Chronic delays have long plagued the international air transportation system. System delays can be reduced either by expanding capacity or restricting system demand. While capacity expansion through infrastructure investment is most often called for to reduce delays, such investments in many cases can be impractical, politically infeasible or too expensive. Thus controlling, in some way, the demand placed on the system can become attractive. However, such measures bring their own set of challenges. Using recent efforts to implement slots controls in the U.S. for background and context, we identify many of the challenges that need to be overcome and questions that need to be answered in order to successfully implement airport congestion management measures. In many cases we provide answers to the questions or challenges either based on recent research or on our experience with the recent activities in the U.S.

*Key words*: air traffic management, slot control, congestion pricing, air transportation
Air transportation delays in the U.S. and around the world represent a well-known burden to society and are the subject matter of both technical and public policy debates. A recent study (Ball et al. 2010) estimated the total economic impact of air transportation delays on the U.S. economy in the year 2007 to be $28.9 billion. The most obvious and often called-for actions are investments in the expansion of system capacity either in the form of infrastructure, e.g. new runways and airports, or new capacity enhancing technology and systems. On the other hand, the fact that delays have fluctuated, sometimes substantially, with yearly demand variations suggests that controlling, in some way, the demand placed on the system could also yield (possibly substantial) reductions in delay. The U.S. National Airspace System (NAS), as well as the world wide air transportation system are queuing systems. To be sure, these are very complex queuing systems. However, they exhibit classic queuing system features in that they have capacities and as demand approaches capacity delays increase at a greater than linear rate. Thus, analyzing the “delay problem” at a very basic level, one can consider two possible solutions: increase capacity or reduce demand. While increasing capacity can be a very expensive, politically controversial, and/or technically challenging proposition, controlling demand can involve controlling the behavior of individuals or companies operating in a competitive, free market environment.

Thus, while the latter approach may superficially seem much cheaper, it can involve much higher
social and political hurdles. Not surprisingly the national approach to controlling demand can vary depending on underlying national and social norms. For example, slot controls, which restrict the level of scheduled demand at an airport, exist at virtually all large European airports but are relatively rare in the U.S.

This chapter broadly addresses certain issues related to both the justification and design of market mechanisms for airport access control. Section 11.2 both discusses some of the basic issues and reviews recent attempts in the U.S. to implement slot controls with provisions for some allocation taking place via auctions. Ultimately these proposals were not implemented. While many political forces were at work in this outcome, it can also be argued that the political discussions had not been informed by scientific research related to the economic tradeoffs underlying the implementation of slot controls. Recent research by ourselves and others have attempted to rectify this situation. In Sections 11.3 and 11.4, we review this work. In Section 11.5, we consider important design questions that arose during the deliberations in the U.S. and we provide at least partial answers to these questions.

11.2. Background

11.2.1. Airport Operations and Slot Controls

The number of operations (arrivals and departures) per hour that can be supported by an airport is limited by its system of runways. Thus, as the number of scheduled operations (demand) approaches the airport’s runway capacity, classic queuing system phenomena are exhibited. That is, when demand reaches to point when it persistently and severely exceeds capacity, delays become excessive. On the other hand, airline competitive behavior and other factors can lead to a socially sub-optimal level of operations. Specifically, classic economic phenomena, e.g. “the tragedy of the commons”, can lead to misuse of the available capacity. In this context, the airlines and other flight operators form a community of users of a common resource, airport capacity. The users gradually start to overuse that resource. In this case, the detrimental effect of overuse is high delays. However, no single user can unilaterally solve the problem: A reduction in use by one user will have minimal
impact on overall delay and, in any case, capacity freed up will quickly be used by the remaining users. Thus, while collective actions by all users would improve overall social welfare, there is no incentive for individuals to take actions that would lead to a better overall solution. This type of phenomenon is observed in many settings and, when it exists, it generally indicates the need for a coordinating policy or mechanism.

The need for, and possible forms of, policy innovation to reduce flight volumes at congested airports has been studied for several decades. Several authors have computed marginal delay costs for flights in busy periods, and found these costs to be higher than the actual charges paid for these operations (Carlin and Park 1970, Morrison 1983, Hansen 2002, Ashley and Savage 2010). This body of research provides strong economic arguments for some external controls on the use of airport capacity. At virtually all major European airports, airport slot controls have been implemented using the International Air Transport Association (IATA) administrative rules – which are adapted and complemented by EU Regulation 95/93 and its several amendments, most importantly, 793/2004. The interested reader is referred to Chapter 10 for a detailed description of the EU slot allocation practice. On the other hand, in the U.S., slot controls have historically been used very sparingly (currently slot controls exist at four U.S. airports). They were first imposed in 1969, in response to growing congestion at major U.S. airports. Under the so-called high density rule, FAA imposed hourly quotas on IFR operations at five airports, including the three commercial airports serving the New York region, plus Chicago O’Hare and Washington National. Slots were allocated by a scheduling committee composed of representatives from the airlines serving the slot controlled airports. The slot limits were seen as a temporary fix when congestion was severe and capacity expansion could not provide timely relief. The limits were continually extended, however, while allocations methods were adapted to industry developments such as airline deregulation and the emergence of new entrant carriers. They have come to be accepted as an acceptable (though not desirable) means of managing congestion at the five original high density airports, but nowhere else.
The reluctance to use slot controls in the U.S. is also evidence in the hourly limits on the number of slots. While the level of operations at European airports is set at a level consistent with poor weather conditions (IMC - instrument meteorological conditions), for those airports with slot controls in the U.S., it is set close to the good weather capacity (VMC - visual meteorological conditions). This liberal policy allows more maximum exploitation of available capacity on good days, but at the cost of severe disruptions on days with adverse weather. Thus, comparing U.S. and Europe leads immediately to two questions:

1. Should more U.S. airports have slot controls or alternatively are slot controls used too extensively in Europe?

2. How should the level of operations be set at slot controlled airports?

A recent paper, Odoni et al. (2009), sheds light on these questions by providing a detailed comparison of Newark Liberty International Airport (EWR) and Frankfurt/Main International Airport (FRA). These two airports have strong physical similarities so the comparison seeks to highlight differences that are generic to differences in the U.S. and European approaches. Evidence certainly indicates that traffic levels at EWR have been allowed to grow too large (in fact, the study considered a time period prior to the most recent imposition of slot controls). The general approach in the U.S. of having fewer slot controls at higher levels is termed “laissez-faire”. It is argued that this approach had led to excessive delays at times. On the other hand, the more conservative European approach could at times had led to under-utilization of available capacity. When considering the alternatives of setting slot levels close to VMC capacity vs IMC capacity, one must also take into account differences in climate, i.e. the percentage of time such conditions exist.

