-period flights arises if the network density benefits are important. As a consequence, the optimal peak-leisure output is a function of peak-business output.

Ultimately, the social optimum reflects a tradeoff between asymmetric network density benefits and congestion costs. The model therefore substantially contrasts with Brueckner's (2002a) case, where only symmetric direct flight benefits are weighted against flight delays. In this respect, also note that the network density benefits do not increase peak-period flight demand relative to the socially efficient level for their own sake. Rather, they raise the overall level of the socially optimal output in comparison to a setting without network density benefits. In contrast to the congestion externality and market power, thus, the network density benefits do not constitute an output distortion.

9.3 Allocation Efficiency

The results from the generic model show that unconstrained equilibrium output is distorted both by market power and by non-internalized congestion. On the one hand, the traditional market power effect decreases output and increases flight fares against the social optimum. On the other hand, the external part of congestion allows total peak-flight volume to be higher than justified by marginal social costs alone. In analogy to Brueckner (2002a), the dual distortion hence applies to the equilibrium. Consequently, the deviation of the equilibrium outputs from their socially optimal values depends on the size of the market power distortion relative to the congestion externality.

In addition to overall output, efficiency also requires the provision of the optimal network size. The optimal network size is computed by the social planner, who maximizes the sum of network density benefits for all peak-business passengers (see 7.4). However, the social optimum target function differs from the airlines' profit maximization also with respect to the network size: as pointed out in Section 8.1, the network benefits for all passengers monotonously increase with network size. By contrast, the business airline optimizes its network value rather than the benefits integrated across all passengers (see Section 6.4). Therefore, the network value is a concave function of output (see Section 8.2). Consequently, the market power effect in terms of endogenous pricing also applies within the network premium and thus affects also the network size. As a result, the concave network value always maintains the network size below its socially optimal value.

Note that this network undersize occurs despite the fact that congestion is partly external in equilibrium (whereas congestion is fully internalized in the social optimum): If the business airline would exclusively serve the peak period, it would choose its monopoly output and thus fully internalize congestion. Both due to the downward sloping demand function and due to the concavity of the network value, this monopoly output must naturally remain smaller than the social optimum. However, as both airlines serve the peak-period in equilibrium, the overall peak-period output will be higher than the business airline's monopoly output (see Mas-Colell et al., 1995, pp.384, 385 and 393). As a consequence, the business airline's duopoly peak-period output needs to be lower than its own monopoly output. This result arises from the conjectural variation because the higher overall duopoly output both decreases the flight fares and increases the negative congestion externality (see Section 8.3). In comparison to the social optimum, therefore, the network must always remain undersized in equilibrium.

As the monotonicity of the network density benefits yields that the leisure airline must never serve the peak period in the social optimum, the above result allows us to reduce the optimality criteria for efficiency to one single condition: reaching the optimal peak-period output of the business airline in order to optimize the provision of network utility against the delay costs from congestion. Because both airlines serve the peak-period and the network size always remains inefficiently low, allocation efficiency in the unconstrained equilibrium is thus degraded in two dimensions. As a consequence of those considerations, only two distinct outcomes may arise in equilibrium: Firstly, both overall output and the network size may be inefficiently low, which happens if the market power distortion is large in relation to the congestion externality. Secondly, overall output may exceed the socially optimal level while the network size would still remain inefficiently low. In this case, the congestion externality would be more important than the market power distortion. Consequently, the network size can never exceed the socially optimal value because the endogenous output adjustment, which caused by the negative externalities from the leisure airline's positive peak-period output, reinforces the market power distortion even if overall output is excessive.

Although the generic model does not allow further revelation of the relative size of the two individual effects within the dual distortion, plausibility indicates the following: As Brueckner (2002a) mentions, the importance of the market power distortion depends on the number of firms in the market. If the number of firms is low, the market power distortion prevails. By contrast, in a highly competitive market, the congestion externality is more important because the overall amount of delay costs accounted for by each firm is very small. Consequently, the low number of firms in the duopoly should manifest in an allocation where the market power distortion is more important than the congestion externality. This signifies that overall output is also likely to be inefficiently low in equilibrium.

Overall, the degree of the inefficiencies arising in the unconstrained equilibrium depends on the dual distortion and on the formalization of the network density benefits. In turn, both the direction and size of the dual distortion as well as the degree of the network undersize depend on the market structure that actually prevails in the flights market. This signifies that the efficiency results of both the equilibrium as well as the corresponding capacity allocation instruments ultimately hinge on the assumption about the market form. As a consequence, the presumption of traditional market power and its relative importance in a duopoly will critically affect the results. The traditional market power assumption with downward sloping demand is discussed as a main limitation in Section 23.2.

9.4 Model Properties

The above discussion shows that the model at hand displays three characteristic model properties: The monotonicity of the network density benefits from a welfare perspective, the concavity of the network value from an airline perspective, and the conjectural variation in equilibrium outputs based on the airlines' reaction functions. Partial market coverage represents a fourth property but not an innate model characteristic, as it ultimately arises by calibration.

The comparison of the equilibrium and the social optimum shows that the network density benefits have distinct properties, depending on whether they concern the passengers' utility from a welfare perspective or the airlines' profits from a stakeholder perspective: From a welfare perspective, utility from network benefits is monotonous in the network size. This finding further specifies the general rule of peak-period leisure output as a function of the business airline's peak-period output; the monotonicity of the network density benefits thus dictates that the leisure airline must never serve the peak period.

By contrast, the benefit of network density with regard to output for the business airline's profits is characterized by concavity. This result is surprising as the density benefits have originally been defined based on the same functional form. As a consequence, the concavity of the network value yields an interior solution for the business airline's network size in equilibrium although density benefits are monotonously increasing with peak-flight preference.

Moreover, the conjectural variation denotes each airline's optimal output adjustments following any competitor's output change according to its reaction function. In equilibrium, this variation obviously is zero. However, if quantity constraints are imposed that are binding asymmetrically (i.e. for one airline) only, this conjectural variation generally differs from zero for the unconstrained airline. Consequently, there is an endogenous output adjustment following the imposition of an asymmetric constraint, which needs to be reflected in the welfare analysis.

Let us once again emphasize that the above three properties arise in addition to the standard features of the oligopoly model with externalities, which typically includes the dual distortion with market power and congestion only. They hence particularly characterize the generic model presented in the study at hand as they are explicitly based on the introduction of the network density benefits and the corresponding airline asymmetry. Therefore, these three model properties are also crucial to the subsequent investigation and welfare analysis of the three distinct allocation schemes in Part II.

Lastly, partial market coverage is ensured by limiting population wealth to a value that yields an overall output smaller than unity. By contrast to the above characteristics, however, this fourth property is directly based on model calibration (by assumption; see 5.3.2). In this respect, the previous investigation shows that for partial market coverage to hold, the population can be wealthier if flight benefits are relatively low and operating and delay costs are relatively high. Inversely, the income of the population considered must be all the lower, the higher the benefits and the lower the costs associated to flight services are. Consequently, the market is fully covered if the population is too rich, while it is completely abandoned if the propensity to consume is too small; in both cases, partial market coverage is not met.

As already mentioned, this calibration is thus important for the quantitative evaluation of the model in Part III. For the analytic investigation of the generic model in Part II, however, it is sufficient to simply assume partial market coverage, which does not have any further implications.

9.5 Dissociation from Brueckner (2002a)

As previously stated, this study's model is formally based on and fundamentally inspired by Brueckner's (2002a) oligopoly setting. Two main innovations, however, qualify this model as a unique contribution to the economic discussion.

From a contextual view, the main innovation consists in the introduction and specification of network density effects as indirect benefits for passengers. The corresponding exogenous airline asymmetry directly follows from the assumption that these network benefits can only be provided by the business airline. As a consequence, the asymmetric indirect travel benefits based on utility from network services introduce product differentiation across the two airlines.

From a technical perspective, product heterogeneity across airlines corresponds to product differentiation by quality and therefore is of the vertical type. By contrast, Brueckner (2002a) considers horizontal differentiation between the two different flight periods, which is based on heterogeneous consumer taste. The firms that provide these benefits, however, are symmetric, so that competition between firms is based on homogeneous goods.

Within the above technical distinction, market coverage is based on two distinct rationales: In the horizontal differentiation setting, where heterogeneity is based on consumer taste, partial market coverage in the peak-period is achieved by definition as long as some individuals prefer the off-peak period in terms of net benefits. Contrastingly, vertical product differentiation requires heterogeneity across consumer income in order for demand to justify the choice of distinct product qualities in equilibrium. As a consequence, market coverage becomes a function of population wealth. For the analytic investigation of the vertical setting, it is also sufficient to simply assume partial market coverage. For a quantitative evaluation, however, the population range needs to be calibrated in order to yield the desired market coverage (see Section 5.3.2).

In addition to the above conceptual distinctions, there are also substantial differences between the two studies with respect to the analysis of the allocation instruments: Brueckner (2002a) considers all different market forms ranging from perfect competition to a price discriminating monopoly but only investigates a congestion pricing scheme in terms of capacity allocation. By contrast, this study also comprehensively investigates a quantity constraint by means of airport quotas and a corresponding secondary trading scheme.

Ultimately, the consideration above shows that this study is formally based on Brueckner's (2002a) seminal work but exhibits substantial differences both in instrumental and conceptual terms by featuring network density benefits. It therefore constitutes a unique contribution to the discussion in the current literature.

10 Review of the Assumptions

This section first comments on the three main presumptions of this model consisting of product differentiation based on product quality, the exogenous airline asymmetry, and product quality based on network density. Subsequently, it discusses the consequences of these presumptions on product quality choice for both airlines. Finally, the specifications of the cost and benefit functions are reconsidered, along with a justification of the partial equilibrium perspective.

10.1 Product Differentiation

Traditionally, airlines have segmented their markets on each route by trip purpose such as business and leisure because separate market segments respond differently to supply variables such as flight frequency, departure times, fares, or external economic conditions. Product differentiation hence allows airlines to exploit their market power by endogenous (i.e., differentiated) pricing (cf. Doganis, 2002, p.188).

10.1.1 Passenger Heterogeneity

According to the theory of imperfect substitutes, product differentiation requires a corresponding demand system with heterogeneous preferences (cf. Vives, 2001, p. 143). For this purpose, air traffic management literature commonly dissociates leisure and business travelers. This distinction is based on different time values: Leisure travelers have low values of time but display high price sensitivity. As an extension to this argument, this study assumes that they only value the direct utility from transportation but not any associated network benefits or services. Business travelers, instead, are characterized by high values of time and hence display much lower price sensitivity. Again, it is furthermore implied that these passengers have preferences for network benefits as they offer ample travel choices and thus greatly increase travel flexibility.

The above demand structure is captured by the discrete choice model with different demand functions presented in Section 5.2. In view of heterogeneous preferences, this model is consistent with the representative consumer approach (Vives, 2001, p.144). A comprehensive overview of further discrete customer demand models for air travel is provided by Garrow (2010).

10.1.2 Exogenous Airline Asymmetry

The airlines are referred to as the leisure and the business airline. This nomenclature is chosen in correspondence with the traveler typology in order to stress that there is not only heterogeneity in consumer tastes but also in product quality. It may sound unfamiliar because it draws on the attributes that usually refer to passengers rather than firms. As an alternative, it may be expressed as the networking and the non-networking airline. Regardless of the nomenclature, however, this typology essentially reflects the market segmentation that arises from the exogenous airline asymmetry.

The exogenous airline asymmetry both represents the innovation of this study as well as the main contrast to Brueckner's (2002a) model. It may be interpreted as follows: While the business airline offers numerous travel options based on its high flight-frequency network structure and associated network services, it may also be assumed to provide full travel flexibility by allowing for free short-term re-booking. By contrast, the leisure airline does not offer any indirect benefits aside from direct utility from transportation as it neither provides a network structure nor offers network services. Correspondingly, one may also suggest that it allows its passengers traveling on the ticketed flight only but does not offer a re-booking possibility.

There might be objections as to how exactly justify such a clear-cut market separation. Especially, such objections seem logical given that new entrant competitors have also been engaging in network structures, targeting business travelers and seeking economies of scale. As a result, the business models of traditional incumbent network airlines and new entrant competitors may have become intermixed.[64] Nevertheless, a certain degree of this strong market separation might be explained by the large traditional network airlines' fundamental resource advantages, which arose from their former status as government controlled and owned monopolists and allowed them gathering an important degree of grandfathering rights and building their network structures over the long run. This notion is supported by the unwarranted but supposedly undoubted claim that large network structures of major airlines only became possible due to their historical dominant airport presence rather than their market entry as a competitor. As a consequence, the business airline as a large, incumbent airline may arguably have accumulated an important first-mover advantage. By contrast, new entrants and competitors simply lack this starting advantage. In addition, they may have been facing substantial barriers of entry and capacity constraints stemming from capacity

[64] I thank referee Sven Maertens (DLR) for this comment at the GARS student researchers workshop 2012, and Armin Schmutzler (University of Zurich) for an extensive discussion of this topic.

shortage and the grand-fathering element of the current airport capacity allocation scheme in practice.

In general, the exogenous airline asymmetry may hence be considered as an extreme point concerning market separation and product differentiation opportunities. Nonetheless, it serves as an illustrative, stylized case of the current real market structure and is further discussed as the second main limitation of this model at the end of the study.

10.2 Network Density Benefits

10.2.1 Network Density as Product Quality

As O'Connell (2006, p.68) states, there is intense competition between network airlines and low-cost carriers. Closely associated with this notion is this model's exogenous airline asymmetry, which is derived from Joppien's (2003, pp.120) distinction of different airline types: high-yield or full-service carriers and low- or no-frills-carriers. The full-service type characterizes an airline that provides a vast network and offers associated services and a broad spectrum of other amenities. These so-called frills may range from ground baggage service up to inflight catering (Joppien, 2003, p.121). Correspondingly, the no-frills airline type constitutes a carrier that focuses on straight transportation services but refrains from offering any additional passenger services. Based on these arguments, I assume that the distinction between the full-service and the no-frills airline type in this model corresponds to the differentiation of identical products as described by Friedman (1983) and relating to Wöckner's (2011) notion of heterogeneity by product image.[65]

As previously discussed, the airline dissociation in this model is extended in that the full-service carrier is assumed to provide a single-hub network with the associated indirect travel benefits from connectivity, whereas the simple no-frills carrier is assumed to provide simple, independent point-to-point connections only. The leisure airline thus represents a no-frills airline and the business airline represents a full-service carrier. As a consequence, the product qualities among the two are distinct, as imposed by their particular business models. Ultimately, this reasoning justifies that peak-period flights of the business airline are high-quality products whereas the same flights of the leisure airline are low-quality products.

[65] The competition analyses of Barbot (2006), Barbot (2006b), Barbot (2008) or Alves and Barbot (2010) assume that low-quality, no-frills airlines serve secondary airports only; thus their passengers are concerned with more cumbersome airport access. Although this complication may more realistically reflect the market structure, it is left out in this study for simplicity.

10.2.2 Density Benefits Function

Generally, there are two options for implementing network density effects. Modeling network density as a function of the networking airline's output or market share corresponds to a simple approach. The corresponding product quality represents an abstract concept, which would account for higher premiums in an implicit way. Put differently, this simple mechanism reflects the notion that a higher market share translates into a higher network value, either in terms of scheduling or travel time advantages but without explicit modeling of the network itself. By contrast, a more advanced implementation would require making the effects of network density explicit. This could be achieved by modeling the flight frequency on particular routes, the flights' time of day, the schedule delays and the associated network optimization and services in detail. Moreover, decomposition of airport capacity into single aircraft movements would allow a more accurate modeling of peak- and off-peak periods as the typical traffic banks of network airlines, which occur at multiple times during a day (see e.g. Brueckner, 2002b, p.4). Lastly, accounting for distinct types of aircraft would allow an examination of the airlines' problem of having to choose between aircraft size and flight frequency (such as, e.g., in Brueckner and Zhang, 2001; Brueckner, 2004; Wei and Hansen, 2007; Givoni and Rietveld, 2009).

On the one hand, an advanced setting would greatly enhance the investigation of the airport capacity discussion; e.g., Hsiao and Hansen (2011) provide an empirically supported model for passenger demand based on distinct "instrumental variables" such as flight fares and frequencies and flight connection properties, which also introduces the indirect utility for potential travelers. On the other hand, these complications increase the complexity of the analysis and substantially increase the scope necessary to cope with the research question at hand. In other words, this degree of detail does not seem to be essential to study the general impact of network density benefits on airport capacity allocation. Moreover, a number of recent contributions from the literature generally seem to support the simple approach: Berry (1990) provided an early model with network density distinguished by airline, based on the respective market shares, referring to this as to the „airport presence" of the airlines. Brueckner (2002b), who extends Brueckner's (2002a) investigation to an airport network still applies the same peak- and off-peak-period model, which does "not explicitly capture the traffic patterns as actual hub airports" (Brueckner, 2002b, p.4). Similarly, Czerny (2010) computes indirect passenger benefits from choosing among flight options as a function of an airport's total flight volume and congestion at an airport. In that, his specification provides a basic formal justification for the existence and quantification of network density benefits, even though it does not distinguish between airlines (see also 4.2).

As a consequence, the above contributions from the literature support the notion of indirect demand-side travel benefits based on the density of a network and justify a simple specification of the corresponding micro-foundation of that density. Moreover, this model's generic benefits function is strictly increasing in the number of network users (see Section 5.2.2). It thus does not contradict the usual S-form of utility from network goods, as presented in Section 3.1.2. Although its functional form remains undetermined, it could easily be specified to replicate the latter. Ultimately, the results above confirm that even a simple, abstract implementation of network density benefits yields an interesting airline asymmetry.

10.3 Product Quality Choice

Product quality choice is characterized by three main properties in this model: Firstly, the business airline can only determine its product quality based on its network density. Network density, in turn, is a direct function of peak-period output. Consequently, product quality depends on the airline's output choice, and thus is not represented by an independent strategic variable. Put differently, the business airline's network quality choice becomes intermixed with its output decision. Secondly, the exogenous airline asymmetry yields that only the business airline can indirectly control its product quality, while the leisure airline must accept a fixed product quality. As a consequence, the model only allows for endogenous product differentiation of the high-quality airline, while the low-quality competitor remains with an exogenous product quality. These three characteristics are explained in more detail below. And thirdly, product quality choice does not imply any quality costs; rather, quality costs are indirect and coincide with the standard operating and congestion costs for output. Lastly, note that the absence of an independent strategic quality variable constitutes the third main limitation of this model.

10.3.1 Business Airline: Dependent Quality Choice

Network density benefits should reflect a high degree of travel flexibility and connectivity within a network. As already discussed, in reality they may consist of, e.g., ample route and re-scheduling choices, high flight frequency on routes or low schedule delays. In this model, however, the density benefits are assumed to abstractly depend on network density, which itself is approximated by the peak-period output of the business airline. This means that an airline's quality choice is not an independent strategic variable but a mere consequence of the airline's output choice. As a consequence, the network airline's output choice becomes an intermixed decision that has to account both for traditional market power and for the network density benefits.

The output decision hence affects prices based both on market power and the network value and thus induces two counter-balancing effects: On the one hand, the output choice based on traditional market power accounts for the endogenous flight fares, which are determined by inverted demand. This effect generally tends to induce an output-contracting tendency as compared to a competitive setting. On the other hand, the network density decision also affects flight fares because the latter also depend on the network value. However, as the network value is a concave function of output, its impact on flight fares is generally ambiguous. Therefore, the airline's decision rationale with regard to product quality cannot ultimately be separated from its output choice based on traditional market power. As a consequence, this dependent or intermixed representation of product quality choice may be deemed unsatisfactory.

When product quality is designed as an independent variable, however, the model becomes a multistage game. Even without quality costs, the equilibrium is very unlikely to yield analytic closed-form solutions. As Lambertini (2006, p.165) points out, convex cost functions lead to polynomial terms within the backward induction and thus require numeric solutions.[66] Convex time and congestion costs also induce this complication when evaluated explicitly. Even linear specifications of the latter may at least lead to quadratic terms if congestion is partly external and thus the congestion costs are multiplied with each airline's own flight share (see the Simulation in Part III). Therefore, I deem this simplified specification as acceptable and suitable for illustrative purposes. This view may be supported by Basso's (2008) setting, which neither distinguishes flight frequency and output.[67]

Despite the rather basic specification of the network density benefits function and the subsequent dependent quality choice, this simple setting provides an illustrative generic representation of indirect network benefits for airline passengers. These benefits, in turn, allow to capture the general properties of a networking airline and their impact on the subsequent welfare analysis of the different capacity allocation schemes. Ultimately, the simplification of product quality as a dependent variable comes at the advantage of closed-form solutions and thus the possibility of analytical investigation of the model.

[66] See, e.g., Motta (1993) for a comprehensive, generic numerical solution both to quantity and price competition. By contrast, both the theoretical models by Lambertini (2006) and Ecchia and Lambertini (2006) yield analytic solutions while providing an explicit, independent single quality variable. In their case, however, there are no externalities and no convexities in the cost functions.

[67] I thank Achim Czerny for this critical remark and the reference to Basso (2008) at the Bremen GARS Workshop 2014.

10.3.2 Leisure Airline: Fixed Product Quality

From a technical perspective, the exogenous airline asymmetry yields that only the business airline can actually influence its product quality, whereas the leisure airline rests with a fixed quality. In other words, the high quality firm faces an endogenous product quality, while the low quality firm has to take its product quality as exogenously given. This contrasts with common vertical differentiation models, where quality levels are either fixed or endogenously determined for both firms.

Nevertheless, the current setting still associates to vertical differentiation models without quality costs (as e.g. Choi and Shin, 1992; Wauthy, 1996), or with exogenous, fixed quality costs (Lambertini, 2006). In these models, the high-quality firm chooses its optimal product quality, while the low-quality firm's quality becomes a best-response function. The first market entrant is simply assumed to become the high-quality firm. This again supports the above notion of the exogenous airline asymmetry being a first-mover advantage. Therefore, the current model can be regarded as a similar setting, where the high-quality business airline chooses its degree of product differentiation by means of the magnitude of its density benefits, and hence product quality. The leisure airline, in turn, represents a corner solution, with the low product quality as a fixed value.

10.3.3 No Cost of Quality

In this model it is assumed that an increased product quality induces no corresponding cost to the airline. Put differently, the fundamental quality difference between business and leisure flights does not imply a cost implication of any sort. At first sight, this assumption may seem to be highly simplifying and therefore quite controversial. This may especially be the case as only one firm has an endogenous quality choice, and hence can yield a competitive advantage for free. As a consequence, the symmetric cost structures may lead to an overestimation of the degree of product differentiation.

As already mentioned above, however, the abstraction from quality costs has been common within the basic vertical differentiation models from literature. It helps to abstract from additional cost side issues, at the benefit of an illustrative analytic solution that focuses on the demand-side heterogeneity. On this matter, note that both convex congestion costs and linear operating costs occur with an output expansion but may not be used to approximate quality costs, as they accrue to both airlines symmetrically.

In models where quality costs are fixed or absent, cost considerations do not enter the first-order conditions. As a consequence, the profits of the high-quality firm are strictly increasing with regard to quality. Therefore, product quality needs to be exogenously bounded in these models, as the high-quality firm's best quality choice would otherwise become infinite (see Lambertini 2006, p.165, Choi and Shin 1992, p.231 and Wauthy 1996, p.348). In this model, by contrast, the network value function represents the value of product quality for the business airline. As it is a concave function of output (see Section 8.2), the business airline's quality choice does not yield infinity or becomes a corner solution with full market preemption. Therefore, the above caveat does not apply, despite the absence of quality costs.

10.4 Airline Cost Functions

10.4.1 Aircraft Size and Seat Load Factor

In the airline context, when firms choose their production quantities, it is crucial to distinguish between the number of passengers and the number of flights: While consumer demand is related to passenger volume, airport demand and capacity constraints are based on the number of flights. Hence, the problem of relating continuous passenger demand to a discrete number of flight movements must be overcome.

On this matter, recent contributions from the literature suggest two distinct types of solutions: Hong and Harker (1992), Brueckner (2002a) and Brueckner (2002b) treat aircraft size as constant. Moreover, they assume that every seat in an aircraft is occupied, which returns a constant seat load factor of 100%. This implies market clearing and omits the crucial airline choice of aircraft size versus flight frequency. Passenger demand is hence simply translated into a discrete number of flights by a single, constant factor. In a distinct approach, Czerny (2010) formulates passenger benefit in terms of the overall number of flights at an airport and hence is not directly concerned with this problem. By contrast, other current studies such as Wei and Hansen (2005, 2007); Brueckner (2004); Pai (2007); Givoni and Rietveld (2009) are dedicated particularly to the airlines' problem of aircraft choice versus flight frequency. These studies model supply as the product of the number of flights on a route and the number of seats per aircraft. The subsequent matching of demand does not strictly require all seats of an aircraft to be occupied.

On the one hand, the above sophisticated type of modeling seems to provide a more realistic approach to reflect actual air transport markets. Its use may be justified especially within the airport capacity allocation context, where capacity is usually allocated in terms of flight numbers but demand is satisfied by means of passenger seats in an aircraft. On the other hand, the aforementioned studies only consider the specific airline problem but do not extend the analysis to actual airport capacity allocation issues. Although a more detailed setting might be tremendously interesting with respect to capacity constraints, it would also increasingly complicate the analysis. Therefore, this study follows the more convenient approach and assumes both aircraft size and the seat load factor to be uniform and constant. In addition, it simplifies the aircraft size to unity. This facilitates the computation in that passenger volume and the number of flights become equivalent. The reason is that the model already abstracts from modeling flight frequency. In this case, adding aircraft size as a separate variable only increases the number of parameters but does not yield more interesting results.

10.4.2 Constant and Symmetric Marginal Costs

Brueckner (2002b, p.11) mentions that a network structure naturally emerged „because of economies of traffic density." He attributes this to decreasing passenger costs per route with higher traffic density and consequently increasing returns at the route level. Harback (2005, pp.2) also notes that hub-and-spoke networks served for clustering flights for the purpose of exploiting economies of scale in aircraft size. However, Brueckner (2002b) subsequently abstracts from cost economies in his analytical approach because the resulting increasing returns would make the analysis „exceedingly cumbersome". As Vives (2001, p.123) confirms, special attention is needed when increasing returns are designed: There might be no equilibrium, or outcomes may turn out to be surprising. Also, the general discussion in the literature is more controversial: While some researchers argue that the trade of aircraft size versus flight frequency paved the way to economies of scale and scope, others assume that no such size effects exists (see Section 3.1.1).

In order to focus on the demand-side effects of network benefits, this study therefore abstracts from modeling cost-side economies from network density. From a technical perspective, this choice is supported by other similar models from recent studies. As already pointed out, the concept of network density effects therefore differs from the common notion of network economies by abstracting from cost considerations and focusing on revenue side benefits from network density. In this respect, it should be stressed that the network benefits also do not constitute increasing returns: Although they monotonously increase utility for passengers, the corresponding network value that governs the airline's profit maximization is concave.

An example of constant but asymmetric marginal cost duopoly is considered in Verhoef (2010). As briefly summarized in Section 4.2, his result is a corner solution, where the less efficient airline is bought out of the market. However, in this model, networking costs are assumed zero. This is also justified by the complications from endogenous costs, as explained in Section 10.3.2. Nonetheless, Verhoef's (2010) result indicates that the investigation of asymmetric costs would provide an interesting topic for further research.

10.4.3 Congestion Costs

As mentioned above, Cook (2007b, p. 97) considers congestion costs both for passengers and airlines as "real, large, but poorly understood quantitatively." Nevertheless, he further specifies distinct instances of congestion costs that differ across airlines and passengers. For airlines, he dissociates tactical and strategic costs of delay: Tactical costs represent irregularity costs that occur due to actual delays, and may, e.g., consist of costly re-bookings on other flights, overnight board and lodging due to missed connections, rerouting of passengers and crews or legal penalties for delayed delivery of passengers or freight. By contrast, strategic costs account for ex-ante contingency planning measures to minimize the impact of actual delays, mainly by loosening operational density through adding buffers and time reserves in the schedules of crews and aircraft. For passengers, delay costs account mainly for the time costs that incur from late arrivals or missed connections.

This study applies delay costs both for airlines and for passengers. An airline's congestion costs, however, only account for the tactical portion. The inclusion of the strategic costs would require a much richer framework reflecting network operations in full detail. Concerning consumer time costs, Daniel (1995) asserts that the "external cost of delay caused by a marginal user" is well-known. Similarly to Brueckner (2002a), Brueckner (2002b) suggests convex congestion costs that increase with the number of peak flights. Czerny (2010) offers a range of different congestion cost functions, from deterministic linear to quadratic marginal cost with stochastic passenger benefits. For the purpose of this study, the above contributions thus justify the acceptability of the approximation of airline congestion costs by a non-decreasing, tactical congestion cost function.

Nonetheless, abstracting from the strategic costs of flight delays may underestimate the impact of congestion on airlines. Moreover, recall that it is assumed that all destinations which the hub connects to are uncongested. This assumption is taken for simplicity and because the analysis focuses on capacity allocation at the central hub airport only. In order for delay costs not to be underestimated, however, a more generalized setting should also take into account airline costs from flight delays incurred at the foreign airports.[68]

[68] I thank Raik Stolletz for this comment.

10.5 Passenger Benefit Functions

10.5.1 Direct Flight Benefits

The preference for peak-period travel illustrates that both traveling in general as well as peak-period travel are increasingly desirable, the lower the individuals' price sensitivity for booking a flight is. These properties can directly be inferred in-line with quality choice from standard vertical differentiation models (cf. e.g. Tirole, 1988, p.296). Along with the general assumptions about the direct flight benefits, it can be illustrated by Brueckner's (2002a) interpretation of consumer taste for peak-period travel:

As Brueckner (2002a, p.1361) suggests, peak-flight preferences may be thought of as reflecting a traveler's "tendency to travel on business." This view is based on the idea that business travel constitutes a "crucial job requirement" so that "both peak and off-peak travel benefits should be high relative to the benefits for a leisure traveler". Moreover, it draws on the common notion that efficient business travel should occur "during the early and late peak hours, to avoid disruptions of the work day." In economic terms, this argument suggests that business trips cause high opportunity costs if they are undertaken during off-peak hours in the middle of the day.

Although Brueckner (2002a) provides a horizontal rather than a vertical differentiation setting, the above arguments reasonably justify that business travelers are generally associated with a higher marginal willingness to pay than leisure travelers. This, in turn, implies that both general air travel as well as peak-period travel should also be increasingly beneficial when the marginal willingness to pay increases and hence the price sensitivity for air travel falls.

10.5.2 Uniform Time Costs

Time costs in this model only depend on the overall traffic volume at the airport and hence are independent of peak-travel preference. This means that at any given level of congestion, time costs induce an identical penalty on peak-period utility, regardless of the passengers' peak-travel preference. This homogenous nature of time costs is adopted from Brueckner (2002a). The main reason for this simplification is the traceability of the analysis, which arises from the additive separability of time costs in the utility function. Although this specification has been quite common in recent theoretical models it is not uncontroversial from an empirical point of view.[69]

[69] I thank Nicole Adler for this critical comment at the GARS Workshop 2014.

Brueckner (2002b) also discusses the issue of homogenous time costs arguably being too simple. However, in his opinion, heterogeneous time costs would increasingly complicate the analysis while not substantially adding to understanding the problem. In turn, Brueckner (2002a, p.1370) briefly investigates non-separable time costs in his model and finds that they „temper the results of the analysis without overturning its main lesson". Therefore, noting that the effect of congestion may be slightly underestimated for business travelers should ultimately justify the use of independent time costs also in this model.

10.6 Partial Equilibrium Analysis

A single-goods market with numeraire residual consumption reflects Marshallian welfare analysis from a partial equilibrium view (Mas-Colell et al., 1995, p.316). The intention behind this perspective is to obtain a traceable analysis. Simplicity is achieved because a partial equilibrium suspends substitution effects across other sectors of the economy by definition. As a consequence, all income effects are captured by the numeraire good, and do not affect equilibrium outputs on the flights market (also see 5.3.1).

This advantage comes at the cost of reduced precision: The model does not consider the general equilibrium of the economy as a whole but only of this sector. Yet, as Vives (2001, p.77) points out, partial equilibrium analysis of an industry is acceptable if its corresponding share of the consumer's budget is small. In this case, income effects are justifiably assumed to be negligible. This, in turn, allows us to both use a utility function that is linear in income as well as to represent the remainder of the economy as an aggregate numeraire. In other words, linear utility and an aggregate numeraire justify partial equilibrium analysis (Vives, 2001, p.145).

This model evidently qualifies for a partial equilibrium perspective, as the negligibility of cross-sectoral substitution effects seems acceptable, particularly with regard to the greatly reduced complexity in comparison to a general equilibrium analysis.

Part II. Application: Specific Allocation Schemes

„(...) transportation economists (...) note the self-evident optimality of pricing solutions, and then sit down waiting for the world to adopt this obviously correct solution. Well, we have been waiting for seventy years now, and it's worth asking what are the facets of the problem that we have been missing."
(Lave 1995 in Rietveld and Verhoef, 1998, p. 285)

11 Quotas (Airport Slots)

As described in Section 3.2.1, a quota allocation aims to reduce congestion to its socially optimal level by restricting output. This section presents two distinct kinds of quota schemes: Firstly, a naive allocation of individual quotas that aims at replicating the social optimum; and secondly, an arbitrary constraint in order to illustrate a grandfathering allocation. The individual quotas are determined based on the efficient market shares from the social optimum, while the arbitrary constraint is supposed to reflect the current administrative allocation scheme as is known from practice (see Section 2.4.2). Although the latter cannot be justified as arising endogenously within the micro-foundation of this model, it serves as a starting point for the subsequent evaluation of a secondary trading scheme.

11.1 Individual Quotas

In order to achieve a first-best solution with individual quotas in a homogenous setting, two general criteria must be met: First, the number of quotas needs to be correctly quantified. Second, the quotas need to be allocated accordingly to the participants. As airport demand usually exceeds supply, allocation efficiency would be reached (cf. Forsyth and Niemeier, 2008, p.66-67).

However, in an asymmetric setting, demand may vary across airlines which complicates the allocation problem (idem, p.68). Moreover, with market power individual outputs may become inefficiently low, which is ultimately indicated by the size and direction of the dual distortion (see Section 3.3). As a consequence, in the asymmetric setting with market power the above optimal number of quotas needs to be distinguished per airline. Moreover, the quotas ultimately need to constitute a binding constraint. If the quotas are not binding they cannot become effective. As a consequence, any welfare improvement could, at best, turn out to be second-best and at worst might even turn out to be adverse. Hence, the third criterion for allocation efficiency is that the quotas need to be individualized and that they must constitute an effective constraint. As previously mentioned, this argument arises from the product heterogeneity on the market.

In the following, the determination and the allocation of the individual quotas, which are individually computed for each airline, is described first. Subsequently, the welfare analysis investigates whether a social planner can reach first-best allocation efficiency when he allocates these individual quotas.

11.1.1 Determination

In this asymmetric model, the determination of the optimal number of quotas follows the social optimum computation from Section 7.1. Recalling that congestion depends on aggregate peak-period output only, the optimal number of airport slots $\hat{q}$ is indirectly determined by

$$\Theta - \hat{q} \equiv \underset{\theta^*}{\operatorname{argmax}}\, W(\theta),$$

where $W(\theta)$ denotes the social welfare function as defined in (14).

The first-order condition for $\hat{q}$ hence corresponds to the social optimum: On the one hand, condition (18) yields overall peak-period output if θ^* is governing. With substitutions $\theta^* = \Theta - N_p = \Theta - \hat{q}$, the optimum overall number of quotas $\hat{q}$ is implicitly defined by

$$B\left(\Theta - \hat{q}\right) - TG(\Theta - \hat{q}) = \hat{q} \cdot TG'(\Theta - \hat{q}). \tag{24}$$

On the other hand, in accordance with the social optimum, condition (19) yields overall peak-period output in the case where θ^D is governing. Moreover, this condition also determines the first-best number of quotas $\hat{q}_B$ that need to be allocated to the business airline in order to yield the optimum network size. This number of quotas is determined by

$$\Theta - \hat{q}_B \equiv \underset{\theta^D}{\operatorname{argmax}}\, W(\theta).$$

With substitutions $\theta^D = \Theta - n_p^B = \Theta - \hat{q}_B$, condition (19) implicitly solves this problem as

$$B\left(\Theta - \hat{q}_B\right) - TG(\Theta - \hat{q}_B) + d(\Theta - \hat{q}_B) = \hat{q}_B \cdot TG(\Theta - \hat{q}_B)' + \int_{\Theta - \hat{q}_B}^{\Theta} \frac{\partial d(\theta, \hat{q}_B)}{\partial \hat{q}_B} d\theta. \tag{25}$$

Monotonicity of the integral and concavity of the network value function imply single crossing and thus a unique non-negative solution. This has been shown correspondingly for condition (19) in Section 7.4.

Generally, whether the leisure airline should be allowed to exhibit a positive peak-period output at all depends on the relative importance of the network density benefits. Corresponding to social optimum condition (20), the substitutions $\theta^* = \Theta - n_p^B - n_p^L = \Theta - \hat{q}_B - \hat{q}_L$ yield the leisure airline's optimal number of quotas

$$\hat{q}_L(\hat{q}_B) = \max\left[0, \frac{B(\hat{q}) - TG(\hat{q})}{TG'(\hat{q})} - \hat{q}_B\right]. \tag{26}$$

The independent determination of $\hat{q}$ and $\hat{q}_B$ from (24) and (25) and the substitution of these results into (26) yields the socially optimal quota allocation.

According to the optimal quota rule that follows the social optimum, the leisure airline may hence either provide some residual peak-period output or may be completely re-allocated to the off-peak period. As the previous social optimum analysis from Section 8.1 reveals, however, the monotonicity of the network density benefits dictates that the business airline needs to exclusively serve the peak period. Therefore, the optimal number of quotas $\hat{q}_L$ for the leisure airline corresponds to the corner solution and is always zero. As a consequence, the business airline becomes the only supplier in the peak period and hence will try to revert to its monopoly output (see e.g. Mas-Colell et al., 1995, p.384ff.). Subsequently, congestion will be fully internalized (cf. e.g. Brueckner, 2002a, pp.1364). This important result will be referred to in the welfare analysis below. Yet, for completeness, in the following both above cases are investigated: The general case where quota rule (26) is governing as a maximum function $\hat{q}_L(\hat{q}_B)$ and this model's specific case where the network density benefits dictate that $\hat{q}_L = 0$.

11.1.2 Allocation

Assume that there is a social planner who is able to compute the correct number of individual quotas. Subsequently, this planner attempts to implement the above target allocation. In the absence of any other instrument, the most straightforward means is a naive imposition of these quotas to the airlines. This means that the quotas are simply distributed to the airlines for free according to the above quota rules.

The welfare result of this naive allocation depends on the different market outcomes before regulation: On the one hand, the overall output efficiency is determined by the dual distortion. On the other hand, however, the network size in equilibrium depends not only on the dual distortion but also on the concavity of the network value (see 8.2). As a consequence, in the asymmetric model, either both airlines may be constrained or the quotas may be asymmetrically binding for the leisure airline only. As mentioned above, however, first-best allocation efficiency requires that the quotas be binding for both airlines. Consequently, the crucial question is whether a social planner can actually yield the first-best optimum by allocating these optimal quotas to the airlines or whether only a second-best or even an adverse welfare effect applies.

For this consideration, recognize first that the asymmetric imposition of individual quotas on one airline yields a secondary output effect: Because the individual outputs provide best answers to each competitor's output based on the mutual reaction functions, an unconstrained airline will endogenously adjust its own output when its competitor's output becomes exogenously restricted. This pattern has been formally revealed as the conjectural variation (see Section 8.3). The welfare impact of the individual quota solution is therefore not straightforward.

The following analysis first describes the endogenous output adjustments as a consequence of the asymmetric quota imposition. Thereafter, it investigates the potential welfare effect of the individual quotas under the assumption that quota rule (26) implies a positive peak-period output for the leisure airline. Based on the model's properties and the different potential outcomes for the overall optimal output and the optimal network size, five different cases may generally occur. Subsequently, the welfare impact is discussed for this model's specific case, where the leisure airline must completely abstain from the peak-period.

11.2 Endogenous Output Adjustments

As stated above, the optimal quota rule may dictate that only one airline is constrained. The fact that airlines select their outputs according to their mutual reaction functions implies that the unconstrained airline will endogenously adjust its output according to its conjectural variation (see Section 8.3). This endogenous output compensation is explained as follows: When one airline's output is restricted asymmetrically, overall peak-period output decreases and congestion is diminished. The unrestricted airline will increase its peak-period output in order to re-adjust its marginal revenues to marginal costs. The output reduction of asymmetrically imposed quotas is hence offset to some extent by an output expansion of the unconstrained airline.

11.2.1 Primary Effects

The endogenous output adjustments can be shown by means of comparative statics in the reaction functions (9) and (10): Firstly, the initial overall output reduction shifts θ^* to the right. Monotonicity implies that net flight benefits B^* strictly increase while time and congestion costs TG^* strictly decrease. As a consequence, the left side of equations (9) and (10) becomes larger. This dictates an increase in the unrestricted airline's output in order to re-balance the right-hand side of (9) and (10). Subsequently, we must still check whether this endogenous output adjustment only partly off-sets the initial output reduction, or whether it might even overcompensate the reduction.

For this purpose, the size of the output compensation relative to the original cutback can be assessed in two ways: On the one hand, each unrestricted airline's profit-maximizing rationale infers that a higher peak-period output could already have been chosen before imposition of the quotas if this had been desirable. However, the unrestricted equilibrium has already balanced delay costs, flight benefits and network size to each airline's optimality. Therefore, an endogenous output compensation can only be worthwhile if it substitutes some of the competitor's output at a lower amount of overall congestion.

On the other hand, the magnitude of the output compensation relative to an airline's initial output diminution can be formally evaluated by comparative statics in the airlines' first-order conditions (9) and (10): To begin with, the output reduction diminishes total congestion and output. As a reaction, the unconstrained airline will increase output in order to re-balance the right-hand side of its equilibrium condition. In contrast to this, the constrained airline will remain with a constant output because the quotas already force it to supply less than its profit-maximizing equilibrium output. Consequently, θ^* again decreases until the higher gross flight benefits at the left-hand side of the respective equilibrium condition equalize the higher peak-period output of the unconstrained airline on the right-hand side.

However, the respective equilibrium condition was originally balanced with a lower number of flights in the unregulated equilibrium. Consequently, the gross flight benefits need to remain above their unconstrained equilibrium value in order to accommodate the now higher number of the unconstrained airline's peak period flights. In turn, this requires that θ^* also remains at a higher value than before regulation. Therefore, overall peak-period output after regulation needs to remain lower than in the unconstrained equilibrium. This shows that the endogenous output compensation cannot be exhaustive but needs to fall short of the initial output reduction caused by the quota constraint. As a consequence, an asymmetric quota constraint reduces overall output by less than actually intended.

11.2.2 Secondary Effects

The secondary effects that arise within this endogenous output adjustment again are implicitly determined by equilibrium condition (10). They can be shown to either dampen or reinforce the output compensation. Consequently, they also have an impact on the effectiveness of the quotas. For the following consideration, it is assumed that only the leisure airline is constrained. An asymmetric constraint on the business airline works in the same way concerning overall congestion and in the opposite way concerning the network density benefits.[70]

Recall that the marginal network value can be positive or negative in equilibrium (see Section 8.2). The marginal network value is negative when network benefits are important relative to delay costs and the natural network size is large. In the opposing case where delay costs are more important than the density benefits, the network size remains small. Consequently, the marginal network value depends on the natural size of the network.

If the marginal network value is negative, the increased network size also decreases the network value. The left-hand side of (10) therefore further diminishes. On the right-hand side, $d'(\theta^D)$ is positive and further increases based on the concavity of the network value (see Section 8.2). An equalization of (10) can thus only be obtained when gross flight benefits $B^* - TG^*$ on the left-hand side further increase. For this, θ^* needs to shift farther to the right. If the natural network size is high and the marginal network value is negative, output compensation is dampened. This would make the quotas more effective, although they still would not be able to reach their goal on an overall congestion level.

If the marginal network value is positive then the network expansion increases the network value. Consequently, the left-hand side of (10) diminishes by a smaller amount than the overall output restriction actually induces. On the right-hand side, the absolute magnitude $\left|d'(\theta^D)\right|$ decreases. However, as $d'(\theta^D) < 0$ has a negative sign in this case, the right-hand side of (10) increases by more than justified by the peak-business output expansion alone (again see Section 8.2). However, the fact that both sides of the equation increase yields ambiguity about the direction of the secondary effect. Therefore, it cannot be determined whether the secondary effect dampens or reinforces the output compensation when network size is small and the marginal network value is positive. Consequently, the effect of the quotas on congestion is ambiguous: It might either be supportive and further reduce output compensation (as in the above case), or it might increase the output compensation and further reduce the effectiveness of the quotas.

[70] As the analysis already revealed, a constraint on the business airline would never occur. This consideration is, rather, theoretical and is provided for completeness only.

Once again, it is worth noting that these adjustments are only secondary effects of the endogenous output adjustment, which is itself a second-order effect of output reduction following the actual quota imposition. Therefore, they cannot overturn the initial effects; otherwise, the corresponding adjustment would have already been undertaken in equilibrium. The above secondary effects may hence either dampen or reinforce the endogenous output compensation depending on the sign of the marginal network value, but overturn neither its direction nor relative magnitude against the original output change as imposed by the quotas.

11.3 Results: General Case

The following analysis investigates the welfare impact of the individual quota allocation according to quota rule (26). For generalizability, it abstracts from this model's particular result that the leisure airline must not serve the peak period at all. The investigation of the latter specific case is presented subsequently in Section 11.4. This analysis separately considers the two typical equilibria that may arise as a function of the dual distortion: The one where the congestion externality is dominant so that overall output is excessive, and the one where the market power distortion prevails and hence overall output is inefficiently low.

11.3.1 Overall Output excessive

In the case where the congestion externality exceeds the market power distortion, the unconstrained aggregate equilibrium output is higher than socially optimal and the quotas are binding on an overall level. In this situation two distinct cases may arrive: either, both airlines' outputs are too high, or, only one airline's market share is excessive but the other one's is too low. The welfare results differ between these two cases.

In the former case, the quotas are binding for both airlines and thus replicate the optimal market shares. Moreover, conjectural variation is zero because the individual constraints are effectively binding for both airlines. Consequently, the individual quotas yield first-best allocation efficiency. In the latter case, where overall output is excessive but only one airline is constrained (because the other airline's output is inefficiently low), the unconstrained airline will increase its own output as a response to its constrained competitor's output decrement. As a result, the individual quotas' impact on allocation efficiency depends on the size of the initial deviations of the equilibrium market shares from the social optimum and on the marginal impact of the network density benefits.

Generally, any endogenous output adjustment based on the conjectural variation is never exhaustive (see 11.2 above). Therefore, such output expansion cannot substitute the constrained airline's output contraction to the full extent. Nevertheless, the unconstrained airline's desired output after the quota imposition may either supersede its individual efficient output or remain lower than its socially optimal market share. In the case where the unconstrained airline's output adjustment equals or exceeds its own socially optimal output, this airline will also become affected by its individual quota constraint. As a consequence, the quotas again become effectively binding for both airlines and first-best allocation efficiency will be established. This case arises when the congestion externality tends to be relatively important.

In the contrasting case, where the congestion externality is less important (but still important enough to cause an excessive overall output), the unconstrained airline's output increases toward its efficient value but remains below its socially optimal market share. Although the constrained airline's output is optimal, the market power distortion becomes prevailing on an overall level, so that overall output becomes inefficiently low.[71] As a consequence, the welfare effect of the individual quotas becomes ambiguous: If the low output inefficiency based on the prevailing market power distortion remains relatively small as compared to the welfare benefit of the congestion cost reduction, the welfare distortion decreases in sum and allocation efficiency is increased. Note, however, that this welfare improvement is second-best because the optimal structure is not fully replicated. If, by contrast, the adverse welfare effect of the low output inefficiency exceeds the efficiency benefits from the reduced congestion, the overall welfare impact of the individual quotas becomes adverse. The welfare caveat of the individual quotas thus increases with an increasing importance of the market power distortion relative to the external effect of congestion. As mentioned above, however, the final welfare result of the individual quota is only revealed when the effect of increasing or decreasing network density benefits has also been taken into account.

The network density benefits substantially complicate the welfare analysis, because they either increase or overturn the above welfare effects from the dual distortion, depending on which airline remains unconstrained: If the networking business airline is not constrained, its endogenous output compensation monotonously increases the network density benefits. Consequently, any positive welfare effect is reinforced while any welfare caveat is reduced. At the limit, the beneficial network effect may even overturn an adverse welfare impact into a second-best welfare improvement if it is important in relation to the market power

[71] Note that this result is valid on an overall output level (as compared to the initial, unconstrained equilibrium) despite the fact that the unconstrained competitor's output inefficiency based on market power actually decreases with the endogenous output increment.

distortion. If the non-networking leisure airline remains unconstrained, its endogenous output adjustment does not affect the indirect utility from the network effects, while the quantity constraint on the business airline would unambiguously decrease the overall network benefits. In the social optimum, however, this decrease is balanced against the beneficial effect arising from lower congestion. Put differently, the network density effects may increase the socially optimal amount of congestion, but do not affect the optimality in itself.

The fact that the network density benefits may yield an additional efficiency gain indicates the potential of an additional second-best welfare improvement by deviation from the optimal quota rule in the case where first-best allocation cannot be reached and the business airline remains unconstrained. In this case, decreasing the leisure airline's number of quotas below the socially optimal amount would induce an additional increase in the network size implying additional net network density benefits. As overall output would further decline, however, this trade-off between a larger market power distortion and higher network density benefits would only be worthwhile as long as the initial welfare effect is positive.

In summary, the general welfare impact of the individual quotas remains ambiguous when the congestion externality is important so that overall output excessive. More specifically, three different cases may arise: Firstly, if both airlines' outputs are excessive, the quotas yield a first-best welfare result. Secondly, if one airline is unconstrained because its initial output had been inefficiently low, a first-best allocation may be reached based on that airline's endogenous output adjustment, as long as the original output deviation was sufficiently small. Thirdly, if the original market shares substantially deviate from the social optimum, either a second-best welfare improvement arises or welfare is adversely affected. The final result depends on the relative importance of the two effects that constitute the dual distortion. In the last case where first-best allocation efficiency cannot be reached, the network density benefits have an additional impact: If the networking business airline remains unconstrained, its endogenous output adjustment increases network density and thus customer value. Consequently, the welfare caveat from an excessive market power distortion is reduced. If the non-networking airline remains unconstrained, there is no additional welfare impact based on network effects.

Lastly, recall that in the specific case of this model the business airline's network size can never be excessive in equilibrium. This upshot arises from the monotonicity of the network density benefits against the concavity of the network value (see Section 9.4). The only case where overall output is excessive in this study is thus the one where the network size is inefficiently low, so that the business airline always remains unconstrained. As a result, the individual quotas may either yield a second-best welfare improvement or a welfare deterioration, whereas first-best allocation efficiency cannot be reached.[72] Due to the asymmetric effectiveness of the quotas, however, any welfare caveat is reduced. This model's specific case is further investigated in Section 11.4.

11.3.2 Overall Output inefficiently low

If the congestion externality is weak in comparison to market power, overall output falls short of the social optimum and the quotas are not binding on an overall level. Nevertheless, if the equilibrium market shares sufficiently deviate from the socially optimal values the individual quotas may be asymmetrically binding for one airline at a time. For the imposition of the individual quotas when output is inefficiently low, three different cases need to be distinguished: The case where peak-leisure output is too low but network size too large, the case where peak-leisure output is too large but network size too small, and the case where both airlines' outputs are inefficiently low.

If the leisure airline's peak-period market share is too large as compared to the social optimum, it is constrained by the initial quota allocation. Based on the endogenous output compensation, the business airline increases its flight volume. On the one hand, this permits the network to grow toward its optimum size. On the other hand, overall output decreases because this output compensation is not exhaustive. Consequently, the market power distortion is increased and congestion decreases even farther below its optimum. Although the network size increases, both overall output and the network size remain inefficiently low. If, in this case, the higher network density benefits can overcompensate the higher deadweight loss, then allocation efficiency is improved at the margin. If, by contrast, the network expansion is unable to counterbalance the increasing welfare loss, then allocation efficiency is reduced. At best, thus, the individual quotas yield a relative welfare improvement; at worst, they induce an adverse welfare effect.

Note that restricting the leisure airline below its optimal market share could again beneficially affect welfare in the above case, provided that the leisure airline remains with a positive number of quotas. The welfare effect of this deviation from the quota rule ultimately depends on the relative importance of market power, the congestion externality and the density benefits at the respective levels of output. As explained above, the trade-off between a lower overall output and a higher network size is beneficial if the higher network density benefits overcompensate the rising deadweight loss.

[72] A first-best solution is ruled out because the network size always remains inefficiently low, even if the business airline exclusively serves the peak period (see 9.3).

In the contrary case, where peak-leisure output is too low but the network size too high, the effects above apply correspondingly but in the opposite direction: The business airline is constrained by the quotas and has to reduce its own output. Consequently, the leisure airline endogenously increases its flight volume. This output compensation again is not exhaustive. As a result, the network size becomes optimal but overall output is further depressed. Again, the higher deadweight loss needs to be balanced by the higher network utility for a beneficial welfare result; otherwise allocation efficiency is diminished.

