Received: 16 May 2018 | Accepted: 13 February 2021
DOI: 10.1111/ecin.12991

ORIGINAL ARTICLE

Economic Inquiry

Heterogeneous price effects and increased price dispersion from quantity-based congestion management

Sunhyung Lee, Associate Research Fellow[1] | **Zexuan Liu, Assistant Professor**[2] | **Haojun Yu, Assistant Professor**[3]

[1]Korea Institute for International Economic Policy, Sejong-si, Korea
[2]Nanjing Audit University, Nanjing, China
[3]College of Business, Shanghai University of Finance and Economics, Shanghai, China

Correspondence
Zexuan Liu, Nanjing Audit University, 86 Yushan West Road, Pukou District, Nanjing, China.
Email: zxliu@nau.edu.cn

Abstract
We examine the effects of quantity-based approaches to airport congestion management such as slot restrictions on price dispersion and on-time performance. From the evidence at Newark Liberty International Airport, we find that slot restrictions increase price dispersion and improve on-time performance. Price dispersion originates from heterogeneous price effects, as ticket prices for high-value passengers increase more than that for low-value passengers. Our novel finding is that the share of passengers who have purchased tickets at a top-end price increase despite the price hike. We use various empirical specifications and an analysis on leads and lags of the policy for robustness.

KEYWORDS
airlines, airport congestion, on-time performance, price dispersion, slot controls

JEL CLASSIFICATION
L11; L43; L93; R41

1 | INTRODUCTION

Flight departure delays in the United States are common because of increasing airport congestion. Managing airport congestion, especially at hub airports, is a delicate task for policymakers because the airports usually have a limited capacity to meet a high demand for air travel. Brueckner (2009) explains two major tools that the Federal Aviation Administration (FAA) can use to manage airport congestion. The first tool is a price-based approach such as the airport charge, in which airlines must pay fees for each flight that uses congested airports. The airport charge would be higher for flights operating during peak hours than during off-peak hours. The second tool is a quantity-based approach such as slot restriction, in which the airport authority fixes a given number of slots for allocation to airlines operating at congested airports.

In June 2008, the FAA implemented slot restrictions at Newark Liberty International Airport (EWR) to manage congestion, which restricted the frequency of takeoffs and landings at EWR. The aim of our article is to analyze the effects of slot restrictions at EWR on price, price dispersion, arrival delays, layover delays, and departure delays using

Abbreviations: ATT, average treatment effect on the treated; BTS, Bureau of Transportation Statistics; DB1B, Airline Origin and Destination Survey; DCA, Reagan National Airport; DID, difference-in-differences; EWR, Newark Liberty International Airport; FAA, Federal Aviation Administration; HDR, High Density Rule; IATA, International Air Transport Association; JFK, John F. Kennedy International Airport; LGA, LaGuardia Airport; ORD, O'Hare Airport; SDID, semi-parametric DID; UA, United Airlines.

quarterly data between 2006 and 2010. We use a difference-in-differences (DID) approach for examining the impacts on price and price dispersion and use a semi-parametric DID approach for the impacts on arrival, layover, and departure delays because of concerns with the common trend assumption. Considering the timing of slot restrictions at EWR, we take 2006:Q1 to 2008:Q2 as the pre-treatment period and 2008:Q3 and 2010:Q4 as the post-treatment period. Both periods contain 10 quarters.

Our first contribution is that we are the first to analyze the effect of slot restrictions on price dispersion and on-time performance. Our finding suggests that slot restrictions increase price dispersion and improve on-time performance. An analysis of leads and lags of the policy shows that the effects were not anticipated prior to the policy implementation and slot restrictions have lasting effects. On-time performance improves as measured by reduction of arrival delays, layover delays, and departure delays because slot restrictions limit the number of flight operations. The source of the increase in price dispersion is a higher increase in the average price for high-value passengers than the increase in the average price for low-value passengers. Regarding the price dispersion, we further find that the increase in the average price is higher for high-value passengers than for low-value passengers.

Our second contribution is that we are the first to show that slot restrictions can protect passengers who have a higher valuation for time. Despite the higher increase in the average price for high-value passengers, we find that the proportion of high-value passengers is increased after the policy implementation. Czerny and Zhang (2011, 2014, 2015) find that the *price*-based approach such as the airport charge can have a positive effect on high-value passengers and increases their demand. Our finding suggests that the *quantity*-based approach such as slot restrictions can also provide a similar protection for high-value passengers.

2 | RELATED LITERATURE

Various factors can affect price dispersion in the airline industry, including price discrimination and peak-load pricing. Namely, price discrimination may lead to price dispersion. Airlines use price discrimination to maximize their profits by capitalizing on heterogeneous consumer preferences. Particularly, Gerardi and Shapiro (2009) highlight that "third-degree price discrimination" is one of the most popular strategies used by the airlines.[1] The airlines place different restrictions on tickets, such as advance-purchase requirements to exploit on passengers' observable characteristics. Because high-value passengers are generally less sensitive to increases in ticket prices than low-value passengers, airlines can take advantage of the demand characteristics and differentiate pricing strategies for each group of passengers.

Another source of increase in price dispersion is peak-load pricing, as shown in Lott and Roberts (1991) and Dana Jr (1999). Peak-load pricing is a strategy to charge higher prices during periods of high demand to manage capacity constraints. Berglas and Pines (1981) show that the user charge, which is equivalent to a peak-load pricing strategy, suffices for maintaining an optimal congestion level by accommodating heterogeneous preferences for community facilities. Similarly, Lipsman (1994) shows that price discrimination in the form of charging different prices to heterogeneous users improves airport utilization within the multi-product club framework. In a more recent study, Basso and Zhang (2008) analyze peak-load pricing in the framework of airports and airlines. In the presence of airport congestion, peak-load pricing is often known as "airport congestion pricing" (e.g., Czerny & Zhang, 2011), which refers to an increased airport charge imposed by the government to manage airport capacity. Practically, the airport charge is one of the main price-based approaches to airport congestion management. Because airport congestion pricing is a policy tool used by the government to reduce negative externalities of airport congestion (e.g., increased arrival and departure delays), this practice is equivalent to a Pigovian tax.

According to Zhang and Czerny (2012), who summarize the congestion pricing literature, congestion pricing refers to the cost incurred from congestion and the airport charge is the most familiar type of congestion pricing. Daniel (1995) defines this pricing strategy as "self internalization," because the airlines internalize the congestion fees. Brueckner (2002) and Mayer and Sinai (2003) find evidence of self internalization in the airline industry. However, self internalization should be approached with caution because of heterogeneous effects and potential miscalculation.[2] Regarding heterogeneity, Brueckner (2005) finds that self internalization depends on the market structure as oligopolists only partially internalize the congestion fees. Similarly, Rupp (2009) addresses a debate regarding the optimal level of congestion and its relation to airlines' pricing strategies, focusing on whether the congestion fee should be universal or carrier-specific. Regarding potential miscalculation consequences, Daniel and Harback (2009) show that a miscalculation of congestion cost can lead to a substantial loss in welfare when examining the effects of congestion fees on major US airports.

Considering that self internalization may not fully explain the pricing behavior in the airline industry, Czerny and Zhang (2011, 2014, 2015) alternatively explain the pricing behavior based on passenger types. High-value passengers and low-value passengers have different preferences for time valuation in the airline market (i.e., a time-valuation effect). Since high-value passengers have a greater time value than low-value passengers, an increase in the airport charge can lead to an increase in the demand of high-value passengers. Therefore, airport charge can protect high-value passengers from excessive congestion. A general conclusion from Czerny and Zhang (2015) is that a socially efficient airport charge can be higher than the existing level as to benefit travelers who have a relatively higher value for time.

While price-based approaches to airport congestion management may be welfare enhancing, congestion pricing is quite complicated as it requires delicate calculations of carrier-specific tolls, in which smaller carriers may pay a higher toll. Because of the complication for price-based approaches, quantity restraints in the form of slot restrictions have been the preferred choice of the FAA to manage congestion at selective US airports, including EWR. The FAA sets a restriction on the quantity of slots and allocates a certain number of slots to each airline. Brueckner (2009) theoretically shows that the quantity-based approaches can serve as an effective strategy for managing airport congestion. Moreover, Reitzes et al. (2015) show that efficiency can be achieved via slot restrictions when slots are transferred from larger airlines to smaller airlines. For empirical analysis, Swaroop et al. (2012) examine slot-controlled airports in the states including EWR and contend that a more aggressive approach in restricting the number of slots can further reduce delays caused by congestion. Similarly, using evidence from the European airports, Pertuiset and Santos (2014) propose a more efficient system to coordinate slot allocations.

Comparing and contrasting the effectiveness of price-based and quantity-based approaches to airport congestion management has been the focus of the recent literature. Brueckner (2009) shows that pricing and quantity restraints are equivalent under the assumption of the absence of market power. Furthermore, Basso and Silva (2015) contend that congestion pricing and quantity restraints are no longer equivalent when market power is included. Similarly, Basso and Zhang (2010) show that the equivalence depends on the ownership status of airports whether the airport is privately owned or self-financed. In terms of welfare, Verhoef (2010) shows that the efficacy of congestion pricing and quantity restraints depends on market power distortions, types of airport charge, and congestion externality within a stochastic model. Last, Czerny (2010) finds that congestion pricing is more advantageous under a linear model of congestion management, whereas quantity restraints may be preferred under a nonlinear model.

The principal objective of our article is how quantity-based approaches to airport congestion management affect price dispersion. While a large literature has examined the impact of price-based approaches on price dispersion by passenger types,[3] the studies on how quantity-based approaches affect price dispersion by passenger types are almost nonexistent. Our article is first to show that slot restrictions increase price dispersion, and the effect is different by passenger types, in which high-value passengers are better off after the policy implementation.[4] Our finding is consistent with Czerny and Zhang (2011, 2014, 2015) who find that high-value passengers are more well off with a higher airport charge under the congestion pricing framework.

3 | SLOT RESTRICTIONS AT EWR

Regulating the number of landing and departure slots was necessary for relieving air traffic congestion at major US airports. The High Density Rule (HDR) of 1968 limited the number of landings and departures at the following airports: EWR, Reagan National Airport (DCA), John F. Kennedy International Airport (JFK), LaGuardia Airport (LGA), and O'Hare airport (ORD). Airlines operating at the airports under the provision of the HDR must reserve a spot for flights during the peak hours, which is called a “slot”. By assigning the number of slots to each airline, the intent of the HDR was to reduce airport congestion at EWR, DCA, JFK, LGA, and ORD. By 1970, however, EWR was exempted from the HDR because the airport congestion was manageable without the government intervention.[5]

Under the original HDR of 1968, a slot scheduling committee determines the allocation of slots for both scheduled and unscheduled flights, and for air carriers other than air taxis. In 1985, the FAA implemented a more systematic procedure for slot allocation by establishing a Buy/Sell Rule. The Buy/Sell rule allows carriers to trade and lease slots at their discretion. To ensure that incumbents do not monopolize the allotment of slots, the FAA complemented the Buy/Sell rule with conditions that would guarantee new entrants a certain number of the slots. Despite the antitrust effort, however, the FAA was constantly criticized by new entrants and small carriers for its inability to guarantee a fair distribution of slots. The new entrants or small carriers could not expand their services at slot-restricted airports because incumbents would hide slots available for trade and refuse to sell or lease slots with a reasonable condition.

Because of the FAA's lack of effort to ensure a fair distribution of slots and lack of transparency, competition at slot-restricted airports was tougher for the new entrants and small carriers.

