A Nested Logit Approach to Airline Operations Decision Process*

Junhua Yu
Department of Economics
East Carolina University
June 24th 2003

Abstract. This study analyzes the role of logistical variables, competition characteristics and weather conditions in a decision model for flight operations by airlines. We demonstrate that a nested logit model is appropriate on both theoretical and empirical grounds. The sample consists of all flights on Fridays in January 2003. We find that there are positive and significant correlation between daily flight frequencies and cancellation rates. The findings revealed that competition does not necessarily lead to better performance and the flights to and from the carrier's hub are less likely to be cancelled. Sever weather is a significant contributor to worse on-time performance.

Keywords: airline decision; on-time performance; nested logit model;

---

* I am especially indebted to Dr. Rupp for excellent research guidance. I also wish to thank Dr. Holmes for helpful suggestions and Dr. Bishop and Dr. Zeager for great support.

1. Introduction

The service quality of airlines in terms of on-time performance has become a growing concern among the U.S. government, media and flying passengers. Flights on major airlines arrived within 15 minutes of schedule 84.9% of the time in January 2003, an improvement from 78.3% in December 2002. However, passengers still stood a 15% chance that their flight was delayed, cancelled or diverted. It is worthwhile to get behind the dataset to understand what factors are contributing to poor airline performance before making policy-related decision in the airline industries.

This study explores potential determinants of the on-time performance decisions by airlines in a nested logit model approach. This research attempts to examine the decision mechanism of the airlines in flight operations. Previous related work on service quality usually focused on one or two of the outcomes. Mayer and Sinai (2002) report that hub carriers have fewer cancellations and longer delays. Mazzeo explores correlation of the service quality and competition. Rupp, Owens, and Plumly (2002) also observe more flight delays on less competitive routes. Borenstein and Rose (2003) analyze whether airline bankruptcies reduce air service quality.

We examine the decision process of flight operations, in the form of delays and cancellations with respect to three classes of independent variables. We have carefully chosen the variables that are identified in previously-cited literature as potential factors. These factors include various logistical variables, competition measures (at both the route and airport level), and weather-related variables. We also include aircraft characteristics, such as aircraft age, manufacture name and the seats of the plane.

The empirical results suggest a number of conclusions that have important public policy implications for service quality in the airline industries. We find that carriers which offer many daily scheduled flights on a route experience more frequent cancellations. Fewer cancellations occur for flights to and from carrier's hub. We also find that severe weather undoubtedly contributes to worse on-time performance.

The paper is structured as follows. The next section discusses the econometric model we select. Section 3 describes the data. Section 4 presents the results from the nested logit estimations. Section 5 concludes the paper.

2. Econometric Model

There are three possible choices that the airline can select in terms of the performance of flights: cancellation, delay, or on-time. The choice set with multiple categories makes the binary discrete choice model inappropriate. Multinomial logit, conditional logit and nested logit models are available to model a decision between more than two alternatives. However, we found solid theoretical and empirical foundations for the employment of nested logit model for this decision process. We assume that a carrier is an agent with the goal of profit maximization. However, the decision is subject to the constraints, such as security concerns, consumer satisfactions and the other non-pecuniary factors. We assume that the outcome of delay, cancellation and being on time for each flight is an equilibrium that is reached by the carriers balancing the goal of profit maximization with the many constraints facing the carriers.

The unordered-choices model can be motivated by a random profit maximization model. As mentioned above, the maximization of profit by a carrier incorporates both profit -related factors and factors indirectly- related to profits. Let $?$, denote the profit, which can be approximated by a linear combination of variables x , which represent flight characteristic. For the flight i with j (j=3) choices, suppose that the profit of choice j is

$$? \text{ (choice } j \text{ for flight } i\text{)} = ?_{ij} = x_i ?_j \; ? \; ?_{ij} \qquad (1)$$

Although a rational agent is assumed to make the choice in a deterministic setting, it would be impossible to observe all the factors that actually influence the choices of a carrier for a flight. We include $?_{ij}$ as the random disturbance term to capture this effect. Thus, the profit level $?$ is determined by the systematic components of profit and the random error. Profit maximization implies that p_{ij}, the probability that the carrier chooses outcome j is determined by

$$p_{ij} \; ? \; \Pr(?_{ij} \; ? \; ?_{ik}) \quad ? j \; ? \; k \qquad (2)$$

2.1 The multinomial logit model.

If the random error term is distributed identically and independently Weibull, then the multinomial logit model (MNL) describes carrier performance choices and the choice probability for alternative k is

$$p_{ij} \; ? \; \frac{\exp(?_{ij})}{\underset{k??_n}{?} \exp(?_{ij})} \; ? \; \frac{\exp(x_i ?_j)}{\overset{J}{\underset{k?1}{?}} \exp(x_i ?_k)} \qquad (3)$$

In equation 3, p_{ij} is the probability that the flight i chooses the alternative operation j. The x is a vector of characteristic of the flight i, and j is the number of unordered alternative operations. $?_j$ measures the contribution of flight characteristic i to the probability of selecting operations j, and $?_k$ measures the contribution of flight characteristic i to the probability of selecting k, that is, a single parameter implies that the attributes does not similarly affect the profit of all flight operations outcome. The time that the flight departure, for instance, may have different effects on the probability of cancellation relative to that of delay.

2.2 The conditional logit model.

The general form of the conditional logit is

$$p_{ij} \ ? \ \exp(z_{ij}?)\Big/ \overset{J}{\underset{k?1}{?}} \exp(z_{ij}?) \qquad (4)$$

Unlike the multinomial logit models, conditional logit models contain regressors that take on different values for each alternative flight operations outcome. That is, attributes contained in the vector vary with both flight i and alternative j. The coefficients are common to all z vectors because the attributes are assumed to influence the profit similarly across the 3 choices. A shortcoming of the conditional logit approach is that it does not allow for parameter estimation of individual flight characteristics because variation in the model is created by the attributes of choices, not the individual flights.

However, as pointed out by Greene (2000), one method to include variables that do not vary with choices is to create a set of dummy variables for the choices and multiply each of them by the variables. Thus, the data set is in such a format that each flight has as many observations as the number of outcomes. Therefore, the total number of observations in this dataset for conditional logit estimation would be three times larger. In addition, the number of variables would also be 3 times of the original number. The data set will be incredibly large and make the estimation process quite slow.

McFadden’s (1978) choice model is closely related to multinomial logistic regression in that if all independent variables are attributes of the individual, then the model is exactly the same as multinomial logit. For the above reasons we discussed, the multinomial logit model is preferred to the conditional logit model.

2.3 The nested logit model.

From the point of view of estimation, an assumption of both multinomial and conditional logit model is that the odds ratio, p_{ij}/p_{ik}, does not depend on the other choices, which is known as the independence of irrelevant alternatives (IIA).[1] The independence of irrelevant alternatives derives from the assumption that disturbances of the profit maximization model are independently and identically distributed. In the model for this application, intuitively speaking, the IIA assumption will be violated if the constraints that would result in both delay and cancellation of the flight are perceived quite similar by the carriers. It is highly possible since cancellation is just an extreme case for a flight delay.

2.4 Econometric Model Testing

We also use the test developed by Hausman and McFadden (1984) to determine if IIA assumption is valid to justify the use of multinomial logit model. They suggest that if a subset of the choice set truly is irrelevant, omitting it from the model altogether will be inefficient but will not lead to inconsistency. But if the remaining odds ratios are not truly independent of these alternatives, then the parameter estimates obtained when these choices are eliminated will be inconsistent. This procedure is the usual basis for Hausman's test. The test statistic has a limiting chi-square distribution. The test for IIA assumption is based on the first specification in Table 4 for the January data. The p-value indicates that the hypothesis is strongly rejected. We also test the residuals from multinomial logit estimation based on the specification in Model 1 in Table 4. We find that the data exhibits strong correlation of the three choices both across carriers and

[1] IIA implies that introducing or omitting another alternative will have same proportional effects on the probability of each alternative.

within carriers on the same day and on the same route. We also find highly significant correlation between days.[2]

If there is unobserved correlation among the alternative choices, multinomial and conditional logit models generate inconsistent parameter estimates because the profit function is no longer statistically independent but is correlated through the error terms across these alternatives. If the independence of irrelevant alternatives assumption fails, then the nested logit model is an appropriate method of estimation. The appeal of the nested model is its ability to accommodate correlation between subsets of alternatives in a choice set. Figure 1 demonstrates two potential nesting features of the decision process. Decision process 1 reveals two branches, which are on-time and not on-time respectively, and the two twigs for the not on-time branches: cancel and delay. By clustering related alternatives into subgroups, the IIA assumption is relaxed between the on-time and not on-time decision, while the IIA assumption is maintained within the subgroups. See Greene (2000) for a complete discussion.

The two-level nested logit model as demonstrated by Greene (2000) and modified to our specification is

$$\text{Prob[cancel | not on time]} = p_{jc} \; ? \; p_{j|c} \; ? p_c \qquad (5)$$

where c refers to on time or not on time choice and j refers to twig choices: cancel or delay.

$$p_{j|c} \; ? \; \exp(x_{j|c} \, ??) \Big/ \; \underset{J?1}{\overset{J_c}{?}} \exp(x_{j|c} \, ?) \qquad (6)$$

---

[2] The null hypothesis of no correlation of the decision choices on flights serving a route is strongly rejected with *p*-value less than 0.0001. The magnitude of this serial correlation varies. For example, in the cancellation equations we find between carriers on the same route the correlation is 0.0374 between carriers and 0.2103 within carriers. Besides, we find the correlation is 0.0485 on the same route between days in January 2003 for cancellation rate and 0.0922 for delay rate.

$$p_c \; ? \; \frac{\exp(x_i ? \;\; ? ?_c I_c)}{\overset{C}{\underset{c?1}{?}} \exp(x_i ? \;\; ? ?_c I_c)} \tag{7}$$

where for the branch c , $I_c \; ? \; \ln \overset{J_C}{\underset{j?1}{?}} \exp(x_{j|k} \; ??)$.

