Preliminary version – Please do not quote

AIRPORT CONGESTION IN EU:
MORE INVESTMENT OR BETTER PRICING?

Ofelia Betancor
Ginés de Rus
Gustavo Nombela (*)

Department of Applied Economic Analysis, and
Economics of Infrastructure and Transport Research Group (EIT)
University of Las Palmas de Gran Canaria (Spain)

September 2003

Abstract

Airport congestion has been generally dealt in the literature in a similar fashion as road congestion. However, the phenomenon is quite different, because entry at airports is not random. Flight delays are a consequence of system overload, which is linked to profit maximization decisions of airports and airlines. Moreover, airport congestion exhibits a cascade-type of effect not present in roads: one single delay may generate congestion accumulating over the next hours. Therefore, congestion costs depend on the time of the day which is studied. This paper analyzes airport congestion from this perspective, arguing that a combination of peak-load pricing and investments in new capacity is probably not the best solution to airport congestion. A theoretical model shows that a pricing system based on congestion fees both for airlines and airports could do a better job to mitigate EU congestion costs. Data from flight delays at Madrid airport are used to illustrate our results.

Keywords: airport congestion, infrastructure, pricing
JEL codes: R41, L93

---

* Corresponding author: *gnombela@daea.ulpgc.es*

Acknowledgements: This paper draws on some results from a case study prepared for the EU Commission (research project UNITE, Unification of Accounts and Marginal Costs for Transport Efficiency). We want to thank our colleagues at UNITE, specially C. A. Nash and P. Tena for their comments and suggestions, and also M.A.Oleaga from Madrid airport management. Remaining errors are authors' sole responsibility, as all opinions expressed in this paper.

1 Introduction

Transport infrastructure congestion was first studied by economists for the case of roads. The seminal concepts can be traced back to Pigou, although the modern literature developed after the contributions of Walters (1961) and Vickrey (1963). Road congestion is caused by multiple individual decisions on entry to a limited-capacity infrastructure. The solution is theoretically simple: users should pay for costs imposed on other users sharing the road, thus eliminating the negative external effects. Congestion tolls can solve the problem (or at least mitigate it) by excluding those users with a lower valuation for their trips. Another frequently quoted result from this literature is that, under perfect divisibility and constant returns to scale, optimal pricing policy would be compatible with infrastructure cost recovery (Mohring and Harwitz, 1962; Strotz, 1965).

Airport congestion is nowadays a serious problem at some world regions, specially US and Europe[1], even more costly than road congestion. Reynolds-Feighan and Button (1999) provide a detailed description of the European airport industry at the end of the 1990's. Most European airports report high percentages of delayed flights, and *average* delays over 20 minutes. Severely disrupted airports at some periods (e.g. holiday seasons) are the more visible part of this problem, but not the most important one. Even short delays are causing enormous economic losses, both for providers of services and travellers.

There are several features of airport congestion quite different from road congestion. Airports' basic infrastructure (runways, aeronautical services) are shared among a relatively small number of agents. Decisions of entry are not random, but scheduled. Agents taking entry decisions (airlines) generate and suffer external costs, as in the case of roads, but at airport there is a large group of people (travellers) who do not cause congestion and only bear costs, being only compensated in special cases.

The literature dealing with the problem of airport congestion is relatively extensive (Levine, 1969; Carlin and Park, 1970; Morrison, 1983; Fisher, 1989; Morrison and Winston, 1989; Oum and Zhang, 1990; Daniel, 1995, 2001; Wolf, 1998; Daniel and Pawha, 2000; Hansen, 2002; Brueckner, 2002a, 2002b). However, most of these works

[1] Some estimates evaluate total air congestion costs around 4.5 billion US$ for the US (1998) and 5.5 billion ECU for Europe (1995) (ATAG, 2000). All estimates of congestion costs are rough approximations, due to the required assumptions on the value of time for passengers and the valuation of indirect impacts on airlines. However, the figures indicate the magnitude of the problem of flight delays.

draw heavily from the analysis of road congestion, and end up proposing similar solutions as for roads: (1) enlarge capacity; (2) manage demand by peak-load pricing.

Capacity enlargements are not necessarily a solution to airport congestion, in a context of rapidly growing demand for air services. We illustrate this point here by using real data from Madrid airport. Peak-load pricing is only a feasible solution if airlines' demand for slots were responsive to prices, and if enough travellers were willing to change their preferred departure/arrival times for other services at non-peak hours. We believe both these assumptions do not currently hold in Europe.

In this paper, we develop a theoretical model to show the distinguishing features of airport congestion, by analyzing the decisions of airports' managers and airlines. Externalities caused among the three groups of participating agents (airports, airlines, travellers) can be easily examined in the context of the model. The next step is to propose a system of fees capable of fully internalizing all those externalities. An approximation to this system could be implementable in practice, and could greatly contribute to alleviate airport congestion problems.

The structure of the paper is as follows: section 2 discusses the concepts of capacity, scarcity and congestion, in the context of airports, and the relationship among them. Section 3 presents the model, and its main results. Section 4 tests the validity of this model to reflect the problem of airport congestion using flight delay data from Madrid airport for the period 1997-2000. Finally, section 5 presents the main conclusions of the work.

2 Capacity, scarcity and airport congestion

Airport capacity is defined as the ability of a component in the airport system to handle aircrafts, and it is usually expressed in terms of plane movements per hour. It is therefore, the maximum number of operations than can be accommodated within an hour, taking into account the prevailing conditions of visibility, air traffic control, aircraft mix and nature of operations (Reynolds-Feighan and Button, 1999). Among all components of airport capacity, runways are usually the main constraint, because they are the key element determining the number of take-offs and landings per hour. Other relevant elements of airport capacity are terminal, apron, gates and air traffic control (ATC) capacities.

A distinction between congestion and scarcity is crucial in the economic analysis of airports' activity. Airport infrastructure is given in the short-run and, in contrast with road infrastructure, only qualified users are allowed to enter the system. The allocation of scarce infrastructure to demand is done through an ex-ante procedure of slot assignment. Following Doll et al (2000), congestion refers to the costs arising from crowding effects (too many users in the system), whilst scarcity is a situation of exclusion of some firms from the system due to lack of capacity. Thus, congestion is related to system overloads once available capacity has been allocated ex ante, while scarcity is related to unattended demand due to capacity constraints: slot demand exceeds slot supply.

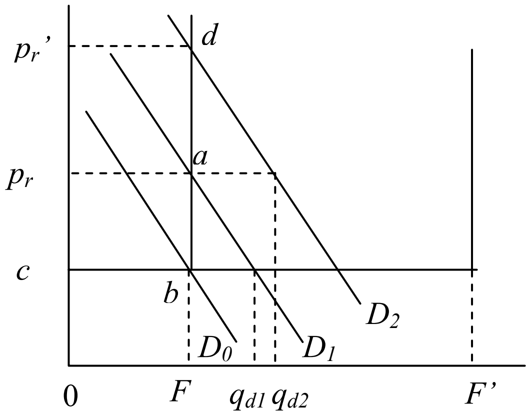
The concept of scarcity is a key difference between air and road transport. It applies to transport systems where a number of firms share some capacity-constrained infrastructure (the concept is also directly transferable to railways or ports). However, scarcity is not present in road transport, where users (motorists, trucks, buses) directly enter the system with their vehicles without any external coordination. It is the impossibility of limiting entry to roads which generates the problem of congestion in the context of roads.

