Working Paper #07-28

September 2007

Congestion and Market Structure in the Airline Industry

Itai Ater
Stanford University

Congestion and Market Structure in the Airline Industry*

Itai Ater

September 30, 2007

Abstract

Empirical research on the relationship between market congestion and the market competitive level largely falsifies the positive relationship predicted by theoretical models. In this paper, I exploit the airline industry network structure and focus on the level of congestion during periods in which passengers cross-connect to their final destinations. About 70% of hub airport flights depart or land during these periods. The empirical analysis establishes a strong positive relationship. Furthermore, based on a simple theoretical model, I am able to quantify the potential time savings from eliminating congestion externalities and find that, on average, a flight can save 2 minutes of flight time at its departing airport and another 1.5 minutes at its destination airport. I also find that airlines choose to pad their schedule particularly on competitive routes, presumably to attract uninformed passengers.

JEL classification: L93; R41;

Keywords: Congestion; Air Transportation;

*I received helpful comments from Ran Abramitzky, Tim Bresnahan, Liran Einav, Mark Hansen, Peter Reiss, Frank Wolak as well as from participants at the 2007 ATRS meeting, Berkeley-Nextor seminar, and the 2007 Aspen Wye Conference on Congestion. I thank the NET Institute www.NETinst.org for financial support. All remaining errors are mine.

1 Introduction

To reduce congestion, economists have long been advocating the implementation of congestion pricing, which would induce airlines to take the costs a flight imposes on other flights into account. The federal government has recently indicated that it is likely to adopt congestion pricing policies for both air and ground delays.[1] Rather surprisingly, the empirical literature on congestion at airports largely contradicts the basic prediction of congestion pricing theory, which posits a positive relationship between the level of competition and the level of congestion at a particular market.[2] To reconcile the (lack of) supporting evidence, economists raised alternative explanations and outlined the necessary policy modifications to be consistent with the existing evidence. In this paper, I offer a simple explanation for the reason such empirical evidence was not found and thereafter document a positive relationship between a market level of competition and the level of congestion in the market.

The explanation relies primarily on a different definition of the relevant market. Following the 1978 deregulation of the airline industry, airlines switched to the hub & spoke network structure to achieve economies of density. In particular, hub carriers transfer passengers arriving from similar origins on the same plane to the hub airport.[3] Following a connecting period, passengers heading to similar destinations fly on the same flight from the hub airport to their destinations. Thus, hub airlines are able to increase the aircraft load factor and reduce the cost per passenger. To minimize connecting passengers' layover time at the airport, hub airlines cluster the arriving and subsequent departing flights in relatively short, high-volume periods of flights known as *banks*. Thus, this paper examines how congestion varies across *banks* with

[1] See, for example, "Bush Plays Traffic Cop in Budget Request." *Wall Street Journal*, 5/2/2007.

[2] The rationale for the prediction follows a typical tragedy of the commons argument: Carriers in competitive environments do not take the impact of an additional flight on overall congestion into account, and consequently over-schedule flights. In a concentrated market, carriers realize that they, at least partially, will bear the cost of an additional flight and thus refrain from doing so.

[3] In the empirical analysis, I define a hub airport is an airport in which more than 50% of a carrier passengers are connecting passengers. This carrier is considered a *hub* or a *dominant* carrier, whereas other carriers at the airport are *fringe* carriers.

different levels of competition. I exploit variation in the competitive environment across banks operating within and across airports as well as variation in the cost of queuing across arriving and departing banks. This approach reveals a strong positive relationship between the bank level of competition and the level of congestion at the bank. Figure 1 displays a representative diurnal pattern of operations at the Detroit airport. Typically, there are about 10 departing and arriving banks at a hub airport and about 70% of a hub airport flights either depart or arrive during these bank periods.

The empirical analysis relies on basic insights from theoretical queuing bottleneck models. In a typical queuing model, the equilibrium is characterized by an initial buildup period, in which the queue gradually develops as the rate of commuters joining the queue becomes larger than the rate of commuters leaving the queue (as dictated by the capacity rate). The queue buildup period ends and a receding period begins as the number of commuters joining the queue drops below the capacity rate. Consequently, the first and last arrivals to the queue wait a relatively shorter period of time at the queue, whereas commuters who have joined the queue in between spend a longer time in the queue. Since commuters ignore the impact of their scheduling decisions on other flights, the resulting equilibrium is inefficient. A social planner would reduce the queuing time by shifting commuters who arrive during the buildup period towards the receding period. To facilitate the empirical specification, I assume that dominant carriers also shift flights towards less congested times during the bank. The incentive of a hub carrier to shorten the queue increases based on its share of flights at the bank. Thus, we should expect a negative relationship between the level of congestion in the bank and the bank concentration level. The underlying framework also implies the following two predictions: first, the effect of a concentrated bank on congestion will be more pronounced closer to the center of the bank. The reason is that the queue is relatively short when a bank begins or ends, regardless of the bank's level of concentration; second, as queuing costs rise, airlines will have larger incentives to avoid congested periods, and consequently congestion would be lower.

Thus, if one minute in the arrival queue is costlier than one minute in the departure queue (because airtime is more expensive than ground time), then we would expect queues to be shorter for the arrival bank than in the departing bank.

To establish the positive relationship between market congestion and market level of competition, I use data for flights in October 2000, focusing on 165,000 flights among 31 of the largest U.S. airports. For each of the 16 hub airports in the sample, I specifically identify the *bank periods* in which connecting passengers arrive at and depart from hub airports. For each flight arriving at or departing from a bank, I derive the following characteristics: its location within the bank period relative to the closer edge of the bank, defined as a flight bank position; an interaction term of the bank concentration with the flight bank position; a congestion measure discussed below. These flight characteristics facilitate a reduced form estimation of a bank's pattern of congestion. The dependent variable is flight congestion; the main regressors are the flight's arriving and departing bank position, and the interaction terms between the flight's bank position and the bank concentration. A negative sign on the interaction term is consistent with a dominant carrier internalizing congestion and suggests that this effect is larger for flights operating at the center of the bank. The estimation results are consistent with the framework predictions' suggesting that concentrated banks experience less congestion and that departing banks are more congested than arriving banks. Moreover, the estimation results enables me to quantify the time savings from transforming a competitive bank into a fully monopolized bank. I find that, on average, a flight departing from a bank and arriving at a bank can save about 2 minutes at the origin and approximately 1.5 minutes of its flight time at the destination.

One mechanism dominant carriers may utilize to reduce bank congestion is to lengthen bank periods, thereby increasing the time passengers spend between connecting flights. I examine this hypothesis by first looking at how bank periods change with bank concentration, and then by investigating how an aircraft's ground time changes with arriving and departing bank concentration levels. The aircraft ground time analysis indicates that aircraft arriving at and departing from concentrated banks remain longer on the ground. The relationship between the length of a bank period and the bank concentration level, however, provides mixed support for the above explanation. An alternative mechanism which hub carrier may undertake, which is based on the higher costs of coordinating bank flights when non-hub carriers operate, is left for future research.

Another distinction between this paper and previous studies on congestion is the congestion measure. The measure used in this paper exploits information possessed by airlines regarding the level of congestion at the origin and destination airports at the time of operation. Since airlines anticipate that flights during peak periods are likely to be longer, they incorporate the predicted excess travel time into their schedules. The airlines' response to the varying congestion levels is the object of the analysis. Thus, I use an airline's scheduling decisions to obtain the congestion measure, defined as the added scheduled time of a flight relative to the fastest schedule time of a flight for the same route. For example, the fastest schedule flight time between San Francisco and Newark in October 2000 was 290 minutes. The scheduled time of another flight flying from San Francisco to Newark is 333 minutes. Thus, the congestion measure of this flight is 43 minutes.

The remainder of the paper is organized as follows. In Section 2, I describe briefly the relevant industry background. In Section 3, I outline a simple theoretical scheduling model and derive testable implications. In Section 4, I describe the data used in the empirical application, provide basic descriptive statistics and explain how I constructed variables used in the empirical estimation. Section 5 includes the estimation results and the counterfactuals. Section 6 provides the ground time analysis and Section 7 includes robustness checks. Section 8 concludes.

1.1 Related Literature

Vickrey (1969) was probably the first to offer a deterministic bottleneck queuing model.[4] These models described how commuters travel to or from their workplace, and did not consider strategic interactions among agents. (Brueckner (2002), Brueckner (2005)) was the first to formalize the idea that concentrated markets should exhibit less congestion as carriers internalize the impact of an additional flight on other aircraft operated by the carrier itself.[5] These models demonstrate that a carrier should be charged according to the external delay imposed on aircraft operated by other airlines.[6] For example, in a Cournot setting with three identical carriers, each carrier should be charged 2/3 of the marginal congestion.

Empirical papers which explored the relationship between delays and market structure include Daniel (1995), who was the first to test whether dominant carriers internalize congestion imposed on their own flights. Daniel relied on Vickery's (1969) theoretical bottleneck model and stochastic queuing models to test alternative dominant carrier behaviors. Based on specifications tests, Daniel rejected an internalization behavior by the dominant carrier. Daniel analyzed data from Minneapolis-St. Paul hub airport, but his methodology does not exploit variation across banks to test how hub airlines behavior changes under varying market structures.[7] I follow Daniel by adopting insights from Vickery's model to describe airline behavior during congested periods. However, I differ from Daniel by exploiting variation across banks and airports to identify the internalization behavior. Mayer & Sinai (2003) showed that hub carriers exhibit longer delays than non-hub carriers flying from the same airport. They argue that hub carrier predominantly fly from and to congested bank periods and that non-

[4]Arnott, Palma and Lindsey (1990), Arnott, Palma and Lindsey (1993) have further developed Vickery's model, but mainly examined how changes in the toll scheme would affect congestion. See also Hendesron (1981) and Hendesron (1985) for another line of models with generally similar predictions.

[5]Brueckner (2002) also presented rudimentary empirical evidence based on annual on-time performance records of 25 U.S. airports which support the internalization hypothesis.

[6]Other papers which examined the effect of airline internalization include Pels and Verhoef (2004), Zhang and Zhang (2006), and Basso and Zhang (2007).

[7]Daniel and Harback (Forthcoming) applied the basic methodology on 27 airports and generally found similar results. The methodology, however, does not use variation across banks or airports to test for the internalization behavior. See also Daniel and Pahwa (2000)

hub carrier flights do not fly during bank periods. Consequently, hub carrier flights are longer than non-hub carrier flights. Conceptually, Mayer & Sinai compared bank flights to non-bank flights, although they do not observe whether a particular flight departs or arrives during a bank period. They also found that the assumed time difference between bank and non-bank flights diminishes as an airport becomes more concentrated.[8] I complement Mayer & Sinai analysis by comparing travel time of different flights operating within a bank. Given the large volume of flights operating during banks and the higher congestion during bank periods, this extension is natural.

More recently, Morrison and Winston (2007) quantified the potential benefits from eliminating congestion at airports. They use calibration and alternative assumptions on the dominant carrier behavior and conclude that the quantitative difference between an internalization behavior and a non-internalizing behavior is immaterial for practical reasons. Their empirical analysis divides down the day into 15 minutes periods, ignoring the important role of banks in hub airports' performance.[9]

2 Background

Air traffic delays have become a major public concern in the U.S., Europe and elsewhere in the world. In the U.S. for example, the total estimated costs of air transportation delays are $9.4 billion annually, and between 2002-2004 more than $4.5 billion were spent annually to reduce flight delays.[10] Though the 9/11 events temporarily mitigated the delay problem, the demand for air travel is steadily increasing and is returning to the pre 9/11 levels. During the first five

[8]To obtain a measure of congestion, Mayer & Sinai used as a benchmark the fastest *actual* flight time for the same route and month. They subtract this benchmark from each flight *actual* flight time to derive their measure of congestion. In Section 7.1, I discuss the advantages of the measure I adopt and repeat the main empirical analysis using this alternative measure of congestion. I obtain qualitatively the same results.

[9]See also Carlin & Park (1970).

