Transportation Research Part E 70 (2014) 245–260

Contents lists available at ScienceDirect

Transportation Research Part E

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

New York, New York: Two ways of estimating the delay impact of New York airports

Lu Hao [a,*], Mark Hansen [a], Yu Zhang [b], Joseph Post [c]

[a] University of California, Berkeley, Berkeley, CA 94720, United States
[b] University of South Florida, Tampa, FL 33620, United States
[c] Federal Aviation Administration, Washington, DC 20591, United States

ARTICLE INFO

Article history:
Received 22 April 2013
Received in revised form 14 July 2014
Accepted 15 July 2014

Keywords:
New York airports
Delay impact
2SLS regression
Simulation model
Models comparison and validation

ABSTRACT

High arrival delay at major airports tends to propagate and generate secondary delay through the National Airspace System (NAS). In the United States, it is widely believed that the major culprits for delay throughout the NAS are the three New York commercial airports – Newark (EWR), LaGuardia (LGA), and John F. Kennedy (JFK). Various estimates of the extent to which the New York airports impact the delay in the NAS have been reported over the years. Yet there is no thorough investigation into the mutual relationship between delays at New York and non-New York airports. In this paper, we take two different approaches to quantify the impact of the three New York airports on delay throughout the NAS. First, we estimate and apply an econometric model using a large historical data-set. The other model is the FAA SWAC model that simulates flights and tracks the daily performance of the system. The counterfactual scenarios in these two models are adjusted to be comparable to each other. There is disparity between the results of the two different models, suggesting the simulation model might not capture all the factors that cause arrival delay. Still both results conclude that the portion of delay in the system caused by New York airports is much less than publicized estimates. Combining econometric and simulation models to address questions of this nature appears to be a promising approach.

1. Introduction

It is widely understood that congestion and delay in transport networks can spread from one part of the network to another. In the case of air transport, this means that airport and terminal airspace congestion in one region can cause delays at quite distant locales. This phenomenon, sometimes known as delay propagation, takes several forms. Downstream propagation occurs because flights involve aircraft, flight crew, cabin crew and payload that are bound for subsequent flights. This creates dependencies between inbound and outbound flights so that delays at a given airport propagate forward to airports downstream. Upstream propagation occurs when flights bound for a given airport are slowed, creating a "backup" that affects traffic bound for other airports, or even delaying departing flights as they await gaps in the en route streams. Various factors contribute to delay propagation, including the hub status of an airport or an airline (Santos and Robin, 2011), slot restrictions (Swaroop et al., 2012), and runway or airspace capacities. Airport delay propagation in the National Airspace System (NAS) has become a widespread interest to analysts and policy makers in the United States.

* Corresponding author. Tel.: +1 510 517 0046.
E-mail addresses: haolu@berkeley.edu (L. Hao), mhansen@ce.berkeley.edu (M. Hansen), yuzhang@usf.edu (Y. Zhang), joseph.post@faa.gov (J. Post).

http://dx.doi.org/10.1016/j.tre.2014.07.004
1366-5545/

High flight volume, coupled with dense airspace and limited capacity, have made the three New York commercial airports—Newark (EWR), LaGuardia (LGA), and John F. Kennedy (JFK)—among the most delay-prone airports in the NAS (FAA, 2010). The poor on-time performance of the three New York airports is well-recognized and is widely believed to be an important source of delay throughout the NAS. Thus the air traffic congestion at New York airports is believed to be not merely a regional issue, but rather a national one. However, there is a lack of solid research regarding the actual extent of the delay propagation impact from New York airports.

While it is often asserted that New York delays contribute heavily to delays elsewhere, FAA has not systematically examined this effect, and researchers in academia and the aviation industry have provided, at best, only limited information on the contribution of New York airports to delay in the NAS as a whole. At a November 2007 White House press conference, then-Secretary Mary Peters stated that "…three-quarters of the flight delays are because the plane went into, out of, or through the New York airspace…" This figure, however, addresses only chronically delayed flights, a small subset of all flight delays (FAA, 2010). The same FAA report also notes that "In April 2010, the Air Transport Association (ATA) noted that the three New York airports represent nearly half of all flight delays among the Nation's largest airports. However their analysis only expresses what portion of nationwide delays occur as a result of congestion at the New York airports and airspace (which includes Philadelphia), not their propagation effect (FAA, 2010)." We were unable to find other documentation concerning the ATA estimate and its methodological basis. In addition, in a 2010 GAO report (GAO report, 2010), OPSNET data for the year 2009 is analyzed by counting the percentage of delayed flights (flights with departure delay over 15 min). It is found that 41.2% of delayed departures can be attributed to the three New York airports. However, this estimate does not consider the severity (length) of delay nor its impact on the rest of the NAS. OPSNET data provide information on what facility the delay was attributed to—that is, which facility instituted a traffic management initiative or other delaying action. However, this approach to delay attribution does not account for propagated delay, since it looks at each flight segment in isolation, rather than as part of a sequence. In another report, the FAA New York Area Program Integration Office (NYAPIO) claims that in the year 2007, approximately 75% of all delays in the NAS occurred because of delays in the NY Metropolitan Area airspace (NYPIO, 2010). The basis for this estimate is not reported, however.

While the above reports claim that New York airports are major causes of delay throughout the NAS, there appear to be no estimates based on clearly documented, sound methodology. Various challenges, such as limitations of existing flight data and analytic methods, have hampered FAA and others within the aviation community in measuring the delay propagation. Thus, a recent FAA report concludes that so far no one fully understands how or to what degree New York delays impact flights nationwide (FAA, 2010).

The aim of this paper is to estimate how much NAS delay can be attributed to delays at the three New York airports, using two very different approaches. First, drawing on previous studies, an econometric model that captures the interaction between New York airport delay and delay at other airports is developed, estimated, and used to assess how NAS delay would change if New York delays, or at least those delays due to congestion in the airport and airspace in the New York region, could be eliminated. Second, a large-scale simulation model of NAS operations is used to predict the NAS delay changes given several scenarios assuming very large increases in New York airport and airspace capacity. The counterfactual scenarios examined in both methods are constructed so that the resulting estimates are as comparable as possible.

The analysis presented here makes several contributions to our knowledge and understanding of the operational performance of air transport systems. Most fundamentally, it contributes to our conceptual and methodological understanding of how congestion and delay in one local region affects conditions in a larger airport system. Second, by assessing this impact for the New York region using two different models and several counterfactual scenarios, our research shows the sensitivity of the results to these factors. Lastly, it offers a thorough and well-documented answer to a specific question about how much New York airports contribute to delay in the NAS. Despite its specificity, the citations above demonstrate that this question is important and has received considerable attention—and a wide range of answers—in the aviation community.

These contributions have important planning and policy implications. Multi-billion dollar expenditures to increase the capacity of the New York airports are currently being considered (Regional Planning Association, 2011), and the business cases for such projects can be substantially affected by the degree of delay propagation from these airports. The research also has a direct bearing on whether flight delay at New York airports is a national problem that requires federal intervention, whether in the form of capital assistance or measures to restrict demand, or a local one best left to the airport operator—the Port Authority of New York and New Jersey—and flight operators to deal with. Although we focus these questions on New York, they are equally relevant to any region with high levels of flight traffic and delay.

The remainder of this paper is organized as follows. Section 2 provides a brief literature review on the study of delay propagation as well as the value of delay. Section 3 provides a preliminary assessment of the delay propagation. Section 4 presents an overview of the methodology, while Section 5 presents the econometric model. The simulation model is presented in Section 6, along with a comparison of the results from the two approaches. Section 7 summarizes our conclusions and discusses future research needs.

2. Literature review

To organize our literature review, we consider what delay propagation mechanisms have been considered, whether explicitly or implicitly, in previous studies. The most commonly studied form is downstream delay that results when an aircraft arrives late at an airport and consequently departs late on its next flight segment. We will refer to this as forward delay propagation due to aircraft, or forward aircraft propagation. For many researchers, delay propagation refers exclusively to this phenomenon. For example, Boswell and Evans (1997) examined delay propagation resulted from late arriving aircraft using Markov chains, and found an overall multiplier of 1.8, meaning that 1 min of initial flight delay generated an additional 0.8 min of delay when propagated downstream. Schaefer and Millner (2001) analyzed aircraft delay propagation by simulating flight itineraries under different weather conditions, represented by different combinations of visual meteorological conditions (VMC) and instrumental meteorological conditions (IMC). They found that the propagation effect of a delay from one flight leg to the next leg was significant, but diminished with each subsequent leg. Further research on forward aircraft propagation by Wang et al. (2003) developed an analytical model to separate controllable factors that influence delays and their propagation in the NAS from other random factors. Finally, Laskey et al. (2006) developed a Bayesian Network model, taking into consideration the dynamic aspects of flight delay, to estimate delay propagation in the NAS. Delay is decomposed to represent different flight phases and each delay component is a dependent variable with previous delay components as independent variables.