On the contrary, in the case of airports, entry is limited to a reduced number of carriers by airports' managers, so as to be able to control for adequate aeronautical safe practices. At large airports, the usual equilibrium is that more operators would like to enter the system but must be excluded, because of lack of capacity. In this situation, scarcity costs arise when a particular slot is assigned to an incumbent but could have had a higher value if used by another carrier.

Airport congestion is then completely independent from the concept of scarcity. It would be possible, for example, that an airport with no excess of demand could suffer from congestion problems (our data on Madrid airport, to be presented below, illustrate this point). Congestion is obviously directly linked to how intensively an airports' capacity is used, but it is not ultimately caused by capacity limits. It could be perfectly possible that flights using an airport operating at full capacity would not experience delays, while another airport with enough available space would force them out of their schedules.

Figure 1, adapted from Starkie (1988), helps to understand better the concepts of congestion and scarcity. In the figure, the marginal cost c of an airport, which has a maximum number of flights per hour *F* for its normal operation, is depicted together with the demand *D(p)* for slots from airlines.

Figure 1. Airport Congestion and Scarcity



When demand is D_0, at the declared runway capacity (*F*), the market equilibrium price is equal to *c*. This price level both keeps demand equal to available capacity and it covers total costs of supply, including a reasonable rate of return on the capital invested (constant average costs are assumed).

Consider that infrastructure capacity is fixed at *F* but assume that demand grows up to D_1. Two options are available for the airport manager: one to increase the price up to p_r; the other consist of allowing an excess of demand (q_{d1}-*F*) to exist, with the associated extra costs to passengers and airlines caused by delays.

Airports do not usually apply price-rationing policies to allocate slots. On the contrary, because a large part of their costs are covered through revenues from non-aeronautical activities (leasing of commercial space at terminals), they even set slot prices below the average cost of supply. In figure 1, with demand D_1 and a price lower than p_r there exists an excess of demand from air travellers. Profit maximising airlines may then increase their fares to accommodate that excess of demand and earn an extra rent, equal to the area $abcp_r$, if the airport price per slot is equal to *c*.[2]

Assume now that price-rationing would be applied by the airport. Even though expected demand D_1 can be accommodated to existing capacity through the market clearing price p_r,

[2] There is some evidence about the size of this economic rent. The RUTCASE report suggested that the difference between p_r and *c* for Heathrow meant a fare premium charged by the airlines of 20 sterling pounds.

actual demand (D_2) might exceed capacity in a particular t hour, due to some flights failing to take-off or landing as originally scheduled in the previous hour, t-1. No scarcity would then be present at the airport (because the excess of demand is eliminated by rising the slot price), but congestion could appear.

An obvious solution to both the problems of scarcity and congestion is to enlarge capacity up to *F'*. However, airport investments require long periods and are extremely costly, so it is not always feasible to invest in new airports or runways to eliminate these problem. This is specially true if the airport is subject to salient peak-periods. In order for the airport to work smoothly during all hours, a large investment would be required, which might render un-economical when considering the low use during non-peak hours.

To avoid inefficiencies in an airport with short-term fixed capacity *F,* as depicted in figure 1, a solution is to apply the right pricing policies. First, excess of demand should be rationed by charging adequate average slot prices p_r. By doing this, the use of infrastructure would be allocated to those airlines with a higher value for it, and scarcity costs would be avoided. Second, internalisation of the external effects that create congestion costs should be pursued. The shift from D_1 to D_2 can be originated by some random factor (weather, incidents), but more generally it will be caused by some variables which can be controlled either by airlines or by airports to improve flights punctuality. The theoretical model presented in the next section shows how this internalization can be achieved.

3 A theoretical model of airport congestion

Consider the case of a single airport with limited capacity *F* (maximum number of flights per some given period of time). For simplicity, it is assumed that only two airlines operate at this airport, and that they have constant shares over the flights operated at all times through the day: airline 1 has a share α, and airline 2 the complementary share 1-α, with 0#α#1.

Demand for airport slots is exogenously given at every period of time within the day, and it is induced by the background demand from passengers. At some periods demand is higher because of passengers' preferences for departures (we may think of this in practice as peaks experienced by most airports during the early hours in the day, e.g. from 7am-10am, and the late afternoon hours 4pm-8pm), while in other periods the demand is weaker. The number of required flights to attend passengers' demand during period i is f_i. But it might easily occur that at some periods the situation f_i > F occurs. The actual number of flights operating at the airport during period i will be denoted by n_i, which will be subject to a capacity constraint, $n_i \# F$, and a demand constraint, $n_i \# f_i$ (only the stricter constraint at each period will be relevant).

The decision on the actual number of allocated slots n_i is taken by the airport, which is a profit-maximizing agent, subject to regulation over its price. The slot price is set at P_A by some regulatory authority. Airport's marginal cost per flight is assumed to be constant and equal to c_A. The actual number of flights n_i is then split between the two airlines according to their respective market shares: $\alpha\, n_i$ flights of airline 1, and $(1-\alpha)\, n_i$ of airline 2. Airlines marginal cost per flight is c_C for both companies, and they obtain a constant revenue per flight P_C. The implicit assumption is that the price per seat is common to both firms and set at some oligopoly equilibrium-level, so that $P_C > c_C$.

Finally, all passengers are considered to have the same reservation utility level $\overline{U}$, expressed in monetary terms, which is high enough so that all of them decide to fly. If their flights are delayed, passengers receive a disutility measured by $v\, t$ (with v representing the value of time, and t the delay length). The number of passengers per flight is D. A passenger's net utility is then $U(t) = \overline{U} - (P_C/D) - vt$.

As it can be observed, the set up considered for this model is intentionally simple. We do not pretend to be representing competition between airlines, or passengers' decisions to fly or not. This is the reason why we introduce the rigidities of exogenous demands both for air travel (from passengers) and for slots (from airlines). We choose to do so in order to focus more clearly on the externality problem of congestion.

3.1 Delays and congestion costs

Empirical studies on airport congestion costs have identified several reasons which generate flight delays:

- airline problems (mechanical failures of aircraft, limits on crews' working hours, baggage checking, etc.)

- airport problems (including both ATC and ground activities, such as aircraft, passenger or baggage handling)

- reactionary (secondary delays experienced after some primary delay occurs, e.g. impact on other flights which must be moved out of their schedule, late arrival of the aircraft to its next leg on a route)

- passengers and cargo

- weather and other major disruptions (e.g. strikes)

Among all these reasons, a great percentage of delay time experienced by flights and passengers can be attributed to the first two groups: problems caused by ATC and airports' assets, and by the airlines. Some large airports (e.g. Madrid) are currently developing information systems capable of determining for each delayed flight the exact amount of time attributable to each reason. Using this information, it is feasible to split delays according to causes.

