Flight Delays and Passenger Preferences: An Axiomatic Approach

John A. Bishop
East Carolina University
Economics Department
Greenville, NC 27858
bishopj@ecu.edu
252-328-6006

Nicholas G. Rupp*
East Carolina University
Economics Department
Greenville, NC 27858
ruppn@ecu.edu
252-328-6821

Buhong Zheng
University of Colorado - Denver
Economics Department
Denver, CO 80217-3364
buhong.zheng@ucdenver.edu
303-556-8375

20 November 2008

JEL Codes: L93, R42

Keywords: Flight delay curve, On-time performance, airport congestion

Abstract

The U.S. Department of Transportation (DOT) defines a flight as "delayed" if it arrives 15+ minutes late. The DOT "flight counting" delay definition is used to rank airline/airport service quality. An obvious caveat of counting flight delays is that the duration of delay plays no role in the delay count. We propose an aggregate delay measure that is sensitive to the distribution of time delayed among passengers. Moreover we derive a delay measure based on passenger preferences rather than its statistical properties. We model passengers preference ordering using the criteria: passengers prefer fewer, shorter, and more equal delay times.

*Corresponding author.

I. Introduction

Airline flight delays, like any other kind of waiting for service, may negatively affect customers (passengers) in many ways. Delays can increase passengers' anger, uncertainty and dissatisfaction with the service provided (Taylor 1994). In addition, flight delays are costly. A recent Joint Economic Committee report estimates that domestic flight delays cost the airline industry and passengers $40.7 billion in 2007.[1] In December 2007, U.S. airline delays reached their highest monthly level since the Bureau of Transportation Statistics began tracking flight delays in 1995 as 32 percent of domestic flights arrived late. Furthermore in 2007, U.S. airline delays reached their highest annual level since 1999, as 24 percent of all domestic flights arrived late. To address this problem, the FAA recently threatened to fine airlines with persistent delays.[2]

In ranking flight delays among airlines and airports, the sole (and official) measure used by the U.S. Department of Transportation is the proportion of flights delayed (i.e., a flight is counted as "delayed" if it arrives fifteen or more minutes behind schedule). This DOT "flight-counting" measure of delays has been adopted by the industry and is widely reported by the media as the de-facto standard to measure on-time performance. In fact, the DOT's *Air Travel Consumer Report* provides a monthly ranking of airlines based on the percentage of on-time arrivals.[3]

There are several flaws with using the DOT standard to measure airline service quality. Foremost is the arbitrariness in assigning fifteen minutes as the delay threshold. Why not ten minutes or twenty minutes? Second, by counting the occurrence of delays, the duration of delay

---

[1] See http://jec.senate.gov report release on 22 May 2008.

[2] For example, see http://www.aviation.com/business/071024-ap-fines-for-delays.html.

[3] The *Air Travel Consumer Report* is available online at: http://airconsumer.ost.dot.gov/

plays no role in the calculation (e.g., no distinction is made between flights delayed sixteen minutes vs. sixty minutes). An implication of using a counting measure for delays is that airlines have no incentive to shorten flight delays for flights that are already considered "delayed". Third, a discrete designation for each flight, either: "on-time" or "delayed" ignores the distribution of flight delays. Even carriers with identical average minutes of delay are likely to be viewed differently if they provide some passengers with severe delays. We believe that extreme delays are viewed as particularly upsetting for travelers (i.e., a one hour delay is more painful for travelers than two thirty minute delays).

Since airline researchers began recognizing the statistical shortcomings of the fifteen minute delay standard; hence they have used a variety of flight delay measures including: counting the number of flight delays (Brueckner 2002), calculating the minutes of travel time on a route in excess of the monthly minimum (Mayer and Sinai 2003), determining the minutes of arrival (Mazzeo 2003) and departure delay (Rupp 2008). Moreover, Bratu and Barnhart (2006) show that when factors such as flight cancellations and missed connections are factored in, actual passenger waiting times are nearly two-thirds higher than minutes of aircraft arrival delay (the DOT reported measure). The unique contribution of our paper is that we derive a delay measure based on passenger preferences, not simply based on a measure's statistical properties. Of course, any measure of airline delays must assert a passenger preference ordering; we model passengers as preferring fewer, shorter, and more equal delay times.

The paper is organized as follows. Section II provides the axiomatic framework for measuring aggregate flight delays. We examine the notion of flight delay and propose a set of axioms governing the measurement of flight delays for a group of airline (or airport) passengers. We then propose a class of decomposable measures of flight delays as well as a partial dominance condition for the rankings of flight delays. In Section III, we apply the proposed measures and dominance condition to measure and rank flight delays of two major U.S. airlines. Section IV provides some extensions and discussion.

II. Measuring Aggregate Flight Delays

Consider a group of N passengers with different delay times x_i, i = 1, 2, …, n. Here the group can be viewed as all passengers of an airline or an airport. Clearly, not all passengers have their flights delayed; some may even depart and arrive early. In this sense, x_i can be positive (delayed), negative (arrived early), or zero (on time). For the group as a whole, we denote $X = (x_1, x_2, ..., x_N)$ as the *flight-delay profile* of the group.

For the passengers as a group, we want to construct a summary measure of delays so that comparisons and rankings among different groups of passengers are feasible. To this end, we define a *measure of flight delays* as a single value function,

$$D = D(x_1, x_2, ..., x_N)$$

which reflects the aggregate level of flight delays for the group as a whole. To characterize $D(\cdot)$, we follow the axiomatic approach that Sen (1976) pioneered in poverty measurement. The similarity between these two measurements suggests that much of the calibration crafted to measure poverty can be applied when measuring flight delays. In this approach, we first lay out the basic ideal properties that an index of flight delays should possess and then generate satisfactory flight-delay measures within the boundaries of the axioms.

II.1. Axioms on $D(\cdot)$

We first require that the flight-delay index be a continuous function of all flight-delay times.

Continuity: $D(\cdot)$ is continuous function of $X = (x_1, x_2, ..., x_N)$.

The second axiom is the anonymity axiom which states that the identities of the passengers play no role in the computation of $D(\cdot)$: if two populations have the same flight-delay profile, then the two groups should have the same level of flight delays. Profiles $X = (x_1, x_2, ..., x_N)$ and $Y = (y_1, y_2, ..., y_N)$ have the same level of flight delay if $Y = PX$ for some permutation matrix P. A permutation matrix is a square matrix with elements zero and one where each row and column sums to one. Formally, the anonymity axiom is stated as follows:

Anonymity: $D(Y) = D(X)$ if $Y = PX$ for some permutation matrix P.