Other forms of forward delay propagation result from other resources, aside from the aircraft itself, used by an arriving flight and required by a departing flight. These include flight crew and cabin crew. AhmadBeygi et al. (2008) consider both aircraft and flight crew in developing an analytical model to study the factors that influence delays and their propagation in the NAS. In their model each flight delay can propagate to one or two subsequent downstream flights. They find that in such cases well-planned slack between flights is crucial to absorb the disruption and thereby reduce propagation. Beatty et al. (1999) consider aircraft, flight crew, and cabin crew. They developed an elaborate table of "delay multipliers" using detailed data derived from a proprietary flight scheduling system developed by American Airlines. The delay multipliers specify the ratio of total delay, including both primary and propagated delay, resulting from a given primary delay. They find that the multipliers are a function of the length and time-of-day of the primary delay.

There are other forms of delay propagation that are not considered in any of the above papers. We will terms these *forward payload propagation*, *downstream demand disturbance*, and *backward propagation*. Payload propagation refers to flights that are delayed in order to await connecting traffic (passengers, bags, or cargo) from delayed arriving flights. Backward propagation occurs when flights bound for a certain destination are intentionally slowed by air traffic controllers, forcing other flights in the same traffic stream to also slow. Finally, downstream demand disturbance is a broader term referring to delays that result from disruptions to demand schedules for a host of resources that can result from a primary delay. For example, a late departing flight may delay the arrival of another flight by occupying its assigned gate. Or, as a result of a primary delay, temporal shifts in the arrival demand or departure demand at a given airport may cause additional delay. Demand disturbances could either increase or decrease delays, since demand could be shifted into more busy or less busy periods.

Researchers have addressed this multiplicity of delay propagation mechanisms by modeling delay propagation at a macroscopic level. Hansen and Zhang (2005) devised an econometric approach based on historical data to estimate delay propagation in the NAS. Their study focuses on the interaction between delay at a single airport and in the rest of the NAS. A simultaneous equation regression model is developed and the two-stage least squares method is used for estimating the parameters. Using LGA airport as a case study, they find a 1-min increase in delay in LGA will result in a 3-min delay increase in the NAS. Zhang and Nayak (2010) then applied a similar methodology to compare the impact of two major airports in the NAS, LGA and Chicago O'Hare (ORD). They introduced additional causal variables and also propose a way of applying the outcomes for NAS capital investment evaluation. Nayak and Zhang (2011) subsequently studied the interaction between delays in different airports and between theses individual airports and the rest of the NAS. A more complicated multiple simultaneous equations regression model was constructed and three-stage least squares method was used for estimation considering the correlation between error terms of the multiple equations. Some of the outcomes appeared counter-intuitive and they concluded that airline network structure needs to be captured in the model to account for flight delay correlation within one airline or between airlines and their alliance partners. Research by Diana (2009) also considered delay propagation from different airports at a macroscopic level based on the discrete Fourier transform. The analysis in that paper is based on "the ratio of the amplitude of delayed arrivals at specific sampled airport ... to the delayed arrivals at all the destinations served from the same sampled airport ..." The results suggest that delay propagation from market-concentrated airports does not have a more significant effect than delay propagation from other airports. Moreover, the analysis includes JFK and LGA, whose estimated delay propagation ratios for 2000, 2007, and 2008 appear to be within the normal range compared to other airports in the sample.

3. Preliminary calculations

While a comprehensive assessment of the contribution of New York airports to NAS delay presents a methodological challenge, some simple calculations shed light on the matter. If New York airports have a disproportionate effect on the rest of the system, one might expect that arrival delays of flights from New York exceed those of flights from other airports. Table 1 shows this comparison for the year 2010. The flights considered in Table 1 are those for airlines reporting to the ASQP data base, which is further discussed below, and whose destination is among a set of busy airports known as the OEP 35, also discussed below. Average arrival delay from flights originating from the New York airports is 2.2 min, or about 20%, greater

Table 1
Arrival delay by flight origin, OEP 35 airports, 2010.

Flight origin	Number of flights	Average arrival delay	Std. dev arrival delay
NY airports	316,012	13.4 min	35.0 min
Other major airports	5,988,643	11.1 min	30.7 min
Difference		2.25 min	4.3 min
Percent difference		20.2%	14.0%

than the delay from those not originating from the New York airports. The standard deviation of the arrival delay from New York airports is also greater, by about 4 min or 14%. Both of these differences are highly statistically significant.

As discussed above, one of the most widely recognized forms of propagated delay is forward propagated aircraft delay, which results when a flight arrives late because the aircraft serving that flight arrived late at the previous airport. The ASQP cause codes for delay include this one (BTS, 2014). Table 2 compares delay causes for flights from New York and flights from other airports for the year 2010. About 30% of flight delay minutes result from aircraft arriving late upstream for both NY and non-NY originating flights, with the NY share slightly lower. Thus, aside from the differences in overall delay revealed in Table 1, there is little indication that this form of propagated delay is a particular concern for the New York Airports.

What do these figures imply about the contribution of New York airports to NAS Delay? One rough answer can be directly obtained from Table 1. Suppose that no flight originating at any New York airport ever suffered an arrival delay, while arrival delays from other OEP 35 airports remained at their 2010 levels. Total arrival delay would then be reduced by $316{,}012 \cdot 13.4 \cong 14.2$ million minutes or 5%. A different basis for comparison is a scenario in which only delays incurred by New York originating flights that result from late arriving aircraft are eliminated. Based on Table 2, we estimate the delay reduction in this case to be 28% of the previous estimate, or about 1.5%.

These calculations are suspect for several reasons. It is not plausible to "blame" New York airports for all of the arrival delay of flights originating from them, as implied by the first estimate. On the other hand it is not appropriate to consider only delays caused by late arriving aircraft as the second estimate does. The first approach overlooks the fact that most arrival delay is caused by conditions at the destination airport, while the second neglects forms of delay propagation that are not related to a late aircraft arrival that were enumerated in the previous section, for example flight crew, cabin crew and payload propagation. An additional problem is that the delays reported in Table 1 are relative to the schedule. However, the schedule itself includes padding in anticipation of delays (Zou and Hansen, 2012; Hao and Hansen, 2012). A more appropriate baseline for assessing delays is against an unimpeded flight time. Finally, the analysis above considers only delays at the first down-line destinations of flights originating from New York airports, whereas delays can further propagate to airports further down-line.

More fundamentally, any analysis of delay propagation must contend with the fact that the effect is mutual. Delay at airport A can propagate to airport B, but delay at airport B can also propagate to airport A. Comparisons like the ones presented above cannot disentangle these effects.

These considerations point to the need for greater conceptual clarity as well as more sophisticated models. Conceptually, our first step is to change our terminology from delay "propagation" to delay "impact." This signifies that our interest is in how much delay is caused by a certain region (New York in our case) rather than how much delay propagates from that region. This distinction is important because delay that propagates from a given region may have been originally caused by delay in another region.

Next we must define more precisely the "impact" of region X's airports on delay in a larger airport system. This is best accomplished by specifying counterfactuals. For a given counterfactual, the delay impact is the change in system-wide delay that would result if the counterfactual were realized. For example, the counterfactual assumed in the first calculation above is no arrival delay for any flight originating in region X. There are other, more meaningful and appropriate, counterfactuals. For example, a counterfactual could be that no delay is ever caused by conditions in region X. A more complete discussion of alternative counterfactuals is provided below.

Whatever the counterfactual is, a model is required to assess its delay impact. The model and the counterfactual must be compatible. This requirement cuts both ways: the model must be capable of predicting the impact of a given counterfactual,

Table 2
Causes of arrival delay by flight origin, OEP 35 airports.

Delay cause	Flights from NY airports (percent of delay minutes)	Flights from other OEP airports (percent of delay minutes)
NAS	2.3	2.4
Carrier	35.2	32.0
Late aircraft	28.1	31.3
Weather	0.4	0.3
Security	0.4	0.1
Not reported	33.6	33.9
Total	100.0	100.0

while the counterfactual should be one whose impact can be modeled. We will see below how two different models of NAS delay lead to different counterfactuals and by extension different definitions of regional impact on system-wide delay.

4. Methodology

In this paper we employ two different methods to estimate the impact of delay at New York airports—consisting, for purposes of this analysis, of EWR, JFK, and LGA, and referred to as NY airports for convenience—on delay in the rest of the NAS. One method is based upon econometric modeling, while a second employs a NAS-wide simulation model.

