UCI-ITS-WP-05-4

Differentiated Road Pricing, Express Lanes, and Carpools: Exploiting Heterogeneous Preferences in Policy Design

UCI-ITS-WP-05-4

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

[1]Department of Economics and
Institute of Transportation Studies
University of California, Irvine; Irvine, CA 92697-3600, U.S.A.
ksmall@uci.edu

[2]The Brookings Institute
Washington, D.C., USA

[3]The Hong Kong Polytechnic University
Hung Hom, Kowloon, Hong Kong

November, 2005

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

DIFFERENTIATED ROAD PRICING, EXPRESS LANES, AND CARPOOLS:

EXPLOITING HETEROGENEOUS PREFERENCES IN POLICY DESIGN

Kenneth A. Small
University of California, Irvine

Clifford Winston
Brookings Institution

Jia Yan
Hong Kong Polytechnic University

November 2005

Presented at the Brookings-Wharton Conference on Urban Affairs, Dec. 8-9, 2005.
Revised version forthcoming in *Brookings-Wharton Papers on Urban Affairs*.

Introduction

The U.S. highway system, largely constructed with public funds from the fuel tax, could be characterized as a public good if it were rarely congested. But like many public goods that are available at little or no charge, its quality has deteriorated with the intensity of use. Today, the nation's road system has turned into a "tragedy of the commons" as road users experience nearly 4 billion hours of annual delay (Schrank and Lomax, 2005).

Historically, the public has had a status-quo bias against economists' recommendations to use the price mechanism to reduce congestion.[1] Policymakers have therefore pursued other approaches that face less political resistance, such as regulations that allocate reserved lanes to vehicles carrying two or more people. But recent evidence indicates that these "high-occupancy-vehicle" (HOV) lanes sometimes carry fewer people than general-purpose lanes, attract family members who would ride together anyhow, and discourage more efficient vanpools because they attract low-occupancy carpools (Orski 2001, Poole and Balaker 2005). As a result, the public has sporadically voiced its disapproval.

A recent innovation is to fill the reserved capacity not used by HOVs with solo drivers willing to pay a toll. These so-called "high-occupancy-toll" (HOT) lanes can be found in the Los Angeles, San Diego, and Houston areas and they are currently under consideration in many others including Denver, Seattle, San Francisco, and Washington, DC.

HOT lanes appeal to a broad set of motorists who are sufficiently inconvenienced by congestion to pay a sizable toll to travel on less congested lanes, either daily or as dictated by their schedules. Although these preferences indicate that the public is not rigidly opposed to congestion pricing, HOT lanes are questionable on welfare grounds for two reasons. By leaving

[1] Small, Winston, and Evans (1989), Mohring (1999), and the papers in Santos, ed. (2004) provide recent discussions of road pricing.

most of the highway unpriced, they continue to allow most motorists to impose high external congestion costs on each other; and by creating a big price differential on the two roadways, they may leave the express lanes relatively underutilized, even if less so than HOV-only lanes (Small and Yan, 2001). Indeed, simulations show that HOT lanes sometimes *lower* welfare compared with keeping all lanes in general use, particularly if they are priced high enough to allow motorists to travel at speeds near-free-flow—a condition that typically results from political considerations or public-relations goals.

In short, HOV and HOT policies do not appear to have answered the long-standing call for an efficient yet politically viable road pricing policy. In this paper, we seek to identify such policies by analyzing the behavior of motorists traveling on California State Route 91 (SR91) in Orange County. These travelers have the option of traveling solo on the general lanes, paying a toll to use the HOT express lanes, or forming a carpool to use the express lanes at a discount. Because travelers are likely to vary in their preferences for speedy and reliable travel, our model of their choices captures observed and unobserved heterogeneity.[2] We find that users of SR91 have high average values of travel time and of the reliability of travel time, and that the distributions of these values exhibit considerable dispersion.

We then show that by designing differentiated pricing schemes for general and express lanes that cater to such varying preferences, it is possible to capture much of the efficiency that HOV and HOT policies sacrifice while generating smaller welfare disparities among road users than more efficient pricing policies — small enough, in fact, to be comparable to policies that already have passed the test of politically acceptability in at least a few urban areas.