The coefficients $?$, $?$ are difficult to interpret, so we also report marginal effects (ME), which is calculated by

$$\text{ME(x)} = \frac{1}{N} \underset{i?1 \ldots N}{?} \frac{?p_i(j)}{?x_i} \tag{9}$$

The question remains whether decision process 1 approximates the true decision mechanism employed by the airline agents.[3] Decision process 2 provides an alternative displayed in Figure 1. The flight operation choices are broken into two branches of performance: cancel and not cancel; the twig of not cancel consists of delay and on-time. We prefer the first decision process primarily for two reasons. First, the construction of the first model is based on the natural assumption that the on-time and not-on time decisions are different, but within the not on-time choice set, cancellation and delay have some similarity. Second, the one-tailed chi-squared likelihood ratio test for selecting the appropriate nesting supports the first decision process. For example, the log likelihood for model (1) based on the decision process 1 is –41815.56, while the log likelihood based on the decision process 2 is –42042.442. The chi-square test statistic is calculated by taking twice the difference of the log likelihood functions. Our test statistics equals

[3] The tree structure is not intended to represent a sequential behavioral model of the decision makers. Logically, we may think of the choice process as that of choosing between the two choice nests and then making the specific choice within a given nest.

452, which easily exceeds the critical 9.2 value. Therefore, the test strongly supports the adoption of decision process 1 as the preferred decision-making process.

3 Data

The data for this analysis were obtained from a variety of sources. We collect the flight on-time performance information for January 2003 from the Bureau of Transportation Statistic (BTS) website [4]. This data set contains departure delays and arrival delays for non-stop domestic flights by major air carriers, and provides such additional items as origin and destination airports, flight numbers, tail numbers, scheduled arrival and departure times, actual arrival and departure times, cancelled or diverted flights, and non-stop distance. The data set is composed by flight specific observations.

The flight dataset includes every flight served by 17 carriers. They are American Airlines (AA), Alaska Airlines (AS), Jetblue Airways (B6), Continental Airlines (CO), Atlantic Coast Airlines (DH), Delta Air lines (DL), Atlantic Southeast Airlines (EV), Airtran Airways Corporation (FL), American West Airlines (HP), American Eagle Airlines (MQ), Northwest Airlines (NW), Sky West Airlines (OO), Continental Express Airline (RU), American Trans Air, Inc. (TZ), United Airlines (UA), US Airways (US) and Southwest Airlines (WN). The largest three carriers, WN, AA and EV account for 14.6 percent, 12.5 percent and 11.1 percent of the monthly flights respectively.

This study uses the definition of "on-time" by the U.S. Department of Transportation (DOT). We let the dummy variable, *ontime* equal to 1 if the flight arrived at the destination airport within 15 minutes of the scheduled time. *Delay* serves as an indicator variable, equal to 1 for flights that arrived more than 15 minutes late after the scheduled time and delay also equals 1 for diverted flights. The data also provides data for cancelled flights. By these definitions, the outcome of the flight can only be in one of

[4] www.bts.gov.

the three categories: cancel, delay or on-time. Table 1 summarizes the flight operations in January 2003. The average cancellation rate is about 1.45 percent while the average delay rate in the sample is 14.5 percent. Hence, the remaining 84% of the flights are on time.

The daily frequency of flights might also affect the probability of the flight being cancelled or delayed. From the January sample, we calculate the total number of daily flights by carrier j on route i. Table 1 shows that the total number of daily flights of a carrier on a route can varies from 1 to 32. Mayer and Sinai (2002) found that congestion also contribute to the worse on-time performance. We calculated the total number of daily airport operations (both arrivals and departures), which serves as a measure of congestion at both origination and destination airports.

There is a strong correlation between competition of airlines and flight performance documented by Rupp (2003). We also include several measures for the competition both at the route level and airport level. *Route market share* is the number of scheduled flights for a carrier on a route r divided by the total number of scheduled flights for all carriers serving route r during the day. *Effective carriers* is the inverse of the Herfindahl index, the sum of the squared market shares for all carriers serving the route. Monopoly equals one for routes served by a single carrier. Airport concentration is defined as the Herfindahl index on the share of flights by the various airlines that serve that airport. It ranges from 0 to 1. Airport concentration measures the competition in airport level.

To investigate the time-of-day effects on flight operations decisions, we constructe a normalized scheduled departure time variable with a scale from 0 to 1. The time is converted into the proportion of minutes elapsed starting from 00:01AM till 11:59 PM. This variable is equal to 0 if the departure time is 12:00AM, and equal to 1 if the departure time is 11:59 PM. For example, if the flight is scheduled to take off at 6:12 AM, then it is calculated by dividing the minutes from 12:00AM, which is 372, by the total number of minutes of 24 hours, which is 1439. The normalized time variable would take the value of 0.25.

An airline's hub is defined as a function of the number of connections by carrier j at the airport. We generate separate dummy variables for hub airlines in origination and destination airports. Airlines with more than 25 connections are considered hub carriers. In addition, in order to differentiate the size effects of hub airlines, we generate a set of dummy variables for small, medium, and large hubs based on the number of connections: 26 to 46, 46 to 70, and 71 or more respectively. In the January 2003 data set, Table 2 shows that approximately 40 percent of the flights originate from an airline's hub (10 percent originate from the large hub, about 15 percent each from the medium hub and small hub). Hub and concentration are included separately for both the origin and destination airports to allow for separate effects for each end of flights.

We are also concerned about weather conditions that might affect the airlines' ability to control the flight operations. National Climate Data Center maintains records of daily atmospheric conditions at various airports throughout the country that is accessible through the Internet (www.ncdc.noaa.gov). We collect the weather information for all domestic airports and then merge it with the flight information data set by the airport name. Thus, each flight specific observation is enriched with useful daily weather elements, such as daily precipitation and the minimum temperature. The weather related variables we include are daily precipitation and freezing. We define *freezing* as 1 if the minimum temperature is less than 33? F and 0 otherwise. In order to investigate effect of the concurrence of both precipitation and freezing, we introduce an interaction term of freezing and precipitation.

The flight level on-time performance data maintained by BTS also includes the tail-number of the aircraft for the flight. The tail number is a unique aircraft identifier that was matched to the U.S. Civil Aviation Registry maintained by the Federal Aviation Administration (www.faa.gov). The FAA data contains manufacturer of the plane and the year it was produced and the maximum number of seats possible on the plane. Merging those aircraft information by the tail number with the flight-level information provides the opportunity to explore whether aircraft characteristics are correlated with flight service quality. Boeing aircraft account for almost 41percent percent of the observations in our sample. Embraer, Mcdonnell and Airbus each averages 10 percent. The remaining manufacturers include Saab, Fornier, etc, accounting for about 30 % percent of the sample. Table 1 shows that the age of the aircrafts ranges from 0 to 47. The maximum number of seats available for the flight in this sample is 495.

Another factor in cancel, delay or on-time decisions by the carrier involve potential flight revenue. We generate the potential revenue variable for each flight by multiplying the number of seats available by the average ticket price for the flight. Another variable used to proxy the potential revenue is yield, which is constructed by dividing the average ticket price by the distance, the non-stop distance is miles between origination and destination.

The estimations are organized as follows. Model 1 (in Table 4) is the nested logit estimation of flight performance outcomes using the three categories of explanatory variables: logistical variables, competition variables and weather variables. Based on model 1, additional estimations are organized as follows: model 2 includes carrier fixed effects for the purpose of identifying the parameters after controlling for unobserved factors for each carrier. In order to investigate the possible hub size effects, we decompose the hub carrier into 3 categories: large, medium and small hub carriers. In model 3, the *origination hub* is replaced by the above three variables. Model 4 also include carrier-fixed effects.

We next explore whether flight performances differ by airport size. We divide our data into three size categories: large, medium and small airports. The “large” airports are those averaging more than 400 flight operations per day, which account for approximately 62% of our January sample flights. One fourth of the flights in the sample originate from the “medium” airports which average between 100 and 400 flights per day. The remaining flights are from small airports that average between 2 and 100 flights per day. The cancellation rate for small airports is 2.34%, which is nearly twice as large as the medium and large airports. (See Table 3) Medium airports have the lowest average delay rate and cancellation rate. It might be worth examining the performance of flight operations separately by airport size due to the variation across our sample. Model 5 is for large-size airports, model 6 is for medium-size airports and model 7 is for small size airports. Carrier-fixed effects are included in each of the three estimations.

As noted above, we also examine whether revenue and aircraft characteristics affect airline service quality. In model 8, capacity and revenue are introduced in the estimation as well as the number of seats and age of the aircraft. Model 9 includes carrier-fixed effects and model 10 includes both carrier-fixed effects and manufacturer-fixed effects.

As mentioned above, the data format required by nested logit estimation would make the dataset too large to manage. Hence, the sample estimated includes only Fridays' flights in January 2003. Figure 2 shows the average flights operations performance for each day of the week in January 2003. Friday's flights are quite representative of the whole month since there is no obvious difference between days in cancellation rates, delay rates and on-time rates.

