ELSEVIER

Available online at www.sciencedirect.com

ScienceDirect

Procedia Social and Behavioral Sciences 17 (2011) 150–168

Procedia Social and Behavioral Sciences

19th International Symposium on Transportation and Traffic Theory

# Estimating time of day demand with errors in reported preferred times: An application to airline travel

Raúl Brey[a], Joan L. Walker[b,]*

[a]*Universidad Pablo de Olavide, Ctra. de Utrera, km. 1 41013, Sevilla, SPAIN*
[b]*University of California at Berkeley, 111 McLaughlin Hall, Berkeley CA 94720, USA*

**Abstract**

An essential element of demand modeling in the airline industry is the representation of time of day demand—the demand for a given itinerary as a function of its departure or arrival times. It is an important datum that drives successful scheduling and fleet decisions. There are two key components to this problem: the distribution of the time of day demand and how preferred travel time influences itinerary choice. This paper focuses on estimating the time of day distribution. Our objective is to estimate it in a manner that is not confounded with air travel supply; is a function of the characteristics of the traveler, the trip, and the market; and accounts for potential measurement errors in self-reported travel time preferences. We employ a stated preference dataset collected by intercepting people who were booking continental US trips via an internet booking service. Respondents reported preferred travel times as well as choices from a hypothetical set of itineraries. We parameterize the time of day distribution as a mixture of normal distributions (due to the strong peaking nature of travel time preferences) and allow the mixing function to vary by individual characteristics and trip attributes. We estimate the time of day distribution and the itinerary choice model jointly in a manner that can account for measurement error in the self-reported travel time preferences. We find that the mixture of normal distributions fits the time of day distribution well and is behaviorally intuitive, and the strongest covariates of travel time preferences are party size and time zone change. The methodology employed to treat self-reported travel time preferences as potentially having error contributes to the broader transportation time of day demand literature, which either assumes that the desired travel times are known with certainty or that they are unknown. We find that the error in self-reported travel time preferences is statistically significant and impacts the inferred time of day demand distribution.


*Keywords: airline demand; time of day demand; schedule delay; discrete choice; latent variables; finite mixture models*

## 1. Introduction

While long a stalwart component of airline schedule and fleet planning, time of day demand is thus far a poorly described phenomenon in the airline industry. Yet, it is critical to understand the passenger choices that drive successful scheduling and fleet decisions; for example, when to schedule a departure in a low frequency market, how to spread flights over the day in a high frequency market, or the impact on market share when a new technology is introduced.

1877–0428 © 2011 Published by Elsevier Ltd. Open access under CC BY-NC-ND license.
doi:10.1016/j.sbspro.2011.04.512

There are two essential components to understanding time of day demand. The first is the choice of itinerary, conditional on the desired travel time. This aspect has received relatively more attention in the literature, and is generally captured by introducing a schedule delay variable in the utility function. The second component is the preferred time of day distribution, which has received considerably less attention, and estimating this distribution is the focus of this paper.

The next section presents background information, which is followed by a literature review and discussion of the contributions of this paper. The proposed modeling framework is then described in detail and an empirical application is presented.

### *1.1. Background*

Common experience suggests that the departure (or arrival) times of alternative flights are important in determining the choice of flights, along with other attributes such as travel time and fare. In terms of utility, we can assert that there is a value associated with the timing of an airline schedule, in the sense that some flights that depart at certain times may be more or less preferred than others. The Boeing Decision Window Model (Boeing, 1996) asserts, for example, that flights that depart within a given window of time are acceptable and have no disutility, while those that do not are unacceptable, that is, have infinite disutility.

It is typical to reflect the attractiveness of certain flights based on time of day using a schedule delay variable, which is a measure of the difference between desired and scheduled departure times. To gain some intuition on the form in which the schedule delay function could take, consider **Error! Reference source not found.**. The horizontal axis is the time of day, and the vertical axis is disutility. Two flights are shown, one in the morning departing at time $t_1$ and one in the evening departing at $t_2$. The desired departure time for the passenger is towards noon, and indicated by $\tau$. The curve represents the disutility of departing at a time different than the desired one, which gets larger as one moves away from the desired departure time and may be asymmetric (in the figure, departing late is more costly than departing early). The function is then parameterized via the functions for earlier than desired and later than desired departure times noted as *GE()* and *GL()*, respectively. *GE()* and *GL()* are functions of the desired time $\tau$, and the scheduled time $t_i$. In addition, there may be a band near the desired departure time where the traveler is indifferent, and this is denoted as the window between *a* and *b* in the figure. The problem then reduces to determining an appropriate functional form for *GE()* and *GL()* and estimating the parameters using travel data.

**Error! Reference source not found.** assumes that the desired departure time ($\tau$) is known. Such information is sometimes collected in passenger surveys and therefore can be incorporated in itinerary choice models, conditional on desired departure times. However, if such models are to be used for airline scheduling, a model of preferred time of day is required. Figure 2 shows what a distribution for preferred time of day might look like. This figure suggests strong morning and evening peaks, although the distribution would be expected to vary based on traveler, trip, and market characteristics. Such distributions are the focus of this paper, and we estimate jointly the time of day distribution and the itinerary choice model with schedule delay. The next section will review the literature on these topics.

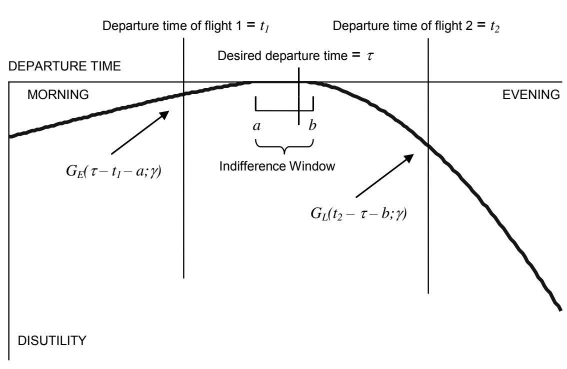



Figure 1: Hypothetical schedule delay curve.

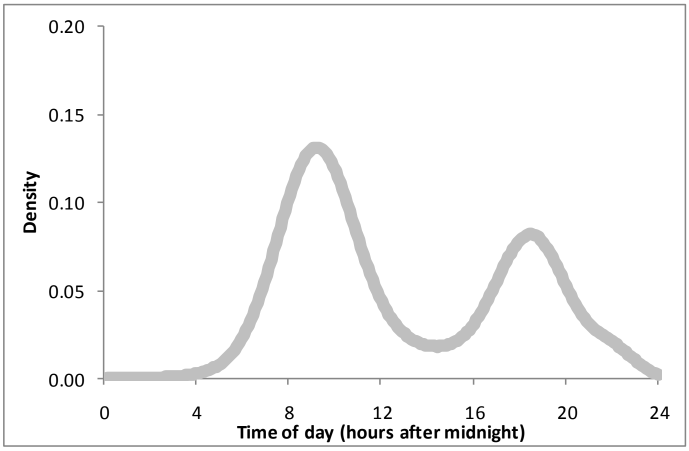



Figure 2: Hypothetical distribution of preferred time of day for travel.

*1.2. Literature*

Temporal dimensions of travel have been a focus in the transportation literature since the 1960s (and probably earlier). Vickrey (1969) introduced the concept that congested conditions in urban settings lead to travelers making tradeoffs between how much time to spend traveling (in traffic) versus arriving either earlier or later at the destination. Abrahams (1983) described early work in the airline industry by Douglas and Miller (1974), Eriksen (1978), and Swan (1979) who developed empirical estimates of passenger delay for an airline schedule, incorporating both the difference between the scheduled flights and when travelers want to fly and the chance of the desired flight being full.

Small (1982) was the first to develop empirical estimates of schedule delay based on individual choices. His focus was on commuters' decisions of when to travel to work. Small specified his "barebones" utility for traveling at a specific time $t$ as

$$U_t = \cdots \alpha(schedule\ delay, early)_t + \beta(schedule\ delay, late)_t + \gamma(late\ arrival\ dummy)_t \quad . \tag{1}$$

This is a linear function with different slopes for late and early arrivals and a discrete fall in utility once one is late (important for the work commute). Estimated parameters $\alpha, \beta,$ and $\gamma$ are assumed to be negative. Small also estimated and found many significant variations, including quadratic schedule delay terms and interactions of the schedule delay variables with household structure, job type, flexibility of arrival time, and an indicator for the carpool mode. Small's specification is used widely in urban passenger transportation (see Bates et al., 2001, and Li et al., 2010, for reviews) and also appears in the freight logistics literature (see Sathaye et al., 2009), and air transportation models (reviewed in more detail below). Important extensions to Small's model include the introduction of an "indifference band" around the preferred travel time (Mahmassani and Chang, 1986), introduction of uncertainty of travel times due to congestion (Noland and Small, 1995), and dynamic variations by time of day where willingness to pay to be at a particular place at a particular time varies throughout the day (Tseng and Verhoef, 2008).