The econometric approach builds upon the work of Hansen and Zhang (2005) and Zhang and Nayak (2010, 2011) described above. Simultaneous equations regression models of daily delay at NY airports and in the rest of the NAS are used to assess the system-wide impacts of NY airport delay. Arrival delays—for NY airports in one equation and the rest of the NAS in the other—are the endogenous variables, and several factors that affect delay, such as flight traffic volume, airport queuing delay, and adverse weather, are included as exogenous variables. These econometric models are estimated on a large sample of historical days as compared to the relatively small number that can be considered using the simulation approach, and thus reflect operational experience in a more comprehensive manner. On the other hand, in contrast to simulation, the econometric models do not explicitly represent the physical and management processes that give rise to the delay, but rather reflect them at an aggregated level. A series of counterfactual scenarios in which the New York delay equation is modified in various ways, and then solved simultaneously with the baseline rest-of-NAS equation, are used to estimate the system-wide impact of New York airport delays.

In the second approach, the impact of New York airports is analyzed using FAA's System-Wide Analysis Capability (SWAC) simulation model. This is a queuing model that incorporates en route and terminal capacity constraints as well as tail number tracking, so that propagation effects can be captured. To assess the impact of NY airport delay, a baseline run is compared to a counterfactual in which the capacity of the NY airports and surrounding terminal airspace is set to be effectively infinite. These runs are performed on a small number of representative days. The reductions in NY airport and NAS delay that SWAC predicts to result from these capacity expansions serve as the basis for the second NY airport delay impact estimate. The SWAC model considers forward aircraft delay propagation and demand disturbance propagation.

In both approaches, we consider the three New York airports together instead of individually. All three airports have substantial traffic and delays, and all three have been alleged to have substantial impacts on the rest of the NAS. Moreover, because of the complex airspace in the New York area there are significant interactions between the airports. While modeling these interactions would be of considerable interest in its own right, to do so here would distract us from the main objective of modeling the interaction between the NY airports and the rest of the NAS.

5. Econometric model

5.1. Model specification

In this econometric study, only flights into the OEP 35 (Operational Evolution Partnership) core airports are considered. The OEP 35 airports are commercial U.S. airports with significant activity and delay. These airports serve major metropolitan areas and also serve as hubs for airline operations. More than 70% of passengers move through these airports. The term "OEP 35" is no longer in official use by FAA, but the set of airports included in the OEP 35 still includes the busiest and most delay-prone ones. Restricting our analysis to this set increases tractability while ensuring that all of the major airport chokepoints in the NAS are considered.

Following the approaches of Hansen and Zhang (2005) and Zhang and Nayak (2010, 2011), two simultaneous models are estimated in this analysis. Daily average arrival delay per flight—including schedule buffer—at the NY airports and the other 32 OEP airports are the dependent variables in the model. (For convenience, we will heretofore refer to the latter as the "OEP 32.") Schedule buffer is an additional amount of time added to the delay taking the airlines' schedule padding into consideration. This will be explained in detail below. The model of NY airport delay includes congestion and weather conditions at the three New York airports, the number of NY airport arrivals, arrival delay at the OEP 32 airports, and convective weather as the explanatory variables. For delay at the OEP 32 airports, the explanatory variables include delay at the three NY airports, congestion and weather conditions at the OEP 32 airports, the number of arrivals at these airports, convective weather, and other factors. Thus the arrival delay on the left hand side of one model also appears on the right hand side in the other model, creating a simultaneous system.

The equation for Model 1 (daily average arrival delay with schedule padding at the New York airports) is thus:

$$D_{NY}(t) = \alpha_{NY} + \beta_1 \times OP_{NY}(t) + \beta_2 \times D_S(t) + \beta_3 \times Q_{NY}(t) + \beta_4 \times Q_{NY}^2(t) + \beta_5 \times I_{NY}(t) + \sum_k \theta_{kNY} W_k(t) \quad (1)$$

Likewise, the Model 2 equation (daily average arrival delay with schedule padding at the OEP 32 airports) is:

$$D_S(t) = \alpha_S + \gamma_1 \times OP_S(t) + \gamma_2 \times D_{NY}(t) + \gamma_3 \times Q_S(t) + \gamma_4 \times Q_S^2(t) + \gamma_5 \times I_S(t) + \sum_k \theta_{kS} W_k(t) \quad (2)$$

where:

$D_{NY}(t)$ = average observed arrival delay plus schedule buffer per flight against schedule at New York airports on day t;
$D_S(t)$ = average observed arrival delay plus schedule buffer per flight at the OEP 32 airports on day t;
$OP_{NY}(t)$ = total operations (arrivals) in New York airports on day t;
$Q_{NY}(t)$ = average arrival deterministic queuing delay in New York airports on day t;
$I_{NY}(t)$ = portion of flights scheduled to arrive under IMC at New York airports on day t;
$W_K(t)$ = thunderstorm portion in region k on day t;
$Q_S(t)$ = average arrival deterministic queuing delay in the OEP 32 airports on day t;
$I_S(t)$ = portion of flights scheduled to arrive under IMC at the OEP 32 on day t;
$OP_S(t)$ = total operations (arrivals) in the OEP 32 on day t;

The α's, β's, γ's, and θ's are coefficients to be estimated. The variables will be further explained in the next section.

The models form a simultaneous system because the delays in NY airports and in the OEP 32 are partially explained by each other. The fact that each model contains variables that are excluded from the other model guarantees there is no identification problem, so both models can be estimated. Our preliminary analysis of the data shows that the error terms of the two models are not correlated, which allows the usage of two-stage least squares (2SLS) to estimate the coefficients in the models. 2SLS is a standard method for estimating simultaneous equations (Pindyck and Rubinfeld, 1998). The application of 2SLS to get the estimates involves two sets of regressions. We first estimate the reduced form models with endogenous variables on the left hand side and all exogenous variables on the right hand side using OLS regression, and obtain fitted values for the endogenous variables. Next, we estimate the original, structural form, model using fitted values for the endogenous variables obtained from the first-stage estimations. By forming predictions for endogenous variables in the second stage based on the exogenous ones, the regression will produce consistent parameter estimates. For a large enough sample size, the 2SLS will produce correct standard errors if the residuals are normally, identically, and independently distributed, as is assumed in typical regressions. However, since time series data is used for the analysis, autocorrelations between the errors are observed and this will complicate the application of statistical tests. Thus, in the regression, we calculate standard errors that are robust with respect to autocorrelation. This guarantees that the standard errors are still reliable when the regression errors are autocorrelated. The statistics package in Stata is employed to estimate parameters and robust standard errors.

In the models presented in Eqs. (1) and (2), many potential factors that could impact delay are included. Some explanatory variables might not have a significant effect. For example, the convective weather in some regions might not significantly affect arrival delay in the New York area, or the OEP 32 airports. For purposes of predicting the delay changes under counterfactual scenarios, it is appropriate to use a more parsimonious model that excludes statistically insignificant variables. This so-called "prediction model" with a smaller set of explanatory variables will be presented in Section 5.4, along with the estimation results.

5.2. Model variables

5.2.1. Arrival delay

The dependent variables in our models are daily average arrival delays at the three NY airports, and the OEP 32 airports, respectively. The arrival delay can vary from time to time in a certain day. Our model, however, focuses on how daily average delay per flight at New York airports affects average flight delay in the OEP 32, and vice versa.

In the model described in this paper, schedule padding is also taken into consideration. Schedule padding is calculated as follows. Using ASQP (Airline Service Quality Performance) data, flights are grouped by OD pair, flight number, airline, and quarter. For each group, the schedule padding is defined as the difference between scheduled flight time and the 20th percentile of the actual flight time. This padding is the anticipated delay that airlines embed in their schedule. In the model, arrival delay is the sum of the reported arrival delay against schedule (which is truncated so that early flights are counted as having zero delay) and the schedule padding across all flights, aggregated to a daily level. This approach is consistent with the simulation model in the next section, which calculates delays relative to unimpeded flight times. The daily total arrival delay is then divided by the total number of arrivals, to obtain daily average delay, in units of min/flight. Note that the ASQP data base only contains domestic flights.

5.2.2. Deterministic queuing delay

Deterministic queuing delay is an indicator of imbalances between operational demand and capacity at an airport. For a certain day, scheduled arrival demand and the estimated capacity are used to construct the daily deterministic queuing delay variable. A given day is decomposed into quarter-hour time intervals. Thus, in each day there are altogether 96 time intervals. The cumulative arrival demand at quarter hour i is the sum of scheduled arrivals minus the number of cancelled flights in that day up until time period i. The arrival count in each quarter hour is assumed to be the minimum of the arrival demand that has yet to be served and the capacity. Let C_i^* be the capacity, based on the AAR (airport acceptance rate); C_i be the actual observed arrival count; and A_i be the as yet unserved arrival demand for quarter hour i. We calculate the arrival count as:

$$C_i = C_i^* \quad \text{for } A_i \geqslant C_i^*$$

$$C_i = A_i \quad \text{for} \quad A_i < C_i^*$$

Under this construction, the cumulative arrival count is always less than the cumulative arrival demand. The total cumulative queuing delay, TQ, in minutes, over the day can be calculated as the summation of the difference between cumulative demand and count over the 96 time periods in that day, as presented below:

$$TQ = 15\sum_{i=1}^{96}(A_i - C_i)$$

The queuing delay is aggregated to a daily average level. Fig. 1 illustrates the daily cumulative queuing delay. The summation in the above formula indicates the area between the demand and actual arrival curves. The daily average queuing delay per flight is calculated by dividing the total delay by the total number of arrivals. We perform this calculation for each of the NY and the OEP 32 airports and for each day in our analysis period.

