Economics of Transportation ∎ (∎∎∎∎) ∎∎∎–∎∎∎

Contents lists available at ScienceDirect

Economics of Transportation

journal homepage: www.elsevier.com/locate/ecotra

The untolled problems with airport slot constraints

Joseph I Daniel *

Department of Economics, University of Delaware, Newark, DE 19716, USA

ARTICLE INFO

Article history:
Received 13 June 2013
Received in revised form
24 January 2014
Accepted 25 January 2014

Keywords:
Airport congestion
Slot constraints
Pricing
Bottleneck
Queuing

ABSTRACT

This paper examines the efficiency and practicality of airport slot constraints using a deterministic bottleneck model of landing and takeoff queues. It adapts this congestion pricing model to determine the optimal timing and quantity of slot permits for any number of slot windows. Aircraft choose their optimal operating times subject to the slot constraints, and airport queues adjust endogenously. The number and length of slot windows affects the congestion levels and efficiency gains. The atomistic bottleneck model is extended to include self-internalizing dominant traffic and atomistic fringe traffic. The model raises questions about the implementation of slot constraints that do not arise in standard congestion models. The theory explains (Daniel's, 2011) empirical findings that slot-constraints at Toronto are ineffective and suggests that recent proposals for slot constraints at US airports would be similarly ineffective. Effective slot constraints require many narrow slot windows, making slot auctions or markets difficult to implement.

1. Introduction

1.1. The issue

Slot constraints are restrictions on the number of landings or takeoffs that airports permit during specified time periods known as slot windows. For a typical example, an airport that can perform sixty landings per hour, issues sixty slot permits for landings during each hour of the day and requires aircraft to operate during their assigned hour. Most major commercial airports throughout the world impose slot constraints to control access to their runways, ostensibly to reduce congestion delays. In the US, however, most major airports that receive federal funding are available on a first-come, first-served basis without requiring slot permits.[1] When airports become severely congested, airport authorities, airline officials, and policy makers generally seem to favor slot constraints over congestion tolls as a means of managing demand. They argue that slot constraints are simpler to implement because the airports need only limit the number of slot permits to the airport's capacity, then sell, auction, or give the permits away. Slot markets can price and allocate the slot permits efficiently. Congestion tolls, they argue, are too difficult (or politically inconvenient) for airports to assess correctly. Moreover, slot permits are supposed to avoid the problem of imposing different toll schedules on dominant airlines that already internalize their self-imposed delays than on fringe airlines that ignore the additional delays they impose on other aircraft.

This case for slot constraints implicitly assumes that airports experience steady-state traffic during slot windows. Actual traffic patterns at major airports, however, exhibit rapid fluctuations that follow regular patterns due to airline scheduling practices. Airlines that operate hub-and-spoke networks schedule flights around passenger interchange periods. Consequently, many airports experience as many as ten substantial peaks during a day (see, Daniel and Harback, 2009). Actual traffic patterns call into question the effectiveness and practicality of slot-constraint systems that are based on steady-state traffic models that underlie most existing policy analysis. If traffic rates and queuing systems are in steady states, then restricting the number of hourly slot permits to the hourly airport capacity might reduce congestion. If congestion is actually caused by traffic rates fluctuating from slack to peak demand within hourly periods, then it is desirable to use a model that captures these features of the problem.

In this paper, I develop a (dynamic) bottleneck model with multiple slot windows in which the airport authority chooses the timing of slot windows and the quantity of landings or takeoffs (operations) to permit during each window. Airlines choose when to operate their aircraft within the slot windows to minimize the costs of their queuing delays and of arriving before or after their most preferred time. A structural model of congestion has state-contingent queues that evolve endogenously in response to traffic adjustments. Congestion

* Tel.: 302 831 1913; fax: 302 831 6968.
E-mail address: danielj@udel.edu

[1] Airports assess landing fees based on aircraft weight, typically between one to five dollars per thousand pounds. These fees have nothing to do with the marginal social cost of serving the aircraft. The social cost consists primarily of the delays aircraft impose externally on other aircraft. Weight-based landing fees encourage too many operations by smaller aircraft compared to marginal-cost congestion tolls that promote efficient levels of operations by internalizing the external delays.

http://dx.doi.org/10.1016/j.ecotra.2014.01.003
2212-0122/

externalities cause traffic rates (i.e., the rates of arrivals at the landing or take off queue) to exceed service rates (i.e., the rates at which airports can perform landings or takeoffs) during a portion of each slot window, even when the number of slot permits per slot window is within the airport's capacity. While airports chose the number of slot permits per slot window, airlines chose aircraft operating times within the slot windows, resulting in equilibrium traffic patterns that exceed capacity during the portions of slot windows that are closest to the preferred operating times. In the bottleneck model, slot constraints cause the queue to empty periodically by preventing airlines from scheduling aircraft too early. This results in smaller peaks during each slot window rather than a single large queue that persists throughout the busy period. As the airport authority adds more slot windows, it constrains aircraft to narrower operating windows and the queue empties more frequently, thus limiting the accumulation of aircraft in the queue and reducing total queuing delay.