Schedule delay appears in the air travel literature in aggregate demand models (Abrahams, 1983), network scheduling decisions (Brueckner and Zhang, 2001), airport operations (Hansen 2002), fleet mix analysis (Ryerson and Hansen, 2010), and disaggregate demand models (Proussaloglou and Koppelman, 1999; Adler et al. 2005; Garrow et al., 2006; Parker and Walker, 2006; and Koppelman et al., 2008). Typically the specification is along the lines of Small (1982) presented above, although without the discrete drop in disutility for being late and often with symmetric response to early versus late (particularly in the aggregate models such as those by Brueckner and Zhang, 2001; and Ryerson and Hansen, 2010). The disaggregate demand models are most relevant to this work. All are itinerary choice models and all are estimated using stated preference data, except for Koppelman et al. (2008) who used bookings data. The stated preference-based works all found differences in schedule delay sensitivity between business and non-business travelers. Proussaloglou and Koppelman (1999) had self-reported information on the preferred travel time window, and they found greater sensitivity to schedule delays outside of the preferred window. Adler et al. (2005) used a mixed logit specification and found that there was significant heterogeneity in the sensitivity to schedule delay. Garrow et al. (2006) used a stepped delay function with 4 windows: 4 or more hours before desired time, between 4 hours before and 2 hours after desired time, 2 to 4 hours after desired time, and 4 or more hours after desired time. Parker and Walker (2006) used a Box-Cox schedule delay function and included an indifference window around the desired time. Koppelman et al. (2008) found that a logistic delay function was superior to other linear and non-linear schedule delay function specifications. From such models one can obtain estimates of the willingness to pay for schedule delay. For example, Proussaloglou and Koppelman (1999) estimated business travelers were willing to pay $60/hour and leisure travelers willing to pay $17/hour, although they suggested an ad-hoc reduction to $40/hour and $10/hour, respectively, due to a presumed endogeneity of the fare variable. Adler et al. (2005) estimated business travelers to value schedule delay at $30.3/hour (with a standard deviation of $22.9) and non-business travelers at $4.8/hour (with standard deviation of $5.7).

The explicit schedule delay function as described above is used when the preferred time of travel is known, for example it is the desired arrival time at work for a commute context or otherwise gathered via a survey. However, often the desired travel time is not known and so the concept of schedule delay is included in a behavioral model in an approximate way by including variables that vary across time. For example, in the urban travel demand literature, it is typical to find time of day demand treated as a choice among discrete blocks of time during the day, such as morning, afternoon, and evening periods (e.g., Bowman and Ben-Akiva, 2001). Rather than dividing the day into discrete periods, continuous approaches have also been employed, for example the trigonometric functions in Abou Zeid et al. (2006). Further, some have considered how the marginal utilities of activities vary over time throughout the day (e.g., Ettema and Timmermans, 2003). A related approach was applied by Mehndiratta and Hansen (1997) for air travel, and found that business travelers' willingness to pay for schedule preferences varied depending on whether the schedule delay cut into work time, leisure time, or sleep time. Koppelman et al. (2008) in their air itinerary share models employed the trigonometric approach of Abou Zeid et al. (2006) to estimate a continuous time of day distribution. They found that this was superior (statistically and behaviorally) to a discrete function with hourly dummy variables. They also found that the distribution varied significantly for passengers on their outbound leg (greater preference to depart in the morning) versus passengers on their inbound leg (greater preference to depart in the afternoon). At the other extreme, Brueckner and Zhang (2001) assumed that desired arrival times were uniformly distributed. The form of the behavioral model itself, within which the preferred time of travel is approximated, varies from simple logit models (as in Abou Zeid et al., 2006), to ordered-GEV (Small, 1987), to error-components logit (de Jong et al., 2003; Hess et al., 2007), to continuous cross-nested logit (Lemp et al., 2010), and to hazard functions (Bhat and Steed, 2002).

Specific to the airline literature, there is a line of work beginning with Douglas and Miller (1974), Eriksen (1978), and Swan (1979) (as described in Abrahams, 1983) that assumed the time of day distribution could be approximated from booked seats on a market with high frequency of flights. The idea was that travelers in such a market were able to find a flight close to their desired departure time, although adjustments were also made based on load factors to reflect the chance of the desired flight being full. Abrahams (1983), Hansen (2002), and Ryerson and Hansen (2010) assumed time of day distributions along these lines. A potential issue from using bookings to infer time of day is that bookings result from both supply and demand forces, and the travelers make tradeoffs between attributes like schedule time and fare (particularly important due to revenue management policies). Koppelman et al. (2008) also based their preferred time of day distribution on bookings; however, they did so within an itinerary share model and so they were able to capture, at least to some extent, such tradeoffs. However, there still may be identification issues from inferring aggregate time of day demand from bookings data as it is a reflection of both supply and demand. Further, other than the outbound and inbound analysis in Koppelman et al. (2008) described above, there is minimal work on how the time of day distribution varies by person, trip, and market variables.

### *1.3. Contributions of this paper*

Our focus is on the time of day distribution, and our objective is to estimate it in a manner that is not confounded with air travel supply; is a function of the characteristics of the traveler, the trip, and the market; and accounts for potential measurement errors in self-reported travel time preferences. To disentangle travel time preference from supply, we use a stated preference itinerary choice dataset. The data were collected via an intercept survey given to persons booking a flight on an internet booking service. Respondents were provided hypothetical itineraries for the city pair on which they were searching and then asked to report their preferred alternative. They also reported, among other things, their preferred times of travel. We use this information to infer the preferred time of day distribution, but we treat preferred travel time as a latent variable in a jointly estimated time of day and itinerary choice model. This specification is motivated by Boeing's concern that statements of preferred travel times may be influenced by the provided schedule (i.e., travelers have a sense of when airlines offer flights and base their stated "preferences" on this knowledge). Therefore there may be measurement error in their statements. Our model specification accounts for such measurement error by using a choice and latent variable model framework (expanded on below). This is distinctly different than the current time of day literature which either assumes that preferred time of travel is known with certainty or is unknown and approximated with time-specific variables. Another advantage of the specification is that it would allow inference of the travel time distribution using a combination of respondents who report a preferred travel time and respondents who do not report such information. Further, the estimated travel time distribution is a function of traveler and trip characteristics and we explore the heterogeneity of the distribution. For the functional form of the preferred travel time distribution, we propose the use of a mixture of normal distributions, which both fits the data well and is behaviorally attractive in the way it reflects the peaked nature of air travel time of day decisions. This multimodal distributional assumption for a continuous latent variable is in contrast to the existing choice and latent variable models in the literature, which assume (in all applications of which we are aware) that the latent variables are normally distributed (see, e.g., Ashok et al., 2002; Ben-Akiva et al., 2002; Bolduc et al. 2008; Temme et al., 2008; Walker et al., 2010; and Yáñez et al., 2010).

## 2. Methodology

### *2.1. Model framework and motivation*

Developing an operational model requires using data collected from airline travelers to estimate the parameters of the time of day curve and the itinerary choice model. There is an interesting issue in terms of how to infer the travel time preferences. It may be possible to infer travel time preferences simply from the observed itinerary choices that passengers make, because the utility of a flight is a function of schedule delay, which is a function of desired travel time. However, this is not likely to provide sufficient information to estimate rich travel time demand functions (e.g., as functions of attributes of the trip or characteristics of the traveler) and so we would gain limited knowledge of the market. Indeed, in the empirical case that is presented later, we could not infer even the most basic travel time distribution from the observed itinerary choices alone. It is also possible to obtain travel time preferences stated by the traveler (which we have in our dataset) and use these to estimate the time of day curves directly. This would enable estimation of richer demand functions. However, as hypothesized above, stated travel time preferences may be a function of when airlines offer flights (that is, current supply conditions) and therefore may be different than when the traveler truly prefers to fly.

Due to the potential error in self-reported travel time preferences, we choose to treat preferred travel time as a latent variable in the itinerary choice model. The framework for such a model is shown in Figure 3 (based on Ben-Akiva et al., 2002; and Walker and Ben-Akiva, 2002). In this figure, the latent variables are in ellipses and observed variables are in boxes, the structural (or behavioral) relationships shown with solid arrows, and the measurement relationships as dashed arrows. The utility of any itinerary $(U)$ is a function of the attributes of the itinerary and characteristics of the trip/traveler $(X)$ and the preferred time of day $(\tau)$, where the latter is considered latent. The stated itinerary preference $(I)$ is then an indicator of these underlying utilities, and the relationship between $U$ and $I$ is a measurement equation (we assume the typical assumption of utility maximization). Analogously, the stated preferred travel time $(d)$ is an indicator of the latent preferred travel time $\tau$, and there is another measurement equation that links these variables. Finally, the latent preferred travel time $(\tau)$ is assumed to be a function of the characteristics of the traveler and trip $(X)$, and this relationship is the time of day distribution model. To summarize, the framework allows for joint estimation of the time of day distribution and the itinerary choice model, where the preferred time of day is a latent variable.