5.2.3. Adverse weather

The effects of adverse weather were introduced into the model in two ways. First, convective weather in the United States is incorporated by using the daily summary of en route weather information obtained from NOAA (National Oceanic and Atmospheric Administration) database. Convective weather increases flight delay by making certain en route airspace unflyable, causing ground holds, lengthening flight paths, and disturbing terminal operations. Each observation from the over 1500 U.S. weather stations has a binary dummy variable indicating whether there is a thunderstorm recorded by the station during that day. To capture geographic patterns of thunderstorm activity, the continental US is divided into regions of 10° latitude by 10° longitude. Since the convective weather near the New York area is of particular interest, the regions near New York are further divided into regions of 5° latitude and longitude as shown in Fig. 2. This yields a total of 25 regions; the proportion of weather stations in each region reporting thunderstorms during each day is computed as an independent variable.

A second important metric for adverse weather is the proportion of arrivals scheduled to arrive under IMC (number of arrivals scheduled to arrive under IMC divided by the total number of scheduled arrivals) for each day. Airports have lower capacity under IMC. While this effect is partly captured by the AAR and thus the queuing delay variable, the actual difference in capacity between IMC and VMC is generally somewhat greater than reflected in the AARs. The IMC metric is based on the scheduled arrivals rather than the actual arrivals because the actual arrival time is partly the result of delays, and is therefore endogenous.

5.2.4. Total flight operations

The total daily flight operations (arrivals) at the three NY airports and the OEP 32 airports were also included as variables. The number of arrivals at an airport affects terminal delay since more arrivals generates more complicated ramp and taxiway operations. Also the New York terminal airspace has capacity constraints that are not fully captured by airport AARs or visibility conditions.

5.3. Data

Our daily dataset covers the years from 2004 through 2010. Much of the data used in this analysis are from the Aviation System Performance Metrics (ASPM) database, maintained by FAA's Aviation Policy and Plans Offices and the Airline Service Quality Performance (ASQP) database, compiled by the Department of Transportation's Bureau of Transportation Statistics. We used ASPM quarter-hourly data on throughput, demand, called arrival rates, and numbers of operations (arrivals) under

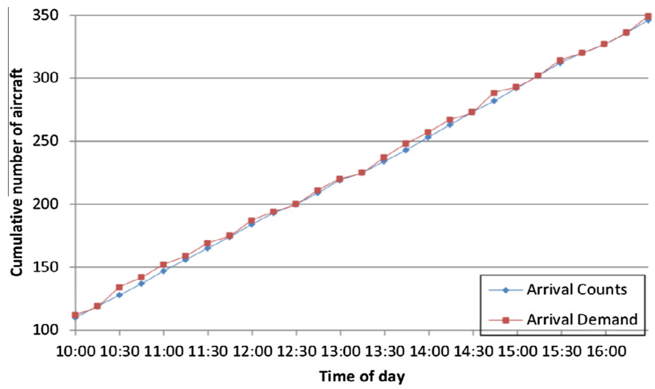

Fig. 1. Cumulative queuing delay for a given airport and day.

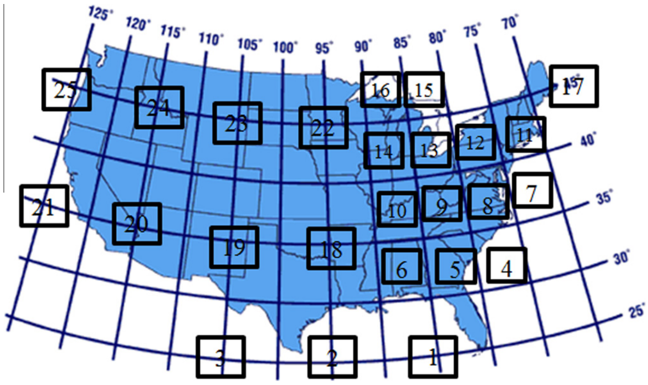

Fig. 2. Convective weather (thunderstorm ratio) regions configuration.

different weather conditions. From the ASQP data, information of each flight operated by a major airline in the United States is captured, including the scheduled and actual arrival times. The delay metric used in this study is calculated based on these values. Both the ASPM and the ASQP datasets are selected based on airports. Only flights into the OEP 35 core airports are considered in this analysis.

We obtained convective weather data from the National Oceanic and Atmospheric Administration (NOAA) Global Summary of the Day dataset. This dataset provides daily surface weather information from over 1,500 U.S. weather stations. The NOAA database characterizes thunderstorm activity in each station by recording whether or not there is a thunderstorm at the station on each day.

5.4. Estimation results

As mentioned in Section 5.1, the "full model" provides a complete picture of how all the relevant factors affect arrival delay more or less. However, for purposes of prediction of delays under counterfactual scenarios we employ a more parsimonious model, which we term the "prediction model". First, the "full model" is estimated on our dataset and insignificant variables are found for both NY and OEP 32 airports. These variables are the convective weather variables in some regions. For the "prediction model", these insignificant variables are removed from the "full model", and the model is estimated again on the dataset. In the "prediction model", all explanatory variables are significant at the 0.1 level. Even though the intercepts in both New York and OEP 32 models are not quite significant, they are retained in the model so that predictions are unbiased. Estimation results for the "prediction model" are shown in Tables 3 and 4.

New York arrival delay model explains about 66% of the variation in average arrival delay, based on the R^2 value from Table 3. The estimated coefficient for IMC scheduled arrivals is 18.86, implying that average delays on all-IMC day would be about 19 min higher than on all-VMC one, all else equal. The coefficient for traffic volume is 0.0074. Since the average total daily arrival volume into the NY airports is 1581 flights per day, the volume effect causes a delay of about 12 min/flight on an average day. The estimated coefficient for average queuing delay is 0.34 while the negative coefficient for the quadratic term shows that the effect diminishes as queuing delay increases. The coefficient for queuing delay is rather small since one

Table 3
Estimation results – New York prediction model.

NY delay model	Description	Estimate	Std. err.	*P*-value
Intercept		−2.36	2.20	0.283
$D_S(t)$	NAS arrival delay	0.22	0.06	<0.001
$I_{NY}(t)$	NY MC arrival	18.86	1.27	<0.001
$Q_{NY}(t)$	NY queuing delay	0.34	0.03	<0.001
$Q^2_{NY}(t)$	Quadratic NY queuing delay	−0.00027	0.00	<0.001
$OP_{NY}(t)$	Total flight operations in NY	0.0074	0.00	<0.001
$W_2(t)$	Thunderstorm ratio in Region2	−32.01	6.80	<0.001
$W_3(t)$	Region3	8.11	4.20	0.054
$W_7(t)$	Region7	41.11	18.73	0.028
$W_8(t)$	Region8	45.29	12.69	<0.001
$W_{11}(t)$	Region11	99.80	10.91	<0.001
$W_{12}(t)$	Region12	51.81	8.48	<0.001
$W_{22}(t)$	Region22	−21.97	4.00	<0.001
$W_{24}(t)$	Region24	−23.95	10.13	0.018
$W_{25}(t)$	Region25	−59.42	28.29	0.036
R-square		0.659		

might expect an additional minute of queuing delay to result in an additional minute of total delay. This could be because queuing delay is highly correlated with traffic volume as well as IMC conditions. Increases in both of these naturally lead to longer queues and more delay. For NY airports specifically, a large portion of this delay results from problems in the terminal airspace, and this may not be fully captured in the queuing delays for the individual airports. Thus, it appears that much of the impact of queuing delay is shifted from the queuing variables to the volume and IMC ones in the NY airport model.

Delays in the OEP 32 airports also have a significant effect on New York arrival delay. One minute of arrival delay per flight at the OEP 32 airports causes about 0.22 min of increased delay per flight for New York arrivals. This is reflects one direction of the two-way interaction between NAS delay and NY airport delay that our simultaneous model is designed to capture.

