Airport Capacity Allocation with Network Airlines

Regulation of Congestion Externalities under Imperfect Competition, with Product Differentiation based on Network Density Effects

Claudio Noto

University of St. Gallen

April 2013

STRC | 13th Swiss Transport Research Conference

Monte Verità / Ascona, April 24 – 26, 2013

University of St. Gallen

Airport Capacity Allocation with Network Airlines

April 2013

Abstract

This study provides a theoretical framework to investigate the efficiency of different airport capacity allocation schemes under congestion externalities and imperfect competition. Its innovation is to consider a dominant network carrier at a hub airport, that endogenously differentiates its product based on passenger benefits from network density. The aim is to formally capture this currently prevailing market structure at large network airports, in order to determine the ambiguities in allocation efficiency that arise with market power and congestion externalities. The results are not expected to strictly favor recently proposed alternative allocation instruments. Hence this study might suggest to unfocus on the latter, but rather investigate a transfer of monopoly regulation schemes known from other network industry sectors. Within the current perspective on the airport capacity allocation problem from previous work, this consideration has not yet received major attention.

Keywords

Airport Capacity Allocation, Market Power, Congestion Externality, Airport Slots, Congestion Pricing, Secondary Trading, Network Airlines, Network Density Effects, Network Operating Licenses

Executive Summary

While major airports still suffer from runway congestion, the current administrative regulation to allocate scarce airport capacity to the airlines has been criticized to be both inefficient and inequitable. Theoretical studies therefore have proposed alternative economic allocation instruments, and shown their efficiency at least for perfectly competitive markets. But when airlines have market power, these results become ambiguous: As imperfect competition already alters output and prices away from efficient levels, additional externalities from congestion lead to a dual distortion in the resource allocation. Consequently, allocation instruments can have an adverse impact on efficiency, if the second-order effect from market power dominates. This may occur with both congestion pricing and secondary trading schemes.

The current market structure at major airports seems to be characterized rather by large, dominant network airlines than by perfect competition. Now, if these dominant airlines can differentiate their products based on network economies, their market power increases, and competition is reduced. On the one hand, this gives rise to the above ambiguity from congestion externalities and market power. But on the other hand, network effects have also been discussed to create additional benefits for passengers. This suggests that the impact of regulation policy on allocation efficiency becomes increasingly complex, when network effects are accounted for. But while quite a number of studies have already analyzed alternative allocation schemes for airport capacity, only few account for airline market power, and mostly with homogeneous goods only. Asymmetries have been considered based on cost differences, but not on demand heterogeneities. And finally, network density benefits for passengers have been discussed, but rarely been formally captured in recent theoretical frameworks. Conclusively, the efficiency of different regulation policies has not yet been illuminated for the case where the natural market structure is characterized by network airlines.

This study provides a theoretical framework to investigate the efficiency of different airport capacity allocation schemes under congestion externalities and imperfect competition. The innovation is that it formally captures the currently prevailing market structure at large network airports: a dominant network carrier, that endogenously differentiates its product based on passenger benefits from network density. The objective is to determine the ambiguities in capacity allocation efficiency, that arise with market power and product heterogeneity.

The results do not favor the proposed alternative allocation instruments. Instead, they suggest further research to focus on monopoly regulation, as known from other industry sectors with networks (e.g. railway infrastructure, power grids, telecommunications or public transport). A tempting idea were to investigate the implementation of network operating licenses, that allowed density benefits to fully develop, both at fair costs for the passengers and at fair privileges for the network airlines. Within the current perspective on the airport capacity allocation problem, this idea has not received attention yet.

Contents

List of Tables

List of Figures

1 Introduction

At major airports, excess demand for operational runway infrastructure has induced congestion (Matthews and Menaz, 2008, p.22). This airport congestion has been recognized as a genuine problem (Brueckner, 2009b, p.11), leading to flight delays that impose significant costs to airlines and passengers (Cook, 2007, p.97). [1] Hence, the "efficient allocation" of the "increasingly scarce airport capacity" has been judged as "one of the key economic issues" in the industry (De Wit and Burghouwt, 2008, p.148). And as most major airports are home to large, dominant carriers, and as large airlines focus on networks (Holloway, 2002, pp.23-24), it seems reasonable to suggest that to a large extent this problem concerns large network hub airports.[2] Because most of the congestion that airlines impose on airport-operations is likely to represent an externality, regulation has been implemented to limit flight volume by means of capacity constraints. Allocation efficiency then presupposes that the level of the constraint is optimally chosen, and that capacity is allocated to the flights according to their overall social value (Matthews and Menaz, 2008, p.25). Both these tasks, however, seem not to be as simple in practice: Determining the optimum number of quota is everything but straightforward (Ulrich, 2008, p.11), and perfect knowledge of all costs, benefits and the social value of all potential flights to be allocated seems to be too strong an assumption (De Wit and Burghouwt, 2008, pp.152). It seems not surprising, therefore, that recent literature has criticized the current, administrative capacity allocation scheme to be both inefficient and inequitable (cf. e.g. Matthews and Menaz, 2008, pp.24).

[1] Operational limitations may concern different parts of the operational infrastructure, but mostly involve runway capacity (Natalie McCaughey and Starkie, 2008, see). Whalen *et al.* (2007, p.7) expect 27 US-airports to be capacity constrained by 2025, and Majumdar (2007, p.65) suspects sixty European airports to be congested as of 2020. Eurocontrol estimates year 2025 European air traffic to reach up to 210 percent of the 2005 volume (EUROCONTROL, 2008, pp.1). (Cook, 2007, p.97) regards congestion costs for both passengers and airlines as "real, large, but poorly understood quantitatively", and dissociates tactical and strategic costs of delay for airlines. Tactical costs represent irregularity costs that occur due to actual delays, where strategic costs arise from contingency planning that aim to minimize the impact of actual delays. For passengers, delay costs amounted mainly to the time costs from late arrivals and missed connections. Generally, Donohue and Zellweger (2001, p.7) emphasize that the constrained aviation system had a large negative effect on both US and EU economic growth. Quantitatively, Morrison (2005, p.418) estimates annual congestion in the US to decrease economic welfare by 4 billion USD in 1988, and Whalen *et al.* (2007, p.6) compute US travelers' time costs for the first quarter of 2007 to amount to 239 million USD.

[2] Based on rudimentary empirical data, Natalie McCaughey and Starkie (2008) shows that at most major slot-constrained airports worldwide, the market share of the main carriers generally reaches 40 to 60 percent in terms of flight movements.

Consequently, a considerable number of studies offers alternative solutions to this classical economic allocation problem. Most prominently, they propose allocation schemes based on market instruments, and internalization of congestion by means of taxation (cf. e.g. Brueckner, 2009a, p.682). In theoretical, perfect competition models, optimality can be shown for all of these alternatives. This is in-line with economic theory, where in absence of market distortions, both internalization of an externality, as well as a market-based allocation equivalently reproduce efficiency (Mas-Colell *et al.,* 1995, pp.356). However, when competition is imperfect, capacity allocation also affects market concentration, and thus the degree of market power among competitors. This, in turn, directly induces second-order effects on efficiency, which may offset or even overcompensate the positive welfare impact of regulation. So, as Starkie (2008b, p.135) puts it, "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". In other words, if allocation efficiency is compromised not only by externalities, but also by market power, the assessment of allocation instruments needs to account for their competitive impact, too. Considering this gap between theory and practice, thus, it seems not surprising that no unified view has been found on the airport capacity allocation problem yet, and that so far alternative policies have rarely been implemented in practice (Madas and Zografos, 2010, p.275). The question hence still seems to be about „the facets of the problem that we have been missing." (Lave, 1995 in Rietveld and Verhoef, 1998, p.285). As this study argues, one of these facets are that competition at large network hub airports is driven by product differentiation based on network density benefits. This implies that their market structure may justifiably be characterized as an asymmetric oligopoly. As mentioned above, the resulting market power should be expected to affect allocation efficiency under all allocation schemes. As laid out below, however, this suggestion has not yet been reflected in theoretical models within the current airport capacity allocation discussion.

Considering capacity allocation at hubs of dominant network airlines, two major gaps attract attention within recent theoretical work in the field:
First, the majority of studies provide perfect competition settings, or at least deliberately abstract from elastic demand. Although this allows to focus on the analysis of the instruments, evidently market-power effects and their second-order impact on efficiency are suppressed. In contrast, the few studies that consider imperfect competition, as suspected, do find potentially

adverse welfare effects from allocation instruments in conjunction with market power. But although they reflect endogenous pricing and allow for second-order effects, still most of the latter only consider flights as homogenous products. This means that there is only one market price, that is determined by total industry output, and is identical across all firms (Vives, 2001, p.94). And as Berry (1990, p.394) puts it, this only represents "traditional market power" that yields profits "by restricting output and driving up prices". Such traditional market power is implemented in Brueckner (2002a), Brueckner (2002b), Basso and Zhang (2010) and Verhoef (2010). But as Gillen and Morrison (2008, pp.178) point out, the airline industry is likely one of product heterogeneities. And as O'Connell (2006, p.54) states, modern airline competition is based on *differentiation strategies*, where successful airlines strategies are able to implement *competitive advantages*. These, in turn, translate into *higher markups*, and thus allow airlines to outperform their competitors (Holloway, 2002, pp.23-24). Thus, at a large network hub, product differentiation may justifiably be characterized as a driver of competition that endogenously affects market power. This would mean that, as Berry (1990, p.394) argues, "both simple cost-reducing and naive market power stories are inappropriate for the airline industry". But although already Brueckner (2002a) mentions the need to account for airline asymmetries other than cost-side differences, market-power that allows for *endogenously differentiated prices* based on demand-side heterogeneities (Berry, 1990, p.394) has not been reflected in recent oligopoly models.

And second, despite the fact that the business model of the hub-and-spoke network airline has faced "near-universal adoption" (Oum *et al.*, 2012, p.432) for multiple decades (Burghouwt, 2007, Zhang *et al.*, 2011, p.803), large network airlines that exploit *network density effects* have not yet been reflected in recent models of airport capacity allocation discussion. But as competitive advantages depend on the market structure (Holloway, 2002, pp.23-24), and a network can be shown to be a dominant strategy in oligopoly competition (Oum *et al.*, 2012, p.432), one might think of network airlines to differentiate their products against competitors based on *network density effects* (Starkie, 2008a, p.197). This would formally capture the idea that networks provide "superior connectivity and wider market coverage" (O'Connell, 2006, p.60), with *hub connectivity* having a "tremendous commercial impact" (Goedeking, 2010, p.21). With density economics being the basic rationale for the prevailing hub-and-spoke structures (Brueckner, 2002b, pp.10-11), competitive advantages on the production side might then be based on network density (Jäggi, 2000, p.271). And because network synergies based on such network density are recognized as *revenue related economies* (Jäggi, 2000, pp.61), for network airlines airport access rights become key business assets (idem, p.272).The notion that network density effects, based on higher market concentration, may constitute heterogeneities that translate into higher markups is strengthened by Aguirregabiria and Ho (2010, p.1), who empirically find differences in prices and hub sizes across airlines, when profit functions account for complementarity. But although network density effects are accounted for in a few models , none of these relate them to competitive advantages, let alone product differentiation. Also Langner (1996, p.15ff.) opinionates that suggestions for alternative instruments did not account for the "network characteristics of flight services", and Aguirregabiria and Ho (2010, p.1) stress that the consideration of airline networks in the context of entry deterrence had been neglected. But following the above reasoning, density benefits as drivers of comparative advantages for network airlines may reasonably be considered to contribute to the "*socially ambiguous nature of much of product differentiation in this industry*", which makes "*welfare analysis particularly difficult*" (Berry, 1990, p.398). The above arguments thus clearly point out the need for a theoretical model that appropriately captures the market structure of a large network hub, by taking into account product differentiation based on network density effects.

This study aims to cover this gap in the theoretical analysis of scarce airport capacity allocation. On this purpose, it proposes a conceptional framework for a theoretical *partial equilibrium model* of an airport-airline network, that allows for *product differentiation* based on network density effects. Horizontal differentiation is based on the idea of distinct product quality, that affects the passengers' customer value from flights with the network airline. Heterogeneous product quality is introduced by means of indirect passenger benefits, that arise from the network density of the network airline. This leads to a demand-side asymmetry in an imperfect airline competition case with market power. *Product heterogeneity* by airline-specific *network density* ultimately allows for *endogenous pricing*, that enables market concentration "to affect both costs *and* demand" (Berry, 1990, p.394). Specifically, the model is based on Brueckner (2002a)'s discrete choice model with symmetric airlines, and further developed according to the general foundations for network goods of Belleflame and Peitz (2010). In general, it reflects the following situation: One airline is supposed to be a *network airline*, where customers enjoy *indirect network benefits* from flight frequency. The other airline is to reflect a straight *point-to-point airline*, that simply offers transportation, without network benefits. Thus, while the network airline endogenously controls market power by horizontal product differentiation and creates network density benefits for passengers, the other airline only provides flights as a homogeneous good. Network density benefits are supposed to represent utility from a high connectivity of the network airline, through the number of destinations and the flight frequency on routes. This connectivity is depicted by network density in the model, which in turn is simply represented by the business airline's flight volume during peak times. Admittedly this is a very basic implementation of the complex concept of network benefits. Nevertheless, it serves for the most simplistic illustration of the case, while it helps to focus on the basic issue of the asymmetry of the airlines. Product differentiation comes along with heterogeneous consumer taste, which is captured by distinct preferences for peak versus off-peak-travel and for network density benefits. This reflects business and leisure travelers. Indirect utility from network density thus only arises from the number of flights of the *business* airline, and only peak-travel sensitive passengers are willing to pay for these benefits. Because network density is based on flight volume, both product differentiation and pricing are *endogenous* to the output decision of the network airline. Ultimately, thus, *endogenous* network density represents a product heterogeneity that yields indirect benefits for passengers. In that, it becomes a comparative advantage that creates endogenous market power for the network airline, and thus also affects airport demand. Hence, it affects allocation efficiency by its potential impact both on congestion and on market concentration.

The innovation of the model thus is that flights are assumed to be *imperfect substitutes*, and that their degree of *differentiation* is *endogenous*. This reflects quantity competition with differentiated prices in oligopoly. Its unique contribution thus is the introduction of *endogenous product differentiation*, that allows to study asymmetric airport demand with a network airline. Subsequently the model is used to investigate the impact of different airport capacity allocation schemes on efficiency. As the *dual distortion* considerably complicates regulation policy, this setup allows to assess whether current and proposed instruments are appropriate for the natural market structure at the large network airlines' hub airports.

The study supports the result from previous imperfect competition models, that the welfare effects of alternative allocation schemes are ambiguous. As the model only reflects generic functions, however, these ambiguities ultimately cannot be resolved. Still, theoretical considerations strongly point towards severe welfare caveats of both a market allocation of constraints, as well as of a tax-based internalization of congestion. One possible implication of this might be that in a network hub context, instead of allocation instruments for individual flights, natural monopoly regulation might have to be considered, that would enable the internalization of externalities while at the same time accounting for market-power problems. Future research therefore might investigate a transfer of monopoly regulation from other sectors to the network airline case, especially from industries that also rely on network structures, such as e.g. telecommunications networks, rail infrastructure, power grids or public ground transport. A tempting proposition were network operating licenses, that allowed the network company to fully develop the network benefits for the customers, while its price might relate to the market power that it provides. At least under the current perspective on the airport capacity allocation problem, such an approach has not yet been considered. In order to emphasize the above theoretical arguments, however, the model should be evaluated for different specific functions, in order to show explicit conditions for the above welfare effects. Moreover, the model parameters could be estimated and calibrated in a real-world context to yield quantitative results. This could be envisioned by further development of this study.

One sensitive issue concerning the market power context of this model should, however, be considered: From a perspective limited to one single market centered around the considered hub airport, it seems natural that the network airline enjoys market power against the leisure airline. However, one might think of other competitors that also offer network benefits from their flights, while being based at other hub airports than the one concerned. So, in a global market, the networks and their dominant airlines themselves would enter into competition, because all non-direct flights between two arbitrary points on earth simply differ by the place where connection is made. Then, apart from other quality differences among distinct networks and carriers, the comparative advantage based on network benefits would in general also be subject to competition. Hence, global competition *between networks* should be considered within the capacity allocation problem, too, although it remains unclear how it relates to market power and to airport demand of a carrier in its network hub.

2 Background and Literature

In Europe and most other world regions, access to large congested airports is restricted by means of airport quota (Ulrich, 2008, p.9). These so-called *Airport slots* constitute access rights, based on an IATA (International Air Transport Association) policy, and legally established by European law. [3] Airport slot allocation takes place in semi-annual, strategic coordination conferences with all airports and airlines concerned. It is mainly based on *grand-fathering rights* of established participants, and on reserving a marginal share for market entrants or expanding airlines (Ulrich, 2008, pp.10). [4] In the US, airport capacity is generally available on a first-come, first-served basis (Madas and Zografos, 2010, p.275). In practice, the administrative allocation process is described as “extremely successful” (Ulrich, 2008, p.10), and found to ensure transparency, fairness and non-discrimination (Bauer, 2008, p.152). However, recent literature has raised severe criticism that it was both *inefficient* and *inequitable* (cf. Matthews and Menaz, 2008, pp.24, or De Wit and Burghouwt, 2008, pp.148): The privilege of established airlines implied entry barriers for business rivals and thus reduced competition (Daniel, 2009b, p.22), and were opposed to the general infrastructure policy goals. [5] On this matter, also legal concerns in terms of commercial freedom of action and equal opportunities are raised (Kost, 2003, pp.108). From an economic perspective, two market distortions arise within the oligopoly context of this study: market power and congestion as an externality: First, market power causes the *market size* to decrease from the social optimum. This induces a *deadweight loss* and excess profits for the firms. Moreover, it determines to what extent congestion is internalized. Second, the congestion externality c.p. leads to a higher output than socially optimal, because only part of the congestion costs are accounted for in the airlines profit-maximizing rationale. In the present context, the impact of market power and the *congestion externality* on *market size* are *opposed* to each other. As at least one of the two is always present, one may

[3] Airport slots must not be confused with tactical "slot" times imposed by air traffic control for operational flow management in case of capacity disturbances during daily operations, technically labeled as CTOT (Calculated Take-Off Time; see Tanner, 2007, p.35).

[4] *Grandfathering rights* refer to the principle that air routes operated regularly during one period are granted airport access at first priority during the next period (Ulrich, 2008, p.12). In general, thus, established connections have priority over the expansion plans of airlines (De Wit and Burghouwt, 2008, p.150). Sieg (2010) thus refers to grandfathering as to *quasi-property rights* for airports access.