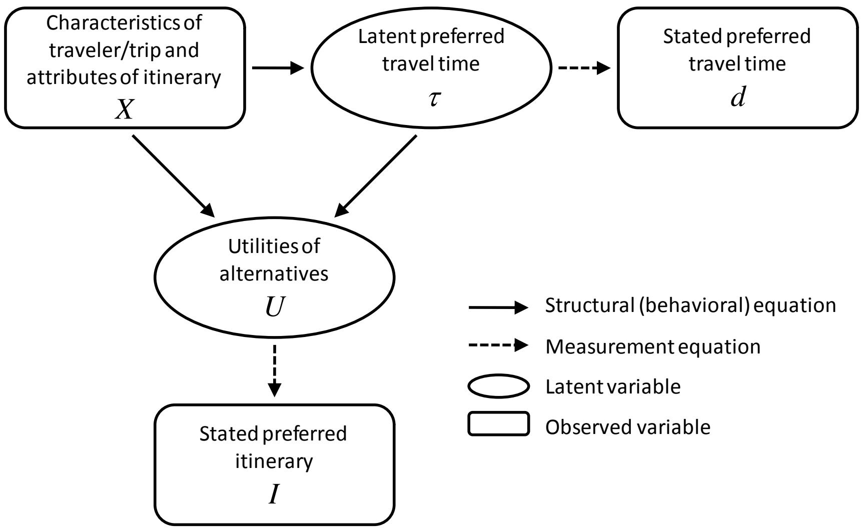



Figure 3: Framework for the jointly estimated travel time distribution and itinerary choice model.

*2.2. Equations*

The framework requires specification of the time of day distribution, the itinerary choice model, and the relationship between them. The time of day curve is denoted

$$f(\tau|X), \tag{2}$$

or the probability density function of the preferred time of day $\tau$, which is conditional on a set of covariates $X$ (characteristics of the traveler/trip). The itinerary choice model is denoted as

$$P(I|X,\tau), \tag{3}$$

or the probability with which a traveler will choose a particular itinerary $I$ conditional on preferred time of travel $\tau$ and a set of covariates $X$ (attributes of the itineraries and characteristics of the traveler/trip). This probability is conditional on a particular choice set (the set of itineraries from which the decision maker chooses), although we do not include such notation in the equations. The joint distribution of travel time preference and itinerary choice is then

$$g(I,\tau|X) = P(I|X,\tau)f(\tau|X). \tag{4}$$

To resolve the estimation issue highlighted above, we treat travel time preferences as a latent variable. Because $\tau$ is known only to a distribution, the unconditional itinerary choice model is obtained by integrating over the time of day curve, or

$$P(I|X) = \int P(I|X,\tau)f(\tau|X)d\tau. \tag{5}$$

The time of day model $f(\tau|X)$ requires the development of a probability density function that is multimodal and can be a function of the covariates $X$. For this purpose, we use a mixture of normal density functions (a finite mixture model), denoted as

$$f(\tau|X) = \sum_{k=1}^{K} prob(k|X)n(\mu_k,\sigma_k), \tag{6}$$

where $K$ is the total number of normal distributions that are mixed, $k\ (= 1,\dots,K)$ is a specific normal distribution, $prob(k|X)$ is the mixing function (specified as a logit probability) and may be a function of covariates $X$, and $n(\mu_k,\sigma_k)$ is a normal distribution with mean $\mu_k$ and standard deviation $\sigma_k$ (both estimated parameters). This provides a parametric, estimable and flexible model of the time of day curve. Figure 4 shows an example of a mixture of two normal distributions in the case of a morning peak centered on 9 AM and an evening peak centered on 6 PM. This example mixing function ($prob(k|X)$) is naive in that it has no covariates; the 9 AM peak has a weight of 60% and the 6 PM peak a weight of 40%. The thick, grey line in the figure is the mixture ($f(\tau|X)$), and is calculated as 60% of a $N(mean = 9, variance = 4)$ and 40% of a $N(mean = 18, variance = 9)$. In the empirical application, we mix over additional (i.e., more than 2) normal distributions to better represent the time of day curve. We also use covariates in the mixing function to explain, for example, what types of trips would have a stronger preference for the morning peak. The obvious alternative specification to the mixture of normals would be a trigonometric function such as in Abou Zeid et al. (2006) and Ben-Akiva and Abou Zeid (2010), and such a specification could be substituted into our framework. We choose to use the mixture of normals because it seems a natural representation of the strong peaking nature of air travel time of day demand. Limitations of the mixture of normals are that the distribution is not constrained to the 24 hour time period that is modeled (the tails extend to $+/-\infty$) and there is a discontinuity at the beginning and end of the day (demand at 0:00 is not constrained to equal 24:00). However, neither of these appears to be a major issue in our estimation results.

The model of itinerary choice follows the convention of random utility models of choice, where the utility is a function of the attributes of the itinerary (e.g., fare, duration, stops, departure/arrival time), of the traveler (e.g., age, income), and of the trip (e.g., purpose, who pays, party size). The key to the time of day aspect in the itinerary choice model is the schedule delay function. Our final models employ a linear schedule delay function in which the slopes are allowed to be different for early schedule delay and late schedule delay. We explored other specifications, but found the linear performed well statistically and the model framework was robust to the schedule delay function.

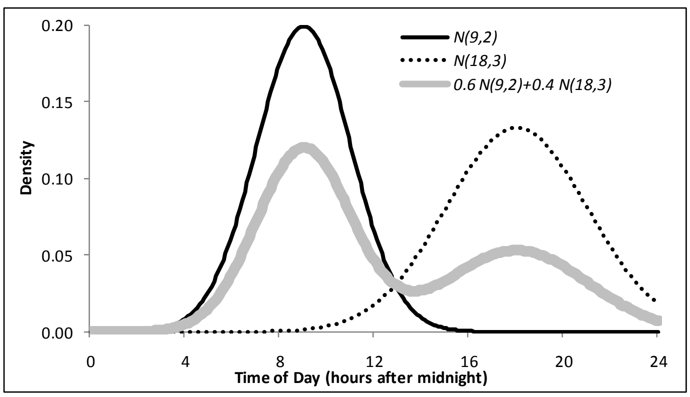



Figure 4: Example of time of day distribution specified as a mixture of normal distributions.

### *2.3. Model Estimation*

We estimate the parameters of the time of day curve and the itinerary choice model using a dataset collected by Boeing through an internet booking service (more detail below). We use a maximum simulated likelihood estimator (programmed in Gauss) with Halton draws.

As described above, the itinerary choice model, $P(I|X, \tau)$, is a logit formulation with an asymmetric, linear schedule delay function and the time of day distribution $f(\tau|X)$ is a mixture of normal distributions as in equation 6. Treating the preferred travel time $\tau$ as a latent variable, the itinerary choice model is $P(I|X) = \int P(I|X, \tau) f(\tau|X) d\tau$. While we could stop here in estimation, such a model is difficult to identify. This is an issue of empirical (rather than mathematical) identification in that the choice indicator alone ($I$) does not contain enough information to infer both the parameters in the utility function and the parameters of the time of day distribution. Further, the self-reported preferred time of travel contains information on the time of day distribution, which we want to make use of. Therefore, we use the stated preferred time of travel as an indicator of the latent travel time preference (as presented in Figure 3). This involves estimating the joint distribution of the observed itinerary choice ($I$) and the stated preferred time of travel ($d$, potentially different than the latent desired travel time $\tau$), where this joint distribution is

$$h(I, d|X) = \int P(I|X, \tau) k(d|X, \tau) f(\tau|X) d\tau. \tag{7}$$

This is the full equation for the model as presented in Figure 3. Such a specification provides a mechanism for correcting bias that may be present in the survey responses regarding preferred travel times.

Note that equation 5 (rather than equation 7) is used when applying the model for analysis or forecasting. The issue is that the measurements (stated preferred travel time) are only known for the estimation sample and will not generally be known for a forecast situation. However, the strength of this approach is that the measurements are not needed for a forecast; once the time of day distribution is estimated then it is the distribution (rather than the self-reported times) that is used directly in the forecast.

## 3. Application

This section details the empirical application. First the data are described and then estimation results are presented.

### *3.1. Data*

The data used in this analysis comes from a stated preference internet choice survey conducted by the Boeing Company in the fall of 2004. Before describing the data in detail, a word about the use of stated preference data. While a significant amount of the airline and time of day literature cited above uses stated preference data, there are also debates over the use of stated preference data for understanding trip timing decisions. Borjesson (2008) in studying both revealed and stated preference data collected for the Stockholm congestion pricing trial, found that there were systematic differences between the stated and revealed responses, suggesting that the stated preference data were less reliable. However, Li et al. (2010) found that stated preference data related to trip timing (also in an urban context) were reliable, although one has to be careful to ask the question in a manner that is consistent to how travelers process data under real scenarios. For example, rather than giving respondents the travel time and travel time variance for a particular trip, it is better to give respondents multiple travel time scenarios to represent the travel time distribution. A beneficial aspect of the data we use in this application is that it is collected via intercepting customers of an internet booking service. The stated preference questions were then designed around the specific city pair on which the traveler was searching. Therefore, the respondents were already in the frame of mind of making the choice presented in the hypothetical scenario, and the data presented are similar to what the booking search engine would return (multiple itineraries with fare, time, airline, etc.).