Of the convective weather variables, nine regions listed in Table 3 are statistically significant and are included in the "prediction model". As expected, the coefficient estimates are larger and more significant for regions near New York area. Specifically, regions 7, 8, 11 and 12 have much larger coefficients than most other regions, and they are all highly statistically significant. As shown in Fig. 2, these four regions surround the New York City region. Some significant convective weather effects are also found for regions quite distant from New York, and in many cases the signs on these coefficients are negative. Similar results have been obtained in past studies (e.g. Hansen and Zhang, 2005). The most likely explanation is that thunderstorm activity is often related to the jet stream, which contributes to atmospheric instability (Mason, 2013). The jet stream in turn can have a substantial impact on flight times. Specifically, if the polar or subtropical jet streams flow over the US, this can lead to convective weather in the US, but also reduce flight times. Eastbound flights can exploit the jet stream through strategic flight routing, while westbound flights can likewise mitigate its impact. Whatever the exact explanation, the convective weather variables are included to help isolate the interactions between delays at NY and non-NY airports, so the interpretation of their coefficients is not of direct relevance to the goal of this research. Moreover, omitting a statistically significant explanatory variable, even one whose coefficient has an unexpected sign, from a regression model can lead to bias in the remaining estimates (Green, 2000).

The results in Table 4 show estimation results for the model of average delay at the OEP 32 airports. The R^2 value indicates that about 66% of the variation in average arrival delay is explained by the model. Again the proportion of IMC arrivals has a major effect on arrival delay and the queuing delay impact, while significant, attenuates as queuing delay increases. Queuing delay has a coefficient of 1.22, which is considerably larger than that in the NY airports model. It appears that for the OEP 32 airports, the airport queuing delay is a stronger explanator because the terminal airspace constraints are not as serious an issue as at the NY airports. This is supported by the far smaller coefficient for the traffic volume variable—0.00027 compared to 0.0074 in the NY airports model. Average arrival volume across the OEP 32 is 16,374 flights per day; thus traffic volume contributes about 4.4 min/flight to delay at the OEP 32, far less than the 12-min contribution for the NY airports calculated above. One minute of arrival delay per flight at the three NY airports causes about 0.08 min of increased delay per flight at the OEP 32 airports. This is the spillover effect through which NY airport delay impacts the rest of the NAS. While the coefficient of NY arrival delay in this model is much smaller than that for OEP 32 delay in the NY model, this is to be expected given the relative numbers of flights in the two sets of airports. The coefficients for convective weather in different regions are more evenly distributed across the regions in the OEP 32 model as compared to the NY airport model. This is to be expected since this model predicts the arrival delay in a set of airports distributed across the US. As in the New York model, there are a few negative signs on the convective weather variables, again presumably as a result of jet stream effects.

5.5. Quantifying the impact of New York airport delay

As discussed above, to assess the system-wide impact of New York airport delays, specific counterfactuals associated with precise definitions of this impact must be defined. As will be shown, the magnitude of the impact obtained strongly depends

Table 4
Estimation results – OEP 32 delay model.

OEP32 delay model	Description	Estimate	Std. err.	*P*-value
Intercept		0.92	0.65	0.157
$D_{NY}(t)$	NY airports arrival delay	0.08	0.01	<0.001
$I_S(t)$	OEP32 average IMC arrival	11.87	1.10	<0.001
$Q_S(t)$	OEP32 queuing delay	1.22	0.08	<0.001
$Q_S^2(t)$	Quadratic OEP 32 queuing delay	−0.025	0.00	<0.001
$OP_S(t)$	Total flight operations in OEP 32	0.00027	0.00	<0.001
$W_6(t)$	Region6	11.80	1.83	<0.001
$W_8(t)$	Region8	18.33	4.68	<0.001
$W_9(t)$	Region9	17.74	3.62	<0.001
$W_{10}(t)$	Region10	−7.00	1.98	<0.001
$W_{14}(t)$	Region14	18.65	2.30	<0.001
$W_{18}(t)$	Region18	11.69	2.09	<0.001
$W_{19}(t)$	Region19	−7.15	2.80	0.011
$W_{20}(t)$	Region20	18.54	6.70	0.006
$W_{22}(t)$	Region22	−9.29	1.86	<0.001
$W_{25}(t)$	Region25	−32.98	7.72	<0.001
R-square		0.655		

on our choices in this regard. In the following scenarios, the delay in the OEP 32 and the NY airports is found by solving two simultaneous equations. By making changes to the New York delay Eq. (1), while keeping Eq. (2), the OEP 32 equation, fixed, new delays for New York and OEP 32 are calculated, and impacts are estimated by comparing these with the baseline delay. As discussed above, the "prediction model" presented in Section 5.4 is used to construct the counterfactual scenarios.

5.5.1. No New York arrival delay (Scenario I)

In this scenario the New York delay is set to be zero. This is an extreme case, since it implies that flights arrive at New York airports on time even if they are delayed upstream. The three NY airports are like "black holes" in the NAS system that cleanse all the delay of any inbound flight. Scenario I is thus called "black hole" scenario. By setting the New York arrival delay in Eq. (2) as zero, there is no contribution from New York to the flight delay in the OEP 32. The assumed NY airports delay and the consequent OEP 32 delay is shown in Eqs. (3) and (4):

$$D_{NY}(t) = 0; \tag{3}$$

$$D_S^I(t) = \alpha_S + \gamma_1 \times OP_S(t) + \gamma_3 \times Q_S(t) + \gamma_4 \times Q_S^2(t) + \gamma_5 \times I_S(t) + \sum_k \theta_{kS} W_k(t) \tag{4}$$

5.5.2. No New York local factors (Scenario II)

In this counterfactual, the potential for reducing arrival delay at the New York airports is circumscribed by dividing exogenous variables in Eqs. (1) and (2) into two categories: local and non-local, and only eliminating the local factors. These are the variables related to local airport operating conditions. For instance, queuing delay is highly related to the traffic volume and capacity of the airport, and is thus considered as a local variable. The non-local factors in this model include the thunderstorm variables across the United States and the arrival delay elsewhere in the NAS. In contrast to the non-local factors, local ones can be mitigated through improvements to the NY airports and airspace. Thus, Scenario II makes considerably more restrictive assumptions regarding the possible reduction in arrival delay that could occur at NY airports. To model this scenario, the thunderstorm variables and intercept term, in addition to the OEP 32 delay term, are retained in Eq. (1) (see Eq. (5) below). The two equations, Eqs. (2) and (5), are solved simultaneously to get the expression for OEP 32 arrival delay in Scenario II, as shown as Eq. (6).

$$D_{NY}(t) = \alpha_{NY} + \beta_2 \times D_S(t) + \sum_k \theta_{kNY} W_k(t) \tag{5}$$

$$D_S^{III}(t) = \frac{\alpha_S + \gamma_1 \times OP_S(t) + \gamma_2 \times \left(\alpha_{NY} + \sum_k \theta_{kNY} W_k(t)\right) + \gamma_3 \times Q_S(t) + \gamma_4 \times Q_S^2(t) + \gamma_5 \times I_S(t) + \sum_k \theta_{kS} W_k(t)}{1 - \gamma_2 \times \beta_2} \tag{6}$$

5.5.3. No New York queuing delay and volume factors (Scenario III)

The next two counterfactuals scenarios focus on variables that directly capture the effect of demand-capacity imbalance on NY airport arrival delay. Queuing delay is clearly one such factor. Moreover, in the simulation model that will be elaborated later, primary delay (that not caused by propagation) is entirely the result of queuing. Thus, it is of interest to have a scenario in which the queuing delay terms are excluded. Furthermore, as discussed above it appears that much of the queuing effect is captured by the arrival volume variable in the NY airport delay model. Consequently, in Scenario IV both the queuing and traffic volumes variables are excluded from the New York delay model and all the other factors—arrival delay in the NAS, IMC arrivals in NY airports, convective weather variables and the intercept—are retained. The resulted equations are similar as those in Scenario III, only with more NY airport local factors, as shown in Eqs. (7) and (8).

$$D_{NY}(t) = \alpha_{NY} + \beta_2 \times D_S(t) + \beta_5 \times I_{NY}(t) + \sum_k \theta_{kNY} W_k(t) \tag{7}$$

$$D_S^{IV}(t) = \frac{\alpha_S + \gamma_1 \times OP_S(t) + \gamma_2 \times \left(\alpha_{NY} + \beta_5 \times I_{NY}(t) + \sum_k \theta_{kNY} W_k(t)\right) + \gamma_3 \times Q_S(t) + \gamma_4 \times Q_S^2(t) + \gamma_5 \times I_S(t) + \sum_k \theta_{kS} W_k(t)}{1 - \gamma_2 \times \beta_2} \tag{8}$$

5.5.4. No New York queuing delay factors (Scenario IV)

In Scenario IV, we only remove the queuing variables from the NY airport delay model. Since the simulation in queuing-based, this scenario is the most direct analogue to the counterfactual scenario considered in that model. It is also the most restrictive scenario because NY airport arrival delay is reduced only through the exclusion of airport queuing delay. The delay equations for this scenario are presented in Eqs. (9) and (10).

$$D_{NY}(t) = \alpha_{NY} + \beta_1 \times OP_{NY}(t) + \beta_2 \times D_S(t) + \beta_5 \times I_{NY}(t) + \sum_k \theta_{kNY} W_k(t) \tag{9}$$

$$D_S^V(t) = \frac{\alpha_S + \gamma_1 \times OP_S(t) + \gamma_2 \times \left(\alpha_{NY} + \beta_1 \times OP_{NY}(t) + \beta_5 \times I_{NY}(t) + \sum_k \theta_{kNY} W_k(t)\right) + \gamma_3 \times Q_S(t) + \gamma_4 \times Q_S^2(t) + \gamma_5 \times I_S(t) + \sum_k \theta_{kS} W_k(t)}{1 - \gamma_2 \times \beta_2} \tag{10}$$

In summary, we have, in addition to the baseline case based on the historical data, four counterfactual scenarios. Scenario I is known as the "black hole" scenario in which, unrealistically, delay is expunged from all flights into the NY airports. Thus in Scenario I New York delay is identically equal to 0. All other scenarios are based on the estimated Eq. (1) but with certain independent variables excluded. Scenario II excludes all factors except non-NY delay and convective weather. In scenario III, only traffic volume and queuing delay are excluded from the NY airport delay equation. Finally, in Scenario IV only the deterministic queuing delay variable is excluded. Therefore, the delay reduction in Scenario IV will be the smallest among all the scenarios.

