Airport Congestion and Noise: Interplay of Allocation and Distribution

Wayne Brough, Edward Clarke, and Nicolaus Tideman

The history of airport takeoff and landing rights is discussed, focusing on how distributive concerns have dominated the resolution of allocative issues. A mechanism is needed to separate allocation and distribution, to permit the simultaneous optimization of the use of scarce resources and minimization of the social cost of unintended redistribution. From this perspective, two allocative mechanisms are compared: a free market in transferable slots and an administered market, using compensated incentive compatibility (CIC), a version of the demand-revealing process. If the only concern is the efficient allocation of a specified number of slots, either mechanism will work. However, if the number of slots is to be optimized endogenously, CIC is needed. An additional important advantage of CIC arises from its capacity to provide optimal control of noise pollution. The CIC mechanism is useful not only for controlling public nuisances such as pollution, for which the mechanism was originally developed, but also for allocating a private good such as airport slots, where the government creates scarcity values in the process of controlling externalities such as noise and congestion. The conclusions extend to other settings in which allocation and distribution are intertwined and for which externalities need to be managed.

The decade following airline deregulation sparked great controversy over the role of the market, particularly in dealing with supply-side constraints (7). The growth in air travel put pressure on existing capacity and created increased congestion.

HISTORY OF SLOT MANAGEMENT

The creation of a market in airport takeoff and landing rights (slots) in 1986 was viewed as a significant, albeit limited, first step toward relieving an important constraint on the supply of airline services and improving the efficiency with which scarce airport capacity is used.

The slot trading (or buy-sell) program replaced a system of quotas for rationing capacity at some airports. The quota system was initially established by FAA in 1969 to allocate capacity at the four busiest airports—O'Hare, Kennedy, LaGuardia, and Washington/National. At these airports, the "high-density" rule established "slots" that represented the right to take off or land within an interval of typically 30 to 60 min. The high-density rule replaced a first-come, first-served rule that is still in effect at the majority of U.S. airports.

At the four high-density airports, slots for trunk carriers were originally allocated by scheduling committees. The scheduling committees met twice a year and were required to reach unanimous agreement, or the FAA would step in and impose its own allocation. Both incumbent carriers and new entrants served on these committees; the result was often extremely protracted negotiations. Trades in slots occasionally were possible, but generally only on a one-for-one basis for another slot at the same airport. In addition to inhibiting the most efficient use of existing slots, the quota system appears to have imposed considerable restraint on new entry.

FAA introduced the buy-sell program when scheduling committees were no longer able to reach unanimous agreement. This program, which became effective in April 1986, "grandfathered" existing users by allocating to them the slots they were using at the four airports as of December 1985. The rule also permitted the holders of slots to sell them.

INTERPLAY OF ALLOCATION AND DISTRIBUTION

The evolution of the buy-sell program, a system of de facto property rights in slots, illustrates the dominance of distributive considerations in the choice of allocative mechanisms. The evolution of the buy-sell rule and the reasons why it was adopted in lieu of other approaches, such as auctions or congestion fees, is discussed by Riker and Sened (2). The evolution of this market also illustrates the conflicts among three competing principles for settling distributive questions: efficiency, equality, and distributive stability.

When slots were not scarce, all three principles could be honored by allowing everyone who wished to do so to take off and land at airports, with charges only for the cost of managing the takeoff and landing process. However, when congestion and delays developed, the principle of efficiency was sacrificed. Equality was preserved. Distributive stability was compromised. It continued to be possible for everyone who wanted to take off or land to do so, but only at a cost of waiting in line.