The internet booking service that was used takes a specific user request for travel in a city pair and interrogates the web sites that offer tickets in that market, returning to the user a compiled list of available options. One to two minutes of time elapse while that interrogation is taking place, during which randomly selected customers were recruited to be surveyed. (Note that while the internet site would allow Boeing to conduct the survey, the company would not share the actual booking data.) Only persons searching for round trips within the continental US were recruited for the survey. The survey was designed by Boeing staff with the assistance of Jordan Louviere of the University of Technology, Sydney. Garrow et al. (2006) provide further information on the experimental design. A typical page of the survey instrument is shown in Figure 5; this is for a customer searching for a flight from Chicago to San Diego. The respondent was offered three choices based on the city pair he or she submitted to the booking service, and was asked to rank the available choices as well as given the option to decline all of the stated options. Only the outbound leg of the flight (home to the destination) is considered in the survey. In addition to the stated choices (one per respondent), demographic data were collected included age, gender, income, occupation, education, and home zip code. Situational variables were also collected, including: travel dates and times (from the booking request), whether departure time or arrival time is more important, and the desired departure or arrival time, trip purpose (business or leisure), who is paying for the trip (self, reimbursed by company, company direct), and the number of people in the travel party.

A total of 3319 valid respondents to the survey were captured. 70.9 percent of the trips were leisure, 17.4% business, 9.3% both business and leisure, and 2.4% attending conference/seminar/training. These trips represent 436 origin-destination city pairs in the continental United States. Each respondent provided both a first choice among the three alternatives and a last choice among the three alternatives, yielding 5460 observed choices.

**Pick Your Preferred Flight**

Three flight options are described for your trip from Chicago to San Diego . These are options that might be available on this route or might be new options actively being considered for this route as well as replacing some options that are offered now. The options differ from each other in one or more of the features described on the left.

Please evaluate these options, assuming that eveything about the options is the same except these particular features. Indicate your choices at the bottom of the appropriate column and press the Continue button.

| FEATURES | Non-Stop (Option 1) | 1 Stop (Option 2) | 1 Stop (Option 3) |
|---|---|---|---|
| Departure time (local) | 6:00 PM | 4:30 PM | 6:00 PM |
| Arrival time (local) | 8:14 PM | 8:44 PM | 9:44 PM |
| Total time in air | 4 hr 14 min | 4 hr 44 min | 4 hr 44 min |
| Total trip time | 4 hr 14 min | 6 hr 14 min | 5 hr 44 min |
| Legroom | typical legroom | 2-in more of legroom | 4-in more of legroom |
| Airline [Airplane] | Depart Chicago Continental Airlines [B737] to San Diego | Depart Chicago Southwest Airlines [A320], connecting with Southwest Airlines [MD80] to San Diego | Depart Chicago Northwest Airlines [MD80], connecting with American Airlines [DC9] to San Diego |
| Fare | $565 | $485 | $620 |

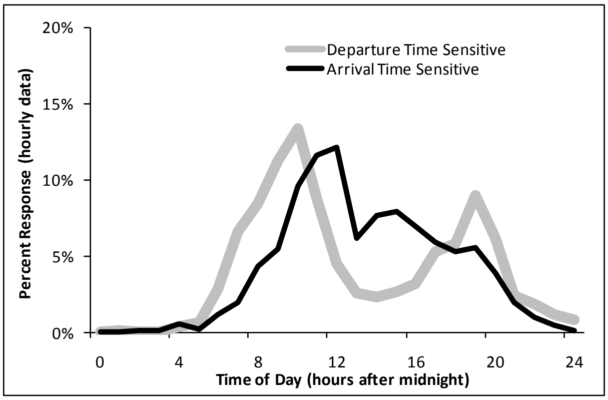
1. Which is MOST attractive? Option 1 Option 2 Option 3

2. Which is LEAST attractive? Option 1 Option 2 Option 3

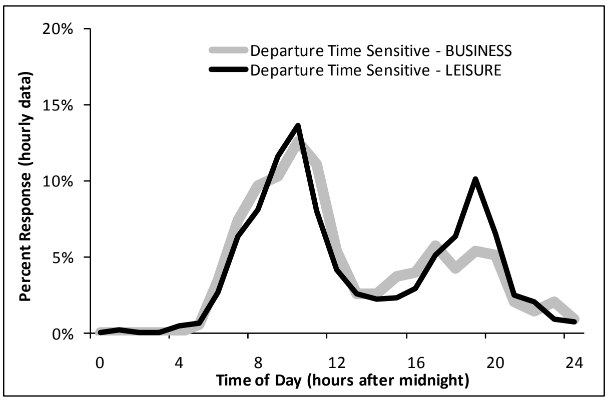
3. If these were the ONLY three options available, I would NOT make this trip by air. Yes No

Figure 5: Example of survey instrument tailored for a person looking to travel from Chicago to San Diego.

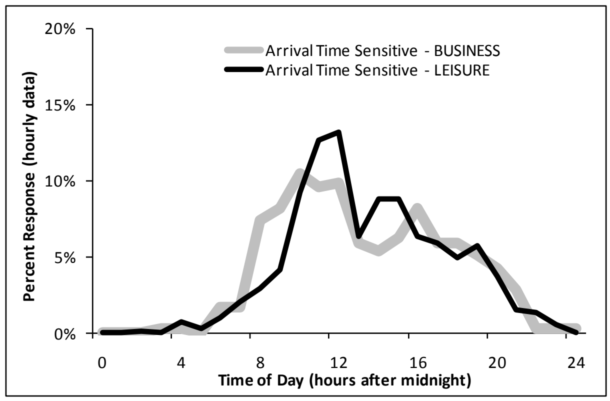
An important aspect for this work is that respondents reported whether they felt the departure time of a flight was more important to a choice of flight option or the arrival time was more important, and then they also provided a preferred travel time (either departure or arrival, whichever applied). Of the respondents, 52% felt departure time was more important, 40% indicated arrival time was more important, and 8% did not answer the question. Figure 6 displays histograms of the stated preferred travel times (aggregated to hourly data), indicating how the distribution varies based on whether someone is departure time or arrival time sensitive, a business or leisure traveler, and the direction of travel. Direction of travel is based on the time zone change between the home and the destination; an origin-destination (OD) pair that is in the same time zone is considered a NORTH-SOUTH flight (direction independent, that is the home can either be north or south of the destination), an OD pair where the home is one or more time zones west of the destination (for example, home in San Francisco and traveling to Chicago) is considered a WEST TO EAST flight, and an OD pair where the home is in one or more time zones east of the destination is considered an EAST TO WEST flight. Figure 6a shows the overall distribution, where departure time sensitive people demonstrate a much stronger peaking behavior, favoring to *depart* either in morning times or evening times. Arrival time sensitive people appear much more attracted to *arriving* midday. Note that these times are only for the outbound flight. Figure 6b and 6c show the distributions categorized by business and leisure. Business travelers have a stronger preference to both depart and arrive in the morning relative to leisure travelers. Figure 6c and 6d show the distributions categorized by direction of travel. The patterns demonstrate distinctive patterns due to the time loss/gain as one moves across the country, especially for those who favor departure time. These figures demonstrate significant heterogeneity in travel time preferences, and a strong relationship to the characteristics of the trip.



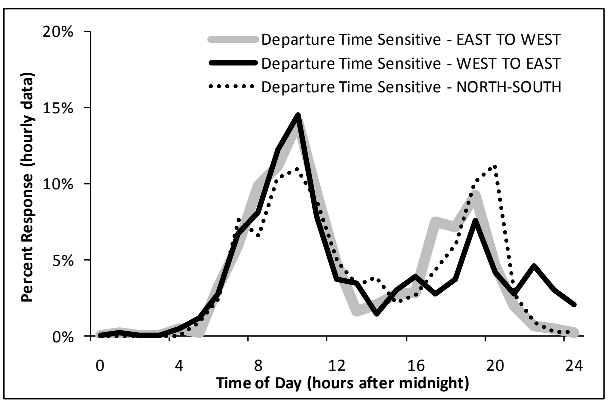

a) All respondents, categorized by departure time sensitive versus arrival time sensitive.



