Title

Uncovering the Distribution of Motorists' Preferences for Travel Time and Reliability: Implications for Road Pricing

Authors

Small, Kenneth A.
Winston, Clifford
Yan, Jia

Publication Date

2002-08-01

UCI-ITS-WP-02--4

Uncovering the Distribution of Motorists' Preferences for Travel Time and Reliability: Implications for Road Pricing

UCI-ITS-WP-02-4

Kenneth A. Small [1]
Clifford Winston [2]
Jia Yan [3]

[1] Department of Economics and Institute of Transportation Studies
University of California, Irvine; Irvine, California 92697-5100,U.S.A.
ksmall@uci.edu

[2] Brookings Institution
1775 Massachusetts Ave., N.W., Washington, D.C. 20036, U.S.A.
cwinston@brook.edu

[3] Department of Economics and Institute of Transportation Studies
University of California, Irvine

August 2002

Institute of Transportation Studies
University of California, Irvine
Irvine, CA 92697-3600, U.S.A.
http://www.its.uci.edu

UNCOVERING THE DISTRIBUTION OF MOTORISTS' PREFERENCES FOR TRAVEL TIME AND RELIABILITY: IMPLICATIONS FOR ROAD PRICING

by Kenneth A. Small, Clifford Winston, and Jia Yan

August 1, 2002

Abstract

Recent econometric advances have made it possible to empirically identify the varied nature of consumers' preferences. We apply these advances to study commuters' preferences for speedy and reliable highway travel with the objective of exploring the efficiency and distributional effects of road pricing that accounts for users' heterogeneity. Our analysis combines revealed and stated commuter choices of whether to pay a toll for congestion-free express travel or to travel free on regular congested roads. We find that highway users exhibit substantial heterogeneity in their values of travel time and reliability. Moreover, we show that road pricing policies that cater to varying preferences can substantially increase efficiency while maintaining the political feasibility exhibited by current experiments. By recognizing heterogeneity, policymakers may break the current impasse in efforts to relieve highway congestion.

Keywords: mixed logit, stated preference, congestion pricing, product differentiation, heterogeneous consumers

Kenneth A. Small*
Department of Economics
University of California
Irvine, CA 92697-5100
ksmall@uci.edu

Clifford Winston
Brookings Institution
1775 Mass. Ave., N.W.
Washington, DC 20036
cwinston@brook.edu

Jia Yan
Department of Economics
University of California
Irvine, CA 92697-5100
jiay@uci.edu

*Corresponding author

Acknowledgment

The authors are grateful to the Brookings Center on Urban and Metropolitan Policy and the University of California Transportation Center for financial support. We thank Edward Sullivan for access to data collected by California Polytechnic State University at San Luis Obispo, with financial support from the California Department of Transportation and the US Federal Highway Administration's Value Pricing Demonstration Program. We also thank David Brownstone, Jerry Hausman, Charles Lave, Steven Morrison, Randy Pozdena, and participants in the UCI Econometrics Seminar for comments.

1. Introduction

On a given weekday, roughly two hundred million people in the United States use a vehicle for work or personal trips. Yet highway authorities have ignored the variety in motorists' preferences for speedy and reliable travel, instead offering uniform service financed by the gasoline tax. The unfortunate result has been greater congestion on urban and intercity highways during ever-expanding peak periods. The standing recommendation of economists—that road pricing could spur motorists to make better use of highway capacity by spreading their travel throughout the day—has gone unheeded by policymakers presumably because it would have a negative impact on road users.[1]

But what if policymakers recognized that motorists are not homogeneous—that their attitudes toward congestion range from loathing to indifference—and offered these motorists differentiated prices that catered to their preferences? Indeed, experience with deregulation of transportation, telecommunications, energy, and other industries has taught us that firms have increased capacity utilization, developed niche markets, and benefited consumers by offering a variety of prices and services that respond to consumer desires (Winston (1998)). Could highway pricing that recognized the heterogeneity in motorists' preferences increase efficiency and curtail its negative welfare effect on motorists?

Recent pricing experiments in the Los Angeles, San Diego, and Houston areas give motorists the option to travel free on regular roads or to pay a time-varying price for congestion-free express travel on a limited part of their journey. These experiments, often called "value pricing," provide rare opportunities to study motorists' preferences in automobile-dominated environments

[1] For homogeneous users, the full price (equal to the combined effect of the toll and travel time savings) must be raised in order to reduce travel and thereby reduce congestion.

where real money is at stake.[2] At the same time, econometric advances are making it possible to identify the varied nature of consumer preferences. These advances include random-parameters models of discrete choice that account for unobserved heterogeneity, error-components models that control for the correlation among repeated choices by a given individual, methodologies that combine the advantages of data generated by consumers' actual and hypothetical choices, and non-parametric techniques that yield plausible characterizations of difficult-to-measure variables such as the reliability of travel time.

This paper measures the preferences of automobile commuters by applying these methodological advances to newly collected data concerning route choices in the Los Angeles-area pricing experiment. Based on their choice of whether to pay a toll and how high a toll to pay to use express lanes, we find that those commuters vary substantially in how they value travel time and travel-time reliability. We then study how the efficiency and distributional effects of road pricing are affected when commuters' heterogeneity is taken into account. Compared with a uniform price, we find that differentiated road prices can significantly reduce the losses in consumer surplus and the distributional disparities between groups of motorists, while still producing sizable efficiency gains. Such prices enhance the political viability of road pricing because only a modest portion of the toll revenues would be necessary to compensate road users.

2. A Brief Methodological Overview

At first blush, our empirical question—how do motorists value travel time and reliability?—is hardly original or one that calls for a sophisticated methodology. A conventional

[2] The term "value pricing" originated as a marketing tool for the first of these experiments. Interestingly the term was found so efficacious that the U.S. Congress substituted it for "congestion pricing" in the 1998 reauthorization of what was then called the "Congestion Pricing Demonstration Program." See *Federal Register*, 63 (192), October 5, 1998, pp. 53487-91.

approach would estimate a model of a commuter's choice of whether to pay a toll to use an uncongested express lane or use a free but congested lane as a function of the toll, the travel times on the lanes, the reliability of travel time on the lanes, and the driver's socioeconomic characteristics. This model, however, is likely to be flawed unless one accounts for commuters' unobserved heterogeneity, the strong negative correlation between road charges and travel time, and the difficulty of obtaining an accurate measure of reliability. Recent advances in econometrics enable us to address these issues.

Unobserved Heterogeneity

Preference heterogeneity may be explained by observable characteristics and unobserved influences. The latter can be captured using models with random coefficients. We will use the mixed logit specification, which extends the random-utility model that underlies multinomial logit (Brownstone and Train (1999), McFadden and Train (2000)).[3] Our version of mixed logit introduces three stochastic utility components: one that has the double-exponential distribution standard for logit models, a second that represents random variation in coefficients (i.e. unobserved heterogeneity in tastes), and a third that represents a panel-type error structure arising from repeated choices by a given commuter.. Choice probabilities are estimated using Monte-Carlo simulation to integrate the computationally difficult parts of the error distribution. Conditional on the Monte-Carlo draws, the probabilities take the logit form.

Revealed and Stated Preferences

Most previous research attempting to determine the value of urban travel time has analyzed revealed preference (RP) data based on the choice between travel by car and public transit. This

[3] Based on data generated by hypothetical questions, mixed logit has been used to estimate the value of time in analyses of long-distance commuting (Calfee, Winston, and Stempski (2001)), urban trucking (Kawamura (2000)), and residential and workplace location (Rouwendahl and Meijer (2001)).

research has found that travelers' value of time varies with trip purpose, income, trip distance, and other observed variables (Small (1992), Wardman (2000)). Some recent studies have analyzed stated preferences (SP) that are elicited from individuals who are faced with hypothetical commuting situations (Calfee and Winston (1998), Hensher (2001)).

Both RP and SP data have drawbacks. Use of RP data is often hindered by strong correlations among travel cost, time, and reliability, and by the difficulty of obtaining accurate values of these variables for all the alternatives faced by each individual. SP data cannot overcome the lingering doubt that the behavior exhibited in hypothetical situations may not apply to actual choices. Methodologies have been developed to combine both types of data, thereby taking advantage of the strengths of each (Ben-Akiva and Morikawa (1990), Hensher (1994), and Bhat and Castelar (2002)).

The key insight of these methodologies is that some parameters or parameter combinations are likely to be identical in the choice functions generating RP and SP choices, whereas others are likely to be different. For example, the variance of the error term describing the choice process is likely to differ across data types, as is the ratio of the coefficients of travel time and cost. The latter difference arises because people commonly overstate the time delays they actually incur, and thus respond more to a given actual time saving than to a hypothetical time saving of the same amount.[4] By combining data sets, one can greatly improve the precision in estimating common coefficients, while allowing important differences in other coefficients to emerge.

Reliability

Travel-time reliability is a potentially critical influence on any mode or route choice, but it can be difficult to measure (Bates, Polak, Jones, and Cook (2001)). Based on data from actual

[4] Sullivan et al. (2000, p. xxiii) provide evidence of this from questions asked of travelers affected by two California road-pricing experiments, including the one used in this study.

driving conditions, we use non-parametric methods to develop plausible characterizations of reliability to include in our RP estimations. We also specify the reliability of trips in our hypothetical (SP) questions.