The Aviation Investment and Reform Act of the 21st Century (AIR-21) was enacted in April 2000 to repeal and replace the HDR. Deregulation of slot restrictions received mixed responses from the airline industry. The major airlines, such as American Airlines and United Airlines, claimed that the elimination of slot restrictions is unfair and ineffective. In contrast, low-cost carriers welcomed the AIR-21, praising it for enhancing fair competition and improving market efficiency. While the new AIR-21 rule may have changed the competition structure for airlines, it also leads to increases of airport congestion and decreases of on-time performance.[6] For instance, the average daily operations at JFK increased by 21% from 2006 to 2007, and on-time performance at JFK decreased from 69% in 2006 to 62% in 2007.

Because of the increased daily operations at JFK and LGA, a substantial number of flights were carried over to EWR, making one of the busiest and most delay-prone airports in the states. Poor infrastructure and frequent severe weather at EWR further aggravated the congestion situation. To address the congestion issue at EWR and other NYC airports, the FAA temporarily returned to limiting the slots of departures and landings based on the airspace and runway capacity in June 2008. Under the policy, EWR is one of the International Air Transport Association (IATA) Worldwide Slot Guidelines Level 3 slot-controlled airport with a runway capacity of 81 operations per hour.[7] The capacity restriction would remain in effect for the summer scheduling season of months between April and October 2008. The FAA renewed the policy for the next summer scheduling season, which would be between April and October 2009.

The effectiveness of slot restrictions at EWR is clear in Figure 1. Using the data from the Bureau of Transportation Statistics (BTS), we track the frequency of flight departures at EWR before and after the policy implementation. The pre-policy period ranges from 2006:Q1 to 2008:Q2 (depicted by the dash line), whereas the post-policy period ranges from 2008:Q2 to 2010:Q4 (depicted by the solid line). The x-axis represents the scheduled departure hours from 06:00 to 24:00, whereas the y-axis represents the frequency of flight departures. For instance, approximately 33 flights are scheduled to depart at 06:00 before the policy, whereas the number for that of the post-policy period is reduced to approximately 28. The top three busiest times at EWR are between 06:00 and 07:00, 08:00 and 09:00, and 15:00 and 16:00 with over 30 flights scheduled to depart at each of those hours before the FAA's intervention. The efficacy of the intervention is clear as the frequency of flight departures has categorically decreased throughout the peak hours during the post-policy period. The slot restrictions do not apply to flight operations after 20:00, so we observe that the two curves converge to an almost identical frequency for hours after 20:00.

FIGURE 1 The frequency of flight departures at EWR before and after slot restriction. The pre-policy period ranges from 2006:Q1 to 2008:Q2, which is depicted by the dash line, whereas the post-policy period ranges from 2008:Q2 to 2010:Q4, which is depicted by the solid line. The x-axis represents the scheduled departure hours from 06:00 to 24:00, whereas the y-axis represents the frequency of flight departures

Source: Authors' own calculations from Bureau of Transportation Statistics

The apparent reduction in scheduled operations at EWR has led to substantial improvements in on-time performance at EWR. According to the FAA,[8] on-time gate arrivals and departures at EWR from 2007 to 2015 have increased by approximately 11% and 3%, respectively. On average, the frequency of arrival and departure delays are both decreased by approximately 33%. Slot restrictions also reduce the duration of each delay, as the delays that are longer than 60 min are down by 37% for arrivals and 38% for departures. Citing these improvements in on-time performance, the FAA removed any slot restrictions at EWR in October 2016 and reclassified EWR to a Level 2 IATA airport.[9]

4 | HYPOTHESIS

Our main hypothesis pertains to the association between the quantity restriction on departure slots at EWR and price dispersion. The hypothesis is derived from a screening model. The screening model is a classical textbook model in which a decision maker designs mechanism to discriminate users with private information by their self-selection. The model is widely used for analysis of nonlinear pricing (Maskin & Riley, 1984), insurance (Rothschild & Stiglitz, 1978), government regulation (Laffont & Tirole, 1986), and auction (Myerson, 1981).[10] For the airline industry, the model is used to analyze consumer's choice of business or economy class tickets (Anderson & Renault, 2011), refundable or non-refundable tickets and timing of ticket purchasing (Escobari & Jindapon, 2014). Empirically, Escobari and Hernandez (2019) test the use of price discrimination in the airline industry. The screening model naturally explains price dispersion. The airline carrier can take advantage of the passengers' differentiated willingness to pay for quality and provide services of differentiated qualities at different prices. The quantity restriction from the policy changes the cost of the carriers and thus the optimal quality-price combination.

Concretely, suppose there are two kinds of passengers: high-value passengers (b) or low-value passengers (l), with fraction λ and $1-\lambda$ on the market.[11] The two types of passengers have a different evaluation of the on-time performance of the flights. Specifically, denote the utility of high-value and low-value passengers from consuming the service of quality (the on-time performance) q as $u(q, \theta_b)$ and $u(q, \theta_l)$, respectively. The utility $u(q, \theta_j)$ is increasing and concave in q for both types of passengers. In addition, high-value passengers always care the quality more than low-value passengers, or $u(q, \theta_b) > u(q, \theta_l), \forall q$ and $u'_q(q,\theta_b) > u'_q(q,\theta_l)$, $\forall q$.

The carrier provides services of quality High (q_H) and Low (q_L) to the passengers.[12] The cost to serve each passenger is $c(q)$, which is increasing and convex.[13] In addition, there is a fixed cost F to operate on the market. The carrier sets prices to discriminate the two types of passengers and to maximize the total profit. From the participation and incentive compatibility condition, the low-value passengers choose q_L and pay $p_L = u(q_L, \theta_l)$; the high-value passengers choose q_H and pay $p_H = u(q_H, \theta_b) - (u(q_L, \theta_b) - u(q_L, \theta_l))$.[14]

Thus, the carrier's profit is

$$\begin{aligned}&\max_{q_H,q_L}\lambda(p_H - c(q_H)) + (1-\lambda)(p_L - c(q_L)) - F\\&= \max_{q_H,q_L}\lambda(u(q_H,\theta_b) - (u(q_L,\theta_b) - u(q_L,\theta_l)) - c(q_H))\\&+ (1-\lambda)(u(q_L,\theta_l) - c(q_L)) - F.\end{aligned} \tag{1}$$

The carrier chooses quality level as

$$u'_q(q_H,\theta_b) = c'(q_H) \tag{2}$$

$$u'_q(q_L,\theta_l) = c'(q_L) + \frac{\lambda}{1-\lambda}\left(u'_q(q_L,\theta_b) - u'_q(q_L,\theta_l)\right) \tag{3}$$

and the prices are $p_H - p_L = u(q_H, \theta_b) - u(q_L, \theta_b)$ and $p_L = u(q_L, \theta_c)$. While the quality provided to high-value passengers is at the socially optimal level (where the marginal utility for passengers equals the marginal cost for the carrier), the quality provided to low-value passengers is below the socially optimal level. As predicted in the screening model, q_L is distorted to reduce the incentive of high-value passengers to mimic the low-value passengers.

Having solved the model, we consider the effect of slot controls on the prices and service quality. The slot control changes the cost for carriers. From FAA's introduction of slot allocation,[15] carriers submit slot requests twice a year and FAA coordinates and responses to their proposals. And the process heavily relies on the "historical slots."

From this description of slot allocation, the fixed cost may have increased, but the marginal cost is reduced. The carrier has to make efforts for the proposal twice a year and this request process increases the fixed cost. Since there is no evidence of slot auction in the slot allocation,[16] the request process does not have a direct effect on the marginal cost. On the contrary, since the slot control limits the number of flights in a given time period, it is easier for the carrier to maintain on-time services, so the marginal cost is decreased.

When $c(q)$ decreases for all q, q_H and q_L increase with general restrictions, and so are p_L and $p_H - p_L$. That is, both the price and the price dispersion increase when $c(q)$ decreases. For better illustration, we use $u(q, \theta_j) = \theta_j q$, $j \in \{b, l\}$ and $c(q) = \frac{c}{2}q^2$ where c is a parameter for cost function. With this model specification, from (2) and (3), $q_H = \frac{\theta_b}{c}$ and $q_L = \frac{1}{c}\left(\theta_l - \frac{\lambda}{1-\lambda}(\theta_b - \theta_l)\right)$. Both q_H and q_L decrease in c when $\theta_l - \frac{\lambda}{1-\lambda}(\theta_b - \theta_l) > 0$ which happens when the fraction of high-value passengers or the incremental preference for service quality is not too large. For prices, we have $p_L = \theta_l q_L$ and $p_H - p_L = \theta_b(q_H - q_L) = \frac{1}{c(1-\lambda)}\theta_b(\theta_b - \theta_l)$. Therefore, both prices and price dispersion decrease in c. The interpretation is intuitive: with lower cost, the carrier provides better service to both high- and low-value passengers and charge higher prices in return. The high-value passengers value the improved quality more than the low-value passengers, so p_H is increased more than p_L.

Figure 2 illustrates the effect of marginal cost reduction in quality, price, and price dispersion. The x-axis and y-axis are the quality and price, respectively. The indifference curve for the high-value and low-value passengers ($\theta_j q - p = constant$) are depicted by green and blue dashed lines. The line for high-value passengers (green line) is steeper for their higher valuation for quality. The orange curves are the iso-profit curve ($p - \frac{c}{2}q^2 = constant$) for the carriers. At optimal prices, the iso-profit curve is tangent to the indifference line for high-value passengers but not tangent to the one for low-value passengers (recall that q_L is distorted to deter high-value passengers to mimic low-value ones). The optimal price-quality (p_H, q_H) and (p_L, q_L) is marked as point A and B. With lower marginal cost, the iso-profit curves become flatter, depicted by the red curves. The curves become tangent to the indifference curve for the high-value passengers at a higher value of q, and the new price-quality is marked as point A' and B'. Comparing points A– A' and B–B', it is straightforward to see that both the qualities and prices are increased with lower cost. Also, the increase of price p_H is higher than that of p_L, leading to a higher price dispersion.

Regarding consumer welfare, the low-value passengers are not affected by the cost reduction as the benefit from better quality is totally seized by the carrier by increased ticket price. For the high-value passengers, the welfare is improved. As q_L is higher, the benefit for high-value passengers to mimic the low-value ones is increased. Thus, the

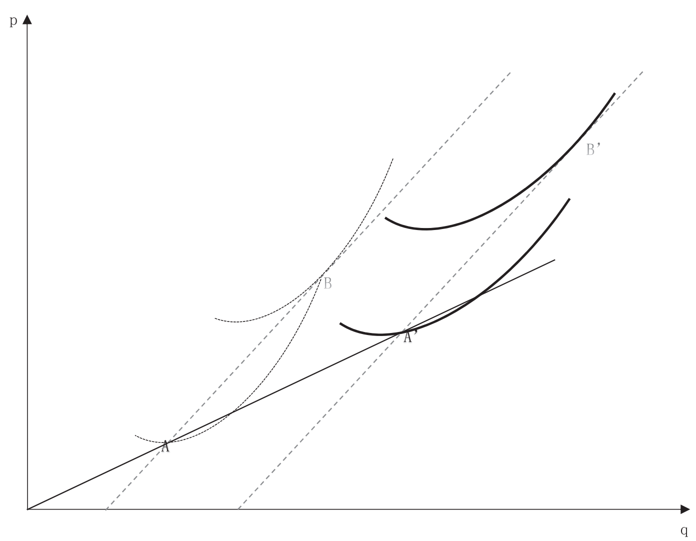

FIGURE 2 The effect of marginal cost reduction on Price and on-time performance. The x-axis and y-axis are the quality and price, respectively. The dashed lines are the indifference curve for the high-value (dashed line) and low-value passengers (solid line), represented by $\theta_j q - p = constant$. The dashed curves are the iso-profit curve for the carrier, represented by $p - c * q^2 = constant$. When the marginal cost decreases, the iso-profit curves become flat (bold solid curve), which leads to an increase in both the quality level and prices. Also, the price dispersion increases. For the welfare of passengers: The welfare for low-value passengers is not affected as the increase in price cancels out the benefit from increased quality (equilibrium quality-price, represented by A and A' in the figure, lies on the same indifference curve). For the high-value passengers, the welfare is improved (the new equilibrium quality-price represented by B' lies on a higher indifference curve), as the benefit from mimicking the low-value passengers increased from improved service quality

increased p_H cannot take away all the benefit of improved q_H. The welfare implication for the high-value passengers is consistent with the empirical findings. One caveat for this welfare implication: our model builds upon the assumption that passengers of the same type have the same evaluation of the service. Thus, the policy change will have the same welfare effect on all passengers of the same type. If we deviate from this homogeneous evaluation assumption, the general principle of marginal/intramarginal passengers difference applies. The increased price will affect passengers differently, and the passengers can benefit or be harmed by the increased price when their disutility from a higher price is lower or higher than the marginal passengers. The heterogeneous evaluation model can predict a welfare change in both directions, depending on a distributional assumption of evaluation. Thus, we leave the complete analysis on welfare to empirical studies in next sections.