Once a decision is made to institute slot controls, one is faced with the questions of exactly how to institute such controls. In both the U.S. and Europe so-called administrative measures have been used. These have granted a form of ownership of the slots to air carriers based on their historical use of the airport. This ownership remains in place as long as those carriers use their slots with a certain minimum frequency (use-it or lose-it rules). Such rules tend to preserve airline
schedules and market shares at an airport, create entry barriers, and over time may make slot use suboptimal. In fact, one argument in favor of the relatively sparing use of slots in the U.S. is that administrative slot controls reduce competition. At the same time, it may be possible to use market mechanisms in conjunction with slot controls, e.g. a combination of slot controls and auctions, to gain the benefits of slot controls while maintaining a competitive airport environment. This approach to demand management was taken in the recently proposed U.S. rule makings for the New York area airports as discussed below.

Congestion pricing has also been proposed as means of curtailing demand at busy times. While the theory of congestion pricing seems as applicable to airports as it is to other transport facilities, in practice there are special challenges to implementing it in this setting. Most proposals have involved setting an additional per flight surcharge for operations at certain times of day. There are legal obstacles to such surcharges because federal law requires that aeronautical charges by airport operators be set on a cost-recovery basis. Conceptually, one should set congestion charges to bring the level of scheduled operations in line with capacity. Doing so requires that one estimate ex ante the relationships between the prices and the level of operations; in general this is a difficult problem. Furthermore, as charges are increased in one time window demand will move to another making the price setting problem multi-dimensional.

Congestion surcharges have been analyzed starting with Levine (1969), who argues this approach would be better than administrative slot allocation. Morrison (1983) proposes a methodology for simultaneously optimizing landing charges and investment levels. Doganis (1991) examines the airline schedule impacts from peak pricing at London Heathrow Airport, and finds that they accorded with expectations. A more recent evaluation by Schank (2005), however, finds that, in practice, peak runway pricing faces significant institutional barriers, in particular small aircraft operators who claim that such policies “discriminate” against them.

Several researchers have compared the effectiveness of pricing and slot controls. Brueckner (2009) finds that atomistic pricing, which charges each flight its marginal congestion cost even though some of that cost is borne by flights of the same airline, is less efficient than slot controls so long
as the number of slots is optimally chosen. Czerny (2010) argues that demand and congestion cost uncertainty, which may lead to a suboptimal choice for the number of slots or the congestion price, favors congestion pricing, i.e. the pricing errors that result from imperfect information are less harmful than errors in setting the number of slots. Ball et al. (2007) report on gaming simulations of congestion pricing and slot auction policies for New York LaGuardia. While the simulation indicated that both schemes are feasible, it confirmed the challenge of setting congestion prices. At the same time, congestion pricing was seen to have certain advantages, including increased carrier scheduling flexibility, and reduced incentive for airlines to hoard slots.

Traditional administrative slot controls, set at the proper level, can certainly lead to a socially optimal level of operations and congestion. However, administrative controls have the potential disadvantage of stifling competition by preserving a pre-existing market structure through grandfather rights. This disadvantage can be eliminated in concept through an effective secondary market. A secondary slot market has existed in the U.S. for many years. Analysis of this marketplace, however, has shown it to be at best, only partially effective (Fukui 2010), and, at worst, a near-failure (Berardino 2007, 2009). In particular, it is shown that nearly all major transactions have been carried out under conditions of bankruptcy or seller distress. Using an analytic model (Verhoef 2010), Verhoef considers both congestion charges and a secondary slot market. He concludes that, since airlines, particularly dominant carriers, internalize some congestion costs, congestion charges should not be set at the true marginal cost level. Further, it is shown that secondary markets must be designed with care as it is possible under broad conditions that monopoly or dominant carriers will be able to increase their dominance. Recently a secondary market for airport slots has been allowed in Europe. Various analyses have been carried out to gain an understanding of the potential impact on this new marketplace (see for example De Wit and Burghouwt (2007)).

The impact of any strategy for reducing peak period airport traffic will be mediated by individual airline scheduling decisions. The decisions are very complicated, since they involve management of expensive and highly constrained resources, including aircraft and flight crew. Nonetheless, the published schedule is widely recognized to be “the single most important product of an airline” and
is usually viewed as the starting point in a sequential process that also includes fleet assignment, maintenance routing, and crew scheduling (Barnhart and Cohn 2004). While offering a schedule that is convenient to passengers is of primary concern, competitive factors can distort this process to some degree, in a manner analogous to the Hotelling (1929) ice cream vendor problem, wherein competing vendors maximize their market shares by both locating at the center of the beach. Borenstein and Netz (1999) and Salvanes et al. (2005) find that schedules in more competitive airline markets tend to be more clustered as competing airlines seek to match one another’s schedules.