In order to represent in our model this complex formation of flight delays, we will consider that all flights operated by airline 1 within a period *i* will experience some delay t_1 which results from adding three components:

$$t_1 = t_A(n) + t_C(e_1) + \beta \bar{t} \quad (1)$$

where $t_A(n)$ is delay time attributable to causes originated by the airport, $t_C(n)$ is delay time induced by the own company, and $\beta \bar{t}$ represents the interaction between a particular flight and the rest of operations performed at the airport during that period.

The function $t_A(n)$ depends on the actual number of operations n, but, most importantly, on how close is n to the maximum airport capacity F. If the number of operations is very small the airport will work smoothly, and then $t_A(n) = 0$. But after some critical threshold, we will have $t_A'(n) > 0$, and possibly also with $t_A''(n) > 0$. The critical threshold of operations will depend on airports' efficiency.

The function $t_C(e_1)$ represents delays originated by the own airline on its flights. It is assumed that the company can exert some level of effort e_1 to reduce these delays, so that $t_C'(e_1) < 0$, and returns to that effort will be decreasing, so $t_C''(e_1) > 0$. For example, an airline can invest in frequent fleet renovation, so that mechanical failures are minimized, and may have some extra crews and planes reserved to be able to respond to unexpected events. The effort in punctuality e_1 imposes a monetary cost on the airline, given by $g(e_1)$, with g'> 0, g''> 0. Finally the amount of delay time experienced by a flight generates an extra cost for airline 1, represented by $G(t_1) = \lambda\, t_1$, where λ is the cost per minute of delay, assumed to be constant.[3]

The term $\beta \bar{t}$ included in equation (1) is intended to reflect the cascade-type of effects that originate at an airport when a flight is delayed. In order to minimize the impact on passengers in that flight, when an operation cannot be performed on time and the airport does not have room to accommodate the plane within the schedule, some other flights are forced out of their scheduled times, or even they are transferred to the next period. The conclusion is that the punctuality of a particular flight is not independent from how the rest of flights within the period are adjusted to their scheduled times. In this model, we choose to incorporate to the delay time of an individual flight some minutes which depend on the average delay time of all flights in the same period. The parameter β represents the importance of this effect. Small values of β indicate that little interaction occurs, while the opposite case applies if β is large, within the limited range 0 # β < 1.

3.2 Equilibrium for one-single period

Assume initially that we are interested in analyzing the equilibrium reached in this model when there is a single period of reference for airport operations. We will extend below the analysis to a two-period scenario, but most of the results can be obtained in the simpler one-period case.

In order to solve for the endogenous variables of this model (n, e_1, e_2), we will analyze separately the maximization problems of the airport and the two airlines. We will compare then the equilibrium with the socially optimal values (n^*, e_1^*, e_2^*).

The airport manager chooses the number of slots to sell in order to maximize its profits:

---

[3] It would probably be more realistic to assume that congestion costs for airlines are not linear, i.e. G''>0, because longer delays can be more costly per minute than shorter ones. We choose the linear form G(t)=λt because it yields more easily comparable outcomes, and the same results hold with the assumption G''>0.

$$\begin{aligned} &\underset{n}{Max}\ \ (P_A - c_A)\, n \\ &s.a.\quad n \le f \ \ (demand\ constraint) \\ &\qquad\ \ n \le F \ \ (capacity\ constraint) \end{aligned} \tag{2}$$

The obvious solution to the optimization problem (2) is to set $n = F$, if $f \geq F$, and $n = f$ otherwise. Given the assumption of a constant margin per operated flight, the airport maximizes profits by selling as many slots as possible.

Airline 1 sets its effort on punctuality e_1 by solving:

$$\begin{aligned} &\underset{e_1}{Max}\ \ \left[P_C - c_C - P_A - \lambda t_1\right] \alpha n - g(e_1) \\ &s.a.\quad t_1 = t_A(n) + t_C(e_1) + \beta \bar{t} \\ &\qquad\ \ t_2 = t_A(n) + t_C(e_2) + \beta \bar{t} \\ &\qquad\ \ \bar{t} = \alpha t_1 + (1-\alpha) t_2 \end{aligned} \tag{3}$$

The first-order condition of this problem implicitly defines the punctuality effort chosen by the company:

$$-(\alpha n)\, \lambda \left[\left(1 + \frac{\alpha\beta}{1-\beta} \right) t'_C(e_1) \right] = g'(e_1) \tag{4}$$

The interpretation of expression (4) is immediate: the LHS represent the marginal benefit obtained by the airline by exerting additional effort e_1. The term in brackets is the reduction of delay time per flight achieved ($dt_1 / de_1 < 0$), which is transformed into monetary terms by the coefficient λ, and multiplied by the total number of flights that the company operates. On the other hand, the RHS is the marginal cost of effort e_1, and the optimal effort is exerted when both sides are equal.

Airline 2 solves a similar problem to (3) to choose its effort level e_2. The solution is also defined implicitly by a first-order condition:

$$-(1-\alpha) n\, \lambda \left[\left(1 + \frac{(1-\alpha)\beta}{1-\beta} \right) t'_C(e_2) \right] = g'(e_2) \tag{5}$$

The solution obtained from (5) will not be typically the same as for the other airline, so in general we will have different values for e_1 and e_2. Only in the case $\alpha = 0.5$, it is obtained that $e_1 = e_2$.

Two interesting results can be extracted from these implicit solutions. First, the effect of the market share α is positive over the effort on punctuality e_1 (correspondingly of 1-α over e_2). This can be seen by applying the implicit function theorem to (4):

$$\frac{de_1}{d\alpha} = -\frac{n\,\lambda\,t'_C(e_1)\left(\dfrac{1+\alpha+\alpha\beta}{1-\beta}\right)}{g''(e_1)+\alpha\,n\,\lambda\left(\dfrac{1+\alpha\beta}{1-\beta}\right)t''_C(e_1)} > 0 \tag{6}$$

The positive sign obtained for expression (6) relies on the assumed convexity for the punctuality function $t_C(e_1)$, the convexity of its associated cost $g(e_1)$, and the punctuality-improving effect of the effort, $t_C'(e_1) < 0$.

The interpretation of the result is clear: the larger is an airline, in terms of total airport operations, the larger the effort in punctuality for its flights that it will choose. This result has been found before by Daniel (1995) and Brueckner (2002), in the context of different models.

We are facing here a classical externality problem: airline 1 knows that by exerting an effort e_1 it can improve the punctuality of its flights, but it also brings in a positive effect over flights of airline 2. However, the cost of punctuality efforts are borne solely by the airline, and it cannot ask for compensation from the other agent. As the market share of an airline is larger, this positive externality is increasingly internalized. The maximum level of effort, compatible with profit maximization, will be exerted by a monopolist airline 1 (a case with α=1).

The second result which can be obtained from (4) is the effect of the parameter β (intensity of interaction between flights) over the punctuality effort. As before, using the implicit function theorem, it is found that the effect is also positive:

$$\frac{\partial e_1}{\partial\beta} = -\frac{\alpha\,n\,\lambda\,t'_C(e_1)\left(\dfrac{\alpha}{1-\beta}\right)}{g''(e_1)+\alpha\,n\,\lambda\left(\dfrac{1+\alpha\beta}{1-\beta}\right)t''_C(e_1)} > 0 \tag{7}$$

As the interaction between flights grows stronger and each flight is more affected by the average delay, the airline has more incentives to invest, in order to reduce the part of flight delays which can be manageable (t_C).