Our model can be modified to accommodate peak load pricing as well. Throughout the model, we assume that the quantity restriction from the policy decreases the marginal cost without reaching the capacity limit to serve the high-value passengers. If not all high-value passengers can be served by the high quality service, the effect is equivalent to a smaller λ. In this scenario, q_H is unaffected but q_L increases. In response, p_L is also increased. The price dispersion is still increased, but by a smaller extent as compared to no capacity limits. The qualitative predictions in quality, price, and price dispersion are not changed by this modification.[17,18]

In summary, we show that when the regulation policy decreases the cost of providing services, the carrier will increase the quality provided to both high-value and low-value passengers. The price level and price dispersion also increase. The model prediction is consistent with the correlation pattern we find in the data, as we will show from the next section.

5 | DATA AND SUMMARY STATISTICS

5.1 | Data source

Four databases are used for our analysis. For observations pertaining to general flight information, we use the Airline Origin and Destination Survey (DB1B) coupon data. The DB1B coupon data is a 10% sample of reported airline tickets, collected by the Office of Airline Information of the Bureau of Transportation Statistics. The DB1B data includes origins and destinations, and other detailed itinerary information related to air travelers measured at a flight level. For information regarding the aircraft and airport capacity, we use the T-100 segment data. The T-100 segment data contains the nonstop segment of domestic flights, which are provided by the US air carriers at a flight level. The flight frequency, aircraft capacity, and load factor for the US airports are also reported in the data. For the flight arrivals and departures, we use the on-time performance data from the Bureau of Transportation Statistics (BTS). The on-time performance data includes departure and arrival delays, tail numbers, and distances for nonstop domestic flights by major airlines. The tail numbers allow us to pin down a unique information for each flight and they are used for identifying a flight network within a given day. The observations are reported at a flight level. Finally, for airlines' financial information, we use the form 41 financial data section of the Air Carrier Financial Reports. The reports contain the balance sheet information at a carrier level for the carriers with more than $20 million in their revenues. Operating expenses, total income, and assets are key variables included in the dataset.

All the observations except the financial variables are measured at a flight level. Therefore, we observe ticket prices and delay information at an individual carrier-route combination. The disaggregated level of observations allows us to observe different prices and delays for each flight. For our empirical analysis, we combine observations from all three databases and merge them into observations at a carrier-route level. These observations are combined with the financial variables and then aggregated to the carrier-route-year-quarter level to make estimation more manageable.

The following procedure explains the general data cleaning process. First, we restrict our sample to nonstop flights.[19] Second, we build our sample from domestic one-way flights. Therefore, the fares for round-trips are divided by two. Third, we only use the sample that has a value of the directional fare greater than $10 to avoid any frequent flier tickets or tickets with a key-punch error. Fourth, we use a relatively permanent route within a quarter by limiting the sample with at least 10 scheduled departures for a given route. This exercise avoids noises from temporary routes. Fifth, routes that do not have the on-time performance information are dropped from the sample. Sixth, we drop the carrier-route combination if either endpoint airport includes LGA or JFK since the flights from these airports are affected by the policy at LGA or JFK.

5.2 | Measurements for prices and delays

Price dispersion refers to a dispersion in ticket prices in the US dollars for each carrier-route-year-quarter combination. Following Borenstein and Rose (1994) and Gerardi and Shapiro (2009), the Gini coefficient is used as a measure for price dispersion. For a carrier-route combination *ij* at quarter *t* with the corresponding ticket price *r* measured at the quarterly level, the Gini coefficient is

$$Gini_{ij,t} = \frac{2}{N^2 \overline{fare}} \sum_{r=1}^{N} \left(R_r - \frac{N+1}{2} \right) fare_r, \tag{4}$$

where $\overline{fare}$ is the average ticket price (i.e., $\sum_{r=1}^{N} fare_r / N$) and R_r is the ranking of ticket price from low to high. If $Gini_{ij,t}$ is equal to 0, then the ticket price for the carrier-route combination *i* is not deviated from the average ticket price, meaning no price dispersion. For the extreme value of $Gini_{ij,t} = 1$, this combination has an absolute deviation from the average fare. Following the interpretation in Gerardi and Shapiro (2009), Equation (4) shows that the Gini coefficient is equal to two times the expected absolute difference between two fares drawn randomly from the entire data set. For instance, the median Gini value from Gerardi and Shapiro (2009) is 0.22. The Gini coefficient implies that the expected difference in the average ticket prices for two passengers on a certain carrier-route combination drawn randomly from the population is 44%.

A flight refers to the aircraft with a unique tail number and this aircraft has its flight network within a given day. A delay is a measure of how much arrival and departure delays exist between each of the routes that the aircraft takes within its flight network. In our analysis, we define three different delays: arrival, layover, and departure. Arrival delays refer to the number of minutes in delays that the flight has accumulated prior to arriving at EWR. Layover delays refer to the number of minutes in delays that the flight has accumulated at EWR alone before departing the airport. Slot restrictions at EWR would directly affect the layover delays. Departure delays refer to the total number of minutes in delays that the departing flight has accumulated during the entire flight network. Therefore, the sum of arrival delays and layover delays is equal to departure delays. We do not directly observe layover delays from our data, but we can calculate layover delays as the difference between departure delays and arrival delays.

For instance, we observe from our data that an aircraft with a tail number N956AT starts from Indianapolis at 05:45 on January 17, 2006, and finishes the trip in Atlanta at 23:59 on the same day. The last two routes for N956AT are from Atlanta to Newark and from Newark to Atlanta. Both flights are nonstop. We also observe the following information from the data: scheduled departure time, actual departure time, scheduled arrival time, and actual arrival time. For the route from Atlanta to Newark, the scheduled arrival time was 19:05, but the flight arrived at 19:39. Therefore, the arrival delay time was 34 min. For the route from Newark to Atlanta, the scheduled departure time was 19:45, but the flight departed at 21:09. Therefore, the departure delay time was 84 min. Given that the arrival delay time was 34 min, the layover delay time was 50 min.

5.3 | Summary statistics

We provide the summary statistics of all relevant variables in Table 1 by the treatment status. We define the treated group as routes that are departing from or arriving at EWR between 2006:Q1 and 2010:Q4. The treated units are divided by the pre-treatment period (i.e., between 2006:Q1 and 2008:Q2) and the post-treatment period (i.e., between 2008:Q2 and 2010:Q4). The control group refers to all flights for which EWR is neither the departure nor arrival airport. The control groups are also divided by the pre- and the post-treatment periods. Out of the total of 97,926 observations in the full sample, there are approximately 2,000 observations for each of the treated before and the treated after groups, whereas there are approximately 47,000 observations for each of the control before and the control after groups. We choose the range of sample periods to avoid confounding factors from other events such as the merger of United Airlines (UA) and Continental Airlines in October 2010. After the merger, UA has been the most dominant carrier at EWR, accounting for 73% of daily operations as of August 2014.[20] If we extend our sample period beyond the fourth quarter of 2010, it would be difficult to identify the effects of slot restrictions because the merger itself would have a significant effect on flights operating at EWR. Therefore, we restrict our full sample ranges to be between the first quarter of 2006 and the fourth quarter of 2010.

TABLE 1 Summary statistics by treatment status

	Full sample		Treated				Control			
			Before		After		Before		After	
	(1) Mean	(2) SD	(3) Mean	(4) SD	(5) Mean	(6) SD	(7) Mean	(8) SD	(9) Mean	(10) SD
Price	210.151	85.249	254.069	73.483	267.459	79.527	205.526	74.101	210.610	94.404
Price Dispersion	0.256	0.071	0.282	0.045	0.314	0.057	0.255	0.067	0.253	0.075
Arrival Delay	10.800	7.109	17.393	9.465	14.390	8.680	11.820	7.018	9.370	6.619
Layover Delay	1.695	7.499	7.606	11.759	6.705	10.668	1.469	7.517	1.472	6.906
Departure Delay	12.495	8.369	24.999	11.800	21.095	10.335	13.289	8.665	10.842	6.920
Distance	774.770	519.955	1,021.455	692.970	958.137	669.358	770.755	515.727	761.132	504.754
$Frequency_{Flight}^{Origin}$	38.625	32.441	37.190	21.752	35.023	21.612	40.108	33.318	37.359	32.225
$Frequency_{Flight}^{Destination}$	37.714	32.613	36.853	21.490	34.745	21.945	39.146	33.607	36.449	32.290
LCCs	0.325	0.468	0.056	0.230	0.060	0.237	0.329	0.470	0.343	0.475
LoadFactor	0.764	0.115	0.760	0.111	0.747	0.121	0.757	0.113	0.772	0.117
Passengers	24.254	27.280	26.539	29.306	24.230	27.615	24.844	27.603	23.574	26.836
ln(Asset)	15.510	1.297	15.209	1.521	14.578	1.857	15.633	1.180	15.439	1.347
Cost/Asset	0.226	0.145	0.432	0.263	0.421	0.243	0.210	0.123	0.225	0.141
Revenue/Asset	0.230	0.131	0.431	0.248	0.395	0.197	0.218	0.117	0.228	0.121
Observations	97,926	97,926	1,959	1,959	1,902	1,902	46,851	46,851	47,214	47,214

Notes: The sample period is between 2006:Q1 and 2010:Q4 and the sample is observed at the carrier-route level. The treated group refers to flights that were affected by slot restrictions at EWR. The control group refers to the rest of the observations that were not affected by the policy. Before refers to the pre-policy periods (i.e., 2006:Q1–2008:Q2) and after refers to the post-policy periods (i.e., 2008:Q2–2010:Q4). $Frequency_{Flight}^{Origin}$, $Frequency_{Flight}^{Destination}$, and *Passengers* are normalized by 1000. The financial variables are observed at the carrier level.

The summary statistics suggest that slot restrictions at EWR have increased both the level and dispersion in prices. The average value of price is approximately $210 for the full sample. For the control group, the average price increases by only five dollars after the treatment. In contrast, the average price is approximately $254 for the pre-treatment group and is $267 for the post-treatment group. From this preliminary analysis, we see that slot restrictions at EWR have increased the level of ticket prices for the treated group. The average value of price dispersion for the full sample is 0.256. The average values before and after the treatment are 0.255 and 0.253, respectively for the control group, which are similar to the average value for the full sample. In contrast, for the treatment group, price dispersion before the treatment is 0.282, which is already higher than the average value for the full sample. After the treatment, price dispersion increases to 0.314. That is, slot restrictions at EWR have increased price dispersion.

Furthermore, there is evidence suggesting that slot restrictions at EWR may have decreased arrival delays, layover delays, and departure delays. The average arrival delay time is approximately 10.8 min for the full sample. The average value for the control group slightly decreased from 11.8 min before the treatment to 9.4 min after the treatment. The treatment group has experienced a much more significant decrease in arrival delay time as the average values decreased from 17.4 to 14.4 min. Layover delay time for the treatment group also decreases from 7.6 to 6.7 min. We notice that EWR was a delay-prone airport as the average values of layover delays of 7.6 min (before 2008:Q2) and 6.7 min (after 2008:Q2) are much higher than the full sample average of 1.7 min. The evidence suggests that slot restrictions have decreased both arrival and layover delays.