[10]See www.fightgridlocknow.gov/docs/conginitoverview070301.htm, the costs of air delays in 1999 in Europe are estimated between EUR 6.6 and EUR 11.5 billions, see www.eurocontrol.int/prc/gallery/content/public/Docs/stu2.pdf

months of 2007 U.S. Airlines' on-time performance, measured as the share of flights arriving less than 15 minutes after their schedule time, was 73.5%, the lowest in seven years.[11] At the same time, infrastructure projects face many financial, environmental and political barriers. Consequently, delays are expected to increase in the coming years. In U.S. airports the order of flights' arrival and departure is based on a first-come first-served process. Landing charges are based on aircraft weight rather than a flight time of operation.

3 Model of Flight Scheduling Decisions

Traditional literature on ground transportation congestion highlights the tradeoff between congestion costs and the costs of arriving before or after the most desired time. This tradeoff applies also to airport banks. Absent capacity constraints, all of the hub carrier flights in an arriving bank would land at the same time. Following a necessary interchange period for passengers to change flights and for aircraft to prepare for their next operation, all flights would depart simultaneously to their final destinations. Capacity constraints, however, affect airlines scheduling decisions and require carriers to spread their flight over a longer time period. A longer bank period reduces the congestion costs hub airlines incur, but at the same time may result in longer layover periods for passengers and lower utilization of the aircraft fleet. Internalization of congestion externalities comes into play since the larger the number of flights operated by the airline the more sensitive it will be to over-scheduling flights within a short time period. In the simple model below, the choice variable is t, the time a flight operates, and congestion costs are a function of the operation times of the other flights in the bank.

3.1 Setup

There are two types of carriers: a hub carrier and non-hub carriers denoted H and F, respectively. The analysis focuses on the scheduling decisions within the congested period and

[11] "Passengers Scowl as Airlines Smile", NY Times, August 4, 2007.

assumes that there are N flights operating during the congested period. I derive the equilibrium scheduling for N=3 for three distinct market structures; a fully monopolized structure in which the hub carrier operates all flights; a concentrated structure in which the hub carrier operates two flights and a competitive setting, in which each flight is operated by a different airline. For each ownership structure, I derive the scheduled congestion of each flight, and then obtain comparative static results within and across ownership structures.

Let the hub airline operational costs be given by $C_H = \beta(t_{\bar{H}} - t_{\underline{H}}) + \alpha \sum_{h=\underline{H}}^{\bar{H}} e^{-\sum_{n=1}^{N} |t_n - t_h|}$. The first term reflects the costs associated with longer bank longer bank periods, such as lower aircraft utilization or the lost revenues from connecting passengers dissatisfaction with longer layovers. β is the marginal cost associated with an increased bank period, and $\underline{H}$ and $\bar{H}$ denote the first and last bank flights operated by the hub carrier, respectively.[12] The second term is the congestion costs which are a decreasing, convex function of the difference in operation time between flights operated within the bank period. α is the marginal cost of a flight additional time of operation. A non-hub carrier which schedules its flight during the bank period has the same congestion cost function. I order the flights in the bank according to their schedule (t_1,t_2,t_3), where $t_1 < t_2 < t_3$ and the airline identity. For example, (H, H, F) implies that the first two flights are operated by the hub carrier and the last flight by the non-hub carrier.

3.1.1 Case 1: Monopolistic Setting, (H, H, H)

The monopolistic hub carrier costs are: $C = \beta(t_3 - t_1) + \alpha e^{-[(t_2 - t_1) + (t_3 - t_1)]} + \alpha e^{-[(t_2 - t_1) + (t_3 - t_2)]} + \alpha e^{-[(t_3 - t_1) + (t_3 - t_2)]}$. Solving for the monopolistic airline optimal schedule, normalizing $t_2 = 0$, we derive that the schedule times, corresponding scheduled congestion and length of the bank as displayed in Table 1.

[12] An alternative interpretation could be to consider the costs of aircrafts operating away from the optimal demand time, where $t_{\bar{H}}$, and $t_{\underline{H}}$ are these time differences.

3.1.2 Case 2: Concentrated Market Setting, (H, F, H)

I assume the non-hub airline schedules its flight at the center of the bank. This can be justified, for example, if the non-hub carrier flight flies into its own bank at the destination airport or arrive from its own hub at its origin airport. Alternatively, scheduling decisions are driven by demand and the non-hub carrier chooses the optimal time of operation. The hub carrier costs are: $C = \beta(t_3 - t_1) + \alpha e^{-[(t_2-t_1)+(t_3-t_1)]} + \alpha e^{-[(t_3-t_1)+(t_3-t_2)]}$. The results of the minimization problem are presented in the relevant row in Table 1.

3.1.3 Case 3: Competitive Setting, (F, F, F)

The intermediate flight operates at the optimal time of operation, $t_2 = 0$ and other flights bear a marginal cost of β for scheduling away from 0. Hence, airlines bear the following costs: $C_1 = \beta|t_1| + \alpha e^{-[(t_2-t_1)+(t_3-t_1)]}$, $C_2 = \alpha e^{-[(t_2-t_1)+(t_3-t_2)]}$, $C_3 = \beta|t_3| + \alpha e^{-[(t_3-t_2)+(t_3-t_1)]}$. Solving for the Nash equilibrium, we obtain the scheduling decisions as displayed in Table 1.

3.2 Comparative Static Results and Testable implications

Before outlining the testable implications, I define as the scheduled congestion curve the curve connecting the scheduled congestion of the flights operating in the congested period for a particular ownership structure. I denote the absolute value of the SC curve slope by S_{SC}. The SC curve and the comparative static results are also shown in Figure 2 and Figure 3. The following are the equilibria and comparative static results:

1. $\frac{\partial SC}{\partial |t|} < 0$. Within a bank - as a flight is scheduled closer to the center of a bank its scheduled congestion increases.

2. $\frac{\partial SC}{\partial \alpha} < 0, \forall t, \quad \frac{\partial (t_3-t_1)}{\partial \alpha} > 0$. Across banks - an increase in the marginal cost of scheduling congestion leads to less congestion and longer banks.

3. $\frac{\partial S_{SC}}{\partial HHI} < 0$. As market concentration rises the scheduled congestion curve becomes flatter.

Table 1: Theoretical Framework Results under Different Market Structures

Market Structure	t_1 , SC_1	t_2 , SC_2	t_3 , SC_3	Bank length
H,H,H	$-\frac{1}{3}(ln\frac{3\alpha}{\beta-\alpha})$, $\frac{\beta-\alpha}{3\alpha}$	0, $(\frac{\beta-\alpha}{3\alpha})e^{\frac{2}{3}}$	$\frac{1}{3}(ln\frac{3\alpha}{\beta-\alpha})$, $\frac{\beta-\alpha}{3\alpha}$	$\frac{2}{3}ln\frac{3\alpha}{\beta-\alpha}$
H,F,H	$-\frac{1}{3}(ln\frac{3\alpha}{\beta})$, $\frac{\beta}{3\alpha}$	0, $(\frac{\beta}{3\alpha})e^{\frac{2}{3}}$	$\frac{1}{3}(ln\frac{3\alpha}{\beta})$, $\frac{\beta}{3\alpha}$	$\frac{2}{3}ln\frac{3\alpha}{\beta}$
F,F,F	$-\frac{1}{3}(ln\frac{2\alpha}{\beta})$, $\frac{\beta}{2\alpha}$	0, $\frac{\beta}{2\alpha}e^{\frac{2}{3}}$	$\frac{1}{3}(ln\frac{2\alpha}{\beta})$, $\frac{\beta}{2\alpha}$	$\frac{2}{3}ln\frac{2\alpha}{\beta}$

4 Data, Variable Construction and Descriptive Statistics

4.1 Data

The data for the empirical analysis were compiled from several sources. The main source is the "On-Time Performance Dataset" which includes data on all scheduled and actual domestic flights operated by airlines carrying more than 1% of U.S. domestic passengers.[13] The ten reporting carriers are: Alaska, America West, American, Continental, Delta, Northwest, Southwest, Trans World, United and US Airways. For each flight the following information is provided: carrier, date of flight, flight origin and destination, scheduled departure and arrival time, actual gate push back time and actual gate arrival time, actual airtime, taxi-in time, taxi-out time, and the aircraft tail number. Using the aircraft tail number, I add data on the following characteristics of the aircraft: the number of aircraft seats, manufacturer, weight, number of engines, and manufacturing year.[14]

I focus on flights departing from and arriving at the 31 largest U.S. airports in October 2000.[15] Measures of the number of hourly landing and departing operations that an airport can handle under different weather conditions were obtained from the "Airport Capacity Benchmark Report."[16] Table 3 displays descriptive statistics of the 31 airports. These statistics include the airport level of concentration, the identity of an airport dominant carrier, the

[13]The database is available at *www.transtats.bts.gov/OT Delay/OT DelayCause1.asp*

[14]The FAA Aircraft Reference File and the Aircraft Registration Master File databases consist these data and can be downloaded from *www.faa.gov/licenses certificates/aircraft certification/aircraft registry/*.

[15]I exclude three airports from the sample: Chicago-O'hare, New York LaGuardia and Washington Reagan. At these slot-constrained airports the process of landings and departures is determined differently.

[16]The full report is available at www.faa.gov/events/benchmarks/.

share of the dominant carrier's enplanements out of the total number of airport enplanements, share of flights operated by the dominant carrier, as well as the airport capacity and the share of dominant carrier's passengers who use the airport to connect to other flights. As can be seen from the Table, 16 of the 31 airports are hub airports, i.e. airports where more than half of a carrier's passengers are connecting passengers. The number of passengers originating from an airport or arriving at an airport as their final destination, defined as O&D passengers, was constructed using the DB1B database, which contains a survey of 10% of all the flight fares sold in the U.S. domestic market.[17] To obtain the total number of an airport enplanements, including both O&D passengers and connecting passengers, I use the T100 database which consists of the total thruput of passengers who used each airport.

4.2 Variable Construction and Variation

4.2.1 Scheduled Congestion Measure

Before continuing to the empirical analysis, an appropriate measure of congestion as the dependant variable should be developed. A natural congestion measure is flight delay. Intuitively, a flight arriving after its scheduled arrival time should be considered *late* or *delayed*.[18] The problem with "actual vs. scheduled" measures as criteria for congestion is that airlines anticipate longer travel times during peak demand periods, and build these excess travel time into their schedules. Thus, airlines' schedules already incorporate the predicted congestion at the time of operation, making "actual vs. scheduled" measures inappropriate.

To address this problem, I adopt an alternative measure of congestion which exploits airline information on the expected congestion at the origin and the destination airports. For each directional route, I derive the fastest *scheduled* time of a flight as a benchmark. I then compute the scheduled travel time for all the other flights with the same route and subtract the relevant benchmark. For example, the fastest schedule time of a flight from San Francisco

[17] I thank Chris Mandel from the Department of Transportation for his help in making these data available.

[18] The FAA defines a *delayed* flight a flight that arrives 15 minutes or more after its schedule arrival time.

to Newark, N.J. is 290 minutes. The scheduled time of another flight flying from San Francisco to Newark is 333 minutes.[19] Hence, the *scheduled congestion* measured by minutes and used as a dependant variable for that particular flight is 43 minutes.[20]

To demonstrate the high correlation between actual flight time and schedule flight time as well as to illustrate the high volatility of actual flights compared to scheduled flight, I present in Figure 8 a time series of flights' average scheduled time and flights' average actual time from leaving the gate at the origin airport till it arrives at the gate in the destination airport from 1995 to 2006.

4.2.2 Bank Structures

A basic distinction between this paper and previous papers on the airline industry is the unit of analysis, where I focus on bank flights as the relevant market. For each of the 16 hub airports, I identify when each of the departing and arriving banks start and end. For example, one of the arriving banks at Memphis International Airport starts at 1131 a.m. and ends at 1210 p.m. For each bank I derive the number of flights operating during the bank and the concentration level, measured by flights HHI. To illustrate the negative relationship between market concentration and market congestion, Figure 4 displays the mean predicted congestion of flights departing from banks at an airport as a function of the airport concentration. The negative relationship is clear suggesting that bank operations at concentrated airports involve less congestion.