The next axiom is the focus axiom which states that an index of flight delays is only concerned with delays; hence arriving early by twenty minutes or by two hours make no difference for the calculation of $D(\cdot)$. That is, recalling that early arrival means $x_i < 0$, in the following statement, an increase in the early arrival time x_i by some ε_i to $y_i = x_i - \varepsilon_i$ has no effect on $D(\cdot)$.

Focus: $D(Y) = D(X)$ if Y is obtained from X via $y_i = x_i$ for all $x_i > 0$ and $y_i \le x_i$ for all $x_i \le 0$.

Contrary to an early arriving flight, if a flight has been delayed, then any further delay will increase the level of aggregate delays. This is the monotonicity axiom to which we alluded earlier in the introduction. In the following statement, a passenger's delay time increases from x_i to $y_i = x_i + \varepsilon_i$.

Monotonicity: $D(Y) > D(X)$ if Y is obtained from X via $y_i = x_i + \varepsilon_i$ for some $x_i > 0$ with some $\varepsilon_i > 0$ and $y_i = x_i$ for all other $x_i > 0$.

While an index $D(\cdot)$ that satisfies the monotonicity axiom reflects the length of a passenger's delay, it may not address the distribution of delays among passengers. To put the necessity of this concern into perspective, consider a total delay of one hour between two flights with an equal number of passengers on a route. In one case, every flight is delayed by thirty minutes, whereas in the other case the outcome alternates between arriving on-time and arriving one hour late. Which case should be considered to have a higher level of passenger flight delays?

A passenger may not mind a delay of ten, twenty or even thirty minutes, but anger, anxiety, uncertainty and boredom mount at an increasing rate as delay prolongs. In this sense, the overall problem of delays in the first case may be considerably smaller compared to the second case. For example, in February 2007, JetBlue Flight 751 was stranded at JFK Airport for more than ten hours. This flight delay would never have become front-page news if JetBlue had evenly distributed ten hours of delay over ten JetBlue flights. Stranded passengers become particularly unhappy when they have to make tight connections, or even worse, miss their connecting flights.

The general idea that spreading the total delay time more evenly across all passengers (or flights) leads to a lower level of aggregate delay can be imposed as an axiom on $D(\cdot)$. In the following statement, passenger *s* experiences a longer delay than passenger *t* ($x_s > x_t > 0$) and from *X* to *Y* passenger *s*'s delay is shortened by ε while *t*'s delay is prolonged by ε (all other passengers' delays are not affected).

Distribution Sensitivity: $D(Y) < D(X)$ if *Y* is obtained from *X* via (1) $y_s = x_s - \varepsilon$, and $y_t = x_t + \varepsilon$ for some $x_s > x_t > 0$ and for some $\varepsilon > 0$ such that $y_s > y_t > 0$; and (2) $y_i = x_i$ for all $i \neq s,t$.

The next axiom that we will impose on $D(\cdot)$ enables the comparison of flights delays between different airlines (or airports) where the number of passengers may differ. The following axiom states that if an airline expands through a simple replication, then the level of flight delays remains unchanged.

Replication Invariance: $D(Y) = D(X)$ if Y is obtained from X via a simple replication, i.e., $Y = (X, X, ..., X)$.

Finally, we introduce a consistency requirement that enables the ranking of flight delays to be independent of the measuring units of time, (e.g., minutes vs. hours).

Unit Consistency: If $D(Y) > D(X)$ then $D(\theta Y) > D(\theta X)$ for all $\theta > 0$.

This last axiom says that if the flight-delay profile Y exhibits more aggregate delay than X when time is measured in minutes, then the conclusion (ranking) remains the same if time is measured in hours or any other units.

II.2. The Implications of the Axioms and Some Examples of $D(\cdot)$

The anonymity axiom implies that we can consider an ordered profile of flight delays, i.e., for each $X = (x_1, x_2, ..., x_N)$ we can assume that $x_1 \geq x_2 \geq ... \geq x_N$. The focus axiom implies that for those passengers whose flights are not delayed (i.e., $x_i \leq 0$), $D(\cdot)$ does not depend upon the specific values of x_i. It follows that we can set all those negative values of x_i to zero – $D(\cdot)$ does not distinguish between those passengers who arrived early and those arriving on time. For each profile X, the anonymity axiom and the focus axiom together allow us to consider the censored profile $\tilde{X} = (\tilde{x}_1, \tilde{x}_2, ..., \tilde{x}_N)$ which sets every negative x_i to zero, i.e., $\tilde{x}_i = \max(x_i, 0)$ for $i = 1, 2, ..., N$, and $\tilde{x}_1 \geq \tilde{x}_2 \geq ... \geq \tilde{x}_N$.

Using our notation, the official measure of aggregate flight delays is

$$D_1(X) = \frac{1}{N}\sum_{i=1}^{N} I(x_i) \qquad (2.1)$$

where $I(x_i)$ is an indicator function which equals one if $x_i > 0$ and zero otherwise. This flight-counting index satisfies only anonymity, replication invariance, and unit consistency. It violates continuity at the point $x_i = 0$ since for any flight with delay – no matter how slight (i.e., x_i is close to zero) – it is counted as one in $D_1(X)$, however, if the delay time is zero then the flight is counted as zero. This problem may be even more intensified with the ambiguity about what constitutes a "delay"? (i.e., how many minutes must the flight be late to be considered "delayed"?).

More importantly, the flight-counting measure violates the monotonicity axiom and the distribution sensitivity axiom. As mentioned in the introduction, the violation of monotonicity implies that once a flight is deemed "delayed" the airline has no incentive to shorten the delay as far as minimizing $D_1(X)$ is concerned. In fact, the airline may have an incentive to prolong the flight delay in order to get other flights on time so that $D_1(X)$ becomes smaller. The violation of the distribution sensitivity means that whether the total delay time is spread evenly among passengers (flights) or is concentrated among a few passengers/flights matters little to the picture that $D_1(X)$ portrays.