The model also addresses the optimal timing of dominant and fringe airline operations, the efficient allocation of slots among dominant and fringe airlines, and the effect of slot constraints on the distribution of surpluses between dominant airlines and atomistic fringe aircraft. A policy section discusses the problems with current and proposed implementations of slot constraints, and the practical issues involved in designing and implementing efficient slot systems. The paper focuses on the more basic issues of timing and quantity of slot permits rather than how to design auctions or markets to distribute slot permits. The literature has largely overlooked these basic issues, but unless policymakers address them, the resulting slot-constraint system may have little effect on congestion at most airports, no matter how elaborately they design slot auctions or markets.

A brief preview of the conclusions I derive from the model is as follows: (1) Effective slot-constraint systems require numerous narrow slot windows that force traffic to spread out over the peak period. Slot-constraint systems that hold the quantity of slot permits to the airport capacity over a single slot widow covering the entire peak period are completely ineffective. (2) In unconstrained equilibria, dominant airlines schedule some of their aircraft to operate at the service rate during the periods just before and after the atomistic traffic. These aircraft fully internalize their delays, while the remaining dominant aircraft join atomistic traffic and ignore the delays that they impose on other dominant aircraft. The fraction of internalizing aircraft varies from one to zero as fringe demand elasticity varies from zero to negative infinity. (3) If the airport authority has complete control over the allocation of slots, it will separate the dominant and fringe operations to enable the dominant airline to fully internalize all self-imposed delays. If the airport is unable to enforce this separation, then the dominant airline will schedule some aircraft atomistically, depending on the elasticity of fringe demand. The first-best optimum requires one slot permit and window for every service interval.

1.2. Background

The International Air Transit Association's (IATA) *Worldwide Slot Guidelines* (*WSG*) provide for slot coordinators at highly congested airports who issue and distribute slot permits on a semi-annual basis. According to the guidelines, the fundamental considerations in allocating slots are preserving historical patterns of use, preventing “confiscation” of incumbent airlines' claims on slots, and allocating slots to new entrants only from new airport capacity. Twice a year, airline representatives submit slot requests that substantiate their past slot usage, and airport slot coordinators make preliminary slot distributions, then airline and airport representatives meet at three-day slot conferences to trade and finalize slot allocations. Slots may be traded in one-for-one exchanges, but may not be sold. The airline industry appears to have designed this slot system as a means of restricting entry.

Recent slot-constraint proposals for the US, however, seek to preserve free entry by providing for issuing of slot permits that would be valid for a period of ten years. Each year, one tenth of an airport's slots would expire and be re-issued. Airlines could resell slots, and presumably markets would develop for trading slots. There are various proposals concerning the type of initial auctions or markets to distribute newly issued slots. Since the values of slots depend on their combination with other slots (including those at other airports), these auctions or markets must be able to value an enormous number of potential slot combinations. Combinatorial auctions to distribute slots would potentially be the largest and most complex ever conducted. The fundamental justification for slot auctions or markets is that they are supposedly simpler than congestion tolling because an auction or market determines the price of slots rather than an administrative agency. To make combinatorial slot auctions or slot markets feasible would require issuing many undifferentiated slot permits with long slot windows to reduce the number of potential slot combinations. Since airlines are free to choose when to operate within the slot windows, however, longer windows with more undifferentiated permits reduce the airport's control over the timing of traffic and lengths of queues. There is a fundamental tradeoff between feasibility of slot auctions or markets and reduction of congestion. Proponents of slot-constraints largely ignore this issue because standard steady-state congestion models implicitly assume that traffic rates are constant within slot windows. Consequently, hourly slot windows appear to be adequate under steady state models.

Any optimal system of congestion management must address two fundamental issues: setting the traffic quantity so that the marginal social cost of operations equals their marginal social benefit; and controlling the timing of traffic to minimize total social cost. The standard congestion models and most real-world slot constraint systems attempt to address the first issue but not the second. Slot-constraint systems proposed in the US, and the systems following the *WSG* all grant authority to operate during specific time windows – usually one-hour intervals. These wide time windows provide flexibility to accommodate random variation in operating times, facilitate exchange of slots, and reduce administrative and compliance costs, but they also make it impossible to achieve optimal traffic patterns and can render slot constraints largely ineffective. For example, Daniel (2011) shows that Toronto's Pearson International Airport experiences significant congestion in spite of its slot-constraint system that follows the *WSG*.

The way airports collect and present traffic data reinforces the standard models' erroneous assumption of constant traffic rates within slot windows. Government and consulting reports typically aggregate airport traffic rates by hour. This practice hides the rapid fluctuations in traffic rates over shorter periods of time that are responsible for a significant amount of airport delays. Airports do not routinely collect data on the actual time spent in landing or takeoff queues. Instead, the primary measures of aircraft delays are on-time arrival and departure statistics. These report the number of aircraft that are more than 15 min late relative to the aircraft's scheduled operating time at the gate. Aside from the obvious issue of not recording delays of less than 15 min, on-time operating statistics do not reflect the additional time that airlines add to their scheduled travel times to allow for queuing delays on takeoff or landing, for traffic jams on the tarmac while trying to access the gates, or any of the other many regular causes of aircraft delay. Atlanta and Chicago O'Hare, for example, average approximately 20 min of queuing delay per aircraft, which is much higher than other US airports, while their on-time arrival statistics are no worse than most other major airports. The hourly data paints a

Please cite this article as: Daniel, J.I, The untolled problems with airport slot constraints. Economics of Transportation (2014), http://dx.doi.org/10.1016/j.ecotra.2014.01.003