Another Look at Airport Congestion Pricing

By Steven A. Morrison and Clifford Winston*

Airline travel in the United States has rebounded sufficiently from the September 11 terrorist attacks that travelers and policymakers have begun to worry once again about gridlocked skies. Economists themselves have been unwavering in their concern about congestion, varying only in the way they would use the price mechanism to allocate existing runway capacity efficiently and thus begin to solve the problem. For decades, beginning with Michael E. Levine (1969) and Alan Carlin and R. E. Park (1970), airline researchers have called for airports to replace their existing landing fees, based on an aircraft’s weight, with efficient landing and takeoff tolls based on an aircraft’s contribution to congestion. In recent years, however, some economists have begun raising questions about the extent to which such pricing would reduce delays. Jan K. Brueckner (2002) has pointed out that because an air carrier bears the cost of delay that it imposes on its other flights, it should be charged only for the delay it imposes on other carriers’ flights. For example, a carrier with a 50 percent share of operations at an airport should be charged for one-half of the delay costs it creates—the delay incurred by other carriers—while the carrier’s smaller competitors, which have many fewer operations, should be charged for all the delay they create. Christopher Mayer and Todd Sinai (2003) apply this idea to hub airports where dominant carriers cluster their operations to provide convenient connections for passengers (while nondominant carriers operate most of their flights at less-congested times). They conclude that optimal tolls at hub airports should be small because most of the delay at those airports is internalized.

Although these authors make theoretically valid points, they do not quantify the welfare gains from modifying congestion tolls to account for internalization and the pattern of airline operations at hubs. The size of the gains is important because commercial airlines and especially general aviation have succeeded thus far in blocking any change in airport pricing policy that would cause them to pay higher landing fees. Accordingly, many air transport operators would be more strongly opposed to efficient congestion pricing, as suggested by Brueckner, and Mayer and Sinai, because they would perceive it as blatantly inequitable: operators with a small share of flights at a congested airport would pay higher tolls than those with a large share, even if both operated during the same time period.¹

In this paper, we develop a model of the net benefits to air travelers from flights to and from US airports and calibrate it with data that account for a large share of the nation’s passenger air travel in 2005. Current delays, which reduce net benefits to travelers, are higher than optimal and would be ameliorated by airport congestion tolls. We estimate and compare the welfare gains from the optimal congestion tolls suggested by Brueckner, and Mayer and Sinai, and from congestion tolls more traditionally recommended by economists, to assess whether the incremental benefits from optimal tolls would justify the requisite effort to overcome the greater political opposition they would likely face. We find a small difference between the net benefits generated by the two congestion pricing policies because the bulk of airport delays are not internalized and because the efficiency loss from pricing internalized congestion is small. Congestion charges traditionally supported by economists would significantly reduce delays at congested US airports and have a greater likelihood than optimal tolls have of actually being implemented.

¹ Brueckner (2005) explicitly shows that asymmetric congestion tolls emerge in a network setting, being low for dominant hub carriers and high for fringe airlines. He focuses, however, on efficiency, and not distributional, issues.

* Morrison: Department of Economics, Northeastern University, Boston, MA 02115 (e-mail: S.Morrison@neu.edu); Winston: Economic Studies Program, Brookings Institution, 1775 Massachusetts Avenue, NW, Washington, DC 20036 (e-mail: CWinston@brookings.edu). We are grateful to Jan Brueckner and to the referees for their helpful comments.
I. Estimating the Welfare Effects of Optimal and Atomistic Airport Congestion Tolls

Currently, landing fees are set by local airport authorities based on an aircraft’s weight, subject to guidelines set by the Federal Aviation Administration (FAA). Because congestion at airports varies by time of day with the volume of traffic, weight-based landing fees bear no relationship to congestion because a plane waiting to take off or land is delayed roughly the same amount of time by a jumbo jet as by a small private plane.