A measure of flight delays which is a modest improvement over $D_1(X)$ would be the following average-time-delayed measure

$$D_2(X) = \frac{1}{N}\sum_{i=1}^{N} x_i I(x_i) = \frac{1}{N}\sum_{i=1}^{N} \tilde{x}_i \ . \qquad (2.2)$$

Compared with $D_1(X)$, the (normalized) average-time-delayed measure $D_2(X)$ satisfies continuity, anonymity, monotonicity and replication invariance, however, it violates the distribution sensitivity axiom. To allow any prolonged delay (i.e., the JetBlue JFK case) to be weighted more than just another delay in the calculation of aggregated delays, $D(X)$ must reflect the axiom of distribution sensitivity.

A measure that satisfies all aforementioned axioms is easy to construct. In fact, we propose a class of such measures. Consider a continuous, increasing, and convex function $\phi(x)$ with $\phi(0)=0$, a member of the class is

$$D_{\phi}(X)=\frac{1}{N}\sum_{i=1}^{N}\phi[x_i I(x_i)]\,. \tag{2.3}$$

It is easy to verify that $D_{\phi}(X)$ satisfies all axioms examined above except the unit consistency axiom. To satisfy unit-consistency, function $\phi(x)$ must also be homogenous (Zheng 2007). An example of the satisfactory $\phi-$function is $\phi(x)=x^{\alpha}$ with $\alpha>1$.

The measures defined in (2.3) are decomposable in the sense that the overall level of flight delays can be written as a (weighted) average of all subgroups' level of delays. This decomposability property is very useful in that it identifies the contribution of the delay from each subgroup (an airline or an airport) to the overall delay of the industry.

II.3. Flight-Delay Dominance

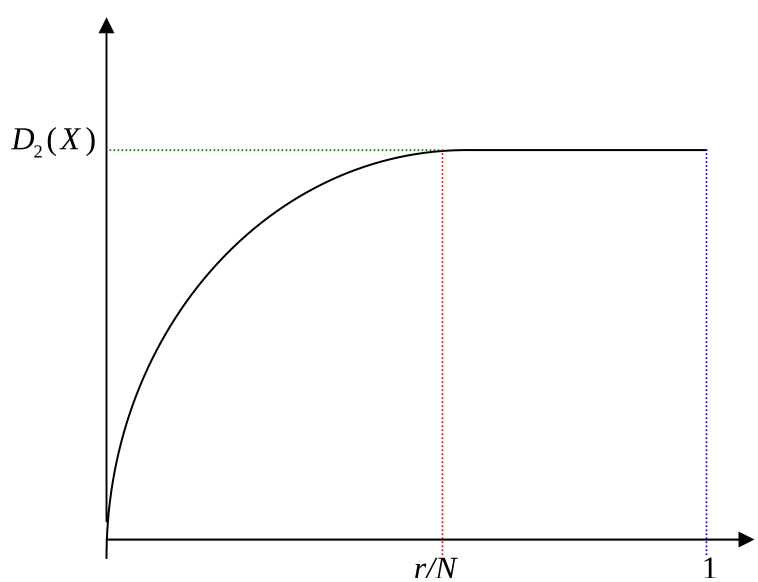
For each $\phi(x)$, we can calculate the corresponding flight-delay measure for each airline or airport. Then we can compare these flight-delay measures among airlines and airports to rank them from the most to the least delayed services. Clearly, the choice of the function $\phi(x)$ is consequential: different functions may lead to different rankings. A natural and important question is under what conditions can we rank one airline as having a higher level of flight delays than another airline for all possible functions $\phi(x)$? In this section, we establish a partial ordering condition and provide a device to enable this unanimous comparison.

Recall that if all measures satisfy anonymity and the focus axiom, then we can consider a censored and decreasingly ordered version of each flight delay profile. Relying on a censored and sorted flight delay profile $\tilde{X}=(\tilde{x}_1,\tilde{x}_2,...,\tilde{x}_r,0,...,0)$ where *r* is the number of passengers delayed, we can construct a flight-delay curve as follows. For each passenger *i* in the sorted profile, we first calculate

$$C(X;i) = \frac{1}{N}\sum_{j=1}^{i}\tilde{x}_j \ . \qquad (2.4)$$

That is, $C(X;i)$ cumulates the first *i* longest delays: $C(X;1) = \frac{\tilde{x}_1}{N}$, $C(X;2) = \frac{\tilde{x}_1 + \tilde{x}_2}{N}$, $C(X;3) = \frac{\tilde{x}_1 + \tilde{x}_2 + \tilde{x}_3}{N}$, … Next, we plot the sequence $\{C(X;i)\}$ against the corresponding cumulative passenger proportion $\{\frac{i}{N}\}$ in a graph with $\frac{i}{N}$ on the horizontal axis and $C(X;i)$ on the vertical axis. The following graph depicts such a curve which is referred to as the *flight-delay curve*.

Figure 1



The flight-delay curve in Figure 1 is concave up to the point $\{r/N, D_2(X)\}$ and then it becomes flat since $x_i = 0$ for $i > r$.

With the flight-delay curve, we can define our *partial flight-delay dominance* relationship as follows: for two flight delay profiles *X* and *Y* with the same number of passengers *N*, *X flight-delay dominates Y* if [4]

$$C(X;i) \le C(Y;i) \tag{2.5}$$

for all $i = 1, 2, ..., N$ and the strict inequality holds for some *i*. Graphically, (2.5) says that the flight-delay curve of *X* lies nowhere above that of *Y* and strictly below over some range.

The important result of this section is the following equivalence between the partial flight-delay dominance and the rankings by all members of flight-delay class of (2.3).

Proposition 1. For any two flight-delay profiles *X* and *Y*, the following two conditions are equivalent:

(1) $D_\phi(X) \le D_\phi(Y)$ for all members of $D_\phi(\cdot)$ and $D_\phi(X) < D_\phi(Y)$ for some members of $D_\phi(\cdot)$;

(2) The flight-delay curve of *X* dominates that of *Y*.

Proof. See Jenkins and Lambert (1998) where the context is of poverty gaps and the curve is known as the TIP curve.