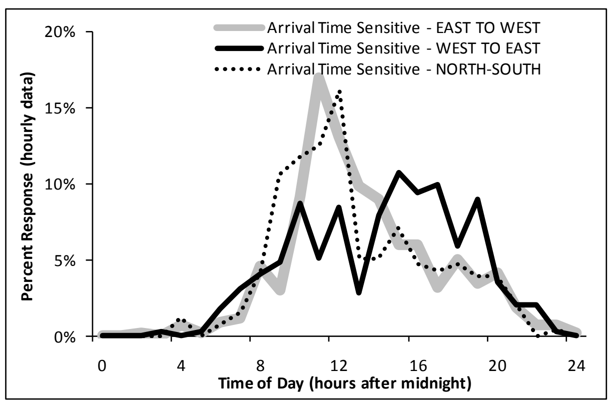

b) Departure time sensitive, categorized by purpose




c) Arrival time sensitive, cateogorized by purpose




d) Departure time sensitive, categorized by direction




e) Arrival time sensitive, categorized by direction

Figure 6: Heterogeneity in departure/arrival time of day histograms from reported preferences.

In this paper, we choose to focus on only a portion of the dataset to reduce the complexity of the problem, although our results capture significant heterogeneity even with the more focused sample. We focus the analysis on those individuals travelling for leisure, who stated departure time as more important in their flight choices. We choose this segment because it is the largest segment in the dataset (many more leisure travelers than business, and more departure time sensitive travelers than arrival time sensitive travelers). This decision reduces the sample size used in this paper to 1366 valid responses.

### *3.2. Model formulation*

Equation (7) as described above is estimated, which requires specification of the three components: the itinerary choice model $P(I|X,\tau)$, the time of day distribution $f(\tau|X)$, and the latent travel time measurement equation $k(d|X,\tau)$.

The explanatory variables of the itinerary choice model $P(I|X,\tau)$ included in the model are fare, travel time, stops, changes in airline, legroom, and schedule delay (the difference between desired and actual departure time). Further, the desired departure time is a latent variable. Each respondent was presented with three itineraries, and was asked to select both the best and the worst, resulting in a full ranking of the three options. Both choices are modeled using a logit for ranking data as follows:

$$\text{Probability}(\text{option } i \text{ is preferred to option } j \text{ is preferred to option } k)$$
$$= \Pr(i \succ j \succ k) = \Pr(i|\{i,j,k\};s_1)\Pr(j|\{j,k\};s_2) = \left(\frac{e^{s_1V_i}}{e^{s_1V_i}+e^{s_1V_j}+e^{s_1V_k}}\right)\left(\frac{e^{s_2V_j}}{e^{s_2V_j}+e^{s_2V_k}}\right) \quad (8)$$

where $V_i$ is the utility for alternative $i$ and $s_1$ is the scale of the first choice and $s_2$ the scale of the second choice. These scales are allowed to differ, since one can expect less precision in the second choice (Chapman and Staeelin, 1982; Hausman and Ruud, 1987). However, both scale parameters cannot be identified, so the estimable parameter is the relative scale of the second level of ranking with respect to the first level of ranking. (Note that while this decomposition assumes iid extreme value errors, this can be relaxed by adding error components to capture correlation.)

The time of day distribution $f(\tau|X)$ is specified using the mixture of normals as presented in equation (6). Each normal distribution (*k=1,...,K*) has a mean and standard deviation to be estimated $(\mu_k, \sigma_k)$. To determine the number of normal distributions $(K)$, we estimate the model with varying numbers of normal distributions, and then used the BIC statistic to determine the appropriate number (a standard approach for latent class models, see Magidson and Vermunt, 2004). We specify the mixing distribution $prob(k|X)$ as a logit model, and explore a number of trip- and person-specific variables as covariates.

The latent variable measurement equation $k(d|X,\tau)$, relates the stated preferred travel time to the latent travel time. The idea is to be able to capture possible measurement error in the self-reported values. We specify the relationship as:

$$d = \tau + \sigma_d \eta \; , \quad (9)$$

where $\eta$ is a normally distributed random variable with mean *0* and variance *1*, and so $d$ (the stated preferred travel time) is normally distributed with mean $\tau$ (the latent preferred travel time) and a variance ${\sigma_d}^2$ (the measurement error of $d$).

### *3.3. Estimation results*

The estimation results are reported in Table 1, displaying parameters for the itinerary choice model, the time of day distribution, and the measurement equation. The first 5 columns of results are for the jointly estimated itinerary choice and time of day distribution, which we will discuss first. The difference between the columns is the number of normal distributions in the mixing function ($K$ in equation 6). First, it is interesting to note that there is tremendous stability of the itinerary choice parameters, regardless of the specification of the time of day distribution. This is true even when it is poorly estimated with only a single normal distribution. All of these parameters have correct signs and are all highly significant. The trade-offs also are reasonable. Table 2 reports the willingness to pay exhibited by the model with 4 normal distributions. Because fare enters the utility with a log transform, the

## Table 1: Estimation results

| Variables | | | Simultaneous estimation | | | | | Sequential Estimation (4 normals) | |
|---|---|---|---|---|---|---|---|---|---|
| | | Model: | 1 normal | 2 normals | 3 normals | 4 normals | 5 normals | Conditional choice | Time of day curve |
| | | | Estimate (p-value) | Estimate (p-value) | Estimate (p-value) | Estimate (p-value) | Estimate (p-value) | Estimate (p-value) | Estimate (p-value) |
| Itinerary choice variables | Legroom (2 inch increments) | | 0.234 (0.00) | 0.233 (0.00) | 0.233 (0.00) | 0.233 (0.00) | 0.233 (0.00) | 0.233 (0.00) | |
| | Includes a stop (vs not) | | -0.862 (0.00) | -0.852 (0.00) | -0.854 (0.00) | -0.858 (0.00) | -0.856 (0.00) | -0.857 (0.00) | |
| | Includes an airline change (vs not) | | -0.141 (0.08) | -0.143 (0.08) | -0.142 (0.08) | -0.139 (0.09) | -0.140 (0.08) | -0.140 (0.09) | |
| | Travel time (hours) | | -0.589 (0.00) | -0.590 (0.00) | -0.591 (0.00) | -0.589 (0.00) | -0.590 (0.00) | -0.590 (0.00) | |
| | Ln(Fare in $) | | -9.224 (0.00) | -9.189 (0.00) | -9.211 (0.00) | -9.195 (0.00) | -9.198 (0.00) | -9.200 (0.00) | |
| | Late schedule delay (hours) | | -0.098 (0.00) | -0.090 (0.00) | -0.091 (0.00) | -0.089 (0.00) | -0.089 (0.00) | -0.089 (0.00) | |
| | Early schedule delay (hours) | | -0.159 (0.00) | -0.159 (0.00) | -0.158 (0.00) | -0.154 (0.00) | -0.154 (0.00) | -0.154 (0.00) | |
| | Scale parameter $s$ (2nd choice vs 1st choice) | | 0.544 (0.00) | 0.546 (0.00) | 0.545 (0.00) | 0.545 (0.00) | 0.545 (0.00) | 0.545 (0.00) | |
| Time of Day Variables: Means $\mu$ and standard deviations $\sigma$ of normal distributions | Normal ~7:00 | $\mu_1$ (hours past midnight) | | | | | 6.882 (0.00) | | |
| | | $\sigma_1$ (hours) | | | | | 0.722 (0.00) | | |
| | Normal ~9:00 | $\mu_2$ (hours past midnight) | | 9.290 (0.00) | 9.340 (0.00) | 9.026 (0.00) | 9.620 (0.00) | | 9.122 (0.00) |
| | | $\sigma_2$ (hours) | | 1.543 (0.00) | 1.740 (0.00) | 1.599 (0.00) | 1.193 (0.00) | | 1.661 (0.00) |
| | Normal ~noon | $\mu_3$ (hours past midnight) | 12.915 (0.00) | | | 12.027 (0.00) | 12.256 (0.00) | | 12.509 (0.00) |
| | | $\sigma_3$ (hours) | 4.810 (0.00) | | | 3.905 (0.00) | 4.038 (0.00) | | 4.150 (0.00) |
| | Normal ~18:00 | $\mu_4$ (hours past midnight) | | 18.472 (0.00) | 18.208 (0.00) | 18.599 (0.00) | 18.543 (0.00) | | 18.546 (0.00) |
| | | $\sigma_4$ (hours) | | 1.727 (0.00) | 1.427 (0.00) | 1.292 (0.00) | 1.551 (0.00) | | 1.417 (0.00) |
| | Normal ~22:00 | $\mu_5$ (hours past midnight) | | | 21.990 (0.00) | 21.838 (0.00) | 21.939 (0.00) | | 22.014 (0.00) |
| | | $\sigma_5$ (hours) | | | 0.012 (0.99) | 1.167 (0.00) | 1.007 (0.00) | | 1.008 (0.00) |
| Time of Day Variables: Mixing function for Normals (base is the normal distribution that is latest in the day) | Normal ~7:00 | Constant | | | | | 6.510 (0.01) | | |
| | | Traveling alone dummy | | | | | -0.816 (0.07) | | |
| | | Time change (hrs westbound) | | | | | 1.589 (0.91) | | |
| | | Time change (hrs eastbound) | | | | | -2.319 (0.01) | | |
| | Normal ~9:00 | Constant | | 0.590 (0.00) | 9.162 (0.01) | 6.473 (0.00) | 7.741 (0.00) | | 7.998 (0.01) |
| | | Traveling alone dummy | | -0.578 (0.00) | -0.396 (0.28) | -0.687 (0.04) | -0.615 (0.10) | | -0.583 (0.12) |
| | | Time change (hrs westbound) | | 0.064 (0.20) | 2.662 (0.95) | -0.677 (0.23) | 1.662 (0.90) | | 2.084 (0.93) |
| | | Time change (hrs eastbound) | | 0.072 (0.21) | -2.620 (0.03) | -1.812 (0.00) | -2.286 (0.01) | | -2.288 (0.03) |
| | Normal ~noon | Constant | | | | 5.742 (0.00) | 7.092 (0.00) | | 7.097 (0.03) |
| | | Traveling alone dummy | | | | 0.221 (0.57) | 0.299 (0.44) | | 0.324 (0.46) |
| | | Time change (hrs westbound) | | | | -1.003 (0.08) | 1.323 (0.92) | | 1.744 (0.94) |
| | | Time change (hrs eastbound) | | | | -1.952 (0.00) | -2.404 (0.00) | | -2.437 (0.02) |
| | Normal ~18:00 | Constant | | | 8.575 (0.02) | 5.988 (0.00) | 7.502 (0.00) | | 7.444 (0.02) |
| | | Traveling alone dummy | | | 0.203 (0.59) | 0.138 (0.70) | 0.144 (0.71) | | 0.200 (0.61) |
| | | Time change (hrs westbound) | | | 2.587 (0.96) | -0.819 (0.15) | 1.527 (0.91) | | 1.972 (0.94) |
| | | Time change (hrs eastbound) | | | -2.924 (0.02) | -2.264 (0.00) | -2.683 (0.00) | | -2.683 (0.01) |
| Measurement Eq. | Measurement error $\sigma_T$ | | 1.089 (0.12) | 1.246 (0.00) | 1.066 (0.00) | 0.043 (0.00) | 0.035 (0.00) | | |
| Goodness of fit | Log-likelihood | | -5806.8 | -5452.6 | -5398.7 | -5367.5 | -5355.0 | -5377.1 | |
| | Akaike Information Criterion (AIC) | | -11635.5 | -10939.2 | -10843.5 | -10793.0 | **-10780.0** | | |
| | Bayesian Information Criterion (BIC) | | -11692.9 | -11027.9 | -10963.5 | **-10944.4** | -10962.7 | | |