Because both arrival and layover delays are decreased, it follows naturally that departure delays, which are the sum of arrival and layover delays, are also reduced. The average value of departure delays is 12.5 min for the full sample. The average value is 25 min for the treatment group before 2008:Q2, which is twice as larger than that of the full sample. The average value decreases to 21.1 min after the policy implementation. A majority of this decrease originates from a reduction in arrival delays as approximately 3 min of the 3.9 min are accounted by arrival delays. To summarize, slot restrictions at EWR improves on-time performance by mostly reducing delays for flights arriving at EWR.

We include the following key variables in the logit regression for generating the propensity score for the SDID estimators. *Distance* refers to the total miles traveled for a trip. The average value of the treated groups is significantly higher than the full sample average, but the distance decreases from the pre-policy periods to post-policy periods. $Frequency^{Origin}_{Flight}$ refers to the total frequency of flights from the origin airport in 1,000s and $Frequency^{Destination}_{Flight}$ refers to the total frequency of flights at the destination airport in 1,000s. Each of these variables measure congestion level at the origin and destination airports. *LCCs* is an indicator variable that equals to 1 if there is a low-cost carrier within the route and 0 otherwise. It is apparent that the low-cost carriers are almost always absent for the treatment group compared to the control group. *LoadFactor* measures the capacity of aircraft, which is the ratio of actual passengers on board over the full capacity of the plane. *LoadFactor* is consistently constant around 76% for all sample groups. Last, *Passengers* refers to the total number of passengers in 1,000s.

In addition, we control the carrier's financial conditions for the difference-in-differences (DID) and semi-parametric DID (SDID) estimators. The data measures the financial variables at the carrier level. $ln(Asset)$ is the log of carriers' total assets, and we also use its squared term in the control variables. $Cost/Asset$ is the ratio of operation cost over total assets of an airline, and $Revenue/Asset$ is the ratio of operation revenue over total assets. While the financial variables are important indicators for airlines' operational capacity, it is difficult to assess how slot restrictions may have affected these variables at the carrier level. For instance, the average value of revenue has decreased after the policy, but it could be because of other confounding variables such as changes in fuel prices, competition, or travelers' demands. Thus, we take the financial conditions of the carriers as control variables without further discussing the endogeneity.

6 | EMPIRICAL ANALYSIS

6.1 | Common trend assumptions

The aim of our article is to analyze the effects of slot restrictions on price, price dispersion, arrival, and departure delays at EWR. We use the DID approach for examining the effects on price and price dispersion and use the SDID approach for the analysis of the delay variables. We base the choice of different estimation methods on validity of the common trend assumption for the DID approach. We find that price variables satisfy the common trend assumption, whereas delay variables fall short of satisfying that assumption. Figure 3 shows the common trend for price, price dispersion, arrival delays, and layover delays. The sample period is between 2006:Q1–2010:Q4 at EWR. The policy was in effect during the second quarter of 2008, and the timing of the policy is depicted by the vertical line. The treated group

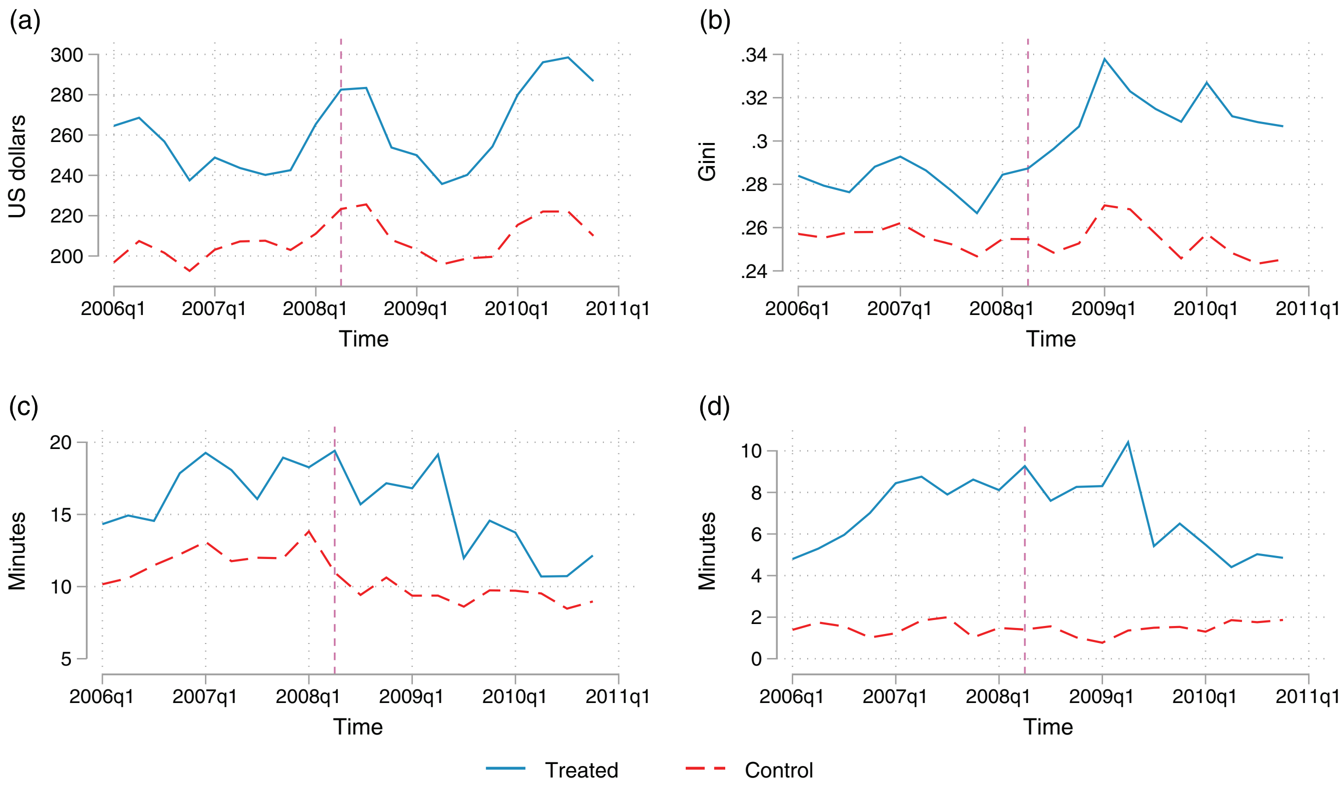

FIGURE 3 Trends in (a) price, (b) price dispersion, (c) arrival delay, and (d) layover delay. The sample period is between 2006:Q1–2010:Q4 at EWR. The policy was in effect during the second quarter of 2008 and the timing of the policy is depicted by the vertical line. The treated group consists of flights that were affected by slot restrictions at EWR and its movements are represented by the straight line. The control group consists of flights that were not affected by slot restrictions and its movements are represented by the dash line

comprises flights that were affected by slot restrictions at EWR and the solid line represents their movements. The control group comprises flights that were not affected by slot restrictions, and the dash line represents their movements.

Panels (a) and (b) show the movements for price and price dispersion, and the evidence suggests that the price variables satisfy the common trend assumption. The movements of price for the treated and control groups are closely followed before slot restrictions at EWR, which implies that price satisfies the common trend assumption. The movements of price dispersion for the pre-policy periods are also parallel for the treated and control groups, which is evidence of satisfying the common trend assumption. The effect of slot restrictions at EWR on price dispersion is apparent as the gap between the treated and control groups significantly widens after the second quarter of 2008.

Panels (c) and (d) show the movements for arrival and layover delays at EWR, and the evidence suggests that the delay variables may not satisfy the common trend assumption. The movements of arrival delay for the treated and control groups are not closely followed before slot restrictions at EWR. For instance, from 2008:Q1 to 2008:Q2 (i.e., the period right before policy implementation), arrival delays increase for the treated group, whereas arrival delays decrease for the control group. The unparalleled movement suggests that the characteristics of the treated group may not be like those of the control group. Thus, the common trend assumption may not hold true for arrival delays. Similarly, the movements of layover delays for the treated and control groups are also unparalleled.

The effects of slot restrictions at EWR on arrival and layover delays are apparent from two time periods, 2008:Q2 and 2009:Q2. The decreases in arrival and layover delays are most noticeable for 2009:Q2. The FAA renewed temporary controls implemented in 2009 during the summer scheduling season for months between April and October. Because slot restrictions aim to reduce congestion during the peak summer season, the significant dips in 2009:Q2 for arrival and layover delays reflect the FAA's effort to manage airport congestion at EWR. Therefore, the evidence suggests that slot controls at EWR had a measurable impact on improving on-time performance at the targeted airport.

To statistically examine the common trend assumptions for the pre-policy observations, we regress the price and delay variables on the treated dummy (i.e., slot-controlled airports), linear time trend, an interaction term between the treated dummy and the linear time trend, and a set of control variables (i.e., financial variables). The interaction term allows us to compare the linear time trend differences between the treated and control groups before the policy implementation. If the coefficient of the interaction term is statistically significant, it implies that there is a substantial difference between the characteristics of the treated and control groups that affect the trend.

Table 2 reports the estimation results for the common trend analysis. The first two columns show the estimates for price and price dispersion. The coefficients of the interaction term for price and price dispersion are marginal in magnitude and are not statistically significant. This finding reinforces the finding that there are no discernible differences between the characteristics of the treated and control groups for the price variables before the policy implementation. Because the common trend assumption holds statistically, we will use the traditional DID estimator for analyzing the impact of slot restrictions at EWR on price and price dispersion.

Columns 3–5 show the estimates for arrival delays, layover delays, and departure delays. The coefficient of the interaction term for arrival delays is not marginal in magnitude but is statistically insignificant. The coefficient for layover delays, however, is statistically significant with a magnitude of 0.396. It implies that the time trends for layover delays are substantially different for EWR and non-EWR airports. In other words, delays at EWR may have been altered prior to slot restrictions, violating the common trend assumption. Additionally, the characteristics of treated and control groups for departure delays are substantially different. The layover delays do not satisfy the common trend assumption, and departure delays also do not satisfy the assumption. Thus, the evidence suggests that a traditional DID approach may not examine the delay variables.

6.2 | Estimation strategy

Based on the examination on the common trend assumption, the DID estimators are used for analyzing price and price dispersion, whereas the SDID estimators are used for the delay variables. For the DID estimators, we compare the carrier-route level changes in price dispersion of the pre- and post-policy periods at EWR with the contemporaneous changes in price dispersion for airports that are not affected by slot controls. The counter-factual would be the flights departing from or arriving at airports that are not affected by slot controls. To calculate the average treatment effect on the treated (ATT) we use the following estimation model:

$$y_{ij,t} = \alpha + \beta_1 \mathrm{A}_{ij} + \beta_2 \mathrm{T}_t + \beta_3 \left(\mathrm{A}_{ij} \times \mathrm{T}_t\right) + \beta_4 X_{i,t} + \epsilon_{ij,t}, \tag{5}$$

where $y_{ij,\,t}$ is the dependent variable of a carrier i on a directional route j at quarter t; A_{ij} is a binary variable for the carrier-route with either endpoint at EWR; and T_t is the post-regulation time dummy variable. β_3, which is the coefficient of the interaction term $\mathrm{A}_{ij} \times \mathrm{T}_t$, captures the ATT of the slot control policy on the dependent variable. Last, $\mathrm{X}_{i,\,t}$ is a set of control variables and $\epsilon_{ij,\,t}$ is the error term. The control variables mostly include financial variables, which are the operation costs, revenues, and the log of assets and its squared term, measured at a carrier level. For all different model specifications, we cluster the standard error at the carrier-route level to account for correlations over time for each carrier-route combination.