This proposition also has an important implication for ranking flight delays when different cutoffs are used to define what is considered "being delayed." Up to this point in our theoretical calibration of measurement, we have assumed that a flight is delayed as long as it is later than scheduled. Now suppose that there are two definitions of delay: one is *s* minutes behind schedule and the other is *t* minutes behind schedule with $0 < s < t$. For example, in our empirical illustration below we consider both five minutes and fifteen minutes delay cutoffs. An

[4] Since $C(X;i)$ satisfies the replication invariance axiom, dominance relation (2.5) can be defined similarly for flight delay profiles with different numbers of passengers.

interesting question to ask is: if one airline has less aggregate delay than another airline when *s*-minute delay cutoff is used, will the airline also have less delay when a *t*-minute delay cutoff is used instead? The following corollary provides a useful guideline for delay comparisons with different delay cutoffs.

Corollary 1. For any two flight-delay cutoffs *s* and *t* with $s < t$, and two pairs of flight-delay profiles (X_s, Y_s) and (X_t, Y_t), if the flight-delay curve of X_s dominates that of Y_s then the flight-delay curve of X_t dominates that of Y_t.

Proof. The proof of this result can also be found in poverty ordering literature (again, see Jenkins and Lambert 1998). Note that increasing the delay cutoff has the same effect as lowering the poverty line. It is a known result in poverty measurement that if one distribution has less poverty than another distribution for all poverty measures at a given poverty line, then the conclusion holds for all lower poverty lines.

From this corollary, it follows that if JetBlue has less aggregate delay than US Airway (i.e., the flight-delay curve of JetBlue lies below that of US Airway) for the 5-minute delay cutoff, then we can be certain without checking that JetBlue will also have less delay than USAir for any higher delay cutoffs (ten minutes, fifteen minutes, …).

II.4. A Gini-type Measure of Flight Delays

The flight-delay curve lends directly to a Gini-type measure of flight delays. The measure is simply equal to the area beneath the flight-delay curve which is

$$\begin{aligned} D_g(X) &= \frac{1}{2}\left(\frac{1}{N}\right)C(X;1) + \frac{1}{2}\left(\frac{1}{N}\right)[C(X;1)+C(X;2)] + \ldots + \frac{1}{2}\left(\frac{1}{N}\right)[C(X;N-1)+C(X;N)] \\ &= \frac{1}{4N^2}\sum_{i=1}^{N}(N-i+1)(N-i+2)\tilde{x}_i \end{aligned} \tag{2.6}$$

Note that this measure is not decomposable in the sense that we defined above. The Gini-type measure reflects a unique passenger preference about flight delays. In this measure, a passenger cares not only about his/her time delayed but also about the relative position in the delay profile (i.e., how many people have less delay time than the passenger). See Lambert (2001, pp. 122-123) for more detailed discussion on the Gini-type preference in social welfare measurement.

III. An Illustration of the Flight Delay Curve

In this section we apply the flight delay curve developed above to actual flight delay data from July 2005.[5] To illustrate our approach we use Bureau of Transportation Statistics on time performance data for every domestic flight for two carriers, JetBlue and US Airways during the first week of July 2005.[6] We begin by plotting the distribution of arrival delays for every JetBlue and US Airways flight in July 2005 (see Figures 2a and 2b). These figures reveal a wider distribution of arrival delays for JetBlue.

Table 1 provides simple delay counts (standard errors and test statistics) for the two carriers for two time periods in 2005: July 1-7 and July 1-4, and six alternative delay cutoffs. We begin with the DOT definition of a flight "delay" (i.e., flights arriving fifteen or more minutes later than scheduled). For the seven day period we find that JetBlue (29.37%) has significantly fewer official delays than US Airways (31.84%) (z-score = 2.24). For the four day sample we find no significant difference (30.86% vs. 29.96%) in the official delay rate (z-score = 0.62).

---

[5] We select this month since it had the highest proportion of flight delays in 2005.

[6] We exclude both diverted and canceled flights since the length of flight delay is ambiguous. Just 2.2% of JetBlue and US Airways domestic flights in July 2005 were diverted or canceled.

The natural question to ask is: Do these official delay rates accurately describe the two carriers' delay distributions? Our answers are: *perhaps* and *not at all.* To arrive at these conclusions we must first examine the test statistics at all possible delay times. In the seven day case (see Table 1), US Airways has significantly higher delay rates than JetBlue for all delays that exceed ten minutes. We note that for five and ten minute delay thresholds the two carriers have delay rates that are not significantly different.

Figure 3 illustrates the July 1-7, 2005 delays where ten minutes serves as the delay threshold. This figure provides the flight delay curves for JetBlue and US Airways. On the *x*-axis we plot the cumulative proportion of passenger flight delays, beginning with the longest delay. The incidence of delay is given by the length of the flight delay curve's non-horizontal section. As noted in the Table 1 using a ten minute definition for flight delays, the delay rate for both carriers is slightly over 36 percent during the first week of July, 2005. After this point, both curves in Figure 3 become horizontal, denoting an on time arrival using the ten minute standard.

On the y-axis we plot the intensity of delay. The vertical intercept at $p = 1$ is the aggregate delay gap, $D_2(X)$, averaged across all of a carriers' flights. The average delay gap would then be equal to the slope of the ray from the origin to the point where the flight delay curve initially goes horizontal (here at 0.36). Figure 3 shows that JetBlue has a smaller aggregate (and average) delay rate (0.047) than does US Airways (0.051), for the period July 1-7.

The inequality dimension of flight delays is summarized by the degree of concavity of the non-horizontal section of the flight delay curve. If there is equality of delays among the delayed flights, i.e., if the delay gaps were equal, then the ray from the origin would be a straight line with slope equal to z (ten minutes, in this case) minus the average delay time. As noted above the flight delay curve combines all three elements: delay rate, delay gap, and delay inequality.

Returning to Figure 3 we see that the JetBlue *flight-delay curve dominates* US Airways since its flight delay curve (the solid line) lies everywhere inside the equivalent curve for US Airways (the dashed line). Thus, in this case the industry's fifteen minute delay standard (US Airways 31.84% vs. JetBlue 29.37%) gives the correct ordinal delay ranking of these two carriers for all delay measures above ten minutes.

To further illustrate the usefulness of the flight delay curve we consider an alternative time frame for our sample of flights: July 1st through July 4th. Recall that for the fifteen minute delay standard we find no significant difference in delay rates between JetBlue and US Airways. Using a ten minute delay threshold, however, we find that US Airways has a smaller delay rate than JetBlue at the ten percent significance level (z-score = 1.83). Furthermore, for a five minute delay threshold, US Airways has a significantly lower delay rate (z-score=3.02). In contrast, as the delay window is expanded (beyond twenty minutes) we find that JetBlue now has significantly lower delay rates. Clearly, the fifteen minute standard—in case that there is no difference between carriers--does not adequately describe the distributions of flight delays.