Note that in this case the effect of the business airline's constraint is counter-intuitive: Because network size was originally excessive, the contraction actually increases the network density benefits. When the network density benefits monotonously increase with network size as in this model, however, such a scenario cannot arise. This particular initial allocation is hence limited to the general consideration and does not apply in the specific case of this asymmetric model. As already mentioned, the same applies to first case in above Section 11.3.1, where the network size is excessive.

Finally, when both airlines' outputs are inefficiently low, the quotas are not binding for either airline and have no effect on welfare. However, also in this case the opportunity might arise to invoke a second-best welfare improvement by constraining the leisure airline below its actual output. As shown above, the higher network density may overcompensate the higher deadweight loss, so that welfare might improve. Note, again, that this constraint on the leisure airline needs to be more restrictive than the individual quotas based on the optimal quota rule.

11.3.3 Distinct Characteristics of the Asymmetric Model

The above investigation reveals the following two distinct characteristics of individual quotas in this model: First, the welfare effect of the quotas crucially depends on the original size of the dual distortion. Generally, first-best allocation efficiency may be reached if both overall output and the original size of the business airline's network are excessive. Due to the monotonicity of the network density benefits, however, this case cannot arise in the current asymmetric model. As a consequence, at best, the individual quotas yield second-best allocation efficiency; at worst, they adversely affect welfare. However, both the welfare improving secondary effect and the welfare caveat of the asymmetric quotas stem from the airline asymmetry; hence, they would not arise in a symmetric setting with flights as homogenous products.

The second property consists of the endogenous output adjustments according to each airline's conjectural variation, based on its reaction function (as explored in Section 8.3). This asymmetric output compensation occurs because the individual quotas may be binding for one airline only, so that the unconstrained competitor can endogenously increase its individual output as a reaction to the exogenous decrease in congestion. Hence, the asymmetric imposition of the quotas has an output-raising secondary effect on the unconstrained airline, which was shown to always occur but not to be exhaustive. As a consequence, the quotas may not only decrease an airline's individual output but may also increase an unconstrained competitor's output in an indirect way.

This result is surprising, as it contradicts to the general functionality of a quota scheme. Because it originates from the combination of the airline asymmetry and endogenous demand, it arises neither in a symmetric setting, nor under perfect competition, nor under market power with inelastic demand. It signifies that the quotas may increase the individual output of one airline whose output is inefficiently low while at the same time restricting the other airline's excessive output. As a result, in an asymmetric allocation, the individual quotas cannot only achieve a second-best welfare improvement by restricting excessive outputs but also by indirectly increasing individual outputs that are inefficiently low.

11.4 Results: Specific Case

Let us now abandon the assumption of a positive peak-period output for the leisure airline and assume that quota rule (26) dictates that the peak period must be exclusively served by the business airline only. In this case, any positive peak-period output of the leisure airline now is excessive by definition. Moreover, recall the distinction between the concavity of the network value for the airline's profit maximization and the monotonicity of the network density benefits from a welfare perspective. As already mentioned in the previous analysis, this yields that the equilibrium network size always remains too small (see, e.g., the discussion in Section 9.3).

As a consequence, only two equilibria need to be considered in the analysis of the asymmetric model: The equilibrium where overall output is excessive and the equilibrium where overall output is inefficiently low, while the network size always remains undersized. The only difference between the two equilibria is that the congestion externality either exceeds or falls short of the market power distortion. This considerably simplifies the welfare analysis of the individual quota allocation in the asymmetric model.

11.4.1 Excessive overall Output

Overall output is excessive in equilibrium if the congestion externality exceeds the market power distortion, so that only a small amount of total congestion is accounted for by the airlines. As the individual quotas completely suppress the leisure airline's peak-period output, the business airline becomes the only supplier in that period and reverts to its monopoly output. The congestion externality hence vanishes so that overall output is lower than before the imposition of the quotas. Despite the endogenous output adjustment, the network size remains inefficiently low and the quotas are not binding for the business airline. The quotas thus eliminate the congestion externality but foster the market power distortion.

As a result, the impact of the quotas on allocation efficiency becomes ambiguous: On the one hand, overall output is no longer excessive but falls below its optimal level. The welfare effect of this output contraction depends on whether the absolute magnitude of the dual distortion rises or falls: If the congestion externality was large, overall output is likely to remain closer to the social optimum, so that efficiency increases. If the congestion externality was small relative to the market power effect, the absolute magnitude of the dual distortion is likely to increase, diminishing allocation efficiency. On the other hand, the net welfare effect of the network expansion would be positive. Thus, the output effect remains ambiguous while the network effect yields a net positive welfare contribution.

As already established in the general analysis, in this case the welfare impact of the individual quotas ranges from a second-best improvement to an adverse effect. The final result depends on the relative sizes of the effects: If the overall output effect positively affects welfare, then the quotas yield a second-best improvement. If the overall output effect negatively affects welfare, then the quotas are only beneficial when the network effect overcompensates the higher dual distortion; otherwise, they decrease allocation efficiency even when the initially unregulated overall output was excessive.

Lastly, note that the effects based on the dual distortion exactly correspond to Brueckner's (2002a) findings for a congestion tax (see Sections 3.3.2 and 4.2). This model's net positive network effect on welfare, however, contrasts to Brueckner's case; it diminishes the chance of a welfare deterioration and, thus, increases the likelihood that a quota scheme yields a second-best welfare improvement.

11.4.2 Overall Output inefficiently low

In the second case, overall output and the network size are both inefficiently low. The imposition of the quota rule on the leisure airline's peak-period output again induces a monopoly situation, which yields the same effects as above but with one crucial difference: When output is inefficiently low, the congestion externality was already inferior to the market power distortion before the imposition of the individual quotas. Consequently, the output effect decreases overall output farther below the social optimum. This signifies that congestion becomes inefficiently low and market power excessively high as compared to their efficient levels. As a result, the dual distortion increases. The output effect of the quotas therefore unambiguously decreases welfare.

Nevertheless, the business airline's endogenous output adjustment once again yields an increment in the network size. This network effect diminishes the negative welfare contribution of the output effect. Yet, whether the network expansion overcompensates the negative output effect again depends on the relative sizes of the effects. Thus, the output effect corresponds to Brueckner's (2002a) congestion pricing case, whereas the network effect yields a net welfare gain. As compared to the initial allocation where overall output had been excessive, however, the net negative output effect increases the welfare caveat.

11.4.3 Welfare Caveat in the asymmetric Case

The comparison of the above welfare analysis against the general case from Section 11.3 shows that there is a welfare caveat that arises specifically from the optimal allocation in the asymmetric setting: Restricting the leisure airline's peak-period output to zero means that the business airline will naturally return to its monopoly output in the peak period. This output allocation arises because it is the profit maximizing choice in the absence of a competitor. As a result, the delay costs are fully internalized. Consequently, the congestion externality vanishes and market power becomes the only distortion. Despite the endogenous output adjustment, overall peak-period output becomes inefficiently low while the network size remains below its socially optimal level. In other words, completely banning the leisure airline from the peak period always yields an inefficiently small network, although the peak-period output contraction is partly off-set by an endogenous output expansion of the business airline.

The welfare effect of this corner solution is generally ambiguous: On the one hand, the congestion externality vanishes, but on the other hand, the market power distortion is increased.

Moreover, the network size increases but remains below the optimal size that is presupposed by the first-best quota computation. Under these circumstances it remains unclear whether the optimizing rationale of the planner can provide any welfare improvement: If the congestion externality before regulation was relatively important, a second-best welfare improvement is likely. If overall output was already close to the social optimum, an adverse welfare result may occur through further output contraction. This shows that a first-best result cannot be reached when the quota rule dictates $\hat{q}_L = 0$, even if overall output had initially been excessive. Moreover, if $\hat{q}_L = 0$ provokes a large reduction of output, the welfare effect may even become adverse.

Note that in this case, allocation efficiency might be improved by reallocating the business airline's unused quotas to the leisure airline. This would lead to an output expansion that would again reduce the market power distortion. However, this output expansion would also partly reverse the business airline's secondary output compensation that followed the initial restriction of the leisure airline and the network size would again diminish. The net impact of this trade-off depends on the change of network size, market power, and congestion. A beneficial welfare effect is only achieved if the initial quota allocation adversely affected welfare; otherwise, the network effect is dominant against the output effect, so that a deviation from the optimum quota rule $\hat{q}_L = 0$ is not advisable.

11.5 Grandfathering Allocation

In a grandfathering allocation both airlines are constrained by definition. As quota rule (26) shows, however, such an allocation cannot be endogenously replicated within this model. The solution is to introduce an arbitrary constraint, as proposed by Verhoef (2010), that is simply defined to be binding for both airlines. This symmetrically binding constraint overcomes the problem of insufficient airport demand in the presence of market power and thus reflects the scarcity of airport capacity.

11.5.1 Definition

The arbitrary constraint is formally denoted with an overbar superscript in order to point out that the corresponding individual outputs are restricted in any case. The condition for defining an appropriate number of arbitrary arbitrary constraints can thus be written as

$$\begin{aligned} \bar{q}_L &\equiv \left\{ q_L \mid 0 < q_L < n_p^{L*}(n_p^{B*}) \right\} \\ \bar{q}_B &\equiv \left\{ q_B \mid 0 < q_B < n_p^{B*}(n_p^{L*}) \right\} \end{aligned} \quad (27)$$

The above conditions simply dictate that the arbitrary constraints need to be below each airline's respective unconstrained equilibrium outputs n_p^{B*}, n_p^{L*}.

Recall that any endogenous output adjustment of an unconstrained airline would exceed that airline's output from a mutually unconstrained equilibrium. This makes clear that definitions (27) ensure that each constraint remains binding for both airlines even in the face of the conjectural variations. For the subsequent investigation, it is sufficient to leave the absolute magnitude of the constraints undetermined.

11.5.2 Results

As already mentioned, the above arbitrary constraint on both airlines cannot be replicated on the grounds of efficiency in the asymmetric oligopoly model. In other words, a grandfathering allocation has no justification from a welfare perspective within this model. Nevertheless, this analysis serves to show its distinct effects. For this purpose, the analysis compares the arbitrary constraint to the unconstrained market equilibrium.

The impact of a grandfathering allocation is the following: On the one hand, the output constraint reduces congestion and increases the market power distortion. Depending on the initial size of the dual distortion, the resulting welfare effect may be positive or negative. Because in this model the constraints need to be sufficiently low in order to become binding, however, an adverse result is more likely. Nevertheless, the arbitrary constraint should replicate a case where congestion would initially be excessive. Consequently, we may simply assume that the welfare impact of its output effect is positive.

On the other hand, any reduction of the business airline's peak-period output decreases the network size and the network density benefits for all passengers, which adversely contributes to efficiency. In the presence of asymmetric network effects, the welfare effect of a grandfathering allocation hence becomes ambiguous. Welfare is only increased if the reduced congestion overcompensates the higher market power and if the resulting positive effect of the reduced dual distortion in turn overcompensates the loss in terms of network density benefits.

As previously explained, however, the above results showed that constraining both airlines cannot correspond to a socially optimal solution in the asymmetric model. Therefore, an allocation where the business airline is also constrained can only arise if the corresponding number of constraints is chosen to be inefficiently small. Consequently, this welfare analysis remains unsatisfactory but still serves to illustrate that the welfare caveat from an arbitrary constraint remains the same as in the asymmetric individual quota scheme: Firstly, it consist of the potentially adverse effect from increasing the dual distortion, and secondly, of the negative network effect arising from restricting the business airline's output.

11.6 Summary

The most important insight from the above analysis is that first-best efficiency can never be reached by a naive allocation of individual quotas. Generally, a second-best efficiency improvement may be expected, but at worst the welfare effect may even become adverse. This result directly arises from the airline asymmetry: The latter causes the business airline to remain unrestricted and the leisure airline to be completely expelled from the peak period. This means that the quotas are never binding for the business airline, while they are applicable to the leisure airline even if overall output is inefficiently low. As a result, the business airline will endogenously increase its output. Nevertheless, the network size remains inefficiently low because the monopoly output is the profit-maximizing choice of the business airline. The inefficiency of the individual quotas thus arises from the dual distortion, the endogenous output adjustments based on the finite demand elasticity and the monotonicity of the network density benefits.

The welfare effect of the individual quotas depends on the relative size of the above effects. If the absolute magnitude of the dual distortion is decreased by the overall output contraction, the benefit of lower congestion overcompensates the higher market power distortion and the quotas yield a second-best welfare improvement. If the absolute magnitude of the dual distortion is increased because overall output is shifted farther away from (and thus below) the optimum, the welfare effect is ambiguous: It is second-best if the higher network density benefits can overcompensate the increased dual distortion. Otherwise, it becomes adverse.

Based on the above result, allocation efficiency might be improved by a deviation from the first-best quota rule: Allocating a part of the business airline's unused constraint to the leisure airline would increase overall peak-period output. As a consequence, congestion and market power would shift closer to their socially optimal values, whereas the network size would again decrease. Therefore, if the network density benefits are small relative to the dual distortion, this deviation might further increase the second-best welfare improvement of the constraint. At the limit, it might even turn an adverse welfare effect into a second-best efficiency improvement.

With regard to the dual distortion, the above welfare result follows Brueckner's (2002a) outcome for a congestion pricing scheme. however, the additional network density benefits introduced in the model at hand monotonously increase with the business airline's output and thus positively affect welfare. Consequently, the welfare caveat caused by an increasing dual distortion is reduced in the asymmetric model with network density benefits, as compared to a homogenous symmetric setting.

12 Secondary Trading

The investigation of the individual quota scheme has shown that, after the allocation of the quotas, both the network size and overall peak-period output remain inefficiently small. As a result, the business airline owns unused slots while the leisure airline is generally not allowed to provide any residual peak-period output. Two questions arise from this outcome: First, if secondary trading were allowed, would the business airline sell some of its remaining slots to its competitor? Second: If such a trade took place, would it be beneficial in terms of welfare or would it yield an adverse effect?

This section investigates the above two questions based on the asymmetric initial allocation of the individual constraint from Section 11. In addition, in order to enhance the analysis, the case of the grandfathering allocation as presented in Section 11.5, which constrains both airlines symmetrically, is also considered. Moreover, the investigation accounts for strategic airline behavior. For this purpose, the first subsection applies the differentiation between strategic behavior and strategic competition from Section 3.5 to the actual game options of both airlines in this model. The second subsection presents a set of trading rules that determine whether and in which form strategic behavior may occur within the secondary trading scheme. These trading rules are applied in the later analysis.

The investigation of secondary trading considers four distinct settings: The asymmetric allocation of the individual quotas and the symmetric grandfathering allocation of the arbitrary constraint, both with and without the opportunity for strategic airline behavior. For this reason, first, the computation of the trading potentials for both airlines is explored. The trading potentials denote the potential total gains or losses that accrue to the two trading parties. Subsequently, the analysis derives the necessary and the sufficient conditions for a positive market price and a corresponding quota trade based on the above trading potentials. These conditions represent the formal background for the evaluation of the four different trading situations. Although the generic model does not allow us to resolve these conditions analytically, the trading results are found based on inferences from the market structure and from general economic theory. Lastly, the welfare analysis reveals the impact of both the trading rules and secondary quota trading on allocation efficiency.

Ultimately, this section thus resolves whether secondary trading is likely to occur or not and whether it potentially yields a welfare benefit for each of the above four settings. With respect to methodology, this analysis is strongly motivated by Verhoef (2010) and Basso and Zhang (2010). Nonetheless, the formalities have been developed independently because these studies differ both in their models as well as in the cases considered.

12.1 Strategic Behavior and Strategic Competition

As discussed in Section 3.5, the concepts of strategic behavior and strategic competition are not easily distinguished. Therefore, this subsection briefly illustrates the specific game strategies that may arise from strategic airline behavior within this model's secondary trading scheme and dissociates the former from the strategic competition for airport capacity. Whether strategic behavior will actually emerge within the distinct situations depends on the specific trading rules that affect the game options available to the participants. These trading rules are presented in the next subsection.

12.1.1 Strategic Competition

In the context of this study, strategic competition refers to the fact that an airline evaluates whether to make more profits from selling its quotas than from actually using them. Similarly, the airlines may try to increase their market share in a costly manner by buying additional slots if it seems profitable and if slots are available for purchase on the secondary trading market. If the competitor's output reduction and its own output expansion yield a profit that overcompensates the slot price, such a trade is worthwhile. Strategic airline competition for airport capacity thus denotes the challenge for airport slots as a scarce resource and a key business asset. As a consequence, it represents the fundamental driver for secondary quota trading.

12.1.2 Strategic Behavior

In contrast with strategic competition, strategic airline behavior denotes a game strategy that aims at profit maximization by either avoiding a quota trade or by trading quotas not for the purpose of utilization but in order to prevent those quotas from being used by the competitor. Either option arises in this model, depending on the type of initial quota allocation.

In the asymmetric initial allocation, the slot-holding airline decides whether to trade its slots or whether to keep them unused. This rationale was assessed above as strategic competition. However, let us suppose that the regulator defined that unused slots had to be returned after the initial allocation. This means that keeping unused quotas would no longer be an option. At this point, strategic behavior arises: If the slot-holding airline would like to avoid the mandatory hand-back of its unused quotas, it would have to utilize them. However, the unconstrained airline has chosen its flight volume after the initial quota allocation so as to maximize its profits. Consequently, an output expansion only for the purpose of utilizing the excess number of quotas depresses profitability. In the absence of the hand-back obligation, this output expansion would be undesired and is, therefore, rated as inefficient. In the literature, the inefficient slot usage for the purpose of hand-back avoidance has been extensively discussed and referred to as the "babysitting" of slots. The potential slot buyer, in contrast, is constrained to zero peak-period output. Therefore, he is always better off using his purchased quotas than hoarding them. As a consequence, the potential buyer is not involved with strategic behavior in the asymmetric case.

As opposed to the asymmetric case, in the symmetric grandfathering allocation, all quotas are utilized before trading. Therefore, strategic behavior occurs in the way that an airline could buy a number of quotas from its competitor for the purpose of holding them unused. The buyer can thus choose to reduce overall output in order to increase his market power without giving up any of his flight volume. Because the initial allocation is symmetric, this signifies that either of the two airlines can restrict its competitor's output in a costly manner by means of a quota trade. Subsequently, it decides whether it is more beneficial to increase its own market share by using the slot or to reduce overall output by not using the quotas. The latter case occurs if the benefit from higher market power and lower congestion overcompensates the potential gain from an additional unit of output. As a result, the buyer exhibits post-trading strategic behavior by actually utilizing only part or none of the purchased slots. This strategy is henceforth referred to as slot hoarding. Verhoef (2010, p.326) considers a symmetric initial allocation and thus reflects this type of strategic behavior. As opposed to the latter, Basso and Zhang (2010, p.383) assume that airlines behave „non-manipulatively“ and abstract from the possibility of post-trading slot hoarding. Lastly, note that the seller is not concerned with strategic behavior in the symmetric case: His output is restricted and any further unsolicited output contraction without a monetary compensation cannot be worthwhile.

12.1.3 Implications

The above consideration shows that strategic airline behavior arises with the secondary trading option itself. In contrast, strategic competition was already reflected both in the unconstrained equilibrium and in the quota scheme without explicitly being mentioned: In terms of the output decision based on the reaction functions in equilibrium and in terms of non-utilization of the allocated slots in the individual quota scheme. Moreover, the above exploration shows that two distinct types of strategic airline behavior may arise in this model: the inefficient pre-trading quota utilization, referred to as babysitting, and the non-utilization of purchased quotas after trading, referred to as slot hoarding. Which one of the two types occurs depends on whether the initial allocation is symmetric or asymmetric.

These two kinds of strategic airline behavior have distinct characteristics, that again depend on the type of the initial quota allocation: In the asymmetric case, strategic behavior concerns the slot seller. It occurs instead of a potential quota trade and thus prior to actual secondary trading. In the symmetric case, it is the potential quota buyer who may decide to buy some slots in order to keep them unused after the trade. Slot hoarding is thus exhibited by the slot buyer and occurs after quota trading. As a consequence, the way in which the trading rules affect the occurrence of strategic behavior differs between the two cases. This issue is further explored in the next subsection.

12.2 Trading Rules

The occurrence of strategic behavior in a secondary trading scheme depends on the applicability of two trading rules: the hand-back rule and the reallocation rule. The hand-back rule dictates that unused quotas need to be handed back to the coordinator. The reallocation rule determines whether returned quotas will in turn be reallocated to the other airline. The following discussion presents these two trading rules and evaluates the corresponding game options of the two airlines in both the symmetric and the asymmetric initial allocation case.

12.2.1 The Hand-Back Rule ("use-it-or-lose-it")

The first restriction on strategic airline behavior is implemented by means of the hand-back rule. This rule dictates that any slot holder must either use all of his slots or return the unused portion to the regulator. As a consequence, any excessive number of quotas cannot be hoarded by an airline. Conceptually, this rule corresponds to the "use-it-or-lose-it" obligation that is currently imposed in practice.[73]

In the symmetric case, the hand-back rule implies that all purchased quotas must also be used and cannot be hoarded after a trade. Consequently, a potential slot buyer has the following two game options after a trade: He can either utilize the slots or hand them back. Note, however, that in the absence of subsequent quota reallocation, a quota hand-back has the same effects as buying a slot and holding it unused: Overall output is reduced

[73] In practice, a sufficient quota utilization under the hand-back rule is defined by a use ratio of 80%. This means that a slot must be operated on four out of five days. This subtlety is not implemented here as the equilibrium model is static and does not allow for variations over time.

and market power increases while the turnover of the slot buyer remains the same. Put differently, a hand-back exactly corresponds to the strategic option of hoarding the purchased quotas. Although a quota trade with a subsequent hand-back seems a rather theoretical case, it exactly replicates strategic airline behavior and therefore needs to be considered. As a consequence, the hand-back rule cannot prevent strategic airline behavior in the symmetric case if it is applied without a reallocation obligation.

Also in the asymmetric trading case, the effects of the distinct game options are ultimately independent of the isolated application of the hand-back rule. On this matter, recall first that in the asymmetric allocation, all potentially traded slots are initially unused. Consequently, purchasing an unused slot and keeping it unused would only be costly but would not change the overall output allocation. Therefore, neither a quota hand-back nor strategic slot hoarding are attractive for the buyer. A trade may only take place if the buyer intends to use the slot that he has bought. As a consequence, the above motive for strategic quota hoarding does not apply to the buyer.

For the seller, however, the imposition of the hand-back rule signifies that he can no longer hold his excess quotas unused. This generates the decision regarding whether the unused slots should be handed back or should be used in order to avoid the hand-back. However, as long as a hand-back does not result in a subsequent quota reallocation to the competitor and thus in a corresponding output expansion, it has the same effect as simply holding on to the excess quotas. Put differently, the hand-back rule allows the exact replication of the strategy of non-utilization of allocated quotas. Consequently, the seller has to choose between the costly inefficient slot usage in terms of babysitting, or the costless hand-back that has no effect on overall output. Obviously, under this premise a hand-back is more attractive than the inefficient usage of the quotas. The isolated application of the hand-back rule hence also does not change the seller's game strategy under the asymmetric initial allocation. In contrast to the symmetric case, this means that strategic behavior does not occur.

As already mentioned, however, the above reasoning is only valid as long as the unused quotas are not reallocated to the constrained competitor after hand-back. Instead, if the regulator considers reallocating some of the unused quotas in order to increase overall output, the airlines' rationales will change. The case of the reallocation after hand-back rule is discussed next.

12.2.2 The Reallocation Rule (after Hand-back)

In practice, unused slots are reallocated across the competitors by a neutral coordinator in an administrative manner after hand-back. While this administrative decision process is hard to formalize in an environment with several potentially distinct airlines, a quota reallocation in this model's two-airline case is straightforward: It is implemented by the assumption of the reallocation rule, which states that the unused quotas that are handed back are subsequently reallocated to the other airline. In the asymmetric case, the beneficiary of this rule is the leisure airline. In the symmetric case, either airline may become the receiver of the reallocated quotas. In both cases, each potential receiver is output constrained, so that he is generally interested in an output expansion.

The implications of this rule on the game strategies of the airlines are the following: In the asymmetric case, the potential seller knows that a hand-back will yield an overall output increase based on its competitor's output expansion when the reallocation rule is imposed. This overall output expansion increases congestion and decreases market power. Therefore, it induces an externality on the quota seller and induces a cost on the quota holder. The quota hand-back hence no longer replicates the holding of excessive unused slots. Instead, it becomes equivalent to a quota trade at a price equal to zero. As a consequence, the quota holder considers using this slot inefficiently in order to avoid a reallocation. This inefficient quota utilization, however, induces the same effect on flight fares and congestion as the competitor's output expansion after the hand-back.[74] The only difference between the two options is that the slot babysitting increases the slot-holder's turnover as compared to the hand-back. The result is thus that the initial holder of the unused quotas will never opt for a hand-back with reallocation, but will always choose to babysit his excess number of slots. Only an actual quota trade that would monetarily compensate the quota holder for his forgone turnover might become an alternative. In the asymmetric case, the potential seller's rationale hence only involves the options of slot babysitting and of a quota trade at a positive price. A quota hand-back with reallocation will never take place.

Note that in the asymmetric case, the reallocation after hand-back rule thus has a counter-intuitive implication: On the one hand, it inhibits the holding of unused excess slots after the initial quota allocation. This game option has been characterized as strategic competition in Section 12.1 above. On the other hand, the reallocation rule motivates the slot-holder to

[74] Note that as opposed to an unconstrained equilibrium, the seller may not endogenously reduce his flight volume after the overall output increase. The endogenous output adjustment is ruled out by the assumption of the trading rule: As it would again produce unused slots, the latter again had to be returned. Although this generally represents an iteration, it is reasonable to disregard it from an equilibrium perspective that only reflects the final outcome of an allocation but not the adjustment path leading to it.

initially keep his excessive number of slots by using them inefficiently rather than to hand them back. Surprisingly, in the asymmetric case, the reallocation rule thus actually introduces strategic behavior in terms of costly slot babysitting. As pointed out above, however, this kind of strategic behavior differs from the potential seller's strategy of hoarding slots after a trade in the symmetric setting. This means that the two types of strategic behavior need to be distinguished.

Also in the symmetric case, the reallocation rule will correct for the ineffectiveness that the hand-back rule exhibits if it is applied on a stand-alone basis; namely, it alters the potential buyer's rationale in a similar way: Because the reallocation after hand-back now implies an overall output expansion, the hand-back no longer replicates holding the purchased quotas unused after a trade. As a consequence, the buyer will have to utilize all his purchased quotas. A quota hand-back, in contrast, reverts to an inferior strategy as it no longer avoids the overall output expansion and thus becomes detrimental to the slot buyer. Ultimately, the reallocation rule in conjunction with the hand-back rule yields the suppression of the post-trading strategic behavior in terms of quota hoarding. Again, note that this contrasts with the asymmetric case, where the rule actually introduces strategic behavior. As a consequence, the application of the trading rules and the occurrence of strategic behavior directly correlate in the asymmetric allocation, but inversely correlate in the symmetric case.

12.2.3 Effects on Slot Trading

In the asymmetric case, the application of the above two trading rules has the following implications for the willingness of the two airlines to pay for a quota trade: On the one hand, the output effect on market power and congestion always occurs when the trading rules are imposed. As a consequence, the buyer only has to compensate the seller for his forgone turnover, but not for the output effect. Similarly, the quota buyer also always experiences the negative externality from the output effect. The benefit of the buyer's higher turnover from a quota trade thus comes at the full expansion of its market share and is not concerned with the higher congestion and the lower market power. As compared to an unregulated asymmetric trading case, the output effect from the seller's inefficient slot utilization depresses the seller's asking price for quota and increases the buyer's potential bid for a slot trade.

On the other hand, however, in the absence of the trading rules, the slot seller remains with his unconstrained equilibrium output minus the endogenous output adjustment from the external effect caused by the buyer's output expansion. With the trading rules, his damages from the individual output reduction amount to one full unit of output. As a consequence, the buyer not only has to compensate the seller for his forgone turnover from the endogenous output adjustment but also for a full unit of output. This considerably increases the seller's asking price for a slot. The overall effect of the trading rules on both airlines' willingness to pay for a quota trade thus remains ambiguous.

In the symmetric case, the application of the trading rules has the following effects: With strategic slot hoarding, the potential buyer has two options: He can either reduce his own turnover and maintain a constant overall output, or he may reduce overall output and thus increase market power and decrease the congestion externality. In the latter case, however, his turnover remains constant. If the trading rules apply, however, the slot hoarding strategy is no longer viable. The effect of the trading rule on the buyer's willingness to pay for a slot thus depends on which one of the two options has been more profitable: If slot hoarding ha been more profitable than utilizing the quotas after a trade, the potential buyer's bid for a quota trade will decrease with the imposition of the trading rules.

In the opposite case, the willingness to pay does not change. As a consequence, the trading rules will either decrease the bid price for the quotas or leave it unchanged. In contrast, the seller's asking price for a slot is not affected by the imposition of the quota rules. This characteristic arises from the property that strategic behavior in the symmetric case occurs after trading. If the quota hoarding were implicitly included in the market price, it would also be beneficial for the seller, constituting a discount on the slot price. However, for simplicity, it is assumed that the quota seller does not ultimately know with certainty whether the buyer will subsequently use the purchased slot or hoard it. Therefore, the seller will not apply the discount from a positive externality. He will ask the price that compensates for his gross output reduction net of any external advantage.

Ultimately, the effect of the trading rules on the likelihood of a quota trade is either negative or neutral, depending on whether hoarding of the quotas after trading were more or less profitable for the buyer than the utilization of the slot.

12.2.4 Implications

The above consideration of the trading rules yields two important implications: First, the reallocation rule is an essential supplement to the hand-back rule: If it is not applied, the hand-back rule will prove useless. As a consequence, from a regulatory view, the hand-back rule must always be applied in conjunction with the reallocation rule. Secondly, the application of the trading rules affects the occurrence of strategic behavior as a function of the initial allocation type: The trading rules suppress strategic behavior in the symmetric case but foster strategic behavior in the asymmetric case. As a consequence, ruling out strategic behavior requires the imposition of the trading rules in the symmetric case but their omission in the asymmetric case. It is therefore important to distinguish between the application of the trading rules and the occurrence of strategic behavior.

In this respect, the welfare analysis in Section 12.5 will show that in either case it is the application of the trading rules rather than the suppression of strategic behavior that is beneficial for allocation efficiency. This result may seem counter-intuitive but is soundly established in the welfare analysis. It yields that, from the regulator's welfare perspective, the application of the trading rules need to be recognized as a dominant strategy.

12.3 Trading Potentials

Secondary trading reflects a market for airport slots, which allows airlines to endogenously adjust the initial quota allocation based on mutual exchange. The initial quota allocations from the previous section are taken as each airline's endowment. Subsequently, the airlines are presumed to have full property rights over their slots and thus are free to trade. This allows the evaluation of whether there is any trading potential from the initial allocation. Trading is assumed not to involve any transaction costs, and clearinghouse issues are not addressed.

12.3.1 Trading Opportunity: Exogenous Shock

When the opportunity for secondary trading becomes possible, the airlines' profit functions need to account for the output changes of both airlines that may be implied by a potential trade. These optimal outputs still are determined by the airlines' reaction functions but they may deviate from the constrained and unconstrained equilibrium outputs. Moreover, the potential monetary impact of the quota trading price needs to be included in the airlines' cost-benefit rationales. As a consequence, the introduction of the secondary trading possibility constitutes an idiosyncratic, exogenous shock to the airlines' profit functions (cf. Corchon, 2001, p.40).[75]

Formally, this shock implies that the profit function differs from the generic profit function (3) without trading. The potential for a quota trade is revealed based on the first-order conditions of this new profit function. It is referred to as the trading potential. Whether

[75] The opposite of an idiosyncratic shock is a generalized shock. It is defined as to imply the same effect $\mathrm{d}n_p^i = \mathrm{d}n_p^j$ on both airlines (again see Corchon, 2001, p.40).

secondary trading takes place or not is evaluated following a two-step procedure, as proposed by Verhoef (2010, pp.326). This procedure investigates whether the necessary conditions required for potential trading are fulfilled and evaluates whether the sufficient condition for secondary trading is met. For this purpose, the amended profit function with secondary trading is presented first. Thereafter, the trading potential is derived based on the total differential of this profit function. Finally, whether a slot trade actually takes place or not is evaluated based on the corresponding necessary and sufficient conditions for secondary trading.

12.3.2 Airline Profit Function with Secondary Trading

As a starting point, airline profits $\Pi_i\left[n_i(n_j), n_j(n_i)\right]$ from (3) represent each airline's original profit function without trading.[76] The profit function with secondary trading is derived from this original profit function by introducing a potential quota transfer $\triangle q$ and the corresponding peak-period output changes $\mathrm{d}n_p^i(\triangle q)$ and $\mathrm{d}n_p^j(\triangle q)$ as

$$\Pi_i^{ST}\left[n_p^i\left(n_p^j, \triangle q\right), n_p^j\left(n_p^i, \triangle q\right), \triangle q\right] \equiv \Pi_i\left[n_i\left(n_j, \mathrm{d}n_p^i(\triangle q)\right), n_j\left(n_p^i, \mathrm{d}n_p^i(\triangle q)\right)\right] - P \cdot \triangle q. \tag{28}$$

In this profit function, a secondary trade is defined as a transfer of quota volume $\triangle q$ from the seller to the buyer. This trade implies a positive net monetary transfer from the seller to the buyer denoted as trading price P. The sign of total trade profit $P{\cdot}\triangle q$ is negative and inversely related to the sign of $\triangle q$. This means that a purchase of slots implies a positive number of quotas $\triangle q$ for the airline concerned but a negative transfer of funds. Correspondingly, a slot sale includes a negative number $-\triangle q$ of traded quotas, but affects profits in a positive way. The market price $P \in \mathbb{R}^+$ is strictly positive and exogenous to the computation.[77]

The impact of a quota trade $\triangle q$ on outputs is captured by each airline's peak-period output changes $\mathrm{d}n_p^i(\triangle q)$ and $\mathrm{d}n_p^j(\triangle q)$. These output changes need not necessarily correspond to the traded volume $\triangle q$: As developed above, the presumptions about slot usage, hand-back and reallocation subsumed into the trading rules. Application of these trading rules now characteristically determines the distinct strategic options and thus the potential trading outcomes for each airline. Therefore, the respective output changes are determined by the output functions$\mathrm{d}n_p^i(\triangle q)$ and $\mathrm{d}n_p^j(\triangle q)$, rather than by the traded volume$\triangle q$ itself. The output functions are therefore referred to as the trading properties of a secondary trading

[76] For notational brevity, both airlines' individual outputs are simply written as vectors n_i and n_j here.

[77] This computation will ultimately not resolve an actual slot trading price; see the limitations in Section 12.3.7.

scenario. They depend on the respective initial allocations in conjunction with the trading rules that are in force.

The relation of the actual output changes to the trading volume is indirectly determined by the trading properties. As a consequence, the potential trading volume $\triangle q$ affects the profit functions both as an actual output change in the case of trading and in terms of opportunity costs from strategic behavior, regardless of whether a trade takes place or not. This makes clear why the shock on the airlines' profit functions is classified as idiosyncratic: It affects both airlines in a different way, first depending on whether the trade is a sale or a purchase and second as a function of the trading properties based on the different, potentially asymmetric trading rules.

The different strategic options that arise for both airlines under the secondary trading scheme depend on which trading rules are imposed. In other words, the output changes following a trade are specific to each distinct trading scenario. This means that the exogenous shock to the profit function is formally different for each case considered. As a consequence of the distinct shocks, the outcomes also may differ. The impact on secondary trading under the distinct trading rules and initial allocations hence needs to be considered on a case to case basis.

The impact of secondary trading on the constrained equilibrium ultimately arises from the exogenous shock on the airlines' profit functions. Formally, this means that the previous steady-state values $n^i\left(n^j\right), n^j\left(n^i\right)$ from the constrained equilibrium are generally no longer optimal for each airline when trading is allowed. The new equilibrium with the trading possibility is based on peak-period outputs $n_p^i\left(n_p^j, \triangle q\right)$ and $n_p^j\left(n_p^i, \triangle q\right)$, which are endogenous functions of the trading volume. Based on the alteration of the profit function, these outcomes generally differ from the constrained equilibria based on the two distinct initial quota allocations. The impact of secondary trading on these constrained equilibria can be investigated by evaluating the optimality conditions of the new profit functions at the steady state values of the original constrained equilibrium without the trading option. This method is referred to as to comparative statics and is standard in the economic analysis of theoretical models (cf. e.g. Corchon, 2001, p.35). The starting point for the comparative statics is the total differential of each airline's profit function with regard to peak-period outputs.

12.3.3 The Total Differential: Gains and Losses from Quota Trading

For comparative statics, the potential gain or loss of a quota trade can be expressed by the total differential of each airlines profit function (28) with regard to all outputs. It is symmetric across the airlines, and in generic notation reads

$$
\begin{aligned}
\mathrm{d}\Pi_i^{ST} \equiv \frac{\mathrm{d}\Pi_i^{ST}}{\mathrm{d}n_p^i} \cdot \mathrm{d}n_p^i + \frac{\mathrm{d}\Pi_i^{ST}}{\mathrm{d}n_p^j} \cdot \mathrm{d}n_p^j - P \cdot \mathrm{d}q = \\
\left[\frac{\partial \Pi_i^{ST}}{\partial n_p^i} + \frac{\partial \Pi_i^{ST}}{\partial n_p^j} \cdot \frac{\partial n_p^j}{\partial n_p^i}\right] \cdot \mathrm{d}n_p^i + \left[\frac{\partial \Pi_i^{ST}}{\partial n_p^j} + \frac{\partial \Pi_i^{ST}}{\partial n_p^i} \cdot \frac{\partial n_p^i}{\partial n_p^j}\right] \cdot \mathrm{d}n_p^j - P \cdot \mathrm{d}q. \quad (29)
\end{aligned}
$$

The differentiation of the now three-dimensional profit function requires application of the chain rule, because each airline's own output is a reaction function of its competitor's output (cf. e.g. Mas-Colell et al., 1995, p.927). The output changes $\mathrm{d}n_p^i$ and $\mathrm{d}n_p^j$ again denote the trading properties that depend on the distinct trading rules across the settings. The differential term $\mathrm{d}q$ denotes the actual number of traded quotas. It corresponds to $\mathrm{d}q = \triangle q$ for a quota purchase, and to $\mathrm{d}q = -\triangle q$ for a quota sale.

The interpretation of total differential (29) is the following:

The first term denotes the effect of each airline's own output change $\mathrm{d}n_p^i$ on its profits. Examining profit function (3) makes clear that this term itself consists of two distinct parts, as shown in the second line of equation (29). The first partial derivative reflects the direct effect of the output change. This direct effect consists of the change in flight fares, congestion, and turnover based on the output volume that is reduced or increased. The second partial derivative denotes the indirect effect on airline i's profits that arises from the endogenous output adjustment of competitor j. This output adjustment is directly invoked by airline i's output change, as dictated by the Cournot reaction functions that define each airline's output as a best response function to the other airline's output. In turn, this output compensation of the competitor again affects congestion and flight fares and thus also affects airline i's profits. In sum, the first term captures both the indirect and the direct effects on airline i's profits from a change in its own output.

The second term denotes the externality on airline i's profits from a change in its competitor's output $\mathrm{d}n_p^j$. In a similar manner to the first term, it also includes two distinct parts, as shown by the two different partial derivatives in brackets. The first partial derivative reflects the direct external effects of the competitor's output change on flight fares and congestion, which affects both airlines' profits. The second partial derivative captures the corresponding indirect effect of this externality on the output of airline i: Similar but opposite to the above, the alteration in market power and the congestion externality again induces an endogenous output adjustment of airline i, as defined and explored in Section 8.3. This endogenous output adjustment is formally denoted by the partial derivative $\frac{\partial n_p^i}{\partial n_p^j}$ and again affects the profits. The indirect external effect hence transmits the change in flight fares and congestion through the endogenous output adjustment to the profits of airline i. It represents the second-order effect of competitor j's output change on airline i's output, based on the primary change in the dual distortion. Ultimately, the second term includes airline i's own readjusting effect on congestion, flight fares and turnover following an output change of its competitor.

The third term simply denotes the impact of the trading price on profits. As explained above, its impact depends on the sign of the quota transfer, i.e., whether airline i is a seller or a buyer of quotas. Note that the total differential is generally valid for both airlines $i \in \{B, L\}$. The definition of airline i as the seller and airline j as the buyer follows later.

The total differential hence captures all effects on the airlines' profits that arise according to the respective output changes following a quota trade, including the secondary output effects based on endogenous output adjustments as invoked by changes in the dual distortion. The partial derivatives formally denote the marginal profit changes for the respective marginal output changes of each airline. In this respect, it is important to note that both the sign and the absolute magnitude of the partial derivatives in (29) depend on the actual output values at which they are evaluated. Particularly, an airline that remains unconstrained after the initial quota allocation can still maximize its profits based on an endogenous output choice. In this case, the first-order conditions from unconstrained profit maximization dictate that the above partial derivatives equal zero. Consequently, all profit changes for any potential slot trade would be zero. Hence, considering the unconstrained steady state alone does not allow us the application of comparative statics.

The equilibria with the trading option generally differ from the unconstrained equilibria without trading. This difference is based on the fact that the profit functions with secondary trading are distinct from the profit function without secondary trading. In turn, the above partial derivatives from (29) are derived from the profit function with secondary trading, but are evaluated at the original, profit-maximizing steady state before trading as based on the constrained equilibrium. As a consequence, the partial derivatives in (29) generally no longer equal zero at the initial allocation because optimality after the shock is distinct from optimality before the shock. The investigation of the impact of secondary trading on the constrained equilibrium hence becomes available and can be performed by means of comparative statics.

The above consideration makes clear that the partial derivatives in (29) should formally be expressed e.g. as

$$\frac{\partial \Pi_i^{ST}}{\partial n_p^i} \equiv \frac{\partial \Pi_i^{ST}[n_p^{i*}, n_p^{j*}]}{\partial n_p^i}.$$

Outputs n_p^{i*} and n_p^{j*} denote the steady state outputs from the equilibrium after the initial allocation, as defined by profit function (3) without secondary trading. For notational convenience, however, the derivatives will only be printed in short form. Moreover, note that the partial derivatives of equilibria that are constrained by the initial quota allocation are generally distinct from zero already because those constrained steady states do not fulfill the first-order conditions for profit maximization by definition. Nevertheless, these equilibria also need to be evaluated based on the profit function that includes the shock in order to correctly assess the trading potential.

As explained above, the characteristic properties for each trading situation are defined in terms of the corresponding primary output changes $\mathrm{d}n_p^i$ and $\mathrm{d}n_p^j$. The actual size of these output changes depends both on the trading properties and on the volume of the quota trade itself. The evaluation of these output changes in the profit function with secondary trading yields the corresponding potential gains or losses for each airline that would occur with a trade. The gains and losses net of trading price P ultimately represent the trading potential of each airline. The formal derivation of the trading potential is therefore provided in the following subsection. The distinct trading properties and the corresponding actual trading potentials are distinct across the different cases considered. They are presented in the subsequent investigation of the different initial allocation and trade settings.

12.3.4 Necessary Condition: Positive Trading Potential

The trading potential measures the gross potential profit change that becomes possible based on a trade, net of the trading price. This profit change can simply be expressed based on the above total differential (29) by setting $P = 0$ as

$$\mathrm{d}\Pi_i^{ST}\Big|_{P=0} = \frac{\mathrm{d}\Pi_i^{ST}}{\mathrm{d}n_p^i} \cdot \mathrm{d}n_p^i + \frac{\mathrm{d}\Pi_i^{ST}}{\mathrm{d}n_p^j} \cdot \mathrm{d}n_p^j. \tag{30}$$

The substitution $\mathrm{d}n_p^i \to \mathrm{d}n_p^i(\triangle q)$ and $\mathrm{d}n_p^j \to \mathrm{d}n_p^j(\triangle q)$ of the trading properties according to the specific trading rules allow the expression of this total profit change as a function of the potential trading volume $\triangle q$ only. With the trading properties as explicit functions, division of both sides of (30) by trading volume $\triangle q$ yields the trading potential as

$$TP_i(\triangle q) \equiv \frac{\mathrm{d}\Pi_i^{ST}\Big|_{P=0}}{\triangle q}. \tag{31}$$

This means that the trading potential is positive when the gross gain from a secondary trading equilibrium exceeds the profits of the equilibrium without trading. The trading price is set to zero in order to consider the trading potential regardless of the actual market price for a slot. This means that if trading is not worthwhile at a zero market price (i.e. an endogenous output adjustment according to the reaction functions), it would certainly not be so at a positive market price. The fact that the trading properties are characteristic and thus distinct in each trading case also shows that the trading potential formally differs across cases.

It is important to stress again that the trading potentials are defined as potential gains or losses from a slot *purchase*. Therefore, a trade may only become mutually beneficial if both airlines have a *positive* trading potential. This yields the necessary condition for secondary trading as

$$TP_i, TP_j > 0.$$

As described above, the trading potential is defined as each airline's gains and losses based on the corresponding output adjustments net of trading price P. Put differently, a positive trading potential denotes a positive profit change of the output changes following a trade as a function of volume $\triangle q$ of that trade and irrespective of the actual magnitude of the market price. Finally, it is important to note that the trading potential is defined for a slot purchase at price zero, formally denoted as $\mathrm{d}q = \triangle q$ and $P = 0$. The trading potential for a slot *sale* thus requires considering $\mathrm{d}q = -\triangle q$, which is reached by a sign-change $TP_i(-\triangle q) = \frac{\mathrm{d}\Pi_i^{ST}(\triangle q)}{-\triangle q} = -\frac{\mathrm{d}\Pi_i^{ST}(\triangle q)}{\triangle q} = -TP_i(\triangle q)$.

The above necessary condition becomes clearer when a negative trading potential is considered. This would signify that an airline could yield a profit by decreasing its output and thus giving away some of its used quotas. A quota sale following a negative trading potential would even permit the selling airline to pay for that trade. An output reduction, however, can be performed at any time by an airline free of any direct cost. A negative price for slots would hence never occur. Ultimately, a trade at a negative trading potential corresponds to an output adjustment of an airline according to its reaction function. The opportunity cost is the corresponding output compensation of the competitor and the market price remains zero. This means that a negative trading potential of any airline would lead to a trade at price zero, which cannot be considered as a trade.

12.3.5 Sufficient Condition: Positive Market Price

The above necessary conditions dictate that a trade needs to be worthwhile for both the seller and the buyer, which happens if both trading potentials for a slot purchase are strictly positive. The sufficient condition evaluates whether the trading potentials allow for a mutually beneficial slot exchange at a positive market price.

Firstly, let us define that airline i is the potential seller and airline j the potential buyer of slots $\triangle q$. Moreover, let us presuppose that both airlines exhibit a positive trading potential $TP_i, TP_j > 0$. For trading to actually take place, the potential gain from a trade for the potential buyer must at least equal the potential loss of the seller for his output contraction following the trade. If this is the case, trading becomes mutually beneficial and will take place. As a consequence, a positive transfer price emerges that overcompensates both airlines for their profit changes.

Formally, the sufficient condition for trading is developed as follows: First, recall that the trading potential measures the profit gains from a slot purchase, where a trading potential larger than zero signifies that both airlines experience a profit gain from an expansion of their own output. For the potential slot buyer with $TP_j > 0$, the trade becomes worthwhile as long as the profit gain from his output expansion is not compromised by the market price P. In other words, the net gain from the trade incurred by the buyer must be strictly positive. This can be expressed as $TP_j - P > 0$, which implies that the market price needs to satisfy $TP_j > P$.

In contrast, the positive trading potential for a slot purchase implies that the seller actually incurs a loss from his output reduction if he sells some of his quotas. The loss from the sale thus is simply quantified by the negative amount of the trading potential $-TP_i$. This means that for the seller, the benefit of the monetary transfer P (i.e. the market price) must at least compensate him for this profit loss. The seller hence trades if $-TP_i + P > 0$, which dictates that the market price must satisfy $P > TP_i$.

Concerning the market price P, the above two relations between the trading potentials and the market price yield the sufficient condition for secondary trading to take place as

$$TP_j > P > TP_i. \tag{32}$$

The fact that the necessary conditions dictate that both trading potentials need to be positive also shows that any market price from a trade will be strictly positive; or, put differently, the above argument for the non-existence of negative prices shows that any negative trading potential would never yield a trade.

In sum, for trading to occur, buyer j must have at least the same potential profit from increasing his output by $\mathrm{d}n_p^j$ following a trade $\triangle q$ as the profit loss incurred by the seller i from giving away his number of slots $\triangle q$ and the corresponding output reduction $\mathrm{d}n_p^i(\triangle q)$. The occurrence of trading is hence simply revealed by evaluation of equation (32).

12.3.6 Symmetric vs. asymmetric Trading

Recall again that two distinct initial allocations are considered: A symmetric case in which both airlines are a potential buyer or seller, and an asymmetric case where only the constrained airline can be the buyer and the unconstrained airline becomes the potential seller. Generally, the sufficient condition for trading is the same for both cases. Nevertheless, there is a crucial difference in the evaluation of secondary trading, depending on whether the situation is asymmetric or symmetric.

In the symmetric case, the seller and the buyer are revealed by comparison of the trading potentials: For a trade at a positive market price, the airline with the higher trading potential becomes the buyer, and the airline with the lower trading potential becomes the seller. For the airline duopoly considered in this study, the necessary and the sufficient conditions hence imply that trading takes place if

$$TP_B, TP_L > 0 \quad \text{and} \quad TP_B \neq TP_L. \tag{33}$$

This means that secondary trading occurs as long as the two trading potentials are positive but not equal. As a consequence, the sufficient condition (32) is always well ordered.

By contrast, in the asymmetric case, the initial quota allocation dictates that one airline is expelled from the peak period. This already determines that only the unconstrained airline holds slots and may become the seller. The constrained airline does not own any slots and thus can only become the buyer. Sufficient condition (32) hence cannot imply that the higher trading potential becomes the buyer. Instead, the seller and the buyer are already determined, so that their trading potentials need to be substituted into (32) in the proper order. The sufficient condition for trading to take place in the asymmetric case therefore is more restrictive than in the symmetric case. With this modification, both the necessary and the sufficient conditions can formally be combined into a single secondary trading condition for the asymmetric case as

$$TP_L > TP_B > 0. \tag{34}$$

This condition is both necessary and sufficient. As a consequence, the inequalities may not be fulfilled and trading may not take place even if both trading potentials are strictly positive and distinct.

12.3.7 Limitations

Three important limitations emerge from the above computation and evaluation of the trading potentials: First, recall that the market price P denotes the positive monetary transfer that is paid from buyer j to seller i in exchange for the airport slots if the trade takes place. Also note that trading condition (32) does not unambiguously determine the magnitude of this market price P. Rather, the actual market price will emerge somewhere in between the two trading potentials, depending on the bargaining power of the two participants. This makes clear that the model actually abstracts from any specific pricing process. To overcome this problem in a simple manner, the transfer price can be assumed to be the average of the two trading potentials, so that both airlines share the excess profit from the trade.[78]

Moreover, recognize that comparative statics only reveal the trading potential for one specific output allocation. In reality, however, the trading potential changes with every trade, so that the marginal price for every additional slot traded would be different. Because the model is time invariant and because output is continuous and thus slot trades cannot be evaluated sequentially and on a discrete scale, the analysis is restricted to the evaluation of the trading potential for the steady state situations of the initial allocations. In other words, the number $\triangle q$ of slots traded is generally not a discrete quantity but only reflects the marginal trading potential based on the initial quota allocation. As a consequence, the net transfer does not reflect an actual market price for a single slot.

Lastly, the above consideration also implies that the trading potential actually depends on the trading volume: While a slot trade might take place for a small number of quotas, it might not be worthwhile for a large volume of slots. In other words, the trading potential might become exhausted for an excessively large output change and trading would not occur. However, this analysis does not explicitly determine the actual number of traded slots $\triangle q$. Nonetheless, based on the consideration of marginal profit gains or losses from trading, the above problem is solved by simply assuming that the trading volume $\triangle q$ always remains sufficiently small to permit emergence of a positive market price P in the trading conditions whenever the marginal trading potentials are sufficient to allow for a trade. This also circumnavigates the problem of considering whether a sufficient number of slots existed for the trade at all.

Ultimately, the above discussion illustrates that the subsequent analysis is limited to evaluating whether trading would take place at the margin of the corresponding initial allocation. Nonetheless, the above simplifications are sufficient to investigate and evaluate the general

[78] The abstraction from a specific pricing scheme also follows Verhoef (2010) and Basso and Zhang (2010). A pricing scheme without too much complexity could be implemented by linearization of the trading potential around $\bar{n}_p^i$.

welfare effects of a secondary trading scheme after the initial allocation of the individual quota.

12.4 Results: Trading Potentials

In the following, the four cases that arise from the two distinct initial allocations and the application of the above trading rules are investigated for their trading potential. The four cases consist of the asymmetric and the symmetric initial allocation, where each allocation is evaluated with and without strategic behavior.