TABLE 2 Comparison of pre-policy movements for treated and control groups

	Price (1)	Price dispersion (2)	Arrival delay (3)	Layover delay (4)	Departure delay (5)
Slot-controlled airport × time trend	−0.080	0.001	0.194	0.396***	0.590***
	(0.892)	(0.001)	(0.140)	(0.046)	(0.106)
Time trend	2.431***	0.000	0.200	0.012	0.212
	(0.817)	(0.001)	(0.142)	(0.032)	(0.141)
Observations	48,548	48,548	48,548	48,548	48,548
Adjusted R^2	0.906	0.812	0.570	0.587	0.630

Notes: The sample period is restricted to the pre-policy periods of 2006:Q1–2008:Q2. The estimates of the interaction term represent the difference in magnitude between the treated (i.e., flights affected by slot restrictions at EWR) and the control groups before the policy implementation. Each of the dependent variables are regressed on the treated dummy (i.e., slot-controlled airports), a time trend, an interaction term between the treated dummy and the time trend, and set of control variables. Standard errors are clustered by the carrier-route combination level and reported in parentheses.

*$p < 0.1$, **$p < 0.05$, ***$p < 0.01$.

For the SDID estimators, we follow the estimation procedure suggested in Abadie (2005). Since a traditional DID approach imposes a strong assumption on the common trend, a discrepancy in the pre-policy movements may lead to an inaccurate estimator. The FAA's concern about the potential airport congestion at EWR stems from the spillover effect of a temporary lifting of slot restrictions at LGA and JFK. This endogenous policy change is triggered by the congestion situation in 2006, leading to a nonparallel pre-treatment trend between the treated and control. To address the issue, the SDID approach assigns weights to the control group to balance the characteristics between the treated and control groups. The same distribution of the covariates is imposed for the treated and controls by weighing down covariates that are over-represented among the controls, while weighing up covariates that are under-represented among the controls.

Similar to the general framework of Rosenbaum and Rubin (1983), we estimate the average treatment effects on the treated:

$$\begin{aligned} ATT &= \mathbb{E}\left(Y^1(1) - Y^0(1) \mid D = 1\right) \\ &= \mathbb{E}\left(Y^1(1) \mid D = 1\right) - \mathbb{E}\left(Y^0(1) \mid D = 1\right), \end{aligned} \tag{6}$$

where $Y^1(1)$ is the value of $Y^1(ij, t)$ if the participant ij receives the treatment at time t; $Y^0(1)$ is the value of $Y^1(ij, t)$ if the participant does not receive the treatment at time t; and the individual argument ij is dropped for simplicity. D is a binary indicator taking a value 1 if a participant receives treatment at time t.

Because $\mathbb{E}[Y^0(1) \mid D = 1]$, which is the average outcome of the treated group had they not been treated is not observable, we cannot directly calculate ATT. Abadie (2005) shows that the following conditions are necessary for implementing a simple two-step method to estimate the ATT:

$$\mathbb{E}\left[Y^0(1) - Y^0(0) \mid X, D = 1\right] = \mathbb{E}\left[Y^0(1) - Y^0(0) \mid X, D = 0\right] \tag{7}$$

$$\mathbb{P}(D = 1) > 0 \text{ and with probability } \mathbb{P}(D = 1 \mid X) < 1, \tag{8}$$

where X is a vector of pre-treatment characteristics, in our case, the characteristics are based on the second quarter of 2008, which is the last time period before the treatment. $\mathbb{P}(D = 1 \mid X)$ is the probability of being treated or the propensity score conditional on X.

Once we have this model, the weighted average of the differences in the outcome variable recovers the ATT in the following way:

$$\begin{aligned} ATT &= \mathbb{E}\left(Y^1(1) - Y^0(1) \mid D = 1\right) \\ &= \mathbb{E}\left(Y^1(1) \mid D = 1\right) - \mathbb{E}\left(Y^0(1) \mid D = 1\right) \\ &= \mathbb{E}\left[\frac{Y(1) - Y(0)}{\mathbb{P}(D = 1)} \times \frac{D - \mathbb{P}(D = 1 \mid X)}{1 - \mathbb{P}(D = 1 \mid X)}\right]. \end{aligned} \tag{9}$$

For the logit model, we use the likelihood from for all carrier-route pairs in the same group across time. The logit model is based on the second quarter of 2008. Consequently, routes that did not operate during 2008:Q2 are dropped from our observations. We include the following covariates to generate the propensity score: distance between the routes, the total frequency of flights from the origin airport, the total frequency of flights at the destination airport, an indicator variable that shows whether a low-cost carrier serves within the route, the ratio of passengers on board in relation to the capacity of the aircraft, and the total number of passengers.

6.3 | Price and price dispersion

6.3.1 | Main results

Using a traditional DID approach, we examine the effects of slot restrictions on price and price dispersion. We formally present the estimation results in Table 3. Columns 1–3 show estimates for price and columns 4–6 show estimates for price dispersion. Column 1 includes the coefficient of β_1 (i.e., slot-controlled airport), β_2 (i.e., policy time), and β_3 (i.e.,

TABLE 3 The impact of slot restrictions on price and price dispersion

	Price			Price dispersion		
	(1)	(2)	(3)	(4)	(5)	(6)
Slot-controlled airport × policy time	16.241***	9.274***	16.660***	0.030***	0.024***	0.027***
	(3.486)	(3.093)	(2.956)	(0.003)	(0.003)	(0.003)
Policy time	12.383***	12.820***	10.139***	−0.006***	0.001	−0.001
	(1.061)	(0.901)	(0.930)	(0.001)	(0.001)	(0.001)
Slot-controlled airport	54.279***			0.031***		
	(5.698)			(0.003)		
All other control variables			Yes			Yes
Carrier-route fixed effect		Yes	Yes		Yes	Yes
Year-quarter fixed effect	Yes	Yes	Yes	Yes	Yes	Yes
Observations	97,531	97,531	97,531	97,531	97,531	97,531
Adjusted R^2	0.028	0.637	0.638	0.024	0.635	0.636

Notes: The sample period is between 2006:Q1 and 2010:Q4. A DID approach is used to retrieve these estimates. Standard errors are clustered by the carrier-route combination level and reported in parentheses. The baseline specifications for price and price dispersion are reported in columns (3) and (6), respectively. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

slot-controlled airport × policy time) in Equation (5). Columns 2 and 3 include the carrier-route fixed effects, and Column 3 further include control variables. Column 3 includes all control variables, carrier-route fixed effects, and year-quarter fixed effects, and is treated as the baseline specification. Columns 4–6 are similarly specified, and we use column 6 as the baseline specification for price dispersion.

The interaction term between the slot-restricted airport dummy and policy time indicates the effect of the policy on price. As shown in column 3, price increases by \$16.66 on average and the estimate is statistically significant. The impact is also quantitatively significant, accounting for an 8% increase from the average price of \$210. This result is intuitive because it implies that a reduction in the quantity supplied leads to an increase in ticket price. As shown in column 6, slot restrictions at EWR result in an increase of 0.027 in price dispersion and the estimate is statistically significant. As the average price dispersion value for the pre-policy periods was 0.282, the increase in price dispersion is approximately 10%. This economically significant effect of slot restrictions on price dispersion is the main finding of our analysis.[21] The underlying mechanism behind how slot restrictions increases price dispersion is not obvious. The subsequent analysis provides our best attempt to dissect the association between quantity-based approaches to airport congestion and price dispersion.

6.3.2 | Leads and lags

We have found that slot restrictions at EWR significantly increase price dispersion. The next analysis attempts to determine whether the effect was anticipated and whether the effect is lasting.[22] As an extension to Equation (5), we can estimate leads and lags of the policy effect in the following way:

$$y_{ij,t} = \alpha + \beta_1 \mathrm{A}_{ij} + \sum_{\tau=-m}^{+j} \gamma_\tau \mathrm{T}_{t_e+\tau} + \sum_{\tau=-m}^{+j} \omega_\tau \left(\mathrm{A}_{ij} \times \mathrm{T}_{t_e+\tau}\right) + \beta_2 X_{i,t} + \epsilon_{ij,t}, \tag{10}$$

where $y_{ij,\,t}$ is the dependent variable of a carrier i on a directional route j at quarter t; A_{ij} is a binary variable for the carrier-route with either endpoint at EWR; $T_{t_e+\tau}$ are time dummies surrounding the announcement period of slot-restrictions, in which all the time dummies are mutually exclusive. The leads time dummies are assigned a value of one for time periods $t_e+\tau$, where $\tau = \{-8, ..., -2, -1\}$, the lags time dummies are assigned a value of one for time periods $t_e+\tau$, where $\tau = \{1, ..., 7, 8\}$, with t_e being the time in which the policy is announced. These mutually exclusive time dummies cover eight quarters before and after the base period. Therefore, ω_τ, which is the coefficient of the interaction term between A_{ij} and leads/lags, captures the ATE of slot restrictions on the dependent variable relative to the value of the policy announcement period. Last, $X_{i,t}$ is a set of control variables and $\epsilon_{ij,t}$ is the error term. For all different model specifications, we cluster the standard error at the carrier-route level to account for correlations over time for each carrier-route combination.

The timing of slot restrictions at EWR is the second quarter of 2008 (i.e., time t in the leads and lags analysis), and the estimates for t follows the baseline specification in column 6 of Table 3. We express all the estimates in relation to t. For instance, $t+1$ is the one quarter lead of the treatment effect, whereas $t-1$ is the one quarter lag of the treatment effect. Table 4 reports estimates from the leads and lags analysis. Prior to the policy implementation, there is a constant reduction in price dispersion. The movement is significant up to three quarters prior to slot restrictions at EWR. After the policy, however, there is a sharp increase in price dispersion of 0.015. The trend is magnified up to $t-3$ with the coefficient of 0.028. The lags of the policy are continuing up to $t-8$ with the coefficient of 0.022. This finding suggests that the slot control policy has a persistent positive impact on price dispersion.

6.3.3 | Price quantile analysis

Given the evidence of a lasting effect of slot restrictions at EWR on price dispersion, we identify the source of price dispersion in this analysis. As indicated from the literature and our theoretical model for hypothesis, the consumer types play an important role in determining variations in ticket prices after quantity-based approaches to airport congestion management are implemented. Our hypothesis predicts that high-value passengers have a higher willingness to pay for the increased quality provided by the carriers after slot restrictions, which may lead to a higher increase in prices for these passengers than the low-value passengers. Hence, we divide the ticket prices into 10 equal groups to identify the heterogeneous effects of slot restrictions on different levels of flight tickets to determine the origin of price dispersion, whether it originates from a higher increase in high-value passengers than low-value passengers.

TABLE 4 A leads and lags analysis: Slot restrictions and Price dispersion

	Price dispersion (1)		**Price dispersion (2)**
Slot restriction$_{t+8}$	−0.015***	*Slot restriction*$_{t-1}$	0.015***
	(0.003)		(0.003)
Slot restriction$_{t+7}$	−0.013***	*Slot restriction*$_{t-2}$	0.019***
	(0.003)		(0.003)
Slot restriction$_{t+6}$	0.001	*Slot restriction*$_{t-3}$	0.028***
	(0.003)		(0.005)
Slot restriction$_{t+5}$	−0.000	*Slot restriction*$_{t-4}$	0.015***
	(0.003)		(0.004)
Slot restriction$_{t+4}$	−0.005	*Slot restriction*$_{t-5}$	0.011***
	(0.003)		(0.004)
Slot restriction$_{t+3}$	−0.009***	*Slot restriction*$_{t-6}$	0.012***
	(0.003)		(0.004)
Slot restriction$_{t+2}$	−0.013***	*Slot restriction*$_{t-7}$	0.036***
	(0.003)		(0.004)
Slot restriction$_{t+1}$	−0.009***	*Slot restriction*$_{t-8}$	0.022***
	(0.003)		(0.003)
Observations	97,531	Adjusted R^2	0.633

Notes: The sample period is between 2006:Q1 and 2010:Q4. A DID approach is used to retrieve these estimates. Standard errors are clustered by the carrier-route combination level and reported in parentheses. Time t is equal to the second quarter of 2008. All the estimates are expressed in relation to t. For instance, $t+1$ is the one quarter lead of the treatment effect, whereas $t-1$ is the one quarter lag of the treatment effect. The estimates follow the baseline specification in column 6 of Table 3.