Conventional formulations of efficient congestion pricing at airports derived tolls for all aircraft without considering whether a plane operated by a given airline was delaying other planes that it operated or planes operated by other operators. These so-called atomistic tolls effectively treat all airlines as operating (no more than) one flight in each time period. Joseph I. Daniel (1995) empirically explored the extent to which an airline’s internalization of delay costs affected its pattern of operations. Brueckner (2002) and Mayer and Sinai (2003) advanced our conceptual knowledge of optimal pricing by recognizing that economic efficiency calls for carriers to be charged only for the delay they impose on other carriers and not for the delay they impose on themselves.

We estimate the extent that optimal pricing’s benefits exceed the benefits from (nonoptimal) atomistic tolls by developing a model of the net benefits from air travel, composed of airport users’ benefits and the delayed and undelayed flight costs to users and carriers. We then determine the tolls that maximize net benefits when carriers are charged only for the delay that they impose on other carriers and the tolls when carriers are assumed (erroneously) to be atomistic.

We assume that air transportation service is provided at a given airport during \( t \) time periods by \( i \) commercial airline carriers. Let \( P_{it}(Q_{it}) \) be the (inverse) demand for airline \( i \) during period \( t \), where \( Q_{it} \) is the flow of flights by the airline per unit of time. The undelayed flight costs, including the value of passengers’ travel time, are given by \( FC_i \). The cost per unit of delay time is denoted by \( v_i \) and the delay incurred at an airport is given by \( D(Q_t, RW) \), which is a function of the total quantity of flights during period \( t \) and the number of air carrier runways at the airport, \( RW \). Thus, net benefits for all flights at an airport during a given period of time can be expressed as

\[
NB = \sum_t \sum_i \int_0^{Q_{it}} P_{it}(Q_{it}) \, dQ_{it} - Q_{it}[FC_i + v_iD(Q_t, RW)].
\]

Our analysis is carried out for calendar year 2005 for the 74 commercial airports included in the FAA’s Aviation System Performance Metrics (ASPM) database. Among the airports are all those that the major network carriers use as hubs, the 46 busiest based on commercial traffic, and 68 of the busiest 100 commercial airports. Together, the airports in our sample account for 73 percent of commercial carrier operations at all US airports.

The database contains scheduled operations every 15 minutes for 23 specific airlines (22 US network and commuter airlines plus Air Canada) plus one composite “other” category for all other commercial airlines. It also includes a measure of actual delay. We thus incorporate airline operations for 35,040 time periods (96 time periods per day times 365 days). The absence of data on general aviation prevents us from assessing the welfare effects that optimal and atomistic charges would have if they were also applied to these operators. But given that

---

2 Although landing fees vary, a representative landing fee is $2.00 per 1,000 pounds of weight. The maximum design landing weight of a Boeing 757–200 aircraft, with a capacity of approximately 186 passengers, is 198,000 pounds. A fee of $2.00/1,000 pounds results in a fee that is less than $3 per passenger for typical passenger loads.

3 Our simple single product formulation of demand abstracts from more complex multiproduct network formulations that, in our view, would increase realism but not have much effect on our central conclusions.

4 Van Nuys airport is also included in the database, but no airlines operate there.

5 Actual delay is composed of airborne delay and taxi delay. Airborne delay is calculated relative to the flight plan, which is based on free-flow travel. Taxi delay is calculated relative to the best time achieved by each carrier at an airport, evaluated seasonally. Average delay per operation ranged from 0 minutes at Gary, Indiana, to 18.0 minutes at Newark, New Jersey, with a weighted average of 8.8 minutes for all operations in 2005.
general aviation operations are atomistic, our 
results will, if anything, overstate the relative 
benefits of optimal versus atomistic tolls
We assume the following functional forms 
and numerical values to execute our analysis.