12.4.1 Asymmetric Trade without Strategic Behavior

The asymmetric case starts with the individual quota allocation from Section 11.1. It therefore presumes that only the business airline owns airport quotas that can be offered for sale. These quotas are previously unused and thus denote excessive airport capacity. The competing leisure airline is completely expelled from the peak period and can only become the buyer. As defined in Section 12.3.2 above, a secondary trade is defined as a quota transfer from a seller i to a buyer j. In order to maintain generalizability, this analysis again is provided in generic notation, although the seller and the buyer are unambiguously determined in the asymmetric case of this model. The characteristic properties and trading rules for this trading case are summarized in Table 12.1.

In addition, it is assumed that strategic behavior does not occur. As revealed in Section 12.2, this requires that the trading rules of hand-back and reallocation are absent. This ensures that strategic behavior cannot occur. Alternatively, the same setting is reached if a use obligation is in place but not followed by a reallocation rule.

ASY	**Seller** i	**Buyer** j	**Remarks**
Initial Allocation:	$\hat{q}_i > n_p^i(\hat{n}_p^j) > 0$	$\hat{q}_j = \hat{n}_p^j = 0$	$j = L$: constrained
Trading Rules:	None (or: Hand-Back *without* reallocation)		
Trading Options:	- hold unused or hand back - sell unused: $\triangle q = \hat{q}_i - n_p^i$	- none - buy and use	 no Strategic Behavior
Trading Properties:	$\mathrm{d}n_p^i = 0$	$\mathrm{d}n_p^j = \triangle q$	due $\triangle q$ prev. unused
Overall Output:	$0 < \mathrm{d}N_p = (1 + {\partial n^i}/{\partial n_p^j}) \cdot \mathrm{d}n_p^j < \triangle q$		due ${\partial n^i}/{\partial n_p^j} \cdot \mathrm{d}n_p^j < 0$
Conj. Variations:	${\partial n_p^i}/{\partial n_p^j} < 0$	${\partial \hat{n}_p^j}/{\partial n_p^i} = 0$	due $\hat{n}_p^j = 0$
Profit Changes from			
- Output Change:	$\frac{\partial \Pi_i^{ST}}{\partial n_p^i} < 0$	$\frac{\partial \hat{\Pi}_j^{ST}}{\partial n_p^j} > 0$	due j = constrained
- External Effect:	$\frac{\partial \Pi_i^{ST}}{\partial n_p^j} < 0$	$\frac{\partial \hat{\Pi}_j^{ST}}{\partial n_p^i} < 0$	

Tab. 12.1: Asymmetric Trading Properties (without Strategic Behavior)

Above all, the initial allocation in the asymmetric quota allocation dictates that the non-networking airline be completely expelled from the peak period. Its number of slots and thus its peak period output is zero, and it becomes the potential buyer. As a consequence, the networking airline chooses its monopoly output, providing an inefficiently small network. Therefore, some quotas remain unused, so that the business airline is factually unconstrained (see 11.1.2).

When strategic behavior is ruled out, the options for the seller are either to keep his unused slots or sell them. Correspondingly, the potential buyer can either abstain from a trade and remain with zero output or buy the unused slots. If he buys them, he will always use them even if there is no trading rule that obligates him to do so (see Section 12.2).

The trading properties from a quota trade $\triangle q$ thus are the following: If a trade takes place, then buyer j receives a positive number of slots. The corresponding output change amounts to $\mathrm{d}n_p^j = \triangle q_j$. In contrast, the trade asymmetry reflects that the traded quotas was not used by seller i before the trade. Consequently, the seller is left without a direct output change from the trade. His trading property is therefore defined as $\mathrm{d}n_p^i = 0$. Nonetheless, he incurs an endogenous output adjustment as a second-order effect from the buyer's output expansion. This is reflected by property ${\partial n^i}/{\partial n_p^j} < 0$ in Table 12.1. As previously explained, the total differential ensures that this secondary effect is included in the trading potential.

The profit changes from a trade for each airline are explained by the partial derivatives from the profit function with secondary trading, which are evaluated at their pre-trading equilibrium outputs. Because the buyer j is output constrained, he earns a higher profit through any output expansion, denoted as $\frac{\partial \hat{\Pi}_j^{ST}}{\partial n_p^j} > 0$. The profit function is denoted with a hat sign in order to stress that profits are limited by constrained output $\hat{n}_p^j$ and hence are not maximized. In contrast to the buyer, the seller balanced his output against congestion and market power before the introduction of the secondary trading opportunity. With the trading opportunity, he is thus confronted with an increase in overall output, as the formerly unused quotas are now used by the competitor to increase his own output. After trading, congestion will therefore be higher and market power lower. As a consequence, the seller is able to increase (or to reduce the decline of) his profits by decreasing his own output only, so that the partial derivative of his profits with regard to his own output is negative, as denoted by $\frac{\partial \Pi_i^{ST}}{\partial n_p^i} < 0$. Lastly, an output expansion of each airline's competitor induces a negative externality both from higher congestion and lower market power. As a consequence, each airline's profits are inversely related to this output expansion, so that $\frac{\partial \Pi_i^{ST}}{\partial n_p^j} < 0$ and $\frac{\partial \hat{\Pi}_j^{ST}}{\partial n_p^i} < 0$.

The reaction functions determine the direction of the endogenous output adjustments following every competitor's output change. As developed in Section 8.3, these endogenous output adjustments are inversely related to the competitor's output change, so that ${}^{\partial n^i}/_{\partial n_p^j} < 0$ for the seller i. In the asymmetric case, however, there is a special case arising from the fact that the potential buyer j is constrained to zero peak-period output: On the one hand, the buyer cannot physically reduce his own output; on the other hand, he can only increase his output if he purchases a positive number of quotas. This means that he cannot endogenously adjust his output at all. With constrained output denoted as $\hat{n}_p^j$, therefore, ${}^{\partial \hat{n}_p^j}/_{\partial n_p^i} = 0$.[79]

The effect of a trade on output will therefore be the following: The buyer will fully use all the slots purchased, while the seller will endogenously reduce his output due to higher congestion. As a consequence, overall output after a potential trade will be higher than before the trade but not to the full number of quotas traded. Whether the trade takes place thus depends on the trading potentials of the two airlines.

The trading potential for both airlines is revealed by the substitution of above trading properties $\mathrm{d}n_p^i = 0$ and $\mathrm{d}n_p^j = \triangle q$ into (30) and division of both sides by $\triangle q$. This shows that for the seller i, the differential of his own profits with regard to a change in his competitor's output is relevant. However, recognize that this trading potential denotes a slot sale, although (31) dictates that the former is defined as the potential profit gain for a slot purchase. Therefore, to ensure compliance with sufficient condition (32), the trading potential from the corresponding total profit change in the asymmetric case requires a sign change. This converts the seller's potential loss from his competitor's output increase into a positive

[79] Verhoef's (2010) setting contrasts with this case because it assumes demand to be inelastic. Consequently, the endogenous output adjustment is suppressed and the seller's output remains constant. The impact of the trade on congestion is thus higher.

asking price for a corresponding output reduction - which is theoretical, because the buyer cannot reduce output below zero in practice.

The asymmetric seller i's trading potential hence amounts to

$$\mathrm{TP}_i^{ASY} = \frac{\mathrm{d}\Pi_i^{ST}}{\triangle q} = -\left(\frac{\partial \Pi_i^{ST}}{\partial n_p^i} \cdot \frac{\partial n_p^i}{\partial n_p^j} + \frac{\partial \Pi_i^{ST}}{\partial n_p^j}\right), \tag{35}$$

where superscript ASY denotes the characteristic asymmetric case without strategic behavior. The first term denotes seller i's own endogenous output adjustment following the higher congestion. As the partial derivatives in Table 12.1 above show, this negatively affects airline i's profits because it reduces the seller's own turnover. Recalling the sign change in order to denote a theoretical quota purchase, however, this term becomes positive. It hence denotes that the seller's trading potential further increases along with its endogenous output adjustment. Consequently, the second term denotes the external effect of the trade on the seller's profits. For a slot sale, this effect is negative as it reflects the higher congestion and the lower flight fares caused by the competitor's output expansion (see Table 12.1). However, with the sign change, this term also positively affects the trading potential, which shows that the seller's asking price increases with the damage from the externality from a trade. In sum, the sign of (35) is positive, which means that the asymmetric slot seller i exhibits a positive trading potential.

Equivalently, the trading potential of the buyer j is revealed as

$$\mathrm{TP}_j^{ASY} = \frac{\mathrm{d}\hat{\Pi}_j^{ST}}{\triangle q} = \frac{\partial \hat{\Pi}_j^{ST}}{\partial n_p^j} + \frac{\partial \hat{\Pi}_j^{ST}}{\partial n_p^i} \cdot \frac{\partial n_p^i}{\partial n_p^j}. \tag{36}$$

As opposed to equation (35), this total derivative is expressed as a profit change following the buyer's own output change. Consequently, the first term denotes the direct effect of the buyer's output expansion. This term has a positive sign because the buyer's output was constrained before the trade. The second term yields the external effect on buyer j's profits from seller i's endogenous output contraction. Because the latter decreases congestion and increases flight fares, this term is also positive. Hence, overall the trading potential of the asymmetric buyer is unambiguously positive as well.

The two positive trading potentials imply that the necessary condition for secondary trading is fulfilled. Whether trading takes place hence depends on the fulfillment of sufficient condition (34). As delineated in Section 12.3.5, this requires that the buyer's trading potential exceeds the seller's trading potential, so that $\mathrm{TP}_j^{ASY} > \mathrm{TP}_i^{ASY}$. This formally illustrates strategic competition for airport slots as a key business asset: The slot-holder as a potential seller evaluates whether a quota trade with his competitor is more beneficial than leaving the slots unused.

Whether this condition is met ultimately depends on the absolute magnitudes of the trading potentials. However, as the business airline starts from its monopoly output, there is no reason why it should allow for overall output to expand while at the same time endogenously decreasing its own output: The only compensation for its higher congestion damage, lower flight fares and lower turnover would be a positive trading price. However, with an expansion of overall output the total industry profits decline.[80] Consequently, the leisure airline will never be able to afford overcompensating the business airline. Trading will hence never take place in the asymmetric case without hand-back and reallocation. This signifies that a secondary trading scheme cannot correct for the inefficiency arising from the initial allocation of the individual quotas. The market structure therefore remains a monopoly, with the business airline as the only provider.

Note that this result corresponds to Verhoef's (2010) case, where a more profitable airline always buys its competitor out of the market. Subsequently, the slot buying airline reverts either to its monopoly output or to the maximum output allowed by the number of quotas, whichever is lower. As a consequence, some quotas may also remain unused.

12.4.2 Asymmetric Trade with Strategic Behavior: Pre-trading Babysitting

Strategic behavior is introduced in an asymmetric trade if the hand-back rule with reallocation is imposed. As discussed above, this combination implies that the seller has to decide what to do with his excess number of slots: Either he can sell them, which will put them in use in favor of his competitor. As the only alternative, he has to utilize the slots himself but in inefficient manner, which is referred to as babysitting. Hoarding the excess unused slots is no longer an option in this setting. The trading properties are thus $\mathrm{d}n_p^i = \triangle q$ and $\mathrm{d}n_p^j = 0$ after the imposition of the trading rules, and $\mathrm{d}n_p^i = -\triangle q$ and $\mathrm{d}n_p^j = \triangle q$ for a subsequent trade. Initially, overall output increases by $\mathrm{d}N_p = \triangle q$, but remains constant with a trade at $\mathrm{d}N_p = \mathrm{d}n_p^i + \mathrm{d}n_p^j = 0$. This signifies that both trading and babysitting invoke a cost to both airlines (see Section 12.2.2).

[80] This is a standard result in economic theory: The quantity distortion of a monopolist based on the downward sloping demand assumption (cf. Mas-Colell et al., 1995, p.385) yields that any increase in the production quantity above the monopoly output must result both in a lower market price and at the same time in lower overall profits. Otherwise, the monopoly output could not be profit-maximizing. This result is valid for nondiscriminatory pricing only.

ASB	Seller i	Buyer j	Remarks
Trading Rules:	Hand-Back *with* Reallocation		
Airline Reaction:	Babysitting	-	Strategic Behavior
Output Effects:	$\mathrm{d}n_p^i = \triangle q$	$\mathrm{d}n_p^j = 0$, $\partial \hat{n}_p^j / \partial n_p^i = 0$	$\triangle q = \hat{q}_i - n_p^i$
Overall Output:	$\mathrm{d}N_p = \triangle q$		
Trading Options:	sell or hand-back	get (for free) or buy, and use	
Trading Properties:	$\mathrm{d}n_p^i = -\triangle q$	$\mathrm{d}n_p^j = \triangle q$	
Overall Output:	$\mathrm{d}N_p = \mathrm{d}n_p^i + \mathrm{d}n_p^j = 0$		
Conj. Variations:	$\partial n^i / \partial n_p^j = 0$	$\partial \hat{n}_p^j / \partial n_p^i = 0$	

Tab. 12.2: Asymmetric Trading Properties (with Strategic Behavior)

The trading properties of this case are reflected in Table 12.2. The structural difference against the asymmetric case without strategic behavior concerns the endogenous output changes: Because holding slots unused is no longer possible, the seller cannot endogenously reduce his output in order to decrease the output effect when he sells part of his quotas to the competitor. This signifies that the cross-elasticity of outputs is zero for both airlines, which is formally indicated as $\partial \hat{n}_p^j / \partial n_p^i = \partial n_p^j / \partial n_p^i = 0$. For the business airline, this presumption is based on the trading rules as from an equilibrium perspective it is not reasonable to model the adjustment path of a trade following an endogenous output compensation and its iteration. For the leisure airline, it arises from the fact that its peak-period output is already restricted to zero. As a consequence, both airlines will encounter the same output effect whether trading occurs or not: If a trade occurs, it occurs as an external effect from the formerly unused quotas that are now utilized by the competitor. If babysitting takes place, it amounts to the same amount of additional congestion as the latter is not compensated by an output reduction of the competitor. Formally, both airlines hence enjoy a full output compensation from a trade. Therefore, the overall output effect vanishes. The initial allocation and the profit changes remain the same as in the asymmetric case without strategic behavior (see Table 12.1) and are therefore not depicted.

Seller i's decision rationale now involves a two-step decision: First, the quota holder has to evaluate whether babysitting the excessive quotas is more or less expensive than a hand-back of the surplus quotas with a subsequent reallocation to his competitor. In this respect, the cost of babysitting corresponds to the expansion of the slot-holder's own output. On the downside, this decreases flight fares and increases congestion. On the upside, turnover is increased. Nonetheless, the overall effect on seller i's profits would be negative because before the imposition of the trading rules his unconstrained output complied with the first-order condition for profit maximization. By contrast, the cost of the hand-back with reallocation to its competitor amounts to the externality caused by the competitor's output augmentation. This output expansion yields exactly the same increased level of congestion and the same lower flight fares because any endogenous output adjustment is inhibited by the trading rule itself. The externality of the competitor's output change hence exactly equals the damage caused by the seller's own output expansion. However, the difference is that the seller's turnover increases in the babysitting case, whereas it is reduced in the hand-back with reallocation case. This signifies that babysitting is always less expensive for the slot-holder than reducing his own output and coping with the competitor's output expansion. As a consequence, the potential seller will never opt for a hand-back but will always increase his output in order to fully utilize all his quotas. The trading rules themselves hence invoke strategic behavior in the way that seller babysits his excess number of airport slots (see also Section 12.2.2).

Once the potential seller babysits his slots, he may offer his inefficiently used quotas for sale. In that case, he would ask a price that would at least compensate him for his reduced turnover. Observe that this price is net of the external damage from the competitor's output increase because this damage is of the same amount as the indirect babysitting costs. Again, with a sign-change in order to comply with definition (31), the trading potential of seller i amounts to

$$TP_i^{ASB} = {}^{\partial \Pi_i^{ST}}/_{\partial n_p^i} - {}^{\partial \Pi_i^{ST}}/_{\partial n_p^j}.$$

The first term denotes the total effect from an output expansion, which is generally negative because the seller's output was profit maximizing before the trade. In particular, it includes the deteriorating effect of increased congestion and decreased flight fares from the higher output of a theoretical slot purchase. Correspondingly, the second term separately denotes the external benefit from the competitor's output contraction. As already explained, this term does not only correspond to a partial endogenous output compensation but also to a full output compensation. Consequently, the term itself is negative and hence has a positive net effect on the trading potential. The trading potential thus denotes the net gain of a higher turnover at a constant overall output. Therefore, the seller's trading potential must be positive.

Similarly, for the potential buyer j, the trading potential amounts to the profit that can be made from the additional output, including the positive external effect from a full output compensation by the seller. This signifies that the potential buyer would have to pay a lower price than the external damage he caused to the seller because this damage always occurs to both airlines, irrespective of whether trading takes place or not. His trading potential is therefore symmetric and reads

$$TP_j^{ASB} = {}^{\partial\hat{\Pi}_j^{ST}}/_{\partial n_p^j} - {}^{\partial\hat{\Pi}_j^{ST}}/_{\partial n_p^i}.$$

For the buyer, both potential profit changes are beneficial, so that his trading potential is also positive.

Again, secondary trading again takes place if $TP_j^{ASB} > TP_i^{ASB}$. Although the trading potentials cannot be explicitly evaluated based on the generic model's functions, it becomes utterly clear that a quota trade will not occur: As the above computation shows, the trading potentials of the two airlines only consist of the turnover effect of an output expansion following a trade. This turnover effect is hence the net of all external effects from congestion and market power. Moreover, the business airline's peak-period flight fare exceeds the leisure airline's flight fare by the amount of the network value. Therefore, the leisure airline will never be able to compensate the business airline for its forgone turnover after a trade because it will always earn less per unit of output than its competitor. As in the asymmetric case without strategic behavior a quota trade will thus never take place.

As compared to the asymmetric case without strategic behavior, the application of the hand-back and reallocation rules thus has two effects: On the one hand, it increases overall output regardless of whether there is a trade or not because it introduces slot babysitting. As a consequence, the external effects need not be accounted for in the trading potentials. Other things remaining equal, this increases the chances for a trade. On the other hand, the slot babysitting implies that the buyer needs to compensate the seller for a full unit of forgone output. Due to the exogenous airline asymmetry, however, the leisure airline cannot afford this compensation. It will hence never be able to buy access to the peak-period market. As a result, the trading rules make trading impossible.

Nevertheless, the business airline has to utilize all of its slots in order to avoid a hand-back with reallocation. As a consequence, the business airline cannot choose its monopoly output, so that both overall output and the network size are higher than in the case without strategic behavior. This causes an increasing output effect on congestion and on flight fares. More precisely, the full utilization of all slots replicates the socially optimal network size if the optimal number of quotas is correctly determined. As a result, the slot babysitting induced by the trading rules yields a first-best output allocation (also see the welfare analysis in Section 12.5). However, this result may be deemed controversial from a distributional point of view. This concern is discussed in Section 14.2.4.

12.4.3 Symmetric Trade without Strategic Behavior

Let us now consider a symmetric trading case without strategic behavior. This means that every slot traded will reduce seller i's output and increase buyer j's output by the same amount. Denoting again the buyer as airline j, his output change of a slot trade would again be $\mathrm{d}n_p^j = \triangle q$. This is the same condition as in the asymmetric case above. As opposed to the latter, the seller i would now also have a net output effect amounting to $\mathrm{d}n_p^i = -\triangle q$. This also means that there is no endogenous output adjustment of the seller as a secondary effect. As a result, total output would not change, as $\mathrm{d}N_p = \mathrm{d}n_p^j + \mathrm{d}n_p^i = 0$. In addition, because both airlines are constrained by assumption, they both exhibit a positive profit change following an own output expansion. The effect of the arbitrary constraint on outputs and profits is therefore denoted with an overbar superscript. These characteristic trading properties are shown in Table 12.3.

SYM	**Seller** i	**Buyer** j	**Remarks**
Initial Allocation	$\bar{n}_p^i = \bar{q}_i < n_p^i(\bar{n}_p^j)$	$\bar{n}_p^j = \bar{q}_j < n_p^j(\bar{n}_p^i)$	both constrained
Trading Rules:	Hand-Back with Reallocation		
Trading Options: Trading Properties:	sell or use $\mathrm{d}n_p^i = -\triangle q$	buy and use $\mathrm{d}n_p^j = \triangle q$	no Strategic Behavior
Overall Output:	$\mathrm{d}N_p = \mathrm{d}n_p^i + \mathrm{d}n_p^j = 0$		
Profit Changes:	$\frac{\partial \bar{\Pi}_i^{ST}}{\partial n_p^i} > 0$	$\frac{\partial \bar{\Pi}_j^{ST}}{\partial n_p^j} > 0$	due $\bar{q}_j, \bar{q}_i$
Cross-Elasticities:	$\partial \bar{n}^i / \partial n_p^j = 0$	$\partial \bar{n}^j / \partial n_p^i = 0$	due $\bar{q}_j, \bar{q}_i$

Tab. 12.3: Symmetric Trading Properties (without Strategic Behavior)

Because both airlines are constrained, their trading potentials are symmetric. Again using the above trading properties in total derivative (29) yields the trading potential for seller i as

$$\mathrm{TP}_i^{SYM} = \frac{\partial \bar{\Pi}_i^{ST}}{\partial n_p^i} - \frac{\partial \bar{\Pi}_i^{ST}}{\partial n_p^j}. \tag{37}$$

With the sign change, the first term denotes the direct effect of the output expansion after the quota purchase. This direct effect again includes the increase in overall congestion and the decrease in flight fares. However, because these secondary effects also do not take place in a symmetric trade, the second term compensates for these effects. In sum, the trading potential thus accounts for the net effect of the output change on turnover. Correspondingly, simply interchanging index $i \leftrightarrow j$ yields the buyer's trading potential as

$$\mathrm{TP}_j^{SYM} = \frac{\partial \bar{\Pi}_j^{ST}}{\partial n_p^j} - \frac{\partial \bar{\Pi}_j^{ST}}{\partial n_p^i}. \tag{38}$$

The trading potentials thus correspond to the asymmetric case with strategic behavior, except for the fact that both airlines' pre-trading profits are limited by the arbitrary constraint. Therefore, the leisure airline again cannot afford a quota purchase. However, as opposed to the asymmetric case, the leisure airline now has a positive peak-period output arising from the initial grandfathering allocation, and thus may offer some quotas for sale. Consequently, the business airline becomes a real potential buyer, so that in the symmetric case, a quota trade from the leisure to the business airline may effectively take place.

Whether the business airline would actually purchase some slots or even buy the leisure airline out of the market is not formally revealed in the generic model, as the above trading potentials only yield the marginal trading potential after the initial quota allocation but do not resolve the explicit trading volumes and prices (see Section 12.3.7). However, the model's rationales indicate the following: The trading rules imply that overall output and, thus, both congestion and market power would always remain constant, regardless of whether quota trading actually occurs or not. Moreover, from the asymmetric case it is known that the business airline yields a higher gross turnover from any additional unit of output than the leisure airline. As a consequence, the business airline yields maximum profits from completely buying the leisure airline out of the market and avoiding a hand-back by utilization of all purchased quotas. In the symmetric case without strategic behavior, hence, the business airline will preempt the entire peak-period market, utilizing all of its slots because the hoarding of unused quotas is suppressed by the trading rules. As a result, congestion remains unchanged while the network size increases; the welfare analysis will thus show that quota trading in this case has a net positive effect on allocation efficiency.

Note that this result is similar to Verhoef's (2010) investigation of symmetric quota trading, where the more cost-efficient airline buys the less efficient airline out of the market. In absence of strategic behavior, this increases efficiency because overall output is maintained. However, when strategic slot hoarding occurs, the efficient airline reverts to its monopoly output and leaves the remainder of the slots unused (cf. idem, p.326). That setting corresponds to a symmetric trading case with strategic behavior, and thus is discussed next.

12.4.4 Symmetric Trade with Strategic Behavior: Post-trading Slot Hoarding

The case of a symmetric trade that allows for strategic behavior of the buyer arises after the grandfathering allocation of an arbitrary constraint if no particular trading rules are implemented. The buyer is hence not obliged to use his purchased slots and thus is free to hoard them unused after the trade. As an important prerequisite, however, let us assume that the seller has no information on whether the buyer will later use the slot or not but will simply presume that the slot will be used after the trade. This seems reasonable as a trade and the subsequent use would also in reality occur sequentially rather than in parallel.[81] The characteristic properties are reflected in Table 12.4.

SSB	Seller i	Buyer j	Remarks
Trading Rules	None, or: Hand-Back *without* Reallocation		
Trading Options		hoarding unused	Strategic Behavior
Trading Properties	$\mathrm{d}n_p^i = -\triangle q$	$\mathrm{d}n_p^j = \kappa \cdot \triangle q$	$\kappa \in [0, 1]$
Overall Output	$\mathrm{d}N_p = (\kappa - 1) \cdot \triangle q \leq 0$		

Tab. 12.4: Symmetric Trading Properties (with Strategic Behavior)

Under the above assumption, the characteristic properties for the seller i remain the same as under the symmetric trade without strategic behavior. Consequently, the trading potential also remains the same and thus exactly corresponds to condition (37).

For buyer j, however, the case is different. Because he can leave his purchased slots unused, the quantity of traded slots and the quantity of used slots must now be distinguished: The buyer knows that seller i will have to reduce his output by $\mathrm{d}n_p^i = -\triangle q_j$ after a trade based on his number of quotas. However, his own output expansion will amount to $\mathrm{d}n_p^j = \kappa \cdot \triangle q_j$ only, where parameter $\kappa \in [0, 1]$ denotes the degree of strategic behavior on a continuous scale: If the buyer uses all of his slots then he will set $\kappa = 1$ and the full output effect of the trade will apply. If none of the slots are used, $\kappa = 0$ yields that the buyer's output remains constant. Any value $0 < \kappa < 1$ denotes an intermediate degree of strategic behavior, where a part of the purchased slots is utilized and the other part is hoarded. In other words, κ accounts for the second-order effects on overall output depending on how many purchased quotas the buyer will actually put into use after a trade.

[81] The consideration of this rationale would require the introduction of an expectation operator in the computation of the seller's trading potential, or an assumption about a mutual agreement between the airlines. However, both complications do not seem justified within the scope of this analysis.

Again, substituting these characteristic properties into (29) yields the buyer j's trading potential as

$$\mathrm{TP}_j^{SSB} \equiv \frac{\mathrm{d}\bar{\Pi}_j^{ST}}{\Delta q} = \kappa \cdot \frac{\partial \bar{\Pi}_j^{ST}}{\partial n_p^j} - \frac{\partial \bar{\Pi}_j^{ST}}{\partial n_p^i}. \tag{39}$$

The first term denotes the gross effect of buyer j's output expansion, which now involves parameter $\kappa \in [0,1]$. The second term again makes sure that the full compensation of the trading volume on overall output and thus on congestion is accounted for in terms of the seller's output reduction.

As described above, the net effect of strategic behavior on profits evidently depends on the choice of κ: If all quotas are used after the trade then there is no overall output effect and the trading potential corresponds to the symmetric case without strategic behavior from (38). If some quotas are left unused, overall output is reduced. The welfare effect of this decrease depends on the magnitude of overall output: It imposes a positive effect on the buyer's profits if output is already high because the higher flight fares and the lower congestion will overcompensate the forgone turnover. If output is inefficiently low, however, an output reduction has a negative effect because the lower turnover cannot be compensated by lower congestion and higher flight fares.

The occurrence of strategic behavior in trading potential (39) can consequently be described as follows: When buyer j's constrained output is well below his monopoly output, his output elasticity of the profit $\frac{\partial \bar{\Pi}_j^{ST}}{\partial n_p^j}$ will be positive and high. This means that any additional output would significantly increase profits, so that $\kappa = 1$ would be chosen. When his output approaches the monopoly output from below, $\frac{\partial \bar{\Pi}_j^{ST}}{\partial n_p^j}$ decreases toward zero. Depending on its actual value, it might thus be beneficial to reduce $\kappa < 1$ as too high of an output change might change the sign of $\frac{\partial \bar{\Pi}_j^{ST}}{\partial n_p^j}$. If the monopoly output is reached, then $\frac{\partial \bar{\Pi}_j^{ST}}{\partial n_p^j} \leq 0$. In this case, $\kappa = 0$ would be the most beneficial and hence would be chosen. Strategic behavior would occur to its full extent, so that all purchased slots would be hoarded. Ultimately, this signifies that the trading potential of the buyer is higher when $\kappa < 1$ is an option as compared to the case where strategic behavior is not allowed. Any trading potential without the opportunity for post-trading slot hoarding hence also supports a trade with strategic behavior - but not vice versa (also cf. Section 12.2.3).[82]

As the business airline yields a higher turnover with every unit of output than the leisure airline, trading will hence take place. In this respect, note that the business airline again purchases the entire amount of quotas available. The reasoning is the same as in the symmetric

[82] This argument is also brought up by Verhoef (2010, p.327). At the downside, it brings up the problem that the trading potential is underestimated when strategic behavior is not accounted for. This means that a trade might still occur under a weaker use assumption, which is inhibited by the full-use assumption.

case without strategic behavior, where every slot that would not be purchased were utilized by the leisure airline and hence caused a higher congestion and lower flight fares. As opposed to the latter case, however, strategic behavior may subsequently apply. Its occurrence depends on the total amount of slots relative to the business airline's monopoly output and follows the above reasoning. The business airline therefore is better off either by providing all peak-period output on its own. If the total number of available quota is higher than the business airline's monopoly output, this implies that the excess number of purchased slots are strategically hoarded.

As the welfare analysis will show, this finding also corresponds to Verhoef's (2010) result accounting for strategic slot hoarding, where the hoarding diminishes overall output and renders the welfare effect of quota trading ambiguous.

12.5 Welfare Analysis

In the secondary trading analysis, the welfare effect of a trade again depends on two well-known factors: First, on the effect of the quota trade on overall peak-period output, which affects both the size of the market power distortion as well as the congestion externality; and second, on the output change of the networking airline, which directly affects network size and thus the passengers' network density benefits. These effects have already been established in the investigation of airport quota allocation in Section 11. Combining the two separate effects from overall output and from network density, the net welfare effect of secondary trading are subsequently evaluated on a case-to-case basis in Section 12.6.

12.5.1 Output Effects

In order to assess the welfare impact of secondary trading systematically, the output effects of the four above trading situations are therefore briefly reviewed in Table 12.5. The table shows the trading rules and their effect on individual and overall outputs.

As the above investigation has shown, the fundamental difference in the overall output changes depends on whether the trade is symmetric or asymmetric and whether strategic behavior occurs or not. In the asymmetric case without strategic behavior, the traded quotas were formerly unused, which means that a trade increases overall output. In the asymmetric case with strategic behavior, the trading rules require the initially unused quotas being utilized either by the slot holder in terms of babysitting or by the competitor after the purchase

Trade: $\triangle q$	Buyer: $\mathrm{d}n_p^j$	Seller: $\mathrm{d}n_p^i$	Peak Period: $\mathrm{d}N_p$	Network Size: $\mathrm{d}n_p^B$
ASY	$+\triangle q$	$\partial n_p^i/\partial n_p^j \cdot \triangle q < 0$	$0 < \mathrm{d}N_p < \triangle q$	$\mathrm{d}n_p^i < 0$
ASB	$+\triangle q$	$-\triangle q$	0	$-\triangle q$
Trading Rules	0	$+\triangle q$	$+\triangle q$	$+\triangle q$
SYM	$+\triangle q$	$-\triangle q$	0	if B Buyer: $\triangle q$ if B Seller: $-\triangle q$
SSB	$+\kappa \cdot \triangle q$	$-\triangle q$	$(\kappa - 1) \cdot \triangle q \leq 0$ with $\kappa \in [0, 1]$	if B Buyer: $\kappa \cdot \triangle q$ if B Seller: $-\triangle q$

Tab. 12.5: Secondary Trading - Output Effects

or a hand-back with reallocation. Observe that in the latter case, however, it is the trading rules themselves that induce the strategic airline behavior and thus cause the unused quotas to be utilized. This means that the quota utilization is enforced irrespective of whether a trade occurs or not.

In the symmetric case without strategic behavior, both airlines were already constrained before the trade. Therefore, a trade would change the individual outputs proportionally, so that overall output would remain the same. Recall that in this case, the trading rules avoid the occurrence of strategic behavior. The symmetric trading case with strategic behavior hence applies when the trading rules are not in place. It has a variable outcome that is opposite to the corresponding asymmetric case because strategic behavior only occurs after a trade: As the slot buyer may either use the purchased slot or hold it unused, overall output either remains the same or decreases after trading. The buyer can induce a costly reduction of his competitor's output and later decide whether he is better off with the overall lower output or whether he would like to partly or fully compensate that output by increasing his own market share.

With regard to overall output, the welfare impact is the following: If output increases with a trade, the market power distortion decreases and the congestion externality increases. Generally, the effect of this output change would be ambiguous. But because it is known that after quota regulation both the network and the overall peak-period output are inefficiently low, an output increase generally increases allocation efficiency. In other words, welfare strictly increases with increasing peak-period output in a secondary trading scenario because the pre-trading overall output is generally too low. As a consequence, the market power distortion decreases and the congestion externality increases toward its optimum. The opposite is the case when output is reduced after a trade, in which the output effect of secondary trading reduces welfare.

The output effects from Table 12.5 reveal an interesting pattern: In the asymmetric cases, overall output always increases. Without strategic behavior, the trade itself increases total output, whereas the trading rules increase output when strategic behavior occurs. A symmetric trade keeps total output constant when strategic behavior is ruled out and the slots purchased must be utilized. In this case, strategic behavior will either keep overall output constant or diminish it, depending on whether the slots purchased are fully or partly utilized or not after the trade. Briefly summarized, in the asymmetric cases, overall output tends to increase, so that the inefficiencies from market power and congestion decrease. In the symmetric cases, total output tends to remain constant or decrease, which means that trading either has no impact on the dual distortion or even imposes an adverse effect.

12.5.2 Network Density Effects

The impact of quota trading on network density and thus on welfare is determined by the monotonicity of the network density benefits and by whether the traded quotas have been previously unused or used. If the business airline is a quota buyer and utilizes the purchased quotas, welfare is increased. If the business airline is a slot seller and has utilized the quotas before trading, welfare is decreased. The trading of unused quotas, or the non-utilization of quotas that have been traded has no welfare impact. As a consequence, the following effects of secondary trading on network density can be put down:

In the asymmetric case, the business airline is always the seller. When strategic behavior is ruled out, only unused slots are sold. However, observe that the business airline endogenously reduces its output after a trade subsequent to the increased congestion. The network size thus slightly decreases so that there is a small negative effect on welfare based on the reduced network density benefits. When the airline previously babysat the traded slots, the network size diminishes by the full amount of the trade so that welfare is reduced. However, the imposition of the trading rules previously caused the unused slots of the business airline to be utilized. Accounting for this effect implies that there is no net impact of a trade on the network size. As a result, the trading itself negatively affects welfare, but the initial allocation previously caused a welfare improvement.

Finally, the symmetric case is characterized as follows: In the case without strategic behavior, all quotas are always in use. If the business airline is the buyer, welfare is increased because the network size rises. If the leisure airline becomes the buyer, welfare is decreased. If, in addition, strategic behavior takes place, welfare is adversely affected if the leisure airline is the buyer. If the business airline is the buyer, the welfare effect only remains beneficial if at least some of the slots purchased are utilized; otherwise, the network size remains constant but overall output declines.

12.6 Results: Allocation Efficiency

12.6.1 Asymmetric Case without Strategic Behavior (ASY)

In the asymmetric case without strategic behavior, overall output would increase and the network size would decrease with a trade. The overall net welfare effect of a secondary slot trade would hence be generally ambiguous. However, the overall output expansion almost amounts to the full trading volume and is a primary effect. In contrast, the network size reduction is based on an endogenous output adjustment and only represents a secondary effect. It is therefore likely that a secondary trade would increase allocation efficiency based on its overall output effect.

This amelioration would only be of second-best manner because it would be based on an output effect but not on a network effect. The likelihood of a positive effect therefore becomes higher, the less the network density benefits matter relative to congestion and market power. However, even if the welfare effect were beneficial, the above analysis has revealed that the trading potentials would only fulfill the necessary but not the sufficient condition for a trade. This means that secondary trading will never take place and cannot be applied to increase allocation efficiency. As a consequence, the market structure remains a monopoly, with the business airline as the only firm.

If trading were beneficial in the above sense, the positive welfare effect could theoretically be achieved by an administrative reallocation of the unused quotas. From a game-theoretic perspective, however, this would correspond to the implementation of the hand-back and reallocation rules. Therefore, any attempt to reallocate the unused quotas returns the setting to an asymmetric case with strategic behavior. In the absence of the trading rules, such a reallocation could not be implemented in practice.

12.6.2 Asymmetric Case with Pre-trading Slot Babysitting (ASB)

In the asymmetric case with strategic behavior, the imposition of the trading rules induces a positive welfare effect both through an overall output expansion and through an increase in the network density benefits: The business airline has to utilize all of its quotas and therefore replicates the social optimum. The analysis has shown that this strategy is dominant over any secondary trading with the leisure airline. This means that the trading rules themselves induce first-best allocation efficiency before secondary trading may actually occur. If a subsequent quota trade were to take place, the overall output would remain at the socially efficient level. However, the network size would again be reduced. This signifies that an actual quota trade would have a negative impact on welfare. Yet, based on the trading potentials, it can be claimed that trading does not occur at all. As a result, in the asymmetric case with strategic behavior, babysitting always occurs and induces a first-best quota allocation.

In addition, note again that it is the babysitting that generates both the socially efficient overall output and the optimal network size. In other words, this effect does not arise from secondary trading itself but from the imposition of the trading rules. A subsequent trade would again diminish this positive welfare effect. However, the trading potentials reveal that trading never takes place. Consequently, from a welfare perspective, it is optimal and sufficient to impose the quota trading rules. The subsequent imposition of a secondary trading scheme itself is neither necessary nor favorable.

12.6.3 Symmetric Case without Strategic Behavior (SYM)

In a symmetric case without strategic behavior, output remains constant after a secondary quota trade. Therefore, if the business airline were the slot buyer, the network size would increase and the overall welfare result of a trade would be positive. By contrast, if the business airline would be the seller of the quotas, the network size would diminish and welfare would decrease. As the previous analysis showed, however, either airline would only sell some of its utilized slots if the trade covered for the loss from reduced turnover. Also, it revealed that the business airline yields a higher turnover with each additional unit of output than the leisure airline, based on the network premium from the exogenous airline asymmetry, so that the business airline becomes the quota buyer and the leisure airline becomes the quota seller and all quotas remain utilized. As a result, secondary trading will actually occur in the symmetric case and will improve welfare in a second-best manner.

Because overall output in the symmetric case without strategic behavior is determined by the total number of quotas, the network size will increase above the monopoly output as long as the overall number of quotas is sufficient; still, as an artifact of the arbitrary constraint, the network will always remain undersized. Therefore, the welfare improvement can only be second-best. However, the remaining inefficiency ultimately arises by definition and cannot be resolved endogenously in this model. In conclusion, in the case where the imposed quota constraint is of the symmetric type, secondary trading is thus generally welfare improving to its full potential extent.

This case contrasts with the asymmetric trade from above: The business airline can now effectively become an actual buyer because the leisure airline at all disposes of a positive number of quotas. Moreover, both airlines only engage in trading if they are overcompensated for their output changes. Consequently, it is more profitable for the leisure airline to sell its market access to the peak period than to actually provide output. As already pointed out, this result corresponds to Verhoef (2010), where the less efficient airline is bought out of the market for the same reason.

Lastly, observe that although the welfare effect from trading is positive, it introduces more unequal market shares. Both the welfare result and the distributional effect may therefore be perceived as controversial, although they are both unambiguously favorable in formal terms.

12.6.4 Symmetric Case with Post-trade Slot Hoarding (SSB)

Generally, the welfare effects of a trade are the same as in the above symmetric case without strategic behavior. However, in a symmetric case with strategic behavior, the buyer can decide whether or not to use his slots purchased. This type of strategic behavior negatively affects welfare by diminishing overall output, which would increase the market power distortion and depress congestion further below its optimum value. The overall welfare effect depends on the degree of strategic behavior: If at least part of the quotas are utilized by the business airline, the network size is still increased. The increased network density may or may not overcompensate the deterioration arising from the increased dual distortion. By contrast, in the extreme case where all excess unused slots are hoarded after the trade, the welfare compensation from the network effect would vanish, so that only the output effect would occur. As a consequence, the net welfare effect of the slot trade from the leisure to the business airline became adverse.

Overall, the occurrence of strategic behavior in any case diminishes the welfare results of secondary trading in a symmetric setting and may even overturn its positive welfare effect. Moreover, the opportunity for strategic behavior increases the trading potential of the buyer, so that the chances of trading to occur increase. From a welfare perspective, the post-trading strategic behavior must hence be suppressed by implementation of the hand-back and reallocation rules. As a consequence, the welfare result reverts to the symmetric case without strategic behavior and becomes positive in a second-best manner.

13 Congestion Pricing

Congestion pricing represents a tax to internalize the non-internalized portion of congestion (Button, 1993, p.94). The correct congestion toll hence equals the marginal congestion costs which are not accounted for in each airline's profit maximization rationale. In this respect it is important to note that the airlines consider the toll as exogenous (Brueckner, 2002a, p.1367); this signifies that they can only change their tax burden by means of their output choice, while the amount of the tax per flight is perceived as given.

13.1 Determination

Because congestion only occurs in the peak period, the tax depends on the airlines' peak-period outputs only. The external marginal costs of flight delays can directly be extracted from the first-order conditions of the unconstrained equilibrium. Based on Brueckner's (2002a, p.1368) generic example for an asymmetric oligopoly, the tax per flight for airline i with competitor j is denoted by r_i and amounts to

$$r_i = n_p^j \cdot TG^{*'} \tag{40}$$

where $i \neq j \in \{B, L\}$.[83] The relationship $n_p^j = N_p - n_p^i$ in the duopoly shows that the tax is inversely proportionate to each airline's market share. This replicates Brueckner's (2002b, p.22) finding, where the tax is symmetric and only depends on the number of firms. Note that the marginal congestion costs $TG^{*'}$ are considered as exogenous to the airline, as prescribed previously. Therefore, the tax is denoted as a variable rather than as a function. As the tax accrues per flight, the total burden for each airline amounts to $R_i = n_p^i \cdot r_i$. The marginal impact of the tax on each airline thus amounts to $\frac{\partial R_i}{\partial n_p^i} = r_i$, which directly equates to (40) and formally shows the exogeneity of tax r_i.

Consequently, the introduction of this tax as an exogenous cost into the airline's profit functions directly allows us to change the peak-period equilibrium conditions (9) and (10) to

$$\begin{aligned} B^* - TG^* &\geq n_p^L \cdot B^{*'} + \left(n_p^L + n_p^B\right) \cdot TG^{*'}, &(41)\\ B^* - TG^* + d(\theta^D) &\geq n_p^B \cdot \left[B^{*'} + d'(\theta^D)\right] + \left(n_p^L + n_p^B\right) \cdot TG^{*'}. &(42)\end{aligned}$$

[83] This corresponds to Brueckner's (2002a, p.1368) result for $s_1 = s_2 = 1$ and $g_1 = g_2 = g$ where the latter reflect this model's duopoly. In addition, the generic factor $1 - \frac{1}{k}$, which denotes residual supply as a function of the number of firms and hence the external part of congestion, with $N_p = n_p^i + n_p^j$ in the asymmetric duopoly becomes $1 - \frac{n_p^i}{N_p} = \frac{n_p^j}{N_p}$ for airline i.

As a result, the congestion costs of each airline's competitor now also enter the equilibrium conditions, which confirms that both airlines fully internalize the whole impact of their output choice on airport delay and, thus, on delay costs.

13.2 Welfare Analysis

Two implications can be inferred from the imposition of the congestion tax: First, in analogy to the equilibrium, the two above conditions are independent.[84] Moreover, comparing them to the unconstrained first-order conditions (10) and (9) shows that both airlines still enjoy a positive output under the congestion pricing regime. As a result, the tax does not replicate the socially optimal market structure, because both firms still serve the market in equilibrium. Consequently, the network size remains inefficiently low, which causes a deviation farther away from first-best efficiency.

Secondly, the inefficiency of the network size is also indicated by comparison of condition (42) to the socially optimal network size (19): The congestion and market power terms on the left-hand side of (42) indicate that the entire peak-period output is governing and hence depresses the business airline's output choice farther below the social optimum. Moreover, the right-hand side reflects the network value rather than the entity of the network density benefits across all passengers. As the former is concave while the latter monotonously increases, this still indicates an inefficiently low network (see also Section 9.3).

In consequence, two corrective measures could be undertaken in order to establish efficiency. As Verhoef (2010) proposes, on the one hand, the congestion tax may be corrected for the market power distortion. In this case, the tax would have to amount to

$$r_i^{MP} \equiv n_p^j \cdot \left[B^{*'} + TG^{*'} \right].$$

In analogy to the above modification of the airlines' first-order conditions, the inclusion of this tax would consequently make condition (41) equal to social optimum condition (18). As a result, overall output would become efficient.

However, the business airline's network size would still be determined by the concave network value, as shown in (42), and would hence still not replicate the optimal network size as dictated by (19). To overcome this shortcoming, the congestion tax would hence also need to compensate the business airline for its forgone profits from expanding its output to the

[84] This contrasts to the social optimum peak-period conditions, which need to be checked against leap-frogging (see Section 7.5).

socially optimal network size. Naturally, this could only be achieved by turning the business airline's congestion tax into a net subsidy. By contrast, the leisure airline would have to pay both for the marginal congestion from its peak-period output as well as for the loss in density benefits arising from the business airline's corresponding endogenous output decrease. However, by definition, any peak-period turnover from the leisure airline can never be sufficient to compensate for the monotonously decreasing network benefits for all passengers because it does not yield a density premium.

In analogy to Verhoef (2010, p.323)'s result with market power only, such a tax would hence effectively expel the leisure airline from the peak period. Consequently, first-best allocation efficiency would be reached. Yet, this solution could no longer be referred to as a congestion pricing scheme, because the business airline's tax would be negative while the leisure airline's net tax burden would be zero due to its complete absence from the peak-period. Rather, such a scheme would constitute a service obligation for the network airline, with the negative tax representing the corresponding monetary compensation.

13.3 Results: Allocation Efficiency

The standard congestion tax (40) removes the congestion externality and hence reduces the dual distortion to comprise the market power distortion only. The comparison of above condition (41) to social optimum condition (18) shows that such a tax depresses overall peak-period output below its efficient level, also decreasing peak-business output and reducing the network density benefits. Consequently, a first-best solution is not achievable with the tax.

However, a second-best welfare improvement may still arise if two prerequisites are met: The resulting inefficiency from market power needs to become smaller than the original dual distortion, and, in addition, the resulting welfare improvement needs to compensate the corresponding loss in network density benefits. This may only occur when the congestion externality before regulation is more important than the market power distortion, so that peak-period output is considerably higher than the social optimum. In the opposite case, the removal of the congestion externality yields an adverse welfare effect. Based on each airline's congestion internalization of its own market share, the latter case is more likely to occur in a duopoly. As a result, the adverse welfare effect of a congestion tax under market power becomes even worse when network effects are present.

This model's result basically corresponds to the result from Brueckner (2002a) but includes an additional adversity based on the network density benefits. Moreover, as in Brueckner

(2002b, p.23), the tax is also reciprocal to each airline's market share. The business airline hence pays a lower price for airport capacity than the leisure airline. As this property may invoke the same distributional concerns as the secondary trading case, it is further discussed in Section 14.2.4.

14 Discussion

This section briefly reviews and discusses the results of the three allocation schemes investigated in this study. The focus lies on the allocation efficiency results and on the corresponding implications for the application of the schemes in practice.

14.1 Quotas (Airport Slots)

14.1.1 Allocation Efficiency of Individual Quotas

As a starting point, recall that from a social welfare perspective, an increase in the business airline's peak-period flights monotonously increases both the network size and the corresponding benefits. Leaving aside congestion costs, this increases utility in a monotonous manner. This model property was revealed as the monotonicity of the network density benefits in Section 8.1. By contrast, an increase in the peak-period flights of the leisure airline does not affect network density benefits but also increases congestion. As a consequence, a peak-period reduction of the leisure airline's flights only reduces congestion, while constraining the business airline also monotonously decreases the network density benefits for the passengers. As compared to a corresponding decrease in the leisure airline's output, a reduction of the business airline's peak-period output yields a higher counteracting effect on welfare for the same reduction of delay costs. From a social point of view, it is therefore always more beneficial to constrain the leisure airline instead of the business airline.

Subsequently, recall that optimum quota rule (26) states that the leisure airline may allocate some peak-period flight volume if the optimal network size is reached and there is still room for more congestion. However, any additional peak-period output of the leisure airline only increases congestion, whereas every additional unit of peak-business output also increases indirect utility from network density. From a welfare perspective, any airport capacity left in terms of congestion needs to be allocated to the business airline. In other words, based on the monotonicity of the network density benefits, there is no reason to allow for a positive peak-period leisure output. Consequently, the optimum quota rule (26) must dictate $\hat{q}_L = 0$, so that the leisure airline is completely expelled from the peak period.

Next, as explored in Section 8.3, the endogenous output adjustments show that after an asymmetric quota imposition, an unconstrained airline adjusts its profit-maximizing output according to its reaction function. As the leisure airline is completely banned from the peak period, the business airline will expand output and thus its network size. This output compensation is not exhaustive and does not fully replace the leisure airline's withdrawal. In absence of the leisure airline, the peak period becomes a monopoly market. As a consequence, the business airline fully internalizes congestion and market power becomes the only distortion. The peak-period output of the business airline therefore falls short of the socially optimal value, so that the network size remains inefficiently small after the imposition of the individual quotas.

In summary, the quota rule must always dictate that the leisure airline abstain from the peak period. Consequently, a monopoly situation arises where peak-business output falls short of its social goal. The network size never reaches the optimum and remains inefficiently small. As a result, the quota allocation cannot replicate the social optimum. Therefore, first-best allocation efficiency cannot be reached by the individual quota solution even if the optimal number of slots is correctly determined.

This shortcoming is caused by the asymmetric imposition of the quotas. The asymmetry itself is based on the monotonicity of the network density benefits, which reflects the particular properties innate to this model: the exogenous airline asymmetry in conjunction with endogenous demand, and the subsequent endogenous output adjustment of the unconstrained airline. Ultimately, the above result shows the impact of the network density benefits on the efficiency of airport capacity allocation. This illustrates the innovation of this study and its unique contribution to the discussion in the recent literature.

Despite the above general inefficiency, the individual quotas may still yield a second-best welfare improvement. More precisely, the overall welfare result can be either positive or negative. As the results from the analysis have shown, the final outcome depends on the initial size of the dual distortion and on the relative size of the network effect. On the one hand, the individual quotas always increase the business airline's output and enhance the network size. This causes a net positive welfare effect. On the other hand, overall output decreases because the business airline's endogenous output adjustment is not exhaustive.

The welfare impact of the output effect may take either direction: If the congestion externality was large before the quota allocation, the overall output contraction decreases the dual distortion. As a consequence, the quotas' overall effect on welfare is positive, and allocation efficiency improves. Note that the resulting overall output may remain above or below its socially optimal value. The crucial argument for a net positive welfare effect is that the absolute size of the dual distortion is diminished. In the opposite case, the dual distortion is increased after the quota imposition. This happens if output is depressed farther away from the social optimum than before regulation, so that the output effect of the quotas has a negative welfare impact. Yet, a beneficial outcome can still be reached if the higher network density benefits overcompensate the higher deadweight loss. Otherwise, the quotas induce an adverse welfare effect.

As a result, the quotas may improve allocation efficiency in a second-best manner if their effect on the dual distortion is positive, or negative but not excessive in comparison to the higher network density benefits. By contrast, if the output effect of the dual distortion is adverse while the network density benefits are of little relative importance, the individual quotas induce a welfare caveat.

14.1.2 Deviation from the Optimal Quota Rule: a Second-best Trade-Off

In the case where the initial quota allocation yields an adverse welfare result, a trade-off may be attempted to revert the efficiency loss into a net welfare gain. This trade-off requires deviation from the optimal quota rule by allocating a certain proportion of the business airline's unused quotas to the leisure airline. As a consequence, overall peak-period output increases while the network size diminishes, based on the corresponding endogenous output adjustment by the business airline .

From a welfare perspective, the output expansion yields the following effects: On the one hand, congestion increases towards its optimal level. At the same time, the market power distortion is reduced. As a result, the dual distortion decreases, yielding a positive net welfare contribution. On the other hand, the reduction of the network size monotonously decreases efficiency. Generally, the overall welfare effect of this trade-off becomes ambiguous. However, recall that an adverse welfare effect from the initial quota allocation arises if the network effect is inferior to the output effect. Consequently, this trade-off increases allocation efficiency whenever the individual quotas yield a welfare loss, which occurs if the network density benefits are relatively unimportant in comparison to the dual distortion. Whether the net welfare benefit of this re-allocation can overcompensate the initial welfare deterioration thus depends on the relative sizes of the effects and on the respective output quantities and cannot be determined in a generalized manner.

As a result, the deviation from the optimal quota rule $\hat{q}_L = 0$ yields a relative welfare gain whenever the reduction of the dual distortion is important relative to the corresponding loss of network density benefits. In this case, the initial quota allocation previously yielded an adverse welfare effect, and the relative welfare improvement may or may not be able to compensate this initial welfare loss.