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

refer to this as to a *dual distortion*. It is depicted in Figure 1, and has the following effect: [6]

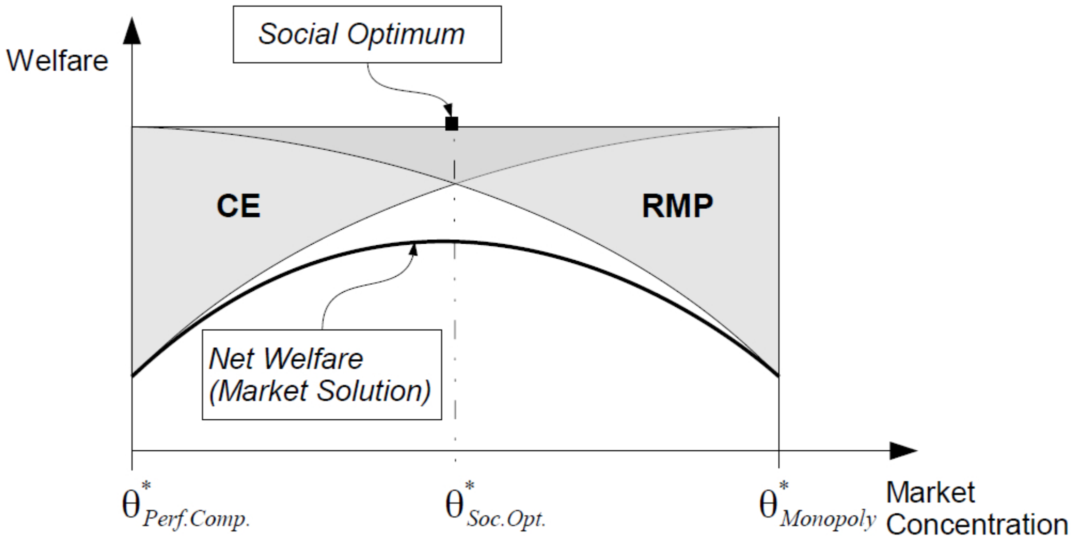

Figure 1: Congestion Externality (CE) and residual Market Power (RMP) [Source: own illustration]

The graph represents the welfare level that can be reached with different degrees of market concentration θ. An increasing market concentration means higher *residual market power* (RMP) but lower *congestion externalities* (CE). The extreme to the right is a non-discriminating monopoly, that fully internalizes congestion but causes a large deadweight loss. The other extreme, at the left, is a perfectly competitive market. It suffers from excessive congestion, as congestion costs are fully external, but is not concerned with market power. The net welfare curve depicts the welfare levels that can be reached by market solutions. The *first-best* welfare level corresponds to the social optimum, where by definition congestion is internalized and output is not distorted by market power. As the graph shows, it can never be reached with a market solution, because the latter always suffers from at least one of both distortions.[7]

In other words, any non-discriminating pricing market equilibrium is inefficient. But because regulation can correct for the congestion externality, at least it can be second-best. Depending on the relative size of the two effects, however, it may also increase the market power distortion, and in sum cause an overall negative welfare effect.

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

[7]The only exception to this is a perfect price discriminating monopoly. It can replicate the social optimum, because it both fully internalizes congestion, and removes the market power distortion at the same time, because there is no deadweight loss. Perfect price discrimination, however, is not treated in here.

Both the actual scheme and the proposed alternative instruments for airport capacity allocation are founded on the three *general economic instruments* to compensate for external effects: *quota*, *taxation* and *decentral bargaining* (cf. Mas-Colell *et al.,* 1995, pp.351). *Quota* reduce externalities to a *socially optimal level* by restricting the activity from which they are caused. With *taxation*, the costs of an activity are adjusted so that they include the external effects caused. Such tax needs to equal the marginal costs of the externality. Finally, *decentral bargaining* is e.g. implemented by perfect competition in a classical, free market allocation.

In theory, quota, taxes as well as a market solution are all equally efficient, if markets are perfectly competitive, and costs and benefits are fully known to the regulator or the market participants, respectively (Mas-Colell *et al.,* 1995, pp.356). The fact that perfect information is a strong requirement, however, already makes evident that efficiency might be difficult to reach in an administrative allocation process. [8]

According to their occurrence in recent studies, the externality taxation scheme known as *congestion pricing* and the market-based allocation of airport slots by means of *secondary trading* are the two most prominent propositions Brueckner (2009a, p.682). As mentioned above, *Congestion pricing* internalizes the congestion externality by means of taxation. *Secondary trading*, in contrast, introduces opportunity costs for slots, which should reflect proper costs of infrastructure scarcity, and thereby also lower barriers of entry and increase competition at the benefit of the consumers (e.g. Gillen and Morrison*,* 2008, p.174).

As explained above, however, in oligopoly with market power their effects become ambiguous: In his theoretical model, Brueckner (2002a) shows that a *congestion tax* can actually *deteriorate* efficiency, if output is already distorted by market power: As the congestion externality tends to *increase* output, but market power *decreases* flight volume against the social optimum, the two distortions are opposed to each other. [9] Consequently, taxation of excess flights works in the same direction as market power. Now, if if the market power distortion is large in relation to the congestion externality, it already *offsets* most of the congestion distortion in terms

[8] This argument strictly refers to the *allocation* of quota, but not to the determination of their optimal size. Otherwise, one might argue that also for optimal taxation, the monetary equivalent of the externalities were difficult to obtain, and that with decentral bargaining, all preferences needed to be known to the market participants. Thus, although *incomplete information* may preclude an efficient solution with all three above options at the quantification stage (Mas-Colell *et al.,* 1995, pp.368), this study only considers the *allocation problem* of pre-determined constraints or externalities, respectively.

[9] Formally, the direction of the residual market power effect depends on the assumptions about consumers' benefits. Brueckner (2002a)'s model requires marginal flight benefits to be higher for business than for leisure travelers, which sounds fairly reasonable. In accordance with the latter, therefore, also in this study market power is presumed to be strictly output decreasing.

of output. A further reduction of output by the congestion tax then might shift the equilibrium further away from the social optimum than without regulation. If, in contrast, the congestion externality is more important relative to the market power distortion, the volume-decreasing effect of the tax still had a positive impact on social welfare. In conclusion, the *ambiguity* of congestion pricing under market power depends on the size of the residual market power effect *relative* to the congestion externality (Brueckner, 2002a, p.1368). In an asymmetric duopoly with product differentiation, however, market power may justifiably expected to be large. But also with *secondary trading*, market power might constitute a caveat to efficiency: In order to increase market concentration through entry deterrence, network airlines might be willing to afford higher prices for access rights than their smaller competitors (Matthews and Menaz, 2008, p.36). This would allow them to achieve network density, and thus scarcity rents based on hub premiums (e.g. Starkie, 1998, p.114 or 2008b, pp.171). [10] But again in contrast to this, the associated network benefits might make passengers willing to pay a network premium, and in terms of welfare offset the hub premium (Berry, 1990, p.394). In other words, both airlines and passengers might profit from "hub networks based on network density benefits" (Starkie, 2008a, pp.193). This conjunction of network benefits and market concentration hence also have ambiguous effects on welfare, and has therefore been brought up as the *dilemma of airport concentration* (Starkie, 2008a). Both market power as well as network density benefits, hence, also disguise the welfare impact of secondary trading. Asymmetries, in turn, might even completely prevent trading opportunities at positive prices. In literature, empirical evidence shows only much limited trading in historical market approaches (Fukui, 2010, Starkie, 2008a and De Wit and Burghouwt, 2008), and adoptions of a congestion pricing mechanism have rarely been reported so far. Moreover, the recent theoretical contributions illustrate the above ambiguities:

In perfect competition models of *single airports*, Daniel (1995), (2001) and (2009a) find socially efficient results for airport congestion pricing compared to administratively allocated slots. For *airport networks*, Hong and Harker (1992) also report positive welfare effects of market-based airport capacity allocation, and Brueckner (2002b) as well as Czerny (2006), (2007) and (2010) find welfare benefits of optimal congestion charges against a slot allocation. As comparative studies, Brueckner (2008), Brueckner (2009a) and Basso and Zhang (2010) find that both congestion pricing and secondary trading schemes may provide first-best allo-

[10] (Morrison, 2005) finds empirical evidence for hub premiums at large network airports.

cation efficiency, but only if strategic airline behavior was ruled out. Although the latter two claim imperfect competition settings, they actually abstract from market power by assuming inelastic demand, and thus face the perfect-competition aggregate output.

In contrast to the above studies, however, Brueckner (2002a), Barbot (2004), (2005) and Verhoef (2010) account for market power. As expected, they find ambiguous results for alternative instruments:

As already mentioned, Brueckner (2002a) shows that a congestion tax can have adverse welfare effects, if the market power distortion is large relative to the congestion externality. Barbot (2004) finds a welfare-decrease after a market re-allocation of airport access rights, because airlines engaged in higher price differentiation, and Barbot (2005) finds that in her setting, *congestion charges* actually decreased welfare. And finally, Verhoef (2010) both confirms Brueckner (2002a)'s above result for congestion pricing, and finds that for secondary trading with a high degree of market power, aggregate demand may be too low to allow trading at positive prices. When the congestion externality is important than market power, in contrast, he shows that the less efficient airline is driven out of the market, yielding a monopoly, which he claims to be potentially worse than no regulation at all. Hence, also a secondary trading scheme can not be claimed to unambiguously yield an efficient allocation, when market power is present.

Still, also the above market-power studies consider flights as homogenous and airlines as symmetric. Only Verhoef (2010) provides an asymmetric airline case and moreover includes market power. His heterogeneity, however, is based on different marginal costs.

Therefore, although the above results point towards potentially hazardous results with alternative schemes, the welfare effect of an *airline asymmetry* based on *demand driven product differentiation* by means of *passenger network benefits* remains to be investigated.

Indirect benefits from network goods have been described and generically formalized in economic network theory (e.g. Belleflame and Peitz, 2010, pp.550). According to their nature, they may be put as to correspond to the prominent concept of *customer value* from the business administration literature. It describes the overall value of a product to the customer that not only includes utility from consumption, but also from product attributes (Woodruff, 1997, p.142). The integration of customer value into the concept of economic utility had already been brought up by Wichers (1996)'s *theory of individual behavior*, and product attributes that enter the utility function had been introduced by Dixit and Stiglitz (1977) and Krugman (1980) in the monopolistic competition literature. Customer Value as a competitive advantage against competitors has been mentioned by Woodruff (1997), and now is a standard argument in marketing theory (Shankar and Carpenter, 2012, cf. e.g.). Horizontal product differentiation, in turn, may thus be stated as its well-known foundation from economics (see e.g. Woeckner, 2011, pp.15). In order to stress the above causality concerning product differentiation and competitive advantages, in the model I refer to the passengers' overall indirect utility from flight and network benefits as to the *Customer Value*. This also helps to avoid confusion between indirect utility referring to network density benefits, while at the same time denoting the formal concept for partial equilibrium analysis.

In contrast to theory, network density benefits as comparative advantage have not been considered in any of the above studies: Czerny (2010) introduces indirect flight benefits based on airline flight frequency. In his case, however, benefits depend on the *overall flight volume* at an airport. And although both Brueckner (2002b) and Hong and Harker (1992) extend their analysis to airport networks, they admittedly neither account for network density effects. However, both Brueckner and Zhang (2001) and Aguirregabiria and Ho (2010) provide formalizations of the central argument of this study: Brueckner and Zhang (2001) *explicitly* model *passenger benefits* from *flight-frequency* within a network structure. They investigate a single monopolistic airline's choice problem of network type and design in a *flight-fare versus frequency* context. Aguirregabiria and Ho (2010) provide a model of a network airline's profit function, that reflects the complementarity of market entry and exit decisions across different routes. With this what they call *supermodular* profit function, they enable the model to „incorporate the entry deterrence motive“ for networks. Both these studies thus provide theoretical support for this study. Nevertheless, both only consider network effects as an optimization problem for one single airline, and thus abstract from competition and from regulation.

3 Model

In the following, first, a simple but innovative, single-airport two-airline model is presented. The model is based on Brueckner (2002a), but is modified to allow for product differentiation between the two airlines. More specifically, product heterogeneity is introduced by means of *additional, indirect passenger benefits* from *network density* of one airline. Because this modification enables to capture *network density effects,* the model is then used to investigate different airport capacity allocation schemes, for the case where airlines are asymmetric, and one of which is a network carrier. The goal of this model is to capture the basic properties of *product heterogeneities* based on *network density benefits* in a stylized airport-airline model, in order to explain airport demand with a *network airline.* Its unique contribution thus is to introduce horizontal product differentiation motivated by product quality, that ultimately leads to an airline asymmetry based on demand, rather than on cost heterogeneities. This, in turn, enables this study to theoretically investigate the efficiency of different airport capacity allocation schemes at a stylized network airline's hub airport. The analysis thus takes a systemic-economic perspective. The allocation instruments considered are selected from recent studies on this topic, and are laid out in the second subsection of this chapter. Subsequently, a table is provided that summarizes all model variables on one single page for quick reference. The model equilibrium and the results of the analysis are presented in Section 4.

3.1 Setting

3.1.1 Network Structure

The model adopts the simple, stylized network structure from Brueckner (2002a), and is shown in Figure 2. It reflects one *congested* hub airport, which connects to several unspecified, *uncongested* destinations:

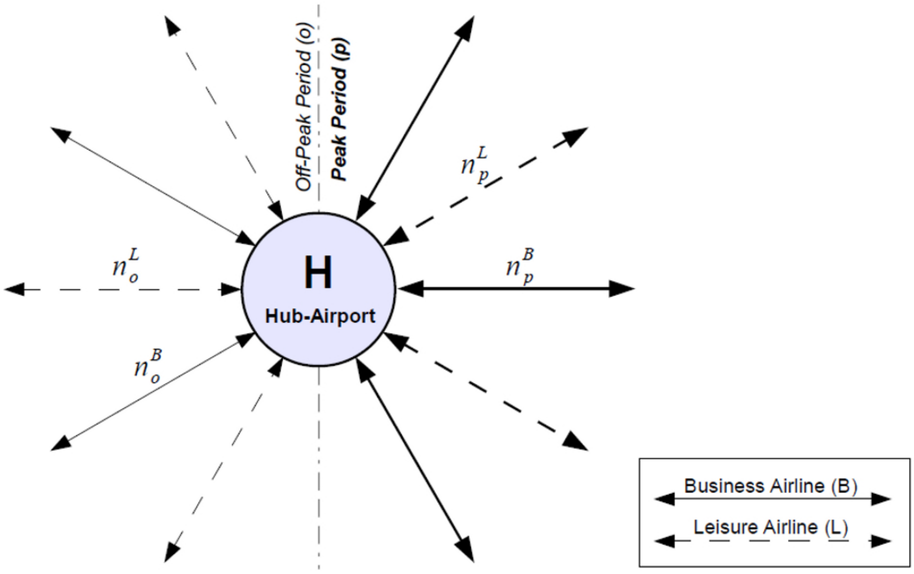

Figure 2: The Network Structure [Source: own]

The hub airport is served by *two airlines*, which operate on all routes between the hub and the destination airports. As the destinations are not specified, neither are the routes, and the number of flights of each airline is simply measured by its flight volume. In contrast to Brueckner (2002a), however, the two airlines are *asymmetric*: One is to reflect a network carrier that is able to exploit *network density benefits* for passengers. It is therefore supposed to target business travelers and thus denoted as *business airline B*. The other airline is thought of as a simple *point-to-point*, *no-frill* carrier, which does not offer advantages from network density. As it is supposed to target leisure travelers - that is, passengers with low schedule preferences - this airline is denoted as *leisure airline L*. The specification of the network density effects is given below. But to start with, one might think of the network airline to provide full travel flexibility, e.g. by allowing for free short-term re-booking across all its flights on the route or in the network, whereas the leisure airline e.g. allowed for travel on the single booked flight only. The introduction of this airline asymmetry based on network density effects represents the main innovation of this study, and it's unique contribution to recent literature.

Congestion occurs at the hub airport only, whereas all destination airports are assumed to be free of flight delays. As Brueckner (2002a, p.1360) points out, this assumption is crucial in order to abstract from cross-effects of congestion at other airports. Also, in order to investigate the allocation of flights between congested and uncongested periods, the model adapts Brueckner (2002a)'s notion that the congested hub has *peak* and *off-peak* periods: In the *peak period,* flights are subject to congestion and thus to delays. Delay is assumed to cause *time costs* to the passengers and *congestion costs* to the airlines, and is an increasing function in *total peak-flight volume* at the hub. In contrast to this, during the *off-peak period* flights do not encounter delays. Hence, for these off-peak flights, no additional costs other than operating costs accrue to the airline. This allows to observe how airlines shift their flights between periods, rather than simply going in and out of the market or adjusting total flight volume.

The investigation then is focused on the hub airport, with flight volume per period and per airline as the unit of analysis. For simplicity, thus, the model does not distinguish whether multiple flights reflect frequency on one single route, or whether every flight serves one single destination and thus their multitude represents destination choice. This simplification is maintained in order to focus on the asymmetry and to keep traceability. It could, however, be removed by model extension if desired so, allowing to analyze the airlines' choice of increasing flight frequency on one route, versus expanding its number of destination.

Of course there might be objections on how to exactly justify such a market separation in today's world, where low-cost airlines have also been engaging into network structures and targeting business travelers, and therefore traditional business models and new entrants' strategies have partly become intermixed.[11] But in fact, the interpretation of the leisure airline is left open to some extent: it may reflect a competing low-cost airline, a leisure air carrier, or a small airline operating to remote destinations. Or, for the sake of this model, it may even be interpreted as *residual supply* from multiple airlines, that offer a homogenous product which only contrasts to the business airline's networking flights: If each would be serving a part of the market on a route basis, this would reflect imperfect competition against the network-airline, but with different flights from distinct but identical airlines, and thus with residual supply as a homogenous product without network effects. Thus, while this issue generally is of importance, it might be abstracted from in this model by the before mentioned *residual supply assumption*. This is also valid for flights that connect at the hub to serve a destination pair, and face competition from airlines serving the same destination pair but on a route connecting at its own hub outside of

[11] I would like to thank referee Sven Maertens from DLR for this comment at the GARS student researchers workshop 2012 in Bremen.

the model, or even by a direct flight, if imperfect competition is assumed on a route basis without those flights actually being reflected within the model. The micro-foundation of air travel supply and demand is explained in the following subsection.