Hub airports vary in the number of daily banks though typically there are several departing and arriving daily banks. Banks within a hub airport vary in the number of flights

[19] A casual examination of the data suggests that Newark, Boston Logan and San Francisco airports are responsible for the longest schedule adjustments. Among hub airports: Atlanta, Houston and Dallas Fort-Worth airports are a major source for long schedule adjustments.

[20] Based on several conversations with airline officials, the scheduled travel time, or block times, are based on historical operational measures of the time it takes a flight from leaving the gate at the origin airport until it arrives at the gate at the destination airport. Usually, airlines compute an average of past operations excluding outliers.

in each bank as well as in their concentration levels. Bank structures, however, change only slightly across the days of the week. Figure 5 displays the relationship between an airport concentration and the average concentration of the airport banks. As expected, banks are more concentrated than the entire airport concentration suggesting that the hub airline predominantly operates during bank periods. Nevertheless, banks concentration as well as the number of flights in each bank vary over the day, probably illustrating the changing demand conditions across the day. Figure 6 shows how bank concentration levels and the number of flights in each bank change over the day in three representative hub airports: Houston International, Minneapolis-St.Paul and St.Louis Airport. Table 4 provides additional characteristics of bank structures at the 16 hub airports.

Flights arriving or departing during the same bank period may differ on several dimensions. For example, a flight might depart from a bank at one hub airport and arrive at an arriving bank at another hub airport. Another flight might arrive during a bank period but depart from a non-hub airport or from a non-bank period. Moreover, a flight might depart or arrive at the center of a bank or towards the end of a beginning of a bank. Departing at the center of a bank potentially entails longer adjustments to the schedule compared to departing at the beginning or toward the end of a bank. To obtain a separate measure for each flight arriving or departing during a bank period, I compute for each flight its bank position, defined as its relative distance from the closer bank edge. I normalize a bank time period to be 1, where the bank midpoint has a bank position of 0.5, and the starting and ending times attain a bank position of 0. Thus, if a bank starts at 8:00 a.m. and ends at 9:00 a.m. then a bank position of a flight operating at 8:30 is $\frac{30}{60} = 0.5$, at 8:10 or 8:50 is $\frac{10}{60} = \frac{1}{6}$, and at 8:00 a.m. or 9:00 a.m. is 0. For flights operating during non-bank periods, the bank position is 0 as well.

Finally, to control for the level of congestion at non-hub airports and at non-bank periods, I compute the number of flights departing from and arriving at each of the 31 airports for each ten minute interval between 6:00 a.m. and 12:00 p.m. Thus, for each flight in the dataset I derive both its bank position at the origin and at the destination airports, the number of departing and arriving operations at the origin, and the number of departing and arriving operations at the destination. Importantly, the bank structures and the operation measures were constructed based on all the flights departing from or arriving at one of the 31 airports. The regression analysis, however, uses flights departing *and* arriving at the 31 airports. Overall, there are more than 165,000 flights used in the basic estimation analysis.[21]

67% of the flights originating from hub airports depart from departure banks, and 15% of these flights are not operated by the origin airport hub carriers. Focusing on flights flying from one hub airport to another, 85% of the flights that depart from banks and are operated by non-hub carriers actually land at banks at the destination airport. Hence, most of the bank flights which are not operated by the bank hub carrier land at a bank at the other end or depart from a bank at their origin airport.

5 Estimation and Results of the Main Empirical Test

The empirical specification utilized below is driven by the theoretical framework developed in Section 3 and illustrated in Figures 2 and 3.[22] I evaluate a flight congestion as a function of (1) flight position within the bank (2) interaction between the flight bank position and the bank concentration (3) bank concentration and bank total flights (4) airport capacity (5) 10 minutes operations (6) airline, airport and aircraft characteristics. An observation is a flight, and the dependant variable is scheduled congestion as explained above. Focusing on the main regressors of interest and abusing some notation the base specification for flight i departing from bank j at the origin airport and arriving at bank k in the destination airport is as follows:

[21]The arrival and departure times of flights operating around the daylight saving time change in October 27 are reported differently by airports. Thus, to avoid measurement errors, I omit flights performed on the 27 and the 28 of October as well as flight flying to or arriving from Phoenix airport during the last 5 days of October.

[22]A qualification to the theoretical framework is that in reality flights could face congestion both at the origin and at the destination airports. The theoretical framework assumed that only the origin or the destination airport is congested. Hence, the empirical specification should control for airport characteristics and congestion both at the origin and the destination airports.

$$
\begin{aligned}
SCHEDULED\text{–}CONGESTION_{ijk} = {} & \beta_1 BANK-POS_{i,j} + \beta_2 BANK-POS_{i,k} + \\
& \gamma_1 CONC_j * BANK-POS_{i,j} + \gamma_2 CONC_k * BANK-POS_{i,k} + \\
& \delta_1 OPER_{i,org/dest} + \delta_2 BANK-FLIGHTS_j + \delta_3 BANK-FLIGHTS_k + \\
& \delta_4 BANK-CONC_j + \delta_5 BANK-CONC_k + \delta_6 CAPACITY_{(org/dest)} + \epsilon_{ijk}
\end{aligned} \tag{1}
$$

The main predictions are tested by the coefficients on a flight bank position and the interaction term of bank position and bank concentration. First, we expect a positive 'level effect': $\beta_1 > 0$, $\beta_2 > 0$ implying that flights scheduled closer to the center of a bank experience longer travel time or more scheduled congestion. These should hold both for departing banks arriving banks. Second, the prediction that concentrated banks exhibit less congestion is captured by the 'slope effect': $\gamma_1 < 0$ and $\gamma_2 < 0$. The distinction between arriving and departing banks is tested by $\beta_1 > \beta_2$, we expect the departing bank position coefficient to be larger than the coefficient on the arriving bank position.

Table 4 presents estimation results for OLS regressions of several specifications. The results support the predictions. In all specifications the coefficient on a flight position within its bank is positive and significant. Furthermore, the coefficient on the departing bank flight position is larger than the corresponding coefficient on the arriving bank. The interaction terms coefficients are also negative and statistically significant for both departing and arriving banks, implying that more concentrated airports exhibit less delays. Other coefficients also have the expected signs. For example, the coefficients on capacity is negative for both origin and destination airports; larger banks, as measured by the number of flights, experience more delays; the Operations coefficients, which represent the number of flights operating within the same ten minutes, divided into flights in the same direction as flight i, and flights operating in the opposite direction are positive as expected. Adding aircraft characteristics in specification 5 entails losing about 1/4 of the observations,[23] but does not change the results: flights operated by older aircrafts are longer, and an additional aircraft engine is correlated with a one minute faster flight.

In the specifications presented in columns 6-7, I add on-route competition proxies; either route concentration level (column 6) or a dummy variable indicating whether the airline operates as a monopoly in the route (column 7).[24] The competition variables are negative and significant in both specifications implying that airlines tend to pad their schedule in competitive routes compared to less competitive routes. An interpretation of these results suggest that airlines facing competition find it cheaper to pad their schedule rather than offer real time savings for passengers.[25] Finally, in the last specification, I add the aircraft sequential number of the flight in the day and the number of remaining daily flights of the aircraft.[26] The negative sign on the flight in day coefficient is consistent with airline padding their early flights to avoid early delays propagating into later flights. The negative sign on the remaining daily flights coefficient is consistent with airline trying to maximize an aircraft utilization by achieving shorter travel time for airplanes which are schedule to operate more flights in the day.

Importantly, the results can be used to obtain an approximate measure of the benefits from introducing congestion pricing. Each market or bank of flights attains its equilibrium level of congestion as a function of its concentration level. If a congestion pricing mechanism can induce airlines to internalize the externality they impose on other airlines then the level of congestion can be reduced to the level of a fully monopolized market. Hence, the thought

[23]There are two main reasons for this loss; first, the FAA registry data does not include the characteristics of aircrafts which are no longer operating or sold to non U.S. entities. Second, some airlines report their aircrafts nose numbers and not tail numbers. As a result the match of these aircrafts with the FAA registry data is not useful.

[24]Non-stop service operates between 788 directional airport pairs. More than half of these routes are served by only one airline. See Mazzeo (2003) for a paper testing the relationship between on-time performance as a proxy for quality and on-route competition.

[25]Airlines have an incentive to avoid padding their schedule since crew members are paid based on the higher between actual flight time and flight schedule time. Thus, reducing schedule time could translate into lower wages in case a flight actually arrives early.

[26]Since each aircraft operate several times per day and delays early in the day may propagate into more delays later in the day airline may try and mitigate this problem by padding early flights.

experiment is to compute the benefits associated with a competitive bank attaining the level of congestion experienced currently at monopolized banks. Clearly, the more competitive the bank the larger the potential time savings. Utilizing the coefficients presented in Column 3, we get that flight i departing from bank j and arriving at bank k could reduce its flight time by $\delta_4 * (1 - HHI_j) + (1 - HHI_j) * (Bank - Pos_i) * \gamma_1$ at its origin, and $\delta_5 * (1 - HHI_k) + (1 - HHI_k) * (Bank - Pos_i) * \gamma_2$ at its destination. Assuming the average flight position of a flight within a bank is 0.25, the average level of bank concentration is 0.8, we obtain that the estimated savings for a departing flight are 1.96 minutes and for an arriving flight are 1.56 minutes. Aggregating separately over all the flights which land in or depart from hub airports during October 2000[27] and over the total number of passengers traveling through hub airports we obtain that estimated savings in terms of hours of flights are nearly 6500 hours, and in terms of passengers hours the savings are approximately 930,000 hours. Focusing on passengers only and using a conservative measure of 30$ as an hour value, we derive that the potential monetary gains from introducing congestion pricing are more than 28 million dollars in October, and 336 million dollars annually.

6 How Hub Carriers Internalize?

In the previous section, the estimation results established a negative relationship between market concentration and market congestion. A related question is what the mechanism through which hub airline lower congestion at concentrated banks. One possible explanation is that hub carriers in concentrated banks spread their flights over a longer time period, thereby lowering the congestion externality each aircraft imposes on other aircraft. An alternative mechanism focuses on the importance of coordination between hub carrier flights in lowering congestion at concentrated banks, where the presence of fringe carrier impedes the ability to coordinate. In this paper, I focus on the first explanation by exploring how the length of banks depends

[27]Including Chicago O'hare, 224458 flights depart from hub airports and 222205 flights land in hub airports.

on bank concentration, and then examine how aircraft ground time between arriving in a bank and departing in the subsequent bank depends on bank concentration.

6.1 Bank Length

I estimate the length of a bank period as a function of the bank concentration level, the number of bank flights, airport capacity, and number of gates leased by the dominant carrier. Note that an observation in the analysis below is a bank. The results, displayed in Table 5, offer inconclusive support for the first explanation. Bank concentration for departing banks is typically negative, suggesting that concentrated banks are shorter than less concentrated banks. For arriving flights, the estimation results lend some support for the prediction that more concentrated banks are longer.

6.2 Ground Time Analysis

I now turn to examine the ground time of flights departing from banks at hub airports. Figure 8 displays the mean ground time for flights departing from banks at the 16 hub airports as a function of the airport concentration level. There is a weak positive relationship between bank airport concentration and mean aircraft ground time. To control for differences across aircraft ground time stemming from their order of arrival or departure in a bank, I compute for each bank flight its relative order of operation in the bank, defined as a flight rank.[28]

The regression analysis contains several specifications in which an aircraft j ground time between flight i and flight $i-1$ is the dependent variable. Flight $i-1$ lands during an arrival bank, k, and flight i departs during a departing bank m. I omit the first daily flight of an aircraft as well as aircraft remaining on the ground longer than 120 minutes. In particular, the following specification is used:

[28] For each bank, I order the flights by their time of operation and then divide their order by the number of bank operations. Thus, the measure for departing bank ranges from above zero to 1.

$$
\begin{aligned}
GROUND\text{–}TIME_{ijkm} = \beta_1 CONC_m + \beta_2 CONC_k + \gamma_1 RANK_{i,m} \\
+ \gamma_2 RANK_{i-1,k} + \delta_1 DIST_i + \delta_2 DIST_{i-1} + \delta_3 BANK - FLIGHTS_m \\
+ \delta_4 BANK - FLIGHTS_k + \delta_5 HUB - CARR_i + \delta_6 REMAIN_j + \epsilon_{ijkm}
\end{aligned} \tag{2}
$$

where the main coefficients of interest are β_1 and β_2. Table 6 presents the results of several specifications. Both β_1 and β_2 are positive and significant implying that aircraft landing in and departing from concentrated banks remain longer on the ground between operations. These results are consistent with the hub airline internalizing congestion by increasing the operation time between flights, thereby increasing aircraft ground time. Other coefficients have the expected signs; for example, longer flights exhibit longer ground time, and aircraft with more remaining flights in the day remains about one less minute on the ground between operations.