14.1.3 Allocation Efficiency of the Grandfathering Allocation

Although the grandfathering allocation is not endogenously replicable within the model, the welfare analysis shows that a symmetric quota allocation is concerned with the same inefficiency found in Brueckner's (2002a) congestion pricing investigation: A negative welfare effect from an increment of the dual distortion based on an excessive output reduction. Moreover, the investigation introduces the impact of asymmetric network effects in a symmetric quota scheme. As a result, the network benefits imply that a grandfathering allocation may even become ambiguous when it decreases the dual distortion but excessively reduces the network size.

As explained above, however, these welfare results are not representative because the symmetric quotas cannot be justified in the asymmetric model on grounds of efficiency. Nevertheless, the grandfathering allocation reflects the current airport capacity allocation from practice more suitably than the asymmetric case. Therefore, it constitutes a starting point for the investigation of a secondary trading scheme that aims at replicating the current initial allocation from practice. This investigation is provided in Section 14.2.2.

14.1.4 Implications for Practice

The welfare analysis implies that the determination of the individual quotas requires perfect and complete information about both airlines' reaction functions. In this respect, the above investigation showed the following results: When network density benefits and market power are present, any excessive reduction of output contains a welfare caveat, while determining the constraints too generously may reduce their effect on congestion but improve allocation efficiency. In parallel to Brueckner's (2002a) investigation of congestion pricing, this insight is also valid when market power is the only distortion. It hence does not directly depend on the airline asymmetry and applies both to individual quotas and an arbitrary constraint.

The above results indicate that care must be taken in the implementation of an individual quota scheme when market power and network density effects are present. The general risk is that output is depressed too far below its optimum, so that the increased network benefits are overcompensated by the higher dual distortion. This insight arises from the fact that the network size always remains inefficiently low both due to market power and due to the concavity of the density benefits. As argued above, the subsequent overall output expansion deviating from the optimal quota rule might either increase allocation efficiency of a second-best solution, or might avoid an adverse welfare effect. As a consequence, a regulator who is in doubt about the optimum number of constraints should choose too small a number of quotas for the non-networking airline rather than strictly applying the optimal quota rule.

Concerning the individual quota allocation, one might argue that, in reality, a social planner could not independently implement just any desired resource allocation.[85] It is also not reasonably arguable that a social planner would constrain an airline when overall output is already inefficiently low. Rather, any informed coordinator would refrain from regulation in this case or would naturally re-allocate some of the unused quotas to the leisure airline. This means that the welfare caveat based on the overall output inefficiency may be seen as a rather theoretical concern. However, the caveat arising from a network size reduction remains valid. This shows that despite a positive output effect, a quota re-allocation to the leisure airline cannot be performed before it is cautiously evaluated against its benefits in terms of a lower dual distortion. Moreover, the above results still emphasize that a first-best allocation is not achievable. This may help the regulator to adjust the naive initial quota allocation according to the corresponding second-best concerns.

In contrast to the individual quotas, a long-run grandfathering allocation of an arbitrary constraint does not formally take into account the first-best criteria for allocation efficiency from an economic perspective. As judged within the asymmetric framework of this study, the administrative allocation is not efficient for two reasons: Firstly, it excessively restricts the size of the network, and secondly, it allows the leisure airline to participate in the peak period. However, this model does not provide a justification on efficiency grounds for such a grandfathering allocation that restricts both airlines. By definition, the welfare result can thus only turn out second-best, ambiguous or even adverse. This might be different in a market structure where market power is absent, so that airlines exhibit excessive outputs, such as, e.g., in a perfect competition setting with a large number of firms (see Brueckner, 2002a, p.1364). Even if the conditions yielding an efficient grandfathering allocation remain undiscovered in this study, the arbitrary constraint from above might yield a substantial positive welfare contribution by decreasing congestion.

In summary, the asymmetric allocation of individual quotas does not reflect the current grandfathering allocation scheme from practice. Moreover, the investigation of the grandfathering allocation remains unsatisfactory because it cannot be motivated by efficiency considerations within this model. Therefore, any inference from the above analysis should be undertaken with caution, respecting the fundamental differences between practice and this model's particular setting. Nevertheless, the above implications seem to be applicable in a general sense.

[85] Achim Czerny is kindly responsible for this comment, which encouraged me to investigate the grandfathering case.

The results of this investigation may thus be judged as a contribution to identify the caveats of a quota solution in practice if when markets are not perfectly competitive and some airlines create indirect benefits from network density.

14.1.5 Market Power vs. Scarce Airport Capacity

The above results yield that network size is generally underdeveloped in the asymmetric equilibrium with network density benefits. Moreover, the market in this study reflects a duopoly with two airlines only. This low number of firms suggests that a large part of airport congestion is internalized. Consequently, the congestion externality is likely to be inferior to the market power distortion in the equilibrium of the current model. As a result, in a market structure that is similar to this model's asymmetric airline duopoly with a dominant network airline, both the overall output as well as the network size may be reasonably supposed to be inefficiently low. However, the current capacity problems at large airports seem to arise from excessive flight volumes rather than from low output inefficiencies. Hence, one might argue that market power does not represent a relevant problem in practice because equilibrium outputs would generally be too high.

However, there is an important caveat of misinterpretation in this inference: First, the regulator might unintentionally choose a wrong target function in the determination of the number of quotas. Most importantly, he might minimize flight delays instead of optimizing the overall level of congestion against market power and network utility. Such a miscalculation may occur, e.g., if passengers perceive congestion as excessive as soon as it actually occurs although from the perspective of maximizing rents, it might be justified or even too low. As Forsyth and Niemeier (2008, p.81) put it, slot managers may simply choose to impose an "arbitrarily determined 'acceptable' level" of flight delays. Similarly, the slot-setting authority may also want to reduce flight volume and thus congestion based on popular but non-efficiency-driven goals such as airport noise reduction, environmental constraints or political activism. These considerations draw on Button's (2005, p.51) notion that the regulator's goals need not correspond to economic efficiency considerations, especially in view of rationales like regional economic development or a politician's own popularity. In practice, the chosen overall number of quotas may therefore simply fall short of the socially optimal output. In this case, airport demand would exceed supply although it would actually be inefficiently low.

As a consequence, the above arguments may encourage the investigation of airport capacity allocation in light of a significant market power distortion and limit any prejudicial doubts about the market power story. Moreover, the above reasoning should stress the importance of focusing on the correct determination of the overall number of quotas in practice, in particular, with reference to the welfare caveat of an output reduction when flight volumes are already inefficiently low. Especially if the information about the optimal number is incomplete or uncertain, the number of quotas chosen should therefore be too generous rather than too sparse. Lastly, this consideration may motivate future research to provide empirical evidence on the output distortion. This may be achieved by a comparison of optimal congestion to real flight delays at regulated airports. Judging the importance of the market power distortion relative to the congestion externality would prove extremely helpful in the efficiency assessment of a quota allocation scheme.

14.2 Secondary Trading

The secondary trading analysis from Section 12 investigates four distinct potential cases for secondary trading: An asymmetric initial allocation of individual quotas, and a symmetric constraint from a long-run, grandfathering allocation. Both cases are evaluated with and without strategic airline behavior.

14.2.1 Secondary Trading from an asymmetric Allocation of Individual Quotas

With individual quotas, the airline asymmetry yields an asymmetric allocation where the leisure airline is completely expelled from the peak period. Therefore, in a secondary trading case, the business airline can only become the seller and the leisure airline can only become the buyer. Due to the market power distortion, the network size is inefficiently low after the initial allocation and some quotas remain unused. A sale of these unused quotas from the business to the leisure airline would induce a positive overall output effect but decrease the network size. A positive welfare effect hence arises if the decrease in the dual distortion overcompensates the negative network effect. Because the latter is only based on an endogenous output adjustment, however, it is likely to be small. Therefore, a quota trade is likely to yield a positive welfare effect.

Yet, the analysis of the trading potentials shows that the dominant strategy of the business airline is to achieve its monopoly output during the peak-period. Consequently, the business airline will not allow the leisure airline to enter the peak-period market, so that trading will never take place in the asymmetric case. Therefore, a secondary trading scheme cannot compensate the inefficiency arising from the initial allocation of the individual quotas.

Nevertheless, two distinct allocations will arise depending on whether the trading rules are enforced or not: In the absence of the trading rules, the business airline will revert to its monopoly output and simply hold its excess number of slots unused. In this case, the inefficiency of the initial allocation remains. By contrast, the imposition of the trading rules in terms of a mandatory hand-back and re-allocation enforce the utilization of all quotas. As trading is never profitable for the business airline, the latter will therefore engage in slot babysitting. This babysitting increases overall output and thus decreases the dual distortion. In addition, the network size increases. As it is assumed that the optimal number of quotas was correctly determined, this allocation replicates the social optimum. A subsequent quota trade from the business to the leisure airline would therefore adversely affect welfare because it would again reduce the network size. However, as mentioned above, the trading potentials yield that such a trade never occurs, while the trading rules factually replicate a service obligation for the business airline by introducing slot babysitting. Although this represents strategic airline behavior, the trading rules yield first-best allocation efficiency.

The analysis of the asymmetric case hence shows that a first-best allocation can be reached. However, this allocation is not induced by secondary trading itself, but by the mere imposition of the trading rules: The trading rules force the business airline to utilize all of its slots, so that both the network size and overall output become optimal. Enabling a secondary trading opportunity thereafter is neither favorable nor necessary as trading would not be beneficial and would not occur. From a welfare perspective, it is therefore sufficient to impose the trading rules consisting of the mandatory hand-back and re-allocation of unused quotas.

This result indicates the oddness of the trading rules, which consists of the paradox that the trading rules induce strategic airline behavior by means of slot babysitting but that, at the same time, this kind of strategic behavior yields first-best allocation efficiency. Put differently, the trading rules yield the highest possible welfare improvement after the asymmetric allocation of the individual quotas by actually introducing strategic airline behavior. The above result may invoke unwarranted distributional concerns, which are discussed in Section 14.2.4.

As a limitation, note that the above result hinges on the model property that overall output after quota imposition is inefficiently low. This initial allocation, in turn, is a direct consequence of the market power assumption in conjunction with the monotonicity of the network density benefits. Consequently, in a setting where the congestion externality is important, market power absent and the network effects excessive, the effect of the above slot babysitting might be reversed. In this respect, note that the above result contrasts to Sieg (2010), where the increased utilization of the allocated quotas due to slot babysitting reduces airline profits but also decreases social welfare. However, his setting applies a monopoly airline at a monopoly airport and hence abstracts from congestion externalities. Despite the output inefficiency arising from market power, the welfare reduction is based on the dual demand structure and hence is not directly comparable to this model (see Section 4.2.2).

14.2.2 Secondary Trading from a symmetric grandfathering Allocation of an arbitrary Constraint

In the symmetric case, both airlines are constrained but have positive outputs, so that trading may occur in either direction. The welfare impact of a trade crucially depends on which airline is the seller and which is the buyer and on whether the airlines exhibit strategic behavior or not.

If strategic behavior is suppressed by the trading rules, in any quota trade the output changes are mutually offset across the two airlines. Consequently, overall output and thus the dual distortion remain constant, and secondary trading does not yield an output effect on welfare. If the business airline is the buyer, the network size increases and the welfare effect is positive. If the leisure airline is the buyer, network size diminishes and there is an adverse welfare effect.

If the trading rules do not apply, strategic airline behavior may occur. This signifies that the buyer decides whether to utilize or to hoard the purchased slots. If slot hoarding occurs, the individual outputs are not completely offset. As a consequence, overall output is reduced after a quota trade, and welfare is adversely affected. In the case where the business airline is the buyer, the initially positive welfare effect of a trade may be overturned to yield an adverse result. In the case where the leisure airline is the buyer, the adverse effect is reinforced.

Consequently, trading in the symmetric case should only be allowed if the networking airline becomes the buyer. In addition, the trading rules should always be imposed in order to suppress post-trading strategic behavior. In contrast to the asymmetric case, an actual quota trade increases welfare.

Specifically, this model's exogenous airline asymmetry yields that the business airline exhibits a higher willingness to pay for a slot trade in the symmetric case because it achieves a higher profit from an additional unit of output than the leisure airline. This higher profitability is based on the network premium. As a result, the business airline will become a buyer. Because its marginal profit is higher for all units of output and congestion initially remains constant, it will completely buy the leisure airline out of the market. Whether all quotas are subsequently utilized or not depends on whether the trading rules are imposed or not:

In they are, the full peak-period output is provided and network size increases, yielding a positive welfare effect. If the purchased quotas may be hoarded, in contrast, the business airline reverts to its monopoly output. In this case, the welfare effect depends on how many of the purchased quotas are utilized and how many are hoarded. This, in turn, depends on the overall number of quotas relative to the business airline's monopoly output. With strategic behavior, there may hence be a negative output effect which may or may not be compensated by a positive network effect. Therefore, in absence of the trading rules, the welfare effect of quota trading remains ambiguous. This yields that from a social welfare perspective, a quota trade is certainly beneficial as long as strategic behavior is suppressed by the trading rules.

Lastly, recall that a grandfathering initial allocation is not endogenously justifiable in this asymmetric oligopoly setting. The result that welfare improvements may only become second-best is thus an artifact of this model. Nevertheless, the grandfathering allocation reasonably approximates the initial allocation according to the current administrative allocation scheme from practice. Therefore, it represents a more suitable starting point for the investigation of a trading solution than the individual quota scheme.

14.2.3 Strategic Airline Behavior: Two distinct Types

As Verhoef (2010, p.326) notes, the occurrence of strategic behavior depends on the total number of access rights available in relation to the monopoly output of a single airline: If the number of quotas is lower than the monopoly output of one airline, the full utilization of all slots is the „profit maximizing strategy“ of any constrained airline. If more slots are available to one airline, it is optimal to purchase as many slots as possible in order to restrict the competitor's output. Subsequently, some of these slots are only partly used, so that the monopoly output is not exceeded. This allows the restriction of the total output in order to keep congestion low but flight fares high.

In the symmetric case of this model, strategic behavior takes place in the above sense: The quota buyer may decide after the trade whether he will utilize or hoard the slots purchased. If the total number of quotas permits, it will revert to its monopoly output and hoard the remainder of the slots. By contrast, if the total number of quotas remains below the monopoly output, all slots are fully utilized. This type of slot hoarding negatively contributes to welfare because it reduces overall output. Moreover, it increases the trading potential by increasing the buyer's potential payoff from a trade.

However, strategic behavior in the asymmetric case contrasts with this pattern: It occurs prior to slot trading and consists in the babysitting of slots in order to avoid a hand-back and re-allocation. Otherwise, the slots will be re-allocated to and utilized by the leisure airline. As a consequence, the output effect on flight fares and on congestion occurs whether trading takes place or not. For this reason, costly slot babysitting is the dominant strategy for the business airline. As opposed to the symmetric case, the business airline always utilizes all of its slots. Two particular properties characterize this kind of strategic behavior: First, it is introduced by the imposition of the slot trading rules rather than by the secondary trading opportunity itself. And second, it positively affects welfare by causing the business airline's excess number of slots to be utilized. This raises overall output away from the monopoly output and increases allocation efficiency. This type of strategic behavior is thus exactly opposite to strategic behavior in the symmetric case, both in terms of slot usage and in terms of the welfare effect.

From a welfare perspective, strategic behavior should be suppressed in the symmetric case but fostered in the asymmetric case. Surprisingly, the action to reach this goal is the same for both initial allocations: Imposing the quota trading rules that dictate the mandatory hand-back and re-allocation of unused quotas. The trading rules will equally increase welfare in both cases, yet for different reasons: In the asymmetric case, by encouraging the slot-holding airline to babysit its unused slots; and in the symmetric case, by avoiding the hoarding of unused slots after trading.

14.2.4 Effect of Secondary Trading on Market Shares

The effect of trading on market shares may be distinct between the symmetric and the asymmetric case: In the asymmetric case, only the business airline can be the seller. If trading takes place, the leisure airline gains access to the peak period and the resource allocation becomes more balanced. In the symmetric case, trading should only be allowed if the business airline becomes the buyer (which is the case in this model). This means that the inequality within the market shares would actually increase with a trade.

As the welfare analysis has shown, allocation efficiency increases when the business airline increases its output. This happens because both the network size and overall output are increased. From a social planner perspective, increasing the inequality of the market shares increases welfare. This is also true if the inequality is not induced by secondary trading itself but by the imposition of the trading rules. However, imposing and enhancing the inequality in the market may be perceived as an unfair and anti-competitive consequence of the secondary trading scheme.

This concern corresponds to that mentioned by Brueckner (2002b, p.22) for an asymmetric congestion toll, where imposing a low toll on the major carrier and a high toll on the smaller competitor may "appear to be the wrong response to concerns about market power by the dominant hub airline". However, as in Brueckner (2002b), the welfare analysis in this study is comprehensive and includes all distortions that affect allocation efficiency. Moreover, trading only takes place if the trade is worthwhile for both airlines. Consequently, it has to be beneficial from a distributional point of view. In the same sense, any "anti-competitive" concern referring to the inequality of the market shares hence also would be "misplaced".

Nevertheless, this potentially counter-intuitive relationship between market concentration and allocation efficiency should be addressed in any attempt at an implementation of a secondary trading scheme. The only exception is the asymmetric case where the trading rules cannot be implemented, so that the second-best welfare improvement also induces a more equal allocation of market access. Because this solution is inferior to the imposition of the trading rules from an overall perspective, however, it is of a lesser practical significance.

14.2.5 Implications for Practice

The above results imply that an implementation of a secondary trading scheme may only be desirable in one situation: In a symmetric case arising from arbitrary quotas where it is ensured that the networking airline is the buyer. In an asymmetric case with individual quotas, the welfare effect of trading is ambiguous, and trading will never occur. Yet, the mere imposition of the trading rules invokes a first-best allocation. Consequently, the trading rules should be applied in the sense of quota utilization rules after the initial allocation of individual quotas.

In the symmetric case, the networking airline preempts the peak-period market by buying the entire number of quotas. If the trading rules are applied, the business airline subsequently has to utilize all these quotas. As a result, overall output remains constant and the network size increases. Allocation efficiency is hence increased based on the secondary trading scheme in conjunction with the trading rules. However, the welfare result is only second best as the symmetric case presupposes an arbitrary number of quotas that is inefficiently low. If there is no re-allocation of unused quotas, the business airline will only supply its monopoly output. In that case, the welfare result becomes ambiguous because the increased network size may not be able to overcompensate the overall output reduction. In the symmetric case, a secondary trading scheme should hence be applied after the initial allocation of the arbitrary constraints. It must, however, be combined with the trading rules; otherwise, the welfare result not only fails to be first-best but may even become adverse.

In the asymmetric case, the leisure airline could achieve a positive peak-period output with a trade and may account for a second-best efficiency gain by increasing overall output. However, trading will never take place because the dominant network airline is always best off with an exclusive market access. Still, the trading rules themselves invoke a first-best quota allocation because they enforce the networking airline to provide the efficient network size. As a result, the socially optimal market structure is replicated. This yields first-best allocation efficiency. The trading rules hence replicate a service obligation for the business airline and therefore should be applied after the initial allocation of the individual quotas.

In both cases, the first-best welfare allocations require that the networking airline enjoy an exclusive market access. As already pointed out, this market preemption may invoke anti-competitive concerns. Nevertheless, this allocation is optimal both in terms of efficiency and rent distribution. The increased inequality in the market shares can hence be justified both from a welfare and from a policy point of view.

The above findings are similar to Verhoef's (2010) result, where a more efficient airline buys its competitor out of the market and reverts to its monopoly output, based on strategic hoarding of the quotas purchased. Also in that case, the trade is beneficial for both airlines and takes place. The quota hoarding also arises from the absence of a trading rule enforcing the utilization of the quotas and also yields an adverse welfare effect. The difference in the asymmetric setting at hand, however, is that the market power distortion not only reduces overall output but also yields an inefficiently small network size. However, the trading rules can be shown to correct for both these inefficiencies in the arbitrary as well as in the individual quota case.

Four essential implications hence arise from the above considerations: First, the result from this model does not strictly comply with the usual implications for secondary trading in perfectly competitive markets, where secondary trading improves welfare and increases the equality of the market shares: In both the asymmetric and the symmetric case, trading may also yield an adverse welfare effect. Moreover, in the former case, it is the trading rules rather than the trading scheme itself that account for the efficiency gain.

Secondly, in practice, a corresponding slot utilization rule is implemented. Regardless of whether the number of quotas is correctly determined or not, this should at signify that any prevailing market power distortion is overcome.

Thirdly, if a quota regulation were to be established from scratch, airport access rights could also be sold or auctioned to the airlines prior to secondary trading. The slots then would already be allocated according to the willingness to pay of the airlines, so that a trading solution after the initial allocation would become obsolete. The allocation efficiency results, however, should remain the same as in the secondary trading case.

Lastly, it may be deemed unrealistic that a planner first attempts to impose an optimal initial quota allocation and thereafter allows for secondary trading. However, one might argue that the planner is aware that he cannot replicate the efficient market shares based on the naive initial allocation alone. In that case, the implications from above still apply and may crucially affect the optimal design of the corresponding airport capacity allocation scheme.

14.3 Congestion Pricing

14.3.1 Allocation Efficiency

The results from congestion pricing are the least surprising in this model: Evidently, if marginal external delay costs are known, they can simply be internalized by the congestion tax. This removes the congestion externality by definition. Due to the dual distortion, however, output becomes inefficiently low, and network density benefits are reduced. The chance for a positive welfare effect thus decreases with an increasing importance of market power and density benefits. In contrast to the quotas, the tax always reduces output below the optimum regardless of the initial output level. Moreover, it always decreases the network density benefits. Therefore, its potential for a second-best welfare improvement is lower both against a quota solution as well as against a symmetric airline case without network density benefits.

In order to avoid the standard welfare caveat from the output inefficiency, recent studies suggest compensating the tax for the market power distortion (cf., e.g., Verhoef, 2010). In a homogenous setting, the result is that overall output generally becomes efficient. In the asymmetric model, however, the adverse welfare contribution from reducing the network size still remains. Consequently, the compensation of market power would increase the chance for a second-best improvement but could not correct for the network inefficiency. In this respect, also the deviation from the optimal network size induced by the concavity of the network value could be accounted for within the congestion pricing scheme. As a consequence, a first-best allocation would be achieved. In this case the tax would become negative and revert to a subsidy for the network services, reflecting a monetary compensation for the corresponding service obligation in terms of network provision.

Lastly, the investigation of secondary trading showed that both above inefficiencies could also be compensated by an individual quota scheme in conjunction with the trading rules of the secondary trading scheme. In contrast to a subsidy, such a quota solution would replicate the above service obligation free of charge. As a result, it should be preferred both from a political and an economic point of view.

14.3.2 Implications for Practice

In the presence of market power and network density effects, a congestion tax seems not to be the optimal choice for airport capacity allocation. Two reasons account for this conclusion: If the tax does not include its effect on the market power distortion and on the network size, it is likely to yield an adverse welfare effect. If it does compensate for these distortions, it will become a subsidy and thus require the public to pay for a service obligation that could also be implemented free of charge by a quota solution with a use obligation and a re-allocation scheme. In the latter case, in addition, the congestion tax would become a combined instrument for externality and competition regulation and would no longer represent a genuine PIGOUVIAN tax. As noted by De Wit and Burghouwt (2008, p.148), this contrasts with the TINBERGEN principle, which dictates that each market distortion should be treated with its own regulation instrument.[86]

Besides the above arguments, a tax solution implies several practical concerns: In theory and with perfect information, the computation of the correct tax and the resolution of its ambiguity are straightforward. However, the quantification of delay costs is still considered to be difficult (see, e.g., Cook 2007b, p.97 and Matthews and Menaz 2008, p.25). The same problem may reasonably be argued to concern the estimation of the passengers' network density benefits. As a consequence, the practical implementation of the tax into a "manageable system with understandable tariffs" might prove a major challenge (Odoni, 2001, pp.40). Moreover, the inverse relationship of the tax to each airline's market share should be expected to be conceived as politically controversial. As a consequence, a congestion pricing scheme might be misinterpreted as an additional market barrier, although this judgment remains unjustified from an economic perspective (cf. Brueckner, 2002b, p.22). Particularly in a market structure where network density benefits are present and create an strong airline asymmetry, this distributional concern could significantly impede an implementation of this allocation instrument. Together with the above objections from an efficiency perspective, these concerns may raise justified doubts about the practicability and the applicability of this instrument.

[86] More precisely, Tinbergen's "Golden Rule" subsumes that if (at least) one individual regulation scheme is dedicated to each dedicated objective then regulation policy will successfully achieve its aim (Acocella et al., 2012, p.2).

Part III. Extension: Simulation based on parametric Model

For the illustration of the previous results and for further analysis, this part presents a quantitative simulation of the results obtained from the generic model. For this purpose, the generic model from Part I is first specified with parametric functions and thereafter simulated by assigning numeric values to the parameters. In order to obtain comparative results, two distinct sets of numeric parameter values are applied. As a consequence, the specified parametric functions are less generalized than those of the generic model. However, the numeric results confirm the general results as discussed in Section 9 and comprehensively illustrate the properties of this particular model with the network density benefits. In addition, comparative statics yield some particular and surprising insights that could not have been obtained from the generic model alone.

15 Parametric Model

15.1 Method

The specific functions used in the generic model are drawn from the literature. Where possible they are chosen to be linear in order to provide a traceable analysis of the model. However, the delay cost and the network density benefits functions exhibit quadratic forms by nature. Despite these quadratic forms, however, the specified model can be solved analytically.[87] The benefit of linearity is thus that parametric closed-form solutions are obtained. In contrast to the implicit equilibrium conditions of the generic model, these explicit terms can ultimately be investigated by means of comparative statics in terms of all model parameters.

The specified linear version of the generic model is first solved by computational means with the full set of parameters.[88] These parametric results, however, are quite extensive and provide little clarification for an intuitive interpretation. Therefore, these solutions

[87] This shows that non-linear basic functions would yield polynomial functions of higher degrees for the delay cost functions and hence for the externalities. The required numeric solution techniques required would generally not allow us to obtain closed-form parametric solutions for the first-order conditions and equilibrium values (cf. e.g. Motta, 1993).

[88] The extensiveness of the parametric terms and the non-linearity of the delay cost functions motivate the use of the software package Mathematica 10.0 for the analytic computations and the subsequent simulation. This ensures the accuracy of the results and allows us to obtain comprehensive graphs to illustrate the numeric results.

are made accessible in two ways: On the one hand, a parameter reduction is applied to the full parametric solution. This considerably simplifies the explicit terms by providing a comprehensible form of the output quantities. These terms can be used to analytically interpret and explore the basic characteristics of the linear model in a qualitative fashion at least to a limited extent. This solution set is referred to as the parametric solution.

On the other hand, the full parametric solutions are further specified with numeric values for distinct sets of parameters. This allows us to obtain numeric results that are used to investigate the linear model's results in a quantitative manner. In addition, the numeric specification serves in the application of comparative statics. The latter illustrate the linear model's typical characteristics under different parameter settings and indicate the occurrence of the corner solutions against the interior solutions in various cases. In addition, the comparative statics are used to perform a sensitivity analysis, which shows both the continuity and the stability of the numeric model for a reasonable parameter range.[89] This solution is correspondingly referred to as the numeric solution.

15.2 Specification

Passenger utility is specified as follows: Direct utility from flights draws Tirole's (1988) ubiquitous specification for vertical product quality, so that direct flight benefits are defined as

$$b_k(\theta) = \beta_k \cdot \theta.$$

Index $k \in \{o, p\}$ denotes the peak and the off-peak-period. Parameter $\beta_k > 0$ is introduced to dissociate the qualities of the direct flight benefits across the two periods.[90]

Indirect network benefits are also based on this specification but are modified in order to reflect the characteristics delineated in Section 5.2. They are assumed to represent a product of the willingness to pay and peak-period flight volume of the business airline, reading

$$d(\theta, n_p^B) = \delta \cdot \theta \cdot n_p^B.$$

Scale parameter $\delta > 0$ determines the magnitude of the density benefits relative to direct

[89] Continuity addresses the mathematical property that there are no abrupt (i.e. discontinuous) changes in the parametric functions when parameters change by small amounts. This property is required for a function to be differentiable. Stability, in turn, refers to the insensitivity of the results to small changes. This means that reasonable parameter changes do not change or overturn the general meaning of a result.

[90] This specification is a standard in the vertical differentiation literature (cf. e.g. Shaked and Sutton, 1982, and many more). Tirole (1988, p.296) specifies $u(\theta, s) = \theta \cdot s$ where s denotes endogenous product quality. In this model, qualities for direct flight benefits are exogenous, so that $s = \beta_k$ for $k \in \{o, p\}$.

utility from flights and the degree of the exogenous airline asymmetry. Moreover, as in the generic model, the equivalence $\theta^D = 1 - n_p^B$ illustrates the concavity of the network value.

The time costs are linearized as $t(n_p^L, n_p^B) \equiv \tau \cdot \left(n_p^L + n_p^B\right)$ because they depend on aggregate peak-period output. In the linear model, utility from (1) thus becomes

$$u(\theta, x) = \begin{cases} \beta_o \cdot \theta & \text{for } x_0 = 1, \\ \beta_p \cdot \theta - \tau \cdot \left(n_p^L + n_p^B\right) & \text{for } x_p^L = 1, \\ \beta_p \cdot \theta - \tau \cdot \left(n_p^L + n_p^B\right) + \delta \cdot \theta \cdot n_p^B & \text{for } x_p^B = 1. \end{cases} \tag{43}$$

In accordance with the generic model, the assumptions $\beta_o, \beta_p > 0$ and $\beta_p > \beta_o$ still apply.

The congestion costs are parametrized as $g\left(n_p^i + n_p^j\right) \equiv n_p^i \cdot \gamma \cdot \left(n_p^i + n_p^j\right)$ with $\gamma > 0$. In contrast to the time costs, they accrue to each airline according to its own peak-period flight share only. Consequently, they are multiplied with the individual outputs. This illustrates that the externalities yield non-linear convex functions despite the underlying linearity. The generic notation $i \neq j \in \{B, L\}$ is used to specify the symmetric congestion cost function for each airline.

Marginal operating costs $c \cdot \left(n_o^i + n_p^i\right)$ are constant, as specified in the generic model. The airline profit function from (3) thus finally becomes

$$\Pi^i[n_o^i, n_p^i] = f_o \cdot n_o^i + f_p^i \cdot n_p^i - c \cdot \left(n_o^i + n_p^i\right) - n_p^i \cdot \gamma \cdot \left(n_p^i + n_p^j\right) \tag{44}$$

and again is symmetric across airlines.

15.3 Parameter Choice

In the simulation, two different parameter sets are evaluated: a simple set and a balanced set. In order to ensure partial market coverage, the population range is calibrated endogenously, as explained in Section 8.4.[91]

15.3.1 Parameter Sets

The two parameter sets are depicted in Table 15.1, where direct benefits from off-peak-period flights are chosen to have slope unity in both sets for simplicity.

[91] Partial market coverage is an essential ex-ante assumption for the model specification both for technical and contextual reasons; see Section 5.3.2.

<table>
<tr><th>Function</th><th>Specification</th><th>Parameter</th><th colspan="2">Value</th></tr>
<tr><td></td><td></td><td></td><td>Simple</td><td>Balanced</td></tr>
<tr><td>Direct Off-Peak Flight Benefits</td><td>$\beta_o \cdot \theta$</td><td>β_o</td><td colspan="2">1</td></tr>
<tr><td>Direct Peak Flight Benefits</td><td>$\beta_p \cdot \theta$</td><td>β_p</td><td>1.5</td><td>2</td></tr>
<tr><td>Network Density Benefits</td><td>$\delta \cdot \theta \cdot n_p^B$</td><td>$\delta$</td><td>1</td><td>0.7</td></tr>
<tr><td colspan="2">Parameter Change for Corner Solution*</td><td>δ</td><td>1.5</td><td>2</td></tr>
<tr><td>Time Costs</td><td>$\tau \cdot \left(n_p^L + n_p^B\right)$</td><td>$\tau$</td><td rowspan="3">0.5</td><td rowspan="2">0.4</td></tr>
<tr><td>Congestion Costs</td><td>$n_p^i \cdot \gamma \cdot \left(n_p^i + n_p^j\right)$</td><td>$\gamma$</td></tr>
<tr><td>Operating Costs</td><td>$c \cdot \left(n_o^i + n_p^i\right)$</td><td>$c$</td><td>0.2</td></tr>
<tr><td>Population Range</td><td>$[\Theta - 1, \Theta]$</td><td>Θ</td><td colspan="2">1.5</td></tr>
<tr><td colspan="2">Max. Population Range for Partial Market Coverage**</td><td>Θ_{max}</td><td>2</td><td>1.7</td></tr>
</table>

Tab. 15.1: Simulation - Parameter Values

Subsequently, the simple set denotes a parametrization that is chosen to most closely resemble the simplified parametric model. Its aim is to show that this very basic model already provides meaningful results. In particular, the symmetric congestion and the peak-period benefit parameters are arranged so as to cancel each other out. This yields a completely symmetric setting across both periods with regard to congestion and to the symmetric flight benefits, which will be revealed in the sensitivity analysis in Section 16.3.2. As a result, the impact of the airline asymmetry on the output choices per period can be studied while abstracting from the cost-benefit asymmetry across the two periods. For this purpose, the simple direct flight benefits from peak-period flights are set to one and a half times the off-peak-period benefits. This yields a reasonable relative peak-period quality increase. The congestion and time cost parameters amount to half the off-peak-period direct flight benefits. This creates symmetry across periods both in terms of congestion costs and airline-symmetric benefits. For convenience, marginal costs are assumed to have equal value. The asymmetric network density benefits are set to unity in order to provide an additional yet moderate peak-density benefit in the above sense of simplicity.

In contrast to the simple set, the balanced set investigates a more symmetric setting with lower congestion, higher direct flight benefits and less important density effects. Therefore, it maintains the cost-benefit asymmetry across periods but at the same time reduces heterogeneity across airlines. In addition, the increase in the density benefits parameter chosen to invoke the corner solution is higher. This serves in investigating a more extreme change in the airline asymmetry. On this account, the balanced direct peak-period flight benefits are doubled as compared to off-peak-period benefits. At the same time, the delay cost values are lowered against the simple set, which improves the cost-benefit ratio for the peak-period passengers of either airline and hence amplifies the difference between the two periods for the purpose of an increased relative attractiveness of peak-period travel. By contrast, the initial network density benefits are less important than in the simple case. As stressed above, this aims at reducing the degree of product differentiation across the two airlines. The operating costs are further diminished in order to marginalize the impact of production costs on the equilibrium outcome.

In both sets, the value change of δ serves in investigating the case where the business airline's hub dominance may extend from the peak period to the entire market, based on an increased importance of the network density benefits (see * in Table 15.1). As discussed in 6.6, this corner solution occurs as follows: At first, the off-peak-period output choice of the business airline may turn negative when density benefits are large. This choice is invalid, however, because it violates the non-negativity constraint. As a consequence, the corner solution becomes governing, which changes the business airline's off-peak-period output to zero. At the same time, this suspends symmetry condition 11, which ultimately allows overall outputs to become asymmetric across the two airlines in both periods. The specific parameter setting necessary to invoke the corner solution is determined ex-post from the equilibrium outcomes. As a result, each above parameter set yields an interior and a corner solution. Note that the later sensitivity analysis identifies situations where the non-negativity constraint becomes binding for parameter changes other than the density benefits. Specifically, the case of marginal costs and of direct peak-period benefits are evaluated. However, with regard to the central topic of this model, the choice of the density benefits is most natural and interesting for the investigation of an asymmetric hub dominance corner solution.

In summary, the simple set cancels the cross-period asymmetry with regard to congestion and direct flight benefits. The balanced set, by contrast, reduces product heterogeneity between the two airlines and at the same time increases the attractiveness of peak-period travel by introducing a more favorable cost-benefit ratio. Both sets yield an interior and a corner solution based on distinct variations of the importance of the network density benefits. As the simulation will show, both sets yield similar results. In comparison, the simple set yields a more asymmetric allocation across periods and airlines, while the balanced set provides a more symmetric allocation of the flight volumes across airlines. Ultimately, this distinction enhances the illustrative character of the results, although the parameter differences are actually small.

15.3.2 Calibration

Parameter Θ, which scales the population range $[\Theta - 1, \Theta]$, is calibrated as follows: Recall that Θ denotes the highest willingness to pay of the individuals in the population range considered, and that the degree of market coverage is determined by the location of the population range within the consumer continuum. This means, however, that the degree of market coverage can only be determined when the equilibrium outputs are known. The appropriate value of Θ hence needs to be calibrated by means of iteration based on the ex-post analysis of the numeric results. This retrospective evaluation of market coverage is standard in the vertical differentiation literature (cf. e.g. Lambertini, 2006, p.64, and others). Because the population range and density have size unity, partial market coverage is simply assured when the sum of overall output is smaller than one.

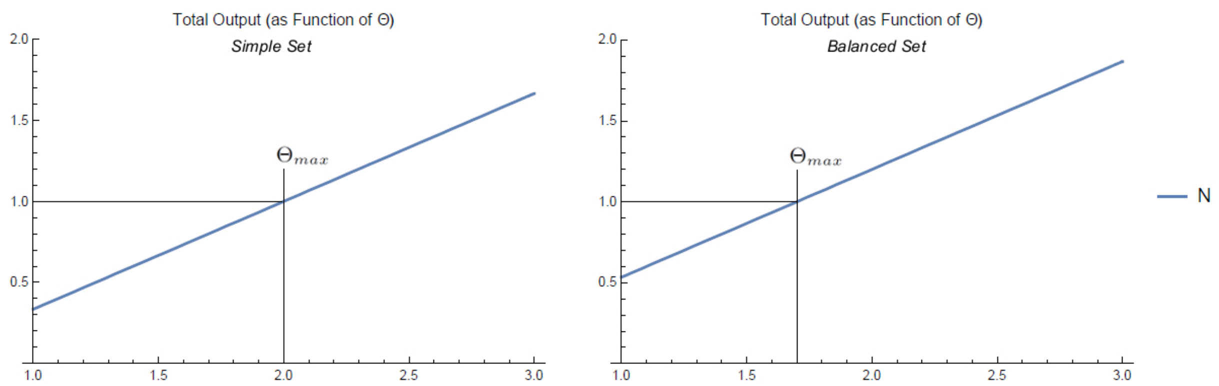

Fig. 15.1: Market Coverage and Population Range

Total output as a function of population range parameter Θ is shown in Figure 15.1. The two graphs illustrate that the chosen values satisfy $\Theta < \Theta_{max}$, so that partial market coverage is assured for both parameter sets. In this respect, the value Θ_{max} indicates the critical value for the willingness to pay where overall outputs revert from partial to full market coverage. It is also shown above in Table 15.1 (see **). In addition, Figure 15.1 reveals that the value of Θ_{max} is slightly lower in the balanced set than in the simple set. This reflects the fact that net utility is higher in the balanced set because direct peak-period benefits are more important but congestion and operating costs are lower. Moreover, the graphs indicate that the relation between overall output and Θ is linear. Linearity in Θ is also observed in similar vertical differentiation models from the literature (cf. e.g. Ecchia and Lambertini, 2006, p.86).[92] Ultimately, this calibration shows that the above choices for Θ ensure partial market coverage and thus the validity of the model specification within the parameter ranges considered.

[92] Ecchia and Lambertini (2006, p.86) find that the relationship between the equilibrium qualities and Θ is linear.

16 Equilibrium

In the following, the equilibrium of the linear model is explored analytically, based on a simplified, closed-form solution of the parametric model, as well as numerically and graphically, based on the results of the simulation and a corresponding sensitivity analysis. The parameter choice for the numeric results follows Table 15.1 in Section 15.3.

16.1 Parametric Solution

As mentioned above, the linear model can be simplified in order to obtain illustrative, analytic solutions. The goal is to simplify the equilibrium terms without introducing too much of a loss in generality. This is achieved by applying a reduction in the number of the scale parameters.

16.1.1 Simplification

First, the direct benefits from peak-period flights are related to off-peak benefits as

$$\beta_p \equiv \beta \cdot \beta_o,$$

so that $\beta > 1$ describes the peak-period benefit relative to an off-peak-period flight. Subsequently, the simplifications

$$\beta_o = 1,\ \tau = \gamma,\ \text{and } c = 0$$

yield that $\beta \cdot \beta_o = \beta$, so that both β_p and β_o are substituted away based on the new parameter β. Moreover, equating $\tau = \gamma$ simply equalizes the importance of congestion and time costs. Lastly, abstracting from marginal costs follows the general model of Ecchia and Lambertini (2006) and should equivalently not have a major impact.

16.1.2 Equilibrium Outputs

With the above simplifications, the parametric off-peak-period reaction functions become

$$n_o^B(n_p^B) = \frac{\Theta}{3} - n_p^B, \qquad \text{and} \qquad n_o^L(n_p^L) = \frac{\Theta}{3} - n_p^L,$$

which correspond to the symmetric reaction functions (8) in the generic model. Correspondingly, the substitution of these conditions into the peak-period reaction functions yields the peak-period reaction functions (9) and (10) as

$$\begin{aligned} n_p^L(n_p^B) &= \frac{1}{2}\cdot\left(\Theta\cdot\frac{\beta-1}{\beta+2\gamma-1}-n_p^B\right), \\ n_p^B(n_p^L) &= \frac{1}{3\delta}\cdot\left(1-\beta-2\gamma+\delta\pm\sqrt{(1-\beta-2\gamma+\delta)^2-3\Theta\delta(1-\beta)+3\delta(1-\beta-2\gamma)\cdot n_p^L}\right). \end{aligned}$$

Cross-substitution and re-arrangement of the above reaction functions reveals the explicit analytic equilibrium outputs as

$$\begin{aligned} n_p^B &= \chi, \\ n_p^L &= \frac{\Theta}{2}\cdot\frac{\beta-1}{\beta+2\gamma-1}-\frac{\chi}{2}, \\ n_o^B &= \frac{\Theta}{3}-\chi, \\ n_o^L &= \frac{\Theta\,(1-\beta+4\gamma)}{6\,(\beta+2\gamma-1)}+\frac{\chi}{2}, \end{aligned}$$

where $\chi \equiv \frac{1}{12\cdot\delta}\left(\sqrt{24\Theta\delta(\beta-1)+(3\beta-4\delta+3)^2}-3(1+\beta+2\gamma)\right)+\frac{1}{3}$ is introduced for notational brevity.

The parametric term χ can be quantified as follows: On the one hand, the scale parameters $\beta > 1$ and $\Theta > 1$ are larger than unity by definition. As the below numeric computation of the simplified model shows, however, both values need to remain small: Depending on the chosen parameter sets, $\beta = 1.5$ yields $\Theta \leq 1.7$ whereas $\beta = 2$ requires $\Theta \leq 2$ in order to ensure partial market coverage. With $\delta > 0$ and $0 < \gamma \leq 0.5$ as per definition, the term below the square-root is hence positive, while the second term within brackets is negative. At the same time, a non-degenerate equilibrium where peak-period output of the business airline is non-negative but smaller than unity requires $0 < \chi < 1$. The bracketed term in χ must therefore be larger than -4δ but remain smaller than $8\cdot\delta$.

The above equilibrium outputs can thus be interpreted as follows: Using quality $\chi = n_p^B$ in the business airline's off-peak-period output condition confirms the one-to-one substitution of output changes across periods. Moreover, this equality implies that the peak-period output of the leisure airline is restricted by the business airline's peak-period output in a linear manner. This linearity also holds for the leisure airline's off-peak-period output but in an opposite sense. This again confirms the proportionate peak-/off-peak-period output compensation. As a result, the absolute values of the equilibrium outputs depend on the parameter sets, while their interdependencies strictly follow the interpretation of the generic model's equilibrium in Section 6.

16.2 Numeric Results

For the simulation, the equilibrium conditions are again solved analytically but without the parameter simplifications from above. The only exception is that the direct peak-period flight benefits β_p are substituted by

$$\beta_p = \beta \cdot \beta_o.$$

This allows us to simplify the impact of the direct flight benefits into two separate effects: a simultaneous increase of the direct flight benefits in both periods, based on β_o, and an increase in the relative attractiveness of the peak period, based on β. In order to obtain quantitative results, the scale parameters are replaced by numeric values. Although the parameter assumptions remain basic, the numeric results are more illustrative than the analytic results. Moreover, a sensitivity analysis shows that these numeric results are not sensitive to parameter changes within a reasonable value range.

16.2.1 Outputs

The equilibrium outputs for the two parameter sets are shown in Table 16.1. From left to right, the table displays the results for both the simple and the balanced set in absolute values. Where applicable, the market shares within each period in percentages are additionally indicated in square brackets. For the corner solutions, the column is split into two half-columns in order to show the corresponding differences: The first half-column displays the theoretical (but invalid) interior solution, which involves a negative off-peak business output $n_o^B < 0$. It represents the first-best solution that the business airline would theoretically choose if optimization were unconstrained. These results are printed in italics to indicate invalidity. The second half-column shows the appropriate corner solution, where $n_o^B = 0$ becomes governing due to the non-negativity constraint. This dual representation is necessary to separate the two effects that arise with the corner solution: First, the increase in the importance of the density benefits that actually induces the corner solution; and second, the application of non-negativity as a binding constraint. Otherwise, a direct comparison of the two solutions would be inaccurate.

Interior Solutions Let us first consider the two interior solutions in Table 16.1: Above all, they confirm that the business airline dominantly serves the peak period. The linear model thus illustratively reflects the hub dominance of the business airline within the peak period. As the expected consequence, the leisure airline has a higher market share in the off-peak-

		Simple Set			Balanced Set		
		Int. Sol.	Corner Sol.		Int. Sol.	Corner Sol.	
Parameter Condition		$\delta = 1$	$\delta = 1.5$, $n_o^B < 0$	$\delta = 1.5$, $n_o^B = 0$	$\delta = 0.7$	$\delta = 2$, $n_o^B < 0$	$\delta = 2$, $n_o^B = 0$
Peak	n_p^B	.314 [77%]	*.384 [87%]*	.365 [84%]	.363 [61%]	*.478 [73%]*	.467 [72%]
	n_p^L	.093 [23%]	*.058 [13%]*	.067 [16%]	.235 [39%]	*.178 [27%]*	.183 [28%]
Subtotal	N_p	.407	*.442*	.432	.598	*.656*	.650
Off-Peak	n_o^B	.019 [08%]	*(-.050)*	0	.070 [26%]	*(-.045)*	0
	n_o^L	.241 [92%]	*.275*	.250	.198 [74%]	*.256*	0.233
Subtotal	N_o	.260	-	.250	.268	-	.233
Total	N^B	.333	-	.365 [54%]	.433	-	.467 [53%]
	N^L	.333	-	.317 [46%]	.433	-	.417 [47%]
Overall	N	.666	-	.682	.866	-	.884

Tab. 16.1: Equilibrium - Individual Outputs

period. The symmetry of total outputs across airlines in both interior solutions is an artifact of the model, as explained in Section 6.5.

Note, however, that the market shares between the two airlines are not inversely proportionate across the two periods; namely, the leisure airline's peak-period output is relatively higher than the business airline's off-peak-period flight volume. Hence, also the leisure airline benefits from the higher willingness to pay of the peak-period passengers although it cannot offer indirect utility from network density. This result follows from the fact that consumer heterogeneity in the vertical differentiation setting is based on each individual's distinct propensity to consume. In other words, peak-period profitability is generally higher than off-peak profitability for both airlines despite congestion and the benefit asymmetry. This increases the opportunity costs to abstain from the congested peak period and therefore puts upward pressure on overall peak-period output and thus on congestion.

Finally, within this cross-period asymmetry, the business airline's off-peak-period output is small in absolute terms. This confirms that the density benefits further increase the profitability of the peak period for the business airline. Consider, however, that any output expansion generally decreases flight fares in both periods. It is noteworthy that the business airline chooses a positive off-peak-period output at all. This shows that additional profits can be made by allocating some output to the uncongested off-peak period despite the negative impact of output expansion on prices.

Corner Solutions As mentioned above, the gross difference of the corner solutions and the interior solutions must be separated into the two independent net effects for the comparison. These consist of the impact of an increase in the importance of the network density benefits and of the application of the non-negativity condition as an effective constraint. The former effect is shown to be evident in either set: The left-hand semi-columns of each corner solution in Table 16.1 indicate that the higher density benefits substantially increase the business airline's hub dominance in the peak period. The higher market share is based both on an output expansion of the business airline and on a corresponding withdrawal of the leisure airline. This withdrawal is undertaken because it reduces the leisure airline's exposure to the increased peak-period congestion. In the generic model, it is referred to as the endogenous output adjustment. Moreover, the occurrence of the corner solution allows total outputs to become asymmetric across airlines. As a result, it is shown that a higher importance of the network density effects fosters the business airline's hub dominance.

The increased hub dominance in the peak period is accompanied by a corresponding counter-balancing effect on the off-peak-period outputs: The leisure airline substitutes its withdrawal from the peak period with an augmentation of its off-peak-period flight volume. The business airline, in turn, reduces its off-peak-period output. As discussed in the generic model, the unconstrained profit maximization would now require the business airline to choose a negative off-peak-period output. However, negative outputs are both theoretically invalid and practically infeasible. Therefore, the non-negativity constraint becomes effective and returns the business airline's off-peak-period output to zero. This signifies that output is no longer a global maximum but reflects the local optimum. The resulting impact on the outputs is shown in the right-hand semi-columns of the corner solution: The business airline slightly reduces its peak-period output to partly offset its inability to reduce total flight volume based on a further off-peak-period output reduction below zero. This output compensation illustrates the traditional market power effect in oligopoly, which implies that output can be adjusted to optimally balance the market price.

For the same reason, the leisure airline reduces its off-peak-period output although it can partly re-balance some of its output toward the peak period. This becomes possible because the business airlines peak-period output adjustment at the margin reduces congestion. As a separate effect, the incidence of the corner solution thus slightly diminishes the business airline's relative output strength against the theoretically invalid interior solution with a negative peak-business output. This represents the second net effect of the corner solution as compared to the interior optimum.

From a quantitative perspective, the effect of the increased density benefits (which actually leads to the corner solution) is large, whereas the isolated effect of the non-negativity constraint is almost negligible. Therefore, the corner solution both fosters the business airline's hub dominance over the leisure airline and extends that dominance from the peak period to overall output. In other words, the commercial impact of important density benefits induces an absolute hub dominance of the business airline across both periods.

16.2.2 Airline Profits, Flight Fares, Passenger Utility and Welfare

The equilibrium results for flight fares, airline profits, net utility per period and overall welfare are shown in Table 16.2. The same above four cases are depicted as in the analysis above: the interior solutions of the simple and of the balanced set, and each corner solution as invoked by the value change in parameter δ.

		Simple Set		**Balanced Set**	
		Interior Sol.	*Corner Sol.*	Interior Sol.	*Corner Sol.*
Parameter Value Change		$\delta = 1$	$\delta = 1.5$	$\delta = 0.7$	$\delta = 2$
Passengers per Period	O	.260	*.250*	.268	*.233*
	P-L	.093	*.067*	.235	*.183*
	P-B	.314	*.365*	.363	*.467*
Flight Fares per Period	O	.833	*.817*	.633	*.617*
	P-L	1.18	*1.13*	1.29	*1.21*
	P-B	1.39	*1.48*	1.46	*1.70*
Airline Profits	**L**	**.124**	***.107***	**.287**	***.234***
	B	**.223**	***.280***	**.400**	***.581***
Net Utility per Period	O	.034	*.031*	.036	*.027*
	P-L	.031	*.020*	.118	*.076*
	P-B	.215	*.265*	.417	*.599*
Total Net Utility		**.280**	***.316***	**.571**	***.702***
Overall Welfare		**.627**	***.703***	**1.258**	***1.518***

Tab. 16.2: Equilibrium - Flight Fares, Profits and Net Utility

Flight Fares and Airline Profits The flight fares in Table 16.2 reflect the expected characteristics: They are higher for the business airline due to the network premium and further increase from the interior to the corner solution. By contrast, the off-peak-period flight fares are both lower and decrease toward the corner solution. In this respect, note that the peak-period flight fares of the leisure airline in the corner solution are theoretical only because the corresponding output is zero. Both airlines' profits develop correspondingly: Firstly, the business airline consistently enjoys higher profits than the leisure airline. Moreover, the business airline can commercialize its peak-period hub dominance in the corner solution while the leisure airline's profits further decrease. These results replicate the predictions from the generic model's formal analysis (see Sections 6.5 and 6.6).

Net Passenger Utility Surprisingly, utility for the business airline's peak-period passengers increases along with a higher importance of the network despite the higher flight fares. This is derived from the comparison of the interior against the corner solutions. Although this increase occurs at the expense of both the peak-leisure and the off-peak-period passengers, overall welfare also increases.

Moreover, as Table 16.2 also shows, the peak-period output of the business airline increases in parallel. As a result, a higher importance of the network density benefits increases both the network size and overall output. This output expansion decreases market power by nature as indicated by the drop in the off-peak-period flight fares. Despite the decreasing market power, however, both the peak-period flight fares and overall profits of the business airline increase.

This result illustrates a core property of this model: The increasing importance of the network density benefits increases the business airline's profitability and flight fares, but also the network size, overall output, the peak-business passengers' utility and total social welfare. This shows that the network density benefits increase the competitive advantage of the networking airline, which contrasts with the traditional market power effect. They hence reflect a demand-side market power effect other than the output inefficiencies based on downward sloping demand, as pointed out in the formal analysis.