Under the above baseline and four counterfactual scenarios, the daily average arrival delays are recalculated using the same set of data and coefficient estimates from Section 5.4. The average delays per flight in the three NY airports and in the OEP 32 airports are computed separately. We perform this calculation over all days in 2010, and average the results across the days. Then, we take a weighted average of the delays in these two sets of airports—New York and OEP 32—to get an overall average delay per flight for the OEP 35 airports. The first part of Table 5 presents the average daily arrival delays. In the baseline case, the average delay per flight (including schedule padding) across the OEP 35 airports is 12.4 min/flight, with NY delays more than 50% greater than non-NY delays. Delays under the four counterfactual scenarios are, as expected, lower by varying amounts.

The average values of delay per flight in different scenarios are then compared with the baseline case to calculate the delay reduction. The middle part of Table 5 presents the amount of delay reduction in each scenario. The delay reductions decline with each successive scenario since scenarios are ordered based on the size of the set of delay-inducing factors that are excluded. In the most extreme Scenario I, the reduction is 3 min/flight if all New York delay is eliminated. More realistically, the overall delay reduction is 2.8 min/flight if NY delays are only caused by delay elsewhere and convective weather. If only the volume and queuing delay factors are excluded from the NY model, as in Scenario III, the delay reduction is 2.4 min/flight. Finally, under Scenario IV, in which only queuing delay is eliminated from the NY airport delay model, we see the lowest delay reduction—0.5 min/flight.

We observe a ratio of about 15–1 in the delay reduction of NY and non-NY airports. This means that, on a per flight basis, delay reductions under the various counterfactuals are heavily concentrated on the flights bound for NY airports. Even under the most extreme counterfactual in which NY airport delay is set to 0, the resulting change in non-NY airport delay is a mere 1.4 min/flight. This challenges the claim that NY airport delay is a dominant source of delay elsewhere in the NAS.

The bottom part of Table 5 illustrates the delay reduction in percentage terms. The system wide percentage reductions in delay range from 24% under the "black hole" Scenario I to a mere 4% if only the NY airport queuing delay variables were eliminated. The NY airport percentage reductions exceed the OEP 32 ones by roughly a factor of 8–10. The reductions to non-NY airport delay under the different scenarios are always less than 12%.

6. Simulation model

6.1. SWAC model

The FAA's System Wide Analysis Capability (SWAC) is a fast-time simulation model used to estimate the potential benefits of new technologies, procedures, and infrastructure in the National Airspace System (NAS). Typical SWAC outputs are flight-by-flight delay (at the gate, on the surface, and in the air) and fuel burn (FAA, 2012). In this study we use SWAC to assess the savings of a hypothetical—and unrealistic—improvement that eliminates capacity constraints at the NY airports and in the terminal area. At its core, SWAC is a discrete-event queuing model. All en route sectors in Contiguous United States (CONUS) airspace, 110 domestic airports, and terminal airspace at the 35 busiest airports (the OEP 35) are included as servers. In order to represent the demand on those servers accurately, each flight must be modeled at a very detailed level.

SWAC begins with actual flight data from the FAA's Enhanced Traffic Management System (ETMS), gathered as the baseline set of flights. These flights are then augmented with visual flight rule (VFR) arrivals and departures from the FAA's OPSNET data. Demand and capacities, as key components to the model, are then predicted under different future scenarios. Various sources of data are used for the estimation, including Eurocontrol's Base of Aircraft Data (BADA) for aircraft performance, and National Convective Weather Diagnostic (NCWD) as well as Meteorological Aerodrome Report (METAR) data. Once initial demand and capacity have been estimated, SWAC runs a module to determine if any ground delay programs need to be implemented (to account for bad weather, for example). This computation allows for more accurate estimates of flight time, fuel usage, and sector congestion, by shifting delay to the surface that might otherwise have been taken in the air.

To capture delay propagation the SWAC model tracks tail numbers and employs random turn times. Given the arrival time of one flight, the departure on an outbound flight by the same aircraft may also be delayed, depending on the realized

Table 5
Average delay and delay reduction per flight under different scenarios.

Scenario Description	Baseline Based on historical data	I “Black hole” scenario	II Delay elsewhere and weather causing delay in NY airports	III Queuing delay and traffic volume are excluded from NY delay	IV Only queuing delay is excluded from NY delay
Average Delay per Flight (min)					
NY	19.20	0	1.21	3.71	14.95
Non-NY	11.59	10.22	10.31	10.51	11.41
Combined	12.39	9.42	9.62	10.02	11.90
Delay Reduction per Flight (min)					
NY	–	19.20	17.99	15.49	4.26
Non-NY	–	1.37	1.28	1.08	0.18
Combined	–	2.97	2.77	2.37	0.49
Delay Reduction per Flight (%)					
NY	–	100.00	93.70	80.66	22.18
Non-NY	–	11.80	11.06	9.31	1.51
Combined	–	23.97	22.35	19.11	3.96

value of the random turn time. Turn time distributions vary according to airport and aircraft type. After the queuing model is run and delays are computed, some additional post-processing is also performed. Excessive airborne delay is shifted to the ground, and flights with excessive ground delays are treated as cancelled flights. The model has recently been enhanced to re-allocate excessive airborne delay and automatically cancel flights. In terms of the terminology introduced in Section 3 above, SWAC captures forward aircraft propagation and downstream demand disturbance.

Each year, the FAA’s system-wide modeling approach is updated to incorporate the latest available operational data and forecasts. Since it is not yet practical to model an entire year with the detail required by SWAC, a set of days is selected with which to represent the year; for this analysis a set of 12 days from Fiscal Year 2010 was used. The FAA uses an optimization method formulated as a Mixed Integer Program (MIP) to select these sample days (Cheng et al., 2011). The objective of the MIP is to minimize the difference, for a set of metrics, between the true population (consisting of 365 observations for FY 2010) and the selected sample of 12 observations, subject to a set of constraints. For this annual update the metrics in the objective function are the operations counts and delays for 30 major airports; and operations counts and flight hours for the 20 domestic en route control centers, Alaska, and the three oceanic centers. This sampling approach assures that good and bad weather days are included. The constraints include requiring three sample days for each calendar quarter.

6.2. Using SWAC to assess the delay impacts of NY airports

In this paper, we used the SWAC model to provide another, independent estimate of the contribution of the three NY airports to NAS-wide delays. To do this, two scenarios are run: a baseline scenario, corresponding to the way the NAS currently functions, and a counterfactual scenario where all constraints associated with the NY airports were relieved. The latter will be referred to as the “Infinite NY Capacity” scenario. Fiscal Year (FY) 2010 data were used in both cases. For each scenario we modeled 12 individual days in 2010. The baseline scenario used current airport and airspace capacity estimates, combined with actual traffic and weather that occurred on the 12 days being modeled. The model produces estimates of gate, surface, and airborne delay for every flight operation in the NAS for each of the 12 days. The results are averaged to provide average daily delay estimates.

For the Infinite NY Capacity scenario the airport capacity restrictions at LGA, JFK, and EWR were removed from the model. This means that these three airports had infinite arrival and departure capacity, or that they acted exclusively as sources and sinks. Additionally, we eliminated any terminal airspace constraints for the three airports. Thus, unless a departing aircraft was assigned an Estimated Departure Clearance Time (EDCT) for a Ground Delay Program (GDP) at another airport, no taxi-out delays were accrued at these airports, and no terminal airspace delays either. Similarly no terminal airspace delays, due either to airspace or runway constraints, were accrued for arriving flights. In addition, no taxi-in delays were accrued. We did leave gate turn time and push-back random variables in the model, so some delay could have been accrued at the gate. Also any upstream delays for arriving flights might be propagated through the airports, depending on the magnitude of the delay, the scheduled turn time, and the random draw for the realized turn time when the model was run. We did not change the capacities of en route sectors adjacent to the New York terminal airspace. While these can constrain traffic (particularly during convective weather), we found that less than 0.3 percent of all airborne delay in the model was attributable to New York and Boston Center airspace.