3. Empirical Setting

The commuter route of interest is California State Route 91 (SR91) in the greater Los Angeles region. It connects rapidly growing residential areas in Riverside and San Bernardino Counties—the so-called Inland Empire—to job centers in Orange and Los Angeles Counties to the west. A ten-mile portion of the route in eastern Orange County includes four regular freeway lanes (91F) and two express lanes (91X) in each direction. Motorists who wish to use the express lanes must set up an account and carry an electronic transponder to pay a toll that varies hourly according to a preset schedule. Tolls on westbound traffic during the morning commute hours covered in this study ranged from $1.65 (at 4-5 a.m.) to $3.30 (at 7-8 a.m. Monday-Thursday).[5] Carpools of three or more received a 50 percent discount.[6] Unlike the regular lanes, the express lanes have no entrances or exits between their end points.

Samples

To enrich our analysis, we draw on two samples of people traveling on this corridor. The surveys generating the data contain sufficiently similar questions and were conducted at nearly the same times, so it has proven feasible to combine them. One is a telephone RP survey composed of SR91 commuters obtained by random-digit dialing and observed license plates on the SR91

---

[5] Morning tolls were slightly lower on Fridays. We accounted for the slight rise in tolls that occurred during our surveys. Tolls were subsequently raised after our analysis was completed to a maximum westbound toll of $3.60, while eastbound tolls are higher, reaching $4.75 at 5-6 p.m. Monday-Tuesday and 4-6 p.m. Wednesday-Friday.

[6] For this reason the express lanes are known as High-Occupancy/Toll (HOT) lanes. Another discount, of $0.75 per trip, was available to people who paid $15/month to be in an "Express Club"; we do not account for the net benefits of this promotion.

corridor. The survey was conducted by researchers at California Polytechnic State University at San Luis Obispo (Cal Poly), under the leadership of Edward Sullivan and with our participation.[7] The Cal Poly data, collected in November 1999, asked participants about their most recent trip on a Monday through Thursday during the morning peak (4-10a.m.) and included questions concerning lane choice (91X or 91F), trip distance, time of commute, vehicle occupancy, mode (drive alone or carpool), and whether they had a flexible work-arrival time. They also provided various personal and household characteristics. The sample we use consists of 438 respondents.

The second sample is a two-stage mail survey collected by us through the Brookings Institution (Brookings), including both RP and SP elements. For the Brookings sample, a market research firm, Allison-Fisher, Inc., mailed a survey custom-designed to our specifications to SR91 commuters who were members of two nationwide household panels, National Family Opinion and Market Facts. A screener was first used to identify motorists who made work trips covering the entire 10-mile segment and thus had the option of using either roadway (91F or 91X). Survey respondents reported on their daily commute for an entire five-day workweek, providing information on the same items as mentioned above. The same people were then asked to complete an SP survey containing eight hypothetical commuting scenarios describing the essential characteristics of express and regular lanes. For each scenario, they were given hypothetical tolls, travel times, and probabilities of delay on the two routes, and asked which they would choose. The values presented in the scenarios were roughly aligned with a respondent's normal commute. An illustrative scenario is shown in Appendix A.

[7] For more details about the Cal Poly sample see Sullivan et al. (2000). The sample also included people who traveled on just a part of route 91F and then exited onto a new toll expressway going to Irvine and southern Orange County; we have not included these people in our analysis.

Due to overestimates of how many respondents would actually face a choice between 91F or 91X, we had to survey three waves of potential respondents—in December 1999, July 2000, and September 2000—to assemble an adequate sample. The final Brookings sample consists of 110 respondents: 84 people providing 377 daily observations on actual behavior (RP), and 81 people providing 633 separate observations on hypothetical behavior (SP), with 55 people answering both surveys.

Summary Statistics

Table 1 summarizes responses from both data sets. Values for the Brookings data are broadly consistent with population summary statistics, indicating that we have a representative sample.[8] The median household income (assigning midpoints to the income intervals) is $46,250. We estimate the average wage rate to be about $23 per hour.[9] The Brookings sample contains information for multiple days and indicates that inertia is a powerful force in route choice behavior because 87 percent of the RP respondents made the same choice every day during the survey week. In fact, about half of the Brookings RP respondents do not have a transponder and thus have committed to not choosing the express lanes on any of our survey days.

---

[8] The distributions of the RP sample's commuting times and route share are close to the ones in 1998 survey data collected by University of California at Irvine (Lam and Small (2001)) and 1999 survey data collected by California Polytechnic State University at San Luis Obispo (Sullivan et al. (2001)). The socioeconomic data are consistent with Census information, and diverge where appropriate. For example, our median income (approximately $46,250) is higher than the average income in the two counties where our respondents lived ($36,189 in Riverside County and $39,729 in San Bernardino County in 1995, as estimated by the Population Research Unit of the California Department of Finance). But this should be expected because our sample only includes people who are employed and commute to work by car. The median number of people per household (which can be expected to be stable across time) is 2.81 and 3.47 in our RP and SP subsamples respectively; these are not far from the 1990 Census figures of 2.85 for Riverside County and 3.15 for San Bernardino County.

[9] Data from the US Bureau of Labor Statistics (BLS) for the year 2000 record the mean hourly wage rate by occupation for residents of Riverside and San Bernardino Counties. We combine the BLS occupational categories into six groups that match our survey question about occupation, then assign to each person in our sample the average BLS wage rate for the appropriate occupational group. We then add 10 percent to reflect the higher wages likely to be attracting these people to jobs that are relatively far away.

The Cal Poly sample's route shares, commuting patterns, respondents' age and sex, and so on are closely aligned with the Brookings sample. Respondents in the Cal Poly sample do have higher household incomes and shorter trip distances than the Brookings respondents; apparently the Brookings sample drew from a wider geographical area including people who reside in lower-priced housing.

Construction of Independent Variables

Obtaining accurate measures of travel conditions facing survey respondents is a challenging part of any travel demand analysis using RP data. Our case is no exception, and is made more difficult by the desire to include travel-time reliability. Our strategy is to use actual field measurements of travel times on SR91 taken at different times during the six-hour morning period covered by our data. Measurements were taken on eleven days, ten of which coincided with the days covered by the second and third waves of the Brookings survey; the eleventh day was two months prior to the first wave of the Brookings survey and one month prior to the Cal Poly survey.

We posit that for any given time of day, observed travel times are random draws from a distribution that travelers know from experience. By asserting that motorists care about trip time and reliability, we maintain that they consider both the central tendency and the dispersion of that distribution.[10] Plausible measures of central tendency include the mean and the median; we find the median fits slightly better (in terms of log-likelihood achieved by the model). Measures of dispersion include the standard deviation and the inter-quartile difference; however, given that motorists—especially commuters—are concerned with occasional significant delays, they are

[10] It seems reasonable for several reasons to assume that motorists' lane choices are based mainly on their knowledge of the distribution of travel times across days, not on the travel time encountered that day. Previous survey results described by Parkany (1999) suggest that whatever information travelers on this road have about conditions on a given day is mostly acquired en route through radio reports, and thus has limited value to them because it cannot affect their departure time. In addition, there is no sign displaying traffic information, and our field observations suggested that the amount of congestion encountered prior to the entrance to the express lanes was not a good predictor of the travel delays along the full 10-mile segment.

likely to pay particular attention to the upper tail of the distribution of travel times. We therefore investigate the upper percentiles of our travel time distributions.

We use non-parametric smoothing techniques to estimate the distribution of travel-time savings from taking the express lanes, by time of day.[11] Details are presented in appendix B, and some results are shown in Figures 1 and 2. Figure 1 shows the raw field observations of travel-time savings. The non-parametric estimates of mean, median, and 80th percentile are superimposed. Median time savings reach a peak of 5.6 minutes around 7:15 a.m.

Figure 2 shows the same raw observations after subtracting our non-parametric estimate of median time savings by time of day. An interesting pattern emerges. Up to 7:30 a.m., the scatter of points is reasonably symmetric around zero with the exception of three data points. But after that time the scatter becomes highly asymmetric, with dispersion in the positive range (the upper half of the figure) continuing to increase until after 8:00 a.m. while dispersion in the negative range decreases. This feature is reflected in the three measures of dispersion, or unreliability, that are also shown in the figure: the standard deviation and the 80th-50th and 90th-50th percentile differences. The standard deviation peaks at roughly 7:45 a.m., the other two between 8:15 and 9:30. The reason for these differences is that traffic in the later part of the peak is affected by incidents occurring either then or earlier. This mostly affects the upper tails of the distribution of travel-time savings and so is most apparent in the percentile differences. The standard deviation, by contrast, is higher early in the rush hour because of days with little congestion—showing up as negative points in Figure 2. Such dispersion is probably less relevant to travelers than dispersion in the upper tails, leading us to prefer the percentile

[11] We never observed any congestion on the express lanes. Thus, to simplify the problem, we assume that travel time on them is equal to the travel time we observed on the free lanes at 4:00 a.m., when there was no congestion: namely, 8 minutes, corresponding to a speed of 75 miles per hour.

differences as reliability measures. These measures are also considerably less correlated with median travel time than is the standard deviation. In our estimations, we obtained the best statistical fits using the 80^{th}-50^{th} percentile difference.[12]

The express-lane toll for a given trip is from the published toll for the relevant time of day, discounted by 50 percent if the trip were in a carpool of three or more.[13] Other potentially important variables include trip distance, annual per capita household income, and a dummy variable that indicates whether the commuter had a flexible arrival time.[14] We also explored a number of others including age, sex, household size, and size of workplace, a few of these variables had little explanatory power and did not influence the other coefficients.