11.2.2. Recent Public Policy Initiatives in the U.S.

While slot controls at certain U.S. airports have existed since the institution of the High Density Rule (HDR) in 1969, the passage of the Wendall H. Ford Aviation Investment and Reform Act (AIR-21) in 2000 marks the starting point for recent policy making. AIR-21 called for the elimination of slot controls at New York’s John F Kennedy International Airport (JFK) and LaGuardia Airport (LGA) by January 1, 2007 and at Chicago O’Hare Airport (ORD) beginning July 1, 2002. Although AIR-21 called for the elimination of slot controls at all New York airports, the Federal Aviation Administration (FAA) anticipated the potential need to replace the HDR rules and caps with an alternative. Ball et al. (2007) describe the results of a broad research program funded by the U.S. Department of Transportation and the FAA that investigated several issues related to airport congestion management. The research, which included two multi-day strategic simulations that brought together major stakeholders from industry and government, defined and analyzed various mechanisms for limiting airport demand including the use of slot auctions.

In a proposed 2006 rulemaking (FAA 2006), the FAA sought to require airlines serving LGA to maintain a certain average gauge (seat capacity). Airlines failing to attain the average gauge standard would lose slots for their smaller-gauge flights until the standard was attained. The proposal was based on the idea that larger aircraft would allow more passengers to benefit from the limited available slots. Several airlines, as well as the operator of the New York airports (the
Port Authority of New York and New Jersey), were strongly critical of this approach, however, arguing that it was overly disruptive and prescriptive, and did not take into account airport-specific constraints.

The FAA next proposed a slot allocation policy for LGA, and soon after JFK and Newark Liberty International Airport (EWR), based primarily on grandfather rights, but with auctioning of a limited number of slots (FAA 2008d,b). In its final rule for LGA, each carrier currently holding slots would have lost 15 percent of its slots in excess of 20 (FAA 2008c). The slots would be relinquished over a five-year period, with two thirds of them auctioned and the remaining one third retired, decreasing the hourly cap from 75 to 71. Similar rules, albeit with relinquishment of 10% of slots in excess of 20 and no retirements, were set forth for JFK and EWR (FAA 2008a). These rules were challenged in court by the Air Transport Association and the Port Authority, who argued that FAA lacked legal authority to conduct slot auctions. The DC Court of Appeals issued a stay delaying the plan, and this, in combination with Congressional action caused FAA to rescind the rule in 2009. An observation from the debate that took place during this time is that the potential benefits of slot controls had not been well-quantified and had not been well-understood by the traveling public. This observation is, in part, the motivation for recent research (Swaroop et al. 2011, Le 2006, Odoni et al. 2009, Vaze and Barnhart 2011), which we review here.

While it was unable to implement the policies just described, the FAA did feel compelled to implement simple caps on the number of operation at each of the three major airports in the New York Region (FAA 2008f,e). These caps remain in effect as of 2011. Neither the setting of caps nor the allocation of slots in these episodes was based on economic analysis. The caps, for example, were not set by comparing the costs and benefits of various slot levels, and slot allocation procedures were not designed to find the highest and best uses of the slots.

11.3. The Fundamental Question: Economic Justification for Slot Controls

While in Europe, slot controls are broadly accepted, in the U.S. there is substantial public debate about the widespread necessity of slot controls and reluctance to set slot levels under the airport’s
VMC capacity. Consequently, it is important to generate compelling economic arguments for the usefulness of slot controls and also to understand the economic tradeoffs for varying the slot levels.

A basic starting point for modeling these tradeoffs is the simple observation that the principal motivation for slot control is reduction of congestion, which is generally manifested by delayed flight departures and arrivals. These are the most evident delays experienced by passengers. We refer to them as queuing delays since they result from classic queuing system behavior, with large delay increases resulting as NAS (or airport) demand approaches NAS (or airport) capacity. Thus the economic benefit of slot control is a reduction in the costs associated with queuing delays. On the cost side there are potentially multiple components involving: (a) passenger disutility; (b) operational adjustments within the airlines, e.g. aircraft and crew re-scheduling; and (c) airline’s actions in the market to combat likely competitive reactions, e.g, re-evaluating load factor and fares on re-scheduled flight.

Probably the most direct cost incurred by slot controls results from the forced reduction in the number of scheduled operations. That is, if the number of available slots in a time window is reduced, then the impacted airlines will either be forced to reduce the number of scheduled operations or move some scheduled operations to less preferred time windows. In the former case, a reduction in the number of scheduled flights implies airlines must reduce the frequency of service they offer in certain markets. Such reductions result in increasing schedule delay.

Schedule delay is a well-known phenomenon in transportation systems – first recognized and modeled as a service quality measure in the airline industry by Douglas and Miller (1974). It measures the degree to which passengers must adjust their planned departure time to accommodate the schedule offered by a transportation service. For example, if a passenger wished to depart at 9 AM but there were only flights offered at 8 AM and 10 AM, then that passenger might choose the 8 AM flight and we would say the passenger suffered one hour of schedule delay. It should be clear that as frequency of service offered in a market decreases, schedule delay increases. It is also intuitive that, in the latter case, when an airline moves a flight to a less desirable time window,
most typically the flight would be moved from a higher demand period to a lower demand period, which also increases overall schedule delay.

Figure 11.1 illustrates the basic tradeoff between queuing delay cost and schedule delay cost. The x-axis is given in units of the fraction of available capacity at which the operations level is set, e.g. by the hourly slot levels. The queuing delay cost increases rapidly as the airport capacity is approached, while the ex-ante schedule delay cost decreases as the higher service offerings increase the airport capacity utilization. In Figure 11.1, the optimal level of operations – from the passenger welfare perspective – is identified by the minimum point on the total cost curve which denotes the sum of the two cost components. If it is the case that many U.S. airports currently are “over-scheduled”, then they would operate at a point far to the right of this optimal point. Verifying this would seem to give strong support for instituting slot control at more airports and for setting the existing slot control levels lower than they are now set. We note that in general it is difficult to estimate schedule delay (and its associated cost), ex post or ex ante, as schedule delay depends on the distribution of preferred passenger departure times. In our related work, Swaroop et al. (2011), we have made progress in this direction. Specifically, we estimate the slopes of the two curves illustrated in Figure 11.1 for several airports using a combination of models. Our results do in fact support the hypothesis that many U.S. airports are over-scheduled. We now summarize these results and the underlying models.