Once the equilibrium at the airport is solved for, the next step is to wonder what would be the optimal solution in social terms, and to compare the two outcomes. In order to find the social optimum (n^*, e_1^*, e_2^*) we define a social welfare function W as the (unweighted) sum of benefits of airport and airlines, plus the consumer surplus obtained by the group of air travellers (EC):

$$W(n,e_1,e_2) = EC + \Pi_A + \Pi_1 + \Pi_2 = n\left[D(\overline{U} - v\bar{t}) - c_A - c_C - \lambda\bar{t}\right] - g(e_1) - g(e_2) \qquad (8)$$

The social optimal solution is achieved by maximizing $W(n, e_1, e_2)$ with respect to the three endogenous variables. The choice of the optimal number of operation yields the following condition:

$$(n) \qquad D(\overline{U} - v\bar{t}) = c_A + c_C + \lambda\bar{t} + n^*(Dv + \lambda)\frac{t'_A(n^*)}{1-\beta} \qquad (9)$$

Expression (9) is the rule to set how many operations should ideally be performed at the airport, taking into account some degree of congestion which would be represented in the expression by the term $\bar{t}$, and determined by the optimal punctuality efforts e_1^* and e_2^*.

The number of flights should be increased up to a point when the last operation induces a social gain equal to the full marginal cost that the flight imposes. The LHS of expression (9) is the net utility that all passengers in the last flight obtain, deducting the value of time spent because of the average delay $\bar{t}$, but not considering the price of the ticket. As the social welfare function has been defined, payments among agents do not affect the equilibrium, because they represent mere transfers of value (observe that neither P_C, nor P_A enter the expressions).

The RHS of (9) represents the full marginal costs of the last flight, including the marginal costs of airline and airport, the congestion cost $\lambda\bar{t}$ borne by the airline (because the last flight will not arrive on schedule as a consequence of the level of congestion $\bar{t}$), and a last term which reflects the external cost caused by that flight. The addition of another operation rises the delay per plane in $t'_A/(1-\beta)$ minutes. This generates a total congestion cost of $n^*\lambda\, t'_A/(1-\beta)$ for airlines and $n^*\, D\, v\, t'_A/(1-\beta)$ for passengers.

As the airport is not obliged to pay for compensation for these congestions costs, it is likely that the number of operations programmed by the airport manager might be typically too high in social terms. This will not be the case if the number of operations is small, since by assumption in that case the delay $t_A(n)$ originated by airport causes will be negligible. However, when the number of operations reaches some point close to capacity limit *F*, it would be optimal to study if an additional flight does not disrupt the rest of operations up to some degree that does not compensate to try to accommodate it at the airport, or it should be denied access.

The other two first-order conditions of the maximization of W are:

$$(e_1) \qquad -\alpha n^*(Dv+\lambda)\frac{1}{1-\beta}t'_C(e_1^*)=g'(e_1^*) \tag{10}$$

$$(e_2) \qquad -(1-\alpha)n^*(Dv+\lambda)\frac{1}{1-\beta}t'_C(e_2^*)=g'(e_2^*) \tag{11}$$

Expressions (10) and (11) can be compared respectively with (4) and (5) (we will discuss here only (10)-(4), the other pair is analogous). It can be proved that, since we may use the results $Dv+\lambda > \lambda$, $1/(1-\beta)>1+\alpha\beta/(1-\beta)$, and remembering that t_C' and g' are, respectively, decreasing and increasing functions in e_1, the social level of effort e_1^* will be generally higher than the solution chosen by the airline when maximizing its private profits.

The conclusion is, therefore, that the congestion costs experienced at the airport, when decisions are taken separately by airport and airlines, are caused partly by the choice of lower than optimal efforts in punctuality by airlines, and partly because the airport is interested in maximizing the number of operations, rather than considering a rule to optimize this number so as to avoid congestion.

Another interesting result drawn from this one-period model, is that congestion costs are not driven down to zero, even when taking a social perspective. At the equilibrium obtained, we may easily find that some given degree of congestion $\bar{t}$ is optimal. The intuition is that the punctuality efforts by airlines have a cost, which must be pondered when deciding to fully eliminate flight delays.

3.3 Internalization of congestion costs

One interesting question which can be answered within the framework of our model is whether it would be feasible to internalize congestion costs. In other words: can the socially optimal solution (n^*, e_1^*, e_2^*) be achieved by decentralized decisions?

The answer requires first the calculation of external costs imposed by each agent on the others, and then to re-write the optimization problems of airport and airlines when a system of congestion fees is used, to assess if the socially optimal solution is reached.

The calculation of congestion costs can be easily performed in our model, due to the additive cost structure chosen for the total delay times t_1 and t_2, and the linear form of the congestion cost function $G(t) = \lambda\, t$.

The airport, with its decision on the number of slots n to be allocated is generating a delay per flight of $t_A(n)/(1-\beta)$ minutes. For passengers, this average delay translates into total congestion costs of $n\ D\ v\ t_A(n)\ (1-\beta)^{-1}$. For airlines, it translates into an average cost per flight of $\lambda\ t_A(n)(1-\beta)^{-1}$.

Airline 1 takes its decision on effort e_1. This generates some delay time over all flights operating at the airport, although the effect is higher for its own flights than for those of the other company (because the indirect effect runs through the interaction among all flights). The average delay of a flight of airline 1 attributable to its decision is $[1+ \alpha\beta/(1-\beta)]\ t_C\ (e_1)$. The average delay imposed on flights of airline 2 is $\alpha\beta(1-\beta)^{-1}t_C(e_1)$. Combining both results, passengers' congestion costs generated by the decision of airline 1 on punctuality effort are $\alpha\ n\ D\ v\ t_C\,(e_1)\ (1-\beta)^{-1}$.

In a similar fashion, the impacts of airline 2's decision on e_2 can be assessed. Average delay imposed on flights of airline 1 is $(1-\alpha)\beta(1-\beta)^{-1}t_C(e_2)$, while on its own flights is equal to $[1+ (1-\alpha)\beta/(1-\beta)]\ t_C\ (e_2)$. Total congestion costs on passengers attributable to airline 2 are $(1-\alpha)\ n\ D\ v\ t_C\,(e_2)\ (1-\beta)^{-1}$.

A summary of total fees to be paid and compensations to be received by all agents is useful to proceed further to solve for the optimization problems:

Agent	Congestion fee payable	Compensation received for congestion costs
Airport	$n\,(Dv+\lambda)\left(\dfrac{t_A(n)}{1-\beta}\right)$	-
Airline 1	$\alpha\, n\left[(1-\alpha)\beta\lambda+Dv\right]\left(\dfrac{t_C(e_1)}{1-\beta}\right)$	$\alpha\, n\,\lambda\left(\dfrac{t_A(n)+(1-\alpha)\beta t_C(e_2)}{1-\beta}\right)$
Airline 2	$(1-\alpha)\,n\left[\alpha\beta\lambda+Dv\right]\left(\dfrac{t_C(e_2)}{1-\beta}\right)$	$(1-\alpha)\,n\,\lambda\left(\dfrac{t_A(n)+\alpha\beta t_C(e_1)}{1-\beta}\right)$
Passengers	-	$n\,D\,v\left(\dfrac{t_A(n)+\alpha t_C(e_1)+(1-\alpha)t_C(e_2)}{1-\beta}\right)$

The airport does not receive any compensation because in this model congestion does not affect to its costs. Similarly, passengers are not subject to any congestion fees, because it has been considered that their impact on average delays is negligible. However, by this system of payments, they would be fully compensated for total congestion costs suffered by flight delays ($n\,D\,v\,\bar{t}$).