As a consequence, this result ultimately provides an answer to the dilemma of hub concentration: An increasing hub dominance of the networking airline as an effect of an increasing importance of the network density benefits induces a net beneficial welfare impact, although it increases both the flight fares as well as the competitive advantage of the networking airline over its competitor. The increased network benefits hence overcompensate the hub premium that the passengers have to pay, so that overall welfare increases.

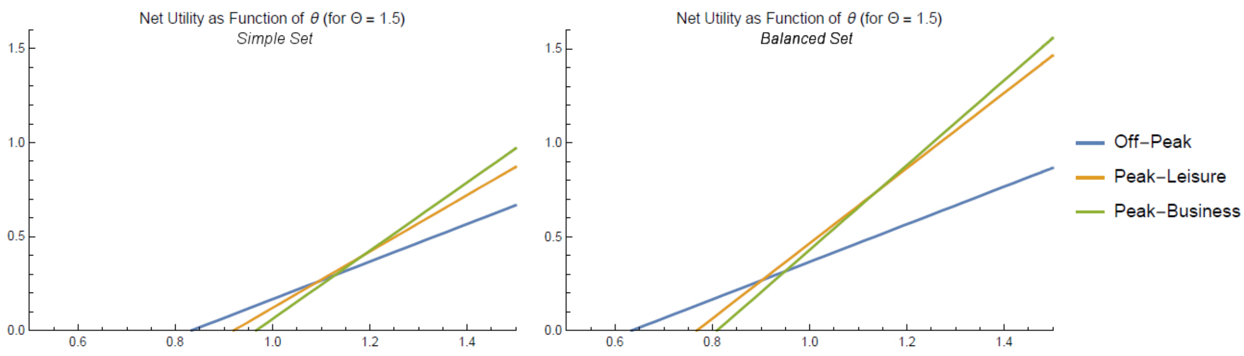

Fig. 16.1: Equilibrium - Net Utility and Peak-Period Preference

Flight Choice as a Function of θ Finally, Figure 16.1 depicts net utility for all types of flights across population range $\theta \in [\Theta,]$. The net utilities consist of the gross utility from flights minus the flight fares and hence directly represent customer value. As a consequence, the individuals' corresponding flight choices can directly be inferred from the two graphs: The left-most individuals have zero (i.e. a negative) value from flights and therefore do not travel. The passengers with a low willingness to pay then start with choosing an off-peak-period flight. In this portion of the population spectrum, the delay costs from peak-period travel overturn the gross benefits, so that both peak-period leisure and business travel are inferior to off-peak travel. Moving farther to the right on the θ-scale, however, this relation is inverted and peak-period leisure travel becomes attractive. This initial advantage owes to the fact that the flight fare of the peak-period business flights still overcompensates the benefits for low-θ travelers. Again, this relation again swaps so that the passengers with higher incomes can be shown to travel on the business airline during the peak period and hence are willing to pay for the network density benefits. The difference between the simple and the balanced graph reflects the higher gross benefits and lower delay costs in the balanced set, so that the overall number of passengers is higher than in the simple set. Ultimately, the two graphs in Figure 16.1 replicate the illustration of the discrete-choice demand system from Figure 6.1 in striking analogy.

16.2.3 Comparison: Simple vs. Balanced Set

Recall that the simple set applies moderate cost parameters and network density effects in relation to the direct flight benefits, while the balanced set applies lower congestion and operating costs but higher peak-period flight benefits. By contrast, the relative importance of the network density effects varies across the two sets: In the interior solution of the balanced set, the density effects are initially less important than in the interior solution of the simple set. However, the increase in δ which triggers the corner solution is higher in the balanced set, so that the previous relations are inverted (see Table 15.1 above). Put differently, as compared to the simple set, the balanced set reduces the importance of all cost functions and increases the importance of the direct peak-period flight benefits. Moreover, it simulates a less significant product heterogeneity in the interior solution but a more prominent airline asymmetry in the corner solution.

The comparison of the results across the two sets shows the following: In general, the outcomes may generally be judged as equivalent in a qualitative sense. Most importantly, both sets reflect an obvious, significant hub dominance of the business airline during the peak period. Also, the extension of the hub dominance to both periods is of similar magnitude. As a significant difference, it can be observed that overall outputs are higher in the balanced set. Undoubtedly, this is owing to the above parameter differences in terms of higher direct flight benefits and lower cost functions. In other words, the balanced set's more favorable cost-benefit ratio yields higher values of utility and thus higher flight volumes.

As a further result of the lower cost-benefit ratio in the balanced set, the business airline's hub dominance is reduced in both the interior and the corner solution as compared to the simple set. Table 16.1 indicates that this happens because the leisure airline can disproportionately increase its peak-period output relative to the business airline in both solutions. Considering the relative increase in parameter δ leading to the corner solution, this result is actually surprising. In the interior solution, it is explained by the lower density benefits in the balanced case, which reduce the airline asymmetry. As the lower cost-benefit ratio encourages an increased usage of the peak-period for both airlines, both a higher peak-period output and a lower hub dominance are the logical consequences. By contrast, the corner solution of the balanced set incorporates the highest absolute parameter value for the network density effects. This should assumedly yield the overall highest hub dominance. However, despite the high value of δ, the hub dominance remains lower than in the simple set. This result reveals that not only the density benefits affect the asymmetry between airlines; surprisingly, it is also the lower cost-benefit ratio that disproportionately fosters the leisure airline's output although it applies symmetrically to both airlines. Consequently, the above result establishes that the asymmetry between the airlines is not only reduced by lower asymmetric network density benefits, but also by a decrease of the symmetric ratio of overall costs to the direct flight benefits. This finding is quite unexpected; at least, it could not have been inferred from the formal analysis of the generic model.

As the sensitivity analysis will show, the direction of the above effect of the cost-benefit ratio accrues to the impact of the higher direct peak-period flight benefits rather than to the lower cost functions: While the direct flight benefits foster both airlines' peak-period output, they have a higher impact on the leisure airline. They hence relatively benefit the latter in an asymmetric way. By contrast, the reduction in the cost-side parameters is actually more favorable for the business airline (see Section 16.3 below). In sum, the overall effect in terms of peak-period output is hence more beneficial for the leisure airline, as the above results show. In general, however, the overall direction of this effect must be expected to depend on the absolute ratio of the cost and direct flight benefit parameters. Whether this effect might also be inverted under different parameter choices remains yet to be proven.

16.3 Sensitivity Analysis

The following analysis investigates the sensitivity of the equilibrium quantities to changes in the parameter choices from Table 15.1. It shows that the results are generally stable and not sensitive to input changes within reasonable parameter ranges.

Beforehand, however, it is important to point out the two boundary conditions that apply to the below analysis: Firstly, the parameter restriction Θ_{max} needs to be observed in order to ensure partial market coverage. As explained in Section 15.3, at full market coverage the results become invalid, because the demand functions are not appropriately defined. In particular, note that any parameter change may generally change the sum of all individual outputs. Consequently, also Θ_{max} may change. In other words, adherence to this limitation applies to all comparative statics, not only to variations in Θ itself. In the analysis below, the validity of the results is checked by verifying that overall output does not exceed unity; if applicable, the corresponding value range is denoted as invalid.

Secondly, if any of the individual outputs fall below zero, then the interior solution becomes invalid. In this case, either the respective corner solution is applied or the respective part of the graph is declared as invalid. Because the theoretical consideration of potential corner solutions from Section 6.6 only offers a foundation for a negative off-peak-period output of the business airline, the analysis will only evaluate the corner solutions for the respective cases. Other corner solutions are either deemed to be degenerate or not interesting; they are thus appropriately denoted by shaded areas in the graphs but not further investigated.

16.3.1 Outputs against Population Wealth

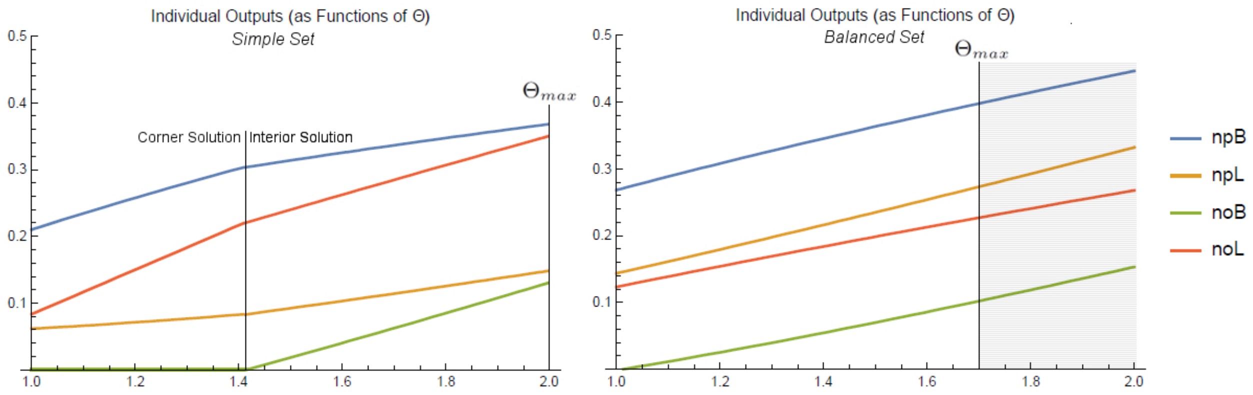

Fig. 16.2: Equilibrium - Individual Outputs against Θ.

Figure 16.2 shows the individual outputs as a function of the highest willingness to pay Θ for both interior solutions. This corresponds to the investigation of the impact of an increasing propensity to consume on demand across all potential customers. The definition of the population range as $[\Theta - 1, \Theta]$ makes clear that the graph must be limited to a value range of $\Theta \in]1, \Theta_{max}]$. The lower bound of this interval is given by the definition of the willingness to pay as $\theta > 0$ from (2) in Section 5.2.2. The upper bound is given by the assumption of partial market coverage and the corresponding specification of demand, as explained in Section 15.3. All values to the right of Θ_{max} are hence invalid. The exact values for Θ_{max} were evaluated numerically and are indicated in Table 15.1.

Above of all, it is evident that all outputs monotonously increase when the population's propensity to consume increases. Also, within each period, this increase is more or less proportionate across both airlines. This means that the business airline does not face higher demand when incomes are higher. The symmetry is quite noteworthy given the business airline's comparative advantage from the network density benefits and the corresponding higher utility.

Moreover, a corner solution is invoked in the simple set for the business airline when Θ and hence incomes are low. This corner solution again arises as the business airline's off-peak-period output theoretically becomes negative in the interior solution. As already discussed further above, it demonstrates the (invalid) technical optimality of increasing peak-period flight fares while holding peak-period output constant. This indicates that at very low prices, output provision of off-peak-period flights is not profitable for the business airline. Interestingly, this is not the case for the leisure airline although both airlines face the same marginal costs. This again confirms the higher profitability that arises from the network benefits. Therefore, the opportunity costs from shifting output toward the off-peak period are much higher for the business airline than for the leisure airline.

Finally, there is a most notable difference between the two sets: Both airline's off-peak outputs rise much more steeply in the simple set, both against their peak-period outputs and against the off-peak outputs in the balanced set. This shows that output growth rather takes place in the off-peak period when high congestion costs prevail (such as in the simple set). By contrast, if the output expansion is relatively unchallenged by congestion, the output growth converges across both periods.

In summary, a higher population's willingness to pay increases all outputs monotonously. If congestion costs are relatively high, this output growth is mainly focused on the off-peak period. Finally, approximate symmetry in this output growth maintains the business airline's hub dominance. The degree of hub dominance is hence more or less independent of consumer wealth.

16.3.2 Outputs against Network Density Benefits

The plots in Figure 16.3 show how the density benefits are key to the airline asymmetry: Starting from the left on the horizontal scale, density benefits are zero. Consequently, both airlines offer homogenous products, and their respective peak-period outputs are equal.

As a special case, observe that in the simple set on the left-hand side, all four individual outputs are identical. This result directly arises from the particular parameter choice, as depicted in Table 15.1: With $\delta = 0$, off-peak and peak-period utility only differ in the disadvantage of the delay costs and the advantage of the peak-period direct flight benefits. As delineated in Section 15.3, however, the respective parameter values $\gamma = \tau = 0.5$ and $\beta_p = 1.5$ cause these two differences to cancel each other out. As a consequence, all outputs are equal in the simple set when network density benefits are zero. By contrast, in the balanced set, the importance of congestion is reduced. The peak-period is hence relatively more attractive, which is reflected in the respective outputs. Nonetheless, symmetry still applies across airlines.

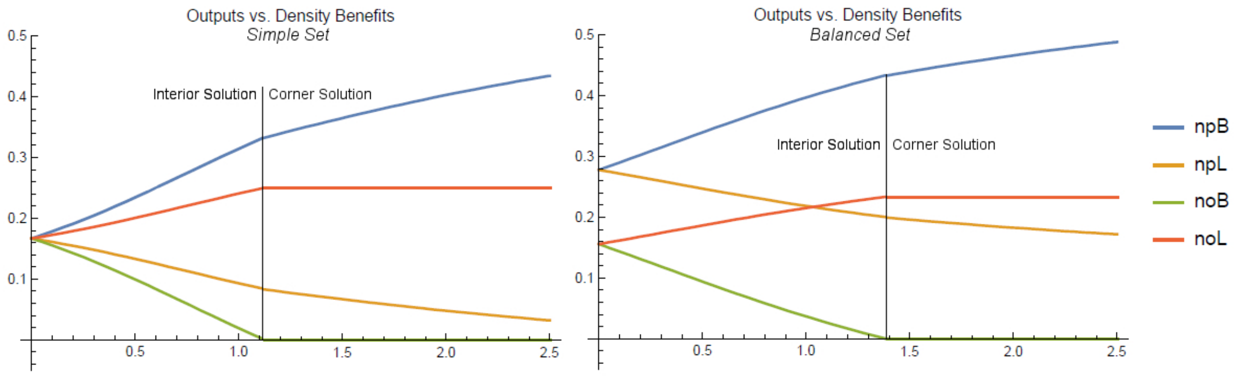

Fig. 16.3: Equilibrium - Outputs against Network Density Benefits

Introducing and increasing network density benefits means moving to the right on the scales in Figure 16.3. The result is as expected and intended: An airline asymmetry emerges, and the peak-period outputs develop in inverse proportion across airlines. As the business airline can offer network density benefits but the leisure airline cannot, the former eventually squeezes its competitor out of the peak period.

The graphs show that, initially, both airlines compensate their output changes across periods on a one-to-one basis. It is thus optimal for the business airline to keep total output constant despite the increasing network density benefits. This directly follows from profit maximization with the possibility of endogenous pricing in oligopoly. When output compensation is no longer possible, however, the corner solution becomes governing. It is only then that the business airline also increases its overall output. The increasing density benefits hence overcompensate the output expansion in terms of flight fares and congestion costs. This result was established in the analysis of the generic model by means of symmetry condition (11) in Section 6.5.

Note that in the corner solution, the leisure airline also omits cross-period output compensation. This happens although technically, the latter could further increase its off-peak-period output, and thus challenge the business airline's absolute market dominance. Nonetheless, it turns out to be more profitable for the leisure airline to accept a lower overall market share, instead of expanding output and further decreasing flight fares.

Overall, the density benefits lead to the intended airline asymmetry, as investigated in the generic model. Finally, the numeric computations and the symmetry condition show that overall output remains constant in the interior solution and increases slightly but remains below unity in the corner solution. The above analysis is hence valid for the full range of parameter δ.

16.3.3 Outputs against Direct Flight Benefits

As delineated above, the simplification $\beta_p = \beta \cdot \beta_o$ allows us investigating two distinct cases: Firstly, an overall increase in direct flight benefits for both periods via β_o, and secondly, a dedicated relative increase in peak-period flight benefits against off-peak-period benefits via parameter β. For both sets, Table 15.1 indicates that $\beta_o = 1$ is the initial parameter setting for the equilibrium.

The two graphs in Figure 16.4 show a variation in the direct flight benefits across both periods within the interior equilibrium. The shaded areas to the left indicate the invalidity of the corresponding results due to the occurrence of corner solutions based on negative output choices, which are considered as degenerate due to their low values of β_o. As the graphs show, the simple set only allows for a minor reduction below the static parameter value in order to maintain strictly positive outputs. This property arises as the relation of direct flight benefits to operating and congestion costs becomes too small with the variation in β_o.

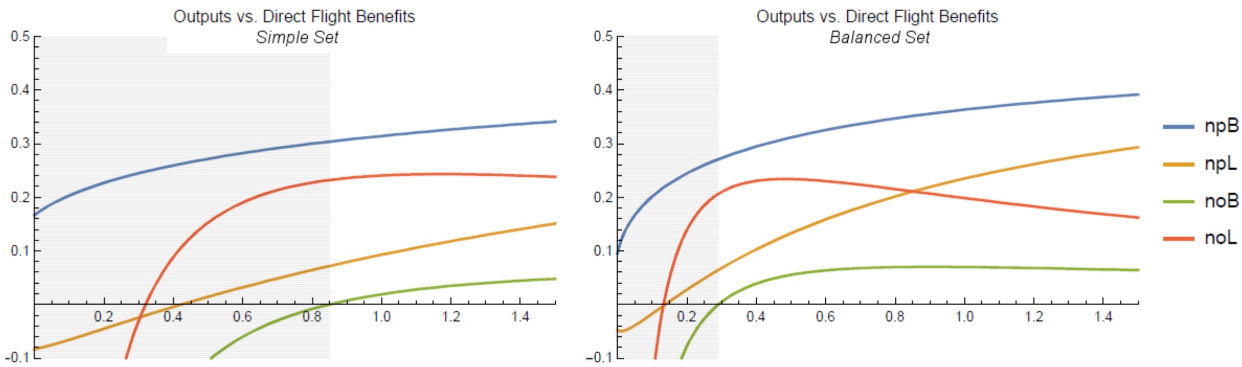

Fig. 16.4: Equilibrium - Outputs against Direct Flight Benefits

The balanced set is more permissive in this respect, which directly follows from the lower cost-benefit ratio as compared to the simple set. In general, a parameter reduction below the equilibrium value reduces both peak-period outputs. This is also straightforward, given the decreasing direct flight benefits against unchanged congestion cost parameters. Interestingly, however, the off-peak-period outputs diverge toward the left-hand side. This means that the business airline reduces its overall output, whereas the leisure airline partly compensates its output reduction with an increase of its off-peak-period output. Hence, the business airline is more concerned with counteracting the decreasing direct flight benefits and hence declining flight fares. By contrast, the leisure airline seems to aim at maintaining a certain level of output in order to maximize profits. Generally speaking, the high-quality firm tends to react to the price dissolution by increasing profit margins, whereas the low-quality firm attempts to maintain its overall output. Note that a similar yet less distinctive tendency is depicted in the simple set.

An increase in β_o above the static parameter choice yields the opposite effect on the equilibrium outputs. In this respect, notice that the leisure airline's peak-period output increases disproportionately, as compared to the business airline. This accompanies a corresponding decrease in its off-peak-period output. Nevertheless, this challenges the business airline's peak-period hub dominance. In other words, a symmetric increase in direct flight benefits is asymmetrically beneficial for the leisure airline's peak-period market share and yields an asymmetric relative reduction in the business airline's hub dominance in the peak period. Recall that this result was established in Section 16.2.3 based on a decreasing cost-benefit ratio. However, the comparison of the graphs in Figure 16.4 against the variable marginal costs from Figure 16.8 now proves that this effect accrues to the increase in direct flight benefits rather than to a decline of the cost-side parameters.

Finally, the graphs in Figure 16.5 show a strong reaction of overall output to a change in direct flight benefits. This result contrasts with the other parameter variations in this sensitivity analysis. Total output increases with increasing benefits and decreases with decreasing benefits. In this respect, the rate of change is slightly higher in the balanced set and increases in both sets toward the left-hand side boundary of the analysis. Nevertheless, the sum of all individual outputs still remains below unity, so that the results are valid.

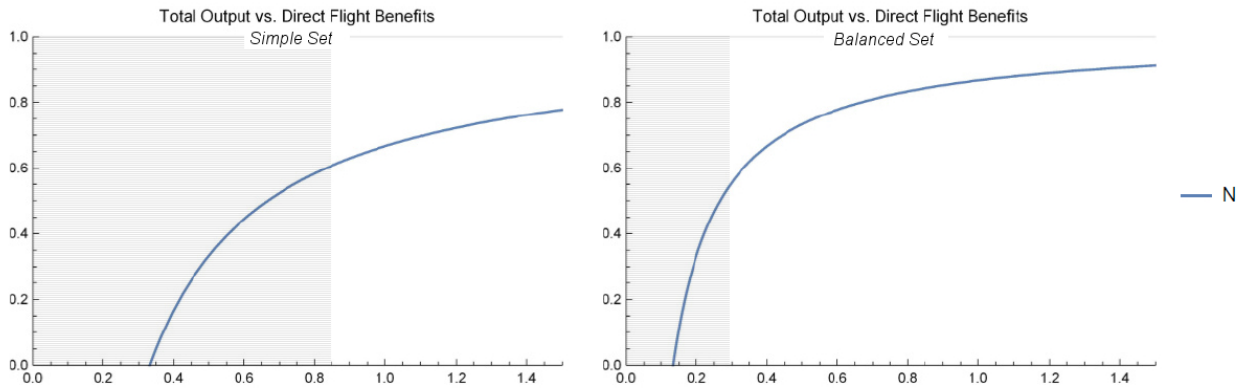

Fig. 16.5: Equilibrium - Flight Volume against Direct Flight Benefits

16.3.4 Outputs against Peak-Period Direct Flight Benefits

The variations in the peak-period direct flight benefits yield similar results to the above variations in overall direct flight benefits. These results are shown in Figure 16.6 for both sets. As the graphs show, an increase in the direct peak-period flight benefits also induces a reduction in the asymmetric peak-period hub dominance of the business airline. This confirms the crucial finding from above.

As opposed to direct flight benefits across both periods, however, two main differences can be observed that exclusively concern the variations in the peak-period benefits: Firstly, there is no divergence in the off-peak-period output adjustments across the two airlines toward the left-hand side of the scale (i.e. for reductions in β_p). Both airlines hence symmetrically shift their outputs toward the uncongested off-peak period when the peak-period flight benefits decrease. In this respect, recall that the symmetry of output compensation across periods for either airline follows from the symmetry condition of the generic model.

Secondly, a corresponding rise in the simple set invokes the corner solution for the business airline. As an explanation, both airlines again attempt to compensate their peak-period output growth through a respective reduction of their off-peak-period outputs. In contrast to the above case, the difference between periods increases and thus provokes more distinctive output quantity compensations. As a consequence, the corner solution again allows the business airline to expand its dominance across both periods, while the above asymmetric effect of a reduction in the peak-period hub dominance nevertheless applies. However, it ultimately remains ambiguous whether the benefit of the higher peak-period direct flight benefits in the corner solution is higher for the business or for the leisure airline.

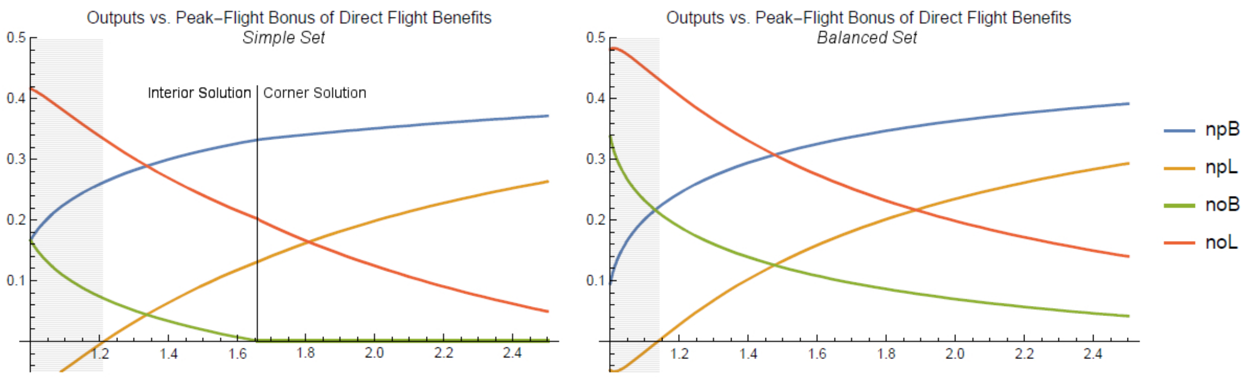

Fig. 16.6: Equilibrium - Outputs against relative Peak-Period Bonus

Overall flight volumes remain constant in both interior solutions but are slightly decreasing with decreasing peak-period benefits in the corner solution of the simple set. As a result, total output complies with the partial market coverage assumption. The respective graphs are shown in Figure C.1 in C.1.1.

16.3.5 Outputs against Congestion and Time Costs

The relationship between outputs and congestion costs is made clear in the graphs in Figure 16.7. The peak-period outputs decrease and off-peak-period outputs increase when congestion costs become more important. When the importance becomes very significant, the per-period outputs even cross over, so that the off-peak period bears the main portion of the overall flight volume.

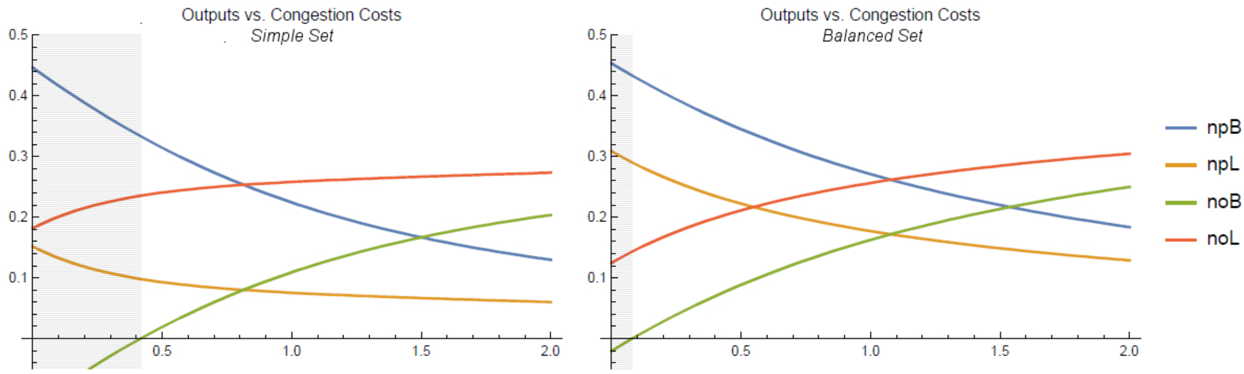

Fig. 16.7: Equilibrium - Individual Outputs against Congestion Costs

Moreover, the graphs show that the rates of peak-period output reduction correlate with the respective market shares: The business airline, which offers most of the peak-period flights, is most concerned with congestion. Therefore, it reduces its peak-period output more progressively than the leisure airline. This leads to an overall decrease in the business airlines peak-period market share. However, as the off-peak-period flight volumes increase correspondingly, the output compensation across periods equalizes the asymmetric declines.

As established in the generic model, in an interior solution the overall outputs need to be symmetric across airlines. An absolute hub dominance across both periods can thus only arise in a corner solution, which does not apply in this case. Moreover, the flights within the off-peak period are homogenous goods. Consequently, an increasing re-balance of outputs toward the off-peak period decreases product heterogeneity between the peak-period flights and hence the airline asymmetry. As a result, we may conclude that increasing congestion costs reduce the business airline's hub dominance even in the interior solution although these costs concern both airlines.

Individual outputs against time costs have exactly the same impact on the equilibrium outputs as the congestion costs. This can already be inferred from the generic equilibrium conditions, where passengers' time costs accrue to the airline to the same amount as congestion. Therefore, the graphs are fully identical and are appended in Appendix C.1.1, Figure C.2.

Finally, market coverage for both congestion and time costs remains constant in the respective interior solutions. This can be inferred from the numerical computations, for which the corresponding graphs are also shown in Appendix C.1.1. In the corner solution, total output slightly increases with decreasing cost parameters. However, the range of the corner solution is not crucial to the above discussion. Moreover, the change is marginal, and output remains below unity. Therefore, partial market coverage is ensured for the full parameter range of the analysis.

16.3.6 Outputs against Marginal Costs

The main lesson from the two graphs in Figure 16.8 is that in the interior solutions, both airlines' peak-period outputs are invariant to a decline in operating costs. An output increase that triggers higher congestion costs is not profit-maximizing, although it is countered by lower operating costs. This is, however, straightforward as a decreasing cost-base affects the flight fares across all periods. This is already made visible by the interdependence of the endogenous flight fares in equation (7) from the generic model (see Section 6.1).

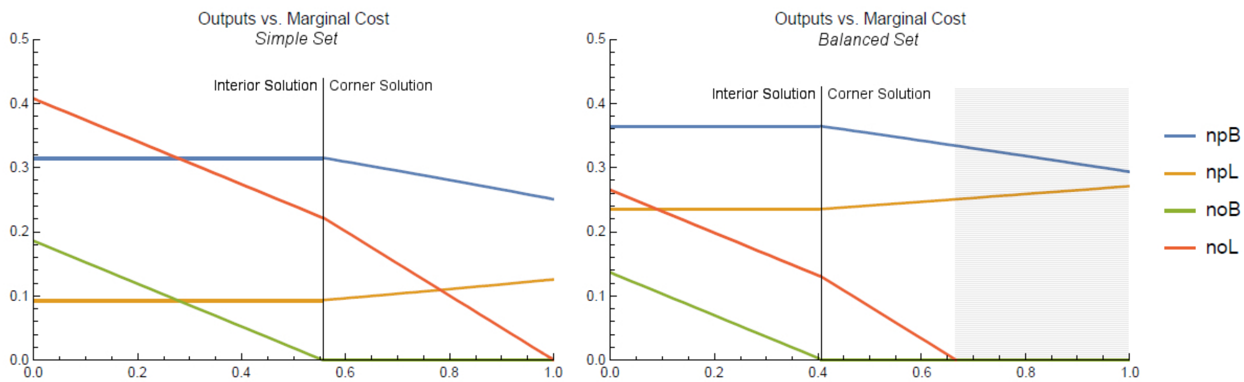

Fig. 16.8: Equilibrium - Individual Outputs against Marginal Costs

Moreover, the results confirm the above finding on the asymmetric reduction in the business airline's peak-period hub dominance based on a reduction in the symmetric cost-benefit ratio: As the graphs show, the cost-side parameters have no impact on the latter in the interior solution as they remain invariant. In the corner solution, the situation is different: In that case, again a reduction in the asymmetric hub dominance can be observed. Note, however, that this relationship is inverted against the above case, as the hub dominance decreases with an increasing cost-benefit ratio.

Ultimately, the following can be concluded: The asymmetric peak-period hub dominance of the business airline is variable with the cost-benefit ratio, albeit the latter is symmetric across the two airlines. In this respect, however, two separate effects need to be distinguished: If the variation is based on a decreasing cost-benefit ratio, for which the sole cause is increasing direct flight benefits, the hub dominance is reduced. If the variation is based on an increasing cost-benefit ratio, which is invoked by the corner solution following an increase of the operating costs, the hub dominance is also reduced. The two above-mentioned effects hence have the same impact on hub dominance, although they have an opposing influence on the cost-benefit ratio. The source of the change in the cost-benefit ratio hence must be accounted for as it is crucial for the direction of the asymmetric hub dominance effect.

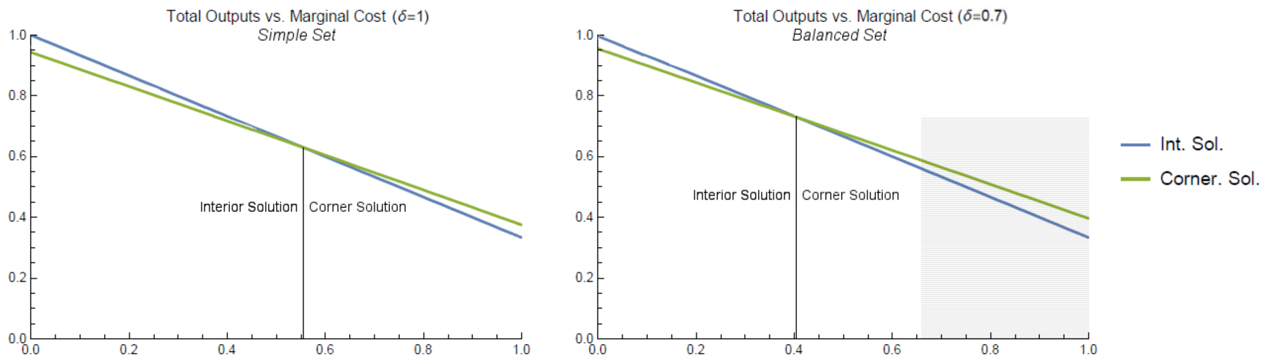

Fig. 16.9: Equilibrium - Flight Volume against Marginal Costs

Finally, the combined graphs in Figure 16.9 confirm that partial market coverage holds across the full range of marginal costs c.

17 Social Optimum

This section first investigates a simplified, analytic solution for the social optimum of the parametric model. Thereafter, it develops the boundary conditions for a computable solution, as this proves slightly more complicated than in the equilibrium. In analogy to the equilibrium investigation, the results of the linear model are subsequently presented both numerically and in a sensitivity analysis.

17.1 Parametric Solution

In order to obtain a parametric solution from the linear model, the simplifications are equivalent to the analytic equilibrium computation from Section 16.1.1 but with one distinction: Marginal costs $c > 0$ are maintained positive for illustrative purposes, as will become clear below. The results for the critical θ's are thus the following:

$$\begin{aligned} \underline{\theta} &= c, \\ \theta^* &= \frac{4\Theta\gamma}{\beta - 1 + 4\gamma}, \\ \theta^D &= \frac{\psi - \sqrt{\psi^2 - 24\Theta\delta\gamma - 9\Theta\delta^2 \cdot \left(\Theta - \frac{12}{9}\right)}}{3\delta}. \end{aligned} \tag{45}$$

Like in the equilibrium condition, the constant $\psi \equiv \beta - 1 + 4\gamma + (3\Theta - 2)\delta$ is introduced to facilitate the above terms.

Firstly, the off-peak-period condition illustrates two model properties concerning the marginal cost function that were revealed within the generic model: On the one hand, the restriction $c < \Theta$ for the parameter choice prevents a degenerate solution $\underline{\theta} = \Theta$ where the market is not served at all. On the other hand, the condition $c \leq \Theta - 1$ denotes the corner solution in the off-peak-period where all individuals travel, so that the market is fully covered. In the opposite case, $c > \Theta - 1$ yields an interior solution where not all individuals travel; this condition ensures partial market coverage as assumed in this model and again illustrates the corresponding parameter constraint.

The peak-period condition for θ^* shows that the optimal peak/off-peak period split in absence of network density effects is determined by the relation of direct flight benefits, congestion costs and the market coverage scale parameter. More precisely, θ^* decreases and thus peak-period traffic increases with higher peak-period flight benefits. This result corresponds to the analysis of the generic model. Recalling, however, that the network yields a corner solution renders this condition insignificant.

The expression for the network size condition θ^D is more cumbersome to interpret. Based on the signs of parameter δ one might assume that θ^D decreases and hence the network size increases with an increasing importance of the network density benefits. Although this relation seems obvious from a contextual view, however, it cannot be analytically derived from the above condition. The analytic value of this parametric solution therefore remains limited.

Lastly, these two conditions have to be compared for the applicability of the corner solution $\theta^* = \theta^D$. While this seems incomprehensible based on the above conditions, the monotonicity of the network density benefits from Section 9.4 warrants that this corner solution always applies.

17.2 Computable Solution

17.2.1 Welfare Function

The social optimum in the parametric model is computed in correspondence to problem (14) of the generic model. In a computable manner, the welfare function (16) can thus be written as

$$\int_{\underline{\theta}}^{\min\left[\theta^*,\theta^D\right]} \left[\beta_o \cdot \theta\right] d\theta + \int_{\min\left[\theta^*,\theta^D\right]}^{\Theta} \left[\beta_o \cdot \theta\right] d\theta + \int_{\theta^D}^{\Theta} \left[\delta \cdot \theta \cdot \left(1 - \theta^D\right)\right] d\theta$$
$$- \left(1 - \min\left[\theta^*, \theta^D\right]\right)^2 \cdot \left[\tau + \gamma\right] - \left(1 - \underline{\theta}\right) \cdot c.$$

The minimum value functions ensure that both peak-period conditions include their effect on congestion and direct flight benefits. In analogy to problem (14), the boundary conditions also need to be observed. This requires the implementation of the generic boundary conditions $\left\{\underline{\theta}, \theta^*, \theta^D\right\} \in [\Theta - 1, \Theta]$ and the fulfillment of the specific ordering rule $\underline{\theta} \leq \theta^* \leq \theta^D$. In this respect, all potential overlapping issues within the interior solutions are inherent.[93]

[93] As already pointed out, leapfrogging in the social optimum may arise from the fact that the optimization of the welfare function treats the critical θ's as independent variables. The ordering rule is hence neither implicit to the optimization nor covered by the non-negativity constraint of the individual outputs.

17.2.2 Boundary Conditions

For the computation, the above two rules can be consolidated into a set of nested minimum and maximum value functions. These nested functions read

$$\begin{aligned} \underline{\theta} &= \min\left[\max\left[\Theta - 1, \widetilde{\underline{\theta}}\right], \theta^*\right], \\ \theta^* &= \min\left[\max\left[\Theta - 1, \widetilde{\theta}^*\right], \theta^D\right], \\ \theta^D &= \min\left[\max\left[\Theta - 1, \widetilde{\theta}^D\right], \Theta\right]. \end{aligned} \tag{46}$$

In these functions, the variables $\widetilde{\theta}$ with a tilde superscript denote the unconstrained solutions for the social optimum problem (14). By definition, they are not yet subject to the boundary conditions.

These minimum and maximum value functions are evaluated as follows: The maximum value function ensures that each critical θ cannot be lower than the lower bound of the population range. If the lower bound is not binding, the minimum value function subsequently constrains any candidate $\widetilde{\theta} > 1 - \Theta$ to either one of two values: the upper bound Θ of the population range in case of θ^D, or to the next higher critical θ according to the above ordering rule in the case of $\underline{\theta}, \theta^*$. If this constraint is also not binding then the unconstrained $\widetilde{\theta}$ becomes the valid interior solution for the social optimum. By contrast, if either of the two constraints applies, the boundary condition automatically returns the candidate $\widetilde{\theta}$ to the respective boundary value. In this case the valid solution is the respective corner solution. Notice, however, that the boundary conditions (46) are not independent. Therefore, they either need to be evaluated in the reverse sequential order θ^D, θ^* and $\underline{\theta}$ or they can be evaluated simultaneously, which requires a nesting in reverse order.

With the nested conditions (46), each critical θ is only concerned with a single, computable constraint. This constraint evaluates the unconstrained solutions against the above ordering rule and the boundaries of the population range. The final result is the valid set of critical theta's $\left\{\underline{\theta}, \theta^*, \theta^D\right\}$, which either consists of valid interior solutions, or the respective corner solutions. For illustrative purposes, Table 17.1 below applies the unconstrained numeric solutions to conditions (46) but leaves them unevaluated. This shows how the validity of the resulting critical $\theta's$ is evaluated.

Before turning to the numeric computation and results, two further issues are important to mention: The difference in the characteristics of the corner solutions between the equilibrium and the indeterminacy of off-peak-period output allocation across the two airlines.

17.2.3 Corner Solutions in the Social Optimum

Above all, note that the general characteristics of the corner solutions in the social optimum are different from those in equilibrium. The reason for this is the absence of equilibrium symmetry condition (12) in the social optimum (see 6.4. Consequently, the social planner is free to allocate output across airlines and is only concerned with the critical theta's. In other words, he may choose asymmetric outputs either across periods or at an overall level for any parameter range.

By contrast, in the interior equilibrium, the symmetry condition dictates that overall outputs need to be symmetric across airlines. A corner solution can thus only occur if it is invoked by specific parameter values. The corner solution allows the business airline to achieve market dominance on an overall level. Put differently, in equilibrium a corner solution formally requires the suspension of symmetry condition (12). Whether this suspension applies is determined by the specific values of the parameter set. In the equilibrium analysis in Section 6.4, this corner solution has been invoked by an increase in the network density benefits.

The above difference has an immediate consequence concerning the occurrence of the corner solution: In the social optimum, the change in δ within the two parameter sets does not necessarily affect the applicability of the boundary conditions. This generally means that a corner solution may have arisen already before the change, or might remain absent despite that change. In the equilibrium, by contrast, this change itself is undertaken for the sole purpose of invoking the corner solution; the occurrence of the corner solutions in the social optimum does hence not generally correspond to the occurrence of a corner solution in the equilibrium. As the results will show, the corner solution always applies both to the optimal off-peak-period output as well as to the peak-period output of the leisure airline, independent of the parameter change in δ.

17.2.4 Indeterminate Allocation of Off-Peak-Period Output

Finally, note that off-peak-period output allocation across airlines is not defined in the social optimum computation: Problem (14) from the generic model in Section 7.1 shows that the welfare function is maximized with regard to the three critical thetas, and not with regard to individual outputs. This is reasonable, as flight fares and utility are equal due to product homogeneity in the off-peak period. In other words, the degree to which the airlines share the off-peak-period market is irrelevant from a welfare point of view.

Off-peak-period flight provision in a theoretically optimal planning solution hence remains solely a distributional issue. For simplicity, I therefore assume in the following that the entire off-peak market share accrues to the leisure airline. As the results will show, the airport presence of the business airline is overwhelming both in the peak-period and in absolute terms. This assumption therefore introduces a reasonable balance. Any interpretation of the social planner's distributional impact, however, requires due consideration of this presupposition and hence the necessary precaution.

17.3 Numeric Results

The computation of the numeric results again draws on the full parametric model. It thus abstracts from the simplifications taken for the parametric solution but for substitution $\beta_p = \beta \cdot \beta_o$, which remains valid. As a result, the parametric off-peak-period condition $\underline{\theta} = c$ from (45) becomes

$$\underline{\theta} = c/\beta_o$$

and will be further referenced below. Beforehand, Table C.4 in Appendix C.2.1 shows that the determinants of the 3x3 Hessian prove strict concavity for all cases. All solution sets presented in the following thus correspond to either global or local maxima, depending on the applicability of the corner solution.

17.3.1 Critical Theta's

Table 17.1 shows the numeric results of the critical θ's in the social optimum both for the simple and the balanced set. In each set the two distinct values for the network density benefits parameter δ are applied. This yields four distinct cases, which are separated by double lines. The first row in the table depicts the boundary conditions (46) in parametric form. They serve for comprehensibility of the interior solution validation: Each paragraph first presents the three candidate values for the unconstrained interior solution. Subsequently, this endogenous result is applied to the corresponding boundary conditions, which determine whether the corner solution applies or whether the interior solution is valid. In this respect, the maximum and minimum value functions are shown in unevaluated condition for illustration. The solutions of these numeric boundary conditions ultimately yield the optimal characteristic theta's either as local or global maxima.

Considering first the off-peak period's lower bound $\underline{\theta}$, the boundary condition becomes binding in all four cases. This signifies that corner solution $\underline{\theta} = \Theta - 1$ always applies. As this

	$\underline{\theta}$	θ^*	θ^D
Condition	$\min\left[\max\left[\Theta-1,\underline{\theta}\right],\theta^*\right]$	$\min\left[\max\left[\Theta-1,\theta^*\right],\theta^D\right]$	$\min\left[\max\left[\Theta-1,\theta^D\right],\Theta\right]$
Population Range		$\theta\in[0.5,1.5]$	
SIMPLE SET		$\delta=1$	
Candidate	0.5	1.2	0.94
Condition	$\min\left[\max\left[0.5,0.5\right],0.94\right]$	$\min\left[\max\left[0.5,1.2\right],0.94\right]$	$\min\left[\max\left[0.5,0.94\right],1.5\right]$
Solution	**0.5**	**0.94**	**0.94**
		$\delta=1.5$	
Candidate	0.5	1.2	0.80
Condition	$\min\left[\max\left[0.5,0.5\right],0.80\right]$	$\min\left[\max\left[0.5,1.2\right],0.80\right]$	$\min\left[\max\left[0.5,0.80\right],1.5\right]$
Solution	**0.5**	**0.80**	**0.80**
BALANCED SET		$\delta=0.7$	
Candidate	0.2	0.92	0.75
Condition	$\min\left[\max\left[0.5,0.2\right],0.75\right]$	$\min\left[\max\left[0.5,0.92\right],0.75\right]$	$\min\left[\max\left[0.5,0.75\right],1.5\right]$
Solution	**0.5**	**0.75**	**0.75**
		$\delta=2$	
Candidate	0.2	0.92	0.52
Condition	$\min\left[\max\left[0.5,0.2\right],0.52\right]$	$\min\left[\max\left[0.5,0.92\right],0.52\right]$	$\min\left[\max\left[0.5,0.52\right],1.5\right]$
Solution	**0.5**	**0.52**	**0.52**

Tab. 17.1: Social Optimum - Characteristic θ's and Boundary Conditions

corner solution denotes the left-hand side of the consumer continuum, it indicates that all individuals travel in the social optimum. In more detail, let us first evaluate the left-hand side of the respective boundary condition based on the parametric solution $\underline{\theta} = c/\beta_o$. This yields the maximum value function as $\underline{\theta} = \max\left[c/\beta_o, \Theta - 1\right]$ and thus shows why the interior solution and the boundary condition coincide in the simple set: The unconstrained solution simply equals the lower bound of the population range. As a consequence, the interior solution represents both the global and the local maximum.

As opposed to the simple set, the unconstrained value for $\underline{\theta}$ is lower than the left-hand boundary of the population range in the balanced set. This reflects the lower cost-benefit ratio $c/\beta_o < \Theta - 1$. The maximum value function therefore rejects the interior optimum and makes the boundary condition governing. This solution hence represents a local maximum only. As a result, full market coverage occurs in the social optimum. The full market coverage arises from the fact that the cost-benefit ratio c/β_o is inferior or equal to the population's lowest willingness to pay Θ. Partial market coverage would thus arise either if the cost-benefit ratio increased or if population wealth decreased.

Again in contrast to the above two cases, the valid solution of θ^* indicates that the interior optimum for the peak-/off-peak split never applies: The boundary condition dictates that the leisure airline must not serve the peak period at all. Formally, this is explained as follows: Given that the unconstrained solutions of θ^* do not violate the boundaries of the population range on either side, this boundary condition can be collapsed to $\min\left[\theta^*, \theta^D\right]$. This immediately clarifies that only $\theta^* \leq \theta^D$ can occur because otherwise the peak-period output of the leisure airline would have to become negative. Still, the results show that θ^* always remains within the defined population range. This denotes the fact that the social optimum would imply a non-negative peak-leisure output in the case where the asymmetric network density benefits were absent. In addition, note that the unconstrained θ^* does not change within either set when parameter δ changes. This invariability arises from the independent optimization of the characteristic thetas in the social optimum computation and reflects the fact that the leisure airline is not concerned with the network density benefits.

In turn, the result that the unconstrained solutions always yield $\theta^D < \theta^*$ confirms that peak-period output is always more valuable when it is provided by the business airline instead of the leisure airline. This reflects the fact that the former provides additional network density benefits at the same amount of congestion and hence supports the result concerning the monotonicity of the network density benefits presented as a main model property from Section 8.1. Hence, also in the parametric model, the regulator would need to expel the leisure airline from the peak period. This replicates the generic model's finding from Section 7.5.

Finally, the results for θ^D imply the following: On the one hand, comparing the parameter changes of δ within each set shows that the optimum network size $\Theta - \theta^D$ strictly increases with increasing network density benefits. This shows that a network expansion yields both more peak-density passengers and higher travel utility for all those passengers, which ultimately again illustrates the monotonicity of the network density benefits, where overall utility strictly increases with a larger network. Recall that this monotonicity contrasts with the equilibrium, where the business airline's profitability from the network depends on the network value. As the network value is a concave function of network size, it yields an inefficiently low network size (see Section 8.2). Ultimately, this confirms that the network exhibits increasing returns to scale in terms of utility from a welfare perspective but concave returns to scale from an airline perspective. On the other hand, comparing the outcomes across the two sets shows that the network size also increases with a more favorable cost-benefit ratio. This result arises because the peak-period output generally increases when congestion becomes less costly and peak-period benefits more important. The (invalid) interior solutions of the θ^*'s confirm this result as they yield a higher peak-period output in the balanced set, where the cost-benefit ratio is lower. Consequently, the congestion costs are the only limiting factor that counter-balance the above increasing returns of the network in terms of social welfare.

In summary, the above results show that the network exhibits increasing returns to scale from a welfare perspective that are only counterbalanced by increasing congestion costs from higher output. As a result, the leisure airline needs to be completely expelled from the peak period. The optimal size of the network thus depends on the importance of the network density benefits against the congestion costs and on the cost-benefit ratio of the peak- against the off-peak period.

17.3.2 Outputs, Flight Fares, Profits & Utilities

Table 17.2 now shows the optimal flight volumes, the corresponding flight fares, airline profits, net utility integrated across all passengers and overall social welfare. The flight volumes are directly computed from the critical θ's, where it is presumed that provision of the entire off-peak-period output is allocated to the leisure airline. This assumption resolves the indeterminacy of the off-peak-period output as delineated in Section 17.2.4.

		Simple Set		**Balanced Set**	
	Period	$\delta = 1$	$\delta = 1.5$	$\delta = 0.7$	$\delta = 2$
Passengers	O	.44	.30	.25	.02
	P-B	.56	.70	.75	.98
	P-L	0	0	0	0
Flight Fares	O	.50	.50	.50	.50
	P-B	.94	.87	1.08	.67
	P-L	-	-	-	-
Airline Profits	B	.09	.01	.43	.08
	L	0	0	.07	.01
Totals		**.09**	**.01**	**.50**	**.09**
Net Utility	O	.10	.05	.03	0
per Period	P-B	.57	.83	.90	1.91
	P-L	-	-	-	-
Totals		**.67**	**.88**	**.93**	**1.91**
Welfare		**.76**	**.89**	**1.43**	**2.00**

Tab. 17.2: Social Optimum - Outputs, Flight Fares, Airline Profits and Net Utility

The above results show that the peak-period output of the business airline (and thus the network size) is dominant in comparison to overall output. This dominance increases with a more favorable peak-period cost-benefit ratio. Correspondingly, the leisure airline is eventually expelled from the market. Again, it is the business airline who profits from lower costs and direct flight benefits although the latter are symmetric for both competitors. This result was established in the investigation of the equilibrium in Section 16.3.3.

However, the business airline should extend its market dominance to the off-peak period only in the balanced set with $\delta = 2$. In all other cases, the network size is large yet not overwhelming as compared to total output. As a consequence, the natural market structure as a whole tends toward a network monopoly in both periods if the density benefits are important and the cost-benefit ratio of the direct flight benefits is low. This confirms that in all cases of moderate parameter settings, an absolute hub dominance of the business airline is only justified in the peak period although the asymmetric characteristics of the network density benefits might have been suggesting a natural monopoly across the entire market from the very beginning. Again, this result was established in the investigation of the corner solutions in the generic model (see Section 6.6).

Subsequently, the flight fares reflect the marginal utilities from travel for the respective left-most passenger of each period. For the off-peak period, this implies that the flight fare is determined by either one of the two following parameters: by equalization of marginal utility and marginal costs in the case of a valid interior solution or by the population's lowest possible willingness to pay $\Theta - 1$ if the corner solution applies (see Section 17.2). In the simple set, both these values coincide, so that the interior solution and the corner solution are identical. By contrast, in the balanced set, the unconstrained optimization yields an excessive output because marginal costs remain at the lower boundary of the population range. As a result, the corner solution always becomes governing, so that the flight fares are equally determined by $\Theta - 1$ in both cases.

In the peak period, by contrast, all four flight fares are distinct. This reflects both the distinct cost-benefit ratios of the peak-period and the different network sizes. In this respect, recalling that the network density benefit can only be commercialized for the left-most peak-density traveler shows that the critical peak-business passenger is shifted to the left when the network size increases due to an increase in parameter δ (see Section 8.2). As a result, the flight fares decrease when the network size increases. The size of this change depends on the counter-weighing balance of the congestion costs in the peak period: As compared to the simple set, the network expansion in the balanced set is larger and the drop in flight fares is sharper when δ is increased. This occurs because the balanced congestion costs are less important and the direct flight benefits function is steeper than in the simple case. These arguments thus explain why flight fares are highest in the first case of the balanced set, where both congestion costs and network density benefits are low. Moreover, they illustrate the fact that flight fares decrease within each set when δ is increased. In conclusion, the business airline's peak-period airfares are inversely proportionate to the importance of the density benefits, and they essentially depend on the network size and on the peak-period cost-benefit ratio.