In line with the discussion above Swaroop et al. (2011) describe two alternative models that represent the reaction of airlines to reductions in the available slots in a time window: FlightMove, which brings scheduled operations in line with a slot limitation by moving flights to a less-congested time window, and FlightTrim, which brings scheduled operations in line with a slot limitation by eliminating some flights. We perform both FlightMove and FlightTrim based on an “average” day schedule aggregated across Tuesdays, Wednesdays, and Thursdays in August 2007 for each of the 35 “Operational Evolution Partnership” (henceforth OEP35) airports. The OEP35 airports are the largest 35 airports in the U.S., these are listed in the Appendix (see Table 11.1).
Given that the *FlightMove* process involves the interactive behavior of multiple airlines and potentially many cost and revenue considerations, an exact prediction of which specific flights airlines would move is difficult. As a consequence, in the study we adopt a more agnostic approach by randomly generating a series of small perturbations to the existing schedule that eventually yields a new schedule that conforms to the slot limits. As this process is stochastic, we simulate it multiple times and look at the average results. In each simulation run, we randomly select a time period where scheduled flights exceeds the slot level, and use multinomial draws to determine the number of flights to be moved for each market in that period. Each selected flight is then randomly moved either to the neighboring time periods, or allowed to remain in the original time-period. This is repeated until the new schedule meets the slot control restrictions. In this procedure, we find that the schedule delay cost associated with each move is non-linear in the length of a move and the same flight segment could be moved multiple times over the iterations. Simply adding the cost of each move would not be accurate to represent the total *FlightMove* cost. We instead employ a linear programming transportation model that minimizes the total cost of all moves. The results can be viewed as a lower bound of the total cost given the original and new schedules, and
the most likely way in which the new schedule would be reached. The whole process is repeated multiple times. Further details of the FlightMove can be found in Swaroop et al. (2011).

In the case of FlightTrim, we adopt a successive trimming procedure. We divide one operation day (16 hours) into four 4-hour time windows. For every city-pair market and 4-hour window, we drop one “representative” flight at each step of the trimming procedure if excess demand exists. This flight is a “representative” one since it consists of portions from all markets. The portion for a given market is equal to the proportion of flights in that market over the total number of flights in the same window. Under the assumption of uniform distribution of passengers’ preferred departure (arrival) time in each 4-hour period and that flights are divisible, the associated passenger schedule delay cost increase for each market and each time period can be calculated. We repeat this until the schedule satisfies the slot restrictions. The passenger schedule delay costs are then summed over different trimming steps, markets, and time windows, to produce the total passenger schedule delay cost increment. For the purpose of preserving service to small communities, markets with less than three flights in a given time window are exempted from the FlightTrim process.

The passenger schedule delay cost increment resulting from either FlightMove or FlightTrim must then be compared with the passenger queuing delay savings that result from the schedule adjustments. To quantify these savings an econometric airport delay model is developed. To do this, we first use the FAA Aviation System Performance Metrics (ASPM) database to calculate deterministic queuing delay for each of the OEP35 airports on each day in 2007. Deterministic queuing delay is sensitive to flight schedule changes associated with different slot control policies. We then estimate an econometric model of average flight arrival delay with the deterministic queuing delay and its higher order terms as part of the explanatory variables. Other explanatory variables include the portion of time under Instrument Flight Rules (IFR) conditions and its quadratic term, wind speed, temperature, airport acceptance rate (AAR), the number of airports connected to the airport of interest, and a series of airport and monthly time dummies. The model is estimated using the Prais-Winsten procedure with panel corrected standard errors. The estimated results are consistent with the conventional wisdom, and produces good prediction of
the observed queuing delay. With this estimated model new flight arrival delays can be predicted for each new schedule generated from FlightMove or FlightTrim. Specifically, we use the new schedule to calculate the new deterministic queuing delay, and then use the econometric model to predict the corresponding new average flight delay. Comparing the new average flight delay with the predicted original average flight delay gives the average flight delay saving, which is multiplied by the number of passengers per flight and a standard value of travel time, and then summed over all flights, to yield the total passenger queuing delay cost savings. Interested readers can refer to Swaroop et al. (2011) for more details about the model and delay cost computation.

The passenger schedule delay cost increment and queuing delay cost savings are computed and compared for each OEP35 airport under three scenarios, which set slot levels at 80%, 90%, and 100% of the peak airport capacity averaged over Tuesdays, Wednesdays, Thursdays in August 2007 (this is generally close to the VMC capacity). The peak hours are defined as between 6 AM and 10 PM. The peak airport capacity is measured as the maximum declared Airport Arrival Rate (AAR) for the specified days, as reported in ASPM database of the FAA. We assume that as long as FlightMove is feasible, airlines would always prefer FlightMove to FlightTrim, because FlightMove preserves baseline demand without requiring changes in fleet mix. Four of the 35 airports justify use of FlightTrim model, as explained below. We find that the results from multiple simulation runs in FlightMove are very stable with small cost deviations of both schedule delay costs and queuing delay cost savings, suggesting the robustness of our proposed FlightMove schedule generating procedure.