The high-density rule of unanimously agreed quotas represented a compromise with all three principles. Unanimously agreed quotas were probably more efficient than the earlier congestion, but because only limited subsequent trading was permitted, slots often remained in the hands of carriers other than those that could use them most productively. A kind of equality was achieved by allowing all carriers to participate in the scheduling committees. However, to the extent that participants believed that a deadlock would lead FAA to impose an unequal allocation, the bargains they reached reflected that inequality. Distributive stability was sacrificed to the extent that existing carriers were obliged to agree to the assignment to new entrants of some slots that they had held previously.
The introduction of the buy-sell rule in 1986 improved efficiency while preserving stability and sacrificing equality. The buy-sell rules permitted slots to be transferred to those carriers for whom they were more valuable, thereby promoting efficiency. The principle of stability was honored by establishing initial allocation of slots according to usage at a time 4 months before the start of the buy-sell program. However, the assignment of slots only to those who had previously held them sacrificed the principle of equality. A variation on the buy-sell rules in which slots were initially auctioned would also have promoted efficiency while honoring the principle of equality, but would have sacrificed the principle of stability.

One alternative to the buy-sell program is a program of rationing slots by rental auctions or some other administered market. This has the potential to be at least as efficient as the buy-sell program, and possibly more efficient, because it may be easier under an administered market to vary the number of slots to take account of changes in the optimal number of takeoffs and landings. An administered market can honor the principle of equality by collecting the value of slots socially instead of assigning it to the carriers who had been using slots. However, by requiring carriers to pay for what they had previously received for nothing, an administered market sacrifices the principle of stability. Alternatively, an administered market can honor the principle of stability and sacrifice the principle of equality, if market-clearing fees are collected and allocated among carriers in proportion to their previous use.

The central distributive question with respect to slot management is as follows: When restrictions on takeoffs and landings are introduced at an airport for the sake of efficiency, should the value of exclusive takeoff and landing rights go to the carriers who had been using the airport (thereby honoring the principle of stability) or to the public treasury (thereby honoring the principle of equality), or should there be some compromise between these two principles?

The question may have no easy answer. We maintain that a history of use of common property does not create an exclusive right to privileged access when the opportunity to use such property becomes scarce. At the same time, it is reasonable to permit commitments of slots to carriers for some span of time—for example, in exchange for the carriers’ investments in developing schedules. Accordingly, there could be a transition, at a rate that was appropriate in view of previous commitments, from entitlements based on past usage to social collection of the scarcity value of slots.

The important point here is that there is a tendency for distributive considerations to influence the choice of allocative mechanisms. To be confident that an allocative mechanism is chosen well despite the tendency of distributive considerations to influence the choice, it is desirable that the distributive issue be settled separately. For example, if the span of time for which existing carriers deserve special access is decided first, it is possible to assess alternative allocative mechanisms without the bias that comes from potential distributive consequences. Then, after the allocative mechanism has been chosen, one can adapt it to reflect the previously chosen rights of existing carriers.

ALLOCATION OF SLOTS BY TRANSFERABLE OWNERSHIP

We noted that the evolution of the buy-sell rule was strongly influenced by concerns for distributive stability. Efficiency considerations played an important role as well. The basic efficiency argument for the buy-sell program is that if the owner of a slot is unable to obtain as much value from it as some other carrier, then self-interest will lead the owner to sell it. Because people are also concerned about distributive stability, to many it will seem natural, if slots are to be owned, that initial ownership be assigned to the carriers that have been using the slots. Both efficiency and stability are thus completely satisfied.

This argument is valid for any specified number of slots, but it does not address the question of how many slots there should be. Allowing the wrong number of takeoffs and landings can be very costly. Morrison and Winston (3) have estimated that welfare gains of approximately $4 billion per year could be achieved by optimizing the number of takeoffs and landings. The attempt to schedule more takeoffs and landings in a given interval of time increases the expected delay for all of them. The number of slots in a given hour is optimal if and only if the rental value of a slot in that hour is equal to the cost of the expected additional delay that is caused by trying to squeeze one more slot into that hour. If the marginal delay cost were less than the rental value of a slot, then to achieve efficiency under a system of transferable ownership of slots, the airport authority would need to create new slots. If the marginal delay cost were greater than the rental value of a slot, the authority would need to eliminate one or more existing slots to achieve efficiency.