1366 individual respondents ; 2732 observed choices (2 per respondent)

willingness to pay is a function of the fare. Therefore, the willingness to pay values are reported in Table 2 as the percent of fare as well as in dollars for a $100 ticket and a $400 ticket (the approximate range of fares in the dataset). The numbers are in line with values found in the literature, also reported in Table 2. Note that all of the reported values (including those from our model) are in dollars of the year in which the corresponding surveys were collected (late 1990s to mid 2000s), which makes these values not directly comparable. Nonetheless, these provide a useful picture of the estimated range of these values. Also note that Garrow et al. (2006) was estimated with the same survey as used in this paper.[1]

[1] There are many differences in the Garrow et al. (2006) application that make direct comparison of the results both difficult to do and not particularly relevant to the focus of this paper (including differences in the overall objective, the choice indicator that is modelled, the set of respondents that is used, and the utility specification). Our focus is on the preferred departure time in which we assume the self-reported time is an indicator of a latent variable, whereas Garrow et al. (2006) assumed that the self-reported time was error free.

Table 2: Willingness to pay - 4 normals model relative to values in the literature

| Willingness to pay to: | % of fare | $100 ticket | $400 ticket | Values in the literature: Proussaloglou & Koppelman (1999) | Values in the literature: Adler et al. (2005) | Values in the literature: Garrow et al. (2006) |
|---|---|---|---|---|---|---|
| Reduce trip time by 1 hour | 6.40% | $6.40 | $25.61 | | $31.20 | $19.64 |
| Avoid departing 1 hour EARLIER than desired | 1.67% | $1.67 | $6.69 | $10.00 | $4.80 | |
| Avoid departing 1 hour LATER than desired | 0.97% | $0.97 | $3.89 | $10.00 | $4.80 | |
| Avoid departing 4 hours EARLIER than desired | 6.69% | $6.69 | $26.78 | | | |
| Avoid departing 4 hours LATER than desired | 3.89% | $3.89 | $15.56 | | | |
| Avoid a stop (BEYOND the additional time) | 9.33% | $9.33 | $37.33 | | $23.00 | |
| Avoid an airline change (BEYOND the stop and time) | 1.51% | $1.51 | $6.06 | | | |
| Gain 2 inch of legroom | 2.53% | $2.53 | $10.14 | | | |
| Gain 4 inches of legroom | 5.07% | $5.07 | $20.28 | | | |

Moving on to the parameters of the time of day distribution, the means and standard deviations of each normal distribution are all estimated. As the same peaks in the time of day distribution appear in the successive models, we organized the presentation of the estimation results according to these peaks which occur at approximately 7 AM, 9 AM, noon, 6 PM, and 10 PM. This is indicative of the peaking nature of the distribution. It is interesting to see which peaks are dominant as we successively increase the number of normal distributions: the 9 AM and 6 PM are most important, followed by the 10 PM, then the noon, and finally the 7 AM.

As for the mixing function, we explored many covariates and found the most significant are those related to party size (travelling alone versus travelling with at least one other) and the time change of the trip in terms of hours either westbound (1, 2, or 3) or eastbound (1, 2, or 3). The time change variable is highly correlated with the flight duration (in general, the more time zones that are crossed, the longer the flight), and it also captures the phenomenon of gaining or losing time as one crosses time zones. We find that this time change variable performs better (both statistically and in terms of interpreting the results) than using flight duration directly. The mixing function parameters allow the time of day distribution to vary based on these covariates, with a positive parameter leading to a stronger influence of that normal distribution in the overall distribution and vice versa. This is easiest to see with the aid of plots of the resulting time of day distributions. Using the parameter estimates from the 4 normals model, Figure 7 shows how the time of day distribution varies by trip characteristics. The top plot shows that west coast to east coast travel has a strong peak in the late evening, which is due to the red eye flights that depart around that time. However, the figure also shows that people travelling with someone else (rather than alone) have relatively less demand for the red eye. The middle plot shows extreme weights on the two main peaks (9 AM and 6 PM): two people travelling westbound across the country have a stronger preference for the morning peak, whereas one person travelling without a time change has a stronger preference for the 6 PM peak. The bottom plot shows the types of trips that act closer to the average distribution.

As for the measurement equation, note from the estimation results that the measurement error continues to decrease as the specification of the travel time distribution improves (i.e., as more normals are added). The parameter is statistically significant ($p\text{-value} < 0.005$) for all of the models, suggesting that there is measurement error in stated preferred travel times. As the units of the parameter are hours, the effective magnitude of the measurement error is quite small. However, we demonstrate below that it does impact the preferred time of day distribution.

Another aspect to discuss in Table 1 is the goodness of fit. Both the AIC and BIC are reported. The AIC suggests that the 5 normals model is superior to the others, whereas the BIC (which imposes a larger penalty on additional parameters) suggests that the 4 normals model is superior to the others. Figure 8 shows how the predicted time of day distributions from the different model specifications fit the reported time of day distribution. One normal distribution is an extremely poor fit. The two and three normal distributions under represent the desire for midday travel. The four and five normal distributions fit similarly, with the 5 normals providing a better fit to the early side of the morning peak. The evening peak is still not fitting the observed data well, even for the 5 normals model, and presumably a sixth normal would improve the fit there.

To investigate the impact of capturing measurement error in the model, the final two columns of Table 1 show