We identify good candidates for slot control as those airports with positive net benefits from schedule adjustments (i.e. queuing delay cost saving - schedule delay cost increase > 0) and with daily passenger queuing delay cost saving above $10,000. This leads to 16 airports (ATL, CLE, CLT, DCA, DTW, EWR, IAD, JFK, LAX, LGA, MSP, ORD, PHL, PHX, SAN, SEA) for which slot control is recommended. Thus, we find that slot control should be far more widespread than it is presently.
Our results further reveal that the number of slots should be limited to a number that is often less than the peak airport capacity. Based on the marginal change in schedule delay cost and queuing delay cost, which in a rough sense is equivalent to the slopes of the corresponding curves in Figure 11.1, we find that the maximum net passenger benefits are achieved when slot limits are set at 100% of the airport capacity level for PHL and SAN; 90% for ATL, DCA, EWR, LGA, LAX, ORD, PHX, and SEA; and 80% for CLE, CLT, DTW, IAD, JFK, and MSP. For all airports except EWR, JFK, LGA, and ORD, the limits can be attained from FlightMove. When slot controls are imposed at 90% level, at least one of the 4-hour time windows at EWR, JFK, LGA, and ORD will encounter insufficient capacity to service the scheduled demand. As a result, FlightTrim is applied at 80% and 90% levels at these airports. For the four airports that currently have slot controls (DCA, JFK, LGA, ORD), our analysis suggests the current slot levels are set too high: a slot limit set a 90% of capacity would be most beneficial for DCA, EWR, and LGA; while at JFK the limit should be 80% of capacity.

It is perhaps surprising to have airports, such as CLE, MSP, SAN, and SEA, which are normally not considered highly congested, on the list of logical candidates for slot control. This occurs because of pronounced peaks at these airports which could be reduced by spreading flights to less congested periods when slot control is imposed. This is illustrated in Figure 11.2, which shows schedules for some highly, mildly, and least congested airports.

As a first-order estimate of annual benefits, we conclude that implementing the best slot control policies at the aforementioned 16 airports would reduce passenger delay cost by $0.8 billion in 2007. The estimate mainly captures reduction in delay that is within the control of the National Airspace System, or NAS delay. Two other types of delay, air carrier delay and aircraft propagated delay, however, are not considered in our analysis. It is evident that air carrier delay, a result of airline internal problems such as aircraft maintenance, is beyond the control of local airports and the FAA. Quantifying aircraft propagated delay would require a different approach which treats delay in a holistic manner. Given an estimate of $4.7 billion for total passenger delay cost in the U.S. in 2007 (Ball et al. 2010) and the fact that the three types of delay account for roughly equal
Figure 11.2  Aug 2007 Aggregated Arrival Schedules for several airports. Congestion levels decrease from top to bottom. The dotted line shows the peak arrival capacity (AAR).

shares of the total number of delayed flights, our results suggest that passenger costs attributable to NAS delays could be reduced up to 50% if slot controls are more pervasively and aggressively implemented at the U.S. airports. We further note that any propagated delay will have as its root cause either NAS delay or air carrier delay. One can then expect that slot control would reduce substantially the (roughly 50%) portion of propagated delay whose root cause is NAS delay. Further since propagated delays grows at a greater than linear rate with its root cause delay one could expect that a 50% reduction in NAS delay would lead to a greater than 50% reduction in the associated propagated delays. Ball et al. (2010) estimate the total direct cost of delay to be $28.9 billion. The analysis we have summarized here involved models that considered passenger delay costs due to delayed flights. Not included were other cost components such as reduction in delays due to missed passenger connections or canceled flights, reduction in the costs airlines incur when delays are reduced as well as other factors. It is clear that slot control would reduce all such costs
as well, although our analysis was not designed to provide a precise estimate of the degree of this reduction.

Implementing slot controls at several airports simultaneously may be a high-risk endeavor. Practical issues like equitable allocation among the airlines of the exact flights to be reduced from the congested time-slots; settlement time for the new schedules to take effect; training of various personnel in the airline industry, airports, and the FAA; adaptation of IT systems and services etc would need to be handled in a careful manner so as to not disrupt the passenger service. Indeed, the entire exercise may seem too daunting to undertake despite the economic benefits, although the same might be said for the alternative of capacity expansion in the case of many airports. Interestingly, the list of airports identified by our research as suitable for slot controls is diverse in many respects: it includes airports spanning the entire geography; of small, medium, and larger capacities; from mildly to highly congested regions; and has recommendations at various slot levels. This could prove useful in de-risking the entire initiative, by phasing the implementation at carefully selected lower-risk pilot airports first. The implementation may be conducted at the pilot airports, and the benefits – as well as challenges – established before taking upon the other airports.

11.4. Other Implications of Slot Controls

While our analysis primarily focuses on passenger schedule and queuing delay costs and benefits, slot control also incurs other consequences, including changes in carrier profitability, load factor, air fare, and aircraft size. Research by Vaze and Barnhart (2011), and Le (2006) attempt to understand the impact of slot controls on these variables, focusing on New York LaGuardia (LGA). These studies, at least implicitly, consider the queuing delay – schedule delay tradeoff as well as other impacts and tradeoffs. Both studies attempt to answer the question, “What would happen if more restrictive slot controls were put in place at LGA?” They consider changes in service frequency but also impact on airline costs and profits.

In Vaze and Barnhart (2011), a game theoretical model based on an S-curve relationship between airlines’ frequency share and market share is developed that explicitly characterizes frequency
competition among airlines servicing LGA. They use the Nash equilibrium concept to predict the outcome of competitive situations under different slot reduction levels at LGA. Two schemes for allocating slots among airlines are considered. The first, Proportionate Allocation Scheme (PAS), distributes slots among different carriers by the same ratio as in the status quo. The PAS is likely to be considered more acceptable by major carriers, but may ignore how efficiently airlines utilize slots. In fact, airlines may differ – often substantially – in the number of passengers carried per flight or per slot. Recognizing this, the authors propose a Reward-based Allocation Scheme (RAS), under which the number of slots allocated to each airline is proportional to the total number of passengers carried by that airline. The RAS provides airlines with incentives to carry more passengers per slot, through either higher load factors or larger planes.