The airline profits are depicted in the next line. They are computed according to profit function (44) and hence equal turnover per period minus operating and congestion costs. The corresponding results thus reflect the different flight fares and cost-benefit ratios of the distinct parameter sets: The leisure airline's profits are zero in the simple set because off-peak-period flight fares equal operating costs. In the balanced set, where marginal costs are lower, a small positive profit remains. Yet, as Table 17.2 shows, the leisure airline's unconstrained output is larger than the corresponding constraint. Its constrained output, however, does not reach the first-best optimality from a welfare perspective, where marginal benefits offset marginal costs and profits are zero. The positive profit therefore represents the second-best nature of the corner solution. The business airline's profits, by contrast, are always positive. In particular, its highest profit arises from the balanced parameter set with $\delta = 0.7$. When the density benefits become very large and the network covers almost the full population range, however, these profits again collapse. This shows that a combination of low congestion costs, high direct flight benefits and relatively low density benefits account for the best profitability.

Furthermore, Table 17.2 shows that from left to right across the four parameter sets considered, peak-period utility strictly increases. This means that the optimal network size needs to be extensive and the cost-benefit ratio moderate in order to provide maximum utility for the passengers. The peak-period utility hence again relates to network size in a proportionate way. By contrast, the off-peak-period utility strictly decreases. However, as congestion and network effects are absent in that period, this merely denotes the vanishing off-peak-period flight supply. In particular, note that utility is particularly high in the rightmost case, where network size is largest. This reflects the fact that utility strictly increases with network size as it is not dependent on the willingness to pay of the left-most network user. The effect is even more prominent as the increasing scale returns in terms of utility are augmented by decreasing airfares when the network size increases. This again illustrates that the network density benefits hence exhibit increasing returns to scale to the users of the network from a social welfare perspective. This contrasts with flight fares and airline profits, which depend on the commercial value of the network. As the latter is concave, the largest network is not the one that is commercially most valuable. Hence, from an airline perspective, profitability follows the concavity of the network value.

Finally, overall welfare denotes the sum of airline profits and net utility for all passengers. Welfare hence also increases strictly from left to right. Again, the lower cost-benefit ratio and the higher network sizes positively affect welfare. In particular, net utility and welfare increase despite the variations in the flight fares. This shows that the actual value of the flight fares only has a distributional impact but does not affect efficiency from a social welfare perspective.

17.4 Sensitivity Analysis

The following analysis investigates the above results by means of comparative statics. It first considers the individual outputs as functions of population wealth Θ and of network density benefits parameter δ. Subsequently, both airlines' flight fares and profits and the passengers' net utilities are evaluated for changes of δ in a range from 0 to 2.5. The same analysis is conducted to evaluate social welfare as a function of the importance of network density benefits. This analysis also separates the two individual components of welfare, which consist in overall utility and total airline profits.

The results illustrate and enhance the findings from the numeric results. Moreover, they confirm that also the social optimum is not particularly sensitive to the parameter choices as it generally yields similar typical results for the entire parameter range considered.

17.4.1 Individual Outputs against Population Wealth

Let us first consider the outputs as a function of population wealth Θ, which are depicted in Figure 17.1 below. Beforehand, note that from each set the solutions with a moderate value for the network density benefits were chosen. The graphs of the extreme values for δ are similar but less illustrative.

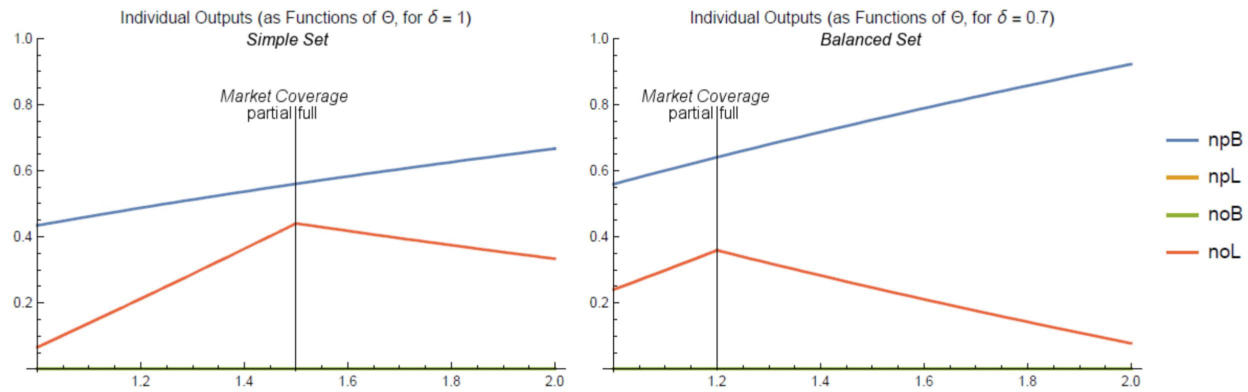

Fig. 17.1: Social Optimum - Outputs against Population Wealth

First and foremost, both graphs show that outputs initially are proportionate to the population's propensity to consume. Off-peak-period output of the business airline remains zero, based on the distributional assumption from above. For partial market coverage, proportionality generally corresponds to the equilibrium case from Section 16.3.1.

The most important insight from Figure 17.1 is that the monopolistic peak period prevails across the full range of Θ. This means that as expected, the willingness to pay has no impact on the general property of the social optimum. Furthermore, notice that off-peak-period output increases until the overall market is fully covered. Thereafter, the network grows further, while off-peak flights decrease in compensation. The turning point is indicated by solid vertical lines in both graphs. The respective value of Θ corresponds to the critical value as given by equation (23).

Increasing the network size hence consistently yields a higher additional value than increasing the number of off-peak-period flights. On the one hand, this property depends on the overall cost-benefit ratio and might turn out differently for a less favorable parameter set. Nonetheless, it is clear that the congestion and time costs become less important and the higher priced network more affordable when population wealth increases. On the other hand, it is again worth mentioning that this result is particularly exposed to the simplified definition of the time costs: Recall that for simplicity, time costs are defined as functions of overall peak-period output (and thus congestion) but not as a function of θ. As has already been explained, however, individuals with higher income and a higher preference for peak-period travel should be expected to have a higher time value and therefore higher time costs. Overall, this means that peak-/off-peak-period compensation in the above case for full market coverage might turn out to be less prominent than in the graph above. Although the cost-benefit considerations become more complex in that case, in the end it remains unlikely that the above result is overturned.

Lastly, observe that network size rises less steeply than off-peak-period output in the simple set. This is owing to the higher congestion cost parameters, which make the off-peak period relatively more attractive. Correspondingly, in the balanced set, output growth is approximately parallel and the turning point where full market coverage is reached occurs earlier.

17.4.2 Individual Outputs against Network Density Benefits

Next, both outputs are depicted against the network density benefits in Figure 17.2. These results confirm that the corner solution in the peak period arises irrespective of the importance of the density effects: It is imperative because peak-business flights yield higher utility than peak-leisure flights while marginal and social costs are equal (see Section 8.1). The discontinuity for the extreme case of $\delta = 0$, however, is not depicted in Figure 17.2: If network density benefits vanish, both airlines become symmetric. In this case, the blue line indicates the optimum overall peak-period output for both airlines. As in the off-peak period, the social optimum does not prescribe which airline provides this output. The case for $\delta = 0$ might therefore illustrate a symmetric output case with a positive peak-period output of the leisure airline as well.

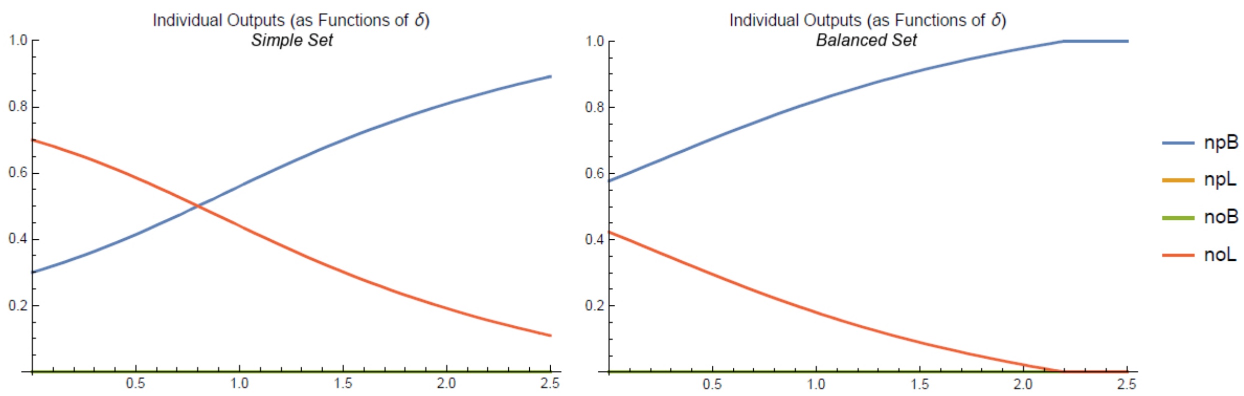

Fig. 17.2: Social Optimum - Outputs against Network Density Benefits

In all other cases, the outputs are functionally equivalent across the two sets while they reflect the different cost-benefit ratios: As presented in the numeric results section, in the simple set where congestion costs are more important, the initial network size is smaller and off-peak-period output is higher. Subsequently, the network grows in proportion with the value of δ. Off-peak-period output thus is reduced based on the one-to-one output compensation because full market coverage applies across the whole range considered. In the balanced set, the network eventually achieves a complete hub dominance across both periods.

The comparison of the two graphs shows that this occurs at a particular value of δ, which depends on the relation of the operating and social costs to the density benefits. As pointed out previously, however, this value of δ is remarkably high. The asymmetric nature of the density benefits might lead one to deduce that the complete hub dominance should arise at a much lower relative importance of this network effect.

17.4.3 Flight Fares against Network Density Benefits

The flight fares that result from the socially optimal outputs with variations in δ for both the simple and the balanced set are depicted in Figure 17.3. The lines are denoted correspondingly but without upper- and lowercase notation of the indices. Note that the theoretical flight fare for peak-period leisure flights can be computed in the simulation based on the respective costs and benefits but is not applicable because peak-leisure output is zero.

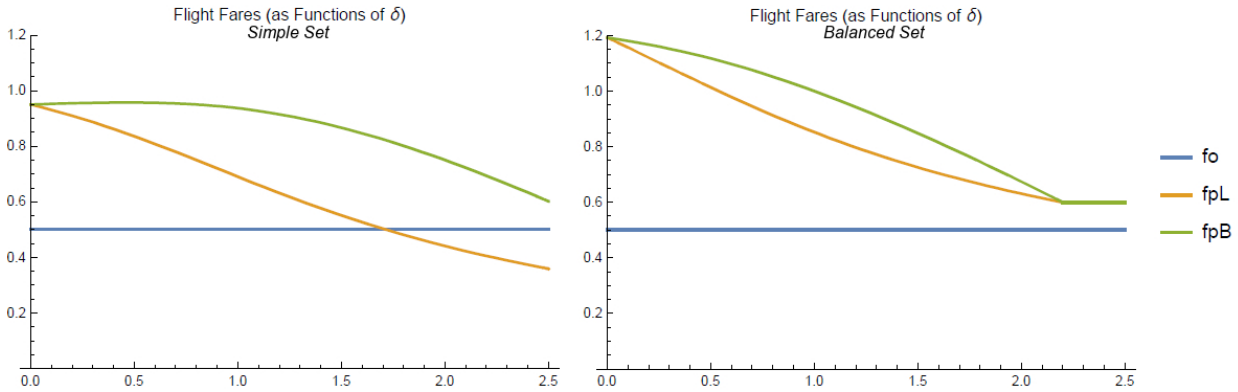

Fig. 17.3: Social Optimum - Flight Fares against Network Density Benefits

The graph shows that the off-peak-period flight fare remains constant at 0.5. This value corresponds to the marginal benefit of the left-most passenger. The revelation in Figure 17.2, where full market coverage occurs irrespective of the value for δ, hence explains that the flight fare must equal the lower boundary $\Theta - 1$ of the population range. Although the number of off-peak flights varies with δ (see above), the off-peak-period flight fare remains constant.

The peak-period flight fares of the business airline behave similarly across the two sets. Above all, they reflect the concavity of the network value: The larger δ is, the larger the network becomes but the lower the marginal willingness to pay of the critical peak-business passenger will be. Moreover, with a larger network the congestion increases. As already suspected above in numeric results section 17.3.2, flight fares hence generally decrease with a network expansion.

However, notice the particularity of the simple set in the graph on the left-hand side of Figure 17.3 where flight fares initially increase. Recalling from Figure 17.2 that peak-period output is quite low at the very onset of δ reveals that flight fares may initially *rise* with network size when the network size is small. While this property follows the concavity of the network value, it was not exhibited in the numeric set of solutions above. Moreover, it does not rise in the balanced set where initial output is higher. This lets us suggest that an initial upward move in the flight fares can only be observed when network benefits are low and congestion costs relatively high, and thus when the network is generally small.

As a second difference between the two sets, observe that the flight fare for a network flight flattens out toward the right side of the δ-scale in the balanced set (the right graph in Figure 17.3). As the network expansion in Figure 17.2 shows, however, this is easily explained by the fact that the network accounts for full market coverage when its corresponding network benefits are highly important. This means that output and congestion no longer change. Consequently, the flight fare also remains constant.

17.4.4 Airline Profits against Network Density Benefits

The most interesting graph in this sensitivity analysis is provided on the left-hand side of Figure 17.4, which depicts both airlines' profits as a function of parameter δ: In the simple set, the business airline's profits initially follow the concavity of the peak-period flight fare from Figure 17.3 but eventually become negative for large values of δ. This result is surprising given that the peak-period airfares of the critical peak-business passenger do not excessively decline. Nevertheless, it may be explained by the decreasing flight fares and the higher congestion in the optimum as compared to the inefficient network size in equilibrium. By contrast, Table 17.2 revealed that the leisure airline's profits are zero under the simple set. This result arises due to the interior solution for market coverage, which dictates that marginal benefits need to equal marginal costs in the off-peak period. Correspondingly, the graph shows that profits are nil across the full range of parameter δ.

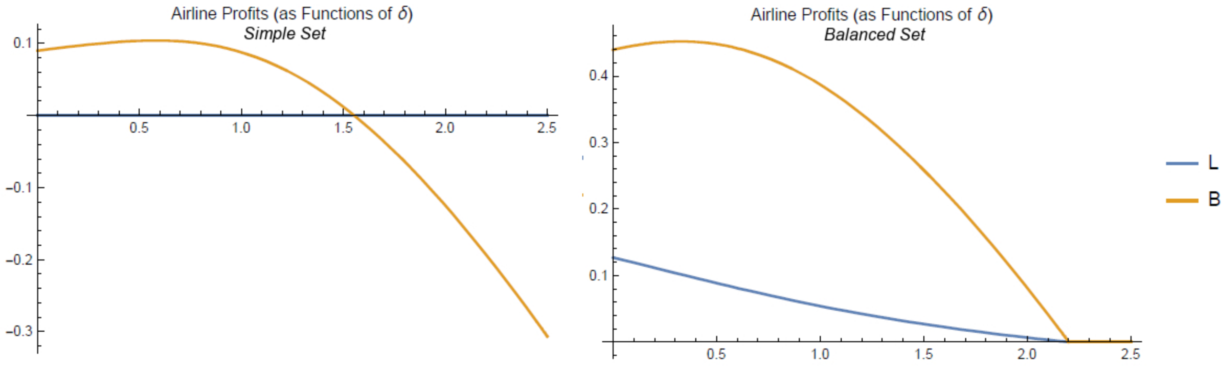

Fig. 17.4: Social Optimum - Airline Profits against Network Density Benefits

The large profit loss from the network expansion again illustrates the concavity of the network returns against the monotonicity of the density benefits. As a result, it is favorable to expand the network when network density benefits are important even if this actually infers considerable losses to the networking airline. This result illustrates the rationale for the network undersize from yet another viewpoint. However, the fact that increasing the network size is favorable from a welfare perspective but detrimental to the business airline's profitability reveals a distributional concern: As already revealed, a network expansion based on the higher density benefits might require to compensate the business airline for these losses if it could not be achieved based on strategic competition (such as, e.g., in the case of the reallocation rule). Although this result already occurred when the network underprovision was discovered, it still narrows the perspective for any output-raising regulation scheme.

The result for the balanced set is uncontroversial yet instructive: The functional forms generally correspond to the above result, with two distinctions. Firstly, the business airline's profits never become negative, and secondly, the leisure airline initially yields positive profits. The reason for these two differences may be found in the more favorable cost-benefit ratio of the balanced parameter set. In addition, for the leisure airline, the marginal flight benefits may exceed marginal operating costs in the corner solution for market coverage. Nevertheless, both airlines' profits still decline sharply with an increasing δ, which provokes the same regulation problem as above: the decreasing network returns, which indicate a natural market structure with both an undersized network and a low overall output.

Lastly, note that this graph dispels any apparent discontinuity in the airline profits that may have been expected from the numeric results in Table 17.2.

17.4.5 Net Utility against Network Density Benefits

The two graphs in Figure 17.5 illustrate net utility, which amounts to the difference of direct and indirect utility from a flight minus the respective flight fare. As in the numeric results section, the utilities in the graphs are integrated across all passengers within the respective period.

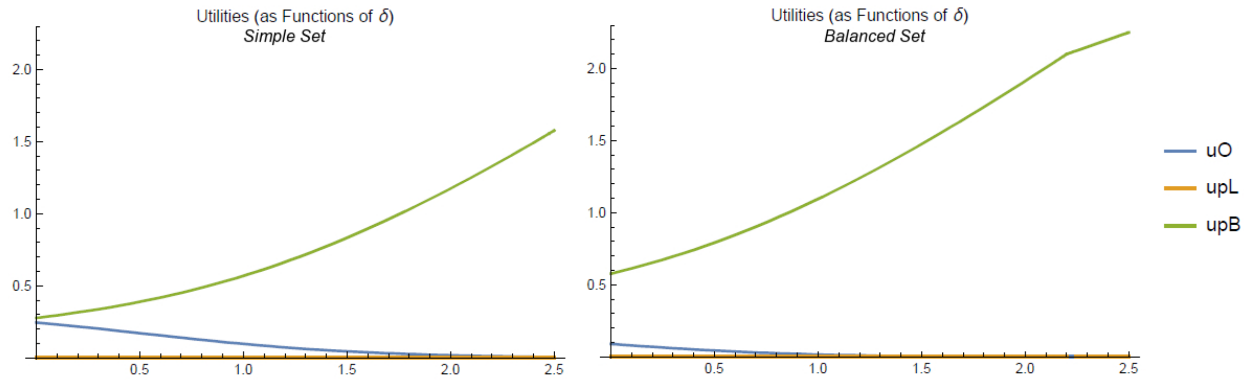

Fig. 17.5: Social Optimum - Net Utility against Network Density Benefits

Generally, net utility for peak-period passengers rises indirectly proportionately to the fall in airline profits. This follows from the facts that network size increases and flight fares decrease with an increase in δ. A higher number of network users, a larger network size and a lower flight fare for the use of that network hence need to yield a monotonously increasing, convex utility curve. Off-peak-period utility evolves in contrary fashion and remains at a very low level, just as the off-peak flight fares and outputs from above would suggest. Put differently, hence, the sharply rising utilities are the reason for the increasing network size in the social optimum, especially as on the downward side the airline profits may even become negative. Furthermore, note that the curve only becomes linear at the right-hand side extreme of the balanced set, where network expansion stops and thus all endogenous variables remain constant. Utility further increases linearly with the increase parameter δ while overall utility is higher in the balanced set in absolute terms. Again, these patterns reflect the different cost-benefit ratios.

17.4.6 Welfare against Network Density Benefits

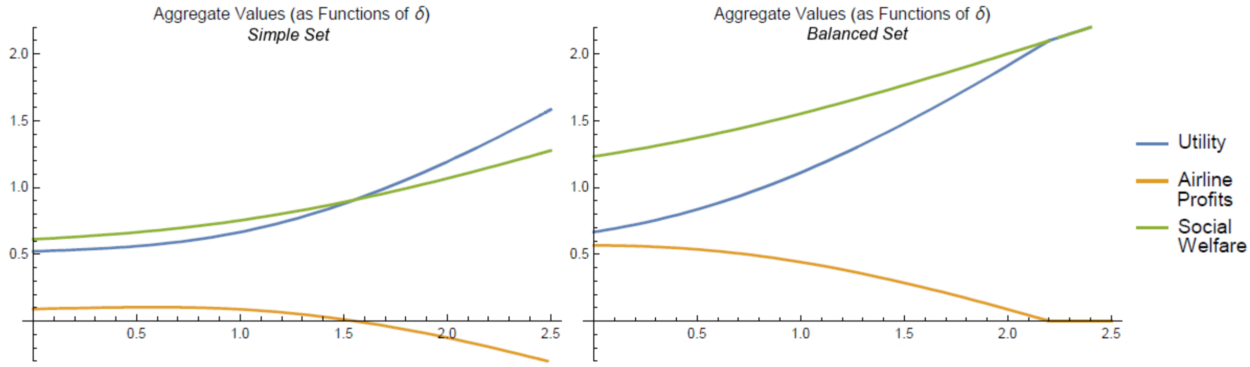

Fig. 17.6: Social Optimum - Welfare against Network Density Benefits

Finally, the graphs of welfare and its components are depicted in Figure 17.6. The two graphs are equivalent in nature and reflect the above findings concerning the original components of welfare: total airline profits and the overall sum of utilities. As discussed above, the increasing δ motivates a higher network size to increase both its benefits and the number of users. Because utility increases in a convex manner, it yields a net welfare gain although airline profits are generally low and dramatically decrease. As a result, the social optimum is straightforward from an economic perspective but controversial from a distributional point of view. This makes for an interesting investigation and discussion of potential regulation benefits.

18 Allocation Efficiency

The formal analysis of the generic model revealed that the equilibrium output allocation is concerned with two kinds of inefficiencies: the overall output distortion based on the market power effect and the network undersize arising from the concavity of the network value. As the size and direction of the overall output distortion depends on the importance of market power in relation to the congestion externality, overall peak-period output can either be too low or too high in equilibrium. By contrast, the network size always falls short of the social optimum.

The following results illustrate the network undersize, indicate that output is inefficiently low for most of the parameter range considered, and show both the equilibrium inefficiencies and the distributional effects of a potential regulation scheme that would shift the airport capacity allocation towards the social optimum. Finally, they allow us to consider some distributional aspects in addition to the usual efficiency analysis.

18.1 Numeric Results

The equilibrium inefficiencies can be numerically quantified based on a comparison of the computable model's results from Sections 16.2 and 17.3. Table 18.1 replicates these results and computes the gains and losses that arise to the different stakeholders in terms of airline profits, integrated net utility across all passengers, and social welfare. In addition, the numeric welfare gains of the social optimum in comparison to the equilibrium are indicated in percentages. For the comprehensiveness of the results, recall that the corner solution in the equilibrium implies that both airlines serve the peak-period while the business airline abstains from the off-peak period. By contrast, the corner solution in the social optimum only reflects the corresponding parameter change in δ while the leisure airline always remains expelled from the peak-period.

From an efficiency perspective, the numeric results in Table 18.1 indicate the following: First and foremost, the social optimum yields considerable welfare gains. While this result is a logical consequence of welfare maximization, the numeric computation allows us to quantify these gains in the last column in Table 18.1, which shows that social welfare increases by 14.5% to 31.8% if a fully efficient capacity allocation can be achieved. These distinct values arise from the different parameter sets and the types of solutions. They indicate that the extremes are based on the balanced set, where the interior solution yields the smallest and the corner solution yields the largest potential benefits. This result arises from parameters $\delta = 0.7$

Set	Sol.	Π_B	Π_L	$\int CV(\theta)$	$W(\theta)$	
		Equilibrium				
Simple	Int.	.223	.124	.280	.627	
	Cor.	.280	.107	.316	.703	
Balanced	Int.	.400	.287	.571	1.258	
	Cor.	.581	.234	.702	1.518	
		Social Optimum				
Simple	Int.	.088	0	.666	.754	
	Cor.	.012	0	.878	.890	
Balanced	Int.	.432	.074	.933	1.440	
	Cor.	.081	.007	1.913	2.001	
		Gains and Losses $(SO - EQ)$				$\triangle W[\%]$
Simple	Int.	-0.135	-0.124	0.386	0.127	20.2%
	Cor.	-0.268	-0.107	0.562	0.187	26.6%
Balanced	Int.	0.032	-0.213	0.362	0.182	14.5%
	Cor.	-0.500	-0.227	1.211	0.483	31.8%

Tab. 18.1: Inefficiencies in Equilibrium

and $\delta = 2$ in the balanced set, which denote the two endpoints concerning the importance of the network density effects (see Table 15.1). In this respect, the relative welfare benefit of the social optimum increases in proportion to the importance of the network density benefits. This signifies that the network density benefits increasingly diminish allocation efficiency.

Moreover, within each set, the relative welfare gain is higher for the corner solutions. Although the welfare levels also increase in absolute terms, this reveals that the equilibrium inefficiencies are more prominent in the corner solutions than in the interior solutions. This implies that a higher hub concentration of the networking airline in the unconstrained equilibrium increases welfare in absolute terms but decreases allocation efficiency relative to the social optimum.

The above result arises from the fact that the corner solution is invoked by a higher importance of the network density benefits, which also increases net utility of the passengers, whereas the corresponding social optimum is based on the same value of δ and reflects the elimination of market power and the network undersize inefficiency. As a consequence, the dilemma of hub concentration arising with higher network density benefits may be assessed to turn out beneficial in terms of absolute airline profits and net consumer value, but adverse in terms of welfare in relation to the social optimum and, thus, of allocation efficiency. This ultimately indicates that the network density effects increasingly aggravate the market power distortion in an unconstrained market solution.

In contrast to the above result, the hub dominance arising in the social optimum is efficient if regulation can reach this optimum by elimination of the network undersize and the market power distortion. This result holds although in the social optimum, where the leisure airline is completely expelled from the peak period, the network airline's hub dominance is actually higher than in the unconstrained equilibrium. Therefore, the dilemma of hub concentration arising from the asymmetric network density benefits is resolved to turn out beneficial if the socially optimal network size would be reached, but adverse if the hub dominance arises in equilibrium based on increased market power and thus a higher output inefficiency.

18.2 Sensitivity Analysis

The following sensitivity analysis evaluates the outcomes for different endogenous variables as functions of network parameter δ, ranging from 0 to δ_{max}. The investigation thus considers a symmetric airline structure that progressively becomes more asymmetric.

18.2.1 Individual and overall Outputs

The graphs in Figure 18.1 compare the peak-period outputs of the two airlines in the unconstrained equilibrium as well as the resulting total peak-period traffic to the socially optimal network size. As in the previous analysis, the left graph shows the results for the simple set, while the balanced set is represented on the right. Both graphs include the interior solution and the corner solution, where the threshold $n_p^B = 0$ is indicated by the separation line. Since the leisure airline is expelled from the peak-period in the optimum, the network size also represents the optimum peak-period output with regard to congestion and thus provides the single benchmark required for this comparison.

Above all, Figure 18.1 indicates two key findings: First, the network is always undersized in equilibrium. This outcome has already been derived in the formal analysis in Section 9.3. As explained above, this result arises from the concavity of the network density benefits and the corresponding concave network returns for the business airline, which contrast with the monotonicity of the network benefits from a social welfare perspective.

The second insight is that overall peak-period output in general is inefficiently low, except for some parameter values. In Section 9.3, it was argued that the dual distortion is expected to reduce overall output below the optimum because the traditional market power is likely

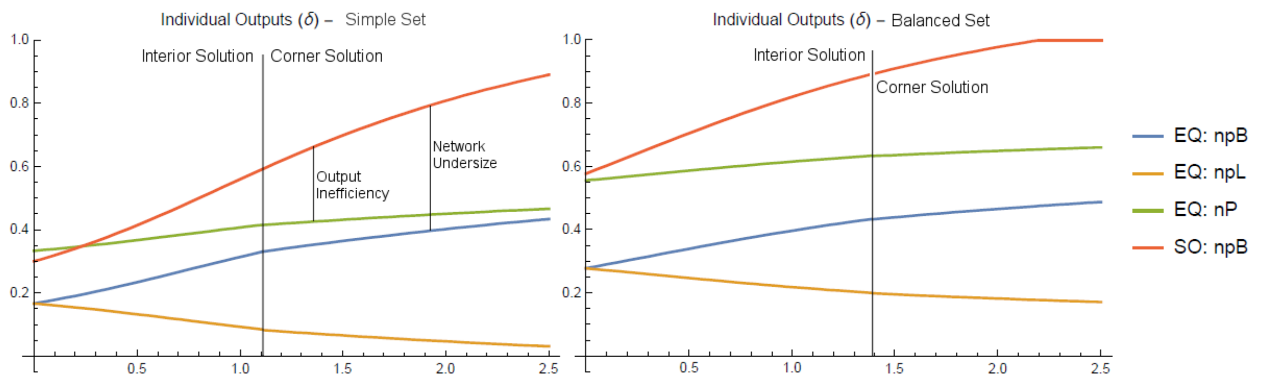

Fig. 18.1: Market Distortions - Output in Equilibrium

to exceed the congestion externality for the small number of firms in this model (i.e., the duopoly). This outlook is thus valid for almost the full parameter range but not for low values of δ in the simple set, where the cost-benefit ratio is higher than in the balanced set. Although this reversal may seem surprising at first sight, it simply reflects the fact that the size of the market power distortion depends on the importance of the direct flight benefits whereas the size of the congestion externality depends on the importance of time and congestion costs; these causalities appeared in the generic model's equilibrium conditions (9) and (10). In correspondence to the specifications from Table 15.1, the higher cost-benefit ratio in the simple set thus overturns the low output inefficiency into an excessive output for the case where the airline asymmetry is not prominent.

In addition to the above key results, also note that the relative network undersize increases with a higher importance of δ in both parameter sets. This tendency indicates that the network density effects aggravate the output distortion arising from market power, at least for the linear model within the considered parameter range. This effect will play an important role in the evaluation of allocation efficiency in the next subsections.

18.2.2 Airline Profits

Figure 18.2 compares both airlines' profits in the equilibrium against the social optimum. The airline profits in the social optimum are known from Figure 17.4: For the business airline, they are positive but declining and become negative in the simple set with increasing values of δ, which signifies that a network expansion is generally unprofitable. The leisure airline enjoys off-peak-period market access only, where flights are priced at marginal costs; its profits thus remain zero in the simple set and are positive but declining in the balanced set due to the more favorable cost-benefit ratio.

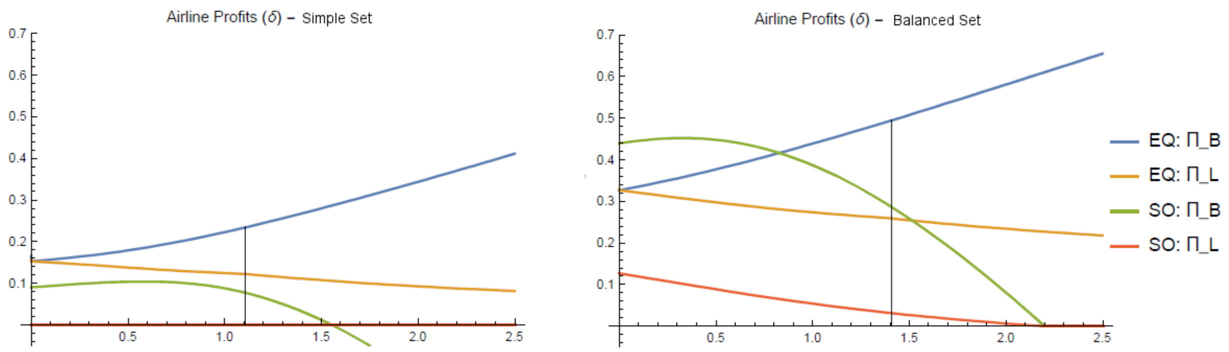

Fig. 18.2: Market Distortions - Airline Profits

The crucial result in Figure 18.2 arises from the comparison of the social optimum profits to the unconstrained equilibrium profits: The left-hand sides of both graphs reveal a small range of parameter δ where the business airline's social optimum profits exceed the corresponding equilibrium profits. This result is surprising at first but is explained by the business airline's exclusive peak-period market access, which allows the airline increasing its turnover without substantially decreasing its flight fares. This outcome indicates that the business airline exhibits an intrinsic motivation for at least a small network expansion (within the relevant parameter range) if it is guaranteed to enjoy exclusive market access. The numeric results in the following section indicate the quantitative magnitude of this motivation, while its implications for regulation policy are discussed in Section 20.5. With an higher importance of the density benefits, the excessive gains again vanish and the graphs reveal a regular pattern for the airline profits, where equilibrium profits are considerably higher than their socially optimal counterparts. From a distributional perspective, this indicates that implementing regulation policies without compensating the foregone profits may prove difficult in practice.

18.2.3 Passenger Utility and Welfare

Figure 18.3 illustrates integrated net utility across all passengers and welfare both in the unconstrained equilibrium and in the social optimum. The separation line indicates separates the interior equilibrium on the left from the corner solution with $n_p^B = 0$ on the right. The airline profits are not depicted in the graphs as they directly correspond to overall welfare minus passenger utility and can be seen in Figure 18.2.

In essence, the two graphs reveal to which extent the passengers may enjoy the rents arising from an increasing relative importance in the network density benefits, formally denoted by an increase in δ. While the network size increases along with the importance of its benefits (see in Figure 18.1), the result shows that both passenger utility and welfare monotonously

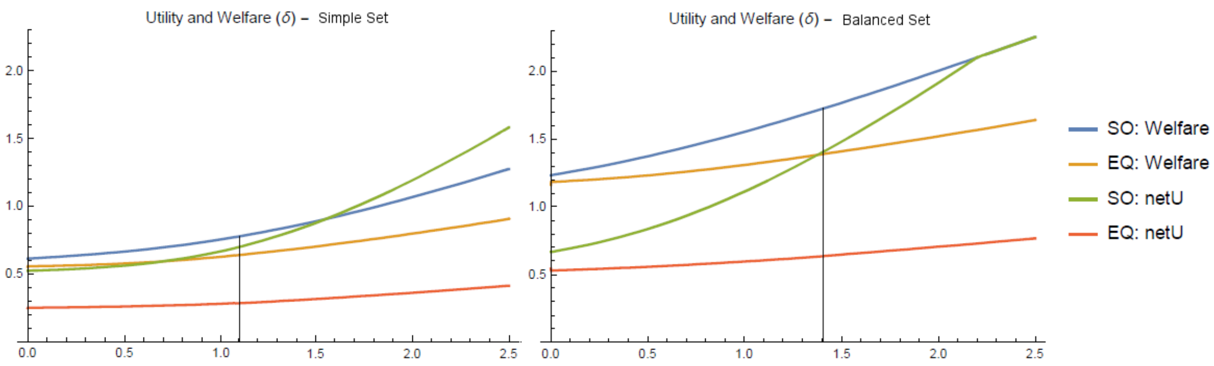

Fig. 18.3: Market Distortions - Passenger Utility and Welfare

also increase with more important network density benefits while also the business airline's profits increase (see Figure 18.2).

As a consequence, one might suggest that the dilemma of hub dominance were resolved for the case of this simulation: Considering an increase in the relative importance of the density benefits, the hub dominance of the business airline increases, which is also increasingly beneficial in terms of both overall social welfare and passenger utility. From a distributional perspective, the higher overall welfare would allow us to issue a monetary compensation to the leisure airline as the only loser. The above results, however, do not imply that the increasing hub dominance in the unconstrained equilibrium yields a higher allocation efficiency. In fact, the contrary is the case: As the overall output plots from Figure 18.1 reveal, with an increasing value of δ the relation of the equilibrium network size to the socially optimal network size progressively decreases. In other words, the inefficiency based on the network undersize increases, despite an increase in the network size in absolute terms. As a consequence, an increasing importance of the network density benefits yields a higher social welfare in absolute terms but an increasing loss in allocation efficiency against the social optimum.

This puzzling insight reveals the two contrasting sides of the dilemma of hub concentration: On the one hand, in absolute terms, an increasing hub dominance based on higher network density effects is socially beneficial even on an overall level. On the other hand, these social benefits increasingly vanish when compared to the socially optimal resource allocation. This ambivalence arises from an increase in the network undersize, which originates both in the traditional market power distortion as well as the concavity of the network value; it shows that both those effects dramatically increase with an increase in the importance of the network density effects, despite the expansion of the network size in absolute terms. This phenomenon is further explored in the next subsection.

18.2.4 Relative Welfare Loss

The relative welfare loss of the unconstrained equilibrium against the social optimum with increasing density benefits is depicted in Figure 18.4. The blue lines indicate the interior solutions, which are valid to the left of the separation line, whereas the amber lines indicate the corner solutions, which prevail on the right hand side.

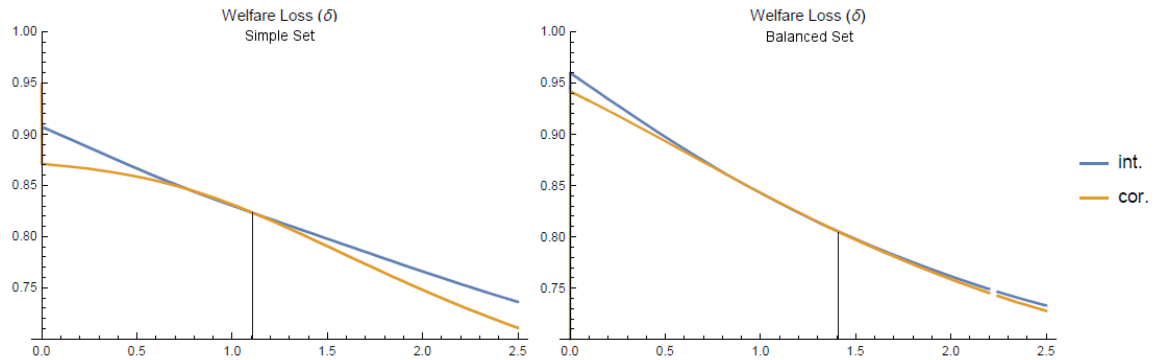

Fig. 18.4: Market Distortions - Relative Welfare Loss

As discussed above, the relative welfare loss strictly increases with δ. As indicated above, this effect reveals that the beneficial welfare effect within the ambiguity of hub concentration is only apparently beneficial: Although the hub concentration is beneficial both in terms of passenger benefits and overall welfare, it actually aggravates the allocation efficiency relative to the optimal market structure with an increase in δ. For the parameter range considered, allocation efficiency decreases from 91% (or 96%) in a homogenous duopoly, where $\delta = 0$, down to 71% (or 73%) for a very pronounced airline asymmetry with $\delta = 2.5$. The relative welfare loss arising from the relative network undersize thus amounts to roundabout 20 percentage points.

This result is particularly counter-intuitive because the overall welfare gain and the increasing passenger benefits disguise both the dramatic increase in market power and the concavity of the network value. One might refer to this puzzling effect as to the ambivalence of beneficial hub dominance, which may ultimately provide a more concise answer to the dilemma of hub concentration than the usual reference saying that the dilemma's outcome depends on the relative size of the effects. Put differently, if this result holds in a more generalized context than this simulation, then only the absolute welfare effect of increasing hub dominance depends on the net effect of higher market power against higher network benefits. By contrast, the impact of hub concentration on allocation efficiency is not ambiguous; rather, it increases with market power and a decreasing network value, and thus is increasingly deteriorating with increasingly important network density benefits. In retrospective, an important foundation for this result was revealed in Section 18.2.1, where the increasing network density benefits showed an aggravating effect on the market power distortion (see Figure 18.1).

18.3 Distributional Effects

The potential welfare gains from the social optimum that have been revealed in this analysis indicate that any regulation scheme that draws the resource allocation towards the optimal outputs should be worthwhile at least in second-best manner. Consequently, the efficiency gains might at least partly be used to finance such regulation. Depending on the accrual of the welfare effects to the distinct stakeholders, however, some participants may have to be compensated for their losses incurred if they are either desired or required to support the corresponding regulation scheme.

On this subject, a brief distributional analysis of the numeric results in Table 18.1 reveals the following insights: Under all four parameter sets, the passengers enjoy welfare gains form the social optimum in terms of net utility. By contrast, both airlines generally suffer from decreased profits when the market power distortion and the network undersize vanish. The only exception to the latter result is the interior solution of the balanced set, where the business airline yields higher profits from the exclusion of the leisure airline than in the unregulated duopoly market. This result is surprising to some extent, as it occurs despite the absence of market power and thus the excessive output from the perspective of the business airline. However, the graphs in Figure 18.2 confirm that this anomaly only arises with both relatively low network density benefits and low operating and congestion costs. In addition, note that this higher profit does not imply that the business airline should increase its output in equilibrium.

As mentioned previously, both airlines' losses need to be compensated in the case where their support is either necessary or desirable for the implementation of a suitable allocation instrument. The total welfare gain, in turn, proves that such compensations are generally affordable. In this respect, the higher net utilities suggest that it is appropriate to finance these compensations by the passengers. Depending on the allocation instrument used, however, the costs and benefits may shift away from the amounts assessed. The way to finance the costs of regulation can thus only be determined when the applicable regulation scheme is ultimately chosen.

19 Allocation Instruments

Although the analytic investigation of the generic model indicates clear implications on the application of the three different allocation instruments, the simulation results on allocation efficiency reveal some additional aspects that are worth considering. For this purpose, first both the administrative and the grandfathering allocation schemes for airport quotas are formalized and simulated with the same parameter sets as the equilibrium and the social optimum. This allows us to quantify the potential welfare caveats and efficiency gains for this linear model. Thereafter, the efficiency concerns of secondary trading and congestion pricing are addressed. In contrast to the quota solutions, those two instruments are not explicitly modeled, because their welfare potential can be assessed from the inefficiency results to a large extent.

19.1 Individual Quotas

The optimal market shares of an individual quota solution are determined by the social optimum. However, as already noted, the social optimum is not likely to be reached with a straight allocation of these quotas to the airlines: As the optimal quota rule dictates that the leisure airline is completely expelled from the peak period, the business airline would rather return to its monopoly peak-period output when facing exclusive market access without a service obligation. Consequently, the network would remain undersized.

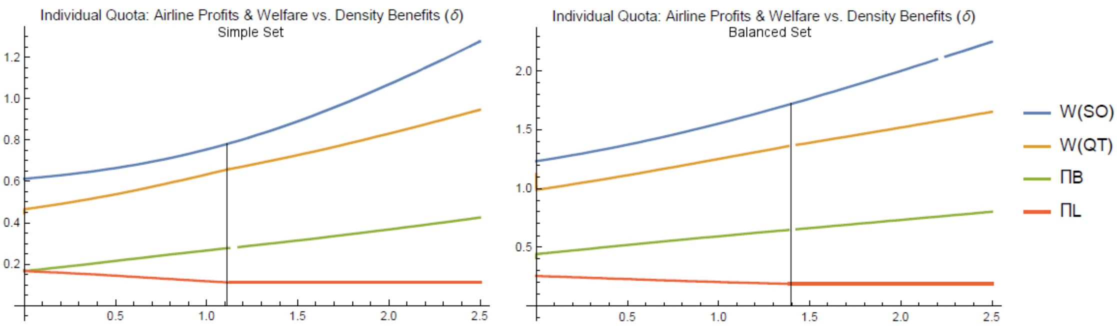

Fig. 19.1: Individual Quotas - Airline Profits and Welfare

Figure 19.1 shows the airline profits and the social welfare levels that arise with the individual quota allocation in the constrained equilibrium. As expected, the comparison of equilibrium welfare under the quota rule $\hat{q}_L = 0$, indicated as W(QT), and welfare in the social optimum, denoted as W(SO), shows that this allocation scheme is not fully efficient. In addition, a comparison of the airline profits with the equilibrium profits from Figure 18.2 shows that the business airline enjoys a higher profitability while the leisure airline yields slightly lower profits than in the unconstrained equilibrium. This result follows the respective market structure.

The efficiency outcome of the individual quotas is generally ambiguous because the unilateral capacity allocation based on $\hat{q}_L = 0$ endogenously diminishes the network undersize but at the same time leads to a decrease in overall output. While the generic analysis revealed that first-best efficiency cannot be reached with this instrument, the question of interest is whether the individual quotas may yield a second-best welfare improvement. Figure 19.2 therefore compares the overall welfare effect of the individual quotas, indicated by W(QT), to social welfare in the unconstrained equilibrium, W(EQ), and the social optimum, W(SO).

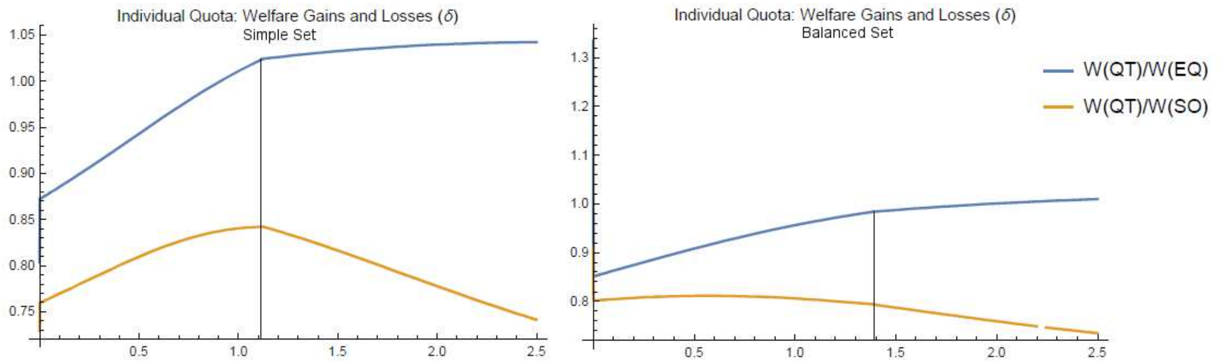

Fig. 19.2: Individual Quotas - Welfare Gains and Losses

The absolute degree of allocation efficiency confirms intuition: As compared to the social optimum, the individual quota allocation yields efficiency levels between 75% and 85% and exhibits a concave functional form that corresponds to that of the network value. Although these numbers may not seem exceedingly unattractive in second-best terms, the two graphs reveal an unexpectedly weak performance of this scheme as compared to the unconstrained equilibrium: The individual quota allocation is only able to exceed the equilibrium welfare under the simple set and for high values of δ. For low values of δ in the simple set, and for its entire value range in the balanced set, allocation efficiency actually decreases against the unconstrained market solution. While the efficiency gain amounts to a low single digit percentage, the adverse effect may decrease welfare up to 15% against the equilibrium. This extreme value is reached at the left hand side of the balanced set, where the airline asymmetry is about to vanish.

This result yields a more detrimental outcome than expected in the generic model analysis, indicating that the welfare caveat from an asymmetric quota solution may be substantial, in particular for a beneficial cost-benefit ratio of the peak period in relation to the off-peak period. By contrast, the potential for a positive welfare result is very weak - keeping in mind, however, that all numeric results crucially depend on the parameters and the linearity of the simulation model.

19.2 Arbitrary Constraints

The arbitrary constraints are modeled as a symmetric proportionate reduction of both airlines peak-period flight volume. The resulting peak-period output thus amounts to $q \cdot n_p^i$ where $q \in [0, 1]$ denotes the reduction relative to each airline's unconstrained equilibrium output n_p^i for $i \in \{B, L\}$.[94]

As an illustrative example, first the arbitrary case of $q = 0.8$ is considered in Figure 19.3. Because this arbitrary, symmetric constraint neither depicts the efficient market shares nor the optimal peak-period output, a negative welfare effect is introduced as per design of this scheme. In this respect, recall that this scheme is introduced to approximate a grandfathering allocation from practice, which is supposedly much more concerned with excessive overall outputs than this model with its low output inefficiency (see Section 11.5). Nevertheless, it is interesting to see that also this scheme seems to depict the concavity of the network value and that its inefficiency against the unconstrained equilibrium amounts to a similar magnitude as the individual quotas from above. Note, however, that this efficiency result is expected to reverse when the congestion externality becomes dominant. Unfortunately, this reversal cannot be shown in this model.

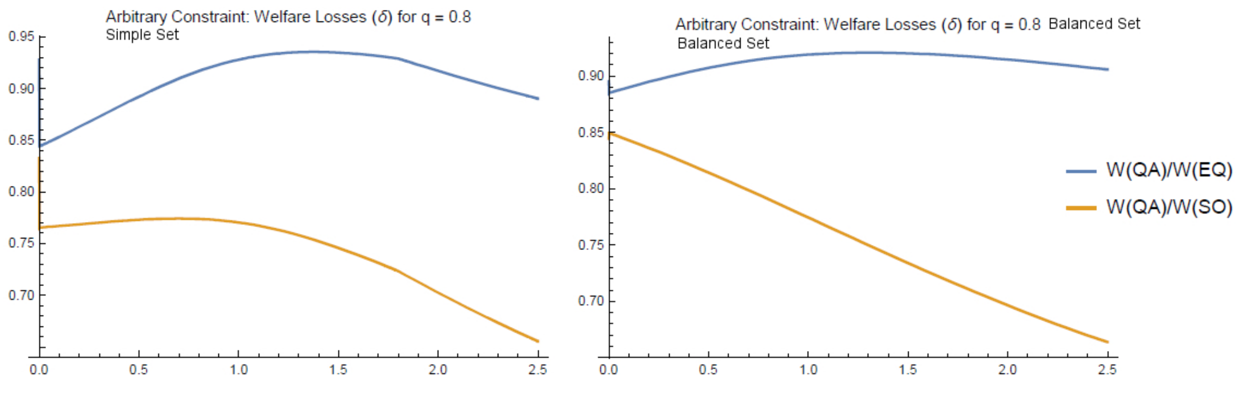

Fig. 19.3: Arbitrary Constraint - Welfare Losses for $q = 0.8$

Figure 19.4 depicts the absolute level of welfare that can be reached with the arbitrary quota as a function of both δ and q, which denotes the size of the constraint in proportion to the equilibrium outputs. The pattern is similar for both parameter sets: Welfare is

[94] For computational reasons, this q does not correspond to the formal arbitrary constraint $\bar{q}$ as introduced in Section 11.5, which denotes an absolute output quantity rather than a proportion.

highest when output is unrestricted (at $q = 1$) and when density benefits are important. Increasing the constraint size reduces welfare in an approximately linear manner whereas reducing the density benefits yields a slightly decreasing reduction of welfare. The plot thus depicts a welfare function that monotonously decreases in both arguments. If the congestion externality became important, however, these functions would become concave. In that case, the graphs would allow us to reveal a second-best solution and thus determine the optimum constraint size depending on the importance of δ.

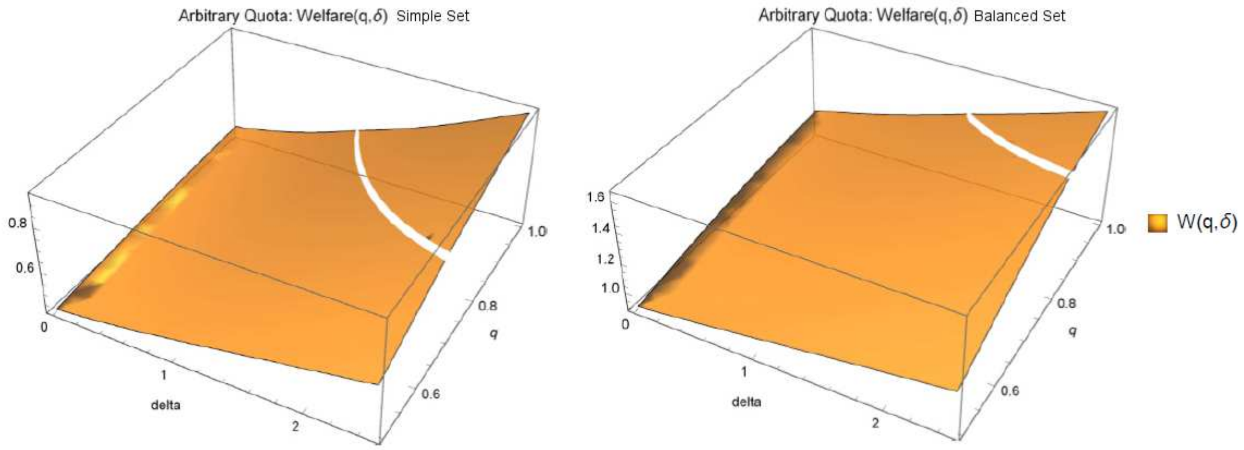

Fig. 19.4: Arbitrary Constraint - Welfare as a Function $W(q, \delta)$

The relative welfare loss in the particular case of this model can also be shown in a three-dimensional graph for the full range of δ. Figure 19.5 denotes the respective welfare levels W_QA(q,δ) for the arbitrary constraint and W_SO(δ) for the social optimum, where the optimum is not concerned with the size of the constraint. The two graphs are again similar for both parameter sets and indicate that the relative loss is minimized in absence of the airline asymmetry, where the network effects vanish and the products become homogenous.

The relative welfare loss monotonously increases both with an increase in δ and an increase in the size of the constraint. These relationships again confirm the particular characteristics of this model, where the former parameter amplifies the market power distortion and, hence, the network undersize whereas the latter increases the overall output inefficiency. As a consequence, these results indicate that an arbitrary constraint becomes increasingly inefficient with an increasing airline asymmetry based on network density benefits. In this respect, the results of the individual quotas may not have been compelling for a homogeneous market structure but in return seem to be more suitable if the airline asymmetry prevails.