3.1.2 Supply

Flight Volumes As shown in Figure 2 above, the peak- and off-peak flights of the two airlines are denoted by subscript o and p, respectively. With uppercase letters used for *aggregate* variables across airlines, the *number of flights* during the *peak* period then is

$$N_p = n_p^B + n_p^L, \tag{1}$$

where n_p^B and n_p^L denote the individual number of flights of each airline during that period. Accordingly, the aggregate *off-peak* flight volume is

$$N_o = n_o^B + n_o^L. \tag{2}$$

Flight volume relates to output in terms of the number of passenger by introducing s as the number of seats per flight. Then, total peak and off-peak-passenger volumes are

$$s \cdot N_o \quad \text{and} \quad s \cdot N_p. \tag{3}$$

For simplicity, the seat load factor assumed to be 100%,[12] and the number of seats per aircraft is held constant and symmetric (as in Brueckner, 2002a). For the equilibrium analysis, then, it can be normalized to unity without loss of insight. Relaxing these restrictions would allow consideration of the airlines' choice problem of flight frequency versus aircraft size, and letting s differ across airlines would enhance the airline asymmetry. However, for traceability of the analysis and to focus on the newly introduced airline asymmetry in the first place, I abstract from these interesting complications.

[12]The seat load factor determines the occupancy of the seats offered per aircraft. This condition ante-cedes the market clearing assumption in the equilibrium computation section (cf. 4.1).

Airline Profits The two airlines behave as profit maximisers. The formulation of airline i's profits Π^i is straightforward and also follows Brueckner (2002a). Modified for the current notation, for $i = \{B, L\}$ it involves s as the number of seats per aircraft, f_o^i and f_p^i as the flight fare dissociated by airline and by period, and generic functions $C_i(N^i)$ for operating costs and $G(N_p)$ for total congestion costs, and yields

$$\Pi^i = s \cdot (f_o \cdot n_o^i + f_p^i \cdot n_p^i) - C_i(n_o^i + n_p^i) - n_p^i \cdot G(n_p^i + n_p^j) \quad \text{for} \quad i, j \in \{B, L\}\,. \tag{4}$$

With heterogeneous goods, fares not only differ between the peak- and off-peak periods, but also across airlines. Still, the profit functions are symmetric, as the network density benefits arise from demand. Congestion costs depend on the overall volume of peak traffic, but only affect the airline via its own peak-flights. In order to concentrate on the network benefits and not to get distracted with different functional specifications for operating costs, let assume constant and symmetrical marginal costs $C_i'(n_o^i + n_p^i) = c$ for $i = \{B, L\}$, and thus $C_i(n_o^i + n_p^i) = c \cdot (n_o^i + n_p^i)$.[13]

The supply-side first-order conditions for the equilibrium are then the respective partial derivatives of the airlines' profit functions for peak and off-peak flights. Notice, however, that the equilibrium conditions depend on the assumption about the market structure: In a perfect competition situation, airlines optimize their output taking market prices as given. With market power, however, their output also endogenously affects the flight fares, and thus price-elasticity of demand needs to be taken into account (see 4.1).

3.1.3 Demand

Above all, passengers are assumed to be utility maximizers. They are presumed to maximize utility from flight services, and from residual consumption of a numéraire good. This abstraction follows Brueckner (2002a) and most theoretical models in the field (e.g. Brueckner, 2002b, Zhang and Zhang, 2006, Czerny, 2010, Verhoef, 2010). The justification is that with one numéraire and one remaining good, demand for the latter only depends on its own price,

[13] Whether airline operations may exhibit economies or dis-economies of scale is controversially discussed in literature. Therefore, increasing and decreasing marginal costs would provide an interesting topic for further research, and their impact could actually be investigated with this model, simply by putting up the corresponding assumptions about the form of the generic cost function.

and on consumer income (Jehle and Reny, 2011, p.50). This means that for the analysis, we obtain the *indirect utility function* that only contains prices and flight benefits, but is not concerned with explicit utility levels. This substantially simplifies the analysis, as the indirect utility captures the relation between relative prices and income *under the premise* that *utility is maximized*. That is, it represents a „maximum-value function corresponding to the consumer's utility maximization problem" (Jehle and Reny, 2011, p.28).[14] With residual consumption as a single numéraire good, however, the model is simplified to a reflect *partial equilibrium* only. This means that it does not consider the general equilibrium of the economy as a whole, but only of this sector. Of course, the intention behind this is to obtain a traceable analysis of the problem at hand. Simplicity is achieved because this abstraction suspends substitution effects across other sectors of the economy, and makes the entire income effects to be captured by the numéraire good only. The main objective of this is that the consumer surplus becomes an „appropriate measure of welfare change", as it „corresponds directly to the indirect utility function" (Vives, 2001, p.77). Of course, this advantage comes at the cost of reduced precision. However, as Vives (2001, p.77) points out, partial equilibrium analysis of an industry is justified, if its corresponding share of the consumer's budget is small. Then, income effects are justifiably assumed to be negligible. This, in turn, justifies a utility function that is linear in income, as well as the representation of the remainder of the economy as an „aggregate numéraire" good. In other words, linear utility and an aggregate numéraire justify partial equilibrium analysis (Vives, 2001, p.145). With the setting from above, evidently both is the case here, and the negligence of cross-sectoral substitution effects seems acceptable in relation to the reduced complexity of the analysis.

Consumer Taste To account for product differentiation, first of all it is essential to reflect heterogeneous consumer tastes in demand. The travel demand rationale thus is based on the discrete choice model as introduced in Brueckner (2002a), as it already accounts for such demand heterogeneities. Discrete choice models represent demand by „the collection of choices made by individuals" across a „finite set of mutually exclusive and collectively exhaustive alterna-

[14]Formally, if $u(x)$ is the utility function, that is maximized under condition that consumption of goods x at prices p cannot exceed income I, $p \cdot x \leq I$, then the *indirect utility function* is the maximum value function $v(I, p)$ that is defined as: $v(I, p) \equiv \max \quad u(x) \quad \text{s.t.} \quad p \cdot x \leq I$. One might think of $v(I, p)$ as of yielding the highest possible indifference curve (and thus utility level) that can be achieved when prices are p and income is I (Jehle and Reny, 2011, p.28). Vives (2001, p.145) refers to this *indirect utility* as to *subutility*.

tives“, and thus allow to capture *how* and *why* consumers take particular decisions - in contrast to aggregate or statistical demand quantities (Garrow, 2010, p.15). Although discrete choice models are usually used to predict decision behavior based on *empirical* data [Garrow, 2010, p.15, Ben-Akiva and Lerman, 1985, pp.2], also the analytical solution of Brueckner (2002a)'s model allows to investigate passenger demand under different conditions.[15] Finally, to adopt the basic idea of this study, Brueckner (2002a)'s demand structure is extended for network density benefits. The passengers' choices according to the general setting of this model are to travel during peak times, to travel during off-peak times, or not to travel at all. Moreover, due to the newly introduced airline asymmetry, passengers also choose which airline they fly. While both airlines generally operate during both periods, they are assumed to differ both in travel benefits and in flight fares. In order for peak travel preferences to be reflected in travel demand, *consumer taste* is denoted by a variable θ. Its value range denotes all the different consumers, and is defined as a continuum in interval $[0, 1]$. Thus, formally, $\theta \in [0, 1]$. The density of $\theta \in [0, 1]$ is assumed to be uniform, which means that the total number of passengers in the model is unity. The meaning of consumer taste then is the following: As Brueckner (2002b, p.6) puts it, low values of θ correspond to *leisure* travelers, that are indifferent about their travel period. In contrast, high values of θ pertain to *business* passengers, which valuate peak travel higher than off-peak travel, precisely due to the business nature of their trips.

Indirect Utility To setup the demand structure as sketched above, I first follow Brueckner (2002b, p.5)'s definition of utility, which is stated as $U = x + B$. B denotes *gross travel benefits*, and x denotes *residual consumption* of the numéraire good. Because the numéraire good has price unity, then, for any given income I residual consumption corresponds to $x = I - f$, where f denotes the flight fare. Substituting this term into the utility equation from above yields $U = I - f + B$. Then, if income is treated as constant, for the analysis it is sufficient to consider the terms of the gross travel benefits and the flight fare, $B - f$. This term then corresponds to the *indirect utility* function, and the setting allows for partial equilibrium analysis (as explained above). In contrast to Brueckner (2002a), however, *gross* flight benefits B in this indirect utility

[15] As Garrow (2010, p.3) points out, discrete choice models have become increasingly important in aviation applications, since they reflect clear-cut alternatives of choice for individuals. Methodically, Brueckner (2002a) represents a *multinomial logit* (MNL) choice model, as it considers three exclusive choices for each customer (cf. Garrow, 2010, pp.46). Product differentiation then adds a fourth option of choice by distinguishing peak-flights between the two airlines, but does not change the model type on methodical grounds.

term now include both *direct* an *indirect* benefits. In order to capture the entire width of utility from flights in this particular model, and according to the foundations from literature as pointed out above, the indirect utility term is referred to as *Customer Value*. With indices left out at first for simplicity, customer value is hence defined as:

$$CV \equiv B - f. \tag{5}$$

Gross Travel Benefits Now, as already briefly stated above, the particularity of this model is that *gross travel benefits* from flight services as reflected in eq. (5) include *two* components: a *direct* and an *indirect* one. On the one hand, benefits to the passengers arise from the bare transportation services of a flight. This type of benefits is straightforward, and is referred to as *direct flight benefits*. On the other hand, travelers are supposed to have *indirect* benefits from the *network density* of their airline. This means that they obtain additional utility from the *supply* of multiple flights at short time intervals on the same route (or, more generally, in the network). Within the simple network structure of this model, this feature can simply be captured by the *flight frequency* on a route. Simply put, this corresponds to the number of flights of an airline during one period. A high frequency then allows both a wide choice of flights for booking before travel, and moreover a flexible schedule change during the day of travel. The latter may be attractive if there are irregularities that require to change flights, or simply if travel plans change on short notice. The important point is that those *indirect benefits* from *network density* are presumed to matter only to passengers that are schedule-sensitive, i.e. that are assumed to have a high value of time and distinctive schedule preferences. In this setting, these two characteristics are not modeled specifically, but reflected by increasing benefits from travel at peak times. Also, only the *business* airline is presumed to be able to provide such benefits from network density, because by assumption it is the only carrier to reflect a network structure. Although the network is of course stylized to the limit in this model, one might argue that the leisure airline by definition targets leisure rather than business travelers (see above), and thus does not try to optimize its schedule with respect to flight frequency. Moreover, one might argue that in the latter case, travelers were detained from changing flights on short notice by high penalty fees, or simply not allowed to re-book.

The conception of gross travel benefits B in (5) therefore is supposed to be as follows: On the one hand, they consist of *direct flight benefits* that arise from the „consumption“ of flights. But on the other hand, as well they should include *indirect benefits* from *network density*. While *direct flight benefits* are replicated from Brueckner (2002a)'s original model, *indirect* benefits from *network density* are not accounted for in the latter, and thus need to be introduced on grounds of this model. The direct and indirect travel benefits are specified in the following.

Direct Flight Benefits Let first consider *direct flight benefits*. Direct travel benefits differ across the two periods, and are defined as a function $b_i(\theta)$ of consumer taste θ, where index $i = \{o, p\}$ denotes the respective period. The two *direct travel benefit* functions $b_o(\theta)$ and $b_p(\theta)$ are depicted in Figure 3 below.[16] First, the graph shows that both functions increase with consumer's preference for peak travel. Second, it shows that business travelers (associated with a high θ) value peak travel more than off-peak travel - and that the opposite is true for leisure travelers. This depicts Brueckner (2002b, p.6)'s notion that business travel were a „crucial job requirement“, and therefore „both peak and off-peak travel benefits should be high“ for business relative to leisure travelers. Moreover, it shows the property that peak travel benefits have to increase relative to the off-peak benefits, when θ is increased. This is because business travel had „to occur during the early and late peak hours to avoid disruption of the work day“. Formally, these characteristics of the direct travel benefits are put down by Brueckner (2002a, p.1361) as the two assumptions $b'_p, b'_o > 0$, and $b'_p > b'_o$ (for all θ). The latter is referred to as the „single crossing assumption“, which gives rise to the „natural property“ that high-θ-individuals travel at peak times, whereas low-θ-passengers are off-peak travelers. As Brueckner (2002a, p.1366) puts it, the importance of the relation $b'_p \gtreqless b'_o$ is as follows: For $b'_p > b'_o$, the *marginal* increase of *direct flight benefits* is larger for peak than for off-peak flights. With $b'_p < b'_o$, the reverse is true. The latter case were somewhat counter-intuitive, as the passengers' increasing preference for peak travel would reverse, in the way that preferential peak-travelers would prefer off-peak-travel and vice-versa. For $b'_p = b'_o$, the peak-/off-peak assignment of passengers were a „matter of indifference“. This model thus takes over the assumption $b'_p > b'_o$ which seems more plausible.

[16]For this example, linear functions are chosen, although linearity is not required in the conditions imposed below.

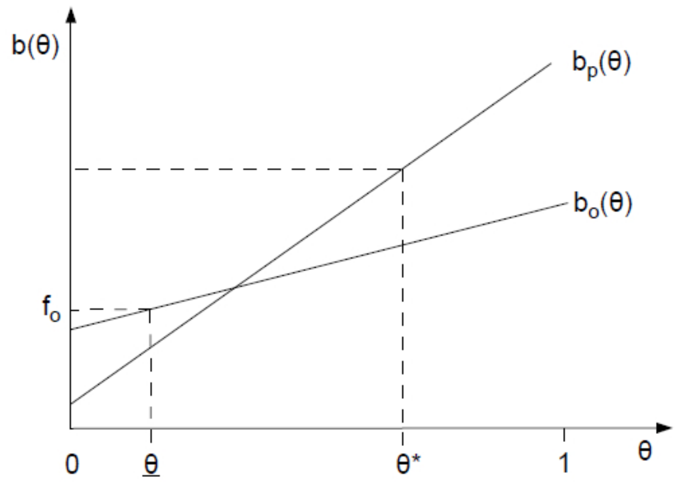

Figure 3: Direct Travel Benefits from Brueckner (2002a) [linear example; own illustration]

From the functions in Figure 3 and the corresponding assumptions from above, we can graphically infer the *characteristic values* of consumer taste θ: θ^* depicts the „relevant" consumer that travels at peak, and $\underline{\theta}$ depicts a lower bound below which individuals do not travel at all. Due to condition $b'_p,\ b'_o > 0$, it follows that all passengers between $\underline{\theta}$ and θ^* *strictly* travel off-peak, where those to the right of θ^* strictly travel at peak (Brueckner, 2002a, p.1362).

Indirect Flight Benefits As Belleflame and Peitz (2010, p.554) put it, „a good exhibits positive network effects when a consumer's willingness to pay for the good depends positively on the size of the corresponding network". Therefore, *indirect* passenger benefits are presumed to arise from network density of the business airline, where density is simply represented by it's peak-flight volume. To model network effects, Belleflame and Peitz (2010, p.554) propose a generic function of the form $U_{ij} = a_i + f_i(n_j)$. In this term, U_{ij} is the „utility to consumer i from belonging to network j", and this utility is assumed to be composed of the two following, additively separable benefits: Variable a denotes the so-called *stand-alone benefit* that arises from „immediate use" of the good. Function f denotes the *network benefit*, that is assumed to be zero for a zero-size network, and rise with the number of network users n_j. Formally, thus, $f(0) = 0$ and $f' > 0$. It is also important to notice that all variables are indexed with the consumer index i, which means that *both* of these benefits can *differ across users*. That is, *both* the stand-alone benefit and the network benefit can depend on consumer taste. Now, the above generic term can be equivalently rewritten to fit this model's variables as $u_i = b_i + d_i(n_j)$. This makes immediately evident two features: First, the first part of this generic term corresponds to the specification of indirect utility from above, where b directly refers to the *direct flight benefits* from above, and index i signifies that flight benefits and thus indirect utility depend on consumers' heterogeneous tastes. Second, it shows that network benefits can be included in that indirect utility function by adding them as a separable term, as suggested by Belleflame and Peitz (2010, p.554), with$d_i(n_j)$ denoting the projected *indirect network benefits*, that depends both on consumer taste and on network density n_j. Recalling that heterogeneous taste is explicitly allowed to affect *both* the stand-alone and the network benefit, it also seems plausible that customers that increasingly like peak travel should also increasingly like network density benefits. And as network density in this model is approximated by the business airline's peak flight volume, n_j exactly corresponds to n_p^B in this model. Finally, consumer index i can equivalently be replaced by consumer taste θ without change the meaning of the equation at all. Thus, the above generic function can be translated to fit the notation of this model as

$$u_p^B(\theta) = b_p(\theta) + d(\theta, n_p^B) \tag{6}$$

where index p denotes the *peak* travel period and superscript B denotes travel with the business airline. In other words, thus, enjoying network benefits in this model means being a user of the business airline's *network of peak-period flights*. This, in turn, is simply achieved by buying any of the business airline's peak-flight tickets. In order to account for the discrete-choice heterogeneous demand, thus, a third critical consumer θ^D is introduced. It denotes the traveler that is indifferent between *peak travel* from the business airline *with* density benefits (but presumably at a higher price), and a *peak flight* from the leisure airline *without* density benefits. Notice that indirect density benefits are *not* included in the direct flight benefits function. Therefore, evidently, θ^D cannot be inferred from the direct flight benefits function, and is thus not depicted in Figure 3. The concept of network density and its corresponding functions are further developed in subsequent Section 3.1.4.

Specification of Gross Travel Benefits Let now relate these *direct* and *indirect* travel benefits to the *gross travel benefits* as presented in eq. (5). Because congestion only occurs in the *peak* period by assumption, gross travel benefits need to be distinguished between the *peak* and the *off-peak* period. In that, this model again follows Brueckner (2002a). During the off-peak period, congestion does not occur. Therefore, no time costs arise to the passengers. Moreover, airlines do not dissociate. Therefore, for travel during the *off-peak period*, gross travel benefits are simply equivalent to direct flight benefits:

$$B_o(\theta) \equiv b_o(\theta). \tag{7}$$

In contrast, peak-period gross travel benefits dissociate from the off-peak based on two additional components: First, during the peak-period, flights are subject to airport congestion. Congestion directly translates into flight delays, and the latter is in turn assumed to cause passenger time costs. This model adopts Brueckner (2002a)'s specification where time costs are *non-decreasing* and *convex*, and have a „*positive range*" that is „*relevant*". Time costs are denoted as a function $t(N_p)$ that directly depend on the *aggregate* number of flights during the peak period, N_p. Notice, however, that time costs do *not* depend on consumer taste, and thus are identical both across passengers and across airlines.[17] G*ross travel benefits* from the *leisure* airline's flights during the *peak period* thus are composed of *direct travel benefits* from transportation minus *time costs* from congestion and read

$$B_p^L(\theta, N_p) \equiv b_p(\theta) - t(N_p). \tag{8}$$

Second, as already put down above, the peak flight supply of the business airline is supposed to create additional, indirect travel benefits. This, in turn, means that gross travel benefits dissociate *by airline* during the peak period. Use of the above specification of the *indirect* flight benefits then yields the gross travel benefits for passengers of the *business* airline as:

$$B_p^B(\theta, N_p, n_p^B) \equiv b_p(\theta) - t(N_p) + d(\theta, n_p^B). \tag{9}$$

This again makes clear that airlines do not dissociate by *direct* travel benefits, but only by presence (or absence) of *network density benefits*.