4 Results

4. 1 Results for logistical variables

From model 1 to model 4, we consistently find positive and significant correlation between daily flight frequencies and cancel rates. In contrast, the delay rate is decreased by 0.06-0.4 percentage due to an additional flight increase as suggested by the first 4 models. Move to the estimation for different size airports, the pattern persists only for the large airports. The coefficient switches its sign only in estimation 8 with aircraft characteristic included, however, after controlling for carrier-fixed effects and manufacturer-fixed effects in model 10, the parameter regains the positive sign and 1% significance level. A stand error increase (four flights) in daily frequency of flight would increase the cancel rates by 0.3 percentage points as indicated in model 2, representing a 20% increase. There are two possible explanations for the positive correlation. First, the airline might opportunistically cancel flights with lower passenger load and re-book these passengers on later flights. Second, flights cancelled on routes with less daily frequency would have less flight substitutes for passengers and cancellation would incur great inconvenience for passengers. In this case, airline would manage to prevent cancellation of the flights on routes with fewer flights and to maintain the satisfaction of consumers.

Another factor that might influence the operations of flights is the scheduled departure time. Results from Model 1 to Model 4 suggest that flights scheduled to departure late in the day are less likely to be on time and typically have fewer cancellations. The marginal effects suggest that flights that departure two hours later, on average, according to the definition of the scale, have 0.9 percentage points higher likelihood of late arrival, representing 7 percentage increase relative to the average level.

Figure 3 plots the average on-time rates for each two-hours period during the day. It shows that on-time performance began to decline starting from 7:00 am. A potential explanation might be that carriers have less ability to adhere to the schedule as air traffic increases since any single cancellation or extended delay can cause ripple effects throughout the rest of the day.

We calculated the total number of flights at an airport to proxy airport congestion. The coefficients for daily total flights at origination airport are quite stable. From model 1 to model 4, the coefficients on total daily flight at airport with respect to on-time outcome are all negative and significant at 1%. Even after controlling for the manufacturer-fixed and carrier-fixed effects in model 8 and 9 and 10, the coefficient remains unchanged in both sign and significance and the marginal effect stays around –0.00003. Hence, busier airports have fewer on-time arrivals. A stand error increase in flights from the origination airport is associated with about one percentage point decrease in on-time rate. We also find similar results for destination airports. Compared to the cancellation rate, the coefficients are less stable, however, all the significant coefficients suggest positive correlation between cancellation probability and the number of airport flights. For example, Model 1 suggests busier airports have higher cancellation rates. A similar finding is reported in Model 3 and 4, however, the magnitude of marginal effects is larger with hub carrier size effects considered. In model 4, we get the highest marginal effect on cancel rate, which means that a stand-error increase in flights (316 flights) at the origination airport leads to 0.3 percent increase in cancel rates, which represent a 21 percentage increase. It provides some support to the argument that congestion leads to higher probability of cancellation and delay.

It is worth noting that in model 5, 6 and 7, where the estimation is based on separate airport size, the total number of flights in airport lose the explanatory power toward cancel rates. The reason might be that the variation of number of flights is smaller within each category of airport. Only in large airports, congestion still adversely affects the on-time performance, though with small effects.

4.2 Results for competition variables

We expect that carriers may organize their product inputs with the competitiveness of their route in mind. We expect that more competitive routes provide better service quality since Suzuki (2000) has documented that passengers with experience of cancel or delay are more likely to switch to other carriers. As a carrier's market share increases, the route becomes less competitive. Thus we expect that the sign of market share for cancellation rates is positive and negative for on-time rates.

However, two of the three coefficients on market share for cancel rates that achieves 5% significance are negative in estimations in model 2 and model 7. It might be explained by the positive correlation between market share and total daily flights of the carrier on a route. The effects of competition might be offset by congestions problems at busy airports or are partly captured by the frequency of daily flights. The negative sign might indicate that the congestion effect outweighs the competition effects.

Only after we control for manufacturer characteristics and revenue in model 8, market share take the expected positive sign for cancellation rates. The marginal effects indicate that large market share prohibits the incentive of carriers to provide high-quality service. For instance, a stand error increase in market share would increase the cancel rates by 0.15, representing a ten percent increase. However, it becomes fragile after we include carrier-fixed effects and manufacturer-fixed effects. We also use the alternative measure of route-level competition: effective carriers and monopoly to examine the effects of competition on flights operations. Taken together, the findings for effects of competition is not as distinct as expected.

It is also useful to examine the effects of competition in airports. In model 2 and model 4 with carrier fixed effects, we find that concentration tells almost the same story as market share. We did not find that increase in airport concentration resulted from less competition by carriers adversely affect the performance. On the contrary, among the 20 coefficients for on-time rates from the 10 estimations, three-fourths of them are positive and significant, which shows that larger concentration is associated with better on-time performance (including both measure of concentration of origination and destination). It provides additional support to the points we made above that less concentration results from more competition might incur additional congestion problems.

Results from Model 1 and model 2 shows that flights to and from the carrier's hub are less likely to be cancelled; however, flights destined for hub are more likely to be on time while flights originating from hub are more likely to be delayed. After we decompose the effects of hub by large, medium and small hub size, the results are only robust to cancellation rate, which is also supported by the results from the other estimations. It is useful to compare different hub size effects: in model 3 we find flights to and from large carrier's hub are less likely to be cancelled relative to small and medium carrier's hub. Results in Model 4 with carrier-fixed effects exhibit the same pattern. Flights to and from the small carrier's hub are 0.2-0.5 percentage less likely to be cancelled, in contrast, flights to and from the large hub are 0.6-0.8 percentage less likely to be cancelled, representing 41%-55% decrease. At the same time, both specifications indicate that flights destined for large hub are less likely to be cancelled compared to those originating from large hubs.

These results are consistent with the argument that hub have fewer cancellations due to access to spare parts and replacement aircrafts. Hub carriers have a greater incentive to keep connecting passengers from missing flights to minimize the cost to rebook passengers onto new connections. It is especially the case for flight destined for hub.

When it comes to the estimation based on three different airport sizes, the hub variables achieve no statistical significance, which might be caused by less variation in designation of hub carriers. For instance, none of the flights in the small airport sample originates from the hub. Also, the results are not supported significantly by model 8,9 and 10 when we include revenue and capacity variables.

4.3 Results for Weather variables

As expected, weather variables are significant predictors of on-time performance. From model 1 to model 10, the twenty coefficients on daily precipitation (both in origination and destination) are all negative and highly significant with respect to on-time rate. We find that flights are significantly more likely to arrive late due to precipitation. However, the marginal effects become modest as we included carrier fixed effects and other variables of interest. An additional inch of rain in origination airport is associated with 0.1 to 0.2 percentage point decrease in on-time rate, while destination airports experience a 0.2-0.6 percentage point decease. In contrast, the coefficients with respect to cancellation rate are positive and significant only from model 1 to model 5. It still supports the view that cancellation may be caused by unexpected precipitation. For instance, in Model 1, the estimations indicate that an additional inch of precipitation at both the origination and destination is associated with an increase of 0.08 percentage points in cancellation rate and an increase of 0.5 percentage points in delay rate. It is also interesting to note that when we control for revenue and manufacturer effects in model 8, 9 and 10, only the precipitation in destination airport are statistically significant. The ten estimations are consistent in that the marginal effect of precipitation in destination is relatively larger than that in origination airport without exception.

Compared to precipitation, only results from model 2, 3 and 5 support the view that flights to and from airports with freezing weather have higher probability to be cancelled or delayed. The reasons that cold weather does not consistently contribute to worse performance might be that freezing without precipitation is not severe enough to affect flight performance. This point can be strongly supported by the negative and significant coefficients on the interaction term of freezing and precipitation. For example, in model 7 and 8, the results show that freezing weather is associated with better on-time rates. However, the concurrence of precipitation and freezing in both origination and destination is always associated with worse on-time rates in all ten estimations.

In the group of estimations for three different size airports, the coefficients on most variables become fragile compared to the weather variables, particularly for precipitation, which indicates the robust negative relation between precipitation and worse on-time performance.

4.4 Other findings

We introduce capacity, potential revenue and number of seats in model 8 to explore the issue whether revenue is one of the primary concerns involved in carrier's operations decision. We find that potential revenue provides little support to the arguments that carriers take actions to avoid cancellation of large revenue flights. Rather, the coefficients on capacity in the three estimations are all highly significant and negative for cancellation rates. It seems that the airlines are concerned about the long- term effects of the cancellation because it would result in passengers' dissatisfaction hence carriers cancel flights with fewer passenger. In model 9, the marginal effect shows that a stand error increase in capacity lead to a decrease of 0.12 percentage points in cancel rates, representing a 8 percentage change. However, larger capacity is associated with higher probability of delay. The negative and highly significant coefficient on the number of the seats for cancellation rate provide further confirmation to the view that passengers' satisfaction is more a concern for airlines than is the short-term profit, although the marginal effect in this estimation is rather small. After controlling for the carrier-fixed effects in model 9 and model 10, we find older planes contribute to worse on-time performance. A flight served by an aircraft with additional 10 years of age has a 0.3 percentage point higher likelihood of cancellation and one percentage point higher likelihood of delay.

For all the above models that include carrier-fixed effects, each of carrier dummies are jointly highly significant in all regressions. It is also interesting to note that there exhibits a pattern in the ranking of the estimated coefficients of carrier dummies, which indeed shows that some carriers have the propensity to keep to the schedule or not. For example, AA, WN, NW, AS, DL, CO always stand among the top six of all the 17 carriers in terms of on-time rate. HP and DH always show up in the bottom six with more frequent cancellation rates.