7 Robustness and Endogeneity

In this section, I test the robustness of my results. First, I use an alternative measure of congestion, computed based on flight actual time rather than flight scheduled time. Second, I discuss potential endogeneity issues of the bank structure and present estimation results of several instrumental variables regressions.

7.1 Decomposing Actual Flight Time

The actual congestion measure is obtained, like before, by subtracting a benchmark from a particular flight performance measure. The difference is that both the benchmark and the flight performance measures are obtained from actual performance measures and not scheduled performance measures. A major advantage of actual measures is the existence of distinct components of actual flight time.[29] In particular, the following components are available: Taxi-out time, the time a plane spends on its way from the gate till it leaves the ground; Air time, the time the plane spends between wheels off and wheels on the ground, and finally taxi-in time, the time it takes a plane to arrive from the tarmac to its gate at the destination airport. Thus, 3 separate measures of actual congestion can be constructed for each flight; a total flight actual congestion measure, a taxi-in time congestion measure and taxi-out congestion measure. For example, the actual flight time congestion measure is obtained as the difference between a flight total actual time[30] and the fastest actual flight time for the same route as the corresponding benchmark.[31] I construct similar measures for a flight taxi-in and a flight taxi-out time.[32] To control for the effect of weather conditions on flight performance , I add as regressors several dummies of daily weather conditions both at the origin and at the destination airports. Daily whether conditions are collected and reported by the National Climatic Data Center, which operates weather stations at each of the airports in my sample.[33]

The regressions results for which actual delay measure is used as the dependent variable are presented in columns 1 and 2 in Table 7. In Columns 3-6, I focus on the taxi-in and taxi-out measures.[34] Finally, in the last two columns, I present results of a SUR regression in which the

[29]There are, however, several difficulties with the actual flight time measures. First, these measures are highly sensitive to outliers driven by factors which are unrelated to market structure fundamentals, such as strong tailwind, extreme weather conditions, mechanical problems etc. This problem is exacerbated since the fastest flight time for a route determines the values of the dependent variables for all other flights in the same route. Hence, not only extreme weather conditions may affect the derivation of the dependent variables, but also measurement errors of a specific flight could result in measurement errors of other flights in the same route. For example, in the scheduled flight time data, I discard 4 flights which had unreasonably fast schedule time. See **?** for a critique of actual congestion measures, in particular, its sensitivity to outliers.

[30]obtained by summing up a flight taxi-out, air time and taxi-in time

[31]This is the same approach used in obtaining the schedule delay measure above. Unlike Mayer and Sinai (2003a) I do not incorporate the departure delay, the time a flight is delayed at the gate, as part of an actual flight time.

[32]To derive the dependant variables for the taxi-in and taxi-out regressions I use as benchmarks the measures of unimpeded taxi-in and taxi-out time for each carrier in each airport. These are seasonal measures calculated by the FAA based on all the carrier flights operating in the airport.

[33]The weather data can be found and downloaded at $http://cdo.ncdc.noaa.gov/ulcd/ULCD$. In the estimation results I include several dummy weather coefficients but report only a subset of these coefficients.

[34]For the taxi-out time regression, I take out the origin destination characteristics. For the taxi-in time, I omit the origin airport characteristics.

three previous actual measure equations are estimated jointly. In the SUR regression, I restrict the bank coefficients to be the same across the three equations and display only the coefficients for the actual flight delay regression. The estimation results for the actual congestion measures are consistent, again, with the predictions. Flights scheduled to operate at the center of a bank last longer, but this effect is smaller at concentrated airports. Moreover, this effect is significantly larger at departing banks compared to arriving banks. In fact, in the regressions with airport fixed effects, the arriving banks coefficients are statistically insignificant. The weather coefficients are large and significant.

7.2 Endogeneity of Bank Structure

Hub airlines benefits from operating their flights during bank periods since they are able to maximize network externalities. Non-hub carriers may avoid bank periods due to high congestion costs during these periods. Indeed, Figure 5 demonstrates that banks are more concentrated than the airport concentration level. Nevertheless, non-hub carriers still choose to operate flights during bank periods. One explanation for this behavior is that hub carriers set their bank operations according to local demand conditions, and non-hub carriers wish to offer a competing service. By including the number of flights operating in each bank, I do control, to some extent, for the underlying changing demand conditions. Other unobserved demand conditions, however, raise potential endogeneity problems. For example, if business morning passengers are more time sensitive than afternoon passengers then morning banks will be more congested than afternoon banks, leading to increased level of competition and congestion in morning banks. Furthermore, if international flights arrive at or depart from the hub airport, and carriers like to offer passengers a connecting service to international flights then the measured level of bank concentration is partially driven by unobserved international traffic.[35] In addition, previous empirical papers on congestion have suggested that hub airlines

[35]Future work will include international flights scheduling decisions as well as regional carriers scheduling decisions.

preempt non-hub carrier operations by over-scheduling their own flights during bank periods. Consequently, bank concentration may be correlated with unobserved competitive considerations at the hub airport. To control for unobserved characteristics affecting both the level of competition and congestion, I use three types of instruments. The first type exploits the existence of multiple networks, where carriers operate simultaneously as hub carriers at one airport and as non-hub carriers at other airports. The second type utilizes the correlation between the airport concentration level and the bank concentration level. The third instrument exploits the relationship between the number of gates the hub carrier leases from the airport and the possible time length of the bank. I discuss below the assumptions and rationales for each of the instruments. The results of the estimation employing the instrumental variables are presented in Table 8 The qualitative results are unchanged.

7.2.1 IV - Constrained Bank Flights

Most carriers operate a hub airport and typically more than one. Carriers, looking to maximize network externalities schedule their flights at airports where they do not operate as a hub carrier conditional on the bank structure where they do operate as a hub carrier. For example, the scheduling of a United Airlines flight from Atlanta airport, a non-United hub airport, to Denver, its hub airport depends on the bank structure at Denver. Thus, the underlying assumption is that bank structure at one airport are exogenous to the bank structure at another airport. If this assumption holds then non-hub carrier flights are driven by their own hub airport bank structure. Consequently, variation across banks in the number of non-hub carrier flights, constrained by their own bank operations, can be exploited as an instrument to control for unobserved demand conditions at a particular bank. Hence, for each arriving bank, I derive the share of arriving flights which departed from banks at their origin airport. I repeat the same process for departing bank flights, deriving the share of bank flights heading to bank periods at their destination.[36] The exclusion restriction is satisfied if the share of constrained flights has no direct casual effect on the flight delay. The additional requirement for a valid instrument is that the share of constrained flights is correlated with the bank concentration.[37]

7.2.2 IV - Airport HHI

The second instrument for the bank competitive level is the airport competitive level. Since the concern is that a particular competitive environment is correlated with unobserve determinants of the congestion level, we need an instrument which is correlated with the bank competitive level but not with the unobserved congestion determinants. The airport competitive level is positively correlated with its banks level of competition. The exclusion restriction is satisfied if, for example, airlines decisions to operate from an airport are unrelated to their decisions to operate during a particular bank period.

7.2.3 IV - Hub Carrier Gates

A third instrument focuses on the potential endogeneity of the length of the bank. As been examined in the previous section, hub carrier may choose to lengthen or shorten the bank period as a function of the bank concentration or the desired level of congestion. If unobserved determinants of congestion are correlated with bank length then the bank position variables are biased. To control for this potential endogenity problem, I proceed in two steps. In the first step, I use the bank length regression presented above in subsection 6.1 and Table 5 to construct fitted values of the bank length. Thus, I use the number of gates each hub carrier leases from the airport as an instrument. In the second stage, I use the derived fitted value

[36] I construct an additional instrument by excluding flights operated by a carrier flying from one of its hub airport to another airport where the same carrier operates as the hub carrier. For example, Delta Airlines operated in 2000 banks at Atlanta, Cincinnati, Dallas and Salt Lake City. The results are qualitatively the same as the first instrument which exploits the multiple network structure of the airline industry.

[37] We expect this correlation to be negative, the larger the number of flights constrained by their operations at other airports the lower the share of flights offered by the hub carrier. Indeed, the correlations between the departing and arriving banks and the relevant instrument are -0.1 and -0.058, respectively.

to construct adjusted values of flight bank positions. These adjusted measures are used in the regression analysis instead of the non-adjusted measures presented in section 5, and Table 4.

8 Concluding Remarks

Numerous empirical papers have investigated the consequences of the airline industry deregulation and the emergence of the hub & spoke network. These papers acknowledged the essential role of hub airports and the effect of network structure on welfare and performance. Nevertheless, these papers have not incorporated the structure of banks into the analysis and not investigated airlines decisions and performance in light of the characteristics and structure of banks. Indeed, airlines vary in the number of hubs they use, the number of banks in an airport, the number of flights in a bank, etc. These differences across airlines can help explain, for example, on airlines ability to exploit their network effects. Furthermore, the restructuring of airlines' banks and network following the 9/11 events can shed light on the relationship between network externalities and the state of demand.

In this paper, I utilize variation across banks to document a positive relationship between a market competitive level and market congestion. The variation across banks enables me to identify this strong relationship, establish that concentrated banks are less congested than competitive banks, and that departing banks are more congested than arriving banks. Moreover, the variation across banks enables me also to quantify the potential benefits from eliminating congestion externalities. I find that, on average, a flight departing from and arriving at bank can save 3.5 minutes of its time. Thus, the results suggest introducing congestion pricing at competitive banks or airports could yield the highest savings. Future research should include the relationship between international flights and regional flights on congestion. Furthermore, it should consider how connecting service between domestic and international flights, and how airline alliances affect flights scheduling decisions and congestion.

References

Arnott, R., Palma, A. D. and Lindsey, R.: 1990, Economics of a bottleneck, *Journal of Urban Economics* **27**, 111–130.

Arnott, R., Palma, A. D. and Lindsey, R.: 1993, A structural model of peak - period congestion: A traffic bottleneck with elastic demand, *American Economic Review* **83(1)**, 161–179.

Basso, L. and Zhang, A.: 2007, Congestible facility rivalry in vertical structures, *Journal of Urban Economics* (61), 218–237.

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

Brueckner, J. K.: 2005, Internalization of airport congestion: A network analysis, *International Journal of Industrial Organization* **23**, 599–614.

Daniel, J. I. and Harback, K. T.: Forthcoming, (when) do hub airlines internalize their self-imposed congestion delays?, *Journal of Urban Economics* .

Daniel, J. I. and Pahwa, M.: 2000, Comparison of three empirical models of airport pricing, *Journal of Urban Economics* **47**, 1–38.

Hendesron, V. J.: 1981, The economics of staggerred work hours, *Journal of Urban Economics* **9**, 349–364.

Hendesron, V. J.: 1985, Economic theory and the citites, *Academic Press, Inc.* **2nd edition**.

Mayer, C. and Sinai, T.: 2003a, Network effects, congestion externalities, and air traffic delays: Or why all delays are not evil, *American Economic Review* **93**(4), 1194–1215.

Mazzeo, M. J.: 2003, Competition and service quality in the u.s. airline industry, *Review of Industrial Organization* **22**(4), 275–296.

Morrison, S. and Winston, C.: 2007, Another look at airport congestion pricing, *American Economic Review (forthcoming)* .

Pels, E. and Verhoef, E.: 2004, The economics of airport congestion pricing, *Journal of Urban Economics* (55), 257–277.

Vickrey, W. S.: 1969, Congestion theory and transport investment, *American Economic Review* **59**(2), 251–260.