Most variables in the SP model correspond exactly to variables in the RP model. An exception is the measure of unreliability, because we did not think survey respondents would understand statements about percentiles of a probability distribution. Instead, we specified in our SP scenarios the probability of being delayed 10 minutes or more.[15] In addition, SP respondents indicated whether they were answering the questions as solo drivers or as part of a carpool of a specified size, enabling us to determine vehicle occupancy for the SP choices.

---

[12] In our RP and joint RP/SP models, the 90^{th}-50^{th} percentile difference fit almost as well as the 80^{th}-50^{th} difference (in terms of log-likelihood) and resulted in similar coefficient estimates. The 75^{th}-50^{th} percentile difference, an additional measure, and the standard deviation fit noticeably less well and gave statistically insignificant results for the reliability measure.

[13] Specifically, the toll is for the time of day that the commuter reported passing the sign stating the applicable toll. Respondents who did not provide information on vehicle occupancy are assumed not to have carpooled; to guard against systematic bias from this, we specified a dummy variable identifying these respondents, but it had no explanatory power so it is not included in the models reported here. Due to the uneven quality of answers about carpooling, and lack of knowledge of ages or characteristics of passengers, we did not attempt to prorate the toll among vehicle occupants.

[14] The question was: "Could you arrive late at work on that day without it having an impact on your job?"

[15] The probability was always stated for the trip as a whole. It was given as 0.05 for all trips using 91X, and either 0.05, 0.1, or 0.2 for trips using 91F. The actual statement is: "Frequency of unexpected delays of 10 minutes or more: 1 day in X" where X=20, 10, or 5.

4. Econometric Framework

We assume that a motorist i, facing an actual or hypothetical choice between commuting lanes at time t, chooses the option that maximizes a random utility function. Define the choice variable as y_{it}=1 if the express lanes are chosen and 0 otherwise. Let

$$U_{it} \equiv \theta_i + \beta_i X_{it} + \varepsilon_{it} \tag{1}$$

be the utility difference, so that the express lanes are chosen whenever U_{it}>0. Variables included in X_{it} measure the toll difference C_{it}, travel-time difference T_{it}, and (un)reliability difference R_{it} between the two alternatives. The values of travel time and reliability are defined as:

$$VOT_i = \frac{\partial U_{it} / \partial T_{it}}{\partial U_{it} / \partial C_{it}}; \qquad VOR_i = \frac{\partial U_{it} / \partial R_{it}}{\partial U_{it} / \partial C_{it}}. \tag{2}$$

As the notation indicates, the models are specified so that *VOT* and *VOR* depend on the individual traveler i but not on the time t that a choice is made. However, they may depend on whether a given individual is answering an RP or an SP question.

User Heterogeneity and Panel-Type Data

We can specify the estimable parameters, θ and β, to capture observed and unobserved heterogeneity. Namely,

$$\theta_i = \bar{\theta} + \phi W_i + \xi_i \tag{3}$$

$$\beta_i = \bar{\beta} + \gamma Z_i + \varsigma_i. \tag{4}$$

Observed heterogeneity is captured by variables W_i and Z_i, while unobserved heterogeneity is captured by the random terms ξ_i and ς_i. The scalar ξ_i indicates an individual's unobserved alternative-specific preferences, whereas the vector ς_i represents an individual's unobserved preferences regarding travel characteristics. (These two sources of randomness are called *preference heterogeneity* and *response heterogeneity*, respectively, by Bhat and Castelar (2002)). Thus the derivatives in (2) depend on variables Z and also contain components of ς_i, giving *VOT* and *VOR* both observable and stochastic variability.

Equation (3) also accommodates the panel-type data structure arising from individuals' repeated observations. We assume ξ_i and all the components of ς_i are distributed normally and independently of each other and of ε_{it}:

$$\xi_i \sim N\left(0,\sigma_\xi^2\right), \quad \varsigma_i \sim N(\mathbf{0},\Omega) \tag{5}$$

with Ω diagonal.

Combining Different Sources of Data

We denote our two data sets by superscripts B (for Brookings) and C (for Cal Poly). The Brookings data contain both RP and SP responses (further denoted with superscripts R and S) and contain multiple responses from the same individual. The Cal Poly data are RP only and purely cross-sectional. As described later, in the RP portion of the Brookings data we form one choice variable from a motorist's multiple-day observations.

A number of possible sources of correlation must be accounted for to combine the data sets without introducing bias. We account for the correlation between the RP and SP error terms from the same individual in the Brookings data, for those 55 individuals who answered both the RP and the SP survey. To accomplish this we split the corresponding error terms in (1) into two independent parts:

$$\varepsilon_i^{BR} = \nu_i^{BR} + \eta_i^{BR} \tag{6}$$

$$\varepsilon_{it}^{BS} = \rho\nu_i^{BR} + \eta_{it}^{BS}, \tag{7}$$

where ρ captures the correlation between the error terms for a given individual and the random terms η_i^{BR} and η_{it}^{BS} are assumed independent of each other.

Denoting the error term ε_i in the Cal Poly data by η_i^C, the full joint model is then represented by the following utility differences between express and regular lanes:

$$U_i^{BR} \equiv \theta_i^{BR} + \beta_i^{BR} X_i^{BR} + \nu_i^{BR} + \eta_i^{BR} \tag{8}$$

$$U_{it}^{BS} \equiv \theta_i^{BS} + \beta_i^{BS} X_{it}^{BS} + \rho\nu_i^{BR} + \eta_{it}^{BS} \tag{9}$$

$$U_i^{C} \equiv \theta_i^{C} + \beta_i^{C} X_i^{C} + \eta_i^{C}. \tag{10}$$

In these equations, the θ and β parameters are specified to capture observed and unobserved heterogeneity as in (3) and (4), except that error terms ξ_i^{BR} and ξ_i^C arising from (3) are omitted because, with only one observation per individual, they are redundant given the presence of η_i^{BR} and η_i^C. Index i in (8)-(10) runs through all individuals in the data sets.

We assume that $\nu_i^{BR} \sim N(0,1)$. We also assume that η_i^{BR}, η_{it}^{BS}, and η_i^C are independently logistic distributed, which yields the familiar logit formula for the choice probability conditional on parameters and on ν_i^{BR}.[16] Our treatment of heterogeneity is therefore an example of a mixed logit model.

As is usual in combining RP and SP data sets, we allow the variances of η_i^{BR} and η_{it}^{BS} to differ, indicating that there may be different sources for random preferences over revealed and stated choices. We also let η_i^C have its own variance because the data sets have different questionnaire formats: most importantly, the Brookings choice variable is based on choices over

[16] Equivalently, each η_i is the difference between two independent random variates each having the extreme-value (double-exponential) distribution.

several (often five) days, so we expect η_i^{BR} to have a smaller variance than η_i^{C}. All this is accomplished by normalizing the variance of η_i^{BR} (to $\pi^2/3$ as in the binary logit model) and estimating the ratios

$$\mu^{BS} \equiv \sigma^{BR}/\sigma^{BS} \tag{11}$$

$$\mu^{C} \equiv \sigma^{BR}/\sigma^{C} \tag{12}$$

where each σ is the standard deviation of the corresponding η_i.

Our specification allows considerable generality in how choices are determined in the three samples (BR, BS, C) relative to each other. Of course, a model that combines the samples can improve statistical efficiency only by imposing some constraints. We therefore assume that some coefficients are identical in two or more of the choice processes. This enables us to use the RP responses to eliminate some sources of SP survey bias, while using the SP responses to help identify some key heterogeneity parameters, whose effects would otherwise be obscured in the RP-only data by multicollinearity.

Estimation

The parameters of the model are estimated by Simulated Maximum Likelihood Estimation (SMLE), as outlined for example by Brownstone and Train (1999). Let F represent all the fixed parameters (i.e., those common to all individuals in a given sample), and let Θ_i represent the random components (other than η_i); Θ_i has a joint distribution with parameter vector ψ, represented by density function $f(\Theta_i \mid \psi)$. The likelihood function of our model is thus specified as:

$$L(F,\psi) = \prod_{it} \int_{\Theta_i} P(y_{it} \mid F, \Theta_i) f(\Theta_i \mid \psi) d\Theta_i \quad , \tag{13}$$

where $P(y_{it} \mid F, \Theta_i)$, the individual's conditional choice probability, takes the binary logit form (with parameters $\mu\theta_i$, $\mu\beta_l$), i runs through all three data sets, and t runs through the repeated SP responses where relevant.