A wide range of slot reduction scenarios at LGA is considered in Vaze and Barnhart (2011). Assuming constant aircraft size, slot reduction reduces flight operations and forces airlines to increase load factor to accommodate as many passengers as possible. This results in a profit increase but only to a certain extent, beyond which profit starts to decline, because after achieving the maximum allowed load factor, loss of flight operations also means loss of revenue. At low slot reduction levels, airline profit gain and passenger loss bear similar changing patterns under PAS and RAS. When slot control is more substantial, RAS gives higher profit and less passenger loss than PAS.

The authors then focus on a more realistic scenario which reduces slots by 12.3% – roughly corresponding to scheduling at IMC (“bad-weather”) capacity instead of VMC (currently practiced “good-weather”) capacity at LGA. Consistent with our work, the authors conclude that this would result in a significant reduction in flight and passenger delays, but a very small schedule delay increase. In addition, a small reduction in total allocated capacity can considerably improve operating profit of all incumbent carriers. By allowing aircraft to partially upgauge, the loss of passengers can be reduced significantly. The authors also observe that while profit increase may be different for each individual airline under PAS and RAS, the two schemes produce similar aggregate impacts.
Two of the assumptions made in Vaze and Barnhart (2011) are that air fare and aircraft size would remain unchanged irrespective of slot control policies (although the latter is partially relaxed in their sensitivity analysis). The changes in these two factors are explicitly investigated in Le (2006) using a different approach. In her thesis, Le points out that the causes of airport congestion include (i) the HDR with grandfather rights that allocates limited slots to incumbent airlines; (ii) no incentives for airlines to use larger aircraft due to weight-based landing fees; (iii) slot exemptions granted to small markets served by 70-seat or less aircraft; (iv) the 80% use-it-or-lose-it requirement which forces airlines to fly low load-factor flights. Instead of explicitly modeling airline competition, Le (2006) assumes a single “benevolent” airline that reacts to price elasticities of demand in competitive markets. She develops an optimization model that solves both an airline scheduling sub-problem and an airport allocation problem, and a stochastic queuing network simulation model to quantify the delay effect. Taking profit maximization by the benevolent airline and seat maximization by the government as two separate objectives, Le demonstrates the existence of profitable flight schedules at LGA that is able to accommodate passenger demand while substantially reducing flight delays.

Several metrics, including air fare, aircraft size, flight delay, and total number of seats and markets serviced, are examined by Le under multiple scenarios. She finds that, if profit maximization is the only goal, seat throughput will be significantly lower than under the throughput-maximizing scenario. Carriers tend to consolidate flights, increase aircraft size and fare, resulting in fewer passengers transported and lower airport delays given the same slot control level. Since the goal of profit seeking conflicts with that of maximizing enplanement, Le (2006) further examines two compromise scenarios by setting 80% and 90% of the unconstrained maximum profit as lower bounds on the profit of the benevolent airline, and considers such scenarios to be (i) close enough to the baseline to provide a feasible transition solution; (ii) reasonably close to the optimal profit curve. Compared to the baseline, the compromise scenarios predict positive changes in total number of seats, aircraft size (assuming constant load factor), and negative changes in average fare, flight
traffic, and substantial reduction in flight delay. She also looks at the number of profitable markets on a daily schedule, and finds no penalty in the number of markets with slot allocation at 8 ops/runway/15-min compared to 10 ops/runway/15-min. However, with tighter slot limits below 8 ops/runway/15-min, an increasing number of markets will become unprofitable. She concludes that having aggregate airline schedules at 8 ops/runway/15-min, which is the current IMC operation rate, would significantly reduce the congestion problem at LGA, increasing the predictability of air transportation and improve the quality of service expected by the flying public.

Despite the above efforts in modeling cost and benefits from slot controls, there remain some gaps in the economic justification for slot controls that warrant further research. One is fare and competition effects. Although fare change is captured in an aggregate manner in the preceding study, modeling of the impact of slot control on fare in an explicit competitive environment is rarely seen. Conceptually, if operations are restricted in some way, then resource scarcity may result leading to higher fares. Further, to the extent that such restrictions allow one or more air carriers to increase market power, this could move fares even higher. These effects are often cited as a major detriment of market-based airport access controls. Among other challenges in addressing this question is that the degree to which there is an anti-competitive effect depends very much on how controls are implemented. For example, administrative slot controls that are based primarily on grandfather rights, would tend to preserve existing market structure and would restrict new entrants from entering the markets served by the airport. Mechanisms that allowed for some slot reallocation, e.g. via auctions, would support a more vibrant competitive environment and lower fares.

An often cited concern related to the imposition of slot control is loss of service to small communities or markets. From the schedule delay perspective, loss of service represents the extreme case for increase in schedule delay (frequency goes to 0). In our work we deliberately eschew removing flights in markets that are very sparsely serviced. On the other hand, as mentioned in Le (2006), preserving small community service represents one cause for inefficient use of slots and airport
congestion. Her results also reveal profitability concerns associated with small community markets. All these make small community access an important issue in the design of slot controls.

11.5. Design Issues for Slot Controls

11.5.1. Getting the Slot Level Right

Runway systems of airports have capacities and these capacities are in fact what lead to the queuing effect described previously and the need for access controls. However, the capacity of a runway system depends on many factors which can vary substantially day to day or even hour to hour. The most obvious variation is caused by changes in weather conditions. A very basic division in airport weather conditions is VMC vs IMC as discussed earlier. The degree to which capacities vary between VMC and IMC depends on airport specifics; an extreme case is San Francisco International Airport (SFO), where IMC arrival capacity is approximately one half VMC arrival capacity.