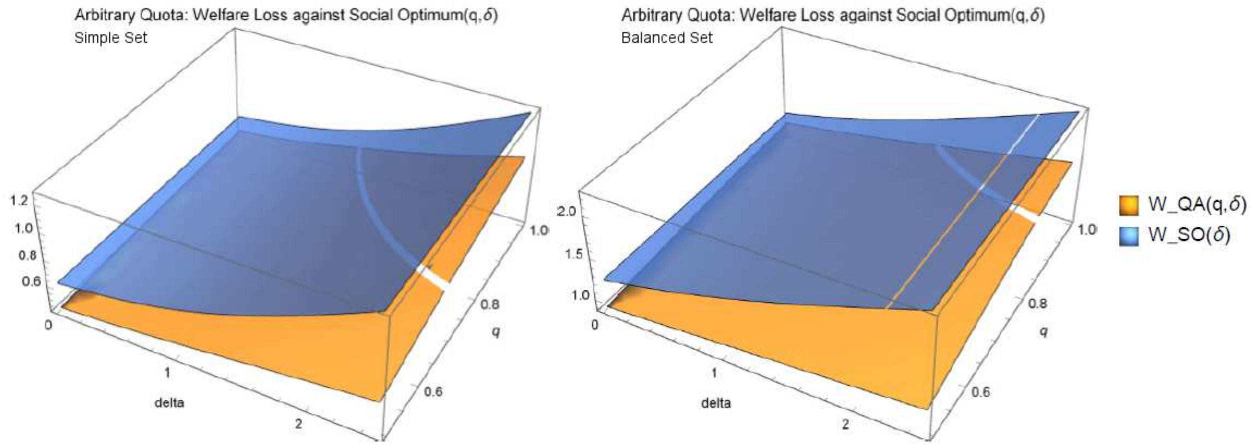

Fig. 19.5: Arbitrary Constraint - Relative Welfare Loss (as a Function of δ, q)

19.3 Secondary Trading

The market structures and the respective allocation efficiencies arising in all four trading cases considered can be inferred from the previous results.

In an individual quota allocation, only the business airline can be the seller. As shown in the formal analysis, however, it has no incentive to commercialize its quota based on a trade (see Section 12.4). In absence of the trading rule, the networking airline does not engage in slot babysitting. Therefore, the effective capacity allocation at the airport exactly equals the allocation under the individual quotas from above and thus exhibits the same efficiency properties as the quota allocation shown in Figure 19.2. In other words, in this case a potential quota market does not have any effect other than the individual quotas themselves.

By contrast, the trading rules were shown to introduce a first-best capacity allocation because they motivate the business airline to use all of its slots. As the number of quotas is assumed to be appropriately based on the social optimum outputs, this allocation exactly replicates the social optimum as investigated in Section 17. The efficiency gains that may be achieved correspond to the welfare deficit of the initial quota allocation with regard to the social optimum, as depicted in Figure 19.1. Thus, this case also does not yield any actual trading activities. However, the introduction of the associated trading rules directly affects the allocation efficiency.

A symmetric trading situation arises from the allocation of arbitrary constraints. By design of these constraints, both airlines are output restricted. In contrast to the asymmetric cases, however, trading does take place and yields that the business airline preempts the entire market. This result is based on the unequal markups that the two airlines can achieve.

In absence of the trading rules, the business airline may hoard some of the slots purchased. Because the business airline aims at producing its monopoly output, hoarding occurs depending on the overall size of the constraint. The allocation thus either replicates the individual quota allocation with the business airline's monopoly output or a lower network size if dictated so by the constraint. Any welfare benefit thus decreases both with a decrease in the overall size of the arbitrary constraint and when the opportunity of slot hoarding exists.

Correspondingly, Figure 19.3 can be used to see whether the welfare effect of trading may or may not be beneficial, and represents the highest available welfare gain. If either the overall number of quotas is lower than the business airline's monopoly output or if it is higher so that slots are hoarded, the welfare result is adversely affected.

Finally, in the symmetric case where the trading rules apply, trading also takes place but all quotas are utilized. The welfare effect thus corresponds to the above symmetric case but without the additional welfare caveat from strategic quota hoarding.

19.4 Congestion Pricing

As explored in Section 13.1, three distinct variations of a congestion tax may be considered: a conventional congestion tax, a tax which accounts for the market power distortion, and a pricing system that additionally incorporates the benefits of network density system and therefore reverts to a subsidy.

In view of both the output inefficiency and the network undersize, a conventional tax that compensates the congestion externality only will generally decrease allocation efficiency in this model because it further distorts the two outputs below the optimum. Only for the particular parameter range for which Figure 18.1 indicates a slightly excessive overall output a marginal welfare improvement may arise. Such an overall output reduction, however, always needs to overcompensate the network contraction for a welfare gain (see Section 13.3). In this model's special case, a positive effect might arrive because the output excess occurs where the network density benefits are relatively unimportant. Nevertheless, given the small output exceedance this effect may thus both be deemed diminutive and unlikely to occur.

By contrast, a tax specification that compensates both the congestion externality and the market power distortion increases both airline's outputs until the socially optimal overall peak-period output is reached. Figure 18.1 shows the corresponding output effect, which amounts to the difference between the equilibrium and the social optimum peak-period outputs. However, as this tax does not consider the network density benefits the market shares would remain similar as in equilibrium. Consequently, the network undersize would still apply and the social optimum were not replicated. The related welfare gain, if any, may therefore not be expected to be substantial. Moreover, the volume of the output expansion indicates that the subsidy required might be significant. According to the corresponding extensive financial losses from Figure 18.2 the subsidy is hence unlikely to be recoverable by the welfare gains. Due to the inefficient market shares the welfare gain does not correspond to the social optimum and thus is not revealed in the numerical results.

Again in contrast to the above two variants, a congestion tax that also covers the network density benefits yields the social optimum. However, in this case the leisure airline's applicable price for peak-period market access effectively prevents market entry, so that the airport does not yield any tax incidence. The subsidy required to compensate the business airline's service obligation, in turn, exactly equals the monetary equivalent of a switch from the equilibrium to the social optimum. This equivalent is generally negative but for one parameter set, as computed in Table 18.1 and illustrated in Figure 18.1. The special case where profits are higher in the social optimum thus creates an intrinsic motivation for an optimal network expansion whereas the leisure airline would still suffer an important loss from this switch. The question hence remains whether even in this particularly suitable case the network airline may be left enjoying the rent from its private market access whereas its competitor faces severe opportunity costs from market exclusion.

20 Discussion

This section briefly recaptures and discusses the simulation results from the parametric model for the equilibrium, the social optimum, the corresponding market inefficiencies, and the potential for the three different allocation instruments.

20.1 Equilibrium

Three main inferences can be drawn from the equilibrium results of the computable model: Firstly, the linear parametric model confirms the characteristics and results of the generic model as defined in Section 5 and discussed in Section 14. Secondly, the numeric equilibrium reveals an unexpected inference: the inverted correlation between the symmetric ratio of cost and the direct flight benefit functions, which is asymmetrically beneficial to the leisure airline. Thirdly, the comprehensive sensitivity analysis shows that the numeric results are generally stable within a reasonable parameter range, and further confirms that the properties of the network density benefits and the corresponding asymmetry correspond to their intended design within the generic model.

In addition, the sensitivity analysis reveals two distinctive findings that could not have been inferred from the analytic investigation: Most importantly, all individual outputs increase proportionally when the population wealth increases. This means that a higher willingness to pay across the entire population range does not only foster the business airline's network but also the residual flight supply. This result is quite surprising, as one may have expected that a higher propensity to consume would mainly increase demand for the high-quality product, which in this model are the peak-period business flights. Furthermore, a symmetric increase in the direct flight benefits, which corresponds to a variation of the cost-benefit ratio between the off-peak and the peak period and thus equally concerns both airlines, reduces the business airline's peak-period hub dominance while increasing the leisure airline's peak-period market share. This implies that the business airline's asymmetric hub dominance does not only depend on the asymmetric network density benefits, but also on parameters that are symmetric across airlines.

Finally, the simulation yields an interesting contrast to Brueckner's (2002a) model: Left aside congestion, the peak period is more profitable for both airlines. This result fundamentally differs from Brueckner's (2002a) horizontal product differentiation setting, where the demand heterogeneity is based on consumer taste, so that there are no commercial differences between the two periods, and the profitability is generally equal across periods for both airlines. As a consequence, the price of the congestion remedies in Brueckner (2002a) might have been underestimated in comparison to this parametric model with vertical product differentiation. Without further provisions, however, all above implications are only valid for the simulation of the parametric model. Therefore, it remains to be confirmed whether they generalize to different functional forms and a more generic case.

20.2 Social Optimum

The simulation of the social optimum confirms that the optimal allocation involves the network corner solution in the peak period. This result has already been derived in the formal analysis based on the monotonicity of the network density benefits. If only the peak period is considered, the corresponding increasing returns to scale thus dictate to provide a natural monopoly of the business airline for social optimality. However, as also already delineated, the business airline faces convex returns to scale in its profit maximization. As the next section will confirm, this indicates that the network is undersized and that overall output is inferior in equilibrium.

However, the network corner solution only concerns the peak period. This signifies that the off-peak period may still be served by either airline in the optimum. In this respect, the model does not define the optimal off-peak market shares between airlines because the off-peak-period flights are perfect substitutes. Nevertheless, this result implies that airport capacity regulation may allocate the spare off-peak capacity either to the business airline or to any other competing airline, as long as this residual off-peak-period supply is considered as homogenous.

20.3 Allocation Efficiency

20.3.1 Dual Distortion

As already supposed in the formal analysis, the simulation confirms that both overall equilibrium output and the network size are inefficiently low against the social optimum but for a small parameter range. While the network undersize has already been shown to unambiguously arise in the generic model, the final direction of the overall output inefficiency based on the dual distortion has not yet been revealed.

The numeric results show that the market power distortion exceeds the congestion externality for all considered parameters. However, the simulation reveals that a small region arises where the congestion externality exceeds the market power distortion, so that overall output is higher than in the optimum. Because the degree both of congestion internalization and of the output distortion depends on the number of firms, the above result seems to reflect an appropriate outcome for an airline duopoly. As a consequence, airport capacity regulation will generally be concerned with the welfare caveat from the low overall output inefficiency rather than with excessive overall outputs in settings that are similar to the one at hand.

20.3.2 Network Undersize

The simulation reveals a puzzling result concerning the dilemma of hub concentration: On the one hand, it resolves this dilemma because social welfare, airline profits and passenger utility are all shown to increase along with the hub concentration in absolute terms when the importance of the network density benefits increases. On the other hand, the increasing network output is opposed by a progressive increment of the socially optimal network size, which signifies that the network undersize is exacerbated with increasing network density benefits, so that allocation efficiency actually decreases in relative terms.

This counter-intuitive result shows that the equilibrium is substantially distorted both by the traditional market power effect and the decreasing returns from the network. The welfare benefits from hub dominance in absolute terms thus risk to disguise the relative adverse effect of hub concentration as compared to the social optimum. As a consequence, the potential of a regulation scheme that would increase the network size and reduce the market power distortion at the same time should be expected to be substantial.

20.3.3 Numeric Results

The numeric results allow us to quantify the inefficiencies arising in the equilibrium for distinct parameter values. Although the absolute magnitude of the welfare results largely depend on the specification of the functions and parameters in the computable model, they nevertheless show the following: The equilibrium inefficiencies caused both by the inferior output and the network undersize cause considerable welfare losses. The corresponding welfare benefits from the social optimum are proportionate to the importance of the network benefits, which signifies that the network density benefits increasingly deteriorate the inefficiency arising from the market power distortion.

20.3.4 A new Dilemma

As a consequence from the above results, the dilemma of hub concentration can be resolved: If the networking airline's hub dominance would result in resolving both the network undersize and the output inefficiency, a beneficial welfare effect could arise. However, in the case at hand, where the hub concentration arises from the exogenous airline asymmetry, it aggravates both the market power distortion and the network undersize in comparison with the respective optimum. Consequently, allocation efficiency decreases along with the network airline's increasing hub dominance.

This insight lets us conclude that the density effects arising from airline network operations increase welfare in absolute terms but negatively affect allocation efficiency in relative terms. This new dilemma must necessarily be considered within future discussions of airport capacity allocation.

20.4 Allocation Instruments

20.4.1 Individual Quotas

The welfare effect of an individual quota allocation in the simulation is less beneficial than expected from the reasoning based on the generic model: Although it had been clear previously that first-best allocation efficiency cannot be achieved, the quotas only yield a welfare gain against the equilibrium with the simple set and a high importance of the network density benefits. Even in this case the benefit only amounts to a few percentage points whereas all other parameter settings yield an adverse effect. This indicates that the welfare caveat from individual quotas should not be underestimated although those quotas aim to account for the asymmetry in the market structure.

20.4.2 Arbitrary Constraint

In contrast to the above finding, the adverse welfare effect of the arbitrary constraint arises by definition. Nevertheless, the simulation is able to show that the detrimental impact monotonously increases with more important network density benefits and with an increasing size of the constraint. The results on the two quota schemes thus imply that an asymmetric market structure with network density benefits requires individual quotas that account for this asymmetry whereas a more homogenous market may be better served with an arbitrary constraint.

20.4.3 Secondary Trading

The results for the secondary trading case recapture the results of the generic model, where in the case of individual quotas trading does not take place and thus does not affect the capacity allocation. However, the application of the trading rules, which implies a use obligation, leads to a first-best allocation that provides substantial welfare gains in this model specification. By contrast, actual trading occurs from a symmetric imposition of an arbitrary constraint. The corresponding welfare effect is beneficial in a second-best manner in absence of strategic behavior whereas it becomes ambiguous if slot hoarding is observed. The absolute size of the welfare effect largely depends on the actual size of the arbitrary constraint.

20.4.4 Congestion Pricing

For a congestion pricing scheme the simulation again indicates the result obtained both from the generic model and from the literature: A potentially detrimental welfare effect originally arising from the market power distortion can only be avoided with certainty if the tax compensates for both this distortion and for the network undersize. In this case, however, the tax becomes a subsidy except for the particular case where the networking airline's profits from an exclusive market access already overcompensate the network expansion. Although the corresponding losses of the leisure airline may invoke distributional concerns this result may provide a starting point for the design of an optimal capacity regulation scheme.

20.5 Distributional Considerations

The distributional analysis based on the numeric results illustrates that the airlines would generally oppose any regulation policy aiming at a resource reallocation towards the social optimum. This implies that any service obligation required to expand outputs or, more concisely, the business airline's network size would necessarily have to be accompanied by a subsidy. By contrast, a compensation for the leisure airline's output reduction is only required if so desired by distributional concerns.

These implications are valid but for the exception where the business airline actually profits from a shift toward the social optimum. Consequently, this exception offers an interesting option from a game-theoretic perspective: If the business airline could choose between an exclusive peak-period market access at the cost of a network expansion and the unregulated duopoly equilibrium, it would choose the exclusive market access. The particularity of this exception hence is that the corresponding profits from this regulated monopoly overcompensate the business airline for its generally unprofitable output expansion. As a consequence, the optimal network size could be implemented based on a service obligation which would come at no additional cost. Although this option only arises within one particular parameter set in this simulation, it still represents an interesting avenue for further investigations of regulation policy in asymmetric network markets with dominant firms.

Conclusions, Limitations and Outlook

This chapter concludes the model analysis from Parts I to III. It first briefly reviews the study design, the model and its results, which serves as the basis for the subsequent presentation of the conclusions in a traceable and transparent manner. Thereafter, the conclusions drawn from the model results are discussed in terms of implications for practice. Finally, I discuss the model's main limitations and provide suggestions for overcoming those limitations. These suggestions thus offer directions for further research.

21 Summary

The summary briefly reviews the model and its results in order to provide a sound traceability of the subsequent conclusions. It is divided into three separate subsections: The design of the study, the general properties of the asymmetric model, and the welfare impact of the three airport capacity allocation schemes.

21.1 Study Design

This study investigates capacity allocation at an airport with a network airline and a non-networking competitor in the case where outputs are distorted by market power and by congestion externalities. The airline asymmetry is based on the notion that the network airline provides additional network density benefits for its passengers, while the other airline provides direct flight benefits from transportation only.

For this purpose, the study presents a theoretical model that reflects a corresponding airline duopoly. This model is based on Brueckner's (2002a) generic airport model but has been modified to account for the asymmetric density benefits. Thereafter, the three most prominent capacity allocation schemes from practice and from propositions in recent economic studies are applied to the model. These three allocation schemes consist of an initial allocation of individual quotas, a secondary trading scheme for the quotas, and a congestion pricing scheme. In order to enhance the practical relevance, a grandfathering allocation of an arbitrary constraint is considered in addition to the individual quotas.

This formal framework is subsequently used for the investigation of allocation efficiency under the three allocation instruments compared to the unconstrained equilibrium and the social optimum. In this respect, the investigation of the secondary trading scheme separately considers the two distinct initial quota allocations and, moreover, accounts for strategic airline behavior within the quota trading process. Lastly, the study presents a parametric linear version of the theoretical model. This linear model serves to illustrate the qualitative model properties both on numerical and graphical grounds. The illustration is achieved through a quantitative simulation of the linear model based on a set of reasonably chosen parameters. The simulation considers both the equilibrium as well as the social optimum, and presents a sensitivity analysis for the most interesting parameters.

In summary, the study provides three contributions to the recent discussion in the literature: Firstly, an instrumental contribution in terms of the modified model that accounts for the exogenous airline asymmetry and the network density benefits; secondly, the investigation of the model's basic properties and of the three allocation schemes in the light of the network density benefits; and lastly, the numeric simulation and graphical illustration of the modified model.

21.2 Generic Model

Above all, the analysis shows that the modified model illustrates the asymmetric market structure of an airline duopoly with a network airline that provides additional network density benefits for its passengers. As intended, the modified model is able to formally depict the central argument of this study: Endogenous market power based on vertical product differentiation, which creates an airline asymmetry based on network density benefits and thus incorporates a demand-side microfoundation for product heterogeneity. Product differentiation is vertical because the network density benefits ultimately reflect product quality. The generic model hence illustratively reflects the equilibrium and the social optimum of an airline duopoly at a hub airport with congestion and endogenous vertical product differentiation. As a result, the networking airline can be shown to enjoy a higher peak-period market share and higher overall profits than its non-networking competitor. This result corresponds to economic theory, where product differentiation decreases competition and increases market power. Consequently, the model provides a suitable formal framework for the subsequent investigation of airport capacity allocation.

In turn, the comparison of the unconstrained equilibrium to the social optimum shows that allocation efficiency depends on two concise arguments: Firstly, the size and the direction of the dual market distortion consisting of the congestion externality and market power, and secondly, the size of the network and the corresponding importance of the network density benefits. As the two market distortions have opposing effects on output, the unconstrained peak-period output may exceed or undershoot the socially optimal level in equilibrium, depending on which distortion prevails. As a special case, both distortions may cancel each other out, so that overall peak-period output becomes equal to its socially optimal level. However, allocation efficiency furthermore requires the size of the business airline's network to be optimal. In this respect, the equilibrium deviates from the social optimum regardless of the overall output distortion.

This inefficiency occurs because the social planner and the profit maximizing airline have distinct target functions for the optimization of the network density benefits: From a social welfare perspective, the network density benefits monotonously increase welfare. Moreover, they are directly proportionate to the business airline's peak-period output. Consequently, the socially optimal allocation requires the leisure airline to be completely expelled from the peak period so that the business airline becomes the sole supplier of peak-period output. By contrast, the networking airline's profitability depends on its network value that represents the additional utility provided to the critical peak-period passenger by the network density benefits of the business airline. This network value exhibits a concave functional form. Therefore, the network size is always inefficiently low in equilibrium, and both airlines serve the peak period.

21.3 Allocation Instruments

The previous analysis of the social optimum revealed the following prerequisites for allocation efficiency in the asymmetric model: The peak period must only be served by the business airline, and the corresponding network size needs to optimally balance the network density benefits against congestion and market power. Regulation theory thus dictates that an optimal capacity allocation scheme must replicate this natural market structure. Yet, the welfare analysis shows that first-best allocation efficiency cannot be reached by any of the tree distinct allocation schemes alone: Generally, the allocation instruments either provide second-best welfare improvements or they may also induce adverse welfare effects. Only the combination of the quota trading rules with an asymmetric initial allocation of individual quota yields first-best allocation efficiency.

21.3.1 Airport Quotas

The investigation of airport quotas considers two types of quotas: Individual quotas based on the socially optimal market shares that lead to an asymmetric restriction of the leisure airline only, and an arbitrary constraint that aims to reflect the current grandfathering allocation from practice. The latter symmetrically restricts both airlines by definition but cannot be justified from a welfare perspective in this model. It is included in order to provide a starting point for symmetric slot trading.

The welfare effect of the individual quotas is generally ambiguous: Above all, the imposition of asymmetric, individual quotas achieves the goal that only the business airline serves the peak period. However, because the latter becomes the only supplier, it reverts to its monopoly output during the peak period. As a result, the network size increases but remains inefficiently small, so that the business airline remains factually unconstrained. The individual quotas thus fail to replicate the social optimum.

Consequently, the impact of the quotas on allocation efficiency depends on whether the dual distortion is increased or decreased in absolute magnitude, and whether the increased network density overcompensates any adverse output effect: If overall equilibrium output was excessive prior to the imposition of the quota, the welfare effect of the dual distortion is ambiguous. The compensation from the net positive network effect hence may or may not be essential for a welfare improvement. In the opposing case where overall output has been inefficiently low, the dual distortion is always increased by the quotas. In that case, allocation efficiency can only improve if the higher network density benefits overcompensate this overall output inefficiency. The individual quotas hence yield a second-best welfare gain if either the dual distortion is decreased or if the latter is increased but is overcompensated by the higher network density. However, if the dual distortion is increased and at the same time this deterioration prevails over the benefit of a higher network density, a net positive effect cannot be achieved and the individual quota scheme yields an adverse welfare effect.

However, any adverse welfare effect that arises may be overcome by deviating from the optimal quota rule: The undersized network yields that some unused quota can be allocated to the leisure airline. This increases overall output and thus again diminishes the dual distortion. As the latter previously could not be overcompensated by the higher network density, allocation efficiency again rises. Depending on the number of re-allocated quotas and on the relative sizes of the effects, this deviation from the optimal quota rule may or may not overturn the above adverse welfare effect into a second-best welfare improvement. Note, however, that this overall output effect will only occur if the second-best number of quotas is allocated to the leisure airline in the initial allocation. If, by contrast, the business airline is confronted with a hand-back and re-allocation scenario when it already owns the entity of all quotas, it will engage in slot babysitting rather than allowing for a quota re-allocation to its competitor. This result has been shown in the investigation of the secondary trading scheme. As a consequence, any quotas that are foreseen for a second-best improvement in the above sense should be made available prior to the initial quota allocation.

An individual quota scheme may hence be favorable if overall output does not undershoot its socially optimal value to a large extent and if the network density benefits are important relative to the market power distortion. In the contrasting case where output is inefficiently low and the network effect is weak, an individual quota scheme invokes an adverse welfare effect. Consequently, the risk of this adverse welfare effect decreases, the higher overall output is and the more important the network density benefits are. In addition, a negative impact may be overcome by deviating from the optimal quota rule and allocating a part of the unused quotas to the leisure airline. This would again decrease the dual distortion and hence increase allocation efficiency. Generally, the outcome in this generic model is undetermined. However, the duopoly leads us to the suspicion that the market power distortion in equilibrium has been large and output inefficiently low. In this asymmetric model, a beneficial welfare result may therefore only be expected if the network density benefits function is relatively important.

The investigation of the grandfathering allocation has shown that overall output is reduced by an arbitrary constraint, so that congestion decreases and market power increases. While the welfare impact of this output contraction generally depends on the initial size of the dual distortion, the market structure in the asymmetric airline duopoly suggests an adverse effect. In addition, the constraint also reduces the network size, which further decreases welfare. However, the arbitrary constraint cannot be justified on efficiency grounds within this model. Therefore, its output effect is assumed to eliminate an excessive level of congestion and thus to be beneficial. This changes the generally negative welfare contribution of the arbitrary constraints into an ambiguous effect. Consequently, a second-best welfare improvement is achieved if the assumed decrement of the dual distortion overcompensates the lower network density benefits. In the contrasting case where the diminution of the network density is important, the grandfathering allocation adversely affect allocation efficiency.

As already stated, this model cannot replicate the grandfathering allocation on efficiency grounds. This consideration thus serves as a starting point for the secondary trading scheme rather than as an investigation of an arbitrary constraint as a stand-alone allocation instrument. Correspondingly, the above result needs to be interpreted with care. Nevertheless, in light of the airline asymmetry, also a grandfathering allocation of arbitrary constraints generally yields an ambiguous welfare result.

21.3.2 Secondary Trading

The investigation of the secondary trading scheme first considers the asymmetric initial allocation of the individual quotas as justified by the social optimum. In order to enhance the analysis, a symmetric trading case is added that is based on the above stylized grandfathering allocation of an arbitrary constraint. The two trading cases are correspondingly referred to as asymmetric and symmetric. Both settings consider quota trading with and without the application of the trading rules, which were originally designed to avoid adverse welfare effects from strategic airline behavior.

In an asymmetric initial allocation, a quota trade would only occur from the business to the leisure airline because the latter does not dispose of any slots. In addition, it would only involve previously unused slots. Without further provisions, secondary trading would therefore increase overall output and reduce the network size. However, the negative network effect would only be marginal, so that the output effect is likely to be dominant. Quota trading hence would presumably be beneficial from a welfare perspective. However, the airlines' trading potentials yield that an actual trade never occurs. As delineated above, the business airline hence maintains its monopoly output. In-line with the welfare effect of the initial quota allocation, allocation efficiency thus either remains inferior or superior to the unconstrained equilibrium.

Both the welfare impact of secondary trading and the resulting allocation efficiency substantially differ against the above case when the trading rules are implemented: Because the business airline is forced to either use its allocated quotas or to hand them back for reallocation, all quotas are always utilized and a quota trade has no impact on overall output. Nevertheless, the business airline remains the only potential seller, so that trading would reduce the network size and hence adversely affect welfare. However, the excess number of quotas would never be traded because the business airline is best-off with inefficiently using its excess number of quotas by increasing its own peak-period output. As a result, both overall output and the network size expand to their optimal volumes.

The imposition of the trading rules in the asymmetric case thus yields a counter-intuitive result: On the one hand, the rules cause the business airline to engage in strategic behavior in the form of slot baby-sitting. On the other hand, this kind of strategic behavior leads to the replication of the social optimum. Consequently, it is the application of the trading rules rather than the secondary trading itself that yields first-best allocation efficiency.

From a regulator perspective, it is therefore optimal to provide an initial allocation of individual quota and thereafter impose the trading rules. The rules force the business airline to fully utilize all of its quotas and hence implement an efficient allocation of airport capacity. In this respect, the trading rules thus replicate a service level agreement based on the business airline's optimal number of quotas. As a result, the socially optimal allocation is best achieved by an initial allocation of individual quota and a corresponding use obligation. In particular, this does neither require nor involve a market place for secondary trading.

In a symmetric case, a trade may be effectuated in either direction because both airlines are allowed to have positive outputs. However, the business airline exhibits a higher willingness to pay than the leisure airline because the network premium yields an additional profit for every unit of its output. As a consequence, the business airline becomes the quota buyer, the leisure airline becomes the seller and trading occurs. Moreover, the network premium may decrease with the network expansion but will always persist. Therefore, the business airline will preempt the entire peak-period market.

Because both airlines are constrained, all traded quotas have already been utilized previously. If the trading rules are imposed, overall output hence remains constant because the business airline is compelled to use all quotas purchased. For the same reason, the network size increases. Yet, based on the determination of the number of quotas overall output by definition remains inefficiently low. In the symmetric case with the trading rules, hence, secondary quota trading takes place and improves allocation efficiency in second-best manner.

In the absence of the trading rules, however, a negative output effect may occur: The buyer may purchase some constraints for the purpose of subsequent slot hoarding. This signifies that he could reduce overall output and thus increase market power at the cost of a quota trade. As the business airline strives toward its monopoly output in the peak period, strategic slot hoarding will occur if the total number of constraints exceeds the business airline's monopoly output. If the overall number of slots is lower, by contrast, it is more profitable for the business airline to utilize all purchased quotas.

If quota hoarding occurs, it has two effects: Firstly, it decreases overall output and hence diminishes allocation efficiency. In addition, it either reduces or completely nullifies the network expansion, depending on whether all purchased quotas are hoarded or whether at least some of them are used. As long as the network effect at least partly emerges, it may overcompensate the negative output effect. In the contrasting case where all traded quotas are hoarded, allocation efficiency unambiguously decreases after a quota trade.

The post-trading type of strategic behavior in terms of quota hoarding hence induces a negative output effect and either reduces or completely suspends the beneficial network expansion.

As a result, the welfare effect of secondary trading depends on the relative size of the two effects and generally becomes ambiguous. If the strategic hoarding concerns all traded quotas, an adverse welfare effect is certain.

Because strategic slot hoarding only emerges if the total number of constraints is larger than the business airline's monopoly output, it need not necessarily occur. However, as an efficient output constraint is likely to be higher than one airline's exclusive provision of the entire market, this welfare caveat may reasonably be expected to actually arise. In conclusion, the secondary trading of a symmetric output constraint improves allocation efficiency in a second-best manner as long as the networking airline becomes the buyer and as long as it is imposed in conjunction with the trading rules. The trading rules again replicate a use obligation and thus suppress the welfare caveat from strategic behavior.

The above result shows that two kinds of strategic behavior occur in this model: In the asymmetric case, it concerns the potential seller and consists of quota babysitting, which causes previously unused quotas to be utilized. In the symmetric case, by contrast, it reflects strategic slot hoarding after a trade, which decreases the degree of slot utilization and reduces overall output and concerns the quota buyer. Surprisingly, the strategic airline behavior needs to be suppressed in the symmetric case, whereas it has to be introduced in the asymmetric case. However, the imposition of the trading rules satisfies both these opposing requirements: In the symmetric case it yields a second-best welfare improvement through the suppression of slot hoarding, and in the asymmetric case it yields a first-best allocation through the introduction of slot babysitting. While this relationship may be considered as a paradox in the asymmetric case, in the symmetric case the impact of the trading rules again becomes intuitive: They increase welfare by inhibiting rather than enabling the opportunity for strategic airline behavior.

Lastly, note that the above welfare results are inversely related to the potential distributional concerns about secondary trading: In the asymmetric case, a trade would allocate some quotas to the leisure airline and hence decrease the inequality of the two airlines' market shares. However, in the absence of the trading rules the welfare effect would only be second-best or might even become adverse. When the trading rules are in place, the rules induce a first-best allocation and the welfare effect of a trade would be detrimental. An equalization of the market shares would hence deteriorate allocation efficiency at least in the first-best solution. By contrast, the imposition of the trading rules invokes strategic slot baby-sitting and further reduces the trading potentials so that trading never occurs. As a consequence, it fosters the inequality of the market shares by increasing the dominant airline's market presence. The distributional concern about the imbalance of a market structure with the business airline as the single supplier is hence dispersed by this market structure's first-best allocation efficiency based on the service obligation.

21.3.3 Congestion Pricing

Congestion pricing removes the congestion externality by internalizing its damage to the airlines using an individual tax. To begin with, it induces the same welfare caveats in this asymmetric model with market power as shown by Brueckner (2002a): The removal of the congestion externality may yield that the remaining market power distortion decreases overall output far below the optimum. In this case, the output deviation from the optimum becomes larger than the initial dual distortion, so that allocation efficiency decreases. This may even occur if the unconstrained overall output previously exceeded its efficient level but the market power distortion has been important relative to the congestion externality.

In addition to this original result, however, in this model the network density effects further increase the risk of an adverse welfare effect: In any case, the tax decreases both airlines' outputs and hence reduces the network benefits for the passengers. As the duopoly suggests that market power is important relative to the congestion externality and as the network density benefits monotonously decrease in proportion to the network size, the welfare effect of congestion pricing in this model is hence very likely to become adverse.

Although both above distortions might be corrected in the computation of the tax, such a remedy would become controversial essentially for two reasons: Firstly, congestion pricing would become both an allocation instrument and an anti-competitive regulation scheme at the same time. This would be critical from a policy point of view. Secondly, the correction would return the tax into a subsidy. However, the same allocation could be replicated free of charge by the previous quota scheme in conjunction with the trading rules.

Besides these two arguments, the computation of the tax and its appropriate implementation in a comprehensible pricing system would face major practical concerns. As a consequence, in light of both network density effects and market power and in view of its practical complications a congestion pricing scheme seems not to be the preferable means for airport capacity allocation.

22 Conclusions

The analysis of the generic model shows that the equilibrium is inefficient due to the underprovision of network services even if the dual output distortion from market power and congestion externalities vanishes. The conclusion from this result is that regulation policy needs not only to account for the market power distortion but also for the asymmetric network effects with its two distinct underlying rationales, consisting of the monotonicity of the network density benefits and the concavity of the network value.

In consequence, the welfare results of the three capacity allocation schemes differ both from the case of perfect competition as well as from a setting with traditional market power but homogenous products. In the introduction, this implication was anticipated and therefore presented as the motivation for this study. The asymmetric model now provides both a formal justification for this reasoning as well as an appropriate foundation for the subsequent investigation of the three capacity allocation instruments.

22.1 Allocation Instrument Choice

The general problem of all three allocation schemes under investigation is that they are unable to replicate the efficient market structure but suffer from a potential welfare caveat based on the market power distortion in conjunction with an undersized network. This concerns both the classical quota scheme with individual quotas or an arbitrary constraint, as well as the proposed alternative instruments of secondary trading and congestion pricing. Consequently, airport capacity regulation may adversely affect welfare if it increases either the output inefficiency, or the network underprovision, or both.

The welfare results of this investigation show that a naive allocation of individual quota without any further provisions cannot replicate the efficient market shares because both overall output and the network size remain inefficiently low. Nevertheless, this scheme may yield a second-best welfare improvement in the case where congestion is excessively large and the network density benefits important. If this relationship is reversed, however, it may even provoke an adverse welfare effect.

Similarly, a secondary trading scheme induces ambiguous results on allocation efficiency: For the case of individual quotas, secondary trading may only increase welfare if the trading rules are not in place; otherwise, a quota trade would provide a welfare deterioration. For either case, however, trading is shown not to take place at all, so that a secondary trading market itself does not yield any effect. By contrast, the imposition of the trading rules induces a first-best capacity allocation in its own right, despite the absence of actual trading. This crucial result arises because the trading rules force the business airline to provide its optimal network size. As this benefit does not arise from actual trading, however, it may not be appraised as a benefit thereof. For the case of a symmetric, arbitrary constraint, secondary trading becomes unambiguously beneficial as long as the trading rules are imposed; in the absence of the trading rules, the welfare effect again becomes ambiguous. It should be kept in mind, however, that the symmetric constraint is not endogenously justifiable in the model at hand but only replicates a presumed grandfathering allocation.

Lastly, a congestion pricing scheme would decrease both airlines' outputs in any case. As the network size always remained inefficiently low, the tax would yield both an ambiguous output effect and a negative network effect. Depending on the initial size of the dual distortion, the welfare result would hence either become second-best or adverse; in light of the low number of firms (i.e., the duopoly) and the presumable importance of the network density benefits, a negative effect on allocation efficiency seems even more likely. If the pricing scheme were designed to account for its two inefficiencies, it would replicate a first-best allocation but become a net subsidy instead of an actual tax. As a such, however, it could be replicated free of charge by an individual quota scheme in conjunction with the trading rules.

In comparison to a homogenous product setting, the consideration of network density benefits thus increases the risk of the welfare caveat for all instruments. In the context of large, dominant network airlines at their hub airports, the potential of airport capacity allocation must hence be expected to be severely limited while containing a serious welfare caveat. The only case where first-best allocation efficiency can be reached in theory is an administrative allocation of individual quotas in conjunction with a use obligation in the sense of a hand-back and re-allocation rule, as presented in the investigation of the secondary trading scheme. As such a scheme in effect closely resembles the current administrative quota allocation from practice, the practical implications of this result are discussed below.

22.2 Reference to previous Research

The efficiency results from recent literature for both proposed alternative instruments are not supported by the asymmetric model at hand, because both the secondary quota trading and the congestion pricing scheme either yield second-best results or adverse welfare effects only, while a first-best result cannot be achieved. Moreover, the investigation of the quota scheme is partly novel, as recent airport models have rarely considered airport slots in a formal manner at all. In this respect, the results of this study may justifiably be deemed not only to replicate, but also extend and challenge the established insights concerning the discussion about airport capacity allocation.

The fact that airport quotas are optimal in a symmetric setting with homogenous airlines and in absence of market power is indicated, e.g., by Forsyth and Niemeier (2008, p.67). Also, these authors stress that allocation efficiency may be difficult to reach if multiple opposing market distortions occur (idem, p.68-69). The corresponding results for the arbitrary constraint in the study at hand thus exactly illustrate this prediction. In addition, the investigation of the individual quotas extends this knowledge by the two essential properties required for an efficient allocation in an asymmetric market structure with congestion, market power and product differentiation: first, the trading rules that replicate a use obligation, and second, the impact of the network density benefits on the optimal number of quotas.

For the congestion pricing scheme, the above ambiguous outcome generally follows Brueckner's (2002a) findings but, in addition, reflects the impact of the network density effects. The consequence in terms of a corrected tax corresponds to Verhoef's (2010, p.323) idea, where first-best efficiency is reached but the tax reverts to a subsidy and hence no longer represents a congestion pricing scheme. Consequently, the study at hand shows that the network density benefits have a similar effect as market power by causing network underprovision, hence increasing the existing welfare caveat.

In the case of secondary trading, the symmetric trading case with an arbitrary constraint replicates Verhoef's (2010) result, where quota trading is beneficial in the absence of strategic behavior but becomes ambiguous when the purchased slots can be hoarded. Moreover, the networking airline replicates the market preemption of Verhoef's more cost-efficient airline and hence reproduces the apparently controversial effect of quota trading on hub concentration. The asymmetric setting with the individual quotas, by contrast, differs from the symmetric case, as secondary trading may only be beneficial in the absence of the trading rules; when the trading rules are imposed, a quota trade would unambiguously decrease allocation efficiency. Such a setting, however, has not been investigated in the recent capacity allocation models and thus lacks comparison. In contrast, the discovery that the trading rules replicate a service obligation, which in conjunction with asymmetric quotas introduce a first-best resource allocation resembles Verhoef (2010), who also applies a use obligation in order to suppress post-trading strategic behavior. In his case, however, the quotas are symmetric and the quota trading process itself increases welfare. The specific insight gained from the model at hand is thus that a use obligation may introduce allocation efficiency in itself, while actual quota trading in the same setting would deteriorate welfare but would never occur.

In consequence, the investigation at hand confirms that a compensation of the two market distortions is not easily achieved when network density benefits and market power are present and interdependent. This insight contrasts with previous perfect competition settings or models featuring inelastic demand, where both alternative allocation schemes are equally efficient. Also, it differs from oligopoly models that incorporate market power but consider flights as homogenous products, where allocation efficiency is only a function of the dual output distortion. Ultimately, the distinction of this study relies on the innovation that its investigation involves the ambiguities arising from the monotonous network density benefits and the concavity of the network value. Thus, as a unique contribution, this study formally shows in which respect the presence of asymmetric network density effects complicates the airport capacity allocation, and how these problems may be overcome.

22.3 Practical Implications

From the theoretical perspective of an asymmetric airport model featuring network density benefits, the above results indicate that airport capacity can be efficiently allocated by the implementation of an individual quota scheme in conjunction with a set of trading rules that replicate a use obligation. Surprisingly, this proposition appears to closely resemble the administrative quota scheme, which is currently in use in practice: Firstly, the current administrative allocation includes a use obligation (in terms of the 80-20-Rule), which in essence equals the hand-back and re-allocation rules. Secondly, the associated grandfathering rights from practice supposedly foster the incumbent network structures and limit the residual supply of new entrant competitors. Although this allocation is severely criticized in the literature and may not exactly replicate the theoretically optimal market shares, it may thus reasonably be presumed to support a market structure that is at least similar to this model's social optimum. The investigation at hand may thus impose an important counterweight to the inequality and inefficiency allegations against the current allocation scheme from recent literature.

As a consequence, if the administrative slot allocation in practice replicates the optimal asymmetric market structure at least to an important degree, it might be deemed superior to the propositions of a secondary trading market and congestion pricing as it might, even inadvertently, take better account of both the output inefficiency and the network density benefits. Although this rationale would probably play an implicit rather than an explicit role in the allocation process, its result might bear a striking similarity to the theoretical first-best allocation of individual airport quota. Also, recall that the theoretical social optimum only requires an exclusive network access during the peak period. This market structure thus enables the incorporation of any non-networking residual supply in the off-peak period, as it seems to be the case with new entrant competitors in practice. By contrast, in the presumed presence of market power and network density effects, both a secondary trading scheme and a congestion tax are not likely to improve allocation efficiency, but may rather induce an adverse welfare effect. In particular, this conclusion may cast a shadow on the recent policy propositions and European legislation changes aiming at promoting a secondary quota trading scheme.

Nevertheless, the administrative quota allocation scheme should still be considered to exhibit some important inadequacies. Most relevantly, the grandfathering allocation is a result of a mutual bargaining process among the airlines, where the airlines' cost-benefit rationales may be determined by market power considerations and concave network returns rather than monotonously increasing network density benefits. Although the use obligation might at least partly overcome any output inefficiencies based on market power, the outcome of this allocation process may still significantly differ from the optimal allocation of the individual quotas in this model. As a consequence, it is left to the coordinator's competency only to balance the airlines' requests against the socially optimal market structure. In this respect, future studies may contribute to the understanding of this potential inefficiency by investigating the airlines' rationales and strategic potentials within the slot coordination process.

Moreover, the above reasoning also leads us to the question of whether regulation schemes from other industry sectors with a similar market structure could be adapted to provide an efficient airport capacity allocation. One might think of network industries with asymmetric, dominant firms or of the extreme case of natural monopolies, where the key property for any adoption would consist of an optimal market structure with network density benefits, involving a high market share for a dominant networking supplier but also allowing other firms to satisfy the residual demand. If such regulation schemes were to permit the maintaining of the benefits of hub concentration while controlling for the adversities of market power in terms of undersupply and overpricing, they might allow the network density benefits to flourish and to unfold their full potential for the passengers at mostly undistorted flight fares. Seeking suggestions to this issue also represents a main topic for further research.

Finally, it is important to stress that the above results and conclusions are only valid within the context of this study with asymmetric network effects based on the exogenous airline asymmetry, monotonous network density benefits and a concave network value. Moreover, they crucially depend on the dual market distortion, where the effects of traditional market power and the congestion externality arise and are opposed to each other, and where each airline internalizes its own portion of flight delays. These limitations are discussed in the last section.

23 Limitations & Outlook

This section presents the most important limitations to the model at hand. These limitations involve both the conceptual properties consequential to the chosen framework and the technical particularities arising from the model design. These properties are either introduced intentionally and are thus based on the essential model assumptions or constitute simplifications deemed to be justified inaccuracies for the benefit of analytical traceability and comprehensibility. As both these properties and characteristics crucially affect the results of the model, they need be given due consideration in the discussion of this study's main conclusions.

23.1 Single Airport vs. Airport Network

Despite Brueckner's (2002b, pp.4-5) suggestion to extend the analysis of airport capacity allocation to a "route structure that more closely resembles a system of actual airline networks," this study, like most other recent work, only considers one single airport. While the results seem to accurately determine allocation efficiency at the single airport, the allocation at an interconnected network hub is likely to be more complicated. In particular, a major congested hub airport should be expected not only to be concerned with non-networking residual supply but also with other networking competitors. Although the competitors may have their hubs at distinct airports, they may connect the same origins and destinations as the home network carrier being considered. From a global perspective, hence, the optimization of customer value needs not only to reflect flight supply at the home airport but also the connections offered by foreign network airlines. This argument ultimately also affects the market power assumption (as delineated in Section 23.2 below).

Moreover, the single airport is characterized as a network hub of a dominant carrier but abstracts from a concise microfoundation of the typical airline network properties such as connectivity, schedule delays and interconnecting routes (see Section 10.2). While this abstraction also admittedly accounts for a weakness in Brueckner's (2002b) network study (see idem, p.4), this model introduces an aggregate formal representation of these network effects in terms of the network density benefits. In this respect, it differs from the models considered in Section 4. This also includes Czerny (2010) and Basso (2008), who include network density as indirect utility but do not dissociate across airlines and only consider density as a function of overall flight volume at an airport.

Therefore, the model may still justifiably claim to introduce the passengers' benefits from airline networks as stipulated by Langner (1996). Moreover, it reflects a corresponding airline asymmetry based on product quality that follows Brueckner's (2002a) suggestion, in which the "passenger valuation of flight frequency is explicitly considered", yet in a more unspecified framework. In this respect, it ultimately responds to Starkie's (2008a) dilemma of airport concentration by providing a formal answer to the valuation of market power against passenger benefits from network operations. These references are specified in more detail in Section 4.5.

Nevertheless, the problem of capacity allocation across multiple networking airlines within several interconnected, constrained airports needs to be addressed by future research. This work might extend, e.g., Brueckner's (2002b) basic network model with the more realistic and detailed supermodular airline profit function as suggested by Aguirregabiria and Ho (2010) and the passengers' explicit evaluation of flight frequency as proposed in Fageda and Flores-Fillol (2013); see Section 4.4.1.

23.2 Traditional Market Power

In light of global competition between major networking airlines, one may arguably doubt whether the assumption of market power may still be supported nowadays. As Pearce (2013, p.17) notes in his IATA industry report, both the "threat of new entrants" and the "rivalry among existing competitors" were high while competitive advantages based on incumbent privileges were limited. Correspondingly, he finds airline profits to be weak on an overall level, so that the industry remained in a "state of extremely poor profitability" (idem, p.10 and p.12).

However, in contrast, Joppien (2003, pp.337) notes that it is the hub airports and the corresponding transport system themselves that provide barriers to entry. Moreover, despite the gradual deregulation of the airline transport market, the global competition between major networks may be presumed to be imperfect because the expansion of international networks seems to be naturally constrained by regulation policy and by airport capacity allocation. This imposes a restriction on the growth of new entrant competitors and on their development of network structures and services. In the same sense, Shepherd (1990, p.449) states that the market entrance for competitors is severely constricted by the scarce resources and their associated regulation, the market presence of the dominant airlines, and their "sharp retaliations in price competition". As a consequence, he deems competition in the airline market to be "immeasurably" limited. More recently, OECD (2014, p.4) acknowledges the above concerns about structural and strategic barriers in terms of airport capacity allocation and "drip pricing strategies". In conclusion, the market for international air travel may thus justifiably be argued to still remain largely incontestable.[95]

Consequently, traditional market power illustrates the notion that a large network carrier enjoys "revenue opportunities by offering travelers a more attractive network with more efficient and frequent schedule options" Brueckner et al. (2010, p.1). In fact, Brueckner et al. (2010) do find significant decreasing price effects when network airline's routes are challenged by low-cost competitors, which may point toward oligopoly rents on the exclusive portions, connections and services of the network.[96] This view is also supported by Verhoef's (2010, p.322) notion that "market power in real aviation markets appears important enough to warrant explicit treatment in the context of congestion regulation". The market power assumption may thus be deemed to be warranted even from a global perspective.

As a result, the market power assumption need not necessarily interfere with global airline competition across networking airlines. A future hub airport model might reflect an oligopoly market structure with multiple networking competitors while accounting for a more elaborate specification of the network properties, as proposed in the previous section.

23.3 Exogenous Airline Asymmetry

The occurrence of network density benefits has been widely discussed in the literature, so that their implementation simply reflects a logical consequence of that discussion. However, the exogenous market separation between the networking business airline and the non-networking leisure airline may be judged as controversial.

The exogenous airline asymmetry implies that the opportunity for product differentiation is asymmetric between firms: Only one airline is assumed to provide indirect network density benefits to its customers in addition to its direct flight benefits. This rationale draws on the common notion that large incumbent airlines enjoy historical first-mover advantages over their new-entrant competitors regarding the airport capacity resource allocation. Based on this notion, it is assumed that this exogenous advantage has allowed the development of sophisticated, optimized network structures and associated services. The airline asymmetry is hence exogenous to the model. It allows the network airline to be dominant over the competing leisure airline by definition. The airline competition thus corresponds to the

[95] For the reference to incontestability, I owe personal thanks to Martin Joppien.

[96] Brueckner et al. (2010) investigate 2008 US data and find price effects between 17.6% and 27.2% when there is a non-stop presence of a low-cost competitor on a route.

situation of two networking high-quality firms and a non-networking low-quality firm that are exogenously dissociated.

This notion may be viewed as uncontroversial in light of actual industry observation, which indicates that dominant network airlines enjoy hub dominance at their network hubs based on incumbent advantages. Moreover, the exogenous quality distinction serves in the investigation of the differences against those recent models, which assume homogenous goods and do not reflect demand-side heterogeneities across airlines. Nonetheless, the strict exogeneity of the airline asymmetry may appear as extreme because after many years of liberalization, also new-entrant airlines may have, over time, established their own - if smaller or less sophisticated - network structures.

Nevertheless, the major network airlines still seem to have a dominant market share at their central network hubs. Their networking competitors may thus be argued to be limited to selected portions of the network by competing on a limited number of routes only. Moreover, in reality they might be compelled to use so-called secondary airports in order to overcome the problem of airport access or to save costs (see Section 10.2.1). Lastly, low-cost carriers by definition limit their associated services based on their cost structures and target customers. For illustrative reasons, product differentiation based on product quality distinctions arising from the exogenous airline asymmetry may still be justified - at least to a crucial extent.

In this respect, it is worth noticing that the previous homogenous product models like Brueckner (2002a) yield completely symmetric market outcomes both in quantity and in price. The results of vertical differentiation models from the literature, however, show that even symmetric firms may reflect an asymmetric market when they have a quality choice: An endogenous quality choice model usually results in a low-quality and a high-quality firm, which are also dissimilar in terms of output quantity and price. As a consequence, relaxing the exogenous airline asymmetry but introducing an explicit quality choice variable would not necessarily yield a symmetric market as in many previous models with flights as homogenous products. Rather, this might allow the investigation of an asymmetric market with endogenous product quality under a less prohibitive assumption. While the exogenous market separation serves as an extreme point for the investigation and may therefore be justified, further research might propose a corresponding vertical differentiation model with a symmetric endogenous quality choice. This suggestion is further explained in Section 23.4 below.

23.4 Dependent Quality Choice

Product quality in this model is specified in terms of indirect network benefits that arise from the density business airline's network. This network density is approximated simply by the business airline's flight volume. As a consequence, the network airline's product quality choice becomes inseparable from its output decision, which, in turn, is based on profit maximization facing endogenous prices from downward sloping demand (see Sections 5.2.2 and 10.2 for a description and justification of this setting, and Section 10.3 for an extensive discussion of this subject). This admittedly affects the transparency of the network airline's two separate but interdependent decisions about its optimal network size and product quality.

In addition, as previously mentioned, for simplicity the horizontal and vertical structures of the network are not explicitly modeled, such as, e.g., by flight frequency and route choice. Instead, the network's depth and width are described by one single variable. This variable reflects network density, which, in turn, is simply assumed to increase with the network size of the airline. As a further shortcut, the network size itself is directly approximated by the business airline's peak-period flight volume. Consequently, the business airline's network density benefits are just a generic function of its output. As a result, product quality is taken into account in the profit maximization rationale but only arises in conjunction with overall output. Although this supports product quality to become endogenous in the model, the quality choice is not modeled as a strategic variable.

On the one hand, this simplification allows the equilibrium computation to abstract from a quality variable in the computation of the first-order conditions. This, in turn, yields analytically traceable closed-form solutions that comprehensively illustrate the characteristic model properties and the investigation of allocation efficiency under different capacity allocation instruments. In comparison, quality choice models usually represent a multistage game that requires numeric solution techniques. This problem especially arises in the light of externalities, which ultimately represent convex cost functions (see Section 10.3.1). The cost of a significant loss in traceability and the increase in computational complexity is therefore not justified within the scope of this study.

On the other hand, however, it does not seem reasonable to completely intermix the quality and size decisions of a network airline. This objection remains valid even if it accounts for the interdependencies that necessarily need to arise between the latter two properties. In this respect, recall that the output effect of a peak-period output expansion in this model yields two distinct effects on the business airline's peak-period flight fare at the same time: a negative contribution based on the reduction of market power in terms of the "traditional" market power effect, but also an ambiguous effect on the network premium based on the concavity of the network value. As a consequence, network quality may turn out to be higher and network size lower in a setting with a separate strategic quality choice. Put differently, the current model might overestimate the equilibrium network size of the business airline and underestimate the problem of the underprovision of network services. Recalling that the welfare caveats of all allocation instruments under investigation increase with higher market power and with more important network density benefits may hence let us presume that the adverse effects of misplaced capacity allocation schemes may be more prominent than indicated by this model.

As proposed above, further research might be undertaken by means of a vertical differentiation model with an explicit network quality choice. Such a model could be based, e.g., on the fundamental contributions of Motta (1993), Ecchia and Lambertini (2006) and Lambertini (2006) but would have to extend the latter by the external effects of congestion. As such, this new model would need to draw on established parametric functions from the economic literature in order to make the quality choice an explicit result. A rudimentary specification has been presented in this study's simulation in Section 15.2. On the downside, however, the externalities ultimately represent convex cost functions, which should be expected to substantially increase the computational complexity of the said basic models.