5. Conclusion

We extend the literature on flight on-time performance by exploring the determinants of airline flights decisions in a nested logit model. This study finds routes with a higher daily frequency are subject to more cancellations. As the day progresses, flight delays become more prevalent. Congestion at busier airports adversely affects the on-time performance of flights. The results in the paper indicates that less competition does not necessarily lead to worse service quality in terms of cancellation and delay records. Hub carriers are less likely to cancel flights, and the effect increases with the size of the hub. Severe weather is a significant contributor to worse on-time performance. The results also suggest that carriers are more concerned about the long-term effects on airline's profitability of flights cancellation than short-term revenue of the flight.

References:

Greene, William H.2000. Econometric Analysis 4th Edition. Upper Saddle River, NJ: Prentice Hall

Mayer, Christopher and Todd Sinai. 2003a. 'Network Effects, Congestion Externalities, and Air Traffic Delays: or Why All Delays Are Not Evil', American Economic Review, forthcoming.

Mazzeo, Michael J. (Updated) 'Competition and Service Quality in the U.S. Airline Industry', North-western University Department of Management & Strategy working paper.

Rupp, Nicholas G. and George M. Holmes. 2003. ' Airline Schedule Recovery After Airport Closures: Empirical Evidence since September 11TH ' working paper 9744.

Rupp, Nicholas G. and George M. Holmes. 2003. ' Why Are There So Many Flights Cancellations?' East Carolina University Department of Economics Working Paper.

Rupp, Nicholas G. and George M. Holmes. 2003. ' Does Competition Influence Air-line On-Time Performance?' East Carolina University Department of Economics Working paper.

Suzuki, Yoshinori. 2000. 'The Relationship between On-Time Performance and Airline Market Share', Transportation Research: Part E: Logistics and Transportation

Review, 36:2, pp.139-54

Thomas A. Knapp. 2001 ‘A Nested Logit Approach To Household Mobility’, Journal of Regional Science, Vol.41, No.1, 2001, pp.1-22