The integration in (13) is performed using Monte Carlo simulation methods. Given a trial value for ψ , we draw Θ_i^r from the assumed distribution $f(\Theta_i \mid \psi)$ and evaluate the likelihood function conditional on Θ_i^r, repeating for $r = 1, \ldots R$. The simulated likelihood function is:

$$SL(F, \Psi) = \prod_{it} \left(\frac{1}{R} \sum_{r=1}^{R} P(y_{it} \mid F, \Theta_i^r) \right). \tag{14}$$

The SMLE maximizes the log of this simulated likelihood function. Lee (1992) and Hajivassilio and Ruud (1994) show that under regularity conditions, the parameter estimates are consistent and asymptotically normal and, when the number of replications rises faster than the square root of the number of observations, asymptotically equivalent to maximum likelihood estimates.

5. Estimation Results

Our primary objective is to estimate distributions of the values of time and reliability based on a joint RP/ SP model. The final specification of this model is sufficiently complex that it will be easier to understand and justify our findings if we proceed in steps. Thus we first present estimates of the separate RP and SP models that form the basis of the joint model.

Revealed Preference Estimates

The Cal Poly RP sample is a simple cross section, but the Brookings RP sample has a panel structure. One way to analyze the Brookings sample is to estimate a binary logit model of lane choice on each observation, including those from the same motorist on different days. We call this a *trip-based* model. An alternative approach is to convert each motorist's multiple observations into one and estimate a model whose dependent variable is the frequency of using toll lanes. We call this a *person-based* model.

The person-based model has certain advantages. First, it correctly assumes that a traveler's decision to get a transponder (a prerequisite for choosing the express lanes) is based on long-run tradeoffs, not on daily considerations. Second, its simpler error structure makes it easier to combine with other data. As noted earlier, few travelers in the Brookings RP sample changed behavior from day to day, so little information is gained from the extra observations contributing to a trip-based model. Indeed, preliminary estimations indicated that the two models yielded similar results; thus, we focus here on the findings from person-based models. We explored alternative ways of specifying the person-based dependent variable, and settled on a binary outcome defined as 1 if the motorist used the express lanes for half or more of reported commuting trips, 0 otherwise.[17] Independent variables are defined as the average value over the days reported.[18]

Because commuters persisted in their route choices from day to day, it seemed reasonable to combine the Brookings person-based observations with the Cal Poly cross-sectional observations. We tested whether the Brookings and Cal Poly respondents react differently to the cost, time, and unreliability variables and found that there were no statistically significant differences. We do allow the random terms (η_i) to have different variances and we specify different alternative-specific constants ($\bar{\theta}$)in the two data sets. (Note that v_i^{BR} in (5) is redundant

[17] We omitted the few respondents who have a transponder but traveled two days or less, because defining a frequency for them involves too much error. Our specification can be thought of as a special case of an ordered logit model that divides the possible [0,1] interval (for fraction of trips made on the express lanes) into *j* sub-intervals. We explored several ways of doing this; based on Vuong's (1989) test for non-nested models, we could not reject any of the specifications we tried in favor of any other one, and all gave similar results for the main parameters of interest.

and can be set to zero in the RP-only case.) As noted, we expect $\sigma^{BR}<\sigma^{C}$ because we average Brookings motorists' choices over (several) days.

We follow convention in the systematic part of the specification by interacting travel cost with income. We also interact median travel time with distance. (We tried interacting unreliability with distance but found no statistically significant effect.)

Estimation results for this RP-only model are shown in the first column of Table 2. Most of the parameter estimates are statistically significant and have the expected signs. Commuters are deterred from the express lanes by a higher toll and from the free lanes by longer median travel times and greater unreliability. (Despite the interaction terms, this remains true throughout the full range of distance in our data.) We also found that women, middle-aged motorists, and motorists in smaller households—who may be more willing to indulge their travel preferences than motorists in larger households—are more likely to choose the toll lanes.[19]

We were unable to satisfactorily identify unobserved heterogeneity through mixed logit (these results are not shown in the table), but *observed* heterogeneity is indicated by preferences that vary in accordance with income and trip distance. Consistent with expectations, motorists with higher incomes are less responsive to the toll. The effect of distance on the time coefficient is captured well by a cubic form with no intercept (i.e., median travel time is not entered by itself). When graphed, the dependence of the value of time on distance is characterized by an inverted U, initially rising but then falling for trips greater than 45 miles. We conjecture that this pattern results from two opposing forces: the increasing scarcity of leisure time as commuting

[18] An exception is the "flexible arrival time" dummy, which is set to one if the respondent indicated flexible arrival for half or more of the reported days. (In fact, only five people reported any daily variation in this variable.)

[19] Parkany (1999), Lam and Small (2001), and Yan, Small, and Sullivan (2001) also find that women are more likely to use toll lanes.

takes up a greater fraction of it, and the self-selection of people with lower values of time to live farther from their workplaces (Calfee and Winston (1998)).

Stated Preference Estimates

The dependent variable is the respondent's choice of whether to use the express lanes in a given scenario. We can successfully estimate unobserved heterogeneity using SP data because the independent variables are, by design, not highly correlated. Mixed-logit estimations were performed using 1,500 random draws for the simulations, assuming that the random parameters for cost, time, and unreliability have independent normal distributions.[20]

The results, presented in the middle column of Table 2, indicate that key parameters have the expected signs and are estimated with good precision. As before, respondents trade off tolls with travel time and unreliability. Surprisingly, income is statistically insignificant, whether entered as a lane-choice shift variable or (as tried but not shown) interacted with the toll. As in the RP model, motorists in smaller households are more likely to select the express lanes; but here neither age nor sex has a statistically significant effect (at conventional levels) on their choices. Surprisingly, the SP estimates indicate that motorists are more likely to use a toll lane if they have a flexible arrival time; we speculate that this variable serves as a proxy for unmeasured job characteristics (e.g., managerial responsibility) requiring punctuality on a given day even though a commuter is not normally constrained to arrive at a particular time.[21]

We allow the coefficient on travel time to differ between people with long or short actual commutes to capture observed heterogeneity (these people received different versions of the SP

[20] This assumption leads to the possibility of a traveler having the "wrong" sign for these coefficients. We tried log-normal and truncated normal distributions for the random coefficients, but were unable to reach convergence—a problem noted by other researchers such as Train (2001), although Bhat (2000) and Calfee, Winston, and Stempski (2001) were successful with the log-normal.

[21] We also explored interactions of this variable with time and reliability, but found that it fit best in the form reported.

survey, as explained in Appendix A), but the difference is negligible. The parameters indicating the standard deviations of random coefficients are comparable in magnitude to the corresponding means, implying considerable unobserved heterogeneity.[22]

Joint RP/SP Estimates

Separate RP and SP models suggest that motorists' values of time and reliability vary in accordance with observed and unobserved influences. By combining the models, we obtain a more precise understanding of preference heterogeneity.

We assume the random components of the cost and time coefficients are the same across RP and SP models, but we cannot invoke this assumption for the unreliability coefficients because the measures of unreliability are constructed in different ways. Instead, we assume that the ratio of the standard deviation to the mean of the unreliability coefficient is the same across samples. Formally, the random parameters in the joint RP/SP model are specified as:

$$\beta_i^k = \bar{\beta}^k + \gamma^k Z_i^k + \varsigma_i \tag{15}$$

$$r_i^k = \bar{r}^k + \pi_i^k = \bar{r}^k \left(1 + \frac{\pi_i^k}{\bar{r}^k} \right) \equiv \bar{r}^k (1 + \omega_i) \tag{16}$$

where β_i refers to the vector of cost and travel time coefficients; Z_i is the matrix of individuals' characteristics including income and trip distance; r_i is the unreliability coefficient with random component π_i; ζ_i and ω_i are assumed independent normal with variances σ_ζ^2 and σ_ω^2 to be estimated; and k=BR, BP, C represent our data sources.[23] This specification allows the RP and

[22] We also estimated models with "inertia effects," in which the SP choice is conditioned on the actual choice of whether to obtain a transponder. Although this improves the goodness of fit considerably, it may introduce bias because the actual choice is not fully exogenous to the SP choice given the likely correlation of their error terms. Models of this type are discussed by Morikawa (1994, pp. 158-159) and estimated by Bhat and Castelar (2002).

[23] We also impose equality of the parameters $\bar{\beta}^k$, γ^k, and $\bar{r}^k$ across the two RP samples, as we did for the RP-only model.

SP values of time and of reliability to differ, but combines the power of the RP and SP data to estimate the random variation in those values. Finally, we assume that the individual characteristics affecting alternative-specific preferences (namely sex, age, flexible arrival, and household size) have the same effects across data sets.

We estimate the joint model by simulated maximum likelihood, using 2000 random draws.[24] Parameter estimates are presented in the last column of table 2. There is clearly a payoff from joint estimation because the coefficients of all the travel characteristics relevant to the RP choice are estimated with greater precision than before. The parameters capturing unobserved heterogeneity in the coefficients of cost, time and unreliability are also precisely estimated, as are the scale and correlation parameters describing the error structure. As expected, the scale parameter μ^c suggests that there is substantially more noise in the Cal Poly responses than in the Brookings RP responses, while the parameter ρ indicates that SP and RP responses from a single respondent are strongly correlated.