A corresponding modification and specification of this study's original generic hub-airport model with an airline duopoly and external congestion would represent network quality choice as a two-stage game that could be made available for both airlines and may include an instance of corresponding quality costs. This would permit the study to present explicit solutions for the endogenous quality choice of the network airline and at the same time abstract from the controversial exogenous airline asymmetry assumption. Ultimately, such a model would not only allow one to investigate airport capacity allocation with a networking competitor, but also with an additional separate independent network from a foreign competitor. Although a specification always induces a loss of generalizability, this disadvantage should be overcome by the more realistic setting.

For future reference, the implementation of a quality choice in terms of network density may draw on the vertical differentiation setting of Baake and Boom (2001), who consider both the dimension of a network and its implicit quality in the specification of their customer value, and on the network externalities modeled in Sarkar (2005) and Kobayashi (2011). Moreover, Herguera et al. (2000) propose a framework based on Motta (1993) with a sequential choice of product quality and output quantities in a Cournot setting with quantity constraints. They find that output restrictions will lead to the under-provision of both product quality and output quantity. The results are higher prices for the goods, higher profits for the firms and a higher deadweight loss that is detrimental to the consumers. Without further provisions, this outcome seems to confirm this study's above inaccuracy in terms of the dependent quality choice.

23.5 Nondiscriminatory Pricing

The model assumes that the airlines cannot engage in price discrimination within each period. Correspondingly, the network size is determined by the concavity of the network value, which yields an underprovision of network services. However, Brueckner (2002a) shows that the output of a monopoly supplier will rise up to the socially efficient value if full price discrimination is accounted for. This result is in line with general economic theory (cf. Mas-Colell et al., 1995, p.387). If price discrimination were allowed in the model, its size would become determined by the integral of network density benefits across all passengers, which is equivalent to the social optimum. As a result, the network size would become efficient. The nondiscriminatory pricing assumption in this model may hence lead to an underestimation of the network size.

However, the literature collectively agrees upon the fact that airlines are perfectly specialized in price discrimination (see, e.g., Varian, 1989, pp.646-647). Consequently, it does not seem reasonable that the business airline does not have an opportunity to skim at least some of the higher-wealth individuals' willingness to pay for its network benefits. Nevertheless, assuming full price discrimination also seems a very strict assumption, so that some degree of imperfect price discrimination should probably be an appropriate assumption. This means that the network size may be expected to still be inefficiently low. In this respect, also recall that the business airline exhibits a dominant market share in either case and that this market dominance implies both a high degree of market power and of congestion internalization. Both these effects work in the same direction and thus prevent excessive outputs. As long as there is imperfect price discrimination in combination with market power, the two effects also affect output in parallel and increase the output inefficiency. As a result, it is reasonable to argue that the unregulated equilibrium network size is located somewhere between the two extremes but still remains inefficiently low. Consequently, the welfare caveats from the output inefficiency could be less important but may still be deemed to be significant.

Nonetheless, the above arguments indicate that at least a certain amount of price discrimination should be accounted for in an airport setting with a dominant airline. This would allow the model to properly reflect the pricing process particularly in regard to high-income passengers. One may hence conclude that the network size has been underestimated based on the nondiscriminatory pricing assumption. Consequently, overall flight volume in equilibrium might be less inefficient than suggested by the market power distortion and the concave network value, so that the adverse welfare impact of the network size in practice may remain smaller than indicated. This, in turn, would increase the likelihood that the congestion externality exceeds market power. As a result, this weakens the suggestion that real flight delays might lead to a perceived level of excessive congestion and strengthen the notion that they are founded on excessive airport demand (see the corresponding discussion in Section 14.1.5). In any case, however, this underestimation would generally balance any overestimation based on the dependent quality choice as found in Section 23.4 above. The net effect of the two errors might therefore remain insignificantly small.

23.6 Generic Functional Forms

Lastly, note that the model relies on generic functional forms, so that it provides analytic solutions with a high degree of generalizability. The price of this generalizability, however, comes at the cost of conditional results. Generally, the generic functions cannot resolve the ambiguities that arise from the dual distortion based on market power and the congestion externality. Therefore, in some cases the analysis can merely provide discussions of the different potential outcomes rather than concise results.

In order to overcome these ambiguities, the generic model might be specified with concise functions from recent vertical differentiation models. Such a specified version would hence be able to assess the magnitude and direction both of the market power distortion and of the congestion externality and thus to resolve the above indeterminacy arising from the dual distortion. This would yield explicit results concerning the welfare gains and adversities of the distinct allocation schemes. Furthermore, an empirical calibration of the functional parameters would subsequently allow the model to quantitatively assess the efficiency results of the distinct capacity allocation schemes. Although the simulation in this study undertakes a rudimentary attempt in this direction and already provides reasonable results illustrating the properties of this model, its parameter specification remains stylized and cannot be compared to a proper empirical calibration.

If these results either confirm the welfare caveats that have been indicated previously or at least indicate only small welfare improvements that might in reality be overturned by, e.g., transaction and setup costs, they may lead future studies to investigate the transfer of regulation policy from other sectors with asymmetric and dominant firms as already proposed in Section 22 above. Based on the monotonicity of the network benefits, it is reasonable to suspect that distinct network industries are concerned with similar allocation problems, especially if they also rely on a fixed invariable infrastructure. In this respect, the generic nature of this framework may be specified with any kind of suitable functions. However, any specification requires its corresponding set of assumptions and hence reduces the general applicability of the results. This also applies to the specification of an explicit quality choice model as proposed in Section 23.4 above.

In addition, the welfare investigation might be enhanced with a distributional analysis because matters of distributional equity become essential when it comes to a political debate about the implementation of regulation instruments. Put differently, as soon as a practicable scheme that will satisfactorily improve allocation efficiency has been decided upon, the stakeholders will naturally engage in a subsequent resource competition. This kind of strategic competition (including, perhaps, strategic behavior) might compromise the policy implementation of an efficient allocation scheme, if policymakers cannot provide well-founded and justified answers to the concerns about distributional equity.

Bibliography

References

Acocella, Nicola; Giovanni Di Bartolomeo; and Andrew Hughes Hallett (2012): *The Theory of Economic Policy in a Strategic Context*. Cambridge University Press.

Aguirregabiria, Victor and Chun-Yu Ho (2010): A Dynamic Game of Airline Network Competition: Hub-and-Spoke Networks and Entry Deterrence. *International Journal of Industrial Organization*, 28(4):377 – 382.

Alves, Carlos F. and Cristina Barbot (2010): Rivalry amongst European low cost airlines: Are shareholders enjoying the game? *Journal of Air Transport Management*, 16(6):304 – 309.

Baake, Pio and Anette Boom (2001): Vertical product differentiation, network externalities, and compatibility decisions. *International Journal of Industrial Organization*, 19:267–284.

Backhouse, Roger E. and Yann Giraud (2010): Circular Flow Diagrams. In: *Famous figures and diagrams in economics*, eds. Mark Blaug and Peter Lloyd, Cheltenham, UK: Edward Elgar, chap. 29, pp. 221–232.

Ball, Michael O.; Lawrence M. Ausubel; Frank Berardino; Peter Cramton; George Donohue; Mark Hansen; and Karla Hoffman (2007): *Market-Based Alternatives for Managing Congestion at New York's La Guardia Airport*. Papers of Peter Cramton 07mbac, University of Maryland, Department of Economics.

Barbot, Cristina (2004): Economic Effects of re-allocating airport Slots: a Vertical Differentiation Approach. *Journal of Air Transport Management*, 10:333–343.

Barbot, Cristina (2005): Effects on Welfare of Different Airport Charge Systems. *Journal of Air Transportation*, 10(2):109–126.

Barbot, Cristina (2006): Low-cost airlines, secondary airports, and state aid: An economic assessment of the Ryanair-Charleroi Airport agreement. *Journal of Air Transport Management*, 12(4):197 – 203.

Barbot, Cristina (2006b): *Entry and Accommodation in Airline Markets: Easyjet Caught in the Middle on the London-Grenoble Route*. CEF.UP Working Papers 0602, Universidade do Porto, Faculdade de Economia do Porto.

Barbot, Cristina (2008): Can low cost carriers deter or accommodate entry? *Transportation Research Part E: Logistics and Transportation Review*, 44(5):883 – 893.

Basso, Leonardo J. (2008): Airport deregulation: Effects on pricing and capacity. *International Journal of Industrial Organization*, 26:1015–1031.

Basso, Leonardo J. and Anming Zhang (2010): Pricing vs. slot policies when airport profits matter. *Transportation Research Part B: Methodological*, 44(3):381–391.

Bauer, Jörg (2008): Do Airlines use Slots Efficiently? In: *Airport Slots: International*

Experiences and Options for Reform, eds. Achim I. Czerny; Peter Forsyth; David Gillen; and Hans-Martin Niemeier, Aldershot, UK: Ashgate, chap. 9, pp. 151–171.

Baumol, William J. (1965): *Welfare Economics and the Theory of the State.* London: G. Bell and Sons Ltd., 2nd edn.

Baumol, William J. (1975): *The Theory of Environmental Policy: Externalities, Public Outlays, and the Quality of Life.* Englewood Cliffs, New Jersey: Prentice-Hall.

Baumol, William J. and Wallace E. Oates (1979): *Economics, Environmental Policy, and the Quality of Life.* Englewood Cliffs, New Jersey: Prentice-Hall.

Bel, Germà and Xavier Fageda (2009): *Intercontinental Flights from European Airports: towards Hub Concentration or not?* Working papers 2009/06, Research Institute of Applied Economics, Universitat de Barcelona.

Bel, Germa and Xavier Fageda (February 2010): *Does Privatization spur Regulation? Evidence from the regulatory Reform of European Airports.* IREA Working Papers 201004, University of Barcelona, Research Institute of Applied Economics.

Belleflame, Paul and Martin Peitz (2010): *Industrial Organization: Markets and Strategies.* Cambrigde University Press.

Ben-Akiva, Moshe and Steven R. Lerman (1985): *Discrete Choice Analysis.* Cambridge, Ma: MIT Press.

Berardino, Frank (2009): New US Airport Slot Policy in Flux. *Journal of Transport Economics and Policy*, 43 (2):279–290.

Berry, Steven T. (1990): Airport Presence as Product Differentiation. *The American Economic Review*, 80(2):394–399.

Blaug, Mark and Peter Lloyd, eds. (2010): *Famous figures and diagrams in economics.* Cheltenham, UK: Edward Elgar.

Borenstein, Severin (1988): On the Efficiency of Competitive Markets for Operating Licenses. *The Quarterly Journal of Economics*, 103(2):357–85.

Brueckner, Jan K. (2002a): Airport Congestion When Carriers Have Market Power. *American Economic Review*, 92(5):1357–1375.

Brueckner, Jan K. (2002b): *Internalization of Airport Congestion: A Network Analysis.* CESifo Working Paper 756, CESifo.

Brueckner, Jan K. (2004): Network Structure and Airline Scheduling. *Journal of Industrial Economics*, 52(2):291–312.

Brueckner, Jan K. (2008): *Slot-Based Approaches to Airport Congestion Management.* CESifo Working Paper 2302, University of California, Irvine, Department of Economics.

Brueckner, Jan K. (2009a): Price vs. quantity-based approaches to airport congestion management. *Journal of Public Economics*, 93(5-6):681–690.

Brueckner, Jan K. and Kurt Van Dender (2008): Atomistic Congestion Tolls at Concentrated Airports: Seeking a Unified View in the Internalization Debate. *Journal of Urban Economics*, 64(2):288–295.

Brueckner, Jan K.; Darin Lee; and Ethan Singer (2010): *Network vs. LCC Competition and Airfares: New empirical Evidence from the US.* Iata key policy issue, International Air Transport Association.

Brueckner, Jan K. and Yimin Zhang (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(2):195–222.

Buchanan, James M. (1965): An Economic Theory of Clubs. *Economica*, 32(125):1–14.

Button, Kenneth (1993): *Transport, the environment and economic policy.* Cheltenham, UK: Edward Elgar.

Button, Kenneth (January 2005): A simple analysis of the rent seeking of airlines, airports and politicians. *Transport Policy*, 12(1):47–56.

Button, Kenneth (2007): What can we learn from Auction Theory for Slot Allocation? In: *How to make Slot Markets Work*, eds. Achim I. Czerny; Peter Forsyth; David Gillen; and Hans-Martin Niemeier, Aldershot, UK: Ashgate, German Aviation Research Seminar Series No.3.

Button, Kenneth (2008): Issues in Airport Runway Capacity Charging and Allocation. *Journal of Transport Economics and Policy*, 42(3):563–585.

Button, Kenneth and Erik Verhoef, eds. (1998): *Road Pricing, Traffic Congestion and the Environment.* Cheltenham, UK: Edward Elgar.

Carlton, Dennis W. and Jeffrey M. Perloff (2005): *Modern Industrial Organization.* Addison-Wesley, 4th edn.

Choi, Chong Ju and Hyun Song Shin (1992): A Comment on a Model of Vertical Product Differentiation. *The Journal of Industrial Economics*, 40(2):p.229–231.

Coase, R. H. (1960): The Problem of Social Cost. *Journal of Law and Economics*, 3:1–44.

Cook, Andrew, ed. (2007a): *European Air Traffic Management.* Aldershot, UK: Ashgate.

Cook, Andrew (2007b): The Management and Costs of Delay. In: *European Air Traffic Management*, ed. Andrew Cook, Aldershot, UK: Ashgate, chap. 4, pp. 97–122.

Corchon, Luis C. (2001): *Theories of Imperfectly Competitive Markets.*

Czerny, Achim I. (2006): *Congestion Pricing vs. Slot Constraints to Airport Networks.* Working Papers 2006-02, Center for Network Industries and Infrastructure (CNI).

Czerny, Achim I. (2007): *Congestion management under uncertainty in a two-airport system.* Working Papers 2007-01, Center for Network Industries and Infrastructure (CNI).

Czerny, Achim I. (2010): Airport congestion management under uncertainty. *Transportation Research Part B: Methodological*, 44(3):371–380.

Czerny, Achim I.; Peter Forsyth; David Gillen; and Hans-Martin Niemeier, eds. (2008): *Airport Slots: International Experiences and Options for Reform.* Aldershot, UK: Ashgate.

Daniel, Joseph I. (1995): Congestion Pricing and Capacity at Large Hub Airports: A Bottleneck Model with Stochastic Queues. *Econometrica*, 63(2):327–370.

Daniel, Joseph I. (2001): Distributional Consequences of Airport Congestion Pricing. *Journal of Urban Economics*, 50(2):230–258.

Daniel, Joseph I. (2009a): *The Untolled Problems with Airport Slot Constraints.* Working Papers 09-09, University of Delaware, Department of Economics.

Daniel, Joseph I. (2009b): *Pricing Canadian Airports.* Working Papers 09-02, University of Delaware, Department of Economics.

Daniel, Joseph I (2014): The untolled problems with airport slot constraints. *Economics of Transportation*, 3(1):16 – 28.

Daniel, Joseph I. and Katherine T. Harback (2005): *Do Airlines that Dominate Traffic at Hub Airports Experience Less Delay?* Working Papers 05-09, University of Delaware, Department of Economics.

Daniel, Joseph I. and Katherine T. Harback (2008): (When) do hub airlines internalize their self-imposed delays? *Journal of Urban Economics*, 63:583–612.

Daniel, Joseph I. and Katherine T. Harback (2009): Pricing the Major US Hub Airports. *Journal of Urban Economics*, 66.

De Wit, Jaap and Guillaume Burghouwt (2007): *The impact of secondary slot trading at Amsterdam Airport Schiphol.* Research Report 957, SEO Economic Research. Report for The Netherlands Ministry of Transport (DGTL).

De Wit, Jaap and Guillaume Burghouwt (2008): Slot allocation and use at hub airports, perspectives for secondary trading. *European Journal of Transport and Infrastructure Research*, 8(2):147–164.

De Wit, Jaap; Jan Veldhuis; Guillaume Burghouwt; and Hidenobu Matsumoto (December 2009): Competitive Position Of Primary Airports In The Asia-Pacific Rim. *Pacific Economic Review*, 14(5):639–650.

Deau, Raphael; Jean-Baptiste Gotteland; and Nicolas Durand (2009): *Airport surface management and runways.* Tech. rep., Eighth USA/Europe Air Traffic Management Research and Development Seminar.

Dixit, Avinash K. (1996): Comparative Statics for Oligopoly. In: *Industrial Organisation*, ed. Anindya Sen, Oxford University Press, pp. 95–114.

Dixit, Avinash K. and Joseph E. Stiglitz (1977): Monopolistic Competition and Optimum Product Diversity. *The American Economic Review*, 67(3):297–308.

Dixit, Avinash K. and Joseph E. Stiglitz (1996): Monopolistic Competition and Optimum Product Diversity. In: *Industrial Organization*, ed. Anindya Sen, pp. 188–211. Originally published in: The American Economic Review, Vol. 67(3), June 1977, pp. 297-308.

Doganis, Rigas (2002): *Flying Off Course.* Routledge: London., 3rd edn.

Donohue, George L. and Andres G. Zellweger, eds. (2001): *Air Transport Systems Engineering*, vol. 193 of *Progress in Astronautics and Aeronautics.* The American Institute of Aeronautics and Astronautics, Inc.

EC (2011): *Regulation of the European Parliament and of the Council on the common Rules* *for the Allocation of Slots at European Union Airports*. Proposal (recast), European Commission (EC).

Ecchia, Giulio and Luca Lambertini (2006): Existence of equilibrium in a vertically differentiated duopoly. In: *The economics of vertically differentiated markets*, Cheltenham, UK: Edward Elgar.

Eichenberger, Jürgen (2010): Nash Equilibrium. In: *Famous figures and diagrams in economics*, eds. Mark Blaug and Peter Lloyd, Cheltenham, UK: Edward Elgar, chap. 11, pp. 89–94.

EUACA (2013): *EUACA Slot Guidelines - Slot Performance Monitoring*. European Airport Coordinators Association. Version as of 15.09.2013.

EUACA (2014a): *EUACA Slot Guidelines - Air Carriers that do not hold a valid operating License*. European Airport Coordinators Association. Version as of 30.03.2014.

EUACA (2014b): *EUACA Slot Guidelines - Transparency and interested Parties*. European Airport Coordinators Association. Version as of 30.03.2014.

EUROCONTROL (2008): *Challenges of Growth*. The European Air Traffic Control Agency.

EUROCONTROL (March 2009): *European Air Traffic Management Master Plan*. Eurocontrol and the European Commission.

Fageda, Xavier and Ricardo Flores-Fillol (2013): *Airport Congestion and Airline Network Structure*. Working paper, Department of Economic Policy, Universitat de Barcelona.

Feess-Dörr, Eberhard (1995): *Umweltökonomie und Umweltpolitik*. München: Verlag Franz Vahlen.

Forsyth, Peter; David W. Gillen; Andreas Knorr; Otto G. Mayer; Hans-Martin Niemeier; and David Starkie, eds. (2004): *The economic regulation of airports*. Aldershot, UK: Ashgate.

Forsyth, Peter and Hans-Martin Niemeier (2008): Setting the Slot Limits at Congested Airports. In: *Airport Slots - International Experiences and Options for Reform*, eds. Achim I. Czerny; Peter Forsyth; David Gillen; and Hans-Martin Niemeier, chap. 5, pp. 63–83.

Friedman, James W. (1983): *Oligopoly Theory*. Cambrigde University Press.

Fukui, Hideki (March 2010): An empirical analysis of airport slot trading in the United States. *Transportation Research Part B: Methodological*, 44(3):330–357.

Gabszewicz, Jaskold J. and J.-F. Thisse (1986): On th eNature of Competition with Differentiated Products. *The Economic Journal*, 96:p.160–172.

Gale, Ian (1994): Competition for scarce inputs: the case of airport takeoff and landing slots. *Economic Review*, (Q II):18–25.

Garrow, Laurie A. (2010): *Discrete Choice Modelling and Air Travel Demand*. Aldershot, UK: Ashgate.

Gödeking, Philipp (2010): *Networks in Aviation*.

Gillen, David (2008): Airport Slots: A Primer. In: *Airport Slots: International Experiences* *and Options for Reform*, eds. Achim I. Czerny; Peter Forsyth; David Gillen; and Hans-Martin Niemeier, chap. 4, pp. 41–59.

Gillen, David and William G. Morrison (2008): Slots and Competition Policy. In: *Airport Slots: International Experiences and Options for Reform*, eds. Achim I. Czerny; Peter Forsyth; David Gillen; and Hans-Martin Niemeier, chap. 10, pp. 173–191.

Giocoli, Nicola (2010): Duopoly reaction curves. In: *Famous figures and diagrams in economics*, eds. Mark Blaug and Peter Lloyd, Cheltenham, UK: Edward Elgar, chap. 17, pp. 137–144.

Givoni, Moshe and Piet Rietveld (June 2009): Airline's choice of aircraft size - Explanations and implications. *Transportation Research Part A: Policy and Practice*, 43(5):500–510.

Graham, Anne (2004): The Regulation of US Airports. In: *The Economic Regulation of Airports: Recent Developments in Australasia, North America and Europe*, eds. Andreas Knorr Otto G. Mayer Hans-Martin Niemeier Peter Forsyth, David W. Gillen and David Starkie, chap. 4, pp. 63–72.

Gruyer, Nicolas and Nathalie Lenoir (Jun 2003): *Auctioning airport slots (?)*. Economics Working Papers 01, LEEA (air transport economics laboratory), ENAC (french national civil aviation school).

Harback, Katherine T. (2005): *Airport Takeoff and Landing Queues, Air Carrier Self-Imposed Congestion, and the Internalization Debate*. Ph.D. thesis, University of Delaware.

Herguera, Inigo; Praveen Kujal; and Emmanuel Petrakis (2000): Quantity restrictions and endogenous quality choice. *International Journal of Industrial Organization*, 18:1259–1277.

Holloway, Stephen (2002): *Airlines: Managing to make money*. Aldershot, UK: Ashgate.

Hong, Sungwook and Patrick T. Harker (August 1992): Air traffic network equilibrium: Toward frequency, price and slot priority analysis. *Transportation Research Part B: Methodological*, 26(4):307–323.

Hotelling, Harold (1929): Stability in Competition. *The Economic Journal*, 39(153):41–57.

Hsiao, Chieh-Yu and Mark Hansen (2011): A passenger demand model for air transportation in a hub-and-spoke network. *Transportation Research Part E*, 47:1112–1125.

IATA (2010): *Worldwide Scheduling Guidelines*. International Air Transport Association, 20th edn.

IATA (2014): *Worldwide Slot Guidelines*. International Air Transport Association, 6th edn.

Jehle, Geoffrey A. and Philip J. Reny (2011): *Advanced Microeconomic Theory*. Pearson Educaction Limited, Essex, England.

Jäggi, Fabian (2000): *Gestaltungsempfehlungen für Hub-and-Spoke-Netzwerke im europäischen Luftverkehr : ein ressourcenbasierter Ansatz*. Ph.D. thesis, Universität St. Gallen.

Jones, Ian; Ivan Viehoff; and Philippa Marks (1993): The Economics of Airport Slots. *Fiscal Studies*, 14(4):37–57.

Joppien, Martin Günter (2003): *Strategisches Airline-Management*. Paul Haupt Verlag, Bern.

Karakus, Selami (2009): *Kapazitätsmanagement im Luftverkehr: ausgewählte europäische Flughäfen im Blickpunkt*. Hamburg: Igel-Verlag.

Kilian, Matthias (2008): The Development of the Regulatory Regime of Slot Allocation in the EU. In: *Airport Slots - International Experiences and Options for Reform*, eds. Achim I. Czerny; Peter Forsyth; David Gillen; and Hans-Martin Niemeier, chap. 14, pp. 251–270.

Kobayashi, Shinji (2011): *Network Externalities, Compatibility, and Product Differentiation*. Working Paper 41, Nihon University.

Kost, Albert (2003): *Die Zuweisung von Start- und Landezeitnischen (Slots) im zivilen Luftverkehr: völkerrechtliche, schweizerische und europarechtliche Grundlagen*. Ph.D. thesis, Rechtswissenschaftliche Fakultät der Universität Bern.

Krugman, Paul R. (1980): Scale Economies, Product Differentiation, and the Pattern of Trade. *American Economic Review*, 70(5):950–959.

Laffont, Jean-Jacques; Patrick Rey; and Jean Tirole (1998): Network Competition: I. Overview and Nondiscriminatory Pricing. *The RAND Journal of Economics*, 29(1):1–37.

Lambertini, Luca (2006): A differential game with quality improvement. In: *The Economics of vertically differentiated Markets*, Cheltenham, UK: Edward Elgar.

Langner, Sabine J. (1996): *The Allocation of Slots in the Airline Industry*. Nomos Verlagsgesellschaft: Baden-Baden.

Levine, Michael E (April 1969): Landing Fees and the Airport Congestion Problem. *Journal of Law and Economics*, 12(1):79–108.

Liang, Diana; William Marnane; and Steve Bradford (2001): *Comparison of U.S. and European Airports and Airspace to Support Concept Validation*, The American Institute of Aeronautics and Astronautics, Inc., vol. 193 of *Progress in Astronautics and Aeronautics*, chap. 3, pp. 27–48.

MacDonald, Mott (2006): *Study on the Ipact of the Introduction of Secondary Trading at Comunity Airports*. Tech. rep., Mott MacDonald's Aviation Group in Association with oxera, Hugh O'Donovan, and Keith Boyfield Associates, UK. Technical Report prepared fo the European commission (DG TREN).

Madas, Michael A. and Konstantinos G. Zografos (March 2006): Airport slot allocation: From instruments to strategies. *Journal of Air Transport Management*, 12(2):53–62.

Madas, Michael A. and Konstantinos G. Zografos (January 2008): Airport capacity vs. demand: Mismatch or mismanagement? *Transportation Research Part A: Policy and Practice*, 42(1):203–226.

Madas, Michael A. and Konstantinos G. Zografos (August 2010): Airport slot allocation: a time for change? *Transport Policy*, 17(4):274–285.

Majumdar, Arnab (2007): Understanding En-Route Sector Capacity in Europe. In: *European Air Traffic Management*, ed. Andrew Cook, chap. 3, pp. 65–96.

Mas-Colell, Andreu; Michael D. Whinston; and Jerry R. Green (1995): *Microeconomic Theory*. Oxford University Press.

Matthews, Bryan and Batool Menaz (2003): *Airport Capacity: The Problem of Slot Allocation.* Working paper, Institute for Transport Studies, University of Leeds.

Matthews, Bryan and Batool Menaz (2008): Economic Perspectives on the Problem of Slot Allocation. In: *Airport Slots: International Experiences and Options for Reform*, eds. Achim I. Czerny; Peter Forsyth; David Gillen; and Hans-Martin Niemeier, chap. 3, pp. 21–39.

Mayer, Christoph and Todd Sinai (2003): Network Effects, Congestion Externalities, and Air Traffic Delays: Or Why Not All Delays Are Evil. *American Economic Review*, 93(4):1194–1215.

Mehndiratta, Shomik Raj and Mark Kiefer (August 2003): Impact of slot controls with a market-based allocation mechanism at San Francisco International Airport. *Transportation Research Part A: Policy and Practice*, 37(7):555–578.

Modrego, Eduardo Goni; Mihai-George Iagaru; Marc Dalichampt; and Roger Lane (2009): *Airport CDM Network Impact Assesment.* Tech. rep., Eighth USA/Europe Air Traffic Management Research and Development Seminar.

Moosecker, Carla (2010): *Kapazitaetsbemessung von Luftverkehrssystemen.* Clausthal-Zellerfeld: Papierflieger Verlag.

Motta, Massimo (1993): Endogenous Quality Choice: Price vs. Quantity Competition. *The Journal of Industrial Economics*, 41 (2):p. 113–131.

NERA (2004): *Study to assess the Effects of different Slot Allocation Schemes.* Report for the european commission (dg tren), National Economic Research Associates (NERA).

NEXTOR (2004): *NEXTOR Congestion Management Project - Interim Report: The Passenger Bill of Rights Game.* Interim report, National Center of Excellence for Aviation Operations Research (NEXTOR). George Mason University, sponsored by the Department of Transportation.

Ng, Yew-Kwang (2010): Consumer Surplus. In: *Famous Figures and Diagrams in Economics*, eds. Mark Blaug and Peter Lloyd, Cheltenham, UK: Edward Elgar, chap. 12, pp. 97–103.

Noto, Claudio and Christian Laesser (2013): Die Allokation von Flughafenkapazitäten - Thesen zu Problematik und Lösungsansätzen. In: *Jahrbuch 2013 - Schweizerische Verkehrswirtschaft*, St. Gallen: IMP-HSG, pp. 83–111.

O'Connell, John F. (2006): Corporate Rivalry and Competition Issues in the Airline Industry. In: *Corporate Rivalry and Market Power*, eds. Achim I. Czerny; Peter Forsyth; David Gillen; and Hans-Martin Niemeier, I.B. Tauris, London, chap. 4, pp. 54–75.

Odoni, Amedeo R. (2001): Congestion Pricing for Airports and for En Route Airspace. In: *New Concepts and Methods in Air Traffic Management*, eds. Lucio Bianco; Paolo Dell'Olmo; and Amedeo R. Odoni, Transportation Analysis, pp. 31–44.

OECD (1988): *Deregulation and Airline Competition.* Report, Organisation for Economic Co-operation and Development.

OECD (2014): *Annex to the Summary Record of the 121st Meeting of the Competition Com-*

mittee held on 18-19 June 2014. Tech. rep., Organisation for Economic Co-operation and Development.

OECD (2015): *The OECD Glossary of Industrial Organisation Economics and Competition Law*. Organisation for Economic Co-operation and Development. Latest Version available from www.oecd.org; undated.

Oster, Clinton V. and John S. Strong (2007): *Managing the Skies*. Aldershot, UK: Ashgate.

Oum, Tae Hoon; Anming Zhang; and Yimin Zhang (2012): Airline Network Rivalry. In: *Transport Economics*, ed. David A. Hensher, Routledge: London, pp. 432–454. Source: Canadian Journal of Economics, 28:4a, 1995, 836-57.

Pai, Vivek (2007): On the Factors that Affect Airline Flight Frequency and Aircraft Size. University of California, Working Paper.

Pearce, Brian (2013): *Profitability and the Air Transport Value Chain*. Economics Briefing No.10, International Air Transport Association.

Pigou, Arthur C. (1924): *The economics of welfare*. Macmillan, London., 2nd edn.

Rassenti, S.J.; V.L. Smith; and R.L. Bulfin (Autumn 1982): A Combinatorial Auction Mechanism for Airport Time Slot Allocation. *Bell Journal of Economics*, 13(2):402–417.

Rietveld, Piet and Erik T. Verhoef (1998): Social Feasibility of policies to reduce externalities in transport. In: *Road Pricing, Traffic Congestion and the Environment*, eds. Kenneth Button and Erik Verhoef, Cheltenham, UK: Edward Elgar, pp. 285–308.

Rupp, Nicholas G. (January 2009): Do carriers internalize congestion costs? Empirical evidence on the internalization question. *Journal of Urban Economics*, 65(1):24–37.

Santos, Georgina and Maël Robin (March 2010): Determinants of delays at European airports. *Transportation Research Part B: Methodological*, 44(3):392–403.

Sarkar, Sumit (July 2005): On Vertical Product Differentiation, Network Externalities and Compatibility Decisions: Existence of Incompatible Networks - An Example. Contemporary Issues and Ideas in Social Sciences, Indian Institute of Management, Kozhikode.

Savage, Ian (2006): Economic regulation of transport: principles and experience. In: *International Handbook of Economic Regulation*, eds. Crew and Parker, pp. 350–369.

SDG (2011): *Impact assessment of revisions to Regulation 95/93*. Final report for the european commission, dg move, Steer Davies Gleeve (SDG), London.

Sentance, Andrew (March 2003): Airport slot auctions: desirable or feasible? *Utilities Policy*, 11(1):53–57.

Shaked, Avner and John Sutton (1982): Relaxing Price Competition through Product Differentiation. *The Review of Economic Studies*, 49(1):3–13.

Shankar, Venkatesh and Gregory S. Carpenter, eds. (2012): *Handbook of Marketing Strategy*. Cheltenham, UK: Edward Elgar.

Shepherd, William G. (1990): *The Economics of Industrial Organization*. Englewood Cliffs, New Jersey: Prentice-Hall, 3rd edn.

Sieg, Gernot (January 2010): Grandfather rights in the market for airport slots. *Transportation Research Part B: Methodological*, 44(1):29–37.

Starkie, David (1998): Allocating Airport Slots - A role for the market? *Journal of Air Transport Management*, 4:111–116.

Starkie, David (2008a): The Dilemma of Slot Concentration at Network Hubs. In: *Airport Slots - International Experiences and Options for Reform*, eds. Achim I. Czerny; Peter Forsyth; David Gillen; and Hans-Martin Niemeier, chap. 11, pp. 193–203.

Starkie, David (2008b): *Aviation Markets*. Aldershot, UK: Ashgate.

Tanner, Graham (2007): The Principles of Flight Planning and ATM Messaging. In: *European Air Traffic Management*, ed. Andrew Cook, chap. 2, pp. 35–64.

Tether, Bruce S. and J. Stan Metcalfe (June 2003): Horndal at Heathrow? Capacity creation through co-operation and system evolution. *Industrial and Corporate Change*, 12(3):437–476.

Tirole, Jean (1988): *The Theory of Industrial Organization*. Cambridge, Ma: MIT Press.

Train, Kenneth E. (1992): *Optimal Regulation*. Cambridge, Ma: MIT Press, 2nd edn.

Ulrich, Claus (2008): How the Present (IATA) Slot Allocation works. In: *Airport Slots: International Experiences and Options for Reform*, eds. Achim I. Czerny; Peter Forsyth; David Gillen; and Hans-Martin Niemeier, chap. 2, pp. 9–20.

Van Houtte, Ben (2007): The Single European Sky - EU Reform of ATM. In: *European Air Traffic Management*, ed. Andrew Cook, chap. 7, pp. 181–198.

Van Reeven, Peran (2005): Transport Policy in the EU. In: *Handbook of Transport Strategy, Policy and Institutions*, eds. D. A. Hensher and K. Button, Elsevier Science Ltd., vol. 6 of *Handbooks in Transports*, chap. 41, pp. 705–724.

Varian, Hal R. (1989): Price Discrimination. In: *Handbook of Industrial Organization*, eds. R. Schmalensee and R. Willig, Elsevier Science Publishers, Amsterdam, chap. 10, pp. 597–654.

Verhoef, Erik T. (March 2010): Congestion pricing, slot sales and slot trading in aviation. *Transportation Research Part B: Methodological*, 44(3):320–329.

Vickrey, William (1969): Congestion Theory and Transport Investment. *American Economic Review Proceedings*, 59:251–260.

Vives, Xavier (2001): *Oligopoly Pricing - Old Ideas and New Tools*. Cambridge, Ma: MIT Press.

Wauthy, Xavier (1996): Quality Choice in Models of Vertical Differentiation. *The Journal of Industrial Economics*, 44(3):345–353.

Wöckner, Bernd (2011): *Strategischer Wettbewerb*. 2nd edn.

Wei, Wenbin and Mark Hansen (2005): Impact of aircraft size and seat availability on airlines' demand and market share in duopoly markets. *Transportation Research Part E*, 41:315–327.

Wei, Wenbin and Mark Hansen (2007): Airlines' competition in aircraft size and service frequency in duopoly markets. *Transportation Research Part E*, 43:409–424.

Whalen, W. Tom; Dennis W. Carlton; Ken Heyer; and Oliver Richard (2007): *Proposal For A Market-Based Solution to Airport Delays*. EAG Discussions Papers 200714, Department of Justice, Antitrust Division.

Wichers, Robert (1996): *A Theory of Individual Behavior*. San Diego: Academic Press.

Wolf, Hartmut (2004): Airport Privatisation and Regulation: Getting the Institutions Right. In: *The Economic Regulation of Airports: Recent Developments in Australasia, North America and Europe*, eds. Peter Forsyth; David W. Gillen; Andreas Knorr; Otto G. Mayer; Hans-Martin Niemeier; and David Starkie, chap. 14, pp. 201–212.

Woodruff, Robert B. (1997): Customer value: The next source for competitive advantage. *Academy of Marketing Science Journal*, 25(2):139–155.

Zhang, Anming and Achim I. Czerny (2012): Airports and airlines economics and policy: An interpretive review of recent research. *Economics of Transportation*, 1:15–34.

Zhang, Anming and Yimin Zhang (2006): Airport capacity and congestion when carriers have market power. *Journal of Urban Economics*, 60(2):229 – 247.

Zhang, Anming; Yimin Zhang; and Joseph A. Clougherty (2011): Competition and Regulation in Air Transport. In: *A Handbook of Transport Economics*, eds. Emile Quinet André de Palma, Robin Lindsey and Roger Vickerman, Cheltenham, UK: Edward Elgar, chap. 35, pp. 797–821.

Appendix

A Table of Variables

Generic Model: Indices	
B, L	Business and Leisure airline, also denoted as $i \neq j \in \{B, L\}$
o, p	Off-peak and peak period at the hub airport
Supply	
$N_o = n_o^B + n_o^L$	Off-peak-period flight volumes
$N_p = n_p^B + n_p^L$	Peak-period flight volumes
$N = N_o + N_p$	Overall output across both periods
Π^L,Π^B	Airline profits
c	Constant marginal costs
$g(n_p^i + n_p^j)$	Airlines' congestion costs
Demand	
$\theta \in [\Theta - 1, \Theta]$	Peak-period preference and population range $[\Theta - 1, \Theta]$
$b_p(\theta)$, $b_o(\theta)$	Passengers' direct flight benefits
$t(N_p)$	Passengers' time costs
$d(\theta, n_p^B)$	Passengers' network density benefits
f_p^L, f_p^B, f_o	Flight fares per period and airline
cv_p^L, cv_p^B, cv_o	Customer value per period and airline
Equilibrium & Social Optimum	
$W(\theta)$	Social welfare function
$\underline{\theta} \leq \theta^* \leq \theta^D$	Critical passenger $\underline{\theta}$ for off-peak-period travel, θ^* for peak-leisure travel and θ^D for peak-business travel
B^*	Relative peak-period flight benefits $b_p(\theta^*) - b_o(\theta^*)$
TG^*	Time and congestion costs $t(N_p) + g(N_p)$ for $N_p = 1 - \theta^*$
$d(\theta^D)$	Business airline's network value $d(\theta, n_p^B)\vert_{\theta=\theta^D}$
$\partial n_p^i / \partial n_p^j$	Airlines' conjectural variation $n_p^i(n_p^j)'$

Allocation Instruments: Quotas	
$\hat{q}$, $\hat{q}_L$,$\hat{q}_B$	Individual quotas
$\overline{q}_L$, $\overline{q}_B$	Arbitrary constraint
Secondary Trading	
$TP_i(\triangle q)$	Trading potential of airline i
$\triangle q$	Trading volume of a quota trade
$\mathrm{d}n_p^i(\triangle q)$	Airline i's output change from a trade
Congestion Pricing	
$R_i = n_p^i \cdot r_i$	Congestion tax for airline i (with r_i as tax per flight)
Simulation: Parametric Model	
β_o, β_p	Direct flight benefits (where $\beta_p = \beta_o \cdot \beta$)
τ, γ	Time and Congestion costs
δ	Network density benefits

Tab. A.1: Quick Reference - Table of Variables

B Second-Order Conditions

Formally, a symmetric $n \times n$-Hessian matrix $D^2F(x)$ of a multidimensional function $F(x)$ is negative definite if and only if

$$(-1)^r \cdot \left|D_r^2\right| > 0 \tag{47}$$

for every $r = 1, ..., n$, where $|D_r^2|$ denotes the determinant of the $r \times r$-submatrix D_r^2 that consists of the first $r \leq n$ rows and columns of D^2 (Mas-Colell et al., 1995, p.936). E.g., for a 3×3 Hessian matrix this yields $|D_1^2| < 0$, $|D_2^2| > 0$, $|D_3^2| < 0$ with $D_3^2 \equiv D^2$.

B.1 Social Optimum

Denoting $D^2W(\theta)$ the 3×3 Hessian of the welfare function (14), the three determinants of the 3×3 Hessian matrix for the interior solution $\theta^{OPT} = (\underline{\theta}, \theta^*, \theta^D)$ can be expressed as

$$\left|D_1^2\right| = -b_o(\underline{\theta}) \tag{48}$$

$$\left|D_2^2\right| = b_o(\underline{\theta}) \left[b_p(\theta^*)' - b_o(\theta^*)' + 2 \cdot TG^{*'} + (1-\theta^*) \cdot TG^{*"}\right] \tag{49}$$

$$\left|D_3^2\right| = -\left|D_2^2\right| \cdot \left[\int_{\theta^D}^1 \frac{\partial^2 d(\theta, n_p^B)}{\partial (n_p^B)^2} d\theta - 3 \cdot d'(\theta^D)\right] \tag{50}$$

For strict concavity, (48) and (50) hence need to turn out strictly negative while (49) needs to be positive.

Firstly, $|D_1^2| < 0$ arises as per model definition. Next, from monotonicity and single-crossing it follows that $|D_2^2| > 0$ because delay costs TG^* are assumed non-decreasing and convex. By contrast, the sign of $|D_3^2|$ is contingent on the two terms in the bracket: On the one hand, the integral denotes the curvature of marginal density benefits to the passengers as a function of network density. This integral is generally non-negative unless network density benefits are concave. On the other hand, as already argued in the equilibrium the marginal network value needs to be negative because otherwise the network benefits could be increased along with the number of peak-business passengers. This yields that its contribution to the bracketed term is positive as well. As a result this returns $|D_3^2| < 0$ for non-concave network density benefits so that the Hessian of the welfare function at θ^{OPT} is negative definite. Considering that the secondary effect of the integral is likely to be small should allow to maintain this conclusion even for concave density benefits. Imposing the non-decreasing specification of the delay costs on the network density benefits thus would make this result hold without restrictions.[97]

B.2 Equilibrium

The proof for concavity of the unconstrained equilibrium follows the proof for the social welfare function in Section B.1. The Hessian of both airlines' profit functions Π_L, Π_B as defined by (3) are correspondingly denoted as $D^2\Pi_L$ and $D^2\Pi_B$. Because these profit functions are two-dimensional only, this implies that the Hessian $D^2 \equiv D_2^2$ is a 2×2 matrix. The requirement for strict concavity thus is that $|D_1^2| < 0$ and $|D_2^2| > 0$. In correspondence with the equilibrium computation two cases need to be considered: The interior solution from Section 6.5 where $n_o^B > 0$ and the corner solution from Section 6.6 where $n_o^B = 0$.

The submatrix D_1^2 concerns the off-peak-period condition and hence is symmetric across the two airlines. In generic notation, its determinant can be expressed as

$$\left|D_{i,1}^2\right| \quad = \quad -b_o(\underline{\theta}) + \left(n_o^i + n_p^i\right) \cdot b_o(\underline{\theta})^{"}, \tag{51}$$

[97] Moreover, recalling that concavity does not require negative definiteness as a necessary condition implies that θ^{OPT} could be a global maximum even if condition (47) could not be proven to hold. This would extend the potential range of the social optimum to, e.g., concave network density benefits in combination with a diminutive marginal network value. However, concavity would still need to be proven.

where $i \in \{B, L\}$. While its first term is negative, the sign of the second term initially remains ambiguous: It depends on the sign of b_o" which is generally unspecified. If the direct flight benefits were assumed non-convex then $b_o" \leq 0$ and (51) were unambiguously negative, whereas a convex specification would yield $b_o" > 0$ and thus ambiguity. However, even in the convex case, this second derivative is supposedly diminutive at the left-hand side of the θ-scale where $\underline{\theta}$ is normally located. Moreover, it is inversely proportionate to $n_o^i + n_p^i$. Therefore, this second-order effect is not likely to overturn the primary benefit $-b_o(\underline{\theta})$ even if outputs are large. As a result, it is reasonable to presume that the right-hand side of (51) remains negative and thus fulfills condition $\left|D_{i,1}^2\right| < 0$ for both airlines.

Subsequently, for the leisure airline the full Hessian matrix is denoted as D_L^2. Its determinant becomes

$$\left|D_L^2\right| \quad = \quad -\left|D_{i,1}^2\right| \cdot \left(2\left[B^{*'} + TG^{*'}\right] + n_p^L\left[TG^{*"} - B^{*"}\right]\right). \tag{52}$$

Both the off-peak-period determinant and the first square bracket hence imply a positive sign for (52), while the second square bracket yields ambiguity: Similar to the above case, the direct peak-period flight benefits could be convex so that $B^{*"} > 0$. However, firstly a negativity of the second square-bracket only arises if $B^{*"} > TG^{*"}$. Moreover, this second-order effect may again be reasonably expected to remain small so that it is unlikely to overturn the primary effect from the first square bracket. Consequently, $|D_L^2| > 0$ may justifiably be concluded for a reasonable specification of functions $b_p(\theta)$, $t(N_p)$ and $g(N_p)$.

Finally, the determinant for the business airline's Hessian D_B^2 can be expressed as

$$\left|D_B^2\right| \quad = \quad -\left|D_{i,1}^2\right| \cdot \left(2\left[B^{*'} + TG^{*'} + d(\theta^D)\right] + n_p^L\left[TG^{*"} - B^{*"} - d(\theta^D)^{"}\right]\right). \tag{53}$$

In analogy to (52), the only ambiguity arises from the second square bracket. However, the concavity of the network value implies $d(\theta^D)^{"} < 0$ as it has been revealed in Section 8.2. Therefore, the presence of the network value both increases the absolute positive value of the first square bracket and at the same time decreases the risk of negativity within the second square bracket. Correspondingly, $|D_B^2| > 0$ follows from $|D_L^2| > 0$. Note that in all three above cases the introduction of a non-convexity constraint on direct flight benefit functions $b_o(\theta)$ and $b_p(\theta)$ would remove all ambiguities, so that strict concavity were not only shown but formally proven.

Concerning the corner solution $n_o^B = 0$, the consideration of above conditions (51)-(53) also makes clear that the corner solution does not overturn the previous concavity result. Rather, inserting the result $n_o^B = 0$ into (51) even fosters the case for concavity in the case where direct flight benefits are convex with $b_o" > 0$. In the peak-period conditions (52) and (53) the off-peak-period outputs do not occur at all.

C Additional Numeric Results

C.1 Equilibrium

C.1.1 Sensitivity Analysis

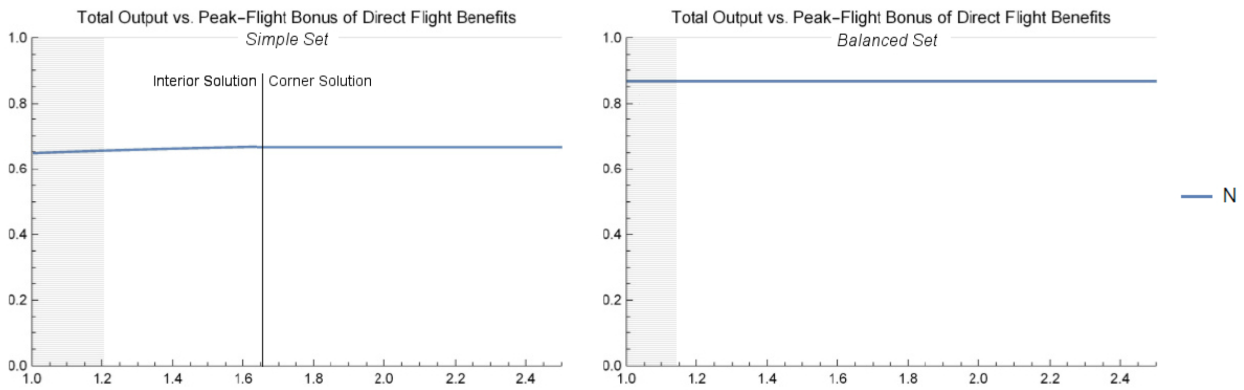

Fig. C.1: Equilibrium - Total Outputs against Peak-Period Flight Benefits

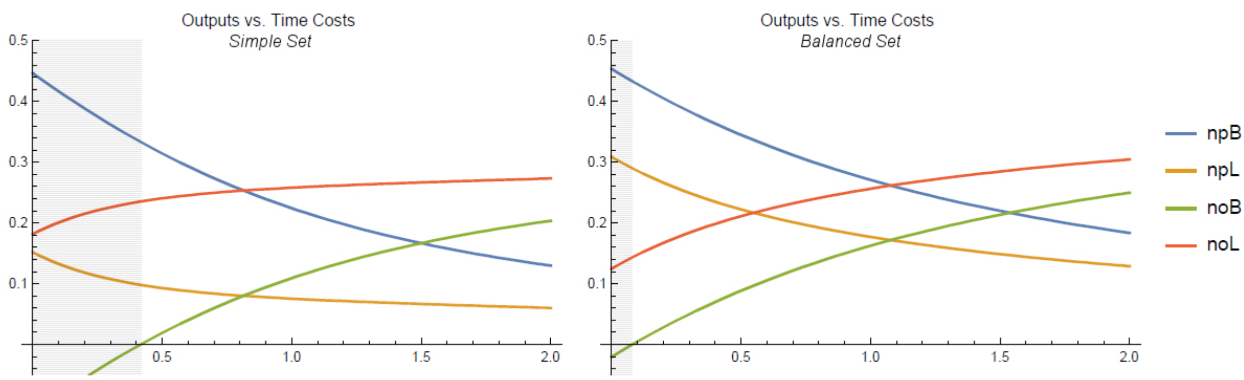

Fig. C.2: Equilibrium - Individual Outputs against Time Costs

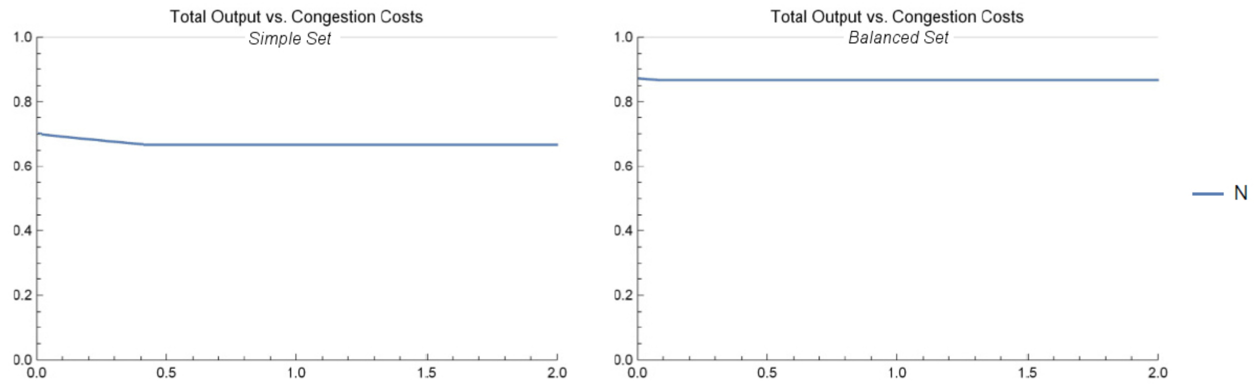

Fig. C.3: Equilibrium - Total Outputs against Congestion and Time Costs

C.2 Social Optimum

C.2.1 Determinants of 3x3 Hessian Matrices

	Simple Set		**Balanced Set**	
	$\delta = 1$	$\delta = 1.5$	$\delta = 0.7$	$\delta = 2$
D1	-1	-1	-1	-1
D2	2.5	2.5	2.6	2.6
D3	-5.4	-6.6	-7.2	-11.6

Concavity if $D1 < 0$, $D2 > 0$, $D3 < 0$.

Fig. C.4: Social Optimum - Determinants of the 3x3 Hessian Matrix

Curriculum

Name	Claudio Giovanni NOTO
born	7. November 1978, in St. Gallen (CH)

Professional Activity

2007-	Airline Pilot, First Officer A330/340, Lufthansa German Airlines

Education

2015	PhD in Management, University of St. Gallen
2007	Airline Transport Pilot License, Lufthansa Flight Training, Bremen
2005	Master of Arts in Economics, University of Zurich (UZH)
2000	Air Force Pilot Certificate, F5-E, Swiss Air Force Pilot School, Emmen
1998	Baccalaureate, Kantonsschule am Burggraben, St. Gallen

Practical & Academic Experience

2001-2005	Student Assistant, Socioeconomic Institute, Chair for Public Finance, UZH
2003-2004	Teaching Assistant, Undergraduate Tutorials in Micro- & Macroeconomics, UZH
2004	ERASMUS Student Exchange Trimester, Universitat Pompeu Fabra, Barcelona
2002	Internship, UBS Investment Bank, Glattbrugg (CH)

Presentations at Conferences

2014	11th Aviation Student Research Workshop, German Aviation Research Society, Bremen.
2013	13th Swiss Transport Research Conference, ETHZ/EPFL/UNISG, Ascona.
2012	9th Aviation Student Research Workshop, German Aviation Research Society, Bremen.

Publications

2013	Noto, Claudio and Christian Laesser: Die Allokation von Flughafenkapazitäten - Thesen zu Problematik und Lösungsansätzen. In: *Jahrbuch 2013 - Schweizerische Verkehrswirtschaft*, St. Gallen: IMP-HSG, pp. 83–111.