A. Delay Function

The generic airport delay function, $D$, is based 
on the arrival and departure delay functions 
estimated by Morrison and Winston (1989):

$$D(Q_i, RW) = \exp(\alpha Q_i / RW),$$ where $\alpha$ is a con-
gestion parameter. The specification of the 
delay function has the desirable properties of 
being homogeneous of degree zero in the (air-
port) use and capacity variables, with marginal 
delay an increasing function of operations and a 
decreasing function of runway capacity. A sep-
arate value of the congestion parameter, $\alpha$, was 
determined for each airport to ensure that the 
average delay we obtain by using the function $D$ 
equals the average delay actually measured at 
these airports by the FAA during 2005. The val-
ues ranged from –2.4 (indicating average delays 
of less than a minute because $e$ is raised to the $\alpha$
power) to 1.6. Although $\alpha$ represents a “techno-
logical” parameter, it is reasonable that it varies 
across airports because of variations in runway 
configurations, meteorological conditions such 
as wind, ceiling, and visibility, and traffic other 
than commercial operations that contributes to 
congestion.

B. Demand Function

The demand for flights by airline $i$ during 
period $t$ takes a simple constant elasticity form

$$Q_{it} = A_i P_{it}^e,$$ where $A_i$ is a scaling parameter, $e$ is 
the passengers’ price elasticity of demand, and 
$P_{it}$ is the full price of the flight, including unde-
layed flight costs, delay costs, and landing fees 
borne by the airline and its passengers.

We can derive a more detailed specification 
of airline $i$’s demand during period $t$ by 
specifying an airline’s profits as its revenues 
minus undelayed flight costs ($FC_i$), landing 
fees ($LF_i$), and delay costs ($v_i D(Q_i, RW)$); 

$$\Pi_i = P_{it} Q_{it} - Q_{it} (FC_i + LF_i + v_i D(Q_i, RW)).$$

We assume price-taking behavior because the 
bulk of the costs of congestion arise on routes 
with significant traffic and robust competition, 
especially from low-cost carriers (Morrison, 
Winston, and Vikram Maheshri 2006), and 
because our focus is on alternative airport pric-
ing regimes. Maximizing profit with respect to 
$Q_{it}$ yields $P_{it} (Q_{it}) = FC_i + LF_i + v_i D(Q_i, RW) + 
v_i Q_{it} \partial D / \partial Q_{it}$.

Using our delay function and substituting the 
profit-maximizing expression for $P_{it}$ into the 
constant elasticity demand equation, we obtain:

$$(2) \quad Q_{it} = A_i [FC_i + LF_i + v_i \exp(\alpha Q_i / RW) (1 + \alpha Q_i / RW)]^e. $$

The values of $A_i$ are calculated so that the 
demand functions based on this expression yield 
the observed number of operations for each car-
rier in each time period. The functional form 
and elasticity of each airline’s demand func-
tion are assumed to be the same, but it should 
be noted that because a carrier internalizing its 
own delay experiences higher incremental costs 
than an otherwise equivalent carrier, it will have 
a greater willingness to pay for a given number 
of operations. We assume the base case value 
of passengers’ demand elasticity, $e$, is –1.5, 
with plausible bounds of –2.3 and –0.7 used for 
sensitivity. Finally, although a carrier’s flights 
transport passengers who are connecting and 
traveling nonstop, all flights between a pair of 
airports, which is our basic unit of observation, 
are nonstop.

C. Undelayed Flight Costs

Undelayed flight costs, $FC_i$, are composed 
of aircraft operating costs and passengers’ 
time costs. Pilot costs are included in data on 
aircraft operating costs, but flight attendant 
costs are not included. Aircraft operating costs 
depend on the type of aircraft, the duration of 
the flight, and the carrier operating the flight. 
At each airport, we assume each airline’s flights 
have the average number of seats, passengers,

$^6$The number of runways at an airport used in the delay 
function is from the FAA’s National Flight Data Center.

$^7$Although the integral of demand is undefined for inelas-
tic constant elasticity demand curves, our interest is in the 
difference in net benefits under different pricing regimes, 
which is defined.
and flight distance and duration that it actually operated to and from that airport in 2005. These variables were obtained from data contained in the US Department of Transportation’s Data Banks 28DS, Domestic Segment data and 28-IS, International Segment data.\(^8\)

Although we have data on aircraft operating costs by carrier and aircraft type, it would be intractable to model each flight by each carrier to conduct a welfare analysis. Thus, we chose to estimate each carrier’s average aircraft operating cost using data on aircraft operating cost per block hour for major and national carriers that are contained in the US Department of Transportation’s Form 41 (from Data Base Products).\(^9\) Specifically, a fixed-airline effects model was estimated with (the logarithm) of aircraft operating costs per block hour for each aircraft type operated by each airline regressed on (the logarithm of) each airline’s average number of seats for the corresponding aircraft type (and carrier-specific constant terms).