Figure 4 presents the flight delay curves for July 1-4 using five minutes as the delay threshold. The first dimension of flight delay preferences, the delay rate, is shown on the horizontal axis. We observe that the US Airways flight delay curve (the dashed line) becomes horizontal at a lower delay rate than does JetBlue's flight delay curve, which reflects US Airways' lower delay rate at five minutes.

The second dimension of flight delay preferences, the intensity of flight delays (i.e., the slope of the ray from the origin where the flight-delay curve becomes horizontal), is shown on the vertical axis of Figure 4. Here we see that JetBlue has the lower aggregate delay rate (0.139 versus 0.148). This example provides a clear conflict between the preference for fewer versus shorter delays. The third dimension of delay preferences, the inequality among flight delays, is reflected in the greater concavity of the flight delay curves. In this example the US Airways flight delay curve shows a larger degree of delay inequality (i.e., greater concavity). In sum, any conflict between passenger preferences (for fewer, shorter, and more equal delays) will result in crossing flight delay curves, as clearly seen in Figure 4. Crossing flight delay curves prohibit an ordinal ranking of carrier flight delays.

There are at least two possible solutions to the delay ambiguity shown in Figure 4. The first approach is to propose a cardinal delay preference function that specifies a tradeoff between the number of flight delays, the length of delays, and the equality of delays. An example of a cardinal preference function is the well-known Gini index of inequality described above.[7] The Gini-type indexes, which reflect the area under the flight-delay curves, are reported in the figure notes. For Figure 4, the Gini-type indexes are 0.0519 for JetBlue and 0.0505 for US Airways. Thus, passengers with Gini-type preferences will prefer US Airways to JetBlue.[8] A second solution is to expand the delay window and check for an ordinal ranking of carriers. Figure 5 illustrates the second option using a ten minute (instead of a five minute) delay window. In this case, JetBlue's flight delay curve lies everywhere below US Airways flight delay curve, implying that passengers will prefer JetBlue to US Airways.

---

[7] In the poverty context, Jenkins and Lambert (1998) note that the preference tradeoffs embodied in the TIP Gini (our flight delay Gini) are equivalent to the modified-Sen index proposed and discussed by Shorrocks (1995).

[8] However, crossing flight delay curves imply that an alternative index can be proposed that reverses this ranking.

IV. Conclusion

Airline economists are well aware of the caveats with using fifteen minutes as a delay standard; hence a variety of alternative flight delay measures have been used in the literature. The unique contribution of our paper is the derivation of a delay measure which is based on passenger preferences, not an arbitrary cut-off decided by the Department of Transportation. We propose a delay ordering based on three widely acceptable preferences--passengers will prefer a carrier that provides fewer, shorter, and more equal delay times. Based on these three preference assumptions we propose the flight delay curve and identify the conditions under which an unambiguous ordering of carriers can be identified. Given the generality of our preference assumptions, the flight delay curve provides only a partial ordering of carriers. In the case of ‘crossing’ flight delay curves we offer several possible solutions.

We illustrate the flight delay curves using actual flight delay data for July 2005. One limitation of this research on passenger delay preferences is that we employ aircraft delay data, rather than the preferred measure of actual passenger delay data. For passengers travelling on non-stop itineraries, these two delay measures are equivalent. For passengers who make connections, however, an aircraft delay can lead to a missed connection. Hence aircraft delays understate passenger delays. We are limited to the publically available DOT data, which only provides information on aircraft delays, rather than passenger delays. Thus, one avenue for future research is to estimate flight delay curves from actual passenger delays.

Our empirical findings suggest that for longer time frames (i.e., a week or a month) aggregate measures of flight delays like the DOT delay definition (proportion of flights delayed fifteen minutes or more) are fairly representative of on-time performance. When we examine shorter time periods, however, the DOT delay definition is less representative of the distribution of flight delays; and hence the flight delay curves provide valuable information that reflect passenger preferences.

V. References

Bikker, Jacob A. and Katharina Haaf. 2002. “Measures of Competition and Concentration in the Banking Industry: A Review of the Literature,” *Economic and Financial Modelling*, 9, p. 53- 98.

Bratu, Stephane and Cynthia Barnhart. 2006. “Flight Operations Recovery: New Approaches considering passenger recovery.” *Journal of Scheduling*, 9:3, p. 279-98.

Brueckner, Jan K. 2002. “Airport Congestion when Carriers have Market Power”, *American Economic Review*, 92:5, p. 1357-75.

Hannan, Timothy. 1997. “Market Share Inequality, the Number of Competitors, and the HHI: An Examination of Bank Pricing,” *Review of Industrial Organization*, 12:1, p. 23-35.

Jenkins, S. and P. Lambert. 1998. “Three I's of Poverty Curves and Poverty Dominance: TIPs for Poverty Analysis”, *Research on Economic Inequality* 8, Greenwich: JAI Press, 39-56.

Lambert, P. 2001. The Distribution and Redistribution of Income, 3rd edition (The Manchester University Press, Manchester and New York).

Mayer, Christopher and Todd Sinai. 2003. “Network Effects, Congestion Externalities, and Air Traffic Delays: or Why All Delays Are Not Evil”, *American Economic Review*, 93:4, p. 1194-1215.

Mazzeo, Michael J. 2003. “Competition and Service Quality in the U.S. Airline Industry”, *Review of Industrial Organization*, 22:4, p. 275-96.

Rhoades, Stephen A. 1995. “Market Share Inequality, the HHI, and Other Measures of the Firm-

Composition of a Market," *Review of Industrial Organization*, 10:6, p. 657-674.

Rupp, Nicholas G. 2007. "Do Carriers Internalize Congestion Costs? Empirical Evidence on the Internalization Question" East Carolina University, Economics Dept. working paper #0713.