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

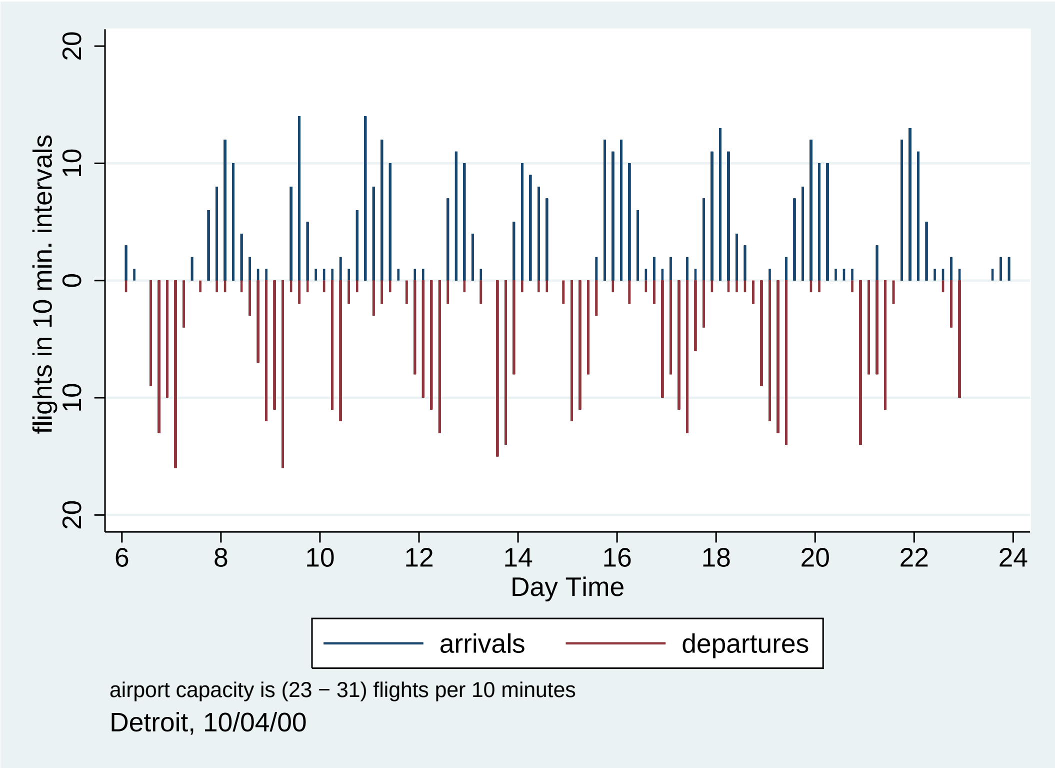

Figure 1: Daily Operations at Detroit Airport

The Figure plots departure and arrival operations in Detroit Airport, October 4, 2000. The number of operations is computed for every ten minutes in the day between 6 a.m. until 12 p.m. The clear sawtooth pattern of arriving and subsequent departing banks is clear. Northwest is the dominant carrier in Detroit, with 58% of its passengers using Detroit airport to connect to other flights on their way to their final destination.

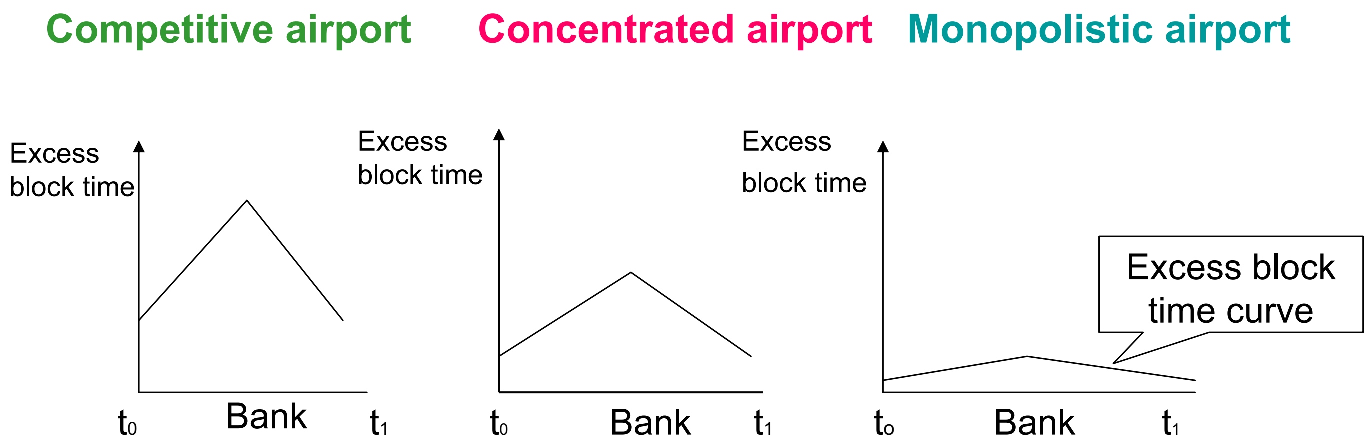

Figure 2: Graphical Illustration 1

The Figure demonstrates how bank or market concentration affects the delay curve. The more concentrated the bank the flatter and lower the curve. The corresponding testable implications are that flights scheduled closer to the center of the bank are longer, but this additional flight time decreases with market concentration.

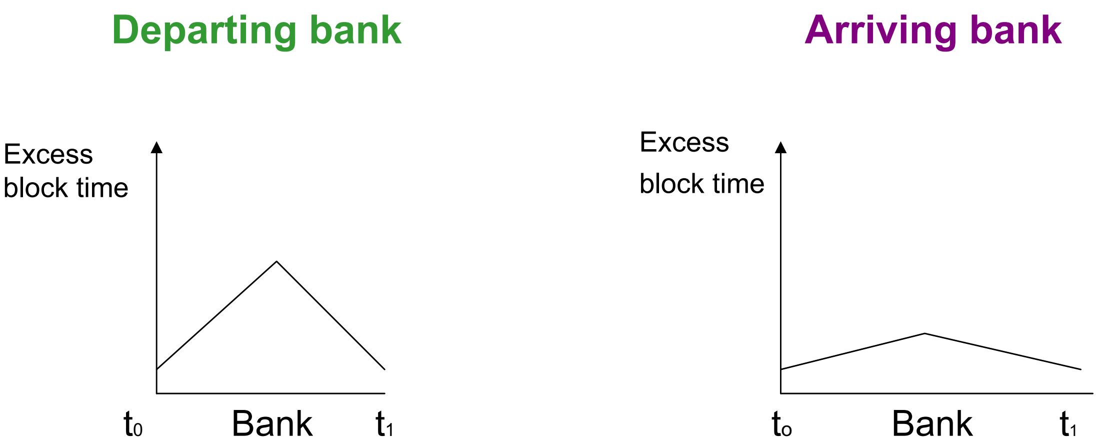

Figure 3: Graphical Illustration 2

An increase in the marginal queuing costs leads to less scheduled congestion since airline ex-ante avoid over-scheduling. The Figure illustrates this result in the context of departing and arriving banks, where it is assumed departing queue time is not as costly as arriving queue time. The relevant testable implication is that the scheduled congestion slope is positive but smaller for arriving banks compared to departing banks.

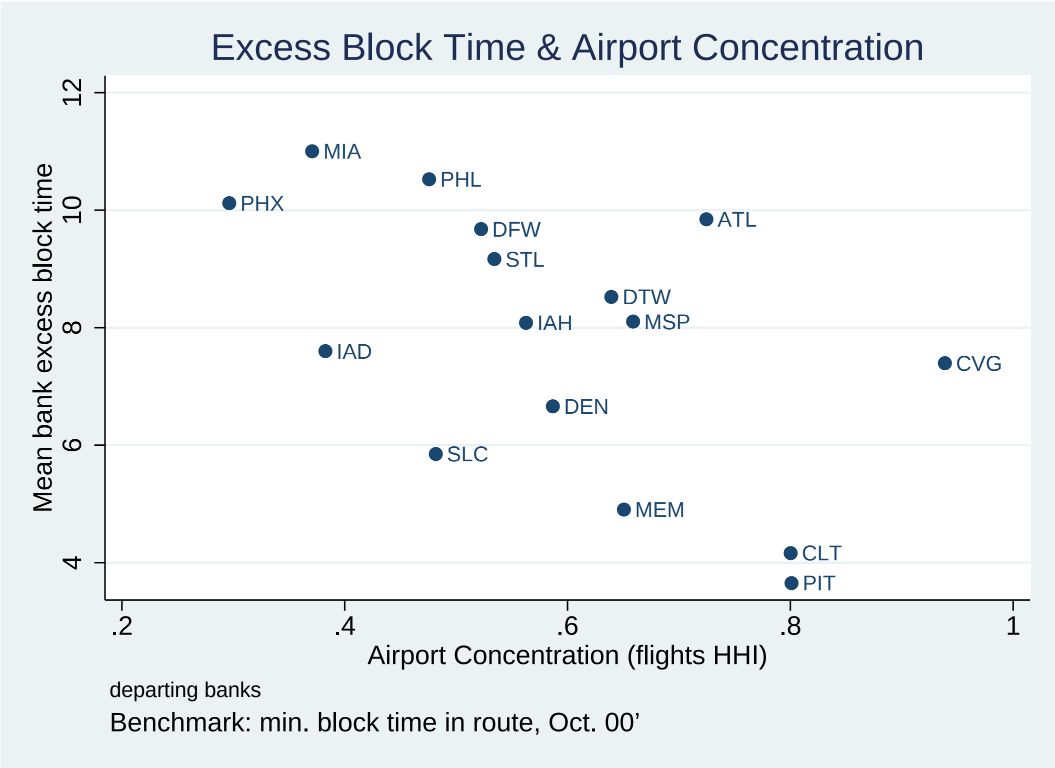

Figure 4: Mean Scheduled Delay as a function of airport Concentration

The Figure displays the mean scheduled congestion of flights departing from banks at the relevant airport and the airport concentration levels. For example, in Philadelphia (PHL) airport, the average scheduled congestion of departing bank flights is more than 10 minutes. In Charlotte, the average scheduled congestion is slightly more than 4 minutes. The figure demonstrates the negative correlation between an airport concentration and scheduled congestion, suggesting an internalization by dominant carriers takes place.