[17]This assumption might be considered as too simple. But as Brueckner (2002b) points out, variable time costs (according to consumer tastes) do not add to understanding of the problem, while they increasingly complicate the analysis. Brueckner (2002a, p.1370) considers *non-separable* time costs and finds these to „temper the results of the analysis without overturning its main lesson". Nevertheless, more realistic time costs differentiated across travelers might be added later as an extension.

Linking Supply and Demand The respective flight volumes that accommodate demand of the above three traveler groups can be rewritten in terms of flight volume by use of the specification from (1) to (3). The discrete variables of flight output directly relate to the continuous spectrum of consumer taste $\theta \in [0,1]$ as:

$$s \cdot n_p^B = 1 - \theta^D, \quad s \cdot n_p^L = \theta^D - \theta^* \quad \text{and} \quad s \cdot N_o \equiv s \cdot (n_o^B + n_o^L) = \theta^* - \underline{\theta}. \tag{10}$$

Re-arranging these equations then yields the three *characteristic values of consumer taste*, $\underline{\theta}$, θ^* and θ^D, in terms of flight volume as:

$$\theta^D = 1 - n_p^B \cdot s, \qquad \theta^* = 1 - s \cdot (n_p^B + n_p^L), \quad \text{and} \qquad \underline{\theta} = 1 - s \cdot (n_o^B + n_o^L + n_p^B + n_p^L) \tag{11}$$

with $n_o^B + n_o^L + n_p^B + n_p^L = N$. These equivalences will be needed when relating flight volume supply to the characteristic consumers, and will enhance understanding of endogenous demand within the network density effects functions.

3.1.4 Network Density Effects

In order to realize the above concept of *Network Density Effects* (NDE) as sketched in the introduction, I define three consecutive functions: the network density function, the density benefits function, and the network value function. The first two functions are depicted in the graph in Figure 4. The network value function is explored below. For linearity, a graphical representation is shown later in Figure 5.

First, at the airline level, the *network density function* denotes network density of the business airline in relation to its number of peak flights. It is defined as

$$D(n_p^B) \in [0, \delta] \tag{12}$$

and depends directly on its peak flight volume. The value range of network density is defined as $[0, \delta]$. The rationale behind this function again is the following: The business airline's network density is determined by its peak flight volume. Network density ranges from 0 if peak-flight volume is zero, to a parametric maximum valueδ, that is achieved when the business airline holds the entire market share of peak flights. The function $D(n_p^B)$ is assumed to be monotonous, continuously differentiable and strictly *increasing* in its argument n_p^B. Formally thus $D'(n_p^B) > 0$. Notice that although n_p^B is not theoretically limited, network density is only defined over the interval of $\theta^D \in [\theta^*, 1]$, because it is applicable to *peak flights* only. For this, notice that due to the equivalence $\theta^D = 1 - n_p^B \cdot s$ from eq. (11), network density can also be written as $D(\frac{1-\theta^D}{s})$. This makes clear that the domain of definition for $D(n_p^B)$ corresponds to the possible range of θ^D. It is thus is naturally restricted to $\theta^D \in [\theta^*, 1]$, and network density obviously is zero for $\theta^D = 1$.

Second, the *indirect network density benefits* that arise to the passengers from the business airline's network are accounted for, based on the generic density benefits function (6) from above. On behalf of this, I define the *Density Benefits Function* for passengers as a two-dimensional function

$$d\,(\theta, D) \in [0, D], \tag{13}$$

that depends both on consumer taste θ and on the business airline's network density function $D(n_p^B)$ from above, which itself is a function of θ^D. Let also assume that each dimension of this function is still continuously differentiable, monotonous and strictly increasing in its arguments. The two above functions are represented in Fig.4, and their interpretation is the following: While $D(n_p^B)$ denotes network density, $d(\theta, D)$ denotes the *value* that accrues to the passengers from this network density. This value depends both on the network size of the business airline's peak flights, and on the respective consumer's preference for peak travel. More precisely, the value that the passengers acclaim from this function starts at a fraction of $D(n_p^B)$ for the first passenger that travels peak-business. Further on, it increases with peak-travel preference θ to the full value $D(n_p^B)$ within the range $[0, \delta]$ for the highest-θ-passenger, as determined by (12). From a game-theoretic perspective, thus, once the maximum level of network density is determined, its value to the passengers then reduces to a one-dimensional function increasing with θ. In other words, network density and thus maximum density benefits are fixed once the business airline has determined its peak flight volume. This is the reason why, formally, network density benefits to the passengers in (13) are no longer written as a function $D(n_p^B)$

but as a value D. Notice that D only depends on the business airline's peak flight volume, and therefore function (13) is consistent with its generic predecessor from (6). Notice that $d(\theta)$

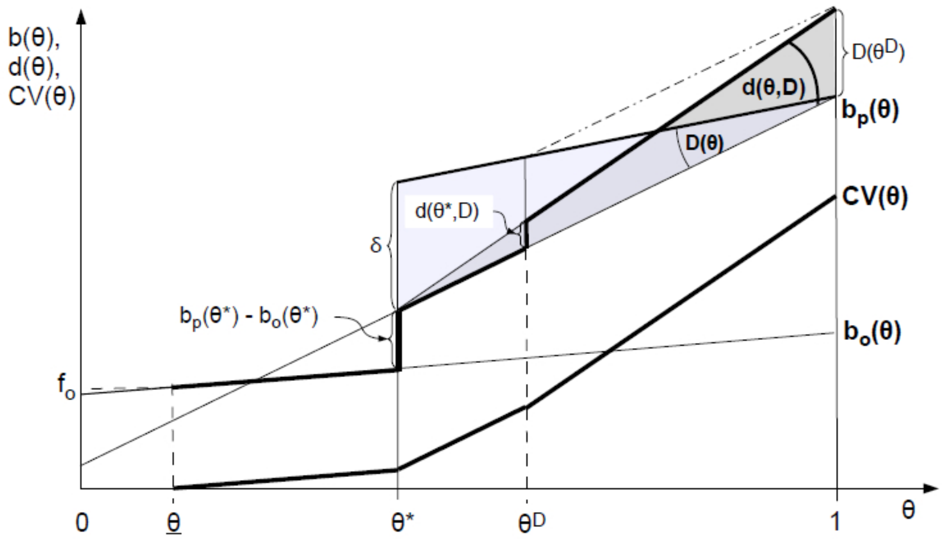

Figure 4: Travel Benefits and Customer Value with *Network Density Effects*
[linear case; own illustration]

increases with θ, but that $D(n_p^B)$ increases with the number of peak flights and is thus decreases in θ by definition. Therefore, these two functions are inversely related in the graph. Overall, the two network effect functions integrate the concept of network density effects into the partial equilibrium model with heterogeneous goods as intended by this study.

Third, again from the airline perspective, these density benefits create additional customer value that is used for pricing of the peak-flights. For this, I introduce a *network value function*, that expresses the value of the passengers' network density benefits in terms of the business airlines' potential pricing markup. This function is defined as

$$V(\theta^D) \equiv d(\theta^D, D) \quad \text{for all} \quad \theta^D \in [\theta^*, 1]. \tag{14}$$

In other words, this value function denotes the amount of the density benefit to the passengers at each potential peak flight volume that θ^D that the airline might choose. This is important because, as we will see later in eq. (22), with endogenous pricing the peak flight premium is determined by the density benefit of the traveler at θ^D, and therefore ultimately by the business airline's peak flight volume. Thus, (14) is the objective function to maximize the business airline's peak premium. It therefore designates the *market value* of the network density to the business airline. This simply means that the density benefits function is evaluated for its maximum, based on the choice of θ^D . This reduces its dimensionality to unity, because for this problem the condition is $\theta = \theta^D$. Notice that the network benefit for all peak-density travelers strictly increases to the right of θ^D. This part of the benefit, however, is part of the consumer rent. Hence, with discretionary pricing, it cannot be commercialized by the airline. The value function is left out in Figure 4 for clarity, but is explicitly shown and analyzed in Figure 5.

3.1.5 Network Density Effects under Linearity

Under the assumption that all relevant functions are strictly linear, some calculus can be performed to reveal parametric solutions for the network effects functions. This serves to illustrate the basic properties of the network density effects. For this, the graph in Figure 4 shows how the two network effects functions can be expressed in terms of the characteristic values of consumer taste. As already mentioned, this requires that the *direct travel benefits* function, the *network density* function as well as the *density benefits* function are strictly linear, exactly as drawn in the sketch of Figure 4. Then, simply by applying the basic theorem of intersecting lines, it is straightforward to express the network density function as

$$D(\theta^D) = \delta \cdot \frac{1-\theta^D}{1-\theta^*}. \tag{15}$$

Consequently, the network density benefits function can be written as

$$d(\theta, D) = \frac{\theta-\theta^*}{1-\theta^*} \cdot D(\theta^D). \tag{16}$$

This shows that the linear density benefits function has the constant slope $d'(\theta, D) = \frac{\partial d}{\partial \theta} = \frac{1}{1-\theta^*} \cdot D(\theta^D)$, which by substituting 15 becomes $d'(\theta, D) = \delta \cdot \frac{1-\theta^D}{(1-\theta^*)^2}$. Moreover, under the above assumption of linearity, also the value function can be specified and expressed in terms of the characteristic θ's. By use of (16) in (14), and imposing condition $\theta = \theta^D$, the value function becomes

$$V(\theta^D) = d(\theta^D, D) = \delta \cdot \frac{\left(1-\theta^D\right)\left(\theta^D-\theta^*\right)}{(1-\theta^*)^2}. \tag{17}$$

Now, the functional form of the value function can be explored by basic analysis. Its general properties are depicted in Figure 5:

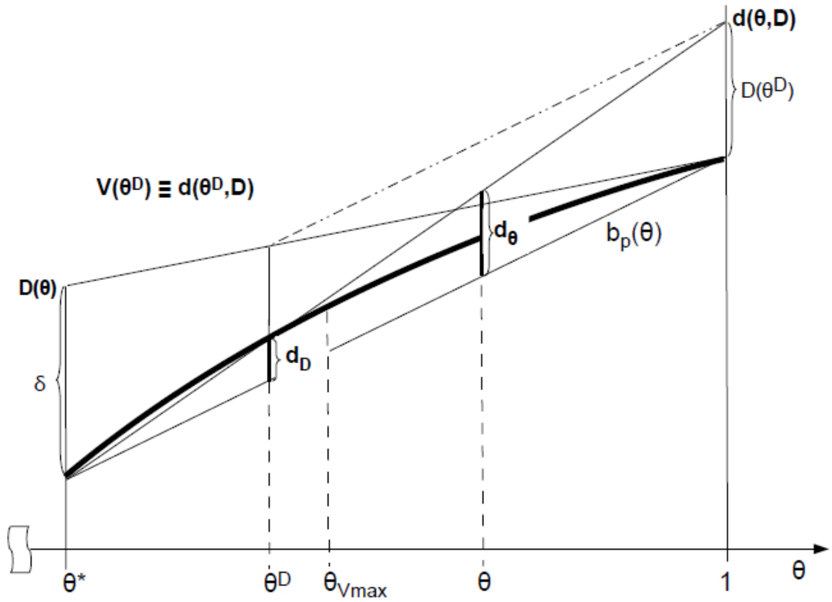

Figure 5: Value Function of Network Density Benefits [linear case; own illustration]

As eq. (17) shows, the *network value* is a quadratic and thus quasi-concave function of θ^D. The local maximum of the network value function is determined by $\partial V \partial\theta^D = \delta \cdot \frac{1-2\theta^D+\theta^*}{(1-\theta^*)^2} = 0$, where solving this equation for θ^D yields the argument that maximizes the value function as

$$\theta^D_{Vmax} \equiv \underset{\theta^D}{argmax}\, V(\theta^D) = \frac{1+\theta^*}{2}$$

In other words, the maximum network value is achieved when $\theta^D = \frac{1+\theta^*}{2}$. Recall that the network value at $\theta^D$is equivalent to the peak-density premium for the business airline. Hence, substituting this value into eq. 17 then yields this maximum network value, which is

$$V_{max} \equiv V(\theta^D_{Vmax}) = \frac{\delta}{4}.$$

This means that despite the maximum network density than can be achieved in the model is δ, which is reached when $\theta^D = \theta^*$, the highest value in monetary terms that the airline can extract from its customers based on network density is $\delta/4$. If the airline chooses a lower θ^D than θ^D_{Vmax}, then it increases network density, but the density benefit for the critical passenger θ^D falls, and with it the peak-premium. Similarly, if it chooses to increase θ^D by lowering its peak-flight volume, it reaches a critical customer with a higher willingness-to-pay, but at the sime time network density decreases, which more than offsets the price-elasticity of demand. Thus, also in this case, the peak premium falls. Nonetheless, the airlines' overall objective function are total profits rather than the network value. And the former also depend on the direct flight benefits and on time and congestion costs, and on flight volume in terms of output. In other words, the local maximum of network value of course only maximizes the business airline's peak premium, but not necessarily its total profits.

Still, these results show two fundamental properties of the product differentiation model with network density effects: First, they show that the maximum markup is achieved at $\theta^D = \frac{1+\theta^*}{2}$, which is not halfway between the peak-off-peak and the peak density split but at some arbitrary, asymmetric value. And because the function is quasi-concave, it is reasonable to assume that this is an interior solution. In other words, the specification of the network effects from above is able to produce an asymmetric equilibrium with an interior solution. Abstracting from linearity, of course the maximum value may shift. For „reasonable" functions that refrain from showing extreme non-linearities, however, it should be reasonable to assume that both the asymmetry and the interiority of the equilibrium remain. Second, the results show that the maximum network value that the airline can potentially commercialize is *one fourth* of the maximum possible network density. In this model where network density incurs no other costs than direct operational costs for the peak flights, this seems less important. However, imagine a situation where network costs occur to install and maintain the network services. Then this might make it questionable whether offering network services might be worthwhile at all, at least under the current network benefit specifications.

3.1.6 Customer Value

Ultimately, the specification of the *Customer Value*-Function as described in (5) is achieved as follows: Dissociating for peak and off-peak travel, substituting travel benefits B for Brueckner (2002a)'s respective *direct travel benefit*-Functions $b_p(\theta)$ and $b_o(\theta)$ from above yields off-peak customer value function

$$CV_o(\theta, f_o) \equiv b_o(\theta) - f_o. \tag{18}$$

As explained above, the off-peak customer value function is quite simple because there are neither congestion nor density effects, and because off-peak flights are homogenous products with the single off-peak travel fare f_0. Furthermore adding time costs $t(N_p)$ and the network benefits function $d(\theta, n_p^B)$ from above, and dissociating travel fares f_p^L and f_p^B for the peak flights of the two *asymmetric* airlines yields

$$CV_p^L(\theta, n_{p,}^L n_p^B, f_p^L) \equiv b_p(\theta) - t(N_p) - f_p^L \tag{19}$$

and

$$CV_p^B(\theta, n_{p,}^L n_p^B, f_p^B) \equiv b_p(\theta) - t(N_p) - f_p^B + d(\theta, n_p^B). \tag{20}$$

These terms denote the *Customer Values* for *peak-flights* as functions of *consumer taste*, *flight fares,* and *flight volumes* of both airlines in both periods. As explained above, $N_p = n_p^L + n_p^B$ is the total volume of peak flights and is important for congestion and time costs. Note that $b_p(\theta)$ is not distinguished by airline, as it denotes the *direct* flight benefits for passengers during the peak, which are identical for both firms and only depend on the consumer taste. Equations (18) to (20) thus represent a demand system of discrete choice under heterogeneous tastes.

Notice how the network density benefits now create an airline asymmetry: They introduce product differentiation, in that the network airline (in contrast to the direct airline) can offer *indirect* density benefits to its customers. Moreover, this differentiation is *endogenous*, by means that the network airline can *endogenously* choose the size of this network effect, based on its flight frequency during the peak period. This heterogeneity, in turn, implies that travel prices differentiate away from the single, uniform market price that were to prevail under a homogeneous goods case. Consequently, also the Customer Value functions need to be dissociated with regard to peak fares and indirect density benefits. Evidently, hence, at this point the model substantially departs from Brueckner (2002a)'s original contribution, by introducing *network effects* and thus generating an *airline asymmetry* - following the latter's own proposition of a

„useful" extension for future research.

A graphical representation of customer value that relates to the flight and network density benefits is depicted above in Figure 4. As the indifference conditions imply (see below), it is a continuous function that strictly increases with θ. However, as the graph shows, due to the demand system characteristics of the discrete choice model, it is not continuously differentiable. This problem will be addressed just below. That is, with each additional benefit, its gradient increases exogenously: First, the low gradient on the left reflects direct benefits from off-peak travel. When characteristic consumer θ^* switches to peak-travel, the steeper peak-travel benefits come into play. But obviously, the customer value function needs to be continuous in θ^*, as this traveler is exactly indifferent between peak and off-peak travel. In other words, the higher benefits at this point need to be compensated away by both a higher flight fare and higher time costs. The same pattern repeats at point θ^D, where density benefits are introduced to peak-business travelers - but are compensated away by an even higher flight fare - the density premium for the business airline.

3.1.7 Equilibrium Conditions from Demand

In models with a continuous and strictly quasi-concave utility function, the indirect utility or value function is well defined, and a *unique solution* exists to the consumer's utility maximising problem (Jehle and Reny, 2011, p.28). Notice that with discrete choice, however, the analytical derivation of the demand functions and their extremes is not helpful. This is because demand consists of multiple utility functions, and thus is not continuously differentiable, due to the discontinuities at the indifference points between two adjacent options of choice (Ben-Akiva and Lerman, 1985, p.44). This can easily be seen by referring to the demand system, eqs. (20) to (18) eq:CVo. In contrast, the equilibrium conditions from demand are found by determining the „critical" consumers, which correspond to the characteristical values of consumer taste. This, in turn, is achieved by considering the indifference conditions between two adjacent options of choice.