Motorists' Preferences and Heterogeneity

We use the estimated coefficients of the joint RP/SP model to calculate motorists' implied values of time and unreliability and indicate the extent of their heterogeneity (table 3). As can be seen from the second column of the table, all estimates are significantly different from zero at a 5% confidence level (one-sided test). The median value of time based on commuters' revealed preferences is $20.20/hour; at 87 percent of the average wage, it is toward the top of the range expected from previous work (Small (1992)). In our data, median time savings at the height of rush hour are 5.6 minutes; thus, the average commuter would pay $1.89 to realize these

[24] We found that the calculated standard errors were sensitive to the number of random draws up to 1500, but did not change when they were increased to 2000.

savings. The median value of reliability is $19.56/hour. Unreliability peaks at 3 minutes; thus, the average commuter would pay $0.98 to avoid this possibility of unanticipated delay. Given these estimates, the actual peak toll of $3.30 would be expected to attract somewhat fewer than half of the total peak traffic—which, in fact, it does.

We are also interested in how much motorists' preferences vary. We use the interquartile difference (the difference between 75th and 25th percentile values) as our heterogeneity measure because it is unaffected by high upper-tail values occasionally found in the calculations of ratios. This measure of heterogeneity exceeds 60% of the median value of time and is greater than the median value of unreliability, indicating that commuters exhibit a wide distribution of preferences for speedy and reliable travel.

It is interesting that the heterogeneity is almost all from unobserved sources, verifying the importance of "taste variation" in motorists' behavior and our attempt to capture it. To be sure, unobserved heterogeneity reflects limitations on empirical work and presumably could be reduced if it were possible to measure all variables that underlie individuals' preferences.

The implied SP values of time are smaller on average than the RP values. This finding may reflect the aforementioned tendency of travelers to overstate the travel time they lose or would lose in congestion. For example, suppose a motorist is in the habit of paying $1.56 to save 10 minutes, but perceives that saving as 15 minutes. That motorist may then answer SP questions as if he or she would pay $1.56 to save 15 minutes—yielding an SP value of time that understates the value used in actual decisions. The SP value of unreliability may be similarly biased, but we have no point of comparison. The median value of $4.17 per incident means that the median motorist in our sample would pay $0.42 per trip to reduce the frequency of 10-minute delays from 0.2 to 0.1.

6. Implications for Road Pricing Policy

Does preference heterogeneity create a strong case for differentiated services? We use Small and Yan's (2001) simulation model to investigate the potential effects of differential road pricing. The model resembles the SR91 road-pricing experiment, in which two 10-mile roadways, Express and Regular, connect the same origin and destination and have the same free-flow travel-time. Users are of two types: high value of time (i=1) and low value of time (i=2). Each user chooses the best option while congestion on each road adjusts endogenously.[25] The two user groups are of equal size when roads are free. Each has a downward-sloping demand as a function of the full price of travel (money cost plus an appropriate value of time and unreliability).

Because SR91 has twice as many regular lanes as express lanes, in equilibrium the express roadway contains only high-VOT users while the regular roadway contains both types of users. Equilibrium is therefore achieved when the full cost to high-VOT users is the same on both roadways, i.e., when the time difference multiplied by the VOT of group 1 equals the toll.

In Small and Yan's model, time and unreliability are not distinguished, but can be assumed to be functionally related. Thus to use the model with the results in this paper, we specify the full price p_{ir} for a user of type i on roadway r to be $p_{ir} = \tau_r + \varphi_i T_r + \delta_i R_r$, where τ is toll, T is travel-time delay (time less free-flow time), and R is unreliability. We assume that for each roadway, R_r / T_r is fixed at a value s=0.3785, which is the ratio of the average R to average T over the 4-hour peak period (5-9 a.m.) in the unpriced lanes in our data set. Thus $p_{ir} = \tau_r + \alpha_i T_r$, where $\alpha_i = \varphi_i + s\delta_i$. For φ_i and δ_i we use the VOT and VOR estimates in table 3 based on RP

[25] The function describing congestion is the Bureau of Public Roads formula T=0.15$T_0(V/C)^4$, where T_0 is free-fow travel time, T is travel time less T_0, and V/C is the volume-capacity ratio.

behavior, taking the two user groups to be represented by the 75th and 25th percentiles.[26] This yields values of $\alpha_1 = \$40.86/hr$, and $\alpha_2 = \$17.62/hr$.

Letting p_i be the lower of the two full prices facing user type i, the demand for trips by these users can be written as $N_i(p_i)$. We assume this function is linear, with parameters calibrated to reproduce real traffic conditions observed on SR91 in summer 1999. Thus each group's money-price elasticity is –0.58, as estimated by Yan, Small, and Sullivan (2001), and the time difference between the lanes is 6 minutes when the price on the express lanes maximizes the operator's profit subject to the regular lanes being free. These assumptions yield a plausible profit-maximizing express-lane toll of $4. This pricing policy may also be regarded as politically feasible because the Express Lanes enjoyed wide public acceptance at that time (Sullivan et al. (2000)).

Based on these parameters, we calculate tolls, travel times, changes in consumer surplus, and social welfare under several alternative pricing policies. Our base case has no toll on either roadway. The improvement in social welfare is the change (from the base case) in the two groups' combined consumer surplus plus toll revenues. Results are presented in table 4.

The first two policies set a price on the express lanes that maximizes social welfare subject to the current constraint that the regular lanes have a zero toll. This "second-best" policy enables road pricing to be politically feasible, but sacrifices efficiency by not pricing all lanes. Nonetheless, when there is user heterogeneity, welfare improves by $0.16 per vehicle, whereas if users are alike there is a negligible gain. Heterogeneity increases the potential efficiency of road pricing because those with a high VOT reap more benefits from the priced option, while those with a low VOT find it more important not to be subjected to policies aimed at average users.

---

[26] The third and sixth rows of table 3 show the difference between 75th and 25th percentiles. The percentiles themselves are: $27.70 and $15.10 for VOT, and $34.79 and $6.66 for VOR.

Suppose we wish to increase the welfare gain beyond that achieved by second-best policies, but we do not want to raise distributional concerns by, for example, disadvantaging low-VOT users more than high-VOT users. Can accounting for heterogeneity help? A first-best differentiated toll, shown in the fourth column of the table, achieves a substantial welfare gain of $0.86 per vehicle; but it also imposes high direct costs on motorists, especially the low-VOT users, undoubtedly creating a major political barrier to implementation. However, we can consider a "limited differentiated toll" that charges tolls on both lanes to maximize social welfare subject to a political feasibility constraint—defined as keeping the largest consumer surplus loss no greater than that in the second-best policy shown in the second column. Compared with the first-best toll, the result is a lower but more sharply differentiated toll that causes substantially smaller losses in consumer surplus for both groups and that narrows the gap between them. It also achieves a welfare gain that is more than one third that of first-best pricing and much larger than that of second-best pricing.

Catering to heterogeneity is apparently the key to softening the distributional effects of more efficient road pricing. This is indicated by a "limited uniform toll" policy shown in the last column of the table, defined to generate the same efficiency gain as the limited differentiated toll. It harms the low-VOT group far more than the high-VOT group. Thus if analysts consider only uniform tolls, they are likely to find that policymakers pay little attention to the efficiency gains because of large distributional disparities.

Traffic on SR91 has increased considerably since 1999. We show the effects of differentiated pricing with greater congestion by recalibrating the simulation model to double the time difference between the lanes that existed in the summer of 1999 (again, assuming that the operator's toll maximizes profit). The results, shown in table 5, indicate that the welfare gains from all the policies are more than doubled with increased congestion, yet the consumer-surplus losses in constrained policies are only about 50 percent greater. If we ignore heterogeneity, distributional concerns also increase as evidenced by the greater disparity among users groups with the limited uniform toll (last column). But this disparity is virtually eliminated by the limited differential toll. As congestion on major highways continues to grow, the case for accounting for heterogeneity will only strengthen.

7. Conclusion

Road pricing has been beloved by economists and opaque to policymakers for decades. Calfee, Winston, and Stempski (2001) rationalized this state of affairs by arguing that, in fact, few long-distance automobile commuters are willing to pay much to save travel time because those with a high value of time spend more on housing to live close to their workplace.

We have applied recent econometric advances to analyze the behavior of commuters in Southern California and found that those with very long commutes have substantially lower values of time, which is consistent with residential selectivity. But we have also found great heterogeneity in motorists' preferences for speed and reliability. One possible explanation is that in very expensive and congested metropolitan areas such as Southern California, consumers face significant constraints in trading off housing expense for commuting time. In such a situation, we find an opportunity to design pricing policies with a greater chance of public acceptance by catering to varying preferences.

Recent "value pricing" experiments have made a start at taking advantage of this opportunity. These experiments offer motorists the option to pay for congestion-free travel. But these experiments leave part of the roadway unpriced, which severely compromises efficiency.

We have demonstrated that pricing policies taking preference heterogeneity explicitly into account can realize substantial efficiency gains and ameliorate distributional concerns. By reducing the adverse direct impact of combined tolls and time savings on consumer surplus, differentiated pricing enhances the political viability of road pricing because policymakers must apportion only a modest fraction of the toll revenues to fully compensate road users. Differential pricing, embedded in both the design and marketing of recent experiments, may thus be the key to addressing the stalemates that impede transportation policy in congested cities.