Figure 1 Schematic of Airline Operations Decision Process

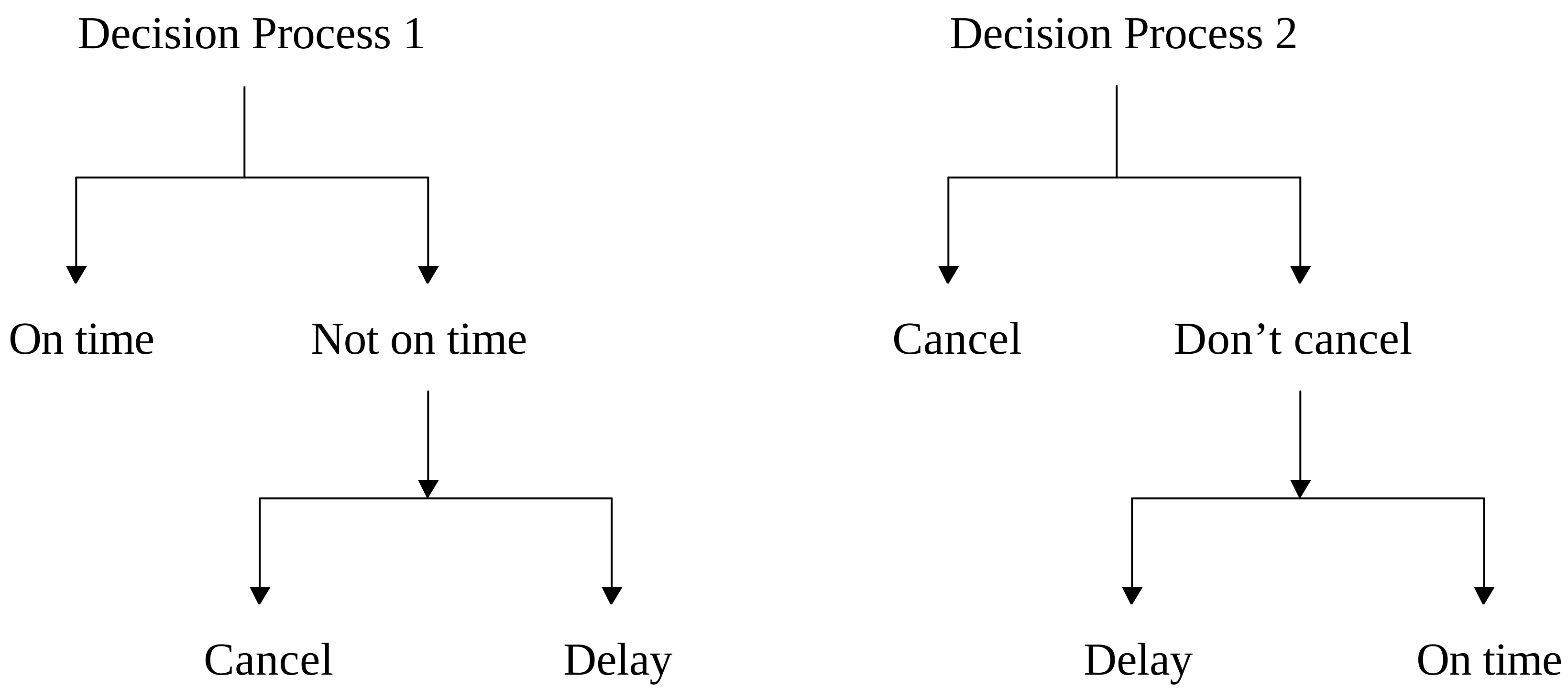

Figure 2: Flight operations outcomes of January 2003

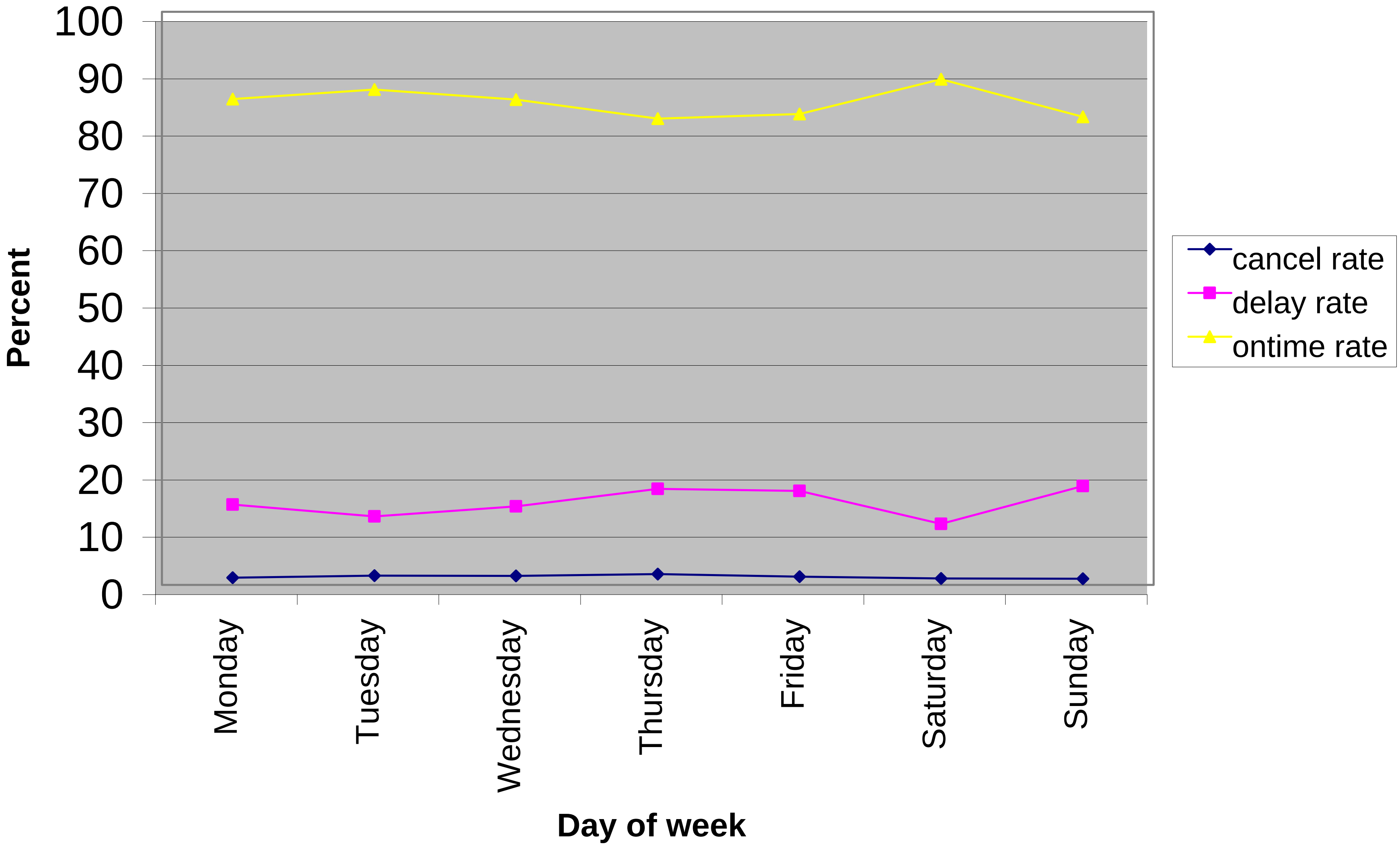

Figure 3:On-time rate & departure time for U.S. domestic flights in January 2003

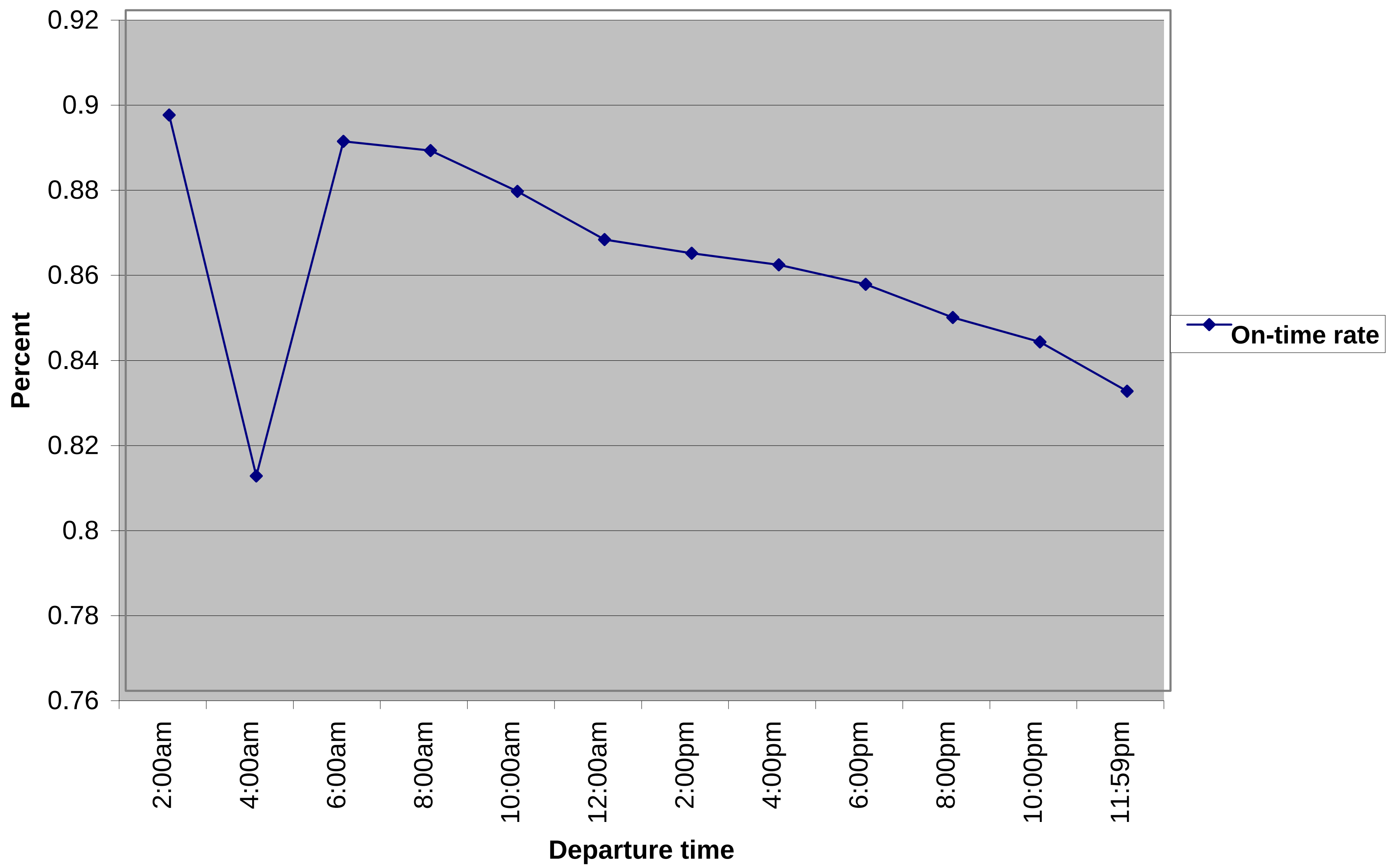

Figure4: Cancelltion, delay rate & departure time for U.S. domestic flights in January 2003

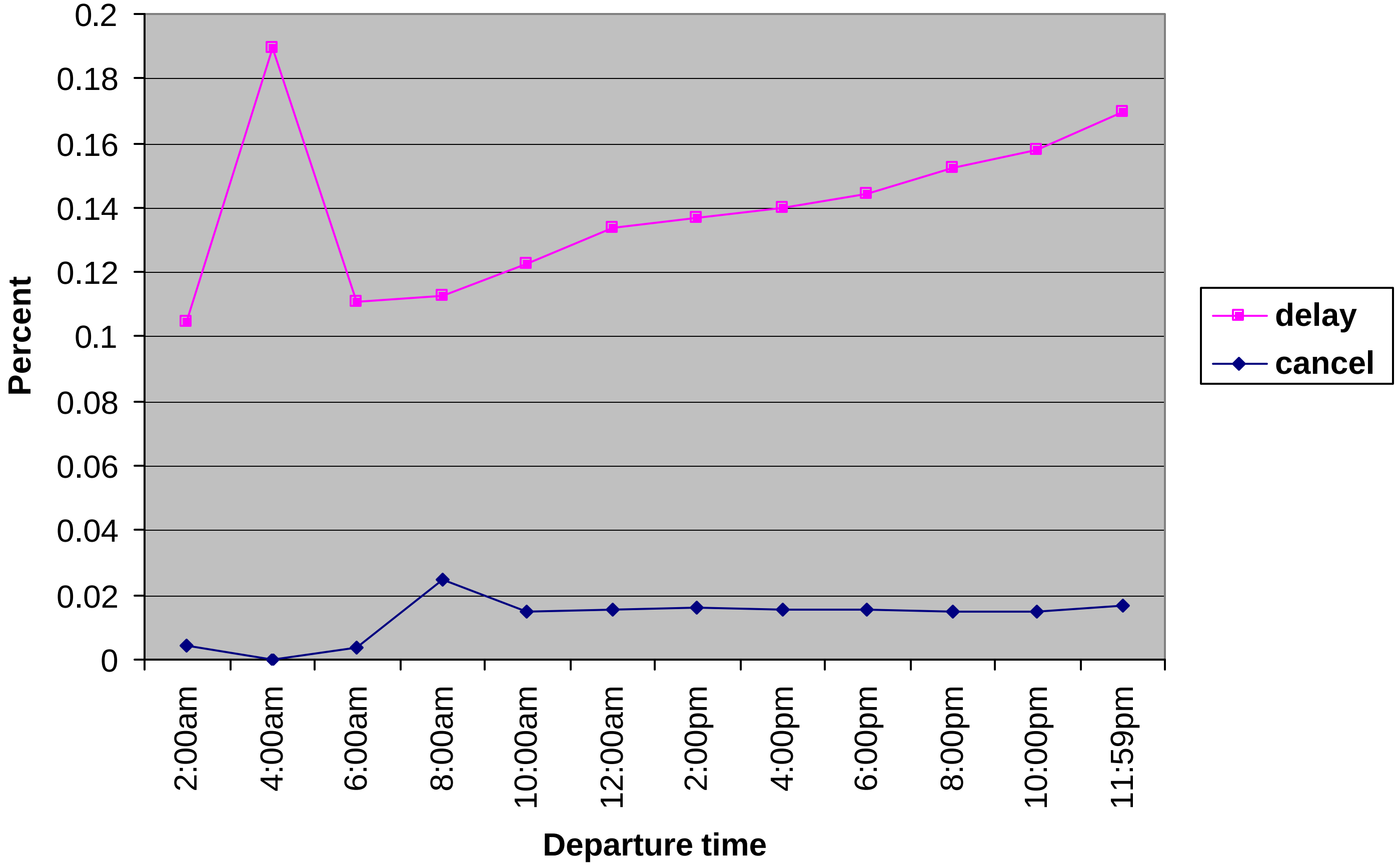

Table 1 : Summary Statistics for Scheduled Flights in January 2003 (n=514805)

Variable	Mean	Std. Dev.	Min	Max
Cancel	0.0145	0.119	0	1
Delay	0.146	0.353	0	1
Ontime	0.84	0.367	0	1
Logistical varibles				
Distance(miles)	710.4	558.7	21	4962
Daily total flights for carrier j on route r	5.668	3.996	1	32
Capacity of the aircraft	0.596	0.133	0	1
Renormed scheduled departure time *	0.568	0.192	0.003	1
Average ticket price ($)	162.84	77.116	0	1304
Yield per flight (Average ticket price/distance)	0.355	0.323	0	4
Daily total flights at origination airport	729.1	634.0	2	2127
Daily total flights at destination airport	740.5	640.6	2	2125
Potential revenue per flight ($)	23095	17998	0	260800
Competition variables				
Route market share	0.723	0.279	0.03	1
Effective carrier	1.6	0.645	1	4.57
Monopoly	0.435	0.496	0	1
Origination hub	0.404	0.491	0	1
Small origination hub	0.145	0.352	0	1
Medium origination hub	0.157	0.364	0	1
Large origination hub	0.102	0.302	0	1
Destination hub	0.406	0.491	0	1
Small destination hub	0.140	0.347	0	1
Medium destination hub	0.164	0.370	0	1
Large destination hub	0.102	0.302	0	1
Concentration at destination airport	0.323	0.177	0.040	1
Concentration at origination airport	0.323	0.177	0.039	1
Weather variables				
Daily precipitation at origination airport(inches to hundredths)	4.063	15.141	0	302
Daily precipitation at destination airport(inches to hundredths)	4.068	15.173	0	302
Freezing at origination airport	0.566	0.496	0	1
Freezing at destination airport	0.566	0.496	0	1
Freezing and precipitation at origination airport	0.132	0.339	0	1
Freezing and precipitation at destination airport	0.132	0.339	0	1
Aircraft characteristics				
Number of seats	133.8	64.5	15	495
Age of aircrafts	9.5	7.5	0	47
Airbus	0.113	0.317	0	1
Boeing	0.409	0.499	0	1
Embraer	0.135	0.342	0	1
Mcdonnell	0.108	0.31	0	1

Note: * The normalized scheduled departure time is a variable with a scale from 0 to 1. The point time is converted into the proportion of minutes elapsed starting from 00:01AM till 11:59 PM. This variable is equal to 0 if the departure time is 12:00PM, and equal to 1 if the departure time is 11:59 PM.