*$p < 0.1$, **$p < 0.05$, ***$p < 0.01$.

To calculate price quantiles, we use the DB1B data of a 10% random draw of the total tickets. We rank each of these tickets from the lowest price to the highest price. We then divide them into 10 distinct groups and calculate the average price for each price quantile. Then the sample is aggregated into the carrier-route-year-quarter level. Table 5 reports the estimates from the quantile analysis of how the average price of each quantile changes after the policy implementation. The estimates are from the baseline specification in Table 3. By construction, column 1 refers to the first decile of the price distribution (i.e., cheapest), column 2 refers to the second cheapest ticket group, and column 10 refers to the highest ticket price group. In this way, lower quantiles represent price groups for low-value passengers, whereas higher quantiles represent price groups for high-value passengers.

The results show that there is a drastic difference in price changes between the different passenger types. The coefficients for the first five quantiles (i.e., columns 1 through 5) are negative and statistically insignificant. Considering that the lower quantiles represent price groups for low-value passengers, the finding suggests that slot restrictions did not affect the ticket prices for the low-value passengers. On the other hand, the coefficients for the last five quantiles (i.e., columns 6 through 10) are positive and statistically significant except for column 6. Since the higher quantiles represent price groups for high-value passengers, the finding suggests that slot restrictions mostly affected the ticket prices for high-value passengers. Particularly, the price increases for the top three most expensive groups are most drastic, with the surge ranging from approximately 36 to 65 dollars. In other words, the increase in price dispersion originates from a higher rise in prices for high-value passengers, while the prices for low-value passengers remain constant. This is consistent with our main hypothesis that the high-value passengers drive up the surge in price dispersion.

Given that the prices increase more for high-value passengers, we examine whether the number of passengers who paid high ticket prices changes. If there are more high-value passengers after the policy, the social benefit for these groups of passengers can increase because of the time-valuation effect as shown in Czerny and Zhang (2015). For the welfare evaluation of slot restrictions, we use a similar empirical approach as in Table 5, in which we divide the DB1B data by 10 distinct groups. Then we calculate the proportion of passengers who fall into each price quantile over the entire passenger size and aggregate the samples into the carrier-route-year-quarter level. Therefore, the dependent variable is the fraction of passengers in each price quantile over the entire passenger size. The number of observations of each quantile is equal to the full sample because we are using the aggregated data.

TABLE 5 The impact of slot restrictions on price: A quantile analysis

	1st (1)	**2nd (2)**	**3rd (3)**	**4th (4)**	**5th (5)**
Slot-controlled airport × policy time	−1.558	−0.766	−0.494	−1.845	−3.307
	(1.161)	(1.565)	(1.858)	(2.043)	(2.371)
Policy time	10.008***	13.232***	11.643***	8.816***	5.612***
	(0.454)	(0.592)	(0.642)	(0.729)	(0.874)
Observations	97,531	97,254	97,237	97,263	97,224
Adjusted R^2	0.744	0.777	0.800	0.808	0.815
	6th (6)	**7th (7)**	**8th (8)**	**9th (9)**	**10th (10)**
Slot-controlled airport × policy time	0.084	13.870***	36.380***	54.454***	65.191***
	(3.006)	(4.234)	(5.113)	(5.624)	(16.203)
Policy time	2.367**	1.537	4.940***	10.559***	33.192***
	(1.041)	(1.245)	(1.550)	(1.880)	(3.326)
Observations	97,370	97,346	97,310	97,357	97,353
Adjusted R^2	0.823	0.832	0.845	0.856	0.223

Notes: The sample period is between 2006:Q1 and 2010:Q4. A DID approach is used to retrieve these estimates. Standard errors are clustered by the carrier-route combination level and reported in parentheses. The estimates show how the average price of each quantile changes after the policy implementation and the estimates follow the baseline specification in Table 3. Column 1 refers to the first decile of the price distribution (i.e., cheapest), column 2 refers to the second cheapest ticket group, and column 10 refers to the highest ticket price group.

*$p < 0.1$, **$p < 0.05$, ***$p < 0.01$.

Table 6 reports the estimation results from the quantile analysis by the proportion of passengers who pay different prices. The estimates show the percent change in the fraction of passengers of each quantile after the policy implementation and the estimates follow the baseline specification used for the results in Table 5. Column 1 refer to the proportion of passengers who purchased the lowest ticket group. Column 2 refer to the proportion of passengers who purchased the highest ticket group. The major difference between columns 1 and 2 is whether the specification includes a time trend. The time trend accounts for any changes in compositions of the cheapest ticket group over time. We can identify the demand change for the cheapest ticket group after the policy when the time trend is included (i.e., column 1). Similarly, column 4 reports the estimates of the demand change for the highest ticket group after the policy.

Column 1 shows that the proportion of passengers who purchased the cheapest tickets is decreased by 1.5% because of slot restrictions. In contrast, column 2 shows that the proportion of passenger who purchased the most expensive tickets has increased by 0.9% after the policy implementation. In other words, the demand for the highest ticket price is actually increased. The finding shows that high-value passengers can be better off after slot restrictions because the proportion of passengers who purchased high price tickets has increased despite the increase in ticket prices. As shown in Czerny and Zhang (2015), this is because of the time-valuation effect, which plays a significant role in determining the movement in demand for high-value passengers.

6.3.4 | Extensions

Given our findings that price dispersion is increased because of a higher increase in ticket prices for high-value passengers than low-value passengers, we further analyze how our estimates compare to the median Gini coefficient values of Gerardi and Shapiro (2009). We present our median Gini coefficient in Table 7. Column 1 shows that the median value for the full sample is 0.255, which is slightly higher than the median Gini coefficient found in Gerardi and Shapiro (2009), which is 0.22 from the sample observed between 1993:Q1 and 2006:Q3. The slight difference in the median Gini originate from different sample periods between the two studies.[23] Nonetheless, our median Gini coefficient value for the full sample is not much different from the existing literature.

While the median Gini coefficient for our sample is similar to that of the previous study, however, the median Gini coefficient value for our treatment group is substantially different. In particular, the treated group of our sample has a much higher value of the median Gini coefficient. Columns 2 and 3 present the median Gini coefficient for the sample before and after the treatment. The value for the pre-policy period is 0.285, which is already higher than the full sample value. The value increases from 0.285 to 0.312 after the policy implementation. In other words, the expected difference in the average ticket prices for two passengers on a certain carrier-route combination drawn randomly from the population is approximately 62% after slot restrictions at EWR.

TABLE 6 The impact of slot restrictions by passenger types: A quantile analysis

	Fraction		Count	
	1st (1)	10th (2)	1st (3)	10th (4)
Slot-controlled airport × policy time	−0.015***	0.009***	−55.209***	11.378***
	(0.004)	(0.002)	(16.014)	(2.555)
Policy time	0.017***	−0.017***	7.569***	−43.128***
	(0.002)	(0.002)	(2.694)	(4.398)
Observations	97,531	97,531	97,531	97,531
Adjusted R^2	0.265	0.796	0.661	0.851

Notes: The sample period is between 2006:Q1 and 2010:Q4. A DID approach is used to retrieve these estimates. Standard errors are clustered by the carrier-route combination level and reported in parentheses. The estimates show the percent and level change in the fraction of passengers of each quantile after the policy implementation and the estimates follow the baseline specification in Table 3. Columns 1 and 2 refer to the percentage change in the fraction of passengers of each quantile. Columns 3 and 4 refer to the number of passenger level change of each quantile.

*$p < 0.1$, **$p < 0.05$, ***$p < 0.01$.

TABLE 7 Median price dispersion and quantile analysis by passenger types

		Treated		Control	
	Full sample (1)	Before (2)	After (3)	Before (4)	After (5)
Median Gini	0.255	0.285	0.312	0.260	0.246
10% price percentile	82.686	99.996	100.284	79.274	84.645
90% price percentile	457.643	548.999	636.481	431.953	472.136
Share of passengers in 10% price percentile	0.121	0.136	0.129	0.120	0.122
Share of passengers in 90% price percentile	0.083	0.060	0.069	0.086	0.082
Observations	97,926	1,959	1,902	46,851	47,214

Notes: The sample period is between 2006:Q1 and 2010:Q4. The treated group refers to flights that were affected by slot restrictions at EWR. The control group refers to the rest of the observations that were not affected by the policy. Before refers to the pre-policy periods (i.e., 2006:Q1–2008:Q2) and after refers to the post-policy periods (i.e., 2008:Q2–2010:Q4).

Additionally, we report the changes in average prices for both the 1st and 10th percentile groups before and after the policy implementation. The lowest price group has the average price of approximately 83 dollars for the full sample, whereas the treated group has a higher average than the full sample. Specifically, both of the pre- and post-policy samples of the treated groups have the average prices of approximately 100 dollars. Consistent with our finding in Table 5, slot restrictions have not affected the lowest price group as much since there are almost no changes between the pre-policy group and the post-policy group. On the other hand, the average prices for the highest price group have drastically changed from approximately 549 dollars to 636 dollars after the policy.

Similarly, we also report the changes in the proportion of passengers within the 1st and 10th percentile groups before and after the policy implementation. Consistent with our finding in Table 6, the share of passengers within the lowest percentile group has decreased from 13.6% to 12.9% after the policy. On the other hand, the share of passengers within the highest percentile group has increased from 6% to 6.9% after slot restrictions at EWR. This pattern is specific to the treated group, as the variations in the control group are trivial before and after the policy. All the extended analysis confirms that price dispersion is driven by the high-value passengers, while the share of high-value passengers has increased after the policy. This analysis implies that the welfare of the high-value passengers may have increased since the policy implementation.

6.4 | Arrival delays, layover delays, and departure delays

6.4.1 | Main results

Using a SDID approach, we examine the effects of slot restrictions on arrival delays, layover delays, and departure delays. We report the estimations results in Table 8. For a comparison purpose, we report both the DID (i.e., columns 1–3) and SDID (i.e., columns 4–6) estimators. Similar to Table 3, and we report the baseline model specifications in column 6, which includes the set of control variables, carrier-route fixed effects, and year-quarter fixed effects. We present the estimation results from arrival, layover, and departure delays in panels A, B, and C, respectively. Panels A and B show the SDID estimates for arrival and layover delays, whereas panel C presents the SDID estimate for departure delays, which is the sum of the estimates from the arrival and layover delays.

The results from panel A show that arrival delays are decreased by 0.956, which is equivalent to a reduction of 1 min, and the estimate is statistically significant. Similarly, the results from panel B show that layover delays are decreased by 0.574 and the estimate is statistically significant. These findings suggest that quantity-based approaches to airport congestion management reduce both the arrival and layover delays at EWR by restricting the frequency of landings and takeoffs. The comprehensive effects are apparent in panel C as departure delays are decreased by 1.529, an approximately 1 min and a half reduction in total delays. The coefficient of the interaction term is also statistically significant.