Appendix A. Stated Preference Survey Questionnaire

Eight hypothetical commuting scenarios were constructed for respondents who travel on SR91. Respondents who indicated that their actual commute was less (more) than 45 minutes were given scenarios that involved trips ranging from 20-40 (50-70) minutes. An illustrative scenario follows:

Scenario 1

Free Lanes	**Express Lanes**
Usual Travel Time: 25 minutes	Usual Travel Time: 15 minutes
Toll: None	Toll: $3.75
Frequency of Unexpected Delays of 10 minutes or more: 1 day in 5	Frequency of Unexpected Delays of 10 minutes or more: 1 day in 20
Your Choice (check one):	
Free Lanes ☐	Toll Lanes ☐

Appendix B. Construction of RP Variables on Travel Time Savings and Reliability

Travel times on the free lanes (91F) were collected on 11 days: first by the California Department of Transportation on October 28, 1999 (six weeks before the first wave of our survey), and then by us on July 10-14 and Sept. 18-22, 2000 (which are the time periods covered by two later waves of our survey).

Data were collected from 4:00 am to 10:00 am on each day, for a total of 210 observations y_i of the travel-time savings from using the express lanes at times of day denoted by x_i, i=1,...210. Our objective is to estimate the mean and quantiles of the distribution (across days) of travel time y conditional on time of day x. To do so, we use non-parametric methods of the class of locally weighted regressions. In these methods, the range of the independent variables is divided arbitrarily into a grid, and a separate regression is estimated at each point of the grid. In our case, there is just one variable, x. For given x, the regression makes use only of observations with x_i near x, the importance of each being weighted in a manner that declines with $|x_i-x|$. The weights are based on a kernel function $K(\bullet)$, and how rapidly they decline is controlled by a *bandwidth* parameter h; typically only observations within one bandwidth of x get any positive weight.

The specific form of locally weighted regression we use is known as *local linear fit*. For each value of x, it estimates a linear function $y_i = a + b(x_i-x) + \varepsilon_i$ in the region $[x-h, x+h]$ by minimizing a loss function of the deviations between observed and predicted y. Denote the p-th quantile value of y, given x, by $q_p(x)$. Its estimator is then:

$$\hat{q}_p(x) = arg\ min_a \sum_{i=1}^{n} g_p[y_i - a - b(x_i - x)] \bullet K[(x_i - x)/h] \qquad \text{(A1)}$$

where $g_p(t)$ is the loss function. Similarly, denoting the mean of y given x by $m(x)$, its estimate is given by the same formula but with subscript p replaced by m.

In the case of mean travel-time savings, we use a simple squared-error loss function, $g_m(t) = t^2$, in which case equation (A1) becomes the *local linear least square regression.* In the case of percentiles of travel-time savings, including the median, we follow Koenker and Bassett's (1978) suggestion and use the following loss function, which is asymmetric except for the median (p=0.5):

$$g_p(t) = \{|t| + (2p-1)t\}/2 \qquad \text{(A2)}$$

With this loss function, equation (A1) defines the *local linear quantile regression* (Yu and Jones, 1997). It can be shown that the estimated percentile values converge in probability to the actual percentile values as the number of observations n grows larger, provided the bandwidth h is allowed to shrink to zero in such a way that $nh \to \infty$. In the case of the median (p=0.5), this is a least-absolute-deviation loss function, and therefore the estimator can be thought of as a non-parametric least-absolute-deviation estimator.

The choice of kernel function has no significant effect on our results. We use the *biweight* kernel function, which has the following form:

$$K(u) = \begin{cases} \frac{15}{16}(1-u^2)^2 & |u| \leq 1 \\ 0 & |u| > 1. \end{cases} \qquad \text{(A3)}$$

The choice of bandwidth, however, is important. We first tried the bandwidth proposed by Silverman (1985):

$$h = 0.9n^{-0.5} \min\left\{ \text{std}(\text{x}), \frac{d}{1.34} \right\} \qquad \text{(A4)}$$

where n is the size of the data set, "std" means standard deviation, and d is the difference between the 75$^{\text{th}}$ and 25$^{\text{th}}$ percentile of x. This bandwidth turns out to be about 0.5 hour for our data. However, there is rather extreme variation in our data at particular times of day, especially around 6:00 a.m., due to accidents that occurred on two days around that time. While these accidents are part of the genuine history and we want to include their effects, they produce an unlikely time pattern for reliability when used with the bandwidth defined by equation (A4) -- namely, one with a sharp but narrow peak in the higher percentiles around 5:30 a.m., followed by the expected broader peak centered around 7:30 a.m. We therefore increased the bandwidth to 0.8 hour in order to smooth out this first peak.

The standard deviation shown in figure 2 of the text is the square root of the estimated variance of time saving, obtained by a similar nonparametric regression of the squared residuals $\left(y_i - \hat{m}(x)\right)^2$ on time of day.

References

Bates, J., J. Polak, P. Jones, and A. Cook (2001), "The Valuation of Reliability for Personal Travel," *Transportation Research E: Logistics and Transportation Review*, 37, pp. 191-229.

Ben-Akiva, M., and T. Morikawa (1990), "Estimation of Travel Demand Models from Multiple Data Sources," in Koshi, M., ed., *Transportation and Traffic Theory*. New York: Elsevier, pp. 461-478.

Bhat, C. R. (2000), "Incorporating Observed and Unobserved Heterogeneity in Urban Work Travel Mode Choice Modeling," *Transportation Science*, 34, pp. 228-238.

Bhat, C. R. (2001), "Quasi-Random Maximum Simulated Likelihood Estimation of the Mixed Multinomial Logit Model," *Transportation Research B*, 35, pp. 677-693.

Bhat, C. R., and S. Castelar (2002), "A Unified Mixed Logit Framework for Modeling Revealed and Stated Preferences: Formulation and Application to Congestion Pricing Analysis in the San Francisco Bay Area," *Transportation Research B*, 36, pp. 593-616.

Brownstone, D., and K. Train (1999), "Forecasting New Product Penetration with Flexible Substitution Patterns," *Journal of Econometrics*, 89, pp. 109-129.

Calfee, J., and C. Winston (1998), "The Value of Automobile Travel Time: Implications for Congestion Policy," *Journal of Public Economics*, 69, pp. 83-102.

Calfee, J., C.Winston, and R. Stempski (2001), "Econometric Issues in Estimating Consumer Preferences from Stated Preference Data: A Case Study of the Value of Automobile Travel Time," *Review of Economics and Statistics*, 83, pp. 699-707.

Hajivassiliou, V., and P. Ruud (1994), "Classical Estimation Methods for LDV Models Using Simulation," *Handbook in Econometrics*, IV, R. Engle and D. MacFadden eds., Elsevire Science B. V., New York.

Hensher, D.A. (1994), "Stated Preference Analysis of Travel Choices: The State of Practice," *Transportation*, 21, pp. 107-133.

Hensher, D.A. (2001), "The Sensitivity of the Valuation of Travel Time Savings to the Specification of Unobserved Effects," *Transportation Research E: Logistics and Transportation Review*, 37, pp. 129-142.

Kawamura, Kazuya (2000), "Perceived Value of Time for Truck Operators," *Transportation Research Record* 1725, pp. 31-36.

Koenker, R., and G. S. Basset (1978), "Regression Quantiles," *Econometrica*, 46, pp. 33-50.

Lam, Terence C., and Kenneth A. Small (2001), "The Value of Time and Reliability: Measurement from a Value Pricing Experiment," *Transportation Research Part E: Logistics and Transportation Review*, 37, pp. 231-251.

Lee, L. (1992), "On Efficiency of Methods of Simulated Moments and Maximum Simulated Likelihood Estimation of Discrete Response Models," *Econometric Theory*, 8, pp. 518-552.

McFadden, D., and K. Train (2000), "Mixed MNL Models for Discrete Response," *Journal of Applied Econometrics*, 15, pp. 447-470.

Morikawa, T. (1994), "Correcting State Dependence and Serial Correlation in the RP/SP Combined Estimation Method," *Transportation*, 21, pp. 153-165.

Parkany, E. A. (1999), *Traveler Responses to New Choices: Toll vs. Free Alternatives in a Congested Corridor*, Ph.D. Dissertation, University of California at Irvine.

Rouwendahl, J., and E. Meijer (2001), "Preferences for Housing, Job, and Commuting: A Mixed Logit Analysis," *Journal of Regional Science*, 41, pp. 475-505.

Silverman, B. W. (1985), *Density Estimation for Statistics and Data Analysis*, London: Chapman and Hall.

Small, K. A. (1992), *Urban Transportation Economics*, Vol. 51 of *Fundamentals of Pure and Applied Economics Series*, Harwood Academic Publishers.

Small, K.A. and J. Yan (2001), "The Value of "Value Pricing" of Roads: Second-Best Pricing and Product Differentiation," *Journal of Urban Economics*, 49, pp. 310-336.

Sullivan, Edward, with Kari Blakely, James Daly, Joseph Gilpin, Kimberley Mastako, Kenneth Small, and Jia Yan (2000), *Continuation Study to Evaluate the Impacts of the SR 91 Value-Priced Express Lanes: Final Report*. Dept. of Civil and Environmental Engineering, California Polytechnic State University at San Luis Obispo, December (http://ceenve.ceng.calpoly.edu/sullivan/SR91/).

Train, Kenneth (2001), "A Comparison of Hierarchical Bayes and Maximum Simulated Likelihood for Mixed Logit," working paper, Dept. of Economics, Univ. of California at Berkeley, May (http://elsa.berkeley.edu/~train/papers.html).