A need to change the number of slots can arise from a change in technology, a change in the number of runways, or a change in the estimate of the spacing required for safety. If carriers own slots, any increase in the number of slots reduces the market value of the assets of carriers, leading to pressure not to increase the number of slots. If efficiency is to be achieved, it is important that such pressure not be effective. Any reduction in the number of slots requires airport authorities to acquire slots from carriers. This would produce financial losses for airport authorities unless there were some provision for taxing the owners of slots to finance required reductions in the number of slots.

One of the consequences of takeoffs and landings is noise pollution. The cost of this pollution varies by time of day, type of day, type of aircraft, and how the aircraft are flown. To achieve proper incentives for reducing the costs of such pollution, those who take off and land must be charged with these costs. Thus efficiency requires that possession of a slot not insulate carriers from noise pollution charges.

To summarize, it is possible in principle to achieve efficiency through transferable slot ownership, but there are potential problems in adjusting the number of slots in each hour, in adapting to the redistributive consequences of changes in the number of slots in each hour, and in implementing efficient noise pollution charges. These problems are not insurmountable, but they demonstrate that transferable slot ownership has costs, which could be higher than the costs of an alternative allocative mechanism. An alternative mechanism could enhance allocative (and procedural) efficiency while making improved trade-offs between the principles of distributive stability and equality, in part by achieving a more effective separation between distributive and allocative issues.

ALLOCATION OF SLOTS BY A COMPENSATED INCENTIVE COMPATIBLE PROCESS

The origins of our proposal are Vickrey’s concept (4) of a ‘‘second-price auction’’ and Clarke’s idea (5) of levying charges on individuals for public goods according to the marginal social costs of accommodating their stated departures from standardized demand
schedules, an idea to which Tideman and Tullock (6) gave the name "the demand-revealing process." The principal virtue of these mechanisms is that they motivate individuals to report their preferences honestly, thereby permitting efficient allocations to be identified and implemented. Dolan (7) described the way in which the demand-revealing process could be employed to optimize the use of a congested facility. Economics, Inc. (8) recommended that FAA use a second-price auction to motivate accurate revelation of willingness to pay for slots.

The authors' proposal involves a variation on second-price auctions and the demand-revealing process that was developed by Tideman (9). This variation, called compensated incentive compatibility (CIC), optimizes the use of a public resource while separating allocative efficiency from the distribution of the revenues generated by efficient pricing (in this case, of slots). Our proposal can also be considered a variant of the zero-revenue auction (ZRA) approach to emissions-rights allocation, which has appeared in the environmental literature (10,11).

In the remainder of this section, we introduce the CIC approach to slot allocation as a refinement of traditional auctions and ZRAs. We then introduce further refinements that take more precise account of congestion costs by better reflecting differences in priority status in a queue, building on Dolan's work (7). Finally, we develop a different application of CIC for dealing with the noise from takeoffs and landings.

Traditional Auctions

Under a traditional approach to the auction of slots, each carrier submits a demand schedule for slots (i.e., a statement of the quantity of slots that the carrier would wish to rent at each possible price). These individual demand schedules are added horizontally to construct an aggregate demand schedule for slots. There is an exogenous limit on the number of slots, set by legislation or regulation. If aggregate demand is less than this quantity at a price equal to the sum of noise and other external costs, then this price is the price of a slot. If aggregate demand is more than the limit at this price, then the price is such that aggregate demand is equal to the limit.

The theory of just distribution that is incorporated in a traditional auction is that slots are public property, so that if they are scarce, any carrier that wants to use slots ought to pay for them according to their market price.

A traditional auction is efficient if the number of carriers is so great that no carrier considers it worthwhile to take account of the effect of his reported demand schedule on the price. When the number of carriers is not so great, there is a modification of the traditional auction that is efficient. This is to require each carrier to pay for each slot it obtains according to the marginal social cost of that slot. That is, the price of a carrier's first slot would be the lowest price at which all other carriers had an aggregate demand for all but one slot, the price of the second slot would be the lowest price at which all other carriers had an aggregate demand for all but two slots, and so on. This method is an application of the principle behind Vickrey's second-price auction (4) and is the method of slot allocation recommended to FAA by Economics, Inc. (8). Because each carrier would pay for slots according to their marginal social cost, each carrier would have an incentive to report its demand schedule accurately. However, if the very considerable revenue that would be generated were to be returned to carriers, then there would be incentives to misstate demands.