With 117 observations (and an R\(^2\) of 0.746), the coefficient of (log) seats [0.6493] has the correct positive sign and is statistically significantly less than one (standard error = 0.1453), reflecting economies of aircraft size (that is, aircraft operating costs rise less than proportionately with the number of seats). The estimated equation is used to predict each carrier’s average aircraft operating costs.

As noted, flight attendant costs are not included in our data on aircraft operating costs. Thus, we calculated these costs per block hour in 2005 for each of the carriers in our sample using the US Department of Transportation’s Form 41 Data Base (from Data Base Products).\(^9\)

Passengers’ time costs depend on their value of travel time, which varies by trip purpose, and on the number of passengers on each flight. For each of the airports in our sample, we used the US Department of Transportation’s National Household Travel Survey to calculate the percentage of trips whose purpose was business, the percentage whose purpose was pleasure, and the average household income for business and pleasure travelers.\(^11\) Consistent with the US Department of Transportation’s (1997) guidelines, we value business travelers’ travel time at 100 percent of the wage and pleasure travelers’ travel time at 70 percent of the wage. Passengers’ values of time in 2005 dollars ranged from $27.13 to $79.16 per hour with a median of $47.97.

The undelayed cost per hour of a given flight is obtained by summing its average aircraft operating costs per block hour, flight attendant costs per block hour, and aggregate passenger time costs per hour. For each carrier’s flights, we then multiplied its undelayed cost per hour by the average duration of its flights (in hours) to obtain carrier-specific undelayed flight costs.\(^12\)

D. Delayed Flight Costs

Delayed flight costs and undelayed flight costs for a given unit of time are identical. Thus, to obtain the cost per unit of delay, \(v_i\), we simply expressed undelayed flight costs for each carrier, as calculated above, on a per-minute basis. We then multiplied \(v_i\) by a carrier’s delay, given by the delay function, \(D\), to obtain carrier and time-specific delayed flight costs.

In sum, the net benefits for all flights at an airport can now be expressed as

\[
NB = \sum_t \sum_i \int_0^{Q_t} (Q_{it}/A_{it})^{-1/1.5} dQ_{it} - Q_t[FC_i + v_i \exp(\alpha Q_i/RW)],
\]

\(^8\)The values of the composite “other” category were calculated for all other carriers in the dataset other than the 23 specific carriers in our sample.

\(^9\)Block hours is the standard measure of the duration of an aircraft operation. A block hour begins when the aircraft pushes back from the gate (wheel blocks out) and ends when the aircraft parks at the gate (wheel blocks in).

\(^10\)If data for a carrier in our sample were not included in the Form 41 data, weighted average data for all carriers were used.

\(^11\) The calculations were based on the Metropolitan Statistical Area (MSA) where the airport was located. If an airport was not located in an MSA or if the MSA was not identified due to its small size, data were calculated for the state in which the airport was located.

\(^12\) Because the average duration of flights from the segment databases includes delay, we have to subtract delay to obtain undelayed flight duration. We used our delay function to estimate the flight delays for each time period at each airport. Because delay varies from period to period, undelayed flight costs will also vary somewhat from period to period.
where $\alpha$ and $RW$ vary by airport, $FC$ and $v$ vary by airport and carrier, and $Q$ and $A$ vary by airport, carrier, and time period.

II. Findings

For each airport and time period, we first determined the net benefits from current (weight-based) landing fees by plugging into equation (3) the actual $Qs$ and other parameter values as derived above. We then determined the number of flights by each airline in each time period that maximizes net benefits as given in equation (3). The optimal airport tolls are a by-product of net-benefit maximization and charge each airline only for the marginal delay that it causes other airlines.