Vuong, Q. (1989), "Likelihood Ratio Tests for Models Selection and Non-nested Hypotheses," *Econometrica*, 57, pp. 307-333.

Wardman, M. (2001), "A Review of British Evidence on Time and Service Quality Valuations," *Transportation Research E: Logistics and Transportation Review*, 37, pp. 107-128.

Winston, C. (1998), "U.S. Industry Adjustment to Economic Deregulation," *Journal of Economic Perspectives*, 12, pp. 89-110.

Yan, J., K. Small, and E. Sullivan (2001), "Choice Models of Route, Occupancy, and Time-of-Day with Value Priced Tolls," *Transportation Research Record*, forthcoming.

Yu, K., and M. C. Jones (1997), "A Comparison of Local Constant and Local Linear Regression Quantile Estimation," *Computational Statistics and Data Analysis*, 25, pp. 159-166.

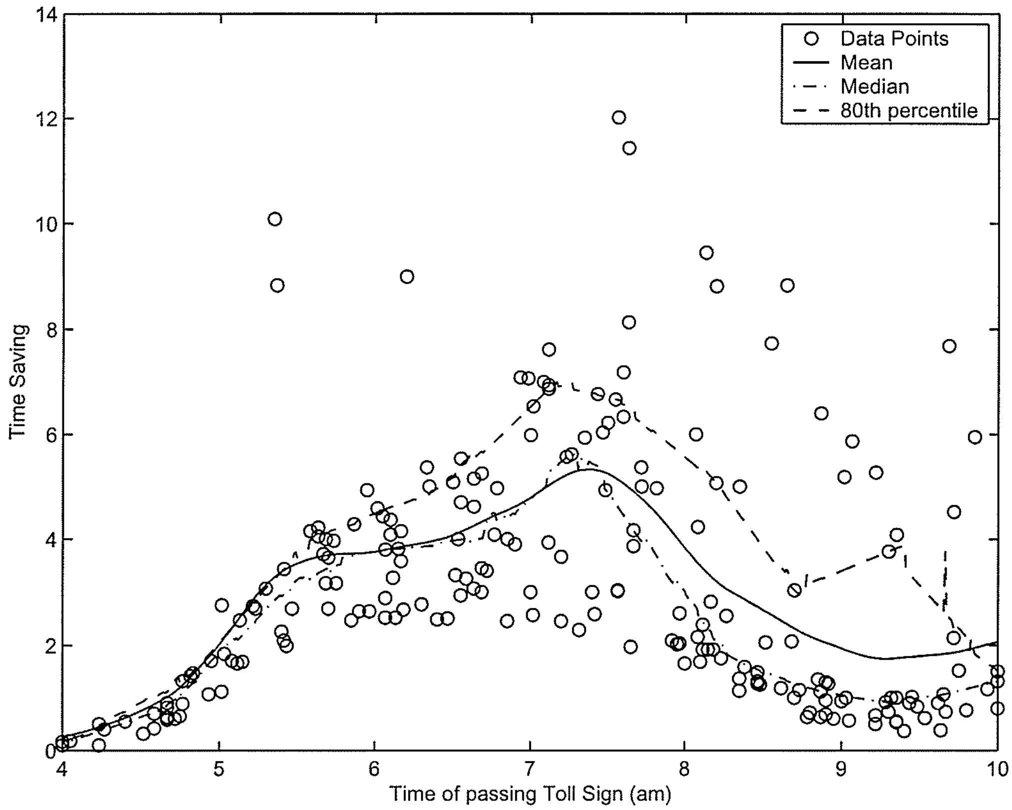

Figure 1. Time Saving

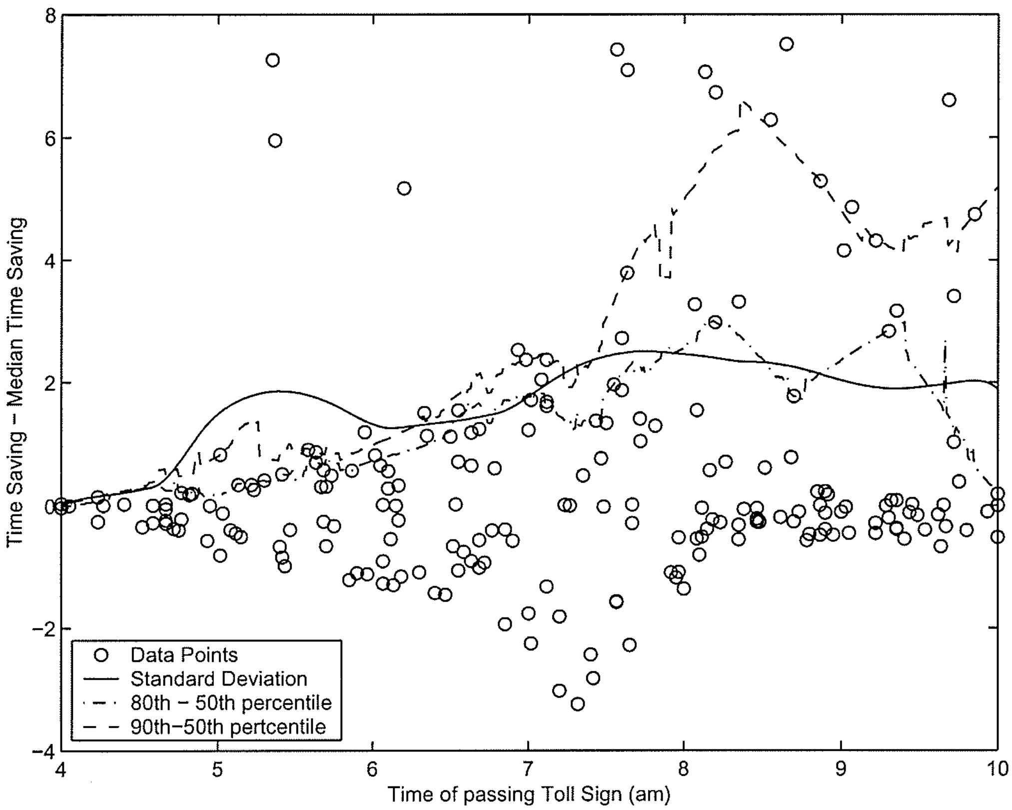

Figure 2. Dispersion of Time Saving

Table 1. Descriptive Statistics

	Value or Fraction of Sample		
	Cal Poly-RP	Brookings-RP	Brookings-SP
Route Share:			
91X	0.26	0.25	
91F	0.74	0.75	
One-Week Trip Pattern:			
Never Use 91X		0.68	
Sometimes Use 91X		0.13	
Always Use 91X		0.19	
Percent of Trips in Each Time Period:			
4:00am-5:00am	0.11	0.15	
5:00am-6:00am	0.22	0.13	
6:00am-7:00am	0.23	0.26	
7:00am-8:00am	0.20	0.21	
8:00am-9:00am	0.14	0.15	
9:00am-10:00am	0.10	0.10	
Age of Respondents:			
<30	0.11	0.12	0.10
30-50	0.62	0.62	0.64
>50	0.27	0.26	0.26
Sex of Respondents:			
Male	0.68	0.63	0.63
Female	0.32	0.37	0.37
Household Income ($):			
<40,000	0.14	0.23	0.24
40,000-60,000	0.24	0.60	0.59
60,000-100,000	0.40	0.15	0.13
>100,000	0.22	0.02	0.04
Flexible Arrival Time:			
Yes	0.40	0.55	0.50
No	0.60	0.45	0.50
Trip Distance (Miles):			
Mean	34.23	44.76	42.56
Standard Deviation	14.19	28.40	26.85
Number of People in Household:			
Mean	3.53	2.91	3.44
Standard Deviation	1.51	1.63	1.55
Number of Respondents	438	84	81
Number of Observations	438	377	633

Table 2. Parameter Estimates

Dependent Variable: 1 if chose toll lanes, 0 otherwise	**Coefficient** (standard error)[a]		
Independent Variable	**RP Only (binary logit)**	**SP Only (mixed logit)**	**Joint RP/SP (mixed logit)**
RP Variables			
Constant: Brookings sub-sample ($\bar{\theta}^{BR}$)	-0.5150 (0.9674)		0.2473 (0.7799)
Constant: Cal Poly sub-sample ($\bar{\theta}^{C}$)	-1.7157 (0.7827)		-1.8389 (0.6860)
Cost ($)[b,c]	-1.3443 (0.5312)		-2.2682 (0.3589)
Cost x dummy for medium household income ($60,000-$100,000)	0.4693 (0.2149)		0.6566 (0.2088)
Cost x dummy for high household income (>$100,000)	0.9047 (0.3096)		1.3147 (0.2794)
Median travel time (minutes) x trip distance (in units of 10 miles)[b]	-0.2618 (0.0917)		-0.4933 (0.1009)
Median travel time x (trip distance squared)	0.0412 (0.0162)		0.0868 (0.0189)
Median travel time x (trip distance cubed)	-0.0017 (0.0007)		-0.0037 (0.0009)
Unreliability of travel time (minutes)[b,d]	-0.5989 (0.2298)		-0.7049 (0.2550)
SP Variables			
Constant ($\bar{\theta}^{BS}$)		-2.3138 (1.6335)	-1.2246 (0.8856)
Standard deviation of constant[e] (σ_{ξ})		4.7793 (0.5239)	0.1284 (0.6669)
Cost[b,c]		-1.5467 (0.3311)	-1.0986 (0.3128)
Cost x dummy for high household income (>$100,000)		0.1625 (0.8980)	0.1915 (0.6469)