ZRAs

In a ZRA, a system administrator assigns each carrier a proportional allocation of slots. As with a traditional auction, all carriers report their demand schedules. The intersection of aggregate demand with an aggregate quantity constraint determines the equilibrium price of slots. Each carrier's final allocation is its quantity demanded at the equilibrium price. Each carrier pays for slots at the equilibrium price and receives back an amount of money equal to the product of the equilibrium price and its provisional allocation. Net payments for all carriers taken together are zero. The theory of just distribution that is incorporated in the ZRA procedure is that each carrier has an entitlement to its initial allocation of slots, but this entitlement is protected by a liability rule rather than a property rule: each carrier may be required to sell some of its slots to any other carrier that values them more highly.

One problem with the ZRA procedure, noted by Hahn and Noll (10) and by others as well, is that when a carrier recognizes that it can affect the equilibrium price of slots, it will have an incentive to report something other than its true demand schedule. This problem can be remedied by the CIC method of auctioning slots.

CIC

Under CIC, a system administrator assigns each carrier a demand schedule as well as an initial quantity of slots. The assigned demand schedule should be the best estimate of that carrier's actual demand that the administrator can make without using information supplied by the carrier itself. (The use of that information would generate an incentive for biased reports.) The administrator can, however, use information supplied by similarly situated carriers to estimate any given carrier's demand. As with the other auction methods, carriers report their actual demand schedules. Assignment of slots to those who value them most highly leads to reallocations of some slots. But for any carrier whose assigned demand schedule is perfectly accurate, these reallocations occur without harm or benefit.

The amount that a carrier pays for its final allocation of slots is shown in Figure 1. Here $D_1$ is the administrator's estimate of the carrier's demand schedule, and $D_2$ is the schedule that the carrier actually reports. The schedule labeled $S$ is the carrier's "derived supply schedule of slots." The quantity on this schedule at any price, $p$, is found by first determining the total number of slots that can be supplied to all carriers at a marginal social cost

![FIGURE 1 Compensated incentive compatibility.](image-url)
of no more than \( p \). (For the case of a regulated number of slots, one would simply use that number.) One then subtracts the quantities demanded at a price of \( p \) by all carriers other than the one whose demand is shown in the figure. The result is the quantity of slots that can be supplied to this carrier at a marginal social cost of no more than \( p \), and the schedule of all such quantities is the derived supply schedule of slots for the carrier. The carrier's initially assigned quantity of slots is labeled \( q_0 \). The quantity of slots that the carrier would receive if it had reported a demand schedule of \( D_1 \) is labeled \( q_1 \). The quantity of slots that it actually receives, given that it reports a demand schedule of \( D_2 \), is labeled \( q_2 \).

The fee that is charged to the carrier whose demand is shown is the shaded area in Figure 1. This price is the sum of the integral of the carrier's assigned demand schedule from \( q_0 \) to \( q_1 \), and the integral of the carrier's derived supply schedule from \( q_1 \) to \( q_2 \). Either integral can be negative, in which case it represents a payment rather than a payment from the carrier.

The reason that two different schedules are used to price different parts of one carrier's reallocation of slots is as follows: The movement from \( q_1 \) to \( q_2 \) is induced by the deviation of the carrier's reported demand schedule from its assigned schedule. To motivate an honest statement of the schedule, this movement must be priced according to marginal social cost, that is, according to the derived supply schedule. The movement from \( q_0 \) to \( q_1 \) is induced by the fact that \( q_0 \) is not the quantity of slots that would be efficient to assign to the carrier if it had reported that its assigned schedule were accurate. Using the assigned schedule to price this movement reduces unintended redistribution for this movement to the minimum amount feasible.