Table 2: Flight operations outcome by hub carriers for all U.S.domestic flights in January 2003

	Hub Carriers				
Percent	Non-hub	Large hub	Medium hub	Small hub	Total
Cancel	1.65	0.95	1.24	1.26	1.46
Delay	13.53	15.47	16.21	16.47	14.58
On-time	84.82	83.58	82.54	82.28	83.97
Total	100	100	100	100	100

Table 3: Flight operations outcome by airports for all U.S.domestic flights in January 2003

	Airport			
Percent	Large airport	Medium airport	Small airport	Total
Cancel	1.38	1.19	2.34	1.46
Delay	15.4	12.4	14.77	14.58
On-time	83.22	86.41	82.88	83.97
Total	100	100	100	100

Table 4: Nested Logit estimations of flight operation decisions for all U.S. domestic flights in Januray 2003 (Fridays only)

Model	(1)		(2)	
Outcome	Cancel	Ontime	Cancel	On-time
Logistical Variables				
Distance(miles)	-0.0003**	0.0002**	-0.0001*	-0.0001**
	(0.0001)	(0.0000)	(0.0001)	(0.0000)
Daily total flights for carrier j on route r	0.0415**	0.0130**	0.0726**	-0.0330**
	(0.0058)	(0.0047)	(0.0072)	(0.0063)
Daily total flights at origination airport	0.0002**	-0.0003**	0.0000	-0.0002**
	(0.0001)	(0.0000)	(0.0001)	(0.0000)
Daily total flights at destination airport	0.0000	-0.0003**	-0.0002**	-0.0002**
	(0.0001)	(0.0000)	(0.0001)	(0.0000)
Renormed scheduled departure time	-1.3167**	-0.2260*	-0.9372**	-0.4050**
	(0.1324)	(0.0919)	(0.1511)	(0.0831)
Competition Variables				
Route market share	-0.1377	-0.0804	-0.2475*	0.0419
	(0.0974)	(0.0588)	(0.1118)	(0.0663)
Concentration at origination airport	-0.0644	0.1155	-0.1344	0.1944*
	(0.1305)	(0.0843)	(0.1387)	0.0000
Concentration at destination airport	-0.0051	0.2796**	-0.4000**	0.5038**
	(0.1329)	(0.0870)	(0.1390)	(0.0937)
Origination hub	-0.3453**	0.0883*	-0.3194**	0.0359
	(0.0702)	(0.0419)	(0.0825)	(0.0495)
Destination hub	-0.2769**	0.2067**	-0.0409	0.0926
	(0.0737)	(0.0438)	(0.0829)	(0.0491)
Weather Variables				
Daily precipitation at origination airport	0.0144**	-0.0236**	0.0142**	-0.0226**
	(0.0008)	(0.0014)	(0.0010)	(0.0015)
Daily precipitation at destination airport	0.0193**	-0.0372**	0.0191**	-0.0354**
	(0.0009)	(0.0016)	(0.0012)	(0.0016)
Freezing at origination airport	0.0579	-0.0387	0.1966**	-0.0924*
	(0.0627)	(0.0352)	(0.0689)	(0.0387)
Freezing at destination airport	0.1656*	-0.1465**	0.2168**	-0.1494**
	(0.0732)	(0.0379)	(0.0739)	(0.0393)
Freezing and precipitation at	-0.2551**	-0.6674**	-0.0924	-0.7495**
origination airport	(0.0626)	(0.0394)	(0.0639)	(0.0395)
Freezing and precipitation at	0.0046	-0.4391**	0.0791	-0.4792**
destination airport	(0.0682)	(0.0411)	(0.0697)	(0.0439)
Constant	-1.8098**	1.8735**	-2.2046**	2.7207**
	(0.1573)	(0.1210)	(0.1917)	(0.1417)
Carrier fixed effects	No		Yes	
Chi-Square test of carrier fixed effects			? $^{2}_{16}$? 214.5 *p*< 0.0001	? $^{2}_{16}$? 657.2 *p*< 0.0001
Log-likelihood	-41,815		-41,010	
Number of observations	85,539		85,539	

Note: Stand errors are reported in parentheses.
**: Significant at 1%; * : Significant at 5%

Table 4a : Marginal Effects for Nested Logit estimations from Table 4

Model	(1)			(2)		
Outcome	Cancel	Delay	On-time	Cancel	Delay	On-time
Logistical Variables						
Distance(miles)	-0.000004	-0.00001	0.00001	0.00000	0.00002	-0.00002
Daily total flights for carrier j on route r	0.000116	-0.00444	0.00433	0.00076	-0.00064	-0.00011
Daily total flights at origination airport	0.000005	0.00003	-0.00004	0.00000	0.00003	-0.00003
Daily total flights at destination airport	0.000004	0.00004	-0.00005	0.00000	0.00004	-0.00004
Renormed scheduled departure time	-0.005802	0.11739	-0.11159	-0.00084	0.11107	-0.11022
Competition Variables						
Route market share	0.000039	0.01940	-0.01944	-0.00182	0.01092	-0.00910
Concentration at origination airport	-0.001723	-0.01013	0.01186	-0.00284	-0.01527	0.01811
Concentration at destination airport	-0.003212	-0.03473	0.03794	-0.00764	-0.03622	0.04386
Origination hub	-0.003200	0.01226	-0.00906	-0.00216	0.01636	-0.01420
Destination hub	-0.004112	-0.00722	0.01133	-0.00122	-0.00878	0.01000
Weather Variables						
Daily precipitation at origination airport (inches to hundredths)	0.000360	0.00198	-0.00234	0.00032	0.00188	-0.00220
Daily precipitation at destination airport (inches to hundredths)	0.000547	0.00337	-0.00392	0.00049	0.00313	-0.00362
Freezing at origination airport	0.000809	0.00095	-0.00176	0.00209	-0.00137	-0.00072
Freezing at destination airport	0.002720	0.00719	-0.00991	0.00281	0.00436	-0.00717
Freezing and precipitation at origination airport	0.005966	0.10101	-0.10697	0.00755	0.09865	-0.10620
Freezing and precipitation at destination airport	0.005022	0.05479	-0.05981	0.00559	0.05406	-0.05966

Table 5: Hub size effects- Nested Logit estimations of flight operation decisions for all U.S. domestic flights in January, 2003 (Fridays only)

Model	(3)		(4)	
Outcome	Cancel	On-time	Cancel	On-time
Logistical Variables				
Distance (miles)	-0.0002**	0.0002**	0.0002**	-0.0002**
	(0.0000)	(0.0000)	(0.0001)	(0.0000)
Daily total flights for carrier j or route r	0.0335**	0.0067	0.0486**	0.0057*
	(0.0053)	(0.0064)	(0.0114)	(0.0029)
Daily total flights at origination airport	0.0002**	-0.0005**	0.0007**	-0.0002**
	(0.0001)	(0.0000)	(0.0001)	(0.0000)
Daily total flights at destination airport	0.0001	-0.0004**	0.0007**	-0.0003**
	(0.0001)	(0.0001)	(0.0001)	(0.0000)
Renormed scheduled departure time	-1.2816**	0.1205	-0.3910*	-0.8287**
	(0.1368)	(0.1699)	(0.1751)	(0.0490)
Competition Variables				
Route market share	-0.0218	-0.1527*	0.1201	-0.0448
	(0.0855)	(0.0756)	(0.1480)	(0.0440)
Concentration at origination airport	0.0405	0.0268	0.5650**	0.2083**
	(0.1150)	(0.1085)	(0.2138)	(0.0633)
Concentration at destination airport	0.0286	0.2068*	0.1626	0.3267**
	(0.1161)	(0.1104)	(0.2217)	(0.0619)
Small origination hub	-0.3625**	0.1825**	-0.5377**	-0.0906**
	(0.0783)	(0.0663)	(0.1375)	(0.0353)
Medium origination hub	-0.1528*	0.0521	-0.2039	-0.3348**
	(0.0798)	(0.0677)	(0.1544)	(0.0416)
Large origination hub	-0.6834**	0.4444**	-0.7448**	-0.2718**
	(0.1147)	(0.1016)	(0.2152)	(0.0561)
Small destination hub	-0.0828	0.1860**	-0.3580**	0.1459**
	(0.0776)	(0.0685)	(0.1382)	(0.0361)
Medium destination hub	-0.3175**	0.2577**	-0.0062	-0.2520**
	(0.0869)	(0.0705)	(0.1517)	(0.0407)
Large destination hub	-0.5430**	0.6167**	-0.6434**	0.0359
	(0.1478)	(0.1231)	(0.2205)	(0.0563)
Weather Variables				
Daily precipitation at origination airport	0.0119**	-0.0254**	0.0117**	-0.0097**
	(0.0010)	(0.0019)	(0.0019)	(0.0006)
Daily precipitation at destination airport	0.0157**	-0.0394**	0.0131**	-0.0127**
	(0.0014)	(0.0021)	(0.0019)	(0.0006)
Freezing at origination airport	0.1307*	-0.1172*	0.1196	0.0095
	(0.0566)	(0.0516)	(0.0941)	(0.0234)
Freezing at destination airport	0.2435**	-0.2370**	0.0801	-0.0119
	(0.0723)	(0.0672)	(0.0955)	(0.0232)
Freezing and precipitation	-0.3522**	-0.5413**	0.7941**	-0.8553**
at origination airport	(0.0569)	(0.0685)	(0.0859)	(0.0262)
Freezing and precipitation	-0.0830	-0.4025**	0.9153**	-0.6007**
at origination airport	(0.0932)	(0.0763)	(0.0863)	(0.0267)
Constant	-2.0288**	1.6131**	-5.1926**	3.2712**
	(0.1437)	(0.2422)	(0.2776)	(0.0741)
Carrier fixed effects	No		Yes	
Chi-Square test of carrier effects			$?^{2}_{16}$? 445.29	$?^{2}_{16}$? 1323.2
			p< 0.0001	p< 0.0001
Log-likelihood	-41,769		-41,421	
Number of observations	85,539		85,539	