TABLE 8 The impact of slot restrictions on arrival, layover, and departure delays

	DID			SDID		
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: *Arrival Delay*						
Slot-controlled airport × policy time	−1.052***	−1.202***	−1.005***	−0.651*	−0.963***	−0.956**
	(0.347)	(0.340)	(0.349)	(0.354)	(0.337)	(0.397)
Policy time	−1.410***	−1.533***	−1.670***	−2.180***	−2.715***	−2.807***
	(0.090)	(0.094)	(0.097)	(0.241)	(0.231)	(0.291)
Slot-controlled airport	5.743***			5.444***		
	(0.502)			(0.512)		
Adjusted R^2	0.112	0.487	0.488	0.230	0.635	0.636
Panel B: *Layover Delay*						
Slot-controlled airport × policy time	−0.884***	−0.653**	−0.559**	−0.910***	−0.666**	−0.574*
	(0.341)	(0.274)	(0.279)	(0.348)	(0.270)	(0.304)
Policy time	0.532***	0.414***	0.308***	0.192	−0.114	−0.460**
	(0.076)	(0.075)	(0.076)	(0.247)	(0.197)	(0.194)
Slot-controlled airport	5.812***			6.207***		
	(0.588)			(0.607)		
Adjusted R^2	0.046	0.501	0.502	0.145	0.706	0.707
Panel C: *Departure Delay*						
Slot-controlled airport × policy time	−1.935***	−1.855***	−1.564***	−1.561***	−1.629***	−1.529***
	(0.416)	(0.435)	(0.440)	(0.430)	(0.432)	(0.499)
Policy time	−0.877***	−1.119***	−1.362***	−1.988***	−2.829***	−3.267***
	(0.119)	(0.120)	(0.122)	(0.248)	(0.241)	(0.295)
Slot-controlled airport	11.555***			11.651***		
	(0.547)			(0.565)		
Adjusted R^2	0.153	0.549	0.550	0.415	0.719	0.721
All other control vars			Yes			Yes
Carrier-route fixed effect		Yes	Yes		Yes	Yes
Year-quarter fixed effect	Yes	Yes	Yes	Yes	Yes	Yes
Observations	97,531	97,531	97,531	83,832	83,832	83,832

Notes: The sample period is between 2006:Q1 and 2010:Q4. A SDID approach is used to retrieve these estimates. Standard errors are clustered by the carrier-route combination level and reported in parentheses. The results from the baseline specification are reported in column 6.
*$p < 0.1$, **$p < 0.05$, ***$p < 0.01$.

6.4.2 | Leads and lags

Considering that most of a reduction in departure delays after the policy implementation originates from a decrease in arrival delays, we investigate the pre- and post-policy movements of arrival delays to examine whether the effect was anticipated and whether the effect is lasting. The leads and lags analysis is equivalent to Section 6.3.2 except that SDID estimators (i.e., column 6 of panel A) are used as the baseline specification in this analysis. We report the estimated coefficients in Table 9. To recap, we express all values in terms of t, in which $t + 1$ is the one quarter lead of the treatment effect, whereas $t - 1$ is the one quarter lag of the treatment effect.

The pre-policy movements of arrival delays are volatile as delays are increased or decreased before slot restrictions at EWR. Particularly, the three periods leading up to slot restrictions have coefficients that are negative and statistically significant, positive and statistically significant, or positive but statistically insignificant. On the other hand, the impacts of slot restrictions on arrival delays are consistently decreasing after the policy implementation. The period right after the policy, the decrease in arrival delays is approximately 2.7 min, and the estimate is statistically significant. The constant decreasing in delays is most apparent in eight quarters after the policy implementation (i.e., $t-8$) with a reduction of approximately 7.3 min.

The only exception to the constant decrease in arrival delays is when four periods after the policy was in effect, at which the arrival delay increases to 0.761 min. This period is the second quarter of 2009 when the summer traveling season has begun. Congestion increases as the busy summer traveling season approaches in 2009 and the FAA reinstated slot restrictions to manage the airport capacity at EWR. The reinstatement of slot restrictions at EWR in the second quarter of 2009 shows an immediate decrease in arrival delays of 5.744 min. It shows the effectiveness of temporary quantity-based approaches to manage airport congestion.

6.4.3 | Extensions

There are several endogenous policy responses from airlines that should concern policymakers and the basis of our analysis. Even though the FAA restricts the per-hour total departure and landing slots at EWR, however, airlines are still in charge of their flight networks. This means that airlines can strategically assign and reassign certain routes at EWR to satisfy the restriction requirements set by the FAA. In other words, the pre- and post-policy route compositions may be different if airlines systematically adjust routes after the policy. The endogenous changes in the route composition can bias the estimation results. Consequently, we limit the sample to the subsample of these consistently operating routes at EWR and compare the estimates from the subsample with the estimates from the full sample. We define the consistently operating routes as the carrier-origin–destination combination that includes EWR for the entire duration of our sample period between 2006:Q1 and 2010:Q4.

Table 10 presents the estimation results using only the sample from the consistently operating routes. Column 6 presents the SDID estimate from the baseline specification, which is the comparable finding with the full sample analysis (i.e., column 6 of panel A in Table 8). In the full sample analysis, arrival delays are decreased by approximately

TABLE 9 A leads and lags analysis: Slot restrictions and arrival delays

	Arrival delay (1)		**Arrival delay (2)**
Slot restriction$_{t+8}$	−1.798***	*Slot restriction*$_{t-1}$	−2.661***
	(0.507)		(0.429)
Slot restriction$_{t+7}$	−2.693***	*Slot restriction*$_{t-2}$	−1.342**
	(0.454)		(0.578)
Slot restriction$_{t+6}$	−0.813*	*Slot restriction*$_{t-3}$	−1.211**
	(0.456)		(0.554)
Slot restriction$_{t+5}$	0.712	*Slot restriction*$_{t-4}$	0.761
	(0.457)		(0.588)
Slot restriction$_{t+4}$	0.515	*Slot restriction*$_{t-5}$	−5.744***
	(0.394)		(0.629)
Slot restriction$_{t+3}$	−1.628***	*Slot restriction*$_{t-6}$	−3.748***
	(0.437)		(0.561)
Slot restriction$_{t+2}$	0.868**	*Slot restriction*$_{t-7}$	−5.180***
	(0.420)		(0.571)
Slot restriction$_{t+1}$	0.139	*Slot restriction*$_{t-8}$	−7.320***
	(0.419)		(0.504)
Observations	83,832	Adjusted R^2	0.629

Notes: The sample period is between 2006:Q1 and 2010:Q4. A SDID approach is used to retrieve these estimates. Standard errors are clustered by the carrier-route combination level and reported in parentheses. Time t is equal to the second quarter of 2008. All the estimates are expressed in relation to t. For instance, $t+1$ is the one quarter lead of the treatment effect, whereas $t-1$ is the one quarter lag of the treatment effect. The estimates follow the baseline specification from column 6 in panel A of Table 8.

*$p < 0.1$, **$p < 0.05$, ***$p < 0.01$.

TABLE 10 The impact of slot restrictions on arrival delay (consistently operating routes only)

	DID			SDID		
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Slot-controlled airport × policy time	−1.243**	−1.170**	−1.220**	−0.954*	−0.895*	−1.169**
	(0.484)	(0.498)	(0.487)	(0.490)	(0.503)	(0.533)
Policy time	−1.627***	−1.622***	−2.039***	−2.681***	−2.657***	−2.770***
	(0.108)	(0.108)	(0.123)	(0.283)	(0.287)	(0.395)
Slot-controlled airport	4.473***			4.077***		
	(0.727)			(0.729)		
All other control vars			Yes			Yes
Carrier-route fixed effect		Yes	Yes		Yes	Yes
Year-quarter fixed effect	Yes	Yes	Yes	Yes	Yes	Yes
Observations	48,700	48,700	48,700	48,700	48,700	48,700
Adjusted R^2	0.127	0.512	0.517	0.231	0.613	0.615

Notes: The sample period is between 2006:Q1 and 2010:Q4. Standard errors are clustered by the carrier-route combination level and reported in parentheses. We retrieve the estimation results from the consistently operating routes only. We define the consistently operating routes as the carrier-origin–destination combination that includes EWR for the entire duration of our sample period. The results from the baseline specification are reported in column 6.
*$p < 0.1$, **$p < 0.05$, ***$p < 0.01$.

TABLE 11 The impact of slot restrictions on arrival delay for leisure routes

	DID			SDID		
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Slot-controlled airport × policy time	−0.290	−0.966	−0.919	0.084	−0.556	−0.610
	(1.462)	(0.730)	(0.741)	(1.478)	(0.744)	(0.826)
Policy time	−0.796***	−0.838***	−0.957***	−2.672***	−2.778***	−2.457***
	(0.245)	(0.242)	(0.247)	(0.620)	(0.687)	(0.638)
Slot-controlled airport	4.985***			3.735***		
	(1.200)			(1.220)		
All other control vars			Yes			Yes
Carrier-route fixed effect		Yes	Yes		Yes	Yes
Year-quarter fixed effect	Yes	Yes	Yes	Yes	Yes	Yes
Observations	11,853	11,853	11,853	10,814	10,814	10,814
Adjusted R^2	0.105	0.375	0.377	0.174	0.543	0.544

Notes: The sample period is between 2006:Q1 and 2010:Q4. A SDID approach is used to retrieve these estimates. Standard errors are clustered by the carrier-route combination level and reported in parentheses. The results from the baseline specification are reported in column 6. We limit the sample to carrier-route combinations with either endpoint is the following tourism cities: Las Vegas, Miami, and Phoenix.
*$p < 0.1$, **$p < 0.05$, ***$p < 0.01$.

0.96 min, whereas the reduction in arrival delays is approximately 1.17 min in the subsample analysis. These estimates between the full sample and subsample are not substantially different from each other. Consequently, we find that changes in the route composition are not driven by slot restrictions, in which routes that are reassigned or the carrier-route combination of new-entrants do not differ from the relatively permanent routes in terms of arrival delays. Although it is difficult to rule out an endogenous change in the route composition, our robustness test suggests that it is not a significant distraction to policy implications.

As an additional analysis, we construct our sample to flights that have an endpoint in leisure-traveler dominated routes to examine whether slot restrictions have any impacts on arrival delays of the leisure-dominated routes. Particularly, we consider the routes that include in the following tourism destinations: Las Vegas, Miami or Phoenix. Table 11 reports the estimation results. The analysis is similar to the one in Table 8, in which column 6 shows the SDID estimator from the baseline specification. The evidence suggests that even though slot restrictions reduce arrival delays for leisure-traveler dominated routes; the finding is marginal as the SDID estimator is statistically insignificant. This finding implies that on-time performance of tourist routes are not improved as much as business-traveler-dominated-routes.

7 | CONCLUSIONS

In this article, we examine the effects of slot restrictions at EWR on price dispersion and on-time performance. We estimate the treatment effects through a traditional difference-in-differences approach for price dispersion and through a semi-parametric difference-in-differences for on-time performance to address issues regarding the common trend assumption. Our findings suggest that slot restrictions increase price dispersion and improve on-time performance for the treated group. The source of the increased price dispersion is a higher increase in the average ticket price for high-value passengers than that for low-value passengers. Even though high-value passengers would have to pay more for ticket prices, we find that the share of passengers who pay for the highest price groups has increased after the policy. The sources of the improvement in on-time performance are reductions in arrival delays and layover delays. For a comprehensive empirical strategy, we provide estimation results from various specifications and robustness checks, and an analysis on leads and lags of the policy. We also compare estimates with the median Gini coefficient values found in Gerardi and Shapiro (2009) for ensuring consistency with the existing literature.

There are several policy implications from our findings. First, the overall improvement in on-time performance suggests that quantity-based approaches to manage airport congestion, such as slot restrictions, are effective for reducing airport congestion and congestion-related delays. Passengers who travel on the policy-affected routes can all benefit from the reduction in arrival delays, layover delays, and departure delays. Second, the higher increase in ticket prices for high-value passengers, which is consistent with our theoretical prediction, combined with the fact that more high-value passengers have purchased the tickets after the policy show that high-value passengers are better off with slot restrictions despite the increase in prices. This finding is consistent with the existing literature that examines the effects of price-based approaches to airport congestion management on the welfare of different passenger types, which contends that the airport charge can be higher to ensure passengers with a greater relative time value can be protected from excessive congestion. Our article is the first to find such an association for quantity-based approaches to airport congestion management.