Consider first the decision, under the system of compensation fees, by airline 1. Profits now are altered, because the company suffers from lower costs on its flights (due to the compensations received), but must pay penalty fees for the congestion costs imposed on others:

$$\underset{e_1}{Max}\left[P_C-c_C-P_A-\lambda\left(1+\frac{\alpha\beta}{1-\beta}\right)t_C(e_1)\right]\alpha n-g(e_1)-\alpha\, n\left[(1-\alpha)\beta\lambda+Dv\right]\left(\frac{t_C(e_1)}{1-\beta}\right)\quad(12)$$

The first order condition of this problem is exactly the same as in expression (10) above. Therefore, we may conclude that this congestion fee system is successful in terms of reaching the socially optimal efforts e_1* and e_2* on flight punctuality from airlines (the analysis for airline 2 follows exactly the same procedure and it is not detailed here). This is hardly surprising, because airlines are now replicating the same problem which is solved by the social planner. The key issue is that, under the penalty system, airlines must take into account the full costs of their decisions.

For the case of the airport, its profit maximization problem now is transformed into:

$$\begin{aligned} &\underset{n}{Max}\ (P_A - c_A)\, n - n\,(Dv + \lambda)\left(\frac{t_A(n)}{1-\beta}\right) \\ &s.a. \quad n \leq f \\ &\qquad\ \ n \leq F \end{aligned} \tag{13}$$

which yields the following first-order condition:

$$P_A - c_A = (Dv + \lambda)\,\frac{t_A(n)}{1-\beta} + n\,(Dv + \lambda)\,\frac{t_A{}'(n)}{1-\beta} \tag{14}$$

The solution achieved by the airport is again a rule of equalization of marginal benefits and marginal costs. If an additional slot is allocated to a company, the airport earns the regulated slot price minus its marginal cost (LHS of expression 14). Without a congestion fee system, the airport faces zero marginal costs associated to congestion, and consequently tries to sell as many slots as possible. However, the introduction of an additional slot now forces the airport to pay for the delay experienced by that flight (first term on the RHS) plus the additional congestion costs generated over all the rest of flights (second term on the RHS). The number of slots determined under the rule set by expression (14) is likely to be smaller than before.

However, it is apparently surprising that, contrary to the behaviour of airlines, the airport will not be generally achieving the socially optimal number of slots $n*$. This statement can be assessed by comparing expressions (14) and (9). Even though the congestion fee system correctly forces the airport to pay for all congestion costs generated, the solution reached is different than the optimal, in terms of number of operations.

The explanation to this result is to be found in the regulation imposed on the airport. Because the regulated price for slots P_A is exogenously determined, the internalization of congestion costs does not automatically guarantee that a profit-maximizing airport will set the same number of operations as the social planner. The reason is that the decision is taken for a given margin (P_A - c_A), which is not necessarily optimal.

This results is a crucial feature of this model: a system of congestion fees imposed on airlines and airport will lead to optimal decisions regarding punctuality efforts by airlines, but not necessarily to the optimal number of operations at the airport, unless regulated slot prices are not arbitrarily set, but coordinated with the whole airport pricing system.

Expression (9) can be re-written to make it more easily comparable to (14):

$$D\overline{U}-(Dv+\lambda)\left(\bar{t}-\frac{t_A(n)}{1-\beta}\right)-c_A-c_C=(Dv+\lambda)\,\frac{t_A(n)}{1-\beta}+n\,(Dv+\lambda)\,\frac{t_A\,'(n)}{1-\beta} \tag{15}$$

Combining (14) and (15), we deduce what should be the optimal regulation rule:

$$P_A=D\overline{U}-c_C-(Dv+\lambda)\left(\bar{t}-\frac{t_A(n)}{1-\beta}\right) \tag{16}$$

The interpretation of expression (16) provides an intuitive result: in order for the airport to set the socially optimal number of slots, it is required that the regulated slot price adequately represents the social value of the use of the infrastructure. The social value is equal to the utility perceived by all travellers using a plane ($D\overline{U}$), minus the airline's normal operating costs (c_C), and minus the part of congestion costs borne both by travellers and airlines for which they are not compensated by the airport. Thus, optimal regulation should take into account the prevailing degree of congestion which is acceptable as normal at the airport, when setting the price of a slot.

3.4 Analysis of a two-period framework

The results obtained with the one-period model described above could be easily extended to a multiple-period context. The basic idea that a well-designed system of fees based on a detailed calculation of congestion costs for all agents involved can achieve the internalization of external effects and socially optimal solutions directly applies. However, the interaction between periods is a real feature which the model must reflect, and consequently both notation and implicit expressions defining the solutions become more complex and will not be presented here.

By extending the framework simply to a two-period model, we want to stress formally a fundamental point that a congestion fee system should pursue consistently: the impact of a flight delay at some time within the day spreads not only over other flights in that period, but also to the next hours of operation at the airport. Consequently, total congestion costs generated by a flight delay heavily depend not only on the period when it occurs, but also importantly on the next closest period.

The implication of this fact for the system of congestion fees proposed here is that an airline causing a delay at a non-peak period followed next by a peak could be liable for payments which could be even higher than congestion fees calculated for a delay at the own peak-period. This is a socially desirable feature which has so far not been previously pointed out in the literature of airport congestion: if an agent must pay for the external costs generated, the fee cannot be calculated automatically as a function, for example, of the number of flights operating at the same one-hour interval. Ideally, it should also take into account which are the congestion costs generated, given the situation at t+1, t+2, etc.

In order to show this, consider the following extension of the model. Assume now the airport operates during two periods, i and i+1. For simplicity, the number of flights is assumed to be the same at both periods, $n^i = n^{i+1} = n$. Airline 1 now must take its decision on a two-dimensional vector of punctuality efforts (e_1^i, e_1^{i+1}). The decision taken at the first period will have some influence over the punctuality of period i but also of period i+1, and will affect both to its flights and to a lesser extent to the flights of airline 2, and consequently will also affect to all passengers.

Delay times for flights of airline 1 are now modelled as follows:

$$t_1^i = t_A(n) + t_C(e_1^i) + \beta_1 \overline{t^i} \tag{17}$$

$$t_1^{i+1} = t_A(n) + t_C(e_1^{i+1}) + \beta_1 \overline{t^{i+1}} + \beta_2 \overline{t^i} \tag{18}$$

Expression (18) reflects how the delay time experienced by a flight of airline 1 at period i+1 depends on how the airport worked on average during the previous i period, through the total average delay time $\overline{t^i}$. The effect is introduced with a parameter β_2, which will be typically smaller than β_1, but which could take a relatively high value if the interaction between the two periods is important. In any scenario, we will have 0 # $\beta_2 < \beta_1 < 1$.