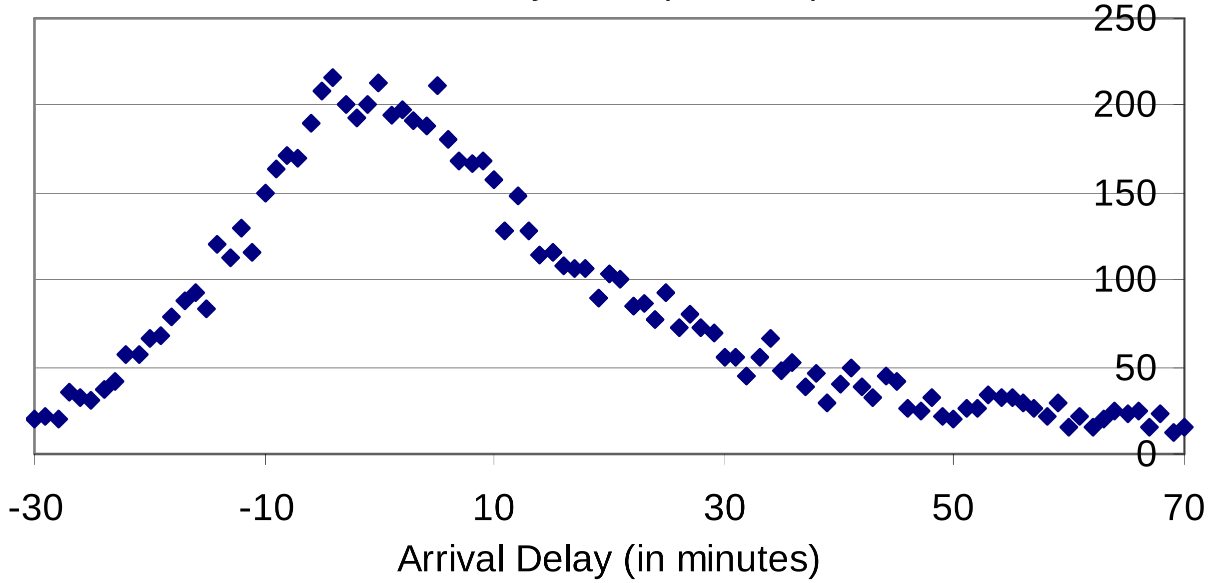
Sen, A. 1976. "Poverty: An Ordinal Approach to Measurement", *Econometrica* 44, 219-231.

Taylor, Shirley. 1994. "Waiting for Service: The Relationship between Delays and Evaluations of Service", *Journal of Marketing* 58, No. 2, 56-69.

Zheng, Buhong. 1997. "Aggregate Poverty Measures", *Journal of Economic Surveys* 11, 123-62.

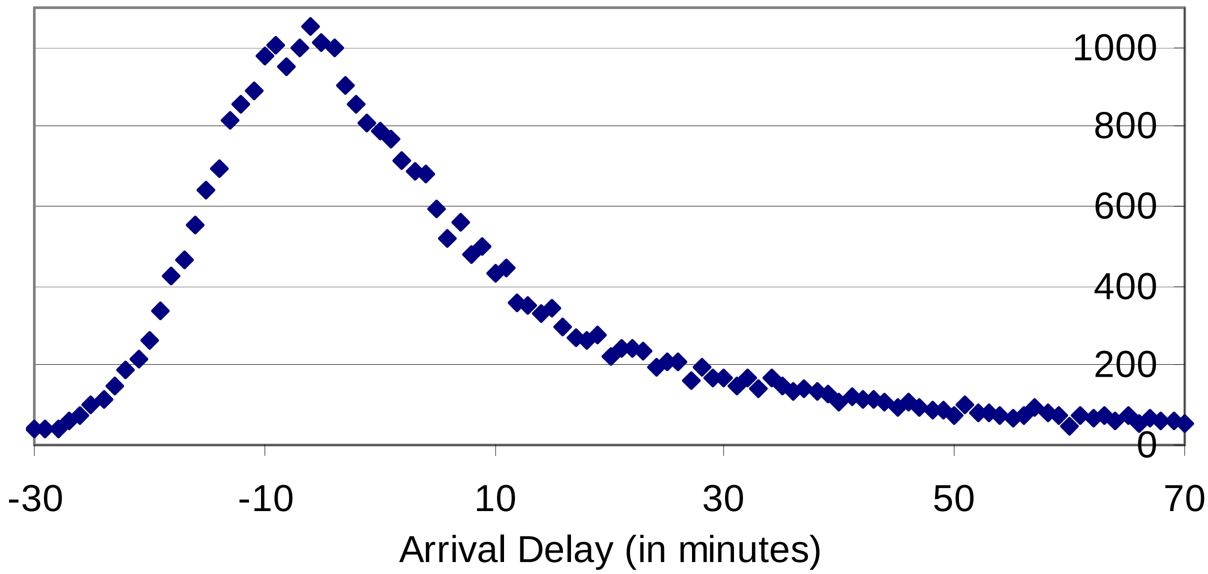
Zheng, Buhong. 2007. "Unit-Consistent Poverty Indices", *Economic Theory* 31, 113-142.

Figure 2a: Distribution of Arrival Delays for all JetBlue Flights[1] in July 2005 (n=9,546)



[1] 148 flights arrived more than 30 minutes early; 881flights arrived 71+ minutes late in July '05.

Figure 2b: Distribution of Arrival Delays for all US Airways Domestic Flights[1] in July 2005 (n=36,707)



[1]105 flights arrived more than 30 minutes early; 3,392 flights arrived 71+ minutes late July '05.