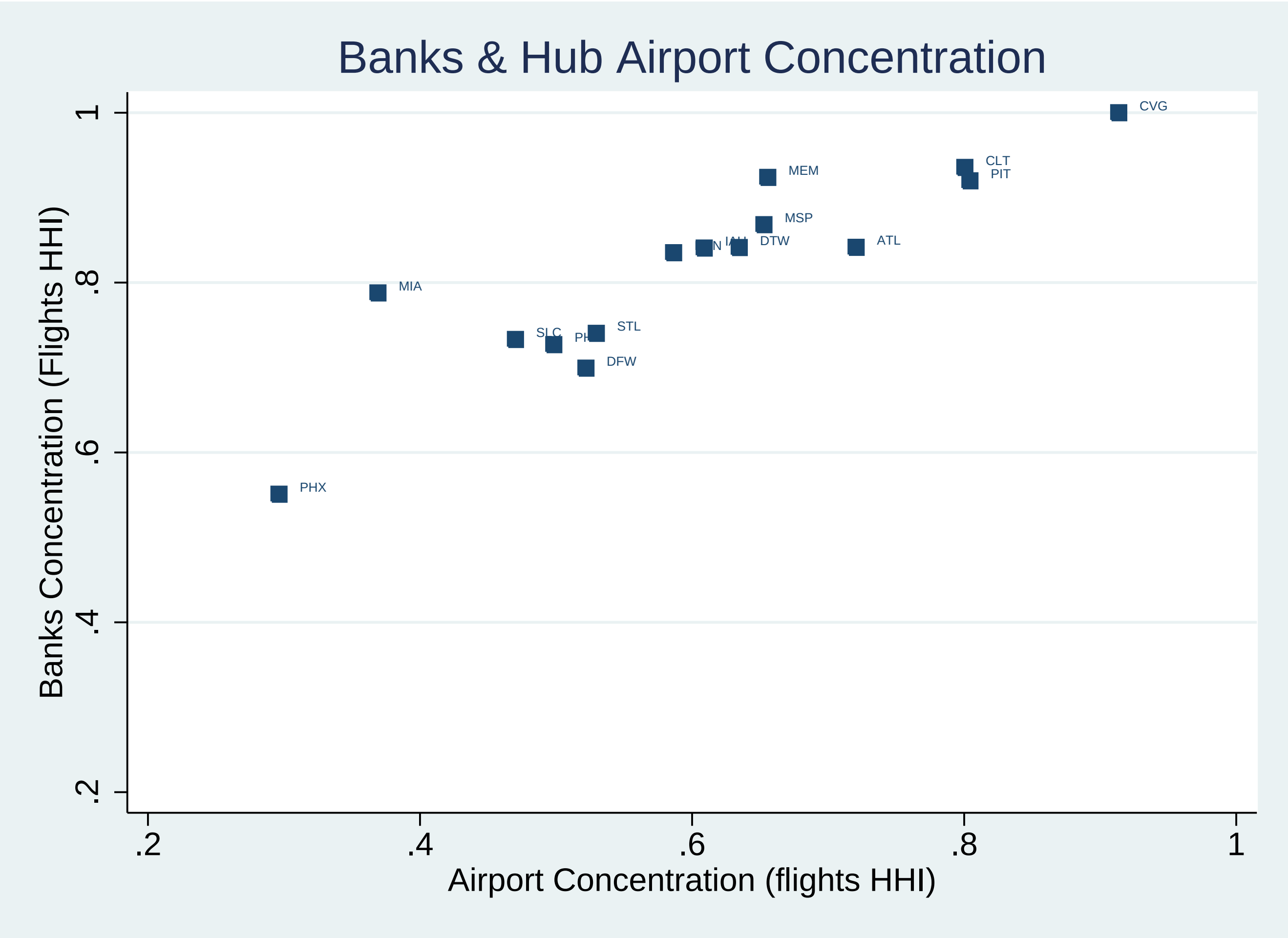

Figure 5: Bank and Airport Concentration

The Figure plots bank concentration levels as a function of the hub airport concentration levels. It demonstrates that banks are more concentrated than the airport they operate in. For example, the concentration level Miami Intl. Airport is less than 0.4, whereas the average bank concentration is nearly 0.8. Thus, hub carrier predominantly operate during banks, and non-hub carrier generally prefer to operate during non-bank periods. Nevertheless, non-hub carrier still operate during bank periods.

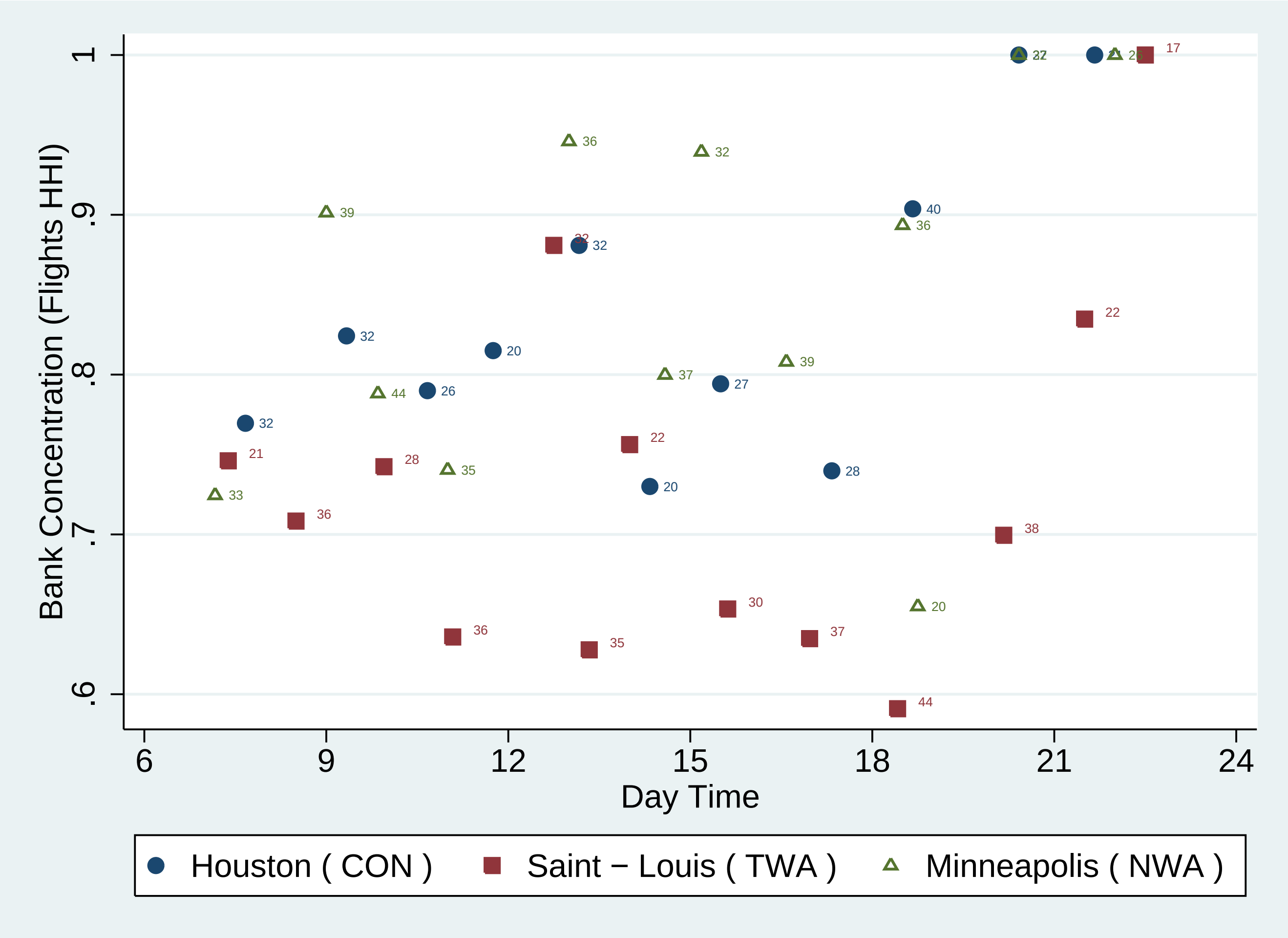

Figure 6: Bank Concentration over the Day

The Figure plots bank concentration levels at three hub airports: Houston International airport, where Continental Airlines operates its hub airport. Saint-Louis airport where Trans World used to operate a hub, and Minneapolis-St.Paul where Northwest Airlines operates as a hub carrier. The numbers represent the number of flights in each bank. Bank Concentration levels vary across the day, and typically late evening banks are more concentrated than morning banks.

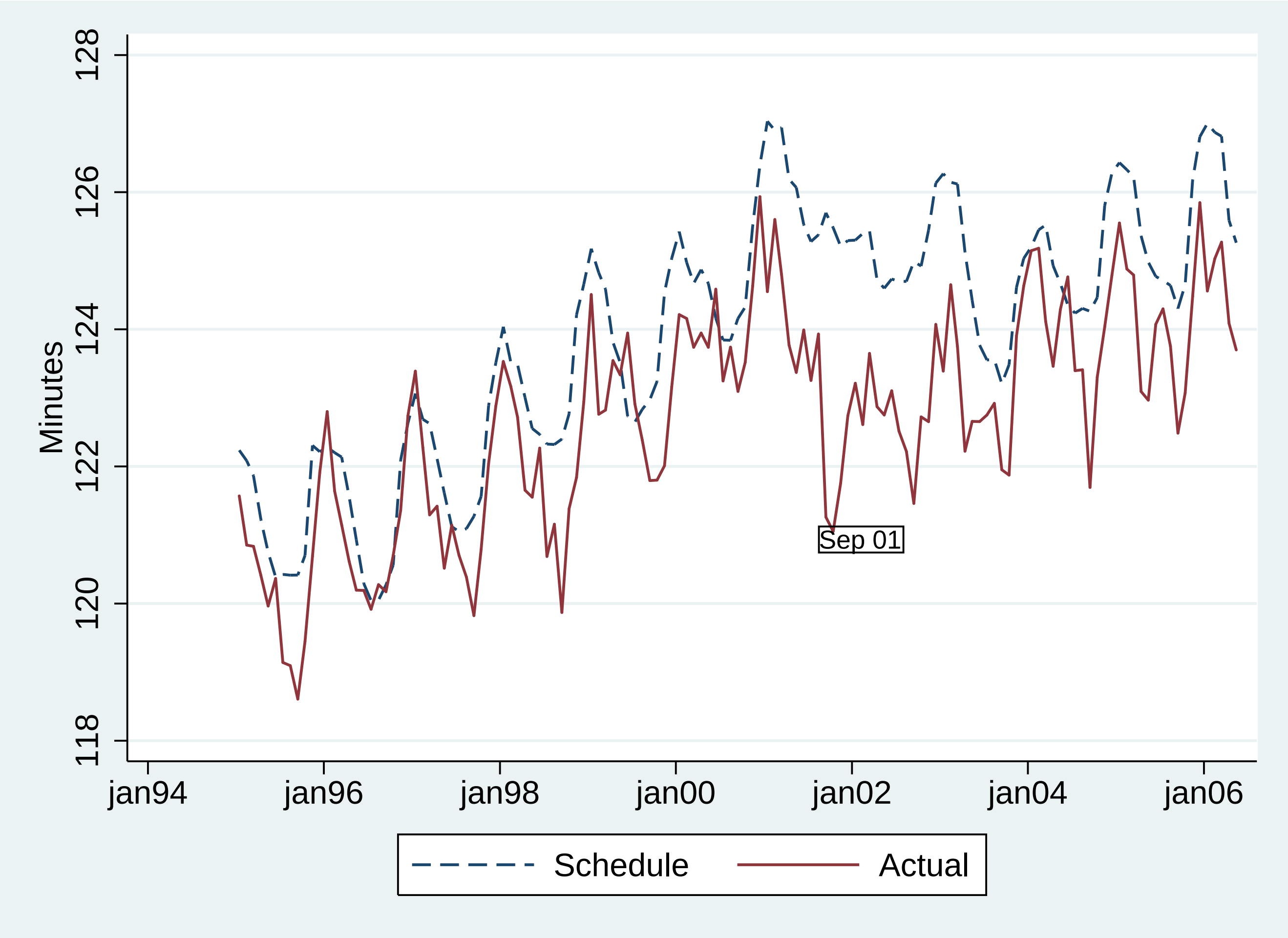

Figure 7: Flight Time

The figure plots monthly mean schedule and actual flight times for all months throughout 1995 to 2006. The flight time represents weighted averages for all domestic airport pairs that have at least 10 flights in every month over the entire time series. The upward time trend and the flight time seasonality is clear. Flights are longer over the years, and flights during Winter months are longer than in the Summer. Before September 2001 the pattern of schedule and actual flight time is similar with larger variation in the actual flight time. Following 9/11 mean actual flight time abruptly falls and the schedule time does not fall by as much. The matching pattern returns as time progresses.