In order to solve the model, it is useful to combine expressions (17) and (18) with the corresponding values for airline 2, and to obtain explicit expressions for t_1^i and t_1^{i+1} as functions only of the endogenous variables:

$$t_1^i = \frac{t_A(n)}{1-\beta_1} + \left(1 + \frac{\alpha\beta_1}{1-\beta_1}\right) t_C(e_1^i) + \frac{(1-\alpha)\beta_1}{1-\beta_1} t_C(e_2^i) \tag{19}$$

$$t_1^{i+1} = \frac{t_A(n)}{1-\beta_1} + \frac{\alpha\beta_2}{(1-\beta_1)^2} t_C(e_1^i) + \left(1 + \frac{\alpha\beta_1}{1-\beta_1}\right) t_C(e_1^{i+1})$$

$$+ \frac{(1-\alpha)\beta_2}{(1-\beta_1)^2} t_C(e_2^i) + \frac{(1-\alpha)\beta_1}{1-\beta_1} t_C(e_2^{i+1}) \tag{20}$$

The profit-maximization problem of airline 1 is now:

$$\underset{e_i^i, e_i^{i+1}}{Max} \; 2\alpha n(P_C - c_C - P_A) - \alpha n\lambda[t_1^i(e_1^i) + t_1^{i+1}(e_1^i, e_1^{i+1})] - g(e_1^i) - g(e_1^{i+1}) \tag{21}$$

with first-order conditions:

$$(e_1^i) \qquad -\alpha n\lambda\left[\left(1 + \frac{\alpha\beta_1}{1-\beta_1}\right) + \frac{\alpha\beta_2}{(1-\beta_1)^2}\right] t_C\,'(e_1^i) - g\,'(e_1^i) = 0 \tag{22}$$

$$(e_1^{i+1}) \qquad -\alpha n\lambda\left(1 + \frac{\alpha\beta_1}{1-\beta_1}\right) t_C\,'(e_1^{i+1}) - g\,'(e_1^{i+1}) = 0 \tag{23}$$

Expressions (22) and (23) are the optimal rules to set punctuality efforts by airline 1. The logic of these rules is exactly the same as in the one-period model. The airline calculates the marginal benefits of efforts and makes them equal to marginal costs. For the case of e_i^1, the effect that it has over the next period i+1 will be considered when computing marginal benefits.

For a relatively small airline, with a low value for α, the second term within brackets in expression (22) will be very low, specially if the interaction between the two periods is not strong. In that case, it can be assumed that $\alpha\beta_2 \to 0$, so the optimal private decision of the company will be to exert the same level of effort on punctuality on both periods, $e_1^i = e_1^{i+1}$.

This behaviour should be corrected by a system of congestion fees, by forcing the airline to realize the greater impact that a flight delay during the period i will have, compared to a similar delay in period i+1. Total congestion costs can be computed to show this extreme. Associated congestion costs to airline 2 and passengers to the decision e_1^i are:

$$C(e_1^i) = \alpha\, n \frac{t_C(e_1^i)}{1-\beta_1}\left[(1-\alpha)\left(\beta_1 + \frac{\beta_2}{1-\beta_1}\right)\lambda + \left(1 + \frac{\beta_2}{1-\beta_1}\right)Dv\right] \tag{24}$$

while total costs associated to e_1^{i+1} are:

$$C(e_1^{i+1}) = \alpha\, n \frac{t_C(e_1^{i+1})}{1-\beta_1}\left[(1-\alpha)\beta_1\lambda + Dv\right] \tag{25}$$

Comparing (24) and (25) it is immediately observed that $C(e_1^{\,i}) > C(e_1^{\,i+1})$, so the company should do a higher effort on punctuality during the period i. Using the calculated total congestion costs, a penalty fee for a flight delay during period i will be much higher than for the same delay during period i+1. Thus, it is possible that in a two-period scenario with different levels of activity ($n^i < n^{i+1}$), congestion fees charged for flight delays during the non-peak period i could be even higher than during the peak-period i+1. The key issue is what happens after the peak-period: if airport activity slows down importantly (consider for example a peak in the late evening) a delay occurring at that time has a lower total congestion cost that a flight delay occurring just before the peak.

4 Empirical results

In order to illustrate the main features of airport congestion discussed in this work, and the validity of assumptions on which the theoretical model is based on, some real data on flight delays at Madrid airport are presented.

Data were obtained from AENA (*Aeropuertos Españoles y Navegación Aérea*), a public institution that owns and manages the system of main Spanish airports. AENA is also responsible for air traffic control in the Spanish airspace. Information is referred to delays experienced by all flights (arrivals and departures) which used the airport during the months of July 1997, July 1998, July 1999 and July 2000. This period is particularly interesting, because the airport significantly expanded its capacity during those years, so the impact on congestion can be studied.

The following table summarizes some indicators of Madrid airport's activity:

	July 1997	July 1998	July 1999	July 2000
Total movements *(% annual change)*	23,002 -	23,036 *(0.1%)*	26,948 *(17.0%)*	31,292 *(16.1%)*
Total passengers *(% annual change)*	3,204,290 -	2,990,599 *(-6.7%)*	3,375,865 *(12.9%)*	3,825,089 *(13.3%)*
Capacity * *(% annual change)*	50 -	50 *(0%)*	68 *(36%)*	68 *(0%)*

* Maximum number of movements (arrivals and departures) per hour.

After computing delays for all observations included in the sample (97,902 flights), some simple descriptive statistics reveal interesting information which helps to understand congestion problems at Madrid airport. Three main findings are remarkable, which are discussed at some length below:

- Arrival and departure delays are highly correlated
- Spill-over effects between one-hour intervals exists
- Expansion of capacity has only slightly eased congestion

4.1 Correlation between arrival and departure delays

Correlation coefficients between arrival and departure delays are extremely high, for each of the four years included in our sample, although in 2000 the link between the two type of series seems to have weakened a litle. Correlations are calculated based on averages of daily data:

1997	1998	1999	2000
0.810	0.960	0.946	0.755

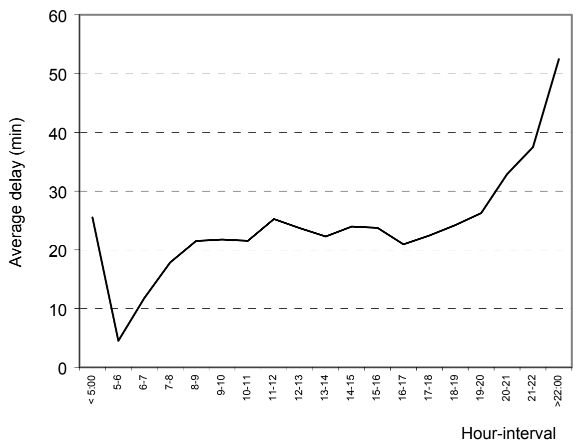
This high positive correlation reveals that whenever there are problems for departures at Madrid airport, the situation is also complicated for arrivals, and vice versa. ATC and ground infrastructure serve simultaneously both to incoming and outgoing flights, thus the result is not surprising.