Starting from the left in the $\theta \in [0, 1]$-continuum, let first consider the lower bound $\underline{\theta}$ which determines the lowest-θ passenger. Most evidently, as soon as there is a positive customer value in conjunction with buying a flight, an individual will travel. Hence, the lower bound is determined by $CV_o(\theta, f_0) \geq 0$. Let now in addition assume, that if an individual is indifferent between travelling and not travelling, because both yields zero customer value, it will travel. Then, equation (18) implies that the equilibrium condition for $\underline{\theta}$ is

$$b_o(\underline{\theta}) = f_o. \tag{21}$$

This relation is depicted in Figure 3. Recall that all passengers to the right of $\underline{\theta}$ will also travel and have a customer value larger than zero, and individuals to the left will not travel, because their direct travel benefit in monetary terms is less than the flight fare. Next is the „relevant consumer“ θ^* that will switch from off-peak to peak-travel, because to him peak-travel offers a higher customer value, despite eventual congestion and possibly a higher flight fare. Obviously, this passenger is determined by his/her indifference between peak and off-peak-travel. Now as this model differentiates peak-flights between the *leisure* and the *business* airline, because the latter in addition offers network density benefits (and thus most probably can charge a higher flight fare), $\theta^*$is defined as the boundary between off-peak travel and peak travel with the leisure airline.The condition to be fulfilled therefore is $CV_o(\theta^*) = CV_p(\theta^*)$, and equating (18) to (19) and re-arranging yields

$$b_p(\theta^*) - b_o(\theta^*) = f_p^L - f_o + t(N_p). \tag{22}$$

With $b_p(\theta^*) - b_o(\theta^*)$ as the surplus between the two direct travel benefit functions for the relevant peak/off-peak passenger. As Figure 3 depicts, this term indicates the vertical distance between the two direct-travel-benefit functions b_o and b_p at θ^*. Then the right side of (22) makes clear that the higher direct flight benefit from peak travel must at least compensate for the higher peak flight fare and the time costs. Ultimately, θ^D denotes the peak-period passenger that is indifferent between a flight with the leisure airline without density benefits, and a flight in the network of the business airline. This indifference is stated in demand terms as $CV_p^L(\theta^D) = CV_p^B(\theta^D)$, and from equating (19) to (20) it follows that

$$d(\theta^D, D) = f_p^B - f_p^L. \tag{23}$$

This shows that with positive network density benefits for passengers, there is a positive mark-up for *density* travel against *peak-leisure* flights. In the following this is referred to as the *peak premium.* Notice that network density benefits to not enter the direct flight benefits function. Therefore, they cancel each other out and do not appear in eq. (23), as opposed to eq. (22). Moreover notice that time costs do not enter the demand condition for *peak-density* travel. This awes to the fact that the density-travel decision is subject to the peak travel decision, and does not directly affect the total number of peak flights, and thus congestion. Indirectly, however, density travel demand might affect the leisure peak-fare, and with it the overall number of peak travelers. This is reflected in (22).

Based on the above considerations, in the following, individuals $[\theta^D, 1]$ are referred to as *peak-density* travelers, as they will fly in the peak with the business airline that offers density benefits. Passengers $[\theta^*, \theta^D]$ also travel on peak but with the leisure airline. To avoid confusion, they will be called *peak-leisure* travelers. Last, *off-peak* passengers simply remain that way, as they are not distinguished by airline. Consequently, let moreover denote the fare difference $f_p^L - f_o$ between *peak-leisure* and *off-peak* flights as the *„peak premium“*, and $f_p^B - f_p^L$ between *peak-leisure* and *peak-density* flights as the *„density premium“*.

As already mentioned above, the following consideration arises with equilibrium condition (23): Notice that b only denotes *direct* travel benefits, and that *indirect* network density benefits d are additively separable and thus do not directly affect b. This means that a *density traveler* is willing to pay the *density premium*, even if his *direct* travel benefits are not higher than from a comparable *peak-leisure* flight. Notwithstanding, it is assumed that the higher the passengers' preference for peak travel (i.e. his θ), the more he will also like - and be willing to pay for - *density travel.* However, if the *density premium* $f_p^B - f_p^L$ is excessive (i.e. if it is not compensated with an appropriate density benefit d, the peak-time traveler will refuse to fly with the business airline but rather buy a less expensive *peak-leisure* ticket with the leisure airline. In contrast, if the fare difference is not important compared to the network density benefits (according to (23)), then the peak traveller will choose the *peak-density* flight from the business airline.

The above issue invokes the question whether it is possible that *all peak demand* will go to the business airline. For this, let first consider whether $b(\theta^D) - b(\theta^*)$ could also turn out to be negative in the model. As eq. (23) shows, this were the case if network benefits were larger than the density premium (or, at the extreme, a zero premium and a positive benefit). Then, density travel would shift *to the left* of peak-leisure travel θ^*. This, however, made no sense, as by model assumption the leisure travelers do not valuate network density, but are only sensitive to the travel prices and to time costs. In other words, θ^* then would shift left with θ^D. Passengers to the left of *peak-density* demand would then simply switch to the *off-peak*. In other words, by model definition, $\theta^D \geq \theta^*$ is assured. Moreover, from (23) and the assumption $b'_p > 0$ from above, it then follows that $b(\theta^D) - b(\theta^*) \geq 0$, and thus that $f_p^B - f_p^L \geq d(\theta, D)$. This also implies that with $D(\cdot) \geq 0$, *peak-density* flights will never be cheaper than *peak-leisure* flights. Nevertheless, the business airline *can* decrease its *peak-density* fare in order to attract more leisure passengers, down to the *peak-leisure* fare. Then, it would receive all peak demand, because peak-density flights offered density benefits at the same price of the straight peak-leisure flights. And because density benefits are an increasing function of consumer taste, starting at zero for the critical peak-passenger, all travelers right of θ^D prefer peak-business over peak-leisure flights. But as pricing is non-discriminate, this would mean that the business airline's peak premium would reduce to zero. In order to maximise profits, hence, it is reasonable to conclude that it is worthwile to charge at least a slightly positive peak premium, loose some customers, but earn more from the entire rest of them. Formally, assuming that the airline's network value function exhibits some degree of concavity assures a non-degenerate, internal solution where $\theta^D > \theta^*$ is strictly valid. This is granted because density benefits are non-negative and start with $d(\theta^*) = 0$ for the critical peak-passenger. Then, concavity of the network value function $d(\theta^D, D)$ requires that at $\theta^* = \theta^D$, $d'(\theta^D, D) > d'(\theta)$. Now, because $d'(\theta^D)$ determines the change of the peak-premium, and $d'(\theta)$ the marginal utility from density benefits, this means that the price-elasticity of peak-density demand at $d(\theta^*)$ is such that a price increase *overcompensates* the output contraction, and thus increases the business airline's profit. This is valid until $d'(\theta^D, D) = d'(\theta)$. Recall that for higher values of θ^D, that target high-θ-passengers, the network density further decreases. Thus, at the upper end, despite the passengers' high preference for peak-travel, the network benefit again collapses, and with it the peak premium. With concavity, hence, an interior solution is granted. And, as the analysis in section 3.1.5 shows, even a simple linear specification of network density and passenger density benefits yields a concave network value function.

Last but not least, the three above equations allow for some basic comparative statics concerning the properties of the individuals' travel choices: Other things equal (ceteris paribus), (22) shows that θ^* increases (i.e. *less* people travel during the peak) with an increase in the peak premium or in time costs. In contrast, as explained above, the change in θ^D with $d(\theta, D)$ depends on the relative change of the density premium. Moreover, as explained above, density benefits cannot directly increase the share of peak travelers. However, if density benefits motivate the leisure airline to decrease its peak-fare, the total number of peak travelers might increase (depending on congestion), and so they still might have an *indirect* effect on θ^*. The lower bound $\underline{\theta}$, in contrast, only depends on off-peak travel-fare f_0. Still, the *number* of off-peak-travelers $\theta^* - \underline{\theta}$ is also a function of the above peak-period variables, but in the opposite sense.

3.2 Allocation Instruments

3.2.1 Quota (Airport Slots)

As already pointed out, quota regulation requires both the determination of the optimal output quantity, and the allocation of the constrained resources to the different stakeholders. And of course, optimality requires an assumption about the target function to be maximized. Inline with recent literature, this study assumes that the airport coordinator sets the number of constraints, and that his goal is to maximize socio-economic welfare.[18] Given this assumption, the *optimal number* of access rights is determined by welfare maximization with respect to total output as an argument. Czerny (2010, p.373) denotes this problem as

$$N_p^{opt} = \underset{N_p}{argmax}\,[W(N)] \tag{24}$$

where N_p^{opt} is the optimum number of slots, and $W(N)$ is social welfare as a function of total output $N = N_o + N_p$.[19] Recalling that in this model, only the peak period is congested, quota regulation only applies to the latter. Airport access in the off-peak period remains unregulated, as it does not lead to congestion. The solution to the above optimization problem simply is the social optimum, as it gives the efficient overall number of peak flights. Notice, however, that

[18] This presumption of course abstracts from the political economy debate, that questions whether the regulating agency and/or it's political supervisors rather pursuit their own interests instead of public welfare (see e.g. Button, 2005).

[19] Czerny (2010) denotes the optimal number of slots as $\hat{q}$, which has correspondingly be changed to N_p^{opt} in this model. Moreover, he applies the expected value operator to welfare, because he introduces uncertainty. This is neither necessary nor suitable here, and has been left out.

in contrast to a homogenous goods, symmetric case, now also the allocation of the quota to the airlines has a potential welfare effect. The investigation of the quota solution is provided in 4.1.1.

3.2.2 Secondary Trading

As Verhoef (2010, p.326) points out, the necessary premise for trading to take place is that the total number of available access rights is lower than *aggregate demand*. Otherwise, supply exceeds demand, and market prices are zero.[20] But under imperfect competition, the market power distortion may cause output to fall short of the social optimum. Therefore, in an oligopoly setting, we will most likely be confronted with this *aggregate demand*-problem. This means that under the above consideration, secondary trading might not take place. If, however, the initial quota endowment of the airlines does not reflect the natural market structure, trading may nevertheless occur. Thus, also the *initial allocation* of slots determines the existence of trading opportunities, as it may considerably deviate from the natural market structure. This may especially arise when airlines exhibit asymmetries, and means that even under imperfect competition, market prices need not be zero. One might think of a symmetric initial allocation but asymmetric airport demand, or an asymmetric initial allocation. Another possibility for the existence of a market equilibrium despite market power is, that aggregate output in absence of constraints is actually *above* the social optimum. This may happen if the congestion externality is large, and thus more important than the market power distortion. It is, however, unlikely to appear in a duopoly, as the analysis of Brueckner (2002a) shows (see Appendix). As far as the *initial allocation* is concerned, airport access rights can either be sold or auctioned to the airlines, or be allocated for free before trading (eg. Verhoef, 2010). This study adopts the presumption of a free allocation, and abstracts from an initial slot sale or auction.

Verhoef (2010) solves the aggregate demand problem by simply assuming that the regulator were interested to reduce *total output*. This is achieved by releasing an *arbitrary number* of slots that is *lower* than aggregate peak demand in the unconstrained market equilibrium. Consequently, the regulated total flight volume is *below* the optimal output of the social planner, and the existence of a trading equilibrium is assured. Notice, however, that this implies that

[20] Negative prices are ruled out, as long as unused slots can be handed back for free (Verhoef, 2010, p.326).

the number of slots is not optimally chosen - unless the target function differs from the maximization of social welfare. Therefore, from an allocation perspective, this study refrains from modeling such an *arbitrary scheme*.

Following the discussion above, secondary trading is investigated under the following premise: Assume that a constraint is introduced at the size of the optimum peak flight volume N_p^{opt}, which is derived from social welfare condition (24). In this case, as discussed above, aggregate demand may either be lower or higher than the constraint: If the market power distortion is prevailing, then $N_p^{opt} \geqslant n_p^B + n_p^L$. Exchange at a positive price then only takes place if the initial endowment differs from the natural market structure. In this case, we are hence left with the endowment problem. If, in contrast, the congestion externality exceeds the market power effect, total equilibrium output is *above* the social optimum, hence $N_p^{opt} < n_p^B + n_p^L$. Then, airport demand assigns a positive value to the access rights in any case. The latter is, however, not a very likely outcome in the current setting, because network density benefits do not increase peak-demand *above* the socially efficient level, but rather *raise* the optimal output.

3.2.3 Congestion Pricing

In general, an efficient congestion toll should equal the marginal externality at the optimum. It is important to notice that airlines take the toll as *exogenous*, which is deemed as „consistent with Cournot behavior" (Brueckner, 2002a, p.1367). If the tax were accounted for in the profit maximizing problem, the behavior then would simply correspond to full internalization. Now notice that the congestion externality is additively separable in the profit function (see eq. (4)). This means that it is straightforward to quantify the marginal externality, once the first-order conditions for the equilibrium under the respective market form is computed. Notice, however, that this study refrains from *including* the market power distortion into the congestion tax, in contrast to Verhoef (2010). Although by doing so, the latter is able to reproduce a second-best solution, this study's approach exactly aims at separating and investigating the interplay between market power and congestion. Following Brueckner (2002a)'s notation, hence, the congestion toll for $i \in \{B, L\}$ is denoted as

$$R^i(n_p^i). \tag{25}$$

In-line with recent literature, it is designed to equal the marginal congestion costs in equilibrium, and it is allowed to differ across airlines. Marginal congestion costs can simply be taken from the subsequent equilibrium conditions.

3.3 Table of Variables (Quick Reference)

Variables	Explanation	Reference	Page
Supply			
$N_p = n_p^B + n_p^L$	Total number *(N)* of *peak* flights *(p)*, as the sum of *business (B)* and *leisure (L)* airlines' peak flights.	eq. (1)	16
$N_o = n_o^B + n_o^L$	*Total number of off-peak flights*, as the sum of *business* and *leisure* airline *off-peak flights*.	eq. (2)	16
Π^L,Π^B	*Net Profit* of *business* and *leisure* airline, respectively	eq.(4)	17
s	*Seat number* per aircraft (normalized to unity)	”	”
$C(N^L)$, $C(N^B)$	Airlines' *direct operating costs*, as a function of each airlines total flight volume	”	”
$G(N_p)$	Airlines' *congestion costs*, depending on *total peak flight volume*	”	”
$D(n_p^B)$	*Network Density*, as a function of *business airline peak flights*	eq.(12)	25
Demand			
θ	*Consumer taste*, in continuum [0, 1] with unit density		18
$B_p(\theta, N_p)$	*Gross travel benefits* during *peak period*, as a function of *consumer taste* and *total number* of *peak flights*	eq.(8)	24
$B_o(\theta)$	*Gross travel benefits* during *off-peak period*, as a function of *consumer taste*	eq.(7)	24
$b_p(\theta)$, $b_o(\theta)$	*Direct flight benefits* (*peak* and *off-peak*), as functions of *consumer taste*	Fig. 3	22
$t(N_p)$	*Passengers' time costs*, depending on *total peak flight volume*	eq.(20)	31
$d(\theta, D)$	Passengers' *network density benefits* as function of consumer taste and of business airline's network density, with $D = D(\theta^D)$	eq.(13)	26
f_p^L, f_p^B	*Peak flight fares* from the *leisure* and the *business* airline		
f_o	*Off-peak flight fare* (identical for both airlines)		
CV_p^B, CV_p^L	*Customer Value*, from *peak flights* of *business* and *leisure* airline, respectively	eq.(20), (19)	31
CV_o	*Customer Value* from *off-peak flights* (symmetric for both airlines)	eq.(18)	31
Allocation Instruments			
N_p^{opt}	*Socially optimal constraint* on *total peak flight volume*	eq.(24)	36
$R^i(n_p^i)$	*Congestion toll* for *airline i* as a function of its *peak flight volume*	eq.(25)	38

Table 1: Table of Variables (Quick Reference)

4 Results

In the following, first, a partial equilibrium analysis of this study's model with heterogeneous products is provided. The general properties of the model are explored by computing both the social optimum and the Cournot duopoly equilibrium. The results are then compared to the homogenous product, symmetric equilibrium, in order to recognize the effects of product differentiation with network effects on the market structure. On this purpose, the symmetric results from Brueckner (2002a)'s model are briefly reviewed in the Appendix. In the second Subsection, the above results are used to investigate the impact of the different capacity allocation schemes on efficiency and competition. Because the model so far draws on generic functions, this will leave us with general conditions regarding the ambiguities from the dual distortion.
In the following analysis, for the sake of traceability, two major simplification are taken. These have already been discussed in the presentation of the model in the previous section. First, airlines' production costs are assumed to be linear, in order to yield *constant marginal costs*. Second, aircraft sizes are presumed to be fixed and symmetric. Thus, the model abstracts from the airline's choice of aircraft size versus flight frequency. Therefore, for simplicity but without loss of insight, the *seat number per aircraft* can be normalized to *unity*.

4.1 Equilibria

4.1.1 Social Optimum

The social optimum is determined by maximising the social welfare function. According to the Marshallian welfare concept, welfare consists of aggregate consumer surplus and firms' net profits (Vives, 2001, p.101).[21] Now, consumer surplus exactly consists of the customer value as expressed in (20), (19) and (18), and firms' net profits are already given in (4). Therefore,

[21] As Vives (2001, p.83) points out, the Marshallian consumper surplus is only a „good approximation" of the „true measure of welfare change", which is the Hicksian consumer. Nevertheless, he comments that if only one price changes, the error is small and in the additive separable case amounts to $\frac{1}{n}$ in percents, for n as the number of goods. This error increases, however, when multiple prices change, because the Marshallian consumer surplus depends on the *sequence* of price changes (Vives, 2001, pp.89).

the function for aggregate social welfare amounts to

$$W = \int_{\underline{\theta}}^{\theta^*} CV_o(\theta)d\theta + \int_{\theta^*}^{\theta^D} CV_p^L(\theta)d\theta + \int_{\theta^D}^{1} CV_p^B(\theta)d\theta + \Pi^L + \Pi^B.$$

Because the flight fares from the customer value and the airline profit functions cancel each other out in the above term, with s normalized to unity the welfare function simplifies to

$$W = \int_{\underline{\theta}}^{\theta^*} b_o(\theta)d\theta + \int_{\theta^*}^{1} b_p(\theta)d\theta + \int_{\theta^D}^{1} d(\theta, D)d\theta - (1-\theta^*)\cdot[t\,(1-\theta^*) + G\,(1-\theta^*)] - C(1-\underline{\theta}), \quad (26)$$

with the equivalences of flight numbers and characteristic θ's from (11), i.e. $1-\theta^* = N_p$ and $1-\theta^D = n_p^B$. Flight volumes are expressed in θ rather than in n to stress that the social planner would determine all three critical consumer taste values in order to maximize welfare. Notice that both functions $t(\cdot)$ and $G(\cdot)$ are not concerned with the integral operator, because they are fixed values evaluated at θ^* and thus no further vary with θ across the consumer continuum. The social optimum is then computed by the three partial derivatives of the above welfare function to the three characteristic θ's and equaling them to zero. With constant marginal costs $C'(\cdot) = c$, they determine the social optimum as follows: First, the derivative for the characteristic threshold between off-peak travel and non-travel is

$$\frac{\partial W}{\partial \underline{\theta}} = b_0(\underline{\theta}) - c = 0. \quad (27)$$

Because θ is non-negative by definition, $\underline{\theta} \geq 0$. The above condition hence states that with marginal costs higher than the lowest θ's direct flight benefit, i.e. $c > b_0(0)$, there is an interior solution to the lowest characetristic theta $\underline{\theta} > 0$, that is determined by marginal costs as

$$b_0(\underline{\theta}) = c.$$

If marginal costs were below the lowest θ's direct flight benefit, i.e. $c < b_0(0)$, then (27) showed that welfare increased with decreasing θ throughout the entire range of $b_0(\theta)$. This would dictate a *corner solution* where all individuals travelled, as a local maximum within the defined range of θ of the welfare function with regard to $\underline{\theta}$. In order to yield an internal solution, we therefore stick to the assumption $c > b_0(0)$. Second, the social welfare derivative for the critical *peak-offpeak-split* is

$$\frac{\partial W}{\partial \theta^*} = b_0(\theta^*) - b_p(\theta^*) - t(N_p) - N_p \cdot \left[t'(N_p) + G'(N_p)\right] - G(N_p) = 0. \tag{28}$$

Here, congestion and time costs are again denoted in terms of flight volumes rather than in consumer taste for notational simplicity. This term means that at the margin, the difference of direct travel benefits between peak and off-peak travel needs to compensate the marginal congestion and time costs from a switch between the two periods. An interior solution $\underline{\theta} < \theta^* < 1$, where peak-passengers are a positive fraction (but not equal to the number) of total travelers, is warranted if the following two assumptions are satisfied (see Brueckner, 2002a, pp.1362): On the one hand, functions b_p and b_o need to intercept at a value $\theta < 1$. Otherwise off-peak travel is always more attractive. On the other hand, also $b_o(\underline{\theta}) > b_p(\underline{\theta})$ must be satisfied. Else the welfare derivative were negative at the point $\theta^* = \underline{\theta}$ where all passengers would travel on peak, indicating a corner solution where letting some passengers travel during off-peak (and thus let θ^* rise) were welfare decreasing.