These considerations imply that careful modeling, such as what was discussed in Section 11.3, should be employed to determine the appropriate slot level; for example, it could easily be the case that the “optimal” slot level would allow some delays during poor weather days Conversely, as shown in Section 11.3, the appropriate slot level is usually somewhat less than the airport capacity under ideal conditions. In general, the slot level should depend on the frequency of poor weather conditions. Thus, the best solution should depend on both the VMC and IMC capacity values and the relative frequency with which such conditions exist. These factors were implicitly considered in the analysis summarized in Section 11.3, since the delay arising from a given schedule is estimated over an entire month, with varying weather conditions and capacity levels. If slot levels are to be treated as essentially constant over many months or years, however, that analysis would need to be extended to cover a commensurate time period. On the other hand, there may also be benefit to varying the slot level by time of day, e.g. by allowing more slots during high demand/high value periods and compensating with “cooling off” periods. (Churchill et al. 2011) describes optimization models for setting slot levels that may vary by time of day taking into account capacity scenarios and their likelihood and also variations in slot value by time of day.
It is also the case, however, that modeling challenges could be insignificant compared to political challenges related to setting appropriate slot levels. The key question is who has the authority and what criteria are used to do so. Since changes in the level of operations could have a substantial economic impact on various air carriers as well as the traveling public, the person or group who sets slot levels should be insulated from political pressures.

11.5.2. Small Community Access

Smaller communities can derive substantial value from regular service to large airports. Many of the FAA proposals for allocating slots have included specific features to protect service to small communities, and those opposing these proposals have routinely cited their adverse impact on small communities (FAA (2006) – see p. 10; Port Authority of New York and New Jersey (2008) – see p. 4). Thus, politically acceptable airport demand management schemes may have to demonstrate their ability to insure adequate access to designated small communities. Both the review of recent public policy initiatives in Section 11.2.2 and the research reviewed in Section 11.3 highlight the importance of this issue.

11.5.3. Where Does the Money Go?

Market-based approaches may generate substantial new revenues. Obviously, how this money is spent will be of great interest to the various stake-holders. Probably the least desirable outcome from the perspective of major stake-holders such as air carriers and airport operators is that the new money goes into the general fund of the Federal Government. Legal assurances that such funds are used to enhance aviation infrastructure certainly would enhance the acceptability of any proposal. Most desirable would be that such funds are used to benefit operations at or around the airport in question. Other desirable features could be that such funds replace existing user fees or taxes such as landing fees.

In its proposal to auction a limited number of slots at LaGuardia, FAA identified two options for spending the proceeds. The first was to use them to "mitigate congestion and delay in the
New York area,” while the second was to give them to the carrier holding the slot that was being auctioned (FAA 2008d, see p. 1).

11.5.4. Federal vs Local Control

One might naturally assume that airport access controls would be implemented by the airport operator and that any associated revenues would go to the airport operator. However, there are many reasons this might not be the case. Airports are natural monopolies and therefore are generally highly regulated. Their revenues are almost always restricted so that only cost recovery is allowed. Thus, market based approaches that may generate revenues of arbitrary size would almost by necessity have to be implemented by another entity – most likely the Federal Government. In the U.S., while the airport operator controls the surface of the airport, the FAA has legal authority to control the airspace and, thus, there are strong arguments that it has right to control airport access using various means including market mechanisms (the extent of this right is subject to legal debate however). Clearly, there is a potential tension related to the question of Federal or local control and this must be managed in devising any access control solution. Taking an objective perspective, the FAA has responsibility for the efficient operation of the entire National Airspace System (NAS) and may need to control access to individual airports with a national perspective in mind. The airport operators has knowledge of its airport’s characteristics and can implement and fine tune local controls. Thus, an ideal solution should take both of these perspectives into account and design an appropriate solution.

This is easier said than done, however. For example, despite consultation between the FAA and the Port Authority of New York and New Jersey in the development of the FAA’s proposal to auction slots a LaGuardia, the Port Authority asserted flight operations using slots obtained through an auction “shall not be conducted” and that it would not consent to the leasing or use of terminal space for flights that used such slots (Port Authority of New York and New Jersey 2008, see p. 5).
11.5.5. Who Can Own Slots?

Assuming the use of a slot-based system one may be faced with the question of who can own slots. In fact, although market mechanisms have never been used to perform a primary allocation of slots from the Federal Government to the airports, a secondary market for slots does exist in the U.S. Not surprisingly slots can have substantial economic value. In fact, certain banks, as a result of airline bankruptcy proceedings, have become the owners of slots. This suggests a more general question of who should be allowed to own slots. For example, if a slot auction were held, should bidders be restricted in any way. It could be that “brokers” may wish to bid on slots and then resell them or lease them over short periods of time. There could in fact be economic value in a market for short term slot leases. Another scenario could involve a local community buying slots so as to insure access between that community and the major airport in question. Thus, there could be sound societal reasons for allowing non-air-carriers to own slots but, on the other hand, there is the potential for various unknown (and unintended) consequences, such as airport opponents purchasing the slots in order to retire them.

11.5.6. International Bilateral Agreements

Scheduled air transportation service between a pair of countries can only be conducted when authorized under a formal bilateral agreement between the countries. Such bilateral agreements typically grant access to an international city pair market to designated air carriers. Slot control regulations must be constructed in a way so as not to violate such agreements.

11.5.7. Infrastructure Investment Incentives

A well-structured air transportation system should provide an appropriate signaling mechanism to indicate when investment in new infrastructure is required. In a system devoid of slot controls or congestion pricing, the typical signal is the presence of (possibly extreme) delays. The implementation of slot controls could eliminate this signal and not replace it at all or replace it with a different (possibly better) signal, e.g. high slot prices. To the extent that recurring fees, e.g. from congestion
pricing, depend on scarce capacity, an incentive is provided to the collectors of the revenue to not invest in new capacity. Obviously, such a situation would be undesirable and should be avoided. Appropriate incentives for infrastructure investment should be an important consideration in the design of any congestion management approach.

11.6. Conclusions

This chapter provided an overview of models that provide economic justification of slot controls and also discussed a number of design issues. We summarized our own research, which considered the tradeoff between queuing delay and schedule delay. This tradeoff is fundamental to determining the need for slot control and the optimal slot control level and our results provide strong justification for the more extensive use of slot control in the U.S. Other related research, specifically focused on LGA airport, supports these results and further considers other issues such as the impact on airline profits and overall passenger performance. More aggressive use of slot control at LGA is shown to substantially improve overall social welfare without disadvantaging any major constituency.