However, this result is useful to reinforce the idea that congestion is originated *at* the airport, both by the interaction of flights among themselves, and by the required intensive use of airports' assets. Whether congestion problems to be imported from other parts of the European airport system–causing a large part of arrival delays– it could then be expected that correlation between the series of arrival and departure delays would be much lower.

4.2 Spill-over effects between one-hour intervals

The analysis of average delays suffered by all flights using the airport within one-hour intervals illustrates how airport congestion develops through the day. Figure 2 shows how departure delays during the first hours in the morning are low (excluding those flights taking-off before 5:00 am, which are probably delayed due to exogenous causes[4]), but delays follow an accumulating pattern as the day evolves. Non-peak hours (from 10:00 to 16:00) serve as a buffer against this effect, but do not seem to eliminate it completely. Average departure delays start to rise from 16:00 onwards, and follow an increasing pattern until the end of the day.

Figure 2. Congestion spill-over effects



This pattern of flight delays at Madrid airport exactly fits our previous discussion about the interaction between airports' operation not only within short periods, but also over some more distant ones. The assumption used in the two-period version of the theoretical model about the average delay time of period *i* having an influence over the next period *i*+1 seems to be empirically justified.

A regression analysis using data on average delays calculated over one-hour intervals (using the >15-minute definition of delay) shows that each delayed flight generates an impact that lasts significantly during at least two hours since the moment the flight is authorized to use the airport. The spill-over effect can be observed, for example, in the following equation estimated with data on arrivals for July 1997:

$$T_H^{av} = -3.30 + 1.325\ N_H + 0.86\ N_{H-1} + 0.54\ N_{H-2}$$
$$\quad (-2.46) \quad (10.51) \quad (6.56) \quad (3.78) \quad (t - ratios)$$

$$R^2 = 0.489$$

where T_H^{av} is the average delay time over the schedule (expressed in minutes) of all flights landing within the one-hour interval *H*, and N_H is the number of flights with delays higher than 15 minutes arriving at the airport during the period *H* (respectively N_{H-1} and N_{H-2}). The interpretation of the coefficients is that each delayed flight arriving at the *H* hour is causing an excess of time of 1.325 minutes to all flights within that *H* hour interval, and will also have an impact of 0.86 minutes for the next hour-interval *H+1*, and 0.54 for the next hour *H+2*.

However, the number of flights and passenger varies at each hour-interval *H*. Therefore the impact of a delayed flight entering the airport at 10:00 am is different to that of a delayed flight at 21:00 pm. The impact is clearly much lower for the latter, because airport activity slows down after 22:00 pm).

4.3 Effects of capacity expansion at Madrid airport

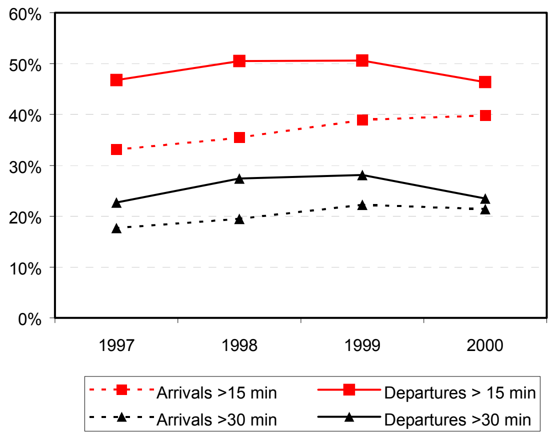
Another point which has been examined is the evolution of congestion during the period of reference (1997-2000). The objective is to check whether the expansion of Madrid airport capacity introduced in 1999 has significantly altered average delays. Figure 3 presents our results separately for arrivals and departures, and using the two definitions of delays used in the industry (>15 minutes and >30 minutes).

The decreasing trend observed between 1999 and 2000 indicates that capacity expansion seems to have had a slight positive effect on congestion problems. This is specially true for the longest delays (more than 30 minutes), where the impact is observed both on arrivals and departures. However, given the magnitude of the capacity increase (36%), it could have been expected that the reduction in the percentage of delayed flights would have been more important.

---

[4] The number of flights using Madrid airport before 5:00 am and after 22:00 pm is small. The arrival/departure of a flight with a long delay may then easily induce the average to be very high. These flights are considered not representative, and have been excluded for the computation of marginal effects.

Figure 3: Percentage of delayed flights over total movements, 1997-2000



Combining this finding with the observed huge increase in the number of flights from/to some particular origins/destinations, our conclusion is that the 1999 capacity expansion at Madrid airport served mostly to accommodate new flights, which the airport could not attend before. Investments on infrastructure thus seems to have solved at least partly the scarcity problems suffered by the airport, but did not have a major impact in solving the serious problems regarding congestion. Data on total number of hours lost by passengers confirms this extreme.

The observation of this behaviour seems to be according with our result obtained in the theoretical model: airport managers are more interested in maximizing the number of operations, rather than in solving congestion.

4.4 Quantification of congestion costs

Congestion costs can be empirically evaluated with flight delays data, by computing total extra time spent by passengers and airlines, and using some assumptions on the value of time for passengers and costs for airlines. Our dataset is sufficiently rich to perform these calculations, because the actual number of passengers at each flight is known, and also the type of aircraft providing the service, which allows to apply differentiated hourly rates per plane to evaluate airlines' congestion costs.

The results obtained are presented in the following tables:

Passenger congestion costs: total and average

	July 1997	July 1998	July 1999	July 2000
Total costs (million €)				
Arrivals [1]	6.42	7.39	8.71	8.84
Departures [1]	7.94	7.64	8.60	7.34
Total flights [1]	***14.36***	***15.03***	***17.31***	***16.18***
Average costs (in €/passenger)				
Arrivals	4.04	4.36	4.51	4.01
Departures	4.91	5.89	5.95	4.53
Total flights	***4.49***	***5.02***	***5.14***	***4.24***

Notes: [1] Total monthly costs

Airlines' congestion costs (million €)

	July 1997	July 1998	July 1999	July 2000
Arrivals	14.7	16.7	20.3	22.0
Departures	18.0	17.6	20.4	17.2
Total flights	***32.7***	***34.3***	***40.7***	***39.2***

Note: Total monthly costs

Passenger congestion costs reveal the magnitude of the problem suffered at Madrid airport. In July 2000, total passenger costs amounted to 16.2 million €. Average costs are estimated between 4.5-5 € per passenger. The process of airport expansion induced some reduction in the average cost per passenger obtained in 2000 compared to previous years.

Adding congestion costs borne by passengers and airlines, the importance of the problem of congestion experienced at Madrid airport can be assessed. Taking July-2000 as a month of reference, total congestion costs in 2000 amounted to 55.4 million € per month (16.2 million corresponding to passenger costs, 39.2 million to airlines). In annual terms, assuming that July could be considered a representative month of the activity of Madrid airport, total congestion costs are evaluated at 664.8 million €.

4.5 Marginal congestion costs

A second exercise performed with our database of delays is to evaluate the marginal effect that each delayed flight causes over the rest of flights. The idea is to provide an approximation of how the system of congestion fees proposed in the theoretical model could be implemented in practice. This can be done, for example, by estimating equations of the form:

$$T_{day} = \alpha + \beta \, N_{day} \qquad (26)$$

where T_{day} is the average extra time of all flights using the airport during each day included in the sample, and N_{day} is the number of delayed flights in the same day.