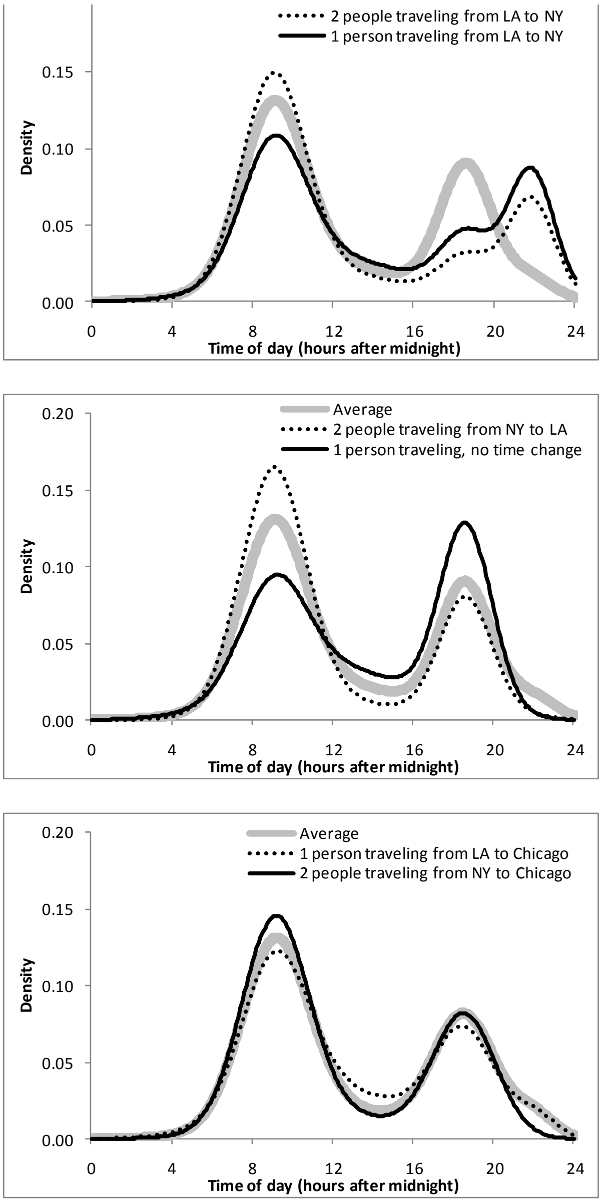



Figure 7: Estimated heterogeneity in time of day distributions (4 normals model)

estimation results for a model that is analogous to the 4 normals model but where the self-reported preferred departure times are considered to be error free. In this case, the itinerary choice model is estimated *given* the stated preferred departure time as an explanatory variable (the simultaneously estimated model is conditional on the *latent* preferred departure time). The mixture of normals for the time of day distribution reported in the last column is estimated directly from the stated preferred travel times. Comparing these results for the simultaneous 4 normals model, the fit decreases significantly in that over 9 log-likelihood points are lost with the removal of the one measurement error parameter. More importantly, comparisons of the parameters of the time of day distribution show differences. First, the time change variable for westbound travel is statistically significant when measurement error is captured. The negative signs on this variable indicate that there is more demand for the evening peak for East to West cross-country trips than would be indicated by the direct responses on desired travel time. This suggests that people may be influenced by the existing schedule when providing their desired travel time, because the last flights

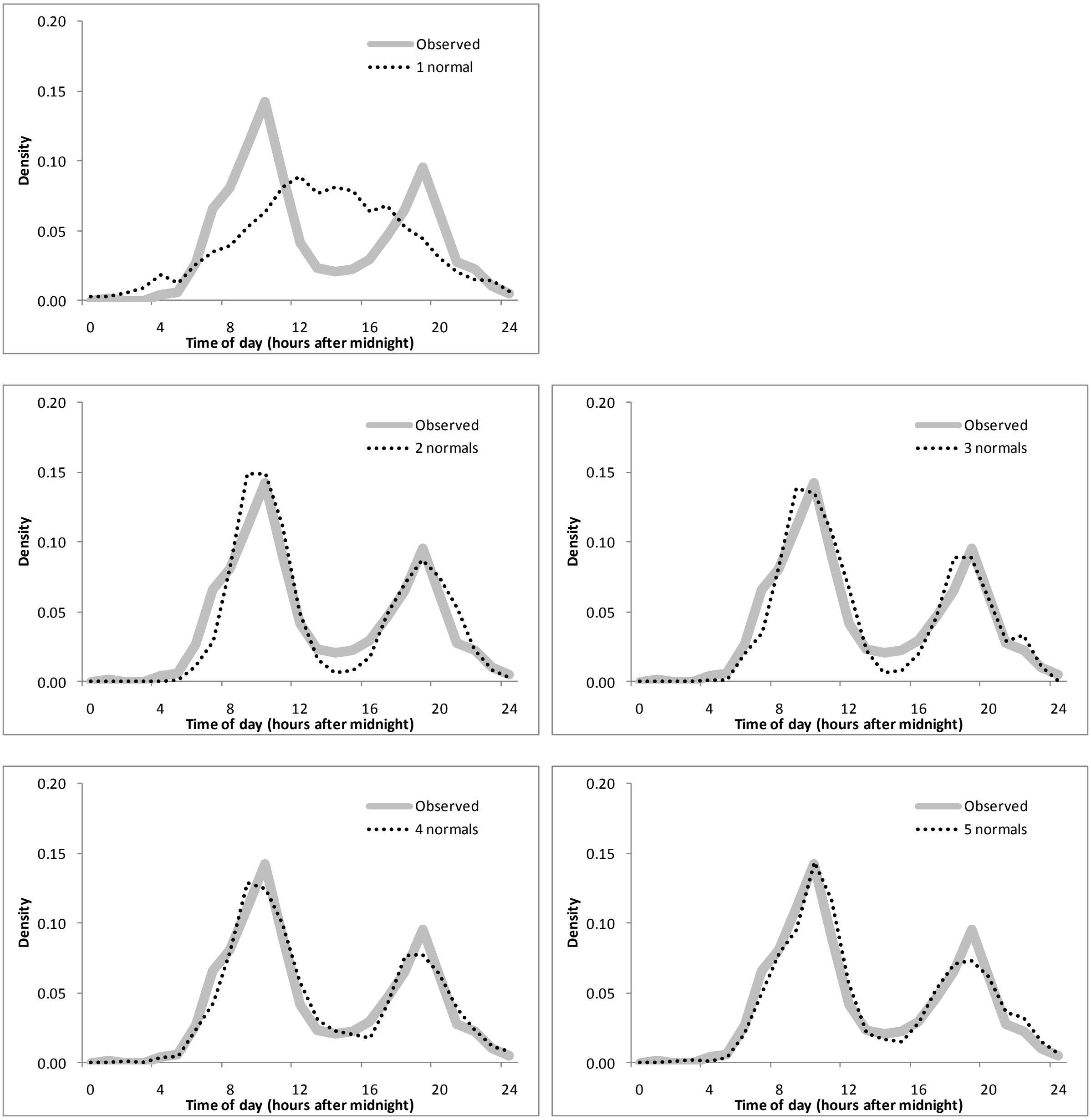
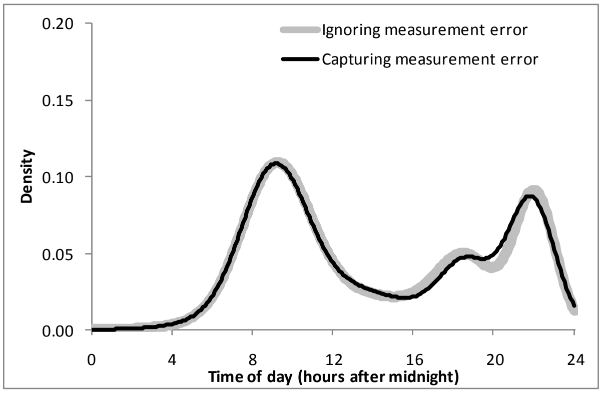
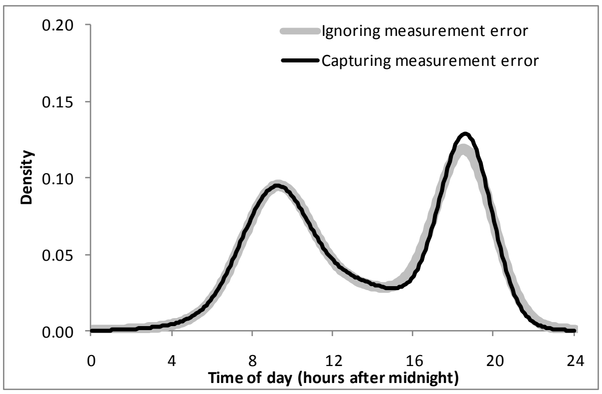



Figure 8: Predicted time of day distribution from different model specifications relative to observed (reported) time of day distribution.

out from the East to West coast are typically around 7:00 PM, whereas the demand captured for the 10 PM peak is greater than stated. Other differences are that the mean of the peaks shift, most notably the midday (~noon) peak that shifts almost 30 minutes earlier. The peaks for all but the late night (~22:00) peaks are sharper in that the standard deviations decrease, suggesting more dense demand for the prime times. While formal statistical tests are not applied, these shifts are large in relation to the estimated standard errors (i.e., shifts of greater than two standard errors). Figure 9 shows two instances in which these changes result in visible differences in the time of day demand. The top figure shows that for travel from the west to east coast the distribution in the late evening is smoother once measurement error is captured (an implication of the larger standard deviation for the late evening peak). The lower figure shows that for travel within the same time zone, there is a sharper peak at around 6 PM (an implication of the smaller standard deviation for the 6 PM peak).




a) Case of 2 people travelling from LA to NY




b) Case of 2 people travelling within a single time zone

Figure 9: Example differences in time of day distribution as a result of capturing measurement error.