The difference in net benefits from current and optimal pricing for the ten airports with the largest net benefits and the ten airports with the smallest net benefits is shown in the first column of numbers in Table 1. (Results for all the airports in our sample are available in a Web Appendix at http://www.e-aer.org/data/dec07/20050873_app.pdf.) The gains from optimal pricing vary by airport and depend on total operations, the average level of congestion, and the delay costs that are not internalized. The largest gains occur at the busiest airports, such as Atlanta, Chicago O’Hare, and New York JFK. Although variation in the extent of internalization exists, its weighted average of 33.1 percent indicates that a substantial amount of aircraft operations do not internalize the delay they cause and that highly congested airports would benefit from efficient pricing. Overall, efficient pricing of takeoffs and landings for commercial air carriers at the 74 airports would yield annual net benefits of $2.7 billion.\footnote{15}

Before Brueckner’s and Mayer and Sinai’s analyses, economists recommended atomistic tolls that did not distinguish between delays an airline caused itself and delays it caused other airlines. Because such tolls are not optimal, their net benefits cannot be determined by maximizing the net benefit function given in equation (3). Instead, we must find the pattern of flights by airlines that would result from an atomistic toll regime and then use the net benefit function to calculate net benefits for that pattern of flights.

We first set takeoff and landing fees to maximize net benefits under the (incorrect) assumption that airlines are atomistic competitors, charging all airlines during period $t$ $Q_v\partial D/\partial Q_v$,\footnote{13} where $v$ is the weighted average value of a minute of delay, evaluated at the $Q_v (=\Sigma Q_v)$ that would maximize net benefits if carriers were atomistic. Carriers are therefore overcharged for congestion because they are charged for delay that they have internalized. The equilibrium pattern of flights for all airlines is then determined by simultaneously solving their demand functions given in equation (2) for each time period at each airport, accounting for how each flight delays other flights through its effect on total operations in the delay function. We then use equation (3) to calculate the net benefits that correspond to the resulting pattern of flights for all carriers.

The second column of Table 1 shows the change in net benefits of charging atomistic tolls, relative to the current practice of weight-based landing fees, for the ten airports with the largest net benefits and the ten airports with

\footnote{14}The number of aircraft operations at Chicago O’Hare, New York JFK and LaGuardia, and Washington Reagan National are constrained by slot controls. If the number of slots is optimal or more than optimal, our methodology overstates the gains from pricing. In all likelihood, however, the number of slots—from the perspective of efficiently reducing congestion—is less than optimal, in which case the direction of the bias is unclear.

\footnote{15}The benefits from optimal pricing vary to some extent with the (absolute value of) demand elasticity—the more elastic demand, the greater the gains from optimal pricing. Assuming a demand elasticity of –0.7 yields $1.9 billion of benefits, and assuming a demand elasticity of –2.3 yields $3.3 billion of benefits.
the smallest net benefits. (The Web Appendix presents results for all airports.) An airport’s benefits from charging atomistic tolls depend on the difference between the noninternalized delay costs that are priced appropriately and the internalized delay costs that are priced inappropriately. As noted, the extent of internalized delay is modest at most congested airports; hence, atomistic pricing generates sizable gains at airports such as Atlanta, Chicago O’Hare, and New York JFK. The total annual welfare gain is $2.5 billion, which is only $0.2 billion less than the gain from optimal pricing.

As noted, the welfare improvement from each pricing regime depends critically on the extent of internal and external delay, and on the fact that tolling external delay increases welfare and tolling internal delay reduces welfare. We can therefore gain insight into the small difference between the welfare gains from the two congestion pricing regimes by quantifying the effects that internal and external delays have on the change in net benefits from pricing. In contrast to optimal tolls, atomistic tolls charge for both types of delay; thus, a regression of the change in net benefits from atomistic tolls at each of our 74 sample airports on the minutes of external delay and the minutes of internal delay can tell us the extent that mispricing internal delay offsets the benefits from pricing external delay. Including a constant term, we find that pricing each minute of external delay increases the change in net benefits $38.32, and pricing each minute of internal delay reduces the change in net benefits $10.01.