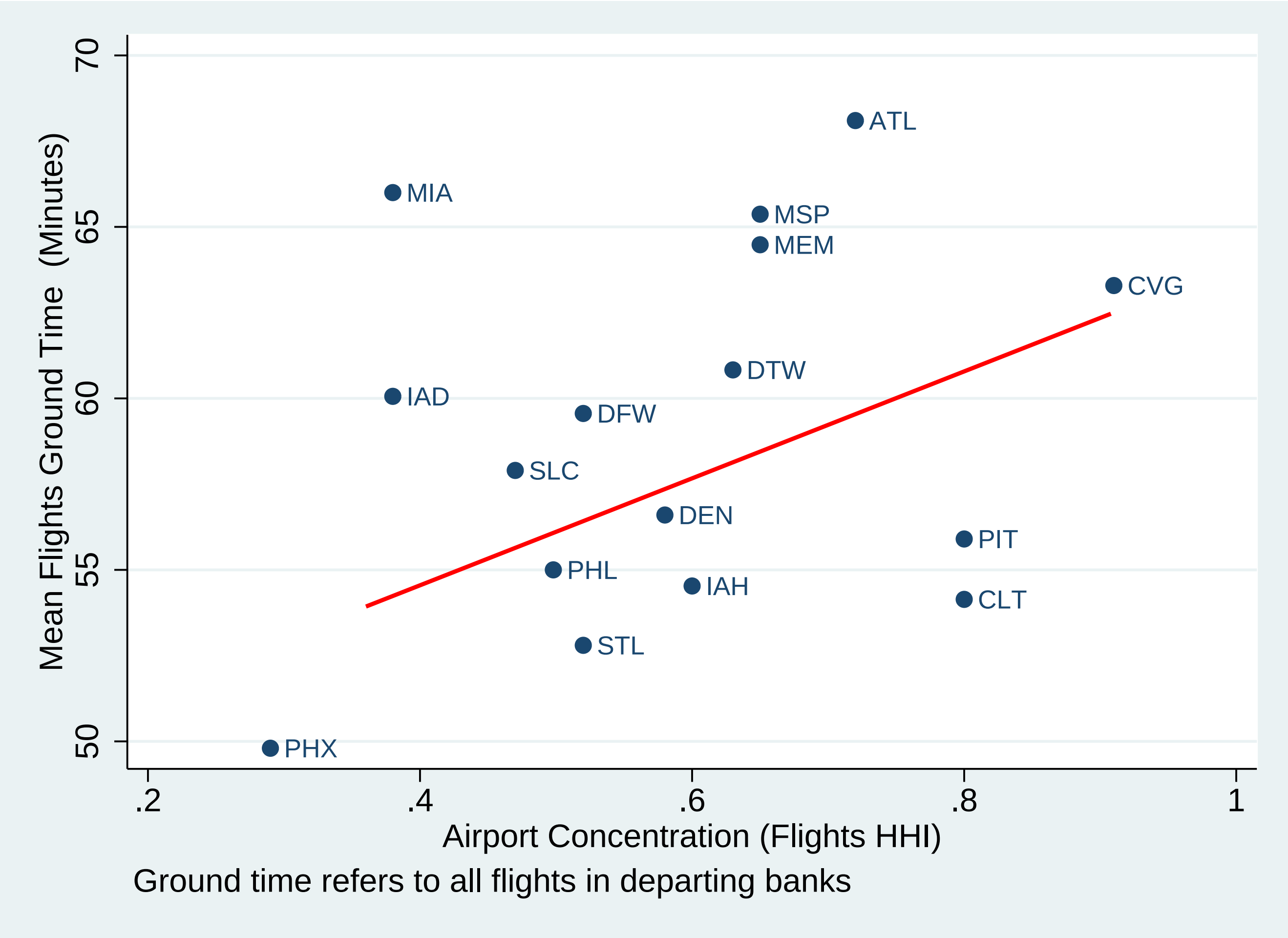

Figure 8: Mean Ground Time as a function of Airport Concentration

The Figure demonstrates a weak positive correlation between an airport level of concentration and the average ground time between operations in that airport. The positive correlation is consistent with banks at more concentrated airports being longer and less congested, requiring aircraft to remain longer on the ground between operations.

Table 2: Airport Characteristics

Airport	Airport concentration*	Dominant carrier	Dominant carrier share of passengers	Share of dominant carrier connecting passengers*	Capacity	
Atlanta	0.71	Delta	0.72	0.66	173	
Boston	0.19	Delta	0.23	0.07	114.5	
Baltimore-Washington	0.29	Southwest	0.41	0.32	86.5	
Cleveland	0.29	Continental	0.37	0.21	74.5	
Charlotte	0.8	US Airways	0.9	0.8	128	
Cincinnati	0.91	Delta	0.64	0.69	122	
Denver	0.58	United	0.64	0.62	194	
Dallas-Fort Worth	0.52	American	0.58	0.63	241.5	
Detroit	0.63	Northwest	0.68	0.58	170.5	
Newark	0.38	Continental	0.54	0.24	80.5	
Fort Lauderdale	0.19	Delta	0.26	0	60.5	
Washington Dulles	0.38	United	0.59	0.51	117	
Houston	0.61	Continental	0.72	0.65	130.5	
New York Kennedy	0.21	American	0.22	0.11	81	
Las Vegas	0.23	Southwest	0.32	0.25	79.5	
Los Angeles	0.18	United	0.28	0.41	129	
Orlando	0.18	Delta	0.28	0.1	138	
Chicago Midway	0.69	Southwest	0.48	0.23	64.5	
Memphis	0.65	Northwest	0.7	0.74	153.5	
Miami	0.37	American	0.53	0.61	111	
Minneapolis-St. Paul	0.65	Northwest	0.71	0.59	113.5	
Portland	0.18	Alaska Air	0.19	0	79.5	
Philadelphia	0.49	US Airways	0.64	0.51	99	
Phoenix	0.29	Southwest	0.31	0.69*	113.5	
Pittsburgh	0.8	US Airways	0.87	0.71	146	
San Diego	0.2	Southwest	0.36	0.03	57	
Seattle	0.22	Alaska Air	0.30	0.07	75	
San Francisco	0.33	United	0.51	0.36	87	
Salt Lake City	0.47	Delta	0.71	0.62	115	
St. Louis	0.53	Trans World	0.72	0.77	93.5	
Tampa	0.17	US Airways	0.2	0.02	92.5	

*Notes: (1) concentration is measured by flights' HHI (2) share of connecting passengers is the dominant carrier share of connecting passengers out of the carrier total passengers (3) in Phoenix, the share of connecting passengers relates to America West, the second largest carrier.

Table 3: Hub Airport Bank Structures

Hub Airport	Hub carrier	# of departing & arriving banks	Share of flights dep. from banks	Average dep. bank period	Average arr. bank period	Average dep. bank concentration	Average arr. bank concentration
Atlanta	Delta	10 , 9	0.8	42.5	59.3	0.84 (0.09)	0.84 (0.09)
Charlotte	US Airways	10 , 10	0.81	29.7	33.6	0.93 (0.06)	0.9 (0.06)
Cincinnati	Delta	9 , 8	0.9	28.3	34.7	1	1
Denver	United	13 , 13	0.72	26.8	27.3	0.8 (0.12)	0.83 (0.15)
Dallas	American	11 , 10	0.88	36.9	29	0.75 (0.18)	0.73 (0.17)
Dallas	Delta	6 , 5	0.88	42.4	40.1	0.62 (0.15)	0.54 (0.07)
Detroit	Northwest	10 , 9	0.86	32.2	37.1	0.81 (0.11)	0.84 (0.16)
Washington Dulles	United	3 , 2	0.37	30	40	0.76 (0.2)	0.7 (0.08)
Houston	Continental	11 , 11	0.78	22.6	33.9	0.83 (0.12)	0.85 (0.08)
Memphis	Northwest	4 , 4	0.79	45	47.25	0.92 (0.06)	0.91 (0.02)
Miami	American	2 , 2	0.16	17.5	33.5	0.79 (0.12)	0.89 (0.15)
Minneapolis	Northwest	9 , 9	0.81	32.1	48.1	0.85 (0.11)	0.81 (0.1)
Philadelphia	US Airways	7 , 7	0.64	35	43.7	0.73 (0.12)	0.77 (0.12)
Phoenix	America West	11 , 12	0.53	22.2	27.8	0.58 (0.15)	0.54 (0.1)
Pittsburgh	US Airways	8 , 8	0.8	31.1	37.6	0.91 (0.07)	0.93 (0.04)
Salt Lake City	Delta	8 , 8	0.68	35.5	31.25	0.73 (0.18)	0.79 (0.1)
St. Louis	Trans World	12 , 12	0.78	28.3	34.5	0.76 (0.12)	0.79 (0.13)

The Table contains characteristics of the 16 hub airports, where a hub airports is defined as an airport in which 50% of a carrier passengers being connecting passengers.

Table 4: **Estimation Results**

Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Bank Pos$_{(org)}$	15.95*** (2.38)	9.54*** (2.4)	16.28*** (2.66)	7.11*** (2.7)	15.31*** (2.87)	14.93*** (2.66)	15.39*** (2.66)	15.18*** (2.64)
Bank Pos$_{(dest)}$	11.41** (2.64)	8.68*** (2.69)	9.1*** (2.86)	6.2** (3)	8.99*** (3.17)	8.04*** (2.85)	8.38*** (2.86)	7.27** (2.84)
Conc*Bank Pos$_{(org)}$	-18.96*** (2.79)	-10.2*** (2.84)	-19.59*** (3.29)	-6.73** (3.33)	-18.68*** (3.62)	-18.04*** (3.29)	-18.45*** (3.28)	-18.47*** (3.28)
Conc*Bank Pos$_{(dest)}$	-14.78*** (3.15)	-9.65*** (3.24)	-11.19*** (3.59)	-6.06 (3.8)	-11.43*** (4.01)	-9.94*** (3.58)	-10.21*** (3.58)	-9.18*** (3.56)
Airport Capacity$_{(dest)}$	-0.003 (0.003)	-0.018*** (0.006)	-0.003 (0.003)	-0.02*** (0.006)	-0.005* (0.003)	-0.003 (0.003)	-0.003 (0.003)	-0.004 (0.003)
Airport Capacity$_{(org)}$	-0.021*** (0.002)	-0.017*** (0.005)	-0.02*** (0.002)	-0.018*** (0.005)	-0.023*** (0.003)	-0.021*** (0.002)	-0.02*** (0.002)	-0.021*** (0.002)
Bank Flights$_{(org)}$	0.074*** (0.007)	0.076*** (0.007)	0.07*** (0.008)	0.09*** (0.01)	0.07*** (0.01)	0.07*** (0.01)	0.07*** (0.01)	0.07*** (0.01)
Bank Flights$_{(dest)}$	0.078*** (0.007)	0.079*** (0.007)	0.08*** (0.01)	0.09*** (0.01)	0.09*** (0.01)	0.08*** (0.01)	0.08*** (0.01)	0.08*** (0.01)
Bank Conc$_{(org)}$			0.14 (0.55)	-1.36** (0.63)	0.66 (0.62)	0.8 (0.55)	0.69 (0.55)	0.89 (0.55)
Bank Conc$_{(dest)}$			-1.1** (0.56)	-1.36* (0.77)	-0.66 (0.62)	-0.42 (0.56)	-0.54 (0.56)	-0.29 (0.56)
Operations$_{(SD,org)}$	0.46*** (0.03)	0.29*** (0.04)	0.46*** (0.03)	0.28*** (0.04)	0.47*** (0.04)	0.43*** (0.03)	0.44*** (0.03)	0.41*** (0.03)
Operations$_{(SD,dest)}$	0.46*** (0.04)	0.32*** (0.04)	0.43*** (0.04)	0.31*** (0.04)	0.45*** (0.04)	0.41*** (0.04)	0.43*** (0.04)	0.39*** (0.04)
Operations$_{(OD,org)}$	0.04** (0.02)	0.01 (0.02)	0.04** (0.02)	0.01 (0.01)	0.03 (0.02)	0.04** (0.02)	0.04** (0.02)	0.04** (0.02)
Operations$_{(OD,dest)}$	0.001 (0.02)	-0.04** (0.02)	-0.003 (0.02)	-0.04** (0.02)	-0.01 (0.02)	-0.002 (0.02)	-0.003 (0.02)	-0.008 (0.02)
Destinations$_{(org)}$	0.02*** (0.007)	-0.02 (0.02)	0.02*** (0.007)	0.09*** (0.009)	0.03*** (0.01)	0.01 (0.01)	0.01 (0.01)	0.014** (0.007)
Destinations$_{(dest)}$	0.02*** (0.007)	0.05*** (0.02)	0.02*** (0.007)	-0.02*** (0.02)	0.03*** (0.01)	0.01 (0.01)	0.01 (0.01)	0.014** (0.007)
No. of Seats					0.005*** (0.002)			
No. of Engines					-0.97*** (0.16)			
Aircraft Age					0.02*** (0.007)			
Competition						-3.85*** (0.37)	-1.72*** (0.19)	-4.15*** (0.375)
Remaining Flights								-0.42*** (0.04)
Flight in day								-0.33*** (0.04)
Airport F.E.	-	+	-	+	-	-	-	-
Carrier F.E.	+	+	+	+	+	+	+	+
Weekday	+	+	+	+	+	+	+	+
Aircraft Char.	No	No	No	No	Yes	No	No	No
R^2	0.17	0.27	0.17	0.27	0.18	0.18	0.18	0.19
N	165538	165538	165538	165538	116572	165538	165538	165538

Standard errors are in parenthesis. Errors are clustered by flight number.