Third and new to the model is the welfare derivative to *density* travel,

$$\frac{\partial W}{\partial \theta^D} = d\left[\theta^D, D(\theta^D)\right] = 0. \tag{29}$$

This simply means that the passengers' network density benefit, evaluated at θ^D, must equal zero. Although this condition may seem odd at first sight, it denotes the two intuitive local extremes of the density benefits function. Namely, the latter equals zero in two cases: First, when the business airline holds the entire peak flight share, so that $\theta^D = \theta^*$. And second, when only the leisure airline operates in the peak period, because then $\theta^D = 1$. Both cases then yield $d = 0$. This means that either all peak passengers have to travel on the business airline, or that the market share of the business airline during peak has to be zero. The explanation and dissociation of these two solutions is straightforward: Because all flights during peak cause congestion, but the density flights of the business airline offer additional density benefits at the same costs, social welfare is *maximised* when all peak flights are operated by the business airline. In the opposite case, none of the peak-time travellers enjoys density benefits, and thus welfare is *minimized.* In other words, the social optimum is reached if the business airlines offers all peak-flights, while the leisure airline operates in the off-peak period only.

Notice that in a market solution, product differentiation yields market power to the airlines and lets them set prices asymmetrically for their peak-flights. Because then density travel causes higher costs for the passengers than peak-leisure travel, it is very likely that the peak-flight share of the leisure airline is not zero. Consequently, a duopoly market solution is not expected to replicate this social maximum. It is explored in the following.

4.1.2 Market Equilibrium: Cournot Duopoly

Cournot competition in general reflects a market with a small number of firms, that *independently* determine their output quantities. In other words, the airlines consider each other's output as given. As this model specifically represents a duopoly situation, this assumption seems to be easily warranted. The consumers, instead, behave passively according to inverse demand, so market clearing occurs by assumption. This simply means that all flights are fully booked, and thus all offered seats are consumed. Then, the situation corresponds to classical demand theory in oligopoly (cf. e.g. Vives, 2001, p.93). Let assume that the firms behave as profit maximizers. In an oligopoly situation, their output has an effect on the market prices. Thus, airlines take into account the price elasticity of demand when determining their output quantities. In other words, the price effect of total output quantity is reflected in airlines' profit maximization. This is achieved by relating the flight output variables from supply side eq. (4) to the characteristic values of consumer taste from the demand side, by substituting the respective terms from equations (11). The optimization problem of the two airlines $i \in \{B, L\}$ thus reads:

$$\max_{n_o^i, n_p^i} \Pi^i \quad \text{s.t.} \quad f_o = f_o(n_o^i, n_p^i) \quad \text{and} \quad f_p^i = f_p(n_o^i, n_p^i). \tag{30}$$

In order to compute the equilibrium, thus, beforehand the *price functions of demand* are derived. These express market prices as functions of the output quantity. For this, demand functions in eq. (21) to (23) are inverted, re-arranged and cross-substituted to yield:

$$f_0(\underline{\theta}) = b_0(\underline{\theta})$$
$$f_p^L(\theta^*) = b_p(\theta^*) - b_o(\theta^*) + f_0(\underline{\theta}) - t(N_p)$$
$$f_p^B(\theta^D) = d(\theta^D, D) - f_p^L(\theta^*) \quad (31)$$

Then, for the derivation of the profit functions that includes price elasticity of demand, the characteristic values of θ in the above equations are rewritten in terms of airlines' output variables n_o^i, n_p^i for $i = \{B, L\}$. Thus, substitution of conditions (11) in (31) yields the *inverted demand functions* that describe flight fares as endogenous variables of output. These are equivalent to 31 but notationally more cumbersome, and therefore not shown . Subsequently, inverted demand is substituted into the airline profit functions from (4). This yields airline profits with endogenous prices, $\Pi^i\left[n_o^i, n_p^i, f_0^i(n_o^i, n_p^i,), f_p^i(n_o^i, n_p^i,)\right]$ for $i = \{B, L\}$ (not shown). Moreover, for the derivation of these profit functions with respect to production quantities, it is also helpful to express the density benefits in terms of flight volumes by help of (11) as

$$d(\theta^D, D) = d\left[1 - s \cdot n_p^B, D(n_p^B)\right].$$

Finally, under profit maximization with endogenous pricing, the first-order equilibrium conditions for the airlines' flight volumes are given by the partial derivatives of these airline profit functions with regard to peak- and offpeak-flights, respectively, set equal to zero: $\frac{\partial \pi^i}{\partial n_o^i} = 0$ and $\frac{\partial \pi^i}{\partial n_p^i} = 0$. Obviously, in the off-peak period, flights are homogeneous because there are no density benefits. With constant and identical marginal costs $C'(N^i) = c$, the partial derivatives thus are symmetric across airlines $i = \{B, L\}$ and read:

$$b_o(\underline{\theta}) - s \cdot \left[n_o^i + n_p^i\right] \cdot b_o'(\underline{\theta}) = c/s \quad (32)$$

In contrast, during the peak period, airlines' flights are heterogenous products, because for the business airline they include network benefits $d(\theta, D)$. Thus, firms become asymmetric and their partial derivatives to peak-flight volume are distinct. Substituting the above condition (32) into these partial derivatives, using the equivalences $s \cdot n_p^B = (1 - \theta^D)$ and $s \cdot n_p^L = (\theta^D - \theta^*)$ and rearranging then yields the two implicit first-order conditions for *peak flight volume* of the two airlines as

$$\left[b_p(\theta^*) - t(N_p) - b_o(\theta^*)\right] - n_p^L \cdot t'(N_p) - \frac{1}{s} \cdot \left[g(N_p) + n_p^L \cdot g'(N_p)\right] \\ - (\theta^D - \theta^*)\left[b_p'(\theta^*) - b_o'(\theta^*)\right] = 0. \quad (33)$$

for the *leisure* airline, and

$$\left[b_p(\theta^*) - t(N_p) - b_o(\theta^*) + d(\theta^D, D)\right] - n_p^B \cdot t'(N_p) - \frac{1}{s} \cdot \left[g(N_p) + n_p^B \cdot g'(N_p)\right] \\ - (1 - \theta^D)\left[b_p'(\theta^*) - b_o'(\theta^*)\right] - (1 - \theta^D) \cdot d'\left[\theta^D, D\right] = 0 \quad (34)$$

for the *business* airline.

4.1.3 Summary: heterogeneous vs. homogeneous Cournot Duopoly

Comparing the above generic equilibrium conditions to the results from Brueckner (2002a)'s homogeneous products model shows some basic properties of product differentiation in the asymmetric model with network density benefits: The two above equations show that in equilibrium, again the impact of direct flight benefits is opposed to the respective time and congestion costs for each airline's peak travel demand. Moreover, in contrast to the social optimum, now also market power affects the number of peak flights: It is reflected by the term $b_p'(\theta^*) - b_o'(\theta^*)$, which denotes the marginal gain of direct benefits between peak- and off-peak travel. Under current assumptions, this term is always positive and thus has a *negatively* enters the equilibrium condition. Although the marginal benefit term itself is identical for both airlines, it is multiplied with each carrier's peak flight volume, and therefore depends on the individual market shares. This means that it works in the same direction as time costs, and thus tends to *decrease* each airline's peak flight volume in equilibrium. Thus, the market power terms in (33) and (34) represent the „traditional“ market power effect.

For the *business airline*, the network density benefits again enter in the same direction as the direct flight benefits: As they increase the value that passengers receive from a *peak-business* flight, they tend to increase the business airline's peak flight volume. Moreover, eq. (31) shows that density benefits also allow the business airline to charge a higher flight fare than the leisure airline at the same flight volume. Notice, however, that for the airline's profit maximization, only the density benefit for the critical passenger at θ^D is important. As the airline cannot discriminate prices but has to stick to one single posted peak-flight price, it has to optimize it's network density in a way to render network benefits *at* θ^D to be profit maximizing, both in terms of flight volume and flight fares. If network benefits are higher to the right of θ^D, they only add to the consumer surplus for density travelers, but are not accessible for the airline to be turned into excess profits. This again reminds that the airline's target function for optimization is the network value function (14), and not the passengers' network density benefits function. Now, recall that marginal density benefits are positive by assumption. Then, $d'(\theta)$ in (34) enters *positively,* due to the inverted relationship between θ and n_p^B. This means that density benefits are increasingly counter-balanced when the peak-business market-share grows, and thus reflects the concavity network value function.[22]

For the *leisure airline*, an output expansion of peak-business flights ceteris paribus would result in a smaller peak market share. This means that in (33), the market power term as well as marginal congestion and time costs would decrease. For the equilibrium to hold, thus, the business airline's additional peak-flights cannot replace the leisure airline's peak flights on a one-to-one basis. Rather, the total number of peak travelers needs to be larger than under symmetry, so that θ^*decreases and allows the relevant direct flight benefits at the left side of (33) to become smaller and thus balance the equilibrium condition. Notice, however, that as a secondary effect, a higher overall peak flight volume also means higher congestion. Other things equal, this opposing effect can only be compensated if the peak-leisure flight volume is again reduced for both airlines. This means that overall, the leisure airline ends up with a slightly lower peak-flight volume than under symmetry, in order to compensate for the overall higher peak-flight volume and subsequent higher congestion. Consequently, equation (31) reveals that both the *increase* of overall peak-flight volume and the subsequent higher time costs will cause leisure-peak travel fares to *fall*.

Generally, thus, we can constitute that positive network benefits tend to *increase* both the business airline's peak flight volume and peak flight fare, versus the homogenous goods case as well as versus the leisure airline without density benefits. This shows that the introduction

[22]Concavity is shown in 3.1.5, but for linearity only. Due to the nature of the optimization problem, however, it may be generally assumed.

of network density benefits causes market power to rise for the business airline - in-line with economic theory, according to which product differentiation decreases competition and hence increases market-power. Notice that this effect contrasts to the „normal“ market-power mechanism, where a higher output causes prices to fall. This is because the introduction of the heterogeneity corresponds to an *exogenous shock*. Once network density benefits have been introduced, the subsequent output changes again are *endogenous* and follow the „traditional“ market-power mechanism, where prices *increase* when output *falls*.

Last but not least, two important conclusion arises from *offpeak*-condition (32): First, as products are not distinguished by airline within this period, flight fares are determined by $b_o(\underline{\theta})$ and are thus identical across airlines. With c/s constant and identical across firms, (32) in turn, requires that the total output of both firms is the same, that is

$$n_o^L + n_p^L = n_o^B + n_p^B.$$

This means that in the unconstrained, first-come, first-served Cournot equilibrium, the output *across both periods* needs to be equal for both airlines, at half the overall market size. Although market concentration during the peak-period is allowed to differ, and is of main interest for the analysis, this represents a major limiting property of this model. Second, in order for (32) to be fulfilled for both airlines, each airline's output change in the peak period must have an exact corresponding offset of output in the off-peak period. Due to this balancing effect, total output across periods and airlines remains constant. This means that $\underline{\theta}$ remains the same relative to the homogeneous case, and consequently, also the off-peak fare determined by $b_o(\underline{\theta})$ in (31) remains unchanged.

Under the above conditions, the conclusion about *airline profits* is also straightforward: Overall, *peak-density* output from the business airline is *higher*, and *peak-leisure* flights from the leisure airline is *lower* than under symmetry. At the same time, each airline offsets its output change one-to-one across periods. Moreover, the *peak-density fare* is higher and the *peak-leasure fare* is lower than under the previous symmetric case. This means that the business airline sells a higher number of flights at the peak price, which in addition is higher than in symmetry, and thus increases total profits. In contrast, the leisure airline suffers both from a lower mark-up as well as a lower peak-leisure airfare, and thus ends up with a lower overall profit. Overall, thus, the presented model seems to succeed in illustrating the central argument of this study: Endogenous market power due to product differentiation based on network density benefits, that creates an airline asymmetry with a micro-foundation from the demand side. Comparing the above Cournot duopoly market equilibrium to the social optimum from 4.1.1 shows that both the peak and the off-peak flight volumes are distorted by market power. From this, the following can be inferred: First, due to the market power terms, the joint peak output thus is lower than under the social optimum. This means that the corresponding characteristic θ's are higher than justified by congestion and time costs alone. This makes the differential direct flight benefits of the critical passengers increase. Moreover, as the resulting flight fares determined by (31) are above marginal costs, they yield a premium for both airlines. Second, also direct benefits for the critical off-peak passenger are higher, and according to (32), also the off-peak flight-fare is higher than marginal costs. And third, as the critical peak-density passenger is expected to be an interior solution, also the leisure airline offers at least some peak flights. Thus, as general market theory predicts, market power results in both *lower overall* peak- and off-peak *output* than in the social optimum. Consequently, all flight fares are above marginal costs and contain a markup. Moreover, the peak-period is not exclusively served by the business airline. Therefore, as expected, the Cournot duopoly market solution does not replicate the socially optimal corner solution as described by (29).

4.2 Regulation

4.2.1 Quota

Given the assumption that the airport coordinator maximizes social welfare, the optimum number of slots according to eq. (4.1.1) is simply given by solving for the social optimum. This means that $N_p^{opt} = N_p^{soc}$, where N_p^{soc} is the overall number of peak flights as revealed by eq. (28) in 4.1.1. Unfortunately, the generic terms from Section 4.1 do not yield explicit solutions for the dependent variables. Nevertheless, from comparison of the social optimum to the Cournot market equilibrium in (33) and (34), we know that the overall Cournot equilibrium peak output is lower than in the social optimum. Formally, thus, $\theta^*_{Soc.Opt.} < \theta^*_{Cournot}$. This means that the quota as a quantity constraint is not binding regarding the overall peak-flight volume.

As already mentioned before, in the asymmetric case, this should turn our attention to the allocation of the quota to the airlines. In general, we can say that any arbitrary assignment of the quota is equivalently fine, as long as it does not infringe the Cournot equilibrium output. If, however, the business airline is constrained in peak-output below its equilibrium market share, welfare is compellingly reduced. This is because congestion is not reduced if business flights are substituted by leisure flights in the peak period. In contrast, customer value decreases due to the decline in network density, which reduces the density benefits for the passengers - at the same congestion and time costs. Even if the total peak flight volume would decrease, because the leisure airline would not use the additional slots received, we know from (29) in 4.1.1 that in the social optimum, the business airline should operate all flights during the peak period. This ultimately means that any decrease of peak-business flights below the social optimum is welfare decreasing. In contrast, recall that the overall equilibrium output is lower than the total number of quota, and that the quota represents a quantity constraint that can *only limit*, *but not increase* an airline's output. Therefore, even if all airport slots are allocated to the business airline, the quota solution cannot restore efficiency in the asymmetric Cournot duopoly.

Conclusively, under the current asymmetric setting with network density benefits, airport quota cannot increase allocation efficiency. At best, they simply allow to replicate the oligopoly market equilibrium. The danger is, however, that they decrease welfare not only below the social optimum but also below the market solution. This happens if their assignment to the airlines does not allow the natural market structure to evolve.

4.2.2 Congestion Pricing

Marginal congestion costs in the Cournot equilibrium are identified by looking for the first derivatives of time and congestion costs in (33) and (34). And because airlines only internalize marginal costs that are represented in the first-order equilibrium conditions, following Brueckner (2002a), the internalized portion of congestion caused by the respective airline can directly be extracted from these equations as $n_p^i \cdot \left[t'(N_p) + \frac{g'(N_p)}{s}\right]$ for $i \in \{B, L\}$. The correct congestion toll to internalize the external part of delay therefore needs to amount to the remaining marginal congestion costs, that are not accounted for in each airline's profit rationale. With $N_p = n_p^B + n_p^L$, the toll for airline i with competitor j thus amounts to

$$R_i(n_p^i) = n_p^j \cdot \left[t'(N_p) + \frac{g'(N_p)}{s} \right].$$

With $s \cdot n_p^B = (1 - \theta^D)$ and $s \cdot n_p^L = (\theta^D - \theta^*)$, and s normalized to unity, the tolls for the two airlines then can be denoted in terms of the characteristic θ's as

$$R_B(n_p^B) = (\theta^D - \theta^*) \cdot [t' + g'] \tag{35}$$

and

$$R_L(n_p^L) = (1 - \theta^D) \cdot [t' + g'] \,. \tag{36}$$

This means that each airlines faces the above toll after having chosen its optimal output, and thus fully internalizes congestion. At this point it is repeated that the airlines do not take the toll into account in their profit maximization, because if they would, they already perfectly internalized congestion in their output choice. Under the above assumption, then, the tax simply enters the airlines profit function as an additional operating cost for peak flights. And because the tax is exogenous, it can directly be included in the equilibrium conditions as an additively separable term. As designed, then, it increases the congestion costs accounted for in the airline profits to their full amount, so that the marginal congestion terms in both equilibrium conditions equivalently become

$$N_p(t' + g'). \tag{37}$$

As already stressed, the efficiency impact of these congestion tolls now depends on their relative size versus the market power distortion. Unfortunately, the implicit equilibrium and social optimum conditions cannot be directly compared. But looking at social optimum condition (28), and comparing this to equilibrium condition (33) nevertheless reveals two features. First, with congestion costs as in (37), the two equations only differ by the market power term. Secondly, the latter only vanishes if peak-leisure flight volume is reduced to zero due to the tax, so that $\theta^D = \theta^*$. This, however, is not plausible, as an interior solution might rather be expected. But even if this were the case, condition (34) would still need to be fulfilled in a market equilibrium. Then at $\theta^D = \theta^*$but $\theta^* < 1$, even with (37) the social optimum condition (28) were not replicated. Rather, as can be seen easily, output of the peak airline were further depressed below the Cournot equilibrium by the congestion tax, which is already below the social optimum, and thus welfare were decreased.