If the administrator were to estimate all demands perfectly and to assign quantities corresponding to the market-clearing ones, then each \( q_i \) would equal the corresponding \( q_i \) and \( q_j \). Thus, there would be no charges on any carrier. The theory of just distribution that is incorporated in this approach to the allocation of slots is that every carrier has a right to receive, without charge, the quantity of slots that it would demand at a market-clearing price.

An efficient variant of ZRA is created by using CIC with any initial allocations of slots that sum to the total available quantity. At the other extreme, if all carriers are assigned quantities of zero and demand schedules that have zero quantity at all prices, then CIC reduces to the efficient auction procedure described above. But every intermediate distributive outcome can also be achieved efficiently through CIC. If one's theory of just distribution says (as the authors' does) that carriers have a temporary right to their initial usage of slots, this can be incorporated into the CIC approach by starting with initial allocations that correspond to prior usage and then reducing the initial allocations to zero quantities, either gradually or all at once after some chosen period of time.

In general, CIC does not achieve budget balance. To achieve the budget balance that must obtain in the end, the budget deficit or surplus could be allocated among carriers in proportion to their assigned quantities. This would induce a slight incentive for misstatements of demands, but it would be difficult to know what misstatements would be profitable. To eliminate the incentive for misstatements all but infinitesimally, the budget surplus or deficit could be assigned to the general funds of the government.

Noise pollution charges can be incorporated into the CIC simply by announcing the charges and asking carriers to report their demands given those charges. To incorporate the distributive premise that carriers are entitled to continue all or part of their past noise for some span of time, one would supplement the noise pollution charges with credits corresponding to those entitlements. With further refinements, priority status within queues could be incorporated as well.

CIC with Priority Pricing

Dolan (7, p. 432) developed a slot allocation method that incorporated congestion costs defined in terms of the demands of carriers (or passengers) for a guaranteed arrival or departure time or a higher priority in a queue. His static queuing model showed that the priority prices that would minimize congestion were equivalent to the taxes generated by a demand-revealing process. A corresponding modification of CIC is to ask each carrier to state, for each slot in which it expresses an interest, not just the amount of money that it is willing to pay for the slot, but rather the schedule of amounts they would be willing to pay for different priority status. The system administrator would assign not just initial quantities and demand schedules, but a priority status for each slot in each carrier's initial allocation and a priority demand schedule for each slot in which each carrier might have an interest. The final allocation would be the set of slot allocations, each with assigned priority status, that maximized aggregate reported value, subject to the constraints of available capacity. As with the earlier version of CIC, each carrier would pay, according to its assigned schedules, for the movement from its initial allocation to the allocation that it would have had, had it reported that its initial allocation was accurate. It would also pay the net cost to all other carriers that resulted from the deviation of its reported schedules from its assigned schedules.

MANAGING NOISE POLLUTION BY USING CIC

Properly set noise pollution charges will motivate carriers to invest efficiently in quieter airplanes, to operate them in such a way as to minimize noise, and to operate at times of the day when noise is least costly. The allocation of noise pollution fees to those who are adversely affected by the noise can neutralize unjustified redistribution that would otherwise result from variations in the level of noise. But what is the right level of charges for noise?

Noise is a "local public bad." That is, the harm that one person suffers from noise does not add or detract from the harm that his neighbors suffer from the same noise. But as with local public goods, the effect abates with distance from the source. The total cost of airplane noise is the sum of the harm experienced over all persons who are harmed by the noise. A CIC process can be used to motivate those who live in the vicinity of an airport to report this harm accurately and motivate carriers to economize on that harm. At the same time, the CIC process also manages distributive effects of airport noise within a prescribed theory of just distribution.

To manage noise, one must begin with a technology for measuring noise. A good measurement technology should achieve cardinal meaningful measurements. That is, if one plane is measured to produce twice as much noise as another at a given location, then all persons in that location should experience twice as much noise from the second plane as from the first. We will assume that such a measurement technology exists. We will also
assume that it is easy to know how much noise any plane produces in any place.