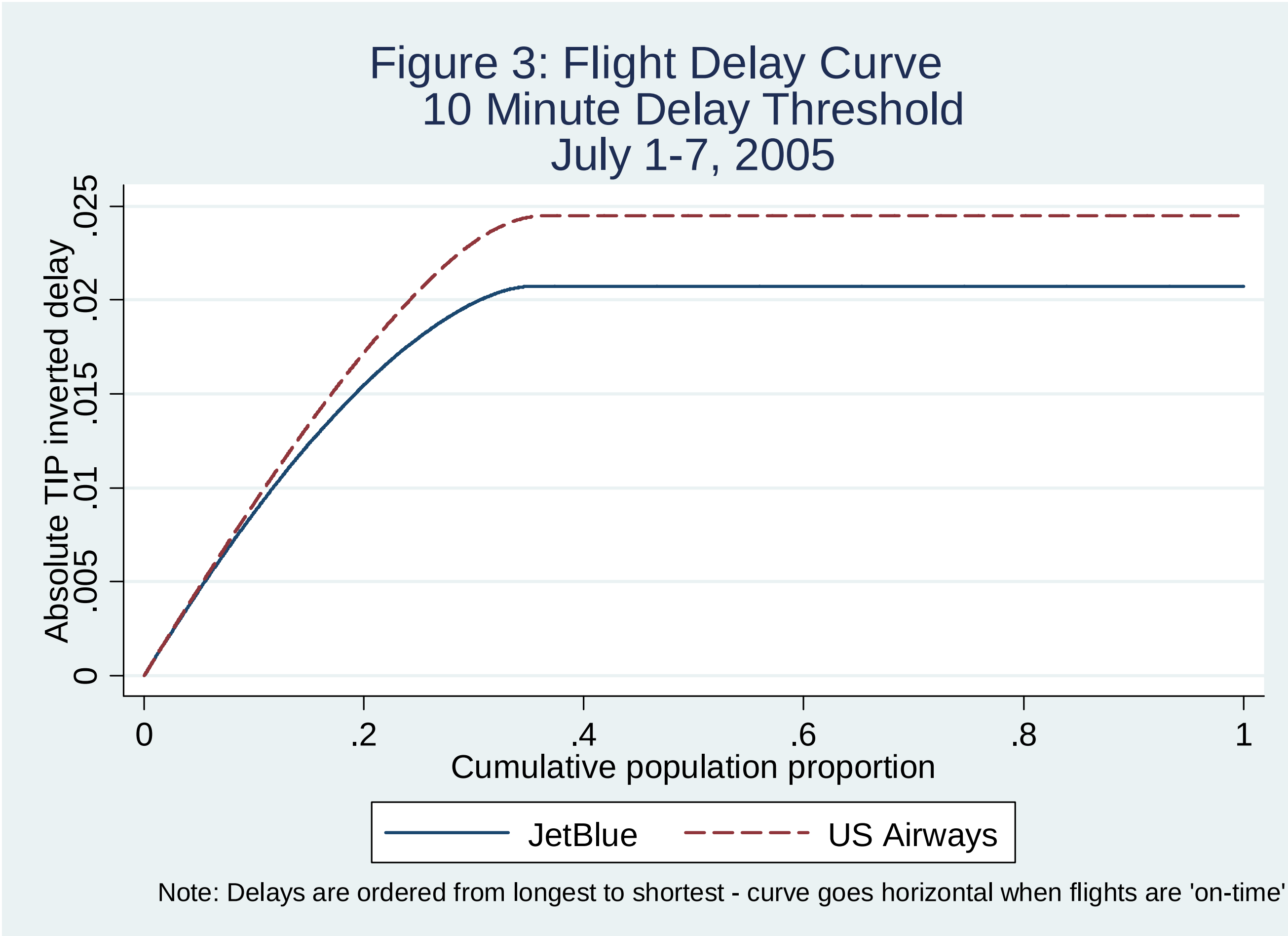


Figure 3: Flight Delay Curve
10 Minute Delay Threshold
July 1-7, 2005
Absolute TIP inverted delay
0
.005
.01
.015
.02
.025
0
.2
.4
.6
.8
1
Cumulative population proportion
JetBlue
US Airways
Note: Delays are ordered from longest to shortest - curve goes horizontal when flights are 'on-time'