Some complexity to this is introduced from the leisure airline due to the asymmetry, however. Namely, as (29) dictates, the leisure airline's peak-flight share needs to vanish in the optimum. As we saw above, even if this was the case, the optimum were not replicated by the tax. But in contrast to the business airline, the leisure airline's peak output reduction by the congestion tax works *towards* its socially optimal peak-flight volume, and not *away* from it. So, congestion in fact *is* reduced while the allocation takes a small step *towards* the optimum, at least from this side. Because in oligopoly, however, this still means that the flight fare is increases, it is highly questionable whether this effect has a substantial offsetting effect on the negative welfare impact of the tax on the business airline, and on overall peak-flight volume.

In sum, we can conclude that based on the implicit equilibrium conditions, the congestion tax is most likely to be welfare decreasing in this setting. This is also confirmed by the expectations from recent theoretical studies. Nevertheless, in order to get an ultimate confirmation and a quantitative result, the equilibrium needed to be made explicit, by replacing the generic functions with specific ones. A simple yet illuminating approach were to assume linearity, as depicted in Figure 4. This should be foreseen as an integral extension of this work.

4.2.3 Secondary Trading

As found out above, in the current duopoly setting with network density benefits, a quota solution can only maintain or infer the welfare level as reached in the market equilibrium. This has the direct implication that a *Secondary Trading* scheme for airport slots can only have two outcomes:

Either, as a first possibility, it can restore the potential welfare loss of a mis-allocation, if the initial endowment of the airlines did not allow to replicate the natural market structure. Recall that this can only be the case if the business airline's peak output is constrained below the Cournot equilibrium. Because the market equilibrium is the profit-maximizing solution for

both airlines, an exchange at positive prices will happen until the market equilibrium is reached. The leisure airline will always be willing to sell capacity until the business airline reaches its unconstrained peak-period market share, because with an excess number of slots, it is always able to produce its own unconstrained, profit-maximizing output. Of course, a slot exchange at positive prices would have a distributional impact on the total profits of the airlines, which depended on their initial endowment, and thus bring up a sensitive political-economy issue about distributional equity. Nevertheless, welfare were not impaired by these distributional effects. Rather, the welfare level of the natural market structure could be restored against an initial mis-allocation. As with a quota solution in general, however, the social optimum could neither be replicated.
Or, as a second option, the initial endowment of the airlines were already such that the oligopoly equilibrium were not constrained. Then, trading would simply not occur, because none of the airline had a motivation to actually spend money on buying airport access rights that exceeded their profit maximizing output.

Notice, however, that the two above results crucially depend on the assumption that the airlines are subject to a service obligation of the slots held. If, in contrast, airlines were free to hold unused slots, then the secondary trading market might offer opportunities for strategic behavior. Namely, it might be more worthwhile for such an airline to hold slots unused at an opportunity cost, than selling it to its competitor, if its own commercial position could be strengthened. In the current oligopoly setting, this would be the case if the potentially slot-selling airline could increase its profitability not by restricting its own output, but by decreasing the competitor's flight volume. In the current model, this were the case if the peak-leisure flight fare would increase with a reduction of the peak-business output, while at the same time the peak-leisure flight volume might be maintained - or, vice-versa, for the business airline. Notice that Cournot behavior requires firms to take the output decision of their competitors as given. In this case, however, the airlines would actually have control over the output of their competitor. Thus, Cournot assumptions should not to be impaired. Moreover, notice that even with an effective service obligation in place, strategic behavior might occur. This were the case if restricting the competitor's output were profitable, even if the slot-holding airline would have to increase its own output above the profit-maximizing, unconstrained market equilibrium, on purpose of utilizing the excess slot. At least, such behavior is suspected to occur in practice.

Although such an investigation were hence crucial for the assessment of secondary trading scheme with market power, its outcomes would essentially depend on the initial endowment of the airlines. Therefore, an extensive analysis of multiple scenarios were necessary. The current framework, however, would need to be enriched in order to enable the airlines' profit maximization, while *endogenously* restricting their competitor's output *subject to* their own airport slot holdings. Nonetheless, this should be foreseen as a high priority extension of this study.

5 Conclusion

This study presents an airline-airport model that captures a hub airport of a large network airline, that differentiates its flight network from the other airline's flights based on network density benefits. Subsequently, it investigates a quota scheme, congestion pricing and a secondary trading market equilibrium to allocate airport capacity, in this asymmetric Cournot duopoly setting with demand-side heterogeneity. The results are the following:

On the one hand, the *network density benefits offset* part of the congestion costs, and thus the *efficient* peak flight volume is higher than in a symmetric case. On the other hand, however, product differentiation increases market power. With a high *internalized* fraction of congestion, then peak-flight volume in the market equilibrium is nevertheless below the social optimum. This means that the residual market power effect is important relative to the congestion externality. Under these conditions, *congestion pricing* is likely to be welfare *decreasing*. With a low overall peak-flight volume, a *quota constraint*, in contrast, would not be binding. hence, it had no effect on welfare, unless the initial endowment of the airlines would not allow the natural market structure to evolve. Under asymmetry, the latter represents a welfare caveat, and thus still underlines the need to focus on the *allocation* of the quota to the airlines. In such a case, a *secondary trading* scheme might help to restore the natural market structure. However, the existence of an equilibrium at positive prices would depend on the initial quota assignment to the airlines. Moreover, trading might cause a considerable *distributional impact* on the two airlines, depending on their initial endowment.

Overall, the above considerations make clear that even if one of the above schemes can replicate the natural market structure, the network still remains *overpriced* and *too small,* impeding the passengers' customer value to flourish to its maximum extent. Notice that this were also true if there was no regulation at all. Consequently, in an *asymmetric oligopoly* setting, regulation less likely needs to be concerned with congestion externalities than with market power. This, in turn, indicates the need for a shift the perspective on the airport capacity allocation problem: Instead of considering economic allocation instruments to lower congestion, one might have to think of regulating mechanisms for the monopolistic market structure of the network good. This might allow to reach maximum density benefits, at fair distributional consequences for both the passengers and the airlines involved.

There are, however, some important limitations to this model: First of all, a confirmation of the results from the implicit equilibrium conditions with calculations from explicit values would substantially help to make the results more transparent. This could be done by replacing the generic functions with specific ones, most easily with a linear version, as already undertaken in the analysis of the network density benefits. Second, the impact of strategic behavior of the airlines should also be reflected in the analysis of the secondary trading equilibrium, both in order to investigate a potential welfare caveat as well as the distributional impact of this scheme. These two issues should certainly have priority in the further development of this study.
Moreover, as mentioned in the introduction, the model does not consider global competition across network airlines through different hub airports. This might decrease the market power from product differentiation, and thus affect the efficiency results, possibly reinstating a positive welfare effect of the allocation instruments considered. And last, the model is considerably simplified with constant marginal costs, and a fixed, symmetric aircraft size. Regarding costs, this helps to keep the model traceable. But because there is a controversial debate whether airline operations exhibit economies of scale, especially in a network context, an analysis with decreasing marginal costs would be interesting. This might include the attempt to differentiate costs between the two airlines, because identical costs for both the network airline and its non-networking competitor might be deemed unrealistic. Last but not least, a variable *aircraft size* would allow to study the airlines' choices of *aircraft size* versus *flight frequency*. Such an extension had already been proposed by Brueckner (2002a, p.1368), but judged there to require a „richer framework, where passenger valuation of flight frequency is explicitly considered".

This model would, however, provide a framework rich enough to allow this modification. Resolving some of the above shortcomings with appropriate modifications certainly provides a strong motivation for further research.

6 References

Aguirregabiria, V. and C.-Y. Ho (2010) A dynamic game of airline network competition: Hub-and-spoke networks and entry deterrence, *Technical Report*, Working Paper.

Barbot, C. (2004) Economic effects of re-allocating airport slots: a vertical differentiation approach, *Journal of Air Transport Management*, **10**, 333–343.

Barbot, C. (2005) Effects on welfare of different airport charge systems, *Journal of Air Transportation*, **10** (2) 109–126.

Basso, L. J. and A. Zhang (2010) Pricing vs. slot policies when airport profits matter, *Transportation Research Part B: Methodological*, **44** (3) 381–391, March 2010.

Bauer, J. (2008) Do airlines use slots efficiently?, in A. I. Czerny, P. Forsyth, D. Gillen and H.-M. Niemeier (eds.) *Airport Slots: International Experiences and Options for Reform*, chap. 9, 151–171, Ashgate.

Belleflame, P. and M. Peitz (2010) *Industrial Organization: Markets and Strategies*, Cambrigde University Press.

Ben-Akiva, M. and S. R. Lerman (1985) *Discrete Choice Analysis*, Cambridge, Massachusetts: MIT Press.

Berry, S. T. (1990) Airport presence as product differentiation, *The American Economic Review*, **80** (2) 394–399, May 1990. Papers and Proceedings of the Hundred and Second Annual Meeting of the American Economic Association.

Brueckner, J. K. (2002a) Airport congestion when carriers have market power, *American Economic Review*, **92** (5) 1357–1375.

Brueckner, J. K. (2002b) Internalization of airport congestion: A network analysis, *Technical Report*, **756**, CESifo Working Paper.

Brueckner, J. K. (2008) Slot-based approaches to airport congestion management, *CESifo Working Paper*, **2302**, University of California, Irvine, Department of Economics.

Brueckner, J. K. (2009a) Price vs. quantity-based approaches to airport congestion management, *Journal of Public Economics*, **93** (5-6) 681–690.

Brueckner, J. K. (2009b) Airport congestion management: Prices or quantities?, *ACCESS: the Magazine of UCTC*, **35**, 10–14.

Brueckner, J. K. and Y. 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) pp. 195–222, ISSN 00225258.

Burghouwt, G. (2007) *Airline network development in Europe and its implications for airport planning*, Aldershot, UK: Ashgate.

Button, K. (2005) A simple analysis of the rent seeking of airlines, airports and politicians, *Transport Policy*, **12** (1) 47–56, January 2005.

Cook, A. (2007) The management and costs of delay, in A. Cook (ed.) *European Air Traffic Management*, chap. 4, 97–122, Ashgate.

Czerny, A. I. (2006) Congestion pricing vs. slot constraints to airport networks, *Working Papers*, **2006-02**, Center for Network Industries and Infrastructure (CNI).

Czerny, A. I. (2007) Congestion management under uncertainty in a two-airport system, *Working Papers*, **2007-01**, Center for Network Industries and Infrastructure (CNI).

Czerny, A. I. (2010) Airport congestion management under uncertainty, *Transportation Research Part B: Methodological*, **44** (3) 371–380.

Daniel, J. I. (1995) Congestion pricing and capacity at large hub airports: A bottleneck model with stochastic queues, *Econometrica*, **63** (2) 327–370.

Daniel, J. I. (2001) Distributional consequences of airport congestion pricing, *Journal of Urban Economics*, **50** (2) 230–258.

Daniel, J. I. (2009a) The untolled problems with airport slot constraints, *Working Papers*, **09-09**, University of Delaware, Department of Economics.

Daniel, J. I. (2009b) Pricing canadian airports, *Working Papers*, **09-02**, University of Delaware, Department of Economics.

De Wit, J. and G. Burghouwt (2008) Slot allocation and use at hub airports, perspectives for secondary trading, *European Journal of Transport and Infrastructure Research*, **8** (2) 147–164.

Dixit, A. K. and J. E. Stiglitz (1977) Monopolistic competition and optimum product diversity, *The American Economic Review*, **67(3)**, 297–308.

Donohue, G. L. and A. G. Zellweger (eds.) (2001) *Air Transport Systems Engineering*, vol. 193 of *Progress in Astronautics and Aeronautics*, The American Institute of Aeronautics and Astronautics, Inc.

EUROCONTROL (2008) *Challenges of Growth*, The European Air Traffic Control Agency.

Friedman, J. W. (1983) *Oligopoly Theory*, Cambrigde University Press.

Fukui, H. (2010) An empirical analysis of airport slot trading in the united states, *Transportation Research Part B: Methodological*, **44** (3) 330–357, March 2010.

Garrow, L. A. (2010) *Discrete Choice Modelling and Air Travel Demand*, Ashgate.

Gillen, D. and W. G. Morrison (2008) Slots and competition policy, in A. I. Czerny, P. Forsyth, D. Gillen and H.-M. Niemeier (eds.) *Airport Slots: International Experiences and Options for Reform*, chap. 10, 173–191, Ashgate.

Goedeking, P. (2010) *Networks in Aviation*, Berlin, Heidelberg: Springer.

Holloway, S. (2002) *Airlines: Managing to make money*, Ashgate.

Hong, S. and P. T. Harker (1992) Air traffic network equilibrium: Toward frequency, price and slot priority analysis, *Transportation Research Part B: Methodological*, **26** (4) 307–323, August 1992.

Jehle, G. A. and P. J. Reny (2011) *Advanced Microeconomic Theory*, Pearson Educaction Limited, Essex, England.

Jäggi, F. (2000) Gestaltungsempfehlungen für hub-and-spoke-netzwerke im europäischen luftverkehr : ein ressourcenbasierter ansatz, Ph.D. Thesis, Universität St. Gallen.

Kost, A. (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, P. R. (1980) Scale economies, product differentiation, and the pattern of trade, *American Economic Review*, **70(5)**, 950–959.

Langner, S. J. (1996) *The Allocation of Slots in the Airline Industry*, Nomos Verlagsgesellschaft: Baden-Baden.

Lave, C. (1995) The demand curve under road pricing and the problem of political feasibility: author's reply, *Transportation Research*, **29A** (6) 464–465.

Madas, M. A. and K. G. Zografos (2010) Airport slot allocation: a time for change?, *Transport Policy*, **17** (4) 274–285, August 2010.

Majumdar, A. (2007) Understanding en-route sector capacity in europe, in A. Cook (ed.) *European Air Traffic Management*, chap. 3, 65–96, Ashgate.

Mas-Colell, A., M. D. Whinston and J. R. Green (1995) *Microeconomic Theory*, Oxford University Press.

Matthews, B. and B. Menaz (2008) Economic perspectives on the problem of slot allocation, in A. I. Czerny, P. Forsyth, D. Gillen and H.-M. Niemeier (eds.) *Airport Slots: International Experiences and Options for Reform*, chap. 3, 21–39, Ashgate.

Morrison, S. A. (2005) Deregulation of us air transportation, in D. A. Hensher and K. Button (eds.) *Handbook of Transport Strategy, Policy and Institutions*, vol. 6 of *Handbooks in Transport*, chap. 23, 405–420, Elsevier Science Ltd.

Natalie McCaughey, P. F. and D. Starkie (2008) Characteristics of slot coordinated airports, 2003-2004, in A. I. Czerny, P. Forsyth, D. Gillen and H.-M. Niemeier (eds.) *Airport Slots - International Experiences and Options for Reform*, chap. Appendix, 411–428, Ashgate.

O'Connell, J. F. (2006) Corporate rivalry and competition issues in the airline industry, in A. I. Czerny, P. Forsyth, D. Gillen and H.-M. Niemeier (eds.) *Corporate Rivalry and Market Power*, chap. 4, 54–75, I.B. Tauris, London.

Oum, T. H., A. Zhang and Y. Zhang (2012) Airline network rivalry, in D. A. Hensher (ed.) *Transport Economics*, 432–454, Routledge: London. Source: Canadian Journal of Economics, 28:4a, 1995, 836-57.

Rietveld, P. and E. T. Verhoef (1998) Social feasibility of policies to reduce externalities in transport, in K. Button and E. Verhoef (eds.) *Road Pricing, Traffic Congestion and the Environment*, 285–308, Cheltenham, UK: Edward Elgar.

Shankar, V. and G. S. Carpenter (eds.) (2012) *Handbook of Marketing Strategy*, Cheltenham, UK: Elgar.

Sieg, G. (2010) Grandfather rights in the market for airport slots, *Transportation Research Part B: Methodological*, **44** (1) 29–37, January 2010.

Starkie, D. (1998) Allocating airport slots - a role for the market?, *Journal of Air Transport Management*, **4**, 111–116.

Starkie, D. (2008a) The dilemma of slot concentration at network hubs, in A. I. Czerny, P. Forsyth, D. Gillen and H.-M. Niemeier (eds.) *Airport Slots - International Experiences and Options for Reform*, chap. 11, 193–203, Ashgate.

Starkie, D. (2008b) *Aviation Markets*, Ashgate.

Tanner, G. (2007) The principles of flight planning and atm messaging, in A. Cook (ed.) *European Air Traffic Management*, chap. 2, 35–64, Ashgate.

Ulrich, C. (2008) How the present (iata) slot allocation works, in A. I. Czerny, P. Forsyth, D. Gillen and H.-M. Niemeier (eds.) *Airport Slots: International Experiences and Options for Reform*, chap. 2, 9–20, Ashgate.

Van Reeven, P. (2005) Transport policy in the eu, in D. A. Hensher and K. Button (eds.) *Handbook of Transport Strategy, Policy and Institutions*, vol. 6 of *Handbooks in Transports*, chap. 41, 705–724, Elsevier Science Ltd.

Verhoef, E. T. (2010) Congestion pricing, slot sales and slot trading in aviation, *Transportation Research Part B: Methodological*, **44** (3) 320–329, March 2010.

Vives, X. (2001) *Oligopoly Pricing - Old Ideas and New Tools*, MIT Press, Massachusetts.

Whalen, W. T., D. W. Carlton, K. Heyer and O. Richard (2007) Proposal for a market-based solution to airport delays, *EAG Discussions Papers*, **200714**, Department of Justice, Antitrust Division.

Wichers, R. (1996) *A Theory of Individual Behavior*, San Diego: Academic Press.

Woeckner, B. (2011) *Strategischer Wettbewerb*, 2nd edn., Heidelberg: Springer-Verlag.

Woodruff, R. B. (1997) Customer value: The next source for competitive advantage, *Academy of Marketing Science Journal*, **25(2)**, 139–155.

Zhang, A. and Y. Zhang (2006) Airport capacity and congestion when carriers have market power, *Journal of Urban Economics*, **60** (2) 229 – 247, ISSN 0094-1190.