Experience and feedback from recent U.S. Federal Government proposed rulemakings was used to identify various political and design challenges related to the implementation of slot control in the U.S. We provide perspectives on these challenges and, in some cases, give approaches to their resolution.

References


Appendix. U.S. Operational Evolution Partnership (OEP) 35 Airports
<table>
<thead>
<tr>
<th>Airport code</th>
<th>Airport Name</th>
<th>City</th>
<th>State</th>
</tr>
</thead>
<tbody>
<tr>
<td>ATL</td>
<td>Atlanta Hartsfield Intl</td>
<td>Atlanta</td>
<td>GA</td>
</tr>
<tr>
<td>BOS</td>
<td>Boston Logan Intl</td>
<td>Boston</td>
<td>MA</td>
</tr>
<tr>
<td>BWI</td>
<td>Baltimore-Washington Intl</td>
<td>Baltimore</td>
<td>MD</td>
</tr>
<tr>
<td>CLE</td>
<td>Cleveland Hopkins Intl</td>
<td>Cleveland</td>
<td>OH</td>
</tr>
<tr>
<td>CLT</td>
<td>Charlotte Douglas Intl</td>
<td>Charlotte</td>
<td>NC</td>
</tr>
<tr>
<td>CVG</td>
<td>Cincinnati-Northern Kentucky Intl</td>
<td>Covington-Cincinnati, OH</td>
<td>KY</td>
</tr>
<tr>
<td>DCA</td>
<td>Washington Reagan Natl</td>
<td>Washington</td>
<td>DC</td>
</tr>
<tr>
<td>DEN</td>
<td>Denver Intl</td>
<td>Denver</td>
<td>CO</td>
</tr>
<tr>
<td>DFW</td>
<td>Dallas-Ft Worth Intl</td>
<td>Dallas-Ft Worth</td>
<td>TX</td>
</tr>
<tr>
<td>DTW</td>
<td>Detroit Metropolitan Wayne County</td>
<td>Detroit</td>
<td>MI</td>
</tr>
<tr>
<td>EWR</td>
<td>Newark Intl</td>
<td>Newark</td>
<td>NJ</td>
</tr>
<tr>
<td>FLL</td>
<td>Ft Lauderdale-Hollywood Intl</td>
<td>Ft Lauderdale</td>
<td>FL</td>
</tr>
<tr>
<td>HNL</td>
<td>Honolulu Intl</td>
<td>Honolulu</td>
<td>HI</td>
</tr>
<tr>
<td>IAD</td>
<td>Washington Dulles Intl</td>
<td>Washington</td>
<td>DC</td>
</tr>
<tr>
<td>IAH</td>
<td>George Bush Intercontinental</td>
<td>Houston</td>
<td>TX</td>
</tr>
<tr>
<td>JFK</td>
<td>John F Kennedy Intl</td>
<td>New York</td>
<td>NY</td>
</tr>
<tr>
<td>LAS</td>
<td>Las Vegas McCarran Intl</td>
<td>Las Vegas</td>
<td>NV</td>
</tr>
<tr>
<td>LAX</td>
<td>Los Angeles Intl</td>
<td>Los Angeles</td>
<td>CA</td>
</tr>
<tr>
<td>LGA</td>
<td>La Guardia</td>
<td>New York</td>
<td>NY</td>
</tr>
<tr>
<td>MCO</td>
<td>Orlando Intl</td>
<td>Orlando</td>
<td>FL</td>
</tr>
<tr>
<td>MDW</td>
<td>Chicago Midway</td>
<td>Chicago</td>
<td>IL</td>
</tr>
<tr>
<td>MEM</td>
<td>Memphis Intl</td>
<td>Memphis</td>
<td>TN</td>
</tr>
<tr>
<td>MIA</td>
<td>Miami Intl</td>
<td>Miami</td>
<td>FL</td>
</tr>
<tr>
<td>MSP</td>
<td>Minneapolis-St Paul Intl</td>
<td>Minneapolis</td>
<td>MN</td>
</tr>
<tr>
<td>ORD</td>
<td>Chicago O’Hare Intl</td>
<td>Chicago</td>
<td>IL</td>
</tr>
<tr>
<td>PDX</td>
<td>Portland Intl</td>
<td>Portland</td>
<td>OR</td>
</tr>
<tr>
<td>PHL</td>
<td>Philadelphia Intl</td>
<td>Philadelphia</td>
<td>PA</td>
</tr>
<tr>
<td>PHX</td>
<td>Phoenix Sky Harbor Intl</td>
<td>Phoenix</td>
<td>AZ</td>
</tr>
<tr>
<td>PIT</td>
<td>Pittsburgh Intl</td>
<td>Pittsburgh</td>
<td>PA</td>
</tr>
<tr>
<td>SAN</td>
<td>San Diego Intl-Lindburgh Field</td>
<td>San Diego</td>
<td>CA</td>
</tr>
<tr>
<td>SEA</td>
<td>Seattle-Tacoma Intl</td>
<td>Seattle</td>
<td>WA</td>
</tr>
<tr>
<td>SFO</td>
<td>San Francisco Intl</td>
<td>San Francisco</td>
<td>CA</td>
</tr>
<tr>
<td>SLC</td>
<td>Salt Lake City Intl</td>
<td>Salt Lake City</td>
<td>UT</td>
</tr>
<tr>
<td>STL</td>
<td>Lambert-St Louis Intl</td>
<td>St Louis</td>
<td>MO</td>
</tr>
<tr>
<td>TPA</td>
<td>Tampa Intl</td>
<td>Tampa</td>
<td>FL</td>
</tr>
</tbody>
</table>

**Table 11.1** List of the U.S. Operational Evolution Partnership (OEP) 35 Airports