Table 5a : Marginal Effects for Nested Logit estimations from Table 5

Model	(3)			(4)		
Outcome	Cancel	Delay	On-time	Cancel	Delay	On-time
Logistical Variables						
Distance (miles)	-0.000003	-0.00001	0.00001	0.000004	0.00002	-0.00002
Daily total flights for carrier j or route r	0.000016	-0.00437	0.00435	0.000478	-0.00137	0.00089
Daily total flights at origination airport	0.000006	0.00003	-0.00004	0.000010	0.00002	-0.00003
Daily total flights at destination airport	0.000005	0.00005	-0.00005	0.000010	0.00003	-0.00004
Renormed scheduled departure time	-0.004873	0.11973	-0.11485	0.003057	0.11083	-0.11389
Competition Variables						
Route market share	0.001675	0.02141	-0.02309	0.001702	0.00414	-0.00585
Concentration at origination airport	-0.000194	-0.00762	0.00782	0.004307	-0.03396	0.02966
Concentration at destination airport	-0.002270	-0.02889	0.03116	-0.001113	-0.04381	0.04492
Small origination hub	-0.003064	0.01529	-0.01222	-0.005047	0.01860	-0.01355
Medium origination hub	-0.001009	0.00956	-0.00855	0.000735	0.04537	-0.04611
Large origination hub	-0.006915	0.01627	-0.00935	-0.005701	0.04442	-0.03872
Small destination hub	-0.002339	-0.01457	0.01690	-0.005180	-0.01394	0.01912
Medium destination hub	-0.003795	0.00115	0.00265	0.002155	0.03223	-0.03439
Large destination hub	-0.008487	-0.02004	0.02853	-0.007313	0.00383	0.00348
Weather Variables						
Daily precipitation at origination airport (inches to hundredths)	0.000322	0.00193	-0.00225	0.000214	0.00109	-0.00130
Daily precipitation at destination airport (inches to hundredths)	0.000491	0.00328	-0.00377	0.000255	0.00145	-0.00171
Freezing at origination airport	0.001688	0.00091	-0.00260	0.001217	-0.00278	0.00156
Freezing at destination airport	0.003358	0.00405	-0.00741	0.000977	0.00048	-0.00146
Freezing and precipitation at origination airport	0.005185	0.10479	-0.10997	0.016178	0.09872	-0.11490
Freezing and precipitation at destination airport	0.004344	0.05909	-0.06344	0.015250	0.06465	-0.07990

Table 6: Airport size effects: Nested Logit estimations of flight operation decisions for all U.S. domestic flights in January 2003 (Fridays only)

Model	(5)		(6)		(7)	
Airport size	Large		Medium		Small	
Outcome	Cancel	On-time	Cancel	On-time	Cancel	On-time
Logistical Variables						
Distance(miles)	-0.0001	-0.0001**	-0.0009**	-0.0001	-0.0004**	0.0009**
	(0.0001)	(0.0001)	(0.0003)	(0.0001)	(0.0001)	(0.0002)
Daily total flights for carrier j on route r	0.0318*	-0.0212*	0.0354	-0.0223**	0.0073	0.0382
	(0.0130)	(0.0088)	(0.0247)	(0.0068)	(0.0180)	(0.0404)
Daily total flights at origination airport	0.0001	-0.0003**	-0.0012	0.0002	-0.0012	-0.0032
	(0.0001)	(0.0001)	(0.0011)	(0.0003)	(0.0017)	(0.0039)
Daily total flights at destination airport	-0.0001	-0.0001*	-0.0002	-0.0004**	-0.0005**	0.0001
	(0.0001)	(0.0001)	(0.0002)	(0.0001)	(0.0001)	(0.0002)
Renormed scheduled departure time	-0.8796**	-0.0386	0.0093	-1.4028**	0.0905	-1.1165**
	(0.1842)	(0.1645)	(0.4573)	(0.1106)	(0.1398)	(0.3450)
Competition Variables						
Route market share	-0.0730	0.0635	-0.1618	0.0179	-0.3614*	0.2067
	(0.1157)	(0.0997)	(0.4227)	(0.1048)	(0.1451)	(0.3343)
Concentration at origination airport	0.2367	-0.1078	0.5121	0.2571	-0.2412	0.3066
	(0.2573)	(0.2206)	(0.5901)	(0.1479)	(0.2014)	(0.4464)
Concentration at destination airport	-0.4072**	0.5475**	-0.1792	0.8299**	-0.7632**	1.6481*
	(0.1433)	(0.1572)	(0.6484)	(0.1731)	(0.2718)	(0.7072)
Origination hub	-0.3813**	0.1519	-0.5958	-0.0316	--	--
	(0.1139)	(0.0847)	(0.5506)	(0.0950)		
Destination hub	-0.0822	0.0470	-0.2615	0.0813	0.1782	0.2567
	(0.0968)	(0.0821)	(0.2978)	(0.0680)	(0.1319)	(0.3269)
Weather Variables						
Daily precipitation at origination airport	0.0134**	-0.0331**	0.0065*	-0.0040**	0.0013	-0.0128*
	(0.0023)	(0.0037)	(0.0035)	(0.0009)	(0.0013)	(0.0051)
Daily precipitation at destination airport	0.0169**	-0.0394**	-0.0079	-0.0143**	0.0087**	-0.0616**
	(0.0023)	(0.0031)	(0.0072)	(0.0013)	(0.0010)	(0.0063)
Freezing at origination airport	0.4828**	-0.3973**	-0.0131	0.0758	-0.1949**	0.5277**
	(0.0732)	(0.1176)	(0.2326)	(0.0540)	(0.0761)	(0.1892)
Freezing at destination airport	0.1784*	-0.1508**	0.0163	-0.0663	-0.3309*	0.1553
	(0.0760)	(0.0520)	(0.2594)	(0.0537)	(0.1351)	(0.2933)
Freezing and precipitation at	-0.0838	-0.7531**	0.2921	-0.5509**	0.0081	-0.7223**

origination airport	(0.0711)	(0.0630)	(0.2504)	(0.0668)	(0.0680)	(0.1627)

Table 6 continued

Model	(5)		(6)		(7)	
Airport size	Large		Medium		Small	
Outcome	Cancel	On-time	Cancel	On-time	Cancel	On-time
Constant	-2.4863**	2.4325**	-2.2462**	3.2884**	-0.9037*	-0.2964
	(0.2008)	(0.2677)	(0.7793)	(0.1782)	(0.3883)	(1.0046)
Carrier fixed effects	Yes		Yes		Yes	
Chi-Square test of carrier effects	$?^{2}_{16}$? 159.8	$?^{2}_{16}$? 281.7	$?^{2}_{16}$? 25.4	$?^{2}_{16}$? 256.1	$?^{2}_{16}$? 32.8	$?^{2}_{16}$? 95.6
	$p < 0.0001$	$p < 0.0001$	$p = 0.0632$	$p < 0.0001$	$p = 0.0018$	$p < 0.0001$
Log-likelihood	-27,000		-8,602		-4,863	
Number of observations	54,139		21,057		9,843	

Note: Large airports are defined as airports with more than 400 daily total flights. Medium airports have between 100 and 400 daily total flights, while small airports have fewer than 100 daily total flights.

Table 6a : Marginal Effects for Nested Logit estimations from Table 6

Model	(5)			(6)			(7)		
Outcome	Cancel	Delay	On-time	Cancel	Delay	On-time	Cancel	Delay	On-time
Logistical Variables									
Distance(miles)	0.00000	0.00003	-0.00003	-0.00001	0.00002	-0.00001	0.00000	0.00000	0.00000
Daily total flights for carrier j on route r	0.00030	-0.00058	0.00028	0.00038	0.00213	-0.00250	-0.00076	-0.00640	0.00716
Daily total flights at origination airport	0.00000	0.00003	-0.00004	-0.00001	-0.00001	0.00002	0.00008	0.00072	-0.00080
Daily total flights at destination airport	0.00000	0.00003	-0.00003	0.00000	0.00004	-0.00004	0.00001	0.00011	-0.00012
Renormed scheduled departure time	-0.00142	0.09760	-0.09618	0.00830	0.14912	-0.15742	0.01500	0.10815	-0.12314
Competition Variables									
Route market share	-0.00084	-0.00060	0.00143	-0.00124	-0.00077	0.00201	0.00524	0.06924	-0.07449
Concentration at origination airport	0.00166	-0.01087	0.00921	0.00207	-0.03093	0.02885	0.00092	0.02633	-0.02725
Concentration at destination airport	-0.00676	-0.02846	0.03522	-0.00612	-0.08701	0.09313	-0.00749	0.00341	0.00408
Origination hub	-0.00244	0.02034	-0.01791	-0.00398	0.00753	-0.00355			
Destination hub	-0.00068	0.00254	-0.00186	-0.00231	-0.00682	0.00913	-0.00808	-0.07642	0.08450
Weather Variables									
Daily precipitation at origination (inches to hundredths)	0.00039	0.00290	-0.00329	0.00007	0.00038	-0.00045	0.00016	0.00116	-0.00132
Daily precipitation at destination (inches to hundredths)	0.00046	0.00335	-0.00381	0.00003	0.00157	-0.00160	0.00074	0.00500	-0.00574
Freezing at origination airport	0.00530	0.00093	-0.00622	-0.00054	-0.00797	0.00851	-0.00355	-0.01172	0.01527
Freezing at destination airport	0.00200	0.00086	-0.00286	0.00050	0.00694	-0.00745	0.00532	0.06740	-0.07273
Freezing and precipitation at origination airport	0.00795	0.10689	-0.11484	0.00528	0.05655	-0.06183	0.01088	0.08304	-0.09391
Freezing and precipitation at destination airport	0.00537	0.04423	-0.04960	0.00832	0.08543	-0.09375	0.02310	0.10129	-0.12439