Zhang, A., Y. Zhang and J. A. Clougherty (2011) Competition and regulation in air transport, in E. Q. André de Palma, Robin Lindsey and R. Vickerman (eds.) *A Handbook of Transport Economics*, chap. 35, 797–821, Cheltenham, UK: Edward Elgar.

A Results from Brueckner (2002a)

In order to compare product differentiation equilibria to previous results with symmetric airlines, this Section presents the main results of Brueckner (2002a).[23] Flights are *homogeneous* across firms, and firms are *symmetric*, and thus there are no firm-specific indices in the equilibrium conditions. Otherwise, the setup is identical to the model presented in Section 3. In the first subsection, the social optimum and market equilibria for monopoly, Cournot oligopoly and perfect competition are shown. In the second subsection, efficiency results for a congestion pricing scheme are shown. Although other instruments are not analyzed, a prediction of expected results for a quota solution is attempted.

The key findings from these results are the following: First, either *market power* or *congestion externalities* always distort the allocation in a market solution. But Cournot oligopoly is the only market form concerned with both distortions at the same time. Second, in Cournot oligopoly a congestion tax may deteriorate welfare, when the market power distortion is large in relation to the externality.

A.1 Equilibria

A.1.1 Social Optimum

The social optimum is developed by maximizing overall welfare, that equals net passenger benefits $B(\theta, n_p) = b(\theta) - t(n_p)$, consisting of direct travel benefits and time costs, minus total airline costs $C(n_o, n_p) = n \cdot c + g(n_p)$. Variable c denotes constant direct operating costs per aircraft, and $g(n_p)$ congestion costs to the airline. The first-order conditions (FOC) for the social optimum are:

$$b_0(\underline{\theta}) \geq c/s \tag{38}$$

[23] Assumptions, formulas, results and interpretations are quoted from Brueckner (2002a) and are not further referenced.

$$\left[b_p(\theta^*) - t(n_p) - b_o(\theta^*)\right] - n_p \cdot t'(n_p) - \frac{1}{s}\left[g(n_p) + n_p \cdot g'(n_p)\right] = 0 \tag{39}$$

with $n_p = (1 - \theta^*)/s$. Condition (38) determines the lower bound of travelers (on the θ- scale), that is: the market size or the overall quantity of people who travel. The inequality shows the dependency of the lower bound of θ from the size of marginal cost per seat: With constant marginal costs, the assumption that marginal costs are higher than the direct travel benefit for the individual at the low end of the θ-scale implies $c/s > b_o(0)$. In this case, it follows from (38) and $b_o' > 0$ that $\underline{\theta} > 0$. Then, the *equality* is binding, i.e. $b_0(\underline{\theta}) = c/s$. With $c/s < b_o(0)$, we get $\underline{\theta} = 0$ and have the *inequality* binding. Then the flight fare is anywhere between marginal cost and marginal benefit, depending on the market structure.

Equation (39) determines the number of *peak travelers*, or equivalently, the *peak-offpeak-split*. It dictates that the marginal benefit of the passenger switching from off-peak to peak travel (the left square bracket) must equal the additional (i.e. marginal) time cost for *all* passengers (the middle term) plus the airlines' congestion costs arising from the additional peak flight volume caused by this user (the right bracket). The most important implication from (38) follows from recalling that $\underline{\theta} = 1 - s \cdot (n_p + n_o)$, where $n_p + n_o$ is the total market size. This illustrates that there is a one-to-one substitution between peak- and off-peak flights. Thus, the *market size* is only determined by the *marginal costs per seat*.

Flight fares are also determined by above conditions (38) and (39): As the willingness-to-pay for the lowest off-peak traveler cannot exceed his benefit under utility maximization, and his benefit cannot be below marginal costs, the off-peak-fare thus must equal marginal costs:

$$f_o = c/s. \tag{40}$$

The peak-flight fare follows from the indifference condition, which states that the fare difference between peak- and offpeak-travel must equal the marginal benefit achieved, minus the associated costs incurred:

$$f_p - f_o = \left[b_p(\theta^*) - t(n_p) - b_o(\theta^*)\right]. \tag{41}$$

Then, from (39) in (41) and rearrangement, it follows that

$$f_p = c/s + n_p \cdot t'(n_p) + \frac{1}{s}\left[g(n_p) + n_p \cdot g'(n_p)\right]. \tag{42}$$

The peak-flight fare thus includes a surcharge over the off-peak fare that exactly equals the marginal time cost and the marginal and total congestion costs caused by one additional peak traveler. This just reflects the complete internalization of congestion costs in the social optimum.

A.1.2 Perfect Competition

Under perfect competition, prices equal marginal costs by standard assumption. Thus, flight fares are determined exogenously as $f_p = \left[c + g(n_p)\right]/s$ and $f_o = c/s$. The first-order condition for $\underline{\theta}$ equals the social optimum condition in (38), with the *equality* binding. Hence, the overall *market size* is *efficient* also under perfect competition. In contrast, the equilibrium condition for the *peak-offpeak-split* becomes

$$\left[b_p(\theta^*) - t(n_p) - b_o(\theta^*)\right] - \frac{1}{s} \cdot g(n_p) = 0. \tag{43}$$

In comparison to equation (39), the terms for the marginal time and congestion costs are missing. This means that an additional peak-traveler does *not* take into account his impact on the additional congestion he causes. Therefore, congestion becomes an externality and leads to *overuse* of the peak period. This result is in-line with economic theory, which mentions that a competitive equilibrium in general is *not* socially optimal when external effects are present (cf. e.g. Mas-Colell *et al.,* 1995, p.353).

A.1.3 Monopoly

For monopoly, two different cases are distinguished: *perfect price discrimination*, and a *non-discriminating* monopolist. Brueckner (2002a, p.1364) calls perfect discrimination an „admittedly strong“ assumption, „despite the airlines’ well-known skill in this practice“, but serving well for illustration purposes.

The *perfectly discriminating equilibrium* is identical to the social optimum. With discrimination, the monopolist can charge all travelers to the full extent of their willingness-to-pay, his profits are maximized at the social optimum. The flight fares are continuous functions of consumer taste, in order to exactly exhaust travel benefits: $f_o(\theta) = b_0(\theta)$ and $f_p(\theta) = b_p(\theta) - t(n_p)$. There is thus no reason for the monopolist to decrease output in order to yield an excess profit. Moreover, he fully takes account of congestion, because he is the only one concerned with it. Therefore, both the *market size* as well as the *peak-off-peak* split remain efficient. As a distributional effect with perfect price discrimination, however, the full amount of the economic rent accrues to the monopolist.

The *non-discriminating case* is different: The monopolist has to decide on one uniform price for each period. As known from standard theory, profit maximization with price-setting implies to set marginal revenue equal to marginal costs. The first-order condition for $\underline{\theta}$ thus becomes

$$b_0(\underline{\theta}) - (1 - \underline{\theta}) \cdot b_0' \geq c/s. \tag{44}$$

With $b_0' > 0$, comparison of (44) to (38) immediately makes clear that $\underline{\theta}$ is higher than in the social optimum, due to the additional market-power term $(1-\underline{\theta}) \cdot b_0'$. Therefore, the market size is less than efficient, which means that *less* people travel in the non-discriminating monopoly case. This is because in a uniform price regime, the monopolist needs to decrease output below the efficient level in order to exploit market power.

For θ^*, the first-order condition reads

$$\left[b_p(\theta^*) - t(n_p) - b_o(\theta^*)\right] - n_p \cdot t'(n_p) - \frac{1}{s}\left[g(n_p) + n_p \cdot g'(n_p)\right] - (1-\theta^*)\left[b_p'(\theta^*) - b_o'(\theta^*)\right] = 0. \tag{45}$$

Also this condition dissociates from (39) by the additional *market-power* term $(1-\theta^*)[b_p'(\theta^*) - b_o'(\theta^*)]$. This means that on the one hand, congestion is still *fully internalized*. But on the other hand, market power also *distorts* the peak/offpeak-split away from its socially efficient level. For $b_p' > b_o'$, the *marginal* increase of *direct flight benefits* is larger for peak than for off-peak flights. In this case the market-power term in (45) is *positive*, and yields a *higher* θ^* than the social optimum. Thus, the peak period is *underused* and congestion is *lower* than optimal.

Pursuant to Brueckner (2002a), this output reduction is referred to as to the *residual market power effect* (rMPE). Corresponding to the marginal benefit of the lowest-θ passenger and to the indifference relation, respectively, the non-discriminating, fixed flight fares are $f_0 = b_0(\underline{\theta})$ and $f_p = f_0 + b_p(\theta^*) - t(n_p) - b_0(\theta^*)$. As both the market size and the peak/off-peak-split are *lower*, it is evident that the monopoly prices are higher than socially efficient prices.

From the monopoly case, we can hence conclude the following: Because the monopolist is concerned with congestion to the full extent, it is completely internalized, regardless of whether prices are discriminated or not. Consequently, the congestion externality does not occur in monopoly. Then, the discriminating monopoly is efficient, as it replicates the social optimum. The non-discriminating monopoly is *not* efficient, as it reduces the market size in both periods below optimum. The reason for this inefficiency is the *residual market power effect*, that aims at increasing the mark-up's by decreasing the output.

A.1.4 Cournot Oligopoly

In Cournot oligopoly, there is a number of k firms that maximize their profits. Aggregate output determines flight fares and the number of travelers. As firms have market-power, they take into account the effect of the aggregate output on the flight fares. Every firm takes the other firms' choices as given.[24] The difference from the Cournot equilibrium versus the monopoly now is the following: firms still internalize congestion and take account of the residual market power effect. However, both of these apply only to each firm's fraction of congestion and output. In other words, each airline is only concerned with its self-imposed congestion. Costs from flight delays that accrue to competitors are not accounted for. Again with the assumption that $c > b_o(0)$ and thus $\underline{\theta} > 0$ and for symmetry, the equilibrium condition for the *market size* is

$$b_0(\underline{\theta}) - \frac{1}{k} \cdot (1 - \underline{\theta}) \cdot b_0'(\underline{\theta}) = c/s. \tag{46}$$

The term again captures marginal costs, *and* a price premium from residual market power. It shows that in oligopoly, the market size depends on the number of firms: For $k \to \infty$, it tends towards the perfect competition equilibrium, and for $k = 1$ it corresponds to the monopoly

[24]From a game-theoretic perspective, this hence represents a Nash-equilibrium. This market form is therefore also referred to as Cournot or *quantity* competition, or Cournot-Nash competition (cf. e.g. Friedman, 1983).

condition. This means that the market size *decreases* and market power *increases,* when the number of firms *diminishes*. The FOC for the number of *peak-flights* reads

$$\left[b_p(\theta^*) - t(n_p) - b_o(\theta^*)\right] - \frac{n_p}{k} \cdot t'(n_p) - \frac{1}{s}\left[g(n_p) + \frac{n_p}{k} \cdot g'(n_p)\right] - \frac{1}{k} \cdot (1-\theta^*) \cdot \left[b'_p(\theta^*) - b'_o(\theta^*)\right] = 0. \tag{47}$$

It contains the same terms as the monopoly case, but now also dependent on k. The brackets or single terms from left to right are*: marginal peak-flight benefits,* marginal *time costs*, carriers' *congestion costs*, and *residual market power*. The condition shows that the oligopoly equilibrium is inefficient due to two opposing distortions: *market power* and *congestion externalities*. With a *large* number of firms, approaching perfect competition, residual market power is low, and so is the internalization of congestion. This means that the peak-flight volume is likely to be *higher* than in the social optimum. Then, mainly the excessive external congestion decreases welfare. In contrast, with a small number of firms, congestion is internalized to a large part but market power increases. Therefore, peak-flight volume tends to be *lower* than in the social optimum. In this case, the welfare constraint stems to a large part from the residual market power effect. Notice that even if the two distortions are balanced to exactly yield the socially efficient level of peak-flights, welfare is still lower than in the social optimum. This is because the *effects* of both distortions remain: market power induces a deadweight loss, and external congestion is welfare deteriorating.

To derive the flight fares in the oligopoly equilibrium, again the presumption is used that with market power, marginal benefits for the left-most traveler are exhausted by the flight fare. This means that the marginal revenue equals marginal flight benefits, as in the non-discriminating monopoly case from above. Then, $f_o = b_o(\underline{\theta})$, and substituting this into (46) and rearranging yields *off-peak* flight fare

$$f_o(k) = c/s + \frac{1}{k} \cdot (1 - \underline{\theta}) \cdot b'_0(\underline{\theta}). \tag{48}$$

This shows that the off-peak flight fare equals marginal cost plus a mark-up. Excess profits increase with market concentration, denoted $\frac{1}{k}$, and decrease with the number of firms. The *peak flight fare* is revealed by using (47) and (48) in indifference condition (41) to get:

$$f_p(k) = c/s+\frac{n_p}{k}\cdot t'(n_p)+\frac{1}{s}\left[g(n_p) + \frac{n_p}{k}\cdot g'(n_p)\right]+\frac{1}{k}\cdot(1-\theta^*)\cdot[b'_p(\theta^*)-b'_o(\theta^*)]+\frac{1}{k}\cdot(1-\underline{\theta})\cdot b'_0(\underline{\theta}) \quad (49)$$

In analogy to the monopoly case, also here the second-last term on the left side represents the mark-up, weighted with the factor $\frac{1}{k}$ that denotes market concentration. The output-decreasing tendency of market power is easily concluded from this term: With $b'_p > b'_o$, the expression in the square bracket evidently increases with a higher θ^* and thus with a *lower* peak-flight share. It is offset by the lower turnover $(1 - \theta^*)$, until their product is maximized. Indifference condition (41) then reveals the relation of flight fares and market power: With a low number of firms, market power is high and so is θ^*. Then, both because peak-flight benefits grow steeper than offpeak-benefits, and because time costs decrease with lower peak flight volume, the fare *difference* grows. In contrast, with a large number of firms and a high peak flight volume, both the fares and their difference are lower. Due to the additive separability of the terms in the indifference condition, the two flight fares change disproportionately. The effect of *market concentration* is hence the following: With a low number of firms, pricing converges towards monopoly mark-ups with a low output volume in both periods. So, the peak is *overpriced* and *underused* relative to the social optimum, because residual market power prevails. With a high number of firms, pricing tends towards marginal costs, and output volume is high. Then, the peak is *overused* but *underpriced* relative to the efficient level, as the congestion externality dominates.

A.2 Regulation

A.2.1 Quota / Access Rights

Quota or access rights, as used under the current administrative allocation scheme in practice, are not explicitly considered in Brueckner (2002a)'s study. Nevertheless, an optimal capacity allocation by use of quota can implicitly be inferred from the social optimum computation: The first-order condition for the efficient number of quota is given by equation (24). The optimum number of slots is therefore implicitly determined by θ^*. Derived from equilibrium condition

(39) it reads

$$N_p = (1 - \theta^*)/s.$$

The allocation problem in this case is straightforward: With symmetric airlines, homogeneous products and constant marginal costs, the quota allocation by the social planner is irrelevant, as long as the total peak-output remains at (or below) the efficient volume. Thus, Brueckner (2002a)'s homogenous model is expected to predict the following: In *perfect competition* with symmetric firms and constant marginal costs, a quota solution should be efficient, as long as the number of access rights for the peak-period is correctly chosen, and airlines have a binding obligation to use them. This is because the quota would constrain peak-output to the social optimum. In an *oligopoly* setting, a quota would either reduce congestion to the socially optimal level, but not be able to correct for the market power distortion (i.e. the deadweight loss). Or, if output were already below the efficient level, it would simply be useless, because it would represent a non-binding constraint. These expectations, however, were only valid for a symmetric equilibrium with identical firms, because then the allocation of the quota would not matter. Of course, this consideration raises the question whether a symmetric market structure is anywhere near to realistic. A possible answer to this question as suggested by this study were that at a large congested hub airport it is presumably *not*. In such a case, it is rather to resemble an asymmetric market structure with one dominant network airline. With such an asymmetry, then, the quota allocation *does* have an impact on allocation efficiency.

A.2.2 Congestion Pricing (CP)

In-line with economic literature, Brueckner (2002a) imposes a congestion toll $R(n_p)$ on each peak flight to internalize the congestion externality. The toll is computed on this premise by use of the equilibrium conditions from above, and leads to full internalization of the external congestion effect. As already pointed out, the tax is computed *after* the airlines have made their output choices. Otherwise their profit-maximization would replicate fully internalizing behavior in equilibrium (cf. 3.2.3).

Monopoly As both the *fully discriminating* and the *non-discriminating* monopoly fully internalize congestion. The deadweight loss as an inefficiency in the latter case stems from the *residual market power effect* only. Hence, there is no externality to be internalized in these market forms and thus no basis for congestion pricing.

Perfect Competition In perfect competition, the externality corresponds to the sum of marginal congestion and time costs. Formally, it equals the difference between (39) and (43), so that

$$R_{PC}(n_p) = s \cdot n_p \cdot t'(n_p) + n_p \cdot g'(n_p).$$

As a result, the peak-flight fare is increased by $R_{CP}(n_p)/s$ per seat, with the higher price yielding the peak/off-peak split of the social optimum. Because there is no residual market-power effect in this case, the tax replicates the social optimum and thus renders the allocation *efficient* (Brueckner, 2002a, p.1364).

Cournot Oligopoly According to first-order condition (47), the congestion pricing tax in the symmetric Cournot oligopoly amounts to

$$R_{OL}(n_p) = \left(1 - \frac{1}{k}\right)\left[s \cdot n_p \cdot t'(n_p) + n_p \cdot g'(n_p)\right].$$

Also this tax fully internalizes congestion. Due to the remaining distortion from the residual market power effect, however, the allocation is still *inefficient.* Exactly as in the non-discriminating monopoly case, the remaining market power distortion is not corrected by the congestion tax.[25] Now, according to the *ambiguity* of congestion pricing, this toll may actually *deteriorate* welfare: Under assumption $b'_p > b'_o$ the residual market power effect tends toward *under*use of the *peak period*, while external congestion tends toward *over*use of the peak period. Now, if the *market power distortion* is large in relation to the *congestion externality*, the tax *further reduces* peak-flights *away* from the social optimum. Then the effect of the congestion

[25] Verhoef (2010, p.322) includes the market power distortion in his congestion pricing computation, in order to „correct for overpricing“, and thus finds congestion pricing to be a *first-best* solution. This, however, has nothing to do with congestion pricing in the sense of a Pigouvian Tax. It rather constitutes an competition instrument to correct for the market power distortion. From an economic point of view, mixing competition and externality issues within a single defined instrument while naming it a congestion tax, however, should at least be questioned.

tax on welfare is *adverse*. This is true even if the resulting *net output* from the two distortions before the tax is *above* the social optimum, but not *sufficiently* above. The ambiguity of congestion pricing hence depends on the *direction* and *size* of the market power effect.