6.3. SWAC results

The SWAC model yielded a total of 1.90 million minutes of delay for the 12 days in the baseline scenario, system wide. This delay includes that taken at the gate (measured as the actual push-back time minus the scheduled push-back time), on the surface (the actual taxi time minus the unimpeded taxi time), and in the air (the actual airborne time minus the unimpeded time, or estimated time en route). For the Infinite NY Capacity scenario the model yielded a total of 1.63 million minutes of delay for the 12 days, a reduction of approximately 14 percent. The average delay is originally 3.80 min/flight and goes down to 3.26 min/flight under the Infinite NY Capacity scenario. The average delay reduction is 0.54 min per flight and the reduction percentage is about 14%.

The statistics reported above reflect the totals for all IMC flights modeled in the SWAC, which are all flights that arrive, depart, or transit U.S. airspace. For a more direct comparison to the regression results reported in Section 5 of this paper, we filter the simulation output to select only those flights arriving at OEP 35 airports. When this is done, we find a total of 1.08 million minutes of delay in the baseline case. Furthermore, in this case the delay is decomposed into NY and non-NY delay, and a combined value is calculated based on the number of operations. In the Infinite NY Capacity case there were a total of 895,000 min of delay, a reduction of approximately 17%. The average delay is 4.94 min/flight originally and 4.10 min/flight under the Infinite NY Capacity scenario, with a delay reduction of 0.84 min/flight.

The SWAC results are summarized in Table 6.

6.4. Result comparison

Table 7 allows us to compare the estimated results of the arrival delay impact of NY airports using the econometric multivariate model (in Table 5) and the SWAC simulation model (in Table 6). To make the comparison more consistent, the analysis in Section 5.5 was repeated specifically on the 12 days that were analyzed with the SWAC model. The results are separately presented in Table 5, with the 12-day analysis results denoted with a suffix 12 in the description column. The two sets of econometric results are quite similar. However to maximize comparability we focus on the 12-day results in the discussion below.

Overall, the average delay in the SWAC model is far less than the values in the econometric model. For the baseline case, SWAC yields an overall delay of 5 min/flight average delay in the OEP 35, whereas the econometric model based on 12 days suggests 12 min (including schedule buffer). The most likely explanation for the difference is that the SWAC simulation model considers almost solely the queuing delay at each airport as the cause for arrival delay. In the econometric model, in contrast, all the possible causes for arrival delay are considered, because they are all reflected in the empirical data. In addition to the recorded arrival delay, the difference between scheduled flight time and the 20th percentile of the actual flight time, defined as schedule padding, is also counted as arrival delay in the econometric model. The choice of the 20th percentile is somewhat arbitrary however, and it may be that part of the difference in delay between the two approaches results from this choice, to which the empirical delay is quite sensitive.

Comparing NY and non-NY airports, the baseline delay at the former is 67% greater according to the SWAC model and 50% greater according to the regression model. It is not surprising that the difference exhibited between NY and non-NY airports is greater under SWAC, since it focuses on delay related to capacity constraints. There is less reason to expect other forms of delay—for example those caused by mechanical problems—to occur disproportionately at the New York airports.

With regard to delay reduction, we see that the average delay reductions obtained from the econometric model, under Scenarios II, III, and IV—2.5, 2.2, and 0.3 min respectively—bracket the 0.8 min/flight from the SWAC model. These three scenarios all assume the elimination of the queuing delay variables. Scenarios III and IV also eliminate other variables that are related to queuing delay: traffic volume and, in the case of Scenario II, IMC conditions. As discussed before, these factors absorb some of the delay impact of congestion in the NY econometric model. Therefore, it is reasonable that the Scenario IV reduction is less than that predicted by the SWAC model. However, it is also notable that the Scenario II and III reductions are more than twice the SWAC estimate. This suggests that much of the delay that is not included in the SWAC model is nonetheless related to the amount of demand at the NY airports.

The relative magnitudes of the NY airport and OEP 32 delay reductions from the SWAC and scenario IV from the econometric model are fairly similar. According to the latter, the ratio of delay reduction for NY and non-NY airports under Scenarios II and III is roughly 14–1. In the SWAC model, this ratio is approximately 10–1. These large ratios again suggest that NY airport delay is mainly a local problem. While the OEP 32 airports feel the NY airports' pain to some extent, the spillover effect is quite modest according to either model.

If overall delay reduction is expressed in percentage terms, scenarios II and III of the econometric model yield estimates quite similar to those from SWAC—21% and 18% respectively versus 17%. Percentage reductions for non-NY airport arrival delays predicted by the two models are also quite close—around 10% in all cases (except Scenario IV for the econometric model)—while there is a somewhat greater difference between the model predictions for the NY airports. As discussed above these similarities in percentages mask larger differences in predictions of baseline delays and absolute delay reduction.

In sum, despite their differences, the econometric and simulation models are consistent in refuting the view that NY airports are dominant source of delay throughout the NAS. To be sure, spillover effects exist, but they are not particularly large. Results from both models thus comport with the plausible if somewhat prosaic idea that benefits from reducing delay at NY airports, either by increasing capacity or reducing demand, are strongly concentrated on operations at these airports.

This agreement notwithstanding, there is substantial disparity between the econometric and simulation model results. The difference in overall delay levels can be explained by the exclusive focus of the SWAC model on queuing delay. But this does not account for the large differences in predicted delay savings from eliminating queuing-related delay at NY airports. The differences suggest that there are significant congestion phenomena not captured by the SWAC model. A plausible

Table 6
Delay and delay reduction results by SWAC simulation model, 12 days.

	Entire NAS		OEP 35 arrivals	
	Baseline	Infinite NY capacity	Baseline	Infinite NY capacity
Average Delay per Flight (min)				
NY	–	–	7.80	3.26
Non-NY	–	–	4.66	4.19
Combined	3.80	3.26	4.94	4.10
Delay Reduction per Flight (min)				
NY	–		4.54	
Non-NY	–		0.47	
Combined	0.54		0.84	
Delay Reduction per Flight (%)				
NY	–		58.21	
Non-NY	–		10.09	
Combined	14.21		17.00	

Table 7
Comparison of econometric model and SWAC model.

	Baseline		Scenarios			
	Econometric	SWAC	Econometrics			SWAC
Description	Based on historical data	Based on simulation	II. Delay elsewhere and weather causing delay in NY airports	III. Queuing delay and traffic volume are excluded from NY delay	IV. Only queuing delay is excluded from NY delay	Infinite NY capacity
Average Delay per Flight (min)						
NY	19.20		1.21	3.71	14.95	
NY(12)	19.43	7.80	3.48	5.38	16.79	3.26
Non-NY	11.59		10.31	10.51	11.41	
Non-NY(12)	11.06	4.66	9.93	10.08	11.00	4.19
Combined	12.39		9.62	10.02	11.90	
Combined(12)	11.89	4.94	9.39	9.71	11.59	4.10
Delay Reduction per Flight (min)						
NY	–	–	17.99	15.49	4.26	
NY(12)			15.96	14.05	2.64	4.54
Non-NY	–	–	1.28	1.08	0.18	
Non-NY(12)			1.13	0.98	0.06	0.47
Combined	–	–	2.77	2.37	0.49	
Combined(12)			2.50	2.18	0.30	0.84
Delay Reduction per Flight (%)						
NY	–	–	93.70	80.66	22.18	
NY(12)			82.11	72.32	13.60	58.21
Non-NY	–	–	11.06	9.31	1.51	
Non-NY(12)			10.25	8.83	0.53	10.09
Combined	–	–	22.35	19.11	3.96	
Combined(12)			21.04	18.30	2.50	17.00

culprit here is airport surface congestion, which is captured (via the flight operations variable) in the econometric model but not by the version of SWAC used here.

7. Conclusion

In this research, two different approaches are used to study the mutual interaction between the three NY airports and the rest of the NAS, and in particular the system impact of NY airport delays. Applying an econometric model, our results confirm that the dependencies do exist. One minute of delay per flight in the three New York airports will cause 0.08 min of delay per flight in the other large airports in the NAS. One minute of delay per flight in the rest of the NAS will also generate 0.22 min of delay in New York. These differing values are reasonable, considering the relative scale of operations for three New York airports compared to the 32 other airports considered in our model. Using the econometric model coefficients, different counterfactual scenarios involving the easing of congestion in New York are assumed to quantify how the delays at the other airports and over the whole system would be affected. In the most extreme case, where all delay disappears once the flights fly into the three New York airports, there would be an overall delay reduction of 3 min/flight, 24.5% of the total delay in the NAS. More realistically, if only "local" sources of delay are eliminated at the New York airports, the resulting delay reduction in the NAS is 2.4–2.8 min/flight, or 19–22%, depending on how we operationalize this concept. These estimates are all fairly close to one another, suggesting that the system delay impact of the New York airports is not strongly sensitive to the specific counterfactual assumed. The one "outlier" scenario, in which only queuing delay variables are eliminated, yields a much lower impact, but this result can be confidently dismissed based on the arguments presented above that other explanatory variables absorb the delay impact and solely eliminating queuing delay therefore captures only a small part of it.