Table 7 :Aircraft Characteristics : Nested Logit estimations of flight operation decisions for all U.S. domestic flights in January 2003(Fridays only)

Model	(8)		(9)		(10)	
Outcome	Cancel	On-time	Cancel	On-time	Cancel	On-time
Logistical Variables						
Distance(miles)	0.0000	0.0000	-0.0000	-0.0000	-0.0001	-0.0001**
	(0.0001)	(0.0000)	(0.0002)	(0.0000)	(0.0002)	(0.0000)
Daily total flights for carrier j on route r	-0.0254**	0.0096*	0.0131	0.0233**	0.0861**	0.0108**
	(0.0077)	(0.0046)	(0.0187)	(0.0033)	(0.0236)	(0.0035)
Daily total flights at origination airport	0.0003**	-0.0001*	0.0004**	-0.0002**	0.0007	-0.0002**
	(0.0001)	(0.0000)	(0.0001)	(0.0000)	(0.0001)	(0.0000)
Daily total flights at destination airport	0.0011**	0.0004**	0.0002*	-0.0003**	0.0007	-0.0003**
	(0.0001)	(0.0001)	(0.0001)	(0.0000)	(0.0001)	(0.0000)
Renormed scheduled departure time	0.6642**	-0.7391**	-1.3641**	-1.0094**	-0.2731	-0.9940**
	(0.1156)	(0.0797)	(0.2991)	(0.0582)	(0.3096)	(0.0582)
Capacity	-0.6021**	-1.3972**	-2.4879**	-0.8761**	-2.3231**	-0.8906**
	(0.1881)	(0.1393)	(0.4779)	(0.0997)	(0.5406)	(0.1059)
Potential revenue Per flight ($)	-5.75E-06	-4.2E-06**	-6E-06	4.51E-07	3.45E-5	-8.4E-07
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Competition Variables						
Route market share	0.5853**	0.2358**	-0.3974	-0.0216	-0.0574	-0.0285
	(0.1041)	(0.0691)	(0.2387)	(0.0491)	(0.2614)	(0.0506)
Concentration at origination airport	-0.0452	0.2394*	-0.0007	0.0828	0.3695	0.1743*
	(0.1139)	(0.1013)	(0.3733)	(0.0718)	(0.4535)	(0.0768)
Concentration at destination airport	-0.0519	0.0848	-0.5625	0.2144**	0.0419	0.3029**
	(0.1791)	(0.1031)	(0.3896)	(0.0707)	(0.4428)	(0.0749)
Origination hub	-0.1296	-0.2597**	-0.2573	-0.1533**	-0.3249	-0.1547**
	(0.0806)	(0.0468)	(0.2023)	(0.0334)	(0.2782)	(0.0374)
Destination hub	-0.0991	-0.0715	0.1613	0.0119	-0.1034	0.0181
	(0.0795)	(0.0486)	(0.1955)	(0.0335)	(0.2711)	(0.0372)
Weather Variables						
Daily precipitation at origination airport	-0.0008	-0.0090**	0.0018	-0.0090**	0.0068	-0.0093**
	(0.0012)	(0.0008)	(0.0041)	(0.0007)	(0.0055)	(0.0007)
Daily precipitation at destination airport	0.0066**	-0.0084**	0.0064*	-0.0121**	0.0132**	-0.0123**
	(0.0009)	(0.0009)	(0.0031)	(0.0007)	(0.0041)	(0.0007)
Freezing at origination airport	-0.0183	0.0318	0.0857	-0.0043	0.0631	0.0281

	(0.0548)	(0.0395)	(0.1469)	(0.0271)	(0.1699)	(0.0278)

Table 7 continued

Model	(8)		(9)		(10)	
Outcome	Cancel	On-time	Cancel	On-time	Cancel	On-time
Freezing at destination airport	0.5931**	0.1045*	-0.2008	-0.0635*	-0.0143	-0.0542*
	(0.0895)	(0.0497)	(0.1476)	(0.0269)	(0.1648)	(0.0276)
Freezing and precipitation at	-0.3819**	-1.0576**	-0.0517	-0.8382**	0.5072**	-0.8292**
origination airport	(0.0616)	(0.0460)	(0.1483)	(0.0303)	(0.1650)	(0.0308)
Freezing and precipitation at	0.8871**	0.0137	0.1682	-0.5724**	0.5329**	-0.5426**
destination airport	(0.0585)	(0.0832)	(0.1458)	(0.0309)	(0.1585)	(0.0314)
Aircraft variables						
Number of seats	-0.0043**	0.0007*	-0.0013	-0.0004	-0.0174**	-0.0005
	(0.0010)	(0.0004)	(0.0027)	(0.0004)	(0.0035)	(0.0004)
Age of aircrafts	-0.0018	-0.0000	0.0465**	-0.0121**	0.0168*	-0.0085**
	(0.0031)	(0.0022)	(0.0117)	(0.0020)	(0.0127)	(0.0023)
Constant	-4.7183**	3.1356**	0.125631	2.8062**		
	(0.2279)	(0.1369)	(0.4592)	(0.0998)		
Carrier fixed effects	No		No		Yes	
Chi-Square test of carrier fixed effects					$?^{2}_{16}$? 3152.8	$?^{2}_{16}$? 1206.1
					p<0.0001	p<0.0001
Aircraft manufacturer fixed effects			Yes		Yes	
Chi-Square test of manufatured fixed effects			$?^{2}_{4}$? 144.5	$?^{2}_{4}$? 274.2	$?^{2}_{4}$? 5964.4	$?^{2}_{4}$? 969.6
			p<0.0001	p<0.0001	p<0.0001	p<0.0001
Log-likelihood	-28,485		-28,468		-27,984	
Number of observations	62,841		62,841		62,841	

Table 7a : Marginal Effects for Nested Logit estimations from Table 7

Model	(8)			(9)			(10)		
Outcome	Cancel	Delay	On-time	Cancel	Delay	On-time	Cancel	Delay	On-time
Logistical Variables									
Distance (miles)	0.000000	0.00000	0.00000	0.000000	0.00001	-0.00001	0.000000	0.00001	-0.00001
Daily total flights for carrier j or route r	-0.000331	-0.00271	0.00304	-0.000028	-0.00305	0.00308	0.000369	-0.00181	0.00144
Daily total flights at origination airport	0.000004	0.00003	-0.00003	0.000003	0.00003	-0.00003	0.000004	0.00003	-0.00003
Daily total flights at destination airport	0.000011	0.00002	-0.00003	0.000003	0.00004	-0.00004	0.000005	0.00003	-0.00004
Renormed scheduled departure time	0.010886	0.13265	-0.14354	-0.002970	0.13617	-0.13320	0.002435	0.12884	-0.13128
Capacity	-0.000312	0.14118	-0.14086	-0.009466	0.12434	-0.11488	-0.007707	0.12534	-0.11763
Potential revenue Per flight ($)	-4.52E-08	1.93E-07	-1.48E-07	-3.36E-08	-3E-08	6.32E-08	1.67E-07	-5.68E-08	-1.11E-07
Competition Variables									
Route market share	0.005492	0.00471	-0.01020	-0.002008	0.00467	-0.00266	-0.000166	0.00393	-0.00377
Concentration at origination airport	-0.001615	-0.03294	0.03455	-0.000352	-0.01064	0.01099	0.001104	-0.02412	0.02302
Concentration at destination airport	-0.000976	-0.01380	0.01478	-0.003871	-0.02489	0.02877	-0.000938	-0.03907	0.04000
Origination hub	-0.000257	0.02520	-0.02494	-0.000716	0.02092	-0.02020	-0.000965	0.02140	-0.02043
Destination hub	-0.000783	0.00321	-0.00242	0.000802	-0.00229	0.00149	-0.000560	-0.00183	0.00239
Weather Variables									
Daily precipitation at origination airport (inches to hundredths)	0.000033	0.00109	-0.00112	0.000048	0.00115	-0.00120	0.000068	0.00116	-0.00123
Daily precipitation at destination airport (inches to hundredths)	0.000115	0.00147	-0.00158	0.000084	0.00152	-0.00160	0.000109	0.00152	-0.00163
Freezing at origination airport	-0.000354	-0.00512	0.00547	0.000472	0.00015	-0.00063	0.000195	-0.00391	0.00372
Freezing at destination airport	0.006186	0.02177	-0.02796	-0.000794	0.00912	-0.00833	0.000136	0.00703	-0.00717
Freezing and precipitation at origination airport	0.000594	0.11122	-0.11182	0.003244	0.10795	-0.11120	0.005529	0.10399	-0.10951
Freezing and precipitation at destination airport	0.009912	0.05059	-0.06050	0.003290	0.07275	-0.07604	0.004574	0.06709	-0.07166
Aircraft variables									
Number of seats	-0.000052	-0.00035	0.00040	-0.000006	0.00006	-0.00005	-0.000081	0.00015	-0.00007
Age of aircrafts	-0.000020	-0.00010	0.00012	0.000296	0.00133	-0.00163	0.000112	0.00102	-0.00113