Cost x dummy for medium household income ($60,000-$100,000)		-0.3033 (0.5840)	-0.0827 (0.2948)
Travel time (minutes) × long-commute dummy (>45 min.)[b]		-0.2893 (0.0503)	-0.1834 (0.0394)
Travel time × (1 – long-commute dummy)		-0.3022 (0.0539)	-0.2127 (0.0590)
Unreliability of travel time (probability)[b,f]		-8.3054 (1.7956)	-5.1686 (1.1195)
Variables Pooled in Joint RP/SP Model			
Female dummy	1.1294 (0.3904)	1.7598 (1.0554)	1.3849 (0.4046)
Age 30-50 dummy	1.1951 (0.4465)	0.0035 (1.1638)	1.3021 (0.3856)
Flexible arrival-time dummy	0.2428 (0.3774)	2.9487 (1.1179)	0.7481 (0.4179)
Household size (number of people)	-0.3847 (0.1846)	-0.8717 (0.4076)	-0.5902 (0.1738)
Standard deviation of coefficient(s) of cost[g] (part of Ω)		0.8774 (0.2162)	0.6577 (0.1826)
Standard deviation of coefficients of travel time[g] (part of Ω)		0.2165 (0.0414)	0.1268 (0.0471)
Standard deviation of coefficient of unreliability[h] (part of Ω)		8.5383 (1.7455)	
Ratio of standard deviation to the mean for coefficients of unreliability[h] (σ_{ω})			0.9886 (0.3136)
Other Parameters			
Scale parameter:[i] Cal Poly sample (μ^{C})	0.5028 (0.1977)		0.3743 (0.0981)
Scale parameter:[i] SP sample (μ^{BS})			1.4723 (0.3585)
Correlation parameter – RP and SP (ρ)			2.5493 (0.4969)

Summary Statistics			
Number of observations	522	633	1155
Number of persons	522	81	548
Log-likelihood	-267.84	-241.32	-501.28
Pseudo R^2	0.3509	0.4500	0.3708

[a] Standard errors are the square root of the corresponding diagonal element in the inverse of the negative Hessian of the simulated log-likelihood function (calculated numerically).

[b] All cost, travel-time, and unreliability variables are entered as the difference between values for toll lanes and for free lanes. In the RP data, the cost for free lanes is zero, travel time for toll lanes is 8 minutes, and unreliability for toll lanes is zero. In the SP data, cost, travel time, and unreliability are specified in the questions.

[c] Value of "cost" for the toll lanes is the posted toll for a solo driver (for RP data) or the listed toll in the survey question (for SP), less 50% discount if car occupancy is 3 or more.

[d] Value of "unreliability" for the free lanes in the RP data is the difference between 80^{th} and 50^{th} percentile travel times (see text).

[e] The estimation of a standard deviation of constant θ_i^{BS}, separate from the standard deviation of the overall random term η_{it}^{BS}, is made possible by the multiple observations for a given individual in the SP data sample (see equation (9)). Hence there is no comparable parameter for the RP samples, where σ_ξ would be redundant with σ_η and so is assumed to be zero (see equations (8) and (10)).

[f] Value of "unreliability" for either set of lanes in the SP data is the probability of unexpected delays of 10 minutes or more, as given in the survey question and applying to the entire trip.

[g] The coefficients on cost and travel time are specified as equation (4) for separate RP and SP models, and as equation (15) for the joint model. The only difference is that in (15) the variance of ζ_i is specified as identical for the RP and SP samples; hence for each variable (cost and time), we estimated separate RP and SP mean coefficients but only a single standard deviation for those coefficients. Note that in our specification for travel time, the "intercept" $\bar{\beta}$ is set to zero, i.e. travel time is entered *only* through interactions with individual characteristics Z_i. In addition, in the RP-only model ζ_i=0 (i.e. the parameters are not random).

[h] The coefficient on unreliability is specified as equation (4) for separate RP and SP models, and as equation (16) for the joint model. Note that in our specification for unreliability, γ=0, i.e., unreliability has *no* interactions with individual characteristics Z_i. In addition, in the RP-only model ζ_i=0 (i.e. the parameters are not random).

[i] Scale parameters are defined in equations (11)-(12). A value less than one means there is more unexplained dispersion in this portion of the data than in the Brookings RP data.

Table 3. Values of Time and Reliability from Joint RP/SP Models

	Median Estimate	90% Confidence Interval[a] [5%-ile, 95%-ile]
RP Estimates		
Value of time ($/hour)		
Median in sample	20.20	[14.72, 25.54]
Unobserved heterogeneity[b]	11.01	[6.48, 16.74]
Total heterogeneity in sample[b]	12.60	[8.30, 18.12]
Value of reliability ($/hour)		
Median in sample	19.56	[8.03, 31.17]
Unobserved heterogeneity[b]	27.67	[11.56, 47.64]
Total heterogeneity in sample[b]	28.13	[11.56, 48.58]
SP Estimates		
Value of time ($/hour)		
Median in sample	9.46	[6.18, 13.53]
Unobserved heterogeneity[b]	13.46	[7.41, 22.02]
Total heterogeneity in sample[b]	13.56	[7.52, 22.99]
Value of reliability ($/incident)		
Median in sample	4.17	[2.37, 6.30]
Unobserved heterogeneity in sample[b]	7.78	[4.36, 12.64]
Total heterogeneity[b]	7.79	[4.36, 12.66]

[a] The confidence interval represents uncertainty due to statistical error, not heterogeneity. It is determined by Monte Carlo draws from the estimated statistical distributions of the parameter estimates. This method is more accurate than approximation formulas based on the standard errors of and correlation between coefficient estimates. The distributions of these ratios are skewed, so the standard deviation would give a misleading characterization of precision. A positive 5th percentile value means the quantity is significantly greater than zero according to a conventional one-sided hypothesis test at a 5 percent significance level.

[b] Heterogeneity is measured as the interquartile difference, i.e., the difference between the 75th and 25th percentile values, computed from Monte Carlo draws. For unobserved heterogeneity, these draws are from the estimated distribution of random parameters ς_i (for the value of time) or ω_i (for the value of unreliability in equations (15)-(16)). For total heterogeneity, the draws are from that distribution and from the relevant RP or SP sample.

Table 4. Simulation Results—Summer 1999 Traffic Conditions

PRICING REGIME[a]	Base case: no toll	Second-best toll: heterogeneity present	Second-best toll: heterogeneity not present	First-best differentiated toll	Limited differentiated toll	Limited uniform toll
Toll:						
Express lanes	0	$1.80	$0.97	$4.51	$1.34	$0.78
Regular lanes	0	0	0	$4.18	$0.47	$0.78
Travel time (minutes):						
Express lanes	14	11	12	10	12	13
Regular lanes	14	15	14	11	14	13
Consumer surplus:[b]						
High-VOT users	0	-$0.45		-$2.41	-$0.44	-$0.40
Low-VOT users	0	-$0.26		-$2.82	-$0.45	-$0.55
Homogeneous users	0		-$0.23			
Social welfare[b]						
All users	0	$0.16	$0.06	$0.86	$0.28	$0.28

[a] Notes on pricing regimes: "Second-best toll (heterogeneity present)" and "Second-best toll (heterogeneity not present)" maximize social welfare subject to the price of the regular lanes being constrained to zero, with and without heterogeneity, respectively. "First-best differentiated toll" maximizes social welfare using differentiated tolls, with no constraint. "Limited differentiated toll" maximizes social welfare using differentiated tolls, subject to the consumer surplus loss of the worst group being constrained the same as in "Second-best toll (heterogeneity presented)". "Limited uiform toll" is the uniform toll providing the same total welfare gain as "Limited differentiated toll".

[b] Consumer surplus and social welfare are measured relative to the no-toll scenario and divided by the number of users in the no-toll scenario to put them in per capita terms. Social welfare is equal to the sum of the two groups' consumer surplus plus revenue, divided by total number of users in the no-toll scenario.

Table 5. Simulation Results—Increased Congestion

PRICING REGIME[a]	Base case: no toll	Second-best toll: heterogeneity present	Second-best toll: heterogeneity not present	First-best differentiated toll	Limited differentiated toll	Limited uniform toll
Toll:						
Express lanes	0	$4.42	$2.68	$8.51	$2.81	$1.43
Regular lanes	0	0	0	$7.93	$0.77	$1.43
Travel time (minutes):						
Express lanes	20	14	15	12	16	18
Regular lanes	20	21	20	13	19	18
Consumer surplus:[b]						
High-VOT users	0	-$0.71		-$2.66	-$0.68	-$0.54
Low-VOT users	0	-$0.42		-$3.31	-$0.71	-$0.89
Homogeneous users	0		-$0.38			
Social welfare[b]	0	$0.48	$0.23	$2.18	$0.67	$0.67

[a,b] See the footnotes of Table 5.