A simulation approach is also conducted to study the same question. Using the SWAC model developed by the FAA, 12 typical days of operations are simulated under a baseline scenario and on which capacity constraints at the New York airports and the New York terminal airspace are eliminated. The difference in average delay across all the OEP airports under these two scenarios is 0.84 min, about a third of the reductions of the most comparable scenarios based on the econometric model. Nonetheless, there is agreement between the two models that delay arising from local congestion-related factors at New York airports is overwhelmingly incurred by New York flight traffic. In contrast to the often-cited claim that New York airport congestion is a major source of delay elsewhere in the NAS, results from both approaches show that, to a very large extent, "what happens in New York stays in New York."

This qualitative agreement belies substantial disparities in the quantitative results. The simulation estimates that eliminating congestion at New York airports would result in a 0.84 min/flight delay reduction NAS wide (including New York) and a 4.5 min/flight reduction at New York airports, while the corresponding values from the econometric model (when applied to the same days as the SWAC model) are 2.2–2.5 and 14–16 min/flight. In terms of percentage difference, the difference in absolute values between the two approaches is offset to some extent, and thus the results are more consistent when expressed in percentage terms—a system-wide reduction in delay of 17–20% from eliminating congestion effects at the NY airports.

The difference between the econometric and simulation results when expressed in absolute terms is nonetheless of some concern, particularly since SWAC is used for investment analysis. At face value, one might conclude that the delay reduction from eliminating congestion effects at NY airports is 3–4 times greater based on the regression model as compared to SWAC. This would have tremendous implications for most business cases involving NY airport investments. This difference could likely extend to other airports in the NAS as well. On the other hand, it may be that SWAC accurately predicts the impacts of the specific improvements it is intended to model, while the regression model captures additional congestion effects that cannot be eliminated by expanding arrival, departure, and terminal airspace capacity. Further research is required to understand the differing magnitudes of the SWAC and econometric results and what they imply about the benefits of expanding NAS capacity.

Both the econometric and the simulation models can be improved. On the econometric side more explicit attention should be given to how New York airspace constraints contribute to delay. The importance of the volume (OP_{NY}) term in the model likely reflects airspace constraints, but does so in a rather coarse manner. Additionally, decomposing the New York airport contribution by airport would be useful since it is possible that some of the airports have a stronger impact on the rest of the system than others. Finally, it would be desirable to consider each major airport region separately, so that the two-equation econometric model becomes a multi-equation model. Again, results of this analysis can be compared to what is obtained from the SWAC model or some enhanced version of it.

On the simulation side, the impacts of surface congestion and gate availability constraints could be more explicitly considered. Also, while the model captures forward propagation associated with aircraft, the impacts of flight and cabin crew-related propagation, and of payload propagation, have yet to be incorporated.

Nonetheless, this paper has established a conceptual and methodological approach to assessing how airports in a particular region contribute to delays in a larger airport system. This approach begins with a clear articulation of the counterfactual scenarios required to define this impact, and then moves to the modeling required to predict how operational performance would change under these counterfactuals. By instantiating this approach using a variety of counterfactuals and two different models, we can assess the extent to which our impact assessment depends on them.

In addition to the methodological contribution, our conclusions about the system impact of New York airports are also notable. Regardless of the model used or the counterfactual scenario assumed, we find that delay at these three airports has a modest impact on the rest of the system. The major investments that are being contemplated to increase capacity and reduce delays at NY airports will primarily affect delays at these airports; their impacts elsewhere should be viewed as a "frosting on the cake." Likewise, other forms of intervention to reduce delay—for example slot controls—should be viewed as primarily local matters. If there is a case for federal involvement, it should not be based on the spillover effects caused by delays at the NY airports. Conversely, it appears that the decisions could best be left to the airport operator (the Port Authority of New York and New Jersey) and flight operators along with other regional stakeholders. While these conclusions pertain specifically to the New York region, we suspect they apply to other regions as well, for which, unlike New York, claims that delay is a national problem are rarely heard.

References

AhmadBeygi, S., Cohn, A., Gua, Y., 2008. Analysis of the potential for delay propagation in passenger aviation flight networks. J. Air Transp. Manag. 14 (5), 280–286.

Beatty, R., Hsu, R., Berry, L., Rome, J., 1999. Preliminary evaluation of flight delay propagation through an airline schedule. Air Traffic Control Quart. 7, 259–270.

Boswell, S., Evans, J., 1997. Analysis of Downstream Impacts of Air Traffic Delay. Project Report ATC-257. MIT Lincoln Laboratory, Cambridge, MA.

Bureau of Transportation Statistics (BTS), 2014. Understanding the Reporting of Causes of Flight Delays and Cancellations. http://www.rita.dot.gov/bts/help/aviation/html/understanding.html (accessed 27.04.14).

Cheng, F., Gulding, J., Baszczewski, B., Galaviz, R., 2011. An optimization-based sample day selection algorithm for future schedule generation. In: Presented at the 11th Aviation Technology, Integration, and Operations (ATIO) Conference, Virginia Beach, VA, September 2011.

Diana, T., 2009. Do market-concentrated airports propagate more delays than less concentrated ones? A case study of selected U.S. airports. J. Air Transp. Manag. 15 (6), 280–286.

FAA, 2010. New York flight Delays Have Three Main Causes, but More Work Is Needed to Understand Their Nationwide Effect. Report Number AV-2011-007, October 28, 2010.

FAA, 2012. The Business Case for the Next Generation Air Transportation System. http://www.faa.gov/nextgen/library (August 2012).

GAO report, 2010. United States Government Accountability Office report: NATIONAL AIRSPACE SYSTEM, Setting On-Time Performance Targets at Congested Airports Could Help Focus FAA's Actions, May, 2010.

Green, W., 2000. Econometric Analysis, 4th ed. Prentice Hall.

Hansen, M., Zhang, Y., 2005. Operational consequences of alternative airport demand management policies: case of LaGuardia airport, New York. Transportation Research Record: Journal of the Transportation Research Board, No. 1915. Transportation Research Board of the National Academies, Washington, DC, pp. 95–104.

Hao, L., Hansen, M., 2012. Block time reliability and scheduled block time setting. In: Working Paper. Berkeley: University of California.

Laskey, K.B., Xu, N., Chen, C.H., 2006. Propagation of delays in the national airspace system. In: Presented at 22nd Conference on Uncertainty in Artificial Intelligence, Cambridge, Mass., July 2006.

Mason, J., 2013, A Rough Guide to the Jet Steam. https://skepticalscience.com/jetstream-guide.html (accessed 12.05.14).

Nayak, N., Zhang, Y., 2011. Estimation and comparison of impact of single airport delay on national airspace system with multivariate simultaneous models. Transportation Research Record: Journal of the Transportation Research Board, No. 2206. Transportation Research Board of the National Academies, Washington, D.C, pp. 52–60.

New York Area Program Integration Office (NYAPIO) report, 2010. What We Do. http://www.faa.gov/about/office_org/headquarters_offices/arc/programs/nyapio/index.cfm?content=what_we_do (July, 2010, last modified).

Pindyck, R.S., Rubinfeld, D.L., 1998. Econometrics Models and Econometric Forecasts. McGraw-hill.

Regional Planning Association, 2011, Upgrading to World Class: The Future of the New York Region's Airports.

Santos, G., Robin, M., 2011. Determinants of delays at European airports. Transp. Res. Part B 45 (1), 392–403.

Schaefer, L., Millner, D., 2001. Flight delay propagation analysis with the detailed policy assessment tool. In: Proc., 2001 IEEE Systems, Man and Cybernetics Conference, Tucson, Ariz., October, 2001.

Swaroop, P., Zou, B., Hansen, M., Ball, M., 2012. Do more US airports need slot controls? A welfare based approach to determine slot levels. Transp. Res. Part B 46 (9), 1239–1259.

Wang, P.T.R., Schaefer, L.A., Wojcik, L.A., 2003. Flight connections and their impacts on delay propagation. In: Digital Avionics Systems Conference, 2003. p. 5.B.4–5.1-9.

Zhang, Y., Nayak, N., 2010. Macroscopic tool for measuring delay performance in national airspace system. Transportation Research Record: Journal of the Transportation Research Board, No. 2177. Transportation Research Board of the National Academies, Washington, DC, pp. 88–97.

Zou, B., Hansen, M., 2012. Impact of operational performance on air carrier cost structure: evidence from US airlines. Transp. Res. Part E 48 (5), 1032–1048.