To create a noise management system, one begins by dividing the total area that is affected by noise into subareas that each can be treated as experiencing a uniform level of noise for any pattern of airplane operation. One must also divide the week into time intervals in which noise has uniform costs. Let the number of subareas be \( S \), and let them be indexed by \( s \). Let \( f_s \) denote the set of individuals who are affected by noise in subarea \( s \). Let the number of time intervals be \( T \), and let them be indexed by \( t \). Let there be \( J \) carriers, indexed by \( j \).

In each subarea, \( s \), in each time interval, \( t \), there is a marginal social cost of noise:

\[
C_{st} = \sum_{j=1}^{J} f_s(Q_{jst})
\]  

(1)

where \( C_{st} \) is the cost of noise to subarea \( s \) at time \( t \) and \( f_s \) is a function that expresses the \( i \)th person’s marginal cost of noise for time interval \( t \) as a function of \( Q_{jst} \), the amount of noise in subarea \( s \) in time interval \( t \). If noise has neither economies nor diseconomies of scale, then the functions \( f_s \) will be constant functions. But these functions can also have regions over which they are increasing or decreasing.

The noise level in any subarea in any interval of time \( Q_{s} \) is the sum of the amounts of noise produced by individual carriers \( q_{jst} \).

That is,

\[
Q_{st} = \sum_{j=1}^{J} q_{jst}
\]  

(2)

Each \( q_{jst} \) is a function of the matrix of prices of noise in all times and places:

\[
q_{jst} = g_{s}(P)
\]  

(3)

where \( P \) is a matrix whose \( st \)th component is the marginal price of noise in subarea \( s \) in time interval \( t \).

The efficient pricing rule is that the carrier operating any plane must pay the incremental noise costs that are caused by its plane. That is,

\[
P_{st} = C_{st}
\]  

(4)

The practice of charging carriers the marginal costs of their noise will motivate them to economize efficiently on the production of noise. But it is also important to motivate individuals who are harmed by noise to report that harm accurately and to economize on the harm they experience. Economizing on harm can be achieved, for example, by insulating houses or, if people are particularly sensitive, by moving away from the airport area. Both the goal of motivating accurate reports and the goal of inducing efficient cost-reducing activity are accomplished by charging individuals for the net marginal costs to all others of the harms that they report. Such a net cost is the difference between the costs to carriers of reducing the amounts of noise they produce and the benefits to other individuals of the reduced amounts of noise.

What has been described so far is an example of an intricate but nevertheless standard application of the demand-revealing process. By utilizing the CIC, individuals can also be given entitlements to nonzero valuations of noise reduction. One reasonable way of doing this is to define those with entitlements as the possessors of land, with entitlements proportional to the assessed value of their land. (Landlords would have an incentive to take account of the concerns of tenants.) The entitlement to noise concern can then be defined as an entitlement to the median cost of noise per dollar of assessed value of land, in each interval of time. Each individual is then charged or credited for the concern he expresses by first computing the reduction in noise that results from the departure of his preferences from median preferences per dollar of assessed value of land. The cost of this reduction in noise to carriers (or benefit in the case of less-than-median concern) is computed, and from this the value of the noise reduction to other individuals is subtracted (or the cost of the increase in noise is subtracted from the benefit to carriers).

In the same way that individuals are given entitlements to noise concern, carriers could be given entitlements to produce noise. An initial level of entitlement could be determined as the level achieved when they have made the transition, for example, to all Stage 3 fleets. Departures from this status quo for further changes in the mix or level of operations would reconcile localized benefits with the costs to carriers expressed through a CIC process. Of course, except as a transitional accommodation, it can be argued that this would generally compromise the principle of equality and give carriers an unjustified privilege. The egalitarian application of CIC would be to give carriers entitlements to the amount of noise that the median person wishes to experience, which would be none.

REFERENCES