Results indicate that the marginal effects of delayed flights have changed over the period 1997-2000, due to the airport enlargement, so it is considered that observations from different years cannot be pooled together. The following table presents results obtained for each year, separately for arrivals and departures:

Estimation of marginal congestion costs

	July 1997	July 1998	July 1999	July 2000
Arrivals *	0.158 (15.03) $R^2 = 0.89$	0.197 (11.67) $R^2 = 0.87$	0.145 (21.99) $R^2 = 0.96$	0.113 (11.05) $R^2 = 0.86$
Departures *	0.151 (11.10) $R^2 = 0.85$	0.201 (12.10) $R^2 = 0.88$	0.167 (13.69) $R^2 = 0.90$	0.117 (5.55) $R^2 = 0.62$
Marginal congestion costs (thousand €)				
Arrivals	**6.59**	**9.25**	**7.88**	**7.07**
Departures	**6.76**	**8.72**	**8.34**	**6.71**

* Coefficients can be interpreted as minutes imposed on average by a delayed flight over all the rest of flights using the airport during the day (between parenthesis, t-ratios).

Estimated coefficients can be used to calculate marginal congestion costs. Multiplying those coefficients by the total number of flights and passengers using Madrid airport on average during one day of each of the periods of reference, it is possible to calculate the amount of total extra time that one delayed flight imposes on airlines and passengers. The conclusions that can be obtained from these results are: (a) each delayed flight at Madrid airport causes congestion costs estimated around 7,000 € in July 2000; (b) marginal congestion costs seem to have improved after the enlargement or airport capacity.

Even though the achievement of full congestion cost internalization would required more refined estimates of marginal congestion costs (specially differentiating by period of the day when delays occur), the exercise presented here aims to illustrate that calculations are relatively easy to perform, and could be used in practice to implement a congestion fees pricing system.

5 Conclusions

Airport congestion tends to be associated with lack of infrastructure to accommodate the required demand for landings and take-offs. Investments in new capacity are frequently demanded as a solution to this problem, and the economic literature has insisted in proposing a more rational pricing system capable of rationing the excess of demand (peak-load pricing).

In this work, we argue that the characteristics of airport congestion are rather different than those of road congestion. Therefore, a policy based on peak-load pricing and investment (although highly recommendable in some particular cases of airports with severe lack of capacity and/or acute peak hours) is probably not a long-term solution to the problem of air congestion. This is specially true in a context of rapidly growing demand for air transport, and passengers with marked preferences for arrival/departure periods. We believe both these factors apply generally in major European airports.

Airport congestion pricing should not be reduced to a peak-pricing problem. Ideally, agents causing delays should pay for the marginal cost of congestion. A theoretical model presented in this paper shows how internalization of congestion costs could be achieved, simply by using a congestion fee system which forces airlines and airports to compensate each other and passengers for the external congestion costs imposed on them by flight delays.

The model also helps to understand a crucial issue: even though descentralized decisions by airlines and airports, in a context of internalization of congestion costs, could lead to a socially optimal solution, it requires that regulation should be coordinated with this kind of policy. Regulation of slot prices cannot be done arbitrarily if congestion fees are to be successful in terms of airports' level of activity. The regulated price of a slot should ideally reflect the full social value of its use, given some optimal degree of congestion reached when airlines internalize the external congestion costs imposed on others.

Currently, some major airports have sophisticated information systems which allow to reflect accurately the causes of flight delays. It is even possible to disentangle a single delay into the time attributable to different causes (airlines' problems, ATC and airport induced delays, passengers and cargo, reactionary delays, and so forth). The combination of this information with estimated marginal congestion costs could be used in practice to implement a congestion fee system.

Ideally, airport congestion fees should reflect the real impacts caused over all agents. Given the observed fact that flights interact over relatively long periods (a delay at one hour t may have an impact up to the t+2 hour), delays at early hours within the day should be more heavily penalized than those at the end of the day, even if they occur during non-peak hours. The usual recommendations of peak-load pricing for airport slots ignore this fact, and thus are not likely to provide airlines (and airports) with the right incentives to put the maximum effort in avoiding flight delays.

6 References

ATAG (Air Transport Action Group) (2000) 'The Economic Benefits of Air Transport', mimeo, www.atag.org

Brueckner, J K (2002a) 'Airport Congestion when Carriers Have Market Power' American Economic Review 92 (5) 1357-75

Brueckner, J K (2002b) 'Internalization of Airport Congestion' Journal of Air Transport Management 8, 141-147

Carlin, A and Park, R E (1970) 'Marginal Cost Pricing of Airport Runway Capacity' American Economic Review 60, 310-319

Daniel, J I (1995) 'Congestion Pricing and Capacity of Large Hub Airports: A Bottleneck Model with Stochastic Queues' Econometrica 63, 327-70

Daniel, J I (2001) 'Distributional Consequences of Airport Congestion Pricing' Journal of Urban Economics 50, 230-58

Daniel, J I and Pawha, M (2000) 'Comparison of Three Empirical Models of Airport Congestion' Journal of Urban Economics 47, 1-38

Doll, C et al (2000) User Costs and Benefits (External) Categories, UNITE Research Project 1999-AM.11157, EU Commission, Brussels

Fisher, J B (1989) 'Managing Demand to Reduce Airport Congestion and Delays' Transportation Research Record 1218, 1-10

Hansen M (2002) 'Micro-Level Analysis of Airport Delay Externalities Using Deterministic Queuing Models: A Case Study' Journal of Air Transport Management 8, 73-87

Levine, M E (1969) 'Landing Fees and the Airport Congestion Problem' Journal of Law and Economics 12, 79-108

Mohring, H and Harwitz, M (1962) Highway Benefits: An Analytical Framework, Northwestern University Press

Morrison, S A (1983) 'Estimation of Long-Run Prices and Investment Levels for Airport Runways' Research in Transportation Economics 1, 103-30

Morrison, S A and Winston, C (1989) 'Enhancing the PErforman of the Deregulated Air Transportation System' Brookings Papers on Economic Affairs, Microeconomics, 1, 61-123

Oum, T H and Zhang, Y (1990) 'Airport Pricing: Congestion Tolls, Lumpy Investment and Cost Recovery' Journal of Public Economics 43, 353-74

Reynolds-Feighan, A J and Button, K J (1999) 'An Assessment of the Capacity and Congestion Levels at European Airports' Journal of Air Transport Management 5, 113-134

Starkie, D (1988) 'Allocating Airport Slots: A Role for the Market?' Journal of Air Transport Management 4, 111-16

Strotz, R (1964) 'Urban Transportation Parables' in Margolis, J (ed) The Public Economy of Urban Communities, Johns Hopkins U. Press

Vickrey W (1963) 'Pricing in Urban and Suburban Transportation' American Economic Review (Papers and Proceedings) 53, 452-65

Walters, A (1961) 'The Theory and Measurement of Private and Social Cost of Highway Congestion' Econometrica 29, 676-99

Wolf, H (1998) 'Airport Regulation: Tackling Congestion and Environmental Problems' Working Paper 876, Institut für Weltwirtschaft, Kiel.