We acknowledge two limitations of our estimation approach despite our efforts to ensure a comprehensive analysis. First, while our sample is observed at a quarterly level, which is similar to many other papers that examine the market dynamics in the US airline industry, a recent study of Chandra and Lederman (2018) have used a more disaggregated level of observations for analyzing the market dynamics in the Canadian airline industry. It would be interesting to see whether a distinct set of data can similarly explain the impact of quantity-based approaches to airport congestion management on price dispersion and on-time performance. Second, the FAA removed slot restrictions at EWR in 2016 and it would be intriguing to examine in a future study how the market dynamics have shifted since the permanent discontinuation of the policy.

ACKNOWLEDGMENTS

We would like to thank Janis K. Pappalardo and two anonymous referees for valuable comments and suggestions. The usual disclaimer applies.

ORCID

Zexuan Liu https://orcid.org/0000-0001-6645-8274

ENDNOTES

[1] In their seminal paper, Gerardi and Shapiro (2009) find a negative association between price dispersion and competition in the airline industry. This finding is in contrast with the finding in Borenstein and Rose (1994) who find a positive association. Moreover, Dai et al. (2014) show that the relationship between competition and price dispersion may be nonmonotonic. The cause of the inverse-U shape stems from market concentration and different degrees of price discrimination. Recently, Chandra and Lederman (2018) revisit the relationship between competition and price dispersion and contend that observations with more detailed characteristics may reconcile the mixed evidence in the literature using observations from Canada's airline industry data.

[2] Morrison and Winston (2007) explore alternative approaches to airport congestion pricing because of these issues.

[3] A recent study by Escobari and Hernandez (2019) separates consumers into two different types and finds that high-type or business travelers are less sensitive to price and have a higher valuation.

[4] Van Dender (2007) finds that the average aeronautical fares are higher and delays are lower at slot-constrained airports. However, the study does not consider ticket price dispersion.

[5] See Berardino (2009) for the detailed history of slots policy in the United States.

[6] Flights that arrive within 15 min of the scheduled arrival times are on time.

[7] Specifics of the rules can be found at the following web page: https://www.gpo.gov/fdsys/pkg/FR-2015-01-08/pdf/2014-30378.pdf.

[8] See https://www.faa.gov/news/updates/?newsId=85309.

[9] The FAA's decision in 2016 to lift the limits on scheduled operations at EWR does not affect our analysis because it is beyond our sample period. However, it is worth considering the implications of such a decision regarding on-time performance and competition in the airline industry as a future research project.

[10] Textbooks like Mas-Colell et al. (1995) and Bolton and Dewatripont (2005) provide more examples. Riley (2001) provides an excellent survey on related works.

[11] The reader can conveniently interpret the types as business travelers and leisure travelers. As we cannot distinguish passenger types in the data, high-value and low-value passengers are more appropriate terms in the context.

[12] The carriers can affect the on-time performance rate by adjusting the flight network, changing the scheduled ground time. Notably, the carriers can provide different levels of quality. For example, a flight can have a different probability of being delayed.

[13] The cost for on-time performance is flight specific. However, the arrangement of flights is related to the number of passengers to serve. Thus, assuming the cost to be per passenger is a simplifying assumption. One may also interpret the cost as the compensation for passengers for failing to provide the on-time performance.

[14] The high-value passengers are indifferent between choosing q_H and q_L. We assume there are no capacity limits so they all choose quality High. If there were capacity limits, some high-value passengers are forced to choose quality Low, which is equivalent to a smaller λ in the model.

[15] For a detailed description of the allocation process, see https://www.faa.gov/about/office_org/headquarters_offices/ato/service_units/systemops/perf_analysis/slot_administration/slot_allocation_process/.

[16] A slot auction was proposed but soon revoked and never implemented after slot restrictions at EWR.

[17] We implicitly assume the capacity is large enough to cover the demand, though not all passengers can choose the desirable quality level. The assumption is reasonable as non-peak hour flights are not restricted.

[18] We assume that the policy changes the supply side by affecting the operation cost while the demand side remains unchanged. In particular, both the total demand and the fraction of high- and low-value consumers are unaffected after the policy. It may violate the assumption. For example, the reduced delay may attract more consumers (specifically more high-value consumers) to take flights after the government policy. Nevertheless, the theoretical framework still applies, and the qualitative results should remain with a slight change of parameters. We leave the focus on the quantitative impact to the empirical studies.

[19] It is important to distinguish the difference between nonstop and direct flights. A nonstop flight from Boston to Los Angeles, for instance, means that the aircraft leaves from Boston and arrives at Los Angeles without stops. A direct flight means that the aircraft could make stops between the origin and destination airports without changing planes.

[20] Even though 2014 is not included in the sample period, this is to show that UA has a dominant presence at EWR. The next two largest airlines are American and Delta Airlines, with 7% and 5% of daily operations, respectively.

[21] Since our measure of price dispersion is bounded between zero and one, we also run a Gini log-odds ratio analysis as in Gerardi and Shapiro (2009), which provides an unbounded statistic. The unbounded statistic analysis is consistent with our main finding and is available upon request.

[22] See Autor (2003) for a similar empirical application. Leads and lags in Autor (2003) are not the same leads and lags in a time-series sense, in which a lead refers to the period after an event and a lag refers to the period before an event. In contrast, leads and lags in Autor (2003) refer to the periods leading up to the policy (i.e., leads) and the periods after the policy (i.e., lags).

[23] Our sample period is later than that of Gerardi and Shapiro (2009) and one reason that our value is slightly higher than their values is that multiple mergers occurred after 2006, which affected the market dynamics of the airline industry. Specifically, airlines' ability to discriminate passengers based on preferences has increased after the mergers, leading to a higher variation in prices.

REFERENCES

Abadie, A. (2005) Semiparametric difference-in-differences estimators. *The Review of Economic Studies*, 72(1), 1–19.

Anderson, S.P. & Renault, R. (2011) Price discrimination. In: *A handbook of transport economics*. Cheltenham, England: Edward Elgar Publishing.

Autor, D.H. (2003) Outsourcing at will: the contribution of unjust dismissal doctrine to the growth of employment outsourcing. *Journal of Labor Economics*, 21(1), 1–42.

Basso, L.J. & Silva, H. (2015) Effects of market power on Pigouvian tolls and permits markets: the case of slot management at congested airports. In: *International Transport Economics Association Conference*, Oslo, Norway.

Basso, L.J. & Zhang, A. (2008) Sequential peak-load pricing: the case of airports and airlines. *Canadian Journal of Economics/Revue canadienne d'économique*, 41(3), 1087–1119.

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

Berardino, F. (2009) New U.S. airport slot policy in flux. *Journal of Transport Economics and Policy*, 43(2), 279–290.

Berglas, E. & Pines, D. (1981) Clubs, local public goods and transportation models: a synthesis. *Journal of Public Economics*, 15(2), 141–162.

Bolton, P. & Dewatripont, M. (2005) *Contract theory*. Cambridge, MA: MIT Press.

Borenstein, S. & Rose, N.L. (1994) Competition and price dispersion in the U.S. airline industry. *Journal of Political Economy*, 102(4), 653–683.

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

Brueckner, J.K. (2005) Internalization of airport congestion: a network analysis. *International Journal of Industrial Organization*, 23(7), 599–614.

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

Chandra, A. & Lederman, M. (2018) Revisiting the relationship between competition and price discrimination. *American Economic Journal: Microeconomics*, 10(2), 190–224.

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

Czerny, A.I. & Zhang, A. (2011) Airport congestion pricing and passenger types. *Transportation Research Part B: Methodological*, 45(3), 595–604.

Czerny, A.I. & Zhang, A. (2014) Airport congestion pricing when airlines price discriminate. *Transportation Research Part B: Methodological*, 65(C), 77–89.

Czerny, A.I. & Zhang, A. (2015) Third-degree price discrimination in the presence of congestion externality. *Canadian Journal of Economics/Revue canadienne d'économique*, 48(4), 1430–1455.

Dai, M., Liu, Q. & Serfes, K. (2014) Is the effect of competition on price dispersion nonmonotonic? Evidence from the U.S. airline industry. *Review of Economics and Statistics*, 96(1), 161–170.

Dana, J.D., Jr. (1999) Equilibrium price dispersion under demand uncertainty: the roles of costly capacity and market structure. *RAND Journal of Economics*, 30(4), 632–660.

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

Daniel, J.I. & Harback, K.T. (2009) Pricing the major U.S. hub airports. *Journal of Urban Economics*, 66(1), 33–56.

Escobari, D. & Hernandez, M.A. (2019) Separating between unobserved consumer types: evidence from airlines. *Economic Inquiry*, 57(2), 1215–1230.

Escobari, D. & Jindapon, P. (2014) Price discrimination through refund contracts in airlines. *International Journal of Industrial Organization*, 34, 1–8.

Gerardi, K.S. & Shapiro, A.H. (2009) Does competition reduce price dispersion? New evidence from the airline industry. *Journal of Political Economy*, 117(1), 1–37.

Laffont, J.-J. & Tirole, J. (1986) Using cost observation to regulate firms. *Journal of Political Economy*, 94(3 Part 1), 614–641.

Lipsman, M.A. (1994) *A theory of transportation clubs with special application to the domestic aviation system*. Technical report, Iowa State University, Department of Economics.

Lott, J.R. & Roberts, R.D. (1991) A guide to the pitfalls of identifying price discrimination. *Economic Inquiry*, 29(1), 14–23.

Mas-Colell, A., Whinston, M.D. & Green, J.R. (1995) *Microeconomic theory*, Vol. 1. New York: Oxford University Press.

Maskin, E. & Riley, J. (1984) Monopoly with incomplete information. *The Rand Journal of Economics*, 15(2), 171–196.

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

Morrison, S.A. & Winston, C. (2007) Another look at airport congestion pricing. *American Economic Review*, 97(5), 1970–1977.

Myerson, R.B. (1981) Optimal auction design. *Mathematics of Operations Research*, 6(1), 58–73.

Pertuiset, T. & Santos, G. (2014) Primary auction of slots at European airports. *Research in Transportation Economics*, 45(C), 66–71.

Reitzes, J.D., McVeigh, B., Powers, N. & Moy, S. (2015) Competitive effects of exchanges or sales of airport landing slots. *Review of Industrial Organization*, 46(2), 95–125.

Riley, J.G. (2001) Silver signals: twenty-five years of screening and signaling. *Journal of Economic Literature*, 39(2), 432–478.

Rosenbaum, P.R. & Rubin, D.B. (1983) The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70(1), 41–55.

Rothschild, M. & Stiglitz, J. (1978) Equilibrium in competitive insurance markets: an essay on the economics of imperfect information. In: *Uncertainty in economics*. Amsterdam, The Netherlands: Elsevier, pp. 257–280.

Rupp, N.G. (2009) Do carriers internalize congestion costs? Empirical evidence on the internalization question. *Journal of Urban Economics*, 65(1), 24–37.

Swaroop, P., Zou, B., Ball, M.O. & Hansen, M. (2012) Do more US airports need slot controls? A welfare based approach to determine slot levels. *Transportation Research Part B: Methodological*, 46(9), 1239–1259.

Van Dender, K. (2007) Determinants of fares and operating revenues at US airports. *Journal of Urban Economics*, 62(2), 317–336.

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

Zhang, A. & Czerny, A.I. (2012) Airports and airlines economics and policy: an interpretive review of recent research. *Economics of Transportation*, 1(1–2), 15–34.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

How to cite this article: Lee, S., Liu, Z. & Yu, H. (2021) Heterogeneous price effects and increased price dispersion from quantity-based congestion management. *Economic Inquiry*, 59:1378–1402. https://doi.org/10.1111/ecin.12991