The findings are illustrated in Figure 1, which shows an average variable cost curve of airport operations (AVC) and its associated marginal cost curve (MC) and two demand curves, each
with the same constant elasticity. The upper demand curve, $D_i$, reflects demand during a period in which there is only one carrier operating and thus all delay is internalized. The carrier would operate at a point where its demand curve crosses the marginal cost curve, resulting in $Q_0$ operations. During another period, operations are conducted by many carriers whose operations are atomistic. These carriers would operate where the (lower) demand curve, $D_a$, intersects the average variable cost curve, also resulting in $Q_0$ operations. Note, by design, that the single internalizing carrier and the atomistic carriers operate the same number of operations per time period; thus, the amount of external and internal delay is equal. The optimal price for the external delay is given by $t_a$, which is the external congestion cost at the optimal quantity of operations where the atomistic demand curve intersects the marginal cost curve at $Q_a$. Charging this toll to the atomistic carriers would generate the (large) welfare gain, shaded with horizontal lines. If this toll ($t_i = t_a$) were also charged to the carrier that internalizes delay, equilibrium would occur at $Q_i$ and would generate the (small) welfare loss, shaded with vertical lines.

We provide additional perspective on our findings by calculating for a representative airport the approximate ratio of the change in net benefits from atomistic pricing to the change in net benefits from (optimally) pricing only external delay. As noted, in our sample internalized delay accounted for 33.1 percent of total delay, while a regression told us the effect of pricing each minute of internalized delay and the effect of pricing each minute of external delay. Let $I$ be the amount of internal delay at the airport. Given that internal delay is 33.1 percent of total delay on average, simple algebra shows that external delay equals $2.02I$. Thus, the change in net benefits from pricing external delay is $2.02I \times 38.32$, and the change in net benefits from pricing both external delay and internal delay is $2.02I \times 38.32 - 10.01I$. The ratio of the change in net benefits, 0.87, is quite close to the ratio of the change in net benefits from atomistic pricing and from optimal pricing that we reported previously ($2.5$ billion/$2.7$ billion = 0.91).

In sum, the small decline in aggregate net benefits from charging atomistic tolls instead of charging net benefit maximizing tolls occurs because, on average, airports experience roughly twice as much external delay as internal delay, and because the relationship between average and marginal costs results in a penalty from pricing internal delay that is a small fraction of the benefits from pricing external delay.

### III. Final Comments

Because we have not accounted for the congestion caused by general aviation, the welfare gains from optimal and atomistic pricing are even greater than what we have found, while

---

**Figure 1. Effect of Pricing Internal and External Delay**
the percentage difference between them is smaller. Morrison and Winston (1989) modeled air carriers as atomistic and estimated that the net benefits from adopting optimal congestion pricing—for all airline carriers, commuters, and general aviation—was $5.7 billion (2005 dollars). In comparison, if the pattern of operations at the 74 sample airports were generated by truly atomistic carriers, optimal pricing would produce annual benefits of $4.0 billion—leaving a sizable gain from pricing general aviation that behaves atomistically.

As air transportation continues to be liberalized throughout the world, it will be useful for future research to explore the extent to which our findings apply to airports in countries such as Japan, which are generally served by a few carriers, and airports in countries such as those in Western Europe, which are served by several carriers.

The GDP implicit price deflator is used to express the gain in 2005 dollars.

In this simulation, we recalculated the values of $A_t$ in the demand functions to account for the assumption that airlines were atomistic and not internalizing delays. An additional issue raised by optimal airport pricing is that it would create a revenue shortfall for airports. Because carriers are not charged for the delay they impose on themselves, we find that optimal pricing raises $4.5 billion less revenue than atomistic pricing raises. In theory, optimal pricing could be modified à la Ramsey pricing to raise the revenue efficiently, but such a mechanism would create additional political opposition because, ceteris paribus, the operators with the least elastic demands would pay the highest tolls.

REFERENCES