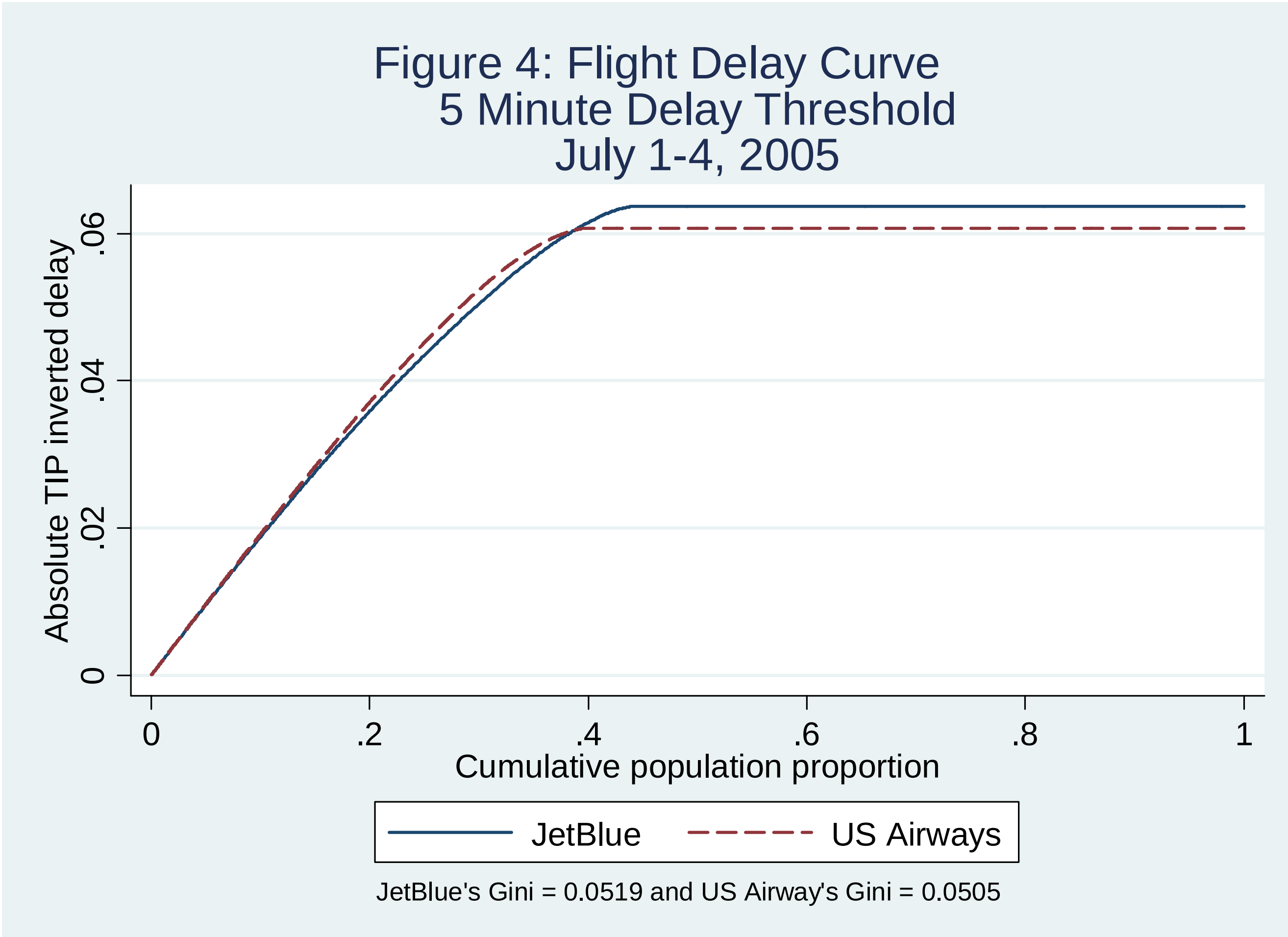


Figure 4: Flight Delay Curve
5 Minute Delay Threshold
July 1-4, 2005
Absolute TIP inverted delay
0
.02
.04
.06
0
.2
.4
.6
.8
1
Cumulative population proportion
JetBlue
US Airways
JetBlue's Gini = 0.0519 and US Airway's Gini = 0.0505

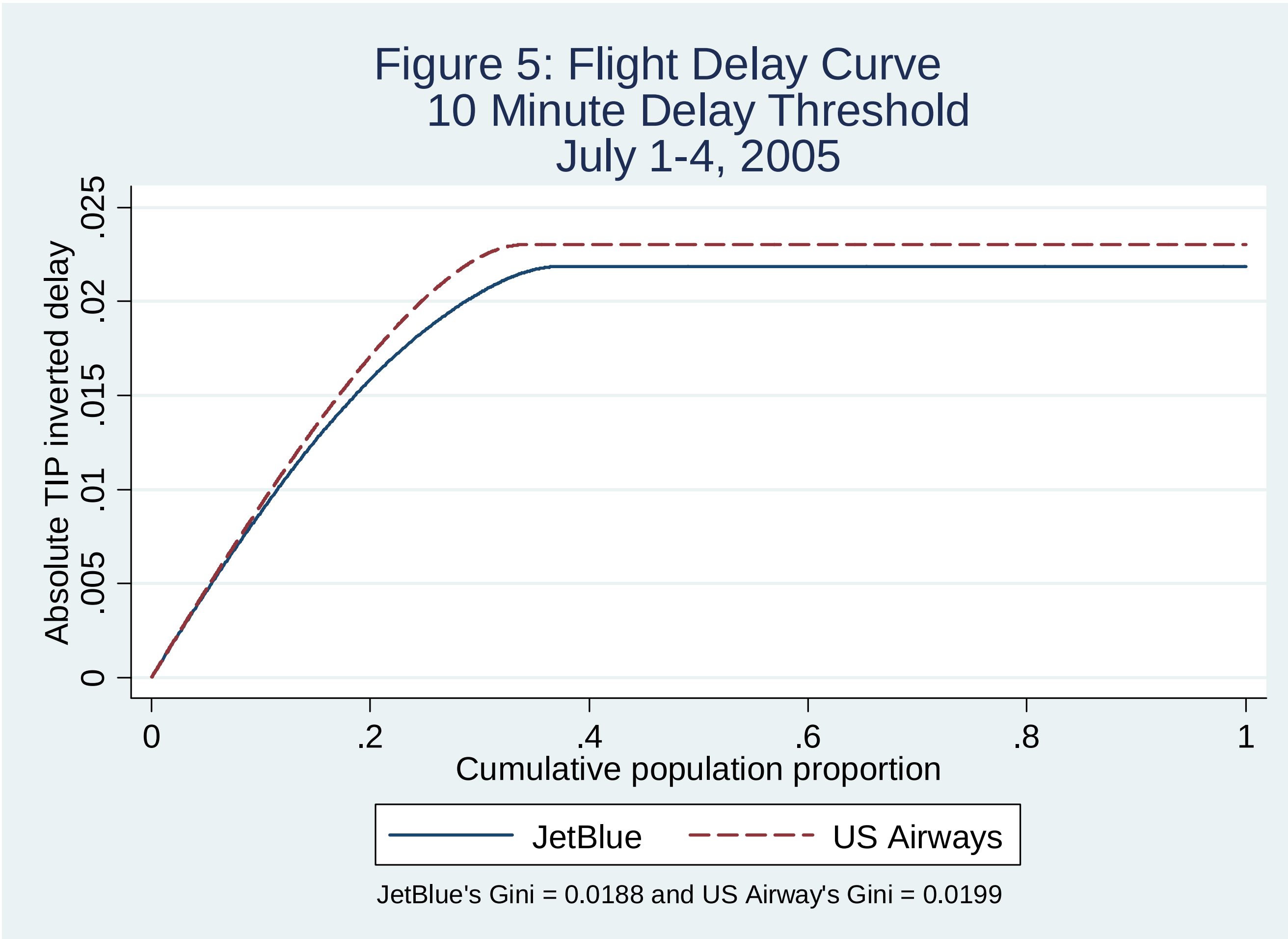


Figure 5: Flight Delay Curve
10 Minute Delay Threshold
July 1-4, 2005
Absolute TIP inverted delay
0
.005
.01
.015
.02
.025
0
.2
.4
.6
.8
1
Cumulative population proportion
JetBlue
US Airways
JetBlue's Gini = 0.0188 and US Airway's Gini = 0.0199

Table 1: Proportion of Flights Delayed

	July 1 - 7, 2005			July 1 - 4, 2005		
Minutes	Jet Blue	US Airways	Z - Score	Jet Blue	US Airways	Z - Score
5	0.4541	0.4363	1.43	0.4580	0.4091	3.02
	0.0111	0.0056		0.0142	0.0077	
10	0.3636	0.3681	-0.38	0.3747	0.3460	1.83
	0.0104	0.0054		0.0138	0.0074	
15	0.2937	0.3184	-2.22	0.3086	0.2996	0.62
	0.0098	0.0053		0.0132	0.0071	
20	0.2326	0.2821	-4.75	0.2392	0.2609	-1.55
	0.0091	0.0051		0.0122	0.0068	
30	0.1548	0.2250	-7.71	0.1657	0.2111	-3.67
	0.0078	0.0047		0.0106	0.0064	
45	0.1026	0.1694	-9.37	0.1200	0.1591	-3.58
	0.0066	0.0027		0.0093	0.0057	
# of Flights	2145	7791		1225	4116	