* significant at 10% confidence level, ** significant at 5% confidence level, *** significant at 1% confidence level.

The results lend support to the testable implications: First, the positive bank position coefficient implies that the closer a flight is located to the center of a bank the longer the flight time; Second, the same effect on flights' duration is more modest in arriving banks than in departing banks; Third, the interaction term of bank position and bank concentration coefficient is negative suggesting that as the airport concentration level rises the impact of locating closer to a bank center becomes smaller.

Table 5: Bank Length Analysis

Variable	(1)	(2)	(3)	(4)
Bank Flights	0.473*** (0.055)	0.53*** (0.067)	0.745*** (0.05)	0.85*** (0.08)
Airport Capacity	-2.73*** (0.98)	-3.37* (1.79)	-5.63*** (1.0)	-7.96*** (1.95)
Hub Carrier Gates	-0.088** (0.043)	0.17 (0.15)	-0.03 (0.04)	-0.189 (0.17)
Bank Conc	-2.92 (4.52)	-16.2*** (5.4)	10.4** (4.9)	7.9 (7.1)
Constant	31.2*** (5.9)	21.7 (13.69)	22.28*** (6.2)	36.8*** (15.7)
Airport FE	-	+	-	+
R^2	0.36	0.59	0.62	0.71
N	140	140	135	135

Standard errors are in parenthesis. Errors are clustered by flight number.

* significant at 10% confidence level, ** significant at 5% confidence level, *** significant at 1% confidence level.

Table 6: Ground Time Analysis

Variable	(1)	(2)	(3)
Bank Conc_{dep}	4.49*** (0.681)	4.51*** (0.75)	3.11*** (0.68)
Bank Conc_{arr}	5.34*** (0.735)	5.94*** (0.76)	6.33*** (0.72)
Flight Rank_{dep}	-23.92*** (0.236)	-24.04*** (0.24)	-23.51*** (0.3)
Flight Rank_{arr}	-2.11*** (0.125)	-2.14*** (0.12)	-2.11*** (0.15)
Flight Distance$^{**}{}_{dep}$	0.388*** (0.011)	0.36*** (0.01)	0.328*** (0.13)
Flight Distance$^{**}{}_{arr}$	0.221*** (0.009)	0.19*** (0.01)	0.18*** (0.01)
Bank Flights_{dep}	0.066*** (0.008)	0.1*** (0.01)	0.09*** (0.01)
Bank Flights_{arr}	0.128*** (0.008)	0.147*** (0.01)	0.11*** (0.01)
Hub Carrier	3.195*** (0.278)	3.72*** (0.26)	3.18*** (0.32)
Remaining Flights	-0.97*** (0.038)	-0.95*** (0.04)	-1.17*** (0.04)
Constant	48.81*** (0.89)	44.03*** (1.29)	49.18*** (1.06)
Carrier FE	+	+	+
Airport FE	-	+	-
Aircraft Char.	-	-	+
R^2	0.41	0.43	0.41
N	95150	95150	63866

Standard errors are in parenthesis. Errors are clustered by aircraft.

* significant at 10% confidence level, ** significant at 5% confidence level, *** significant at 1% confidence level.

Notes: * The regression analysis contains all the flights departing from banks and landing in the previous arrival bank, I also exclude an aircraft first flight in the day. **Flight distance is by hundreds miles. The estimates suggest that an aircraft ground time between operations is longer the more concentrated the bank. Hub carriers schedules additional buffer time of 3 minutes between operations, and for each remaining operation by the aircraft in that day the ground time is shorter by one minute.

Table 7: **Estimation Results - Total Flight Time, Taxi-out and Taxi-in Time**

	Actual Flight		Taxi Out		Taxi In		Actual Flight - SUR	
Variable	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
Bank Pos$_{org}$	26.915*** (3.795)	12.57*** (3.64)	15.29*** (1.72)	9.19*** (1.67)			14.96*** (0.743)	9.12*** (0.797)
Conc*Bank Pos$_{org}$	-32.699*** (4.409)	-12.36***	-16.73*** (1.97)	-8.26*** (1.95)			-16.256*** (0.887)	-8.17*** (0.965)
Operations$_{org,SD}$	0.322*** (0.034)	0.17*** (0.03)	0.25*** (0.02)	0.15*** (0.02)			0.49*** (0.011)	0.176*** (0.011)
Operations$_{org,OD}$	0.084*** (0.028)	0.03 (0.024)	0.05*** (0.01)	0.05*** (0.005)			0.15*** (0.006)	0.033*** (0.009)
Airport Capacity$_{org}$	-0.036*** (0.004)	0.006 (0.007)	-0.02*** (0.002)	0.01*** (0.004)			-0.031*** (0.001)	0.026*** (0.002)
Bank Flights$_{org}$	0.05*** (0.015)	0.089*** (0.015)	0.06*** (0.01)	0.08** (0.01)			0.05*** (0.004)	0.083*** (0.005)
Destinations$_{org}$	0.01* (0.01)	-0.065*** (0.019)	0.02*** (0.005)	-0.003 (0.002)			0.013*** (0.003)	0.058*** (0.006)
Thunderstorm$_{org}$	4.303*** (0.251)	4.77*** (0.233)	3.27*** (0.18)	3.33*** (0.18)			4.49*** (0.168)	4.98*** (0.165)
Heavy Fog$_{org}$	1.104*** (0.187)	0.944*** (0.157)	1.3*** (0.12)	0.82*** (0.11)			1.304*** (0.167)	1.026*** (0.165)
Rain$_{org}$	1.1*** (0.143)	0.622*** (0.121)	0.76*** (0.09)	0.7*** (0.08)			1.1*** (0.121)	0.669*** (0.12)
Snow$_{org}$	-2.304*** (0.492)	-2.823*** (0.461)	1.05*** (0.38)	1.1*** (0.35)			-2.61*** (0.463)	-2.935*** (0.453)
Bank Pos$_{dest}$	21.964*** (3.833)	4.168*** (3.465)			1.52* (0.85)	-0.15 (0.86)	1.58*** (0.35)	-0.163 (0.376)
Conc*Bank Pos$_{dest}$	-28.98*** (4.43)	-2.683 (4.11)			-2.62*** (0.99)	0.16 (1.02)	-2.706*** (0.409)	0.194 (0.447)
Operations$_{dest,SD}$	0.552*** (0.042)	0.28*** (0.038)			0.15*** (0.01)	0.08*** (0.01)	0.491*** (0.011)	0.286*** (0.011)
Operations$_{dest,OD}$	0.119*** (0.025)	0.002 (0.022)			0.04*** (0.005)	0.027*** (0.006)	0.15*** (0.006)	0.058*** (0.006)
Airport Capacity$_{dest}$	-0.033*** (0.003)	0.013** (0.006)					-0.032*** (0.001)	0.007*** (0.002)
Bank Flights$_{dest}$	0.017 (0.012)	0.055*** (0.013)			0.03*** (0.003)	0.03*** (0.003)	0.028*** (0.003)	0.036*** (0.004)
Destinations$_{dest}$	0.005 (0.009)	-0.019 (0.019)			0.07*** (0.01)	-0.005* (0.003)	0.016*** (0.002)	-0.01** (0.005)
Thunderstorm$_{dest}$	2.35*** (0.243)	2.55*** (0.219)			0.54*** (0.06)	0.62*** (0.06)	1.43*** (0.244)	1.43*** (0.13)
Heavy Fog$_{dest}$	1.35*** (0.209)	1.01*** (0.177)			0.28*** (0.05)	0.16*** (0.04)	1.302*** (0.131)	0.726*** (0.129)
Rain$_{dest}$	3.985*** (0.164)	3.23*** (0.136)			0.35*** (0.04)	0.17*** (0.03)	2.819*** (0.095)	2.55*** (0.094)
Snow$_{dest}$	6.69*** (0.744)	7.005*** (.637)			-0.17 (0.12)	-0.19* (0.11)	-1.61* (0.92)	6.48*** (0.353)
Airline FE	+	+	-	-	-	-	+	+
R^2	0.09	0.17	0.06	0.11	0.04	0.07		
N	165270	165270	165403	165403	165404	165404	165270	165270

Standard errors are in parenthesis. Errors are clustered by flight number.

* significant at 10% confidence level, ** significant at 5% confidence level, *** significant at 1% confidence level.

The actual flight time regression results are consistent with the testable predictions. The bank position coefficient is positive and larger in departing banks than in arriving banks. Moreover, the interaction term is negative implying less congestion in concentrated airports. In the SUR regression estimation the bank position and the interaction terms are constrained to be equal across the three equations, and only coefficients of the total flight equation are displayed.

Table 8: **Estimation Results - IV Regression**

	IV - Endpoint Bank Flights		IV - Airport HHI		IV - Bank Length	
Variable	(1)	(2)	(1)	(2)	(1)	(2)
Bank Pos $_{(org)}$	28.8*** (6.87)	22.06*** (6.64)	25.35*** (3.79)	13.99** (4.37)	13.25*** (2.28)	8.01*** (2.23)
Bank Pos $_{(dest)}$	39.64*** (12.52)	36.14*** (11.56)	13.18*** (4.16)	4.37 (4.56)	10.01** (2.24)	7.15*** (2.21)
Conc*Bank Pos $_{(org)}$	-35.81*** (8.49)	-26.12*** (8.25)	-31.47*** (4.62)	-15.94** (5.4)	-16.68*** (2.67)	-9.39*** (2.66)
Conc*Bank Pos $_{(dest)}$	-51.21*** (15.4)	-45.5** (14.29)	-18.32*** (5.1)	-5.81 (5.68)	-13.08** (2.79)	-7.75** (2.78)
Airport Capacity $_{(dest)}$	-0.006** (0.003)	-0.03*** (0.007)	-0.006** (0.003)	-0.02*** (0.006)	-0.003 (0.003)	-0.017*** (0.006)
Airport Capacity $_{(org)}$	-0.022*** (0.003)	-0.023*** (0.006)	-0.022*** (0.003)	-0.019*** (0.006)	-0.02*** (0.002)	-0.017*** (0.005)
Bank Flights $_{(org)}$	0.05*** (0.01)	0.07*** (0.01)	0.05*** (0.01)	0.07*** (0.01)	0.079*** (0.007)	0.08*** (0.007)
Bank Flights $_{(dest)}$	0.07*** (0.01)	0.07*** (0.01)	0.07*** (0.01)	0.07** (0.01)	0.078*** (0.007)	0.079*** (0.007)
Operations $_{(SD,org)}$	0.15*** (0.02)	0.08*** (0.02)	0.15*** (0.02)	0.08*** (0.02)	0.45*** (0.033)	0.287*** (0.037)
Operations $_{(SD,dest)}$	0.18*** (0.03)	0.14*** (0.03)	0.22*** (0.03)	0.16*** (0.03)	0.465*** (0.038)	0.326*** (0.042)
Destinations $_{(org)}$	0.03*** (0.01)	-0.01 (0.02)	0.03*** (0.01)	-0.01 (0.02)	0.02*** (0.007)	0.05*** (0.016)
Destinations $_{(dest)}$	0.03*** (0.01)	0.09** (0.02)	0.02*** (0.01)	0.06*** (0.02)	0.023*** (0.007)	-0.016 (0.016)
Airport F.E.	-	+	-	+	-	+
R^2	0.13	0.24	0.15	0.26	0.17	0.26
N	165538	165538	165538	165538	165538	165538

Standard errors are in parenthesis. Errors are clustered by flight number.

* significant at 10% confidence level, ** significant at 5% confidence level, *** significant at 1% confidence level.

The IV regression results do not affect the qualitative results, which are consistent with the internalization theory. The bank position coefficient is positive and larger in departing banks than in arriving banks. Moreover, the interaction term is negative implying less congestion in concentrated banks.