## 4. Conclusion

This specification of the time of day curve and itinerary choice with schedule delay provides estimates of the willingness to pay for flights traveling at a particular time in the day, and therefore can provide estimates of the market share shifts caused by an adjustment in the schedule. There are relatively few time of day demand models in the air transportation literature, which is surprising given its importance on scheduling and fleet decisions. We have made use of stated preference data, which has the advantage that the results are not susceptible to the potential identification issues inherent in observed bookings data due to the interaction of supply and demand. The disadvantage is the hypothetical nature of the choice environment; however, this is lessened by the design of the survey we employed, which intercepted travelers while they were making the choice in the real world. For the time of day distribution we specify a finite mixture of normal distributions, which fits the data well and is behaviorally intuitive given the peaking nature of the distribution. We found the time of day distribution varied significantly based on the number of people in the party and the number (and direction) of time zones crossed. To address the issue that the self-reported preferred travel times may be influenced by the existing air schedule (and therefore have error), we treat travel time preference as a latent variable and jointly estimate the time of day and itinerary preference model. This represents a distinct difference from existing transport literature on time of day demand in which the preferred time of travel is either assumed to be unknown or assumed to be known with certainty. We found that there is statistically significant error in the self-reported travel times, and that this impacts the inferred time of day distributions in terms of the locations and dispersion of the peak travel periods. This may be particularly important in predicting the demand for new time of day markets opened up due to a change in technology.

**Acknowledgments**

We gratefully acknowledge Boeing, and especially Dr. Roger Parker (formerly Boeing's Chief of Technology and Senior Marketing Scientist with the Marketing and Business Strategy group), for motivating this work and funding and collaborating on the initial development of the ideas. Also we are grateful for funding from the Ministerio de Educación y Ciencia of Spain for partial financial support through a Jose Castillejo mobility grant.

**References**

Abou Zeid, M., Rossi, T.F., and Gardner B. (2006) Modeling time-of-day choice in context of tour- and activity-based models. *Transportation Research Record*, 1981, 42-49.

Abrahams, M. (1983) A service quality model of air travel demand: an empirical study. *Transportation Research Part A*, 17(5), 385-393.

Adler, T., Falzarano, C.S., and Spitz, G. (2005) Modeling service trade-offs in air itinerary choices. *Transportation Research Record*, 1915, 20-26.

Ashok, K., Dillon, W., and Yuan, S. (2002) Extending discrete choice models to incorporate attitudinal and other latent variables. *Journal of Marketing Research*, 39, 31-46.

Bates, J., Polak, J., Jones, P., and Cook, A. (2001) The valuation of reliability for personal travel. *Transportation Research Part E*, 37, 191-229.

Ben-Akiva, M., Walker, J., Bernardino, A.T., Gopinath, D.A., Morikawa, T., and Polydoropoulou, A. (2002) Integration of choice and latent variable models. *In Perpetual Motion: Travel Behavior Research Opportunities and Application Challenges* ed H. Mahmassani, pp. 275-302. Elsevier Science.

Ben-Akiva, M. and Abou Zeid, M. (2010) *Methodological Issues in Modeling Time-of-Travel Preferences*. Working paper, MIT.

Bhat, C.R. and Steed, J.L. (2002) A continuous-time model of departure time choice for urban shopping trips. *Transportation Research Part B*, 36(3), 207-224.

Boeing Company, The (1996) *The Decision Window Path Preference Model: Summary Discussion*. Marketing and Business Strategy, Boeing Commercial Airplanes.

Bolduc, D., Boucher, N., and Alvarez-Daziano, R. (2008) Hybrid choice modeling of new technologies for car choice in Canada. *Transportation Research Record*, 2082, 63-71.

Borjesson, M. (2008) Joint RP-SP data in a mixed logit analysis of trip timing decisions. *Transportation Research Part E*, 44, 1025-1038.

Bowman, J.L. and Ben-Akiva, M.E. (2001) Activity-based disaggregate travel demand model system with activity schedules. *Transportation Research Part A*, 35, 1-28.

Brueckner, J.K. and Zhang, Y. (2001) A model of scheduling gin airline networks. *Journal of Transport Economics and Policy*, 35(2), 195-222.

Chapman, R.G. and Staelin, R. (1982) Exploiting rank ordered choice set data within the stochastic utility model. *Journal of Marketing Research*, 19(3), 288-301.

de Jong, G., Daly, A., Pieters, M., Vellay, C., Bradley, M., and Hofman, F. (2003) A model for time of day and mode choice using error components logit. *Transportation Research Part E*, 39(3), 245-268.

Douglas, G.W. and Miller, J.C. (1974) *Economic Regulation of Domestic Air Transport: Theory and Policy.* Brookings institution, Washington D.C.

Eriksen, S.E. (1978) *Demand Models for US Domestic Air Passenger Markets*. Department of Aeronautics and Astronautics, MIT Report No FTL-R78-2.

Ettema, D.F. and Timmermans, H.J.P. (2003) Modeling departure time choice in the context of activity scheduling behavior. *Transportation Research Record*, 1831, 39-46.

Garrow, L.A., Jones, S.P., and Parker, R.A. (2006) How much airline customers are willing to pay: an analysis of price sensitivity in online distribution channels. *Journal of Revenue and Pricing Management*, 5(4), 271-290.

Hansen, M. (2002) Micro-level analysis of airport delay externalities using deterministic queuing models: a case study. *Journal of Air Transport Management*, 8, 73-87.

Hausman, J.A. and Ruud, P.A. (1987) Specifying and testing econometrics models for rank ordered data. *Journal of Econometrics*, 34(1-2), 83-104.

Hess, S., Polak, J.W., Daly, A., and Hyman, G. (2007) Flexible substitution patterns in models of mode and time of day choice: new evidence from the UK and The Netherlands. *Transportation*, 34(2), 213-238.

Koppelman, F.S., Coldren, G.M., and Parker, R.A. (2008) Schedule delay impacts on air-travel itinerary demand. *Transportation Research Part B*, 42, 263-273.

Lemp, J.D., Kockelman, K.M., and Damien, P. (2010) The continuous cross-nested logit model: Formulation and application for departure time choice. *Transportation Research Part B*, 44, 646-661.

Li, Z., Hensher, D.A., and Rose, J.M. (2010) Willingness to pay for travel time reliability in passenger transport: A review and some new empirical evidence. *Transportation Research Part E*, 46, 384-403.

Mahmassani, H.S., and Chang, G.L. (1986) Experiments with departure time choice dynamics of urban commuters. *Transportation Research Part B*, 20(4), 297-320.

Magidson, J., and Vermunt, J.K. (2004) Latent class models. *The Sage Handbook of Quantitative Methodology for the Social Sciences* ed D. Kaplan, pp. 175-198. Thousand Oaks: Sage Publications.

Mehndiratta, S.R. and Hansen, M. (1997) Analysis of discrete choice data with repeated observations: comparison of three techniques in intercity travel case. *Transportation Research Record*, 1607, 69-73.

Noland, R.B. and Small, K.A. (1995) Travel-time uncertainty, departure time choice, and the cost of morning commutes. *Transportation Research Record*, 1493, 150-158.

Parker, R.A. and Walker, J.L. (2006) Estimating utility of time-of-day demand for airline schedules using mixed logit model. *Transportation Research Board 85th Annual Meeting Compendium of Papers*, Washington, D.C.

Proussaloglou, K. and Koppelman, F.S. (1999) The choice of air carrier, flight, and fare class. *Journal of Air Transport Management*, 5, 193-201.

Ryerson, M.S. and Hansen, M. (2010) The potential of turboprops for reducing aviation fuel consumption. *Transportation Research Part D*, 15, 305-314.

Sathaye, N., Harley, R., and Madanat, S. (2009) *Unintended Environmental Impacts of Nighttime Freight Logistics Activities.* UC Berkeley Center for Future Urban Transport: A Volvo Center of Excellence, Institute of Transportation Studies, UC Berkeley, UCB-ITS-VWP-2009-11.

Small, K.A. (1987) A discrete choice model for ordered alternatives. *Econometrica*, 55(2), 409-424.

Small, K.A. (1982) The scheduling of consumer activities: work trips. *American Economic Review*, 72, 467-479.

Swan, W.E. (1979) *A Systems Analysis of Scheduled Air Transportation Networks*. Department of Aeronautics and Astronautics, MIT Report No. FTL-R79.

Temme, D., Paulssen, M., Dannewald, T. (2008) Incorporating Latent Variables into Discrete Choice Models - A Simultaneous Estimation Approach Using SEM Software. *BuR Business Research Journal*, 1(2), 220-237.

Tseng, Y. and Verhoef, E.T. (2008) Value of time by time of day: a stated-preference study. *Transportation Research Part B*, 42, 607-618.

Vickrey, W. (1969) Congestion theory and transport investment. *American Economic Review*, 59(2), 251-260.

Walker, J. and Ben-Akiva, M. (2002) Generalized random utility model. *Mathematical Social Sciences*, 43(3), 303-343.

Walker, J.L., Li, J., Srinivasan, S., and Bolduc, D. (2010) Travel demand models in the developing world: correcting for measurement errors. *Transportation Letters*, 2(4), 231-243.

Yáñez, M.F., Raveau, S., Ortúzar, J. de D. (2010) Inclusion of latent variables in Mixed Logit models: Modelling and forecasting. *Transportation Research Part A*, 44, 744-753.