[2] Previous empirical studies that allow for heterogeneous preferences among motorists include Calfee, Winston, and Stempski (2001), Hensher (2001), Jiang and Morikawa (2004), Steimetz and Brownstone (2005), Hess, Bierlaire, and Polak (2005), and Small, Winston, and Yan (2005a). Simulation studies incorporating heterogeneity to analyze pricing scenarios include Small and Yan (2001), Verhoef and Small (2004), and De Palma and Lindsey (2004).

Empirical Model of Travel Choices

California State Route 91 is a major limited-access expressway used heavily by long-distance commuters. A 10-mile stretch in Orange County includes four free lanes (91F) and two express lanes (91X) in each direction. Motorists who wish to use the express lanes must set up a financial account and carry an electronic transponder to pay a toll, which varies hourly according to a preset schedule. Carpools of three or more could use the express lanes at a 50 percent discount at the time covered by our surveys.[3] Unlike the regular lanes, the express lanes have no entrances or exits between their end points.

In an earlier paper using the same data, we modeled motorists' lane choice only (Small, Winston, and Yan 2005a). Here we model three decisions by motorists: whether to acquire a transponder, which gives them the flexibility to use the express lanes when they want to; whether to travel on the express or free lanes for the trip in question; and how many people to travel with in their vehicle. These three choices are assumed conditional on related choices including travel mode (car or public transport), residential location, and time of day of travel. (In our context, mode choice is unimportant because public transportation has a very small share of travelers on the corridor served by SR91.) In the empirical analysis that follows, we combine data that describe motorists' actual decisions for their morning commute on SR91 with data indicating hypothetical choices between the express and free lanes under varying travel conditions.

Formally, traveler n faces a choice whether to have a transponder (T) or not (N), whether to travel on a free lane (F) or express lane (X), and whether to travel with 1, 2, or 3 people in the vehicle (where 3 means three or more). The three choice dimensions define 2x2x3=12

[3] As of late May 2003, these carpools can travel for free except weekday afternoons outbound from 4-6 p.m., when they receive a 50 percent discount. Current information is available on the Orange County Transportation Authority's "91 Express Lanes" website at http://www.91expresslanes.com.

alternatives, but only nine of them are legally available because a highway traveler must have a transponder to use the express lane, thereby eliminating combinations *NX*1, *NX*2, and *NX*3.

The random utility of traveler n choosing an alternative j is:

$$U_{jn} = X_{jn}\beta_n + \varepsilon_{jn} \tag{1}$$

where X_{jn} is a vector of attributes associated with alternative j including the toll, travel time, and reliability that apply to the traveler's trip; β_n is a vector of parameters that capture the traveler's preferences for those attributes; and ε_{jn} is an error term capturing unobserved influences. We allow for preference heterogeneity by specifying parameter vector β_n as:

$$\beta_n = W_n\gamma + \mu_n\,, \tag{2}$$

where W_n is a vector of explanatory variables relating to the individual, γ is a vector of parameters to be estimated, and μ_n is a vector of independent normal random variates with variances to be estimated. The term $W_n\gamma$ describes the observed distribution of preferences, while μ_n represents unobserved preference heterogeneity.

If ε_{jn} were iid extreme value, then equations (1) and (2) would constitute a conventional mixed-logit model, whose choice probabilities are expressed as integrals (over the distribution of μ_n) of the multinomial logit choice probability (conditional on β_n). However, we specify the structure of ε_{jn} in detail to account for certain features of the data set, which we describe later, and of the choice set, which we describe now.

For a given individual, it is likely that the alternative-specific preferences for our nine permitted alternatives are not independent of each other, but rather are shaped by the three dimensions of choice in our model. These preferences could be characterized by a nested-logit specification for ε_{jn}; but given that we use a mixed logit model to allow for preference heterogeneity, it is easier — and more flexible in practice — to specify random preferences for groups of alternatives as presented by Brownstone and Train (1999). Thus we let ε_{jn} include four distinct preferences: for a transponder, for the express lane, for a two-person carpool, and for a three-person carpool. Namely:

$$\varepsilon_{jn} = \Delta_j^T \nu_n^T + \Delta_j^X \nu_n^X + \Delta_j^{H2} \nu_n^{H2} + \Delta_j^{H3} \nu_n^{H3} + \eta_{jn} \tag{3}$$

where Δ_j^k denotes a dummy variable equal to one if alternative j is one of those characterized by a transponder (when k=T), by the express lane (when k=X), or by a 2- or 3-person carpool (when k=H2 or H3). The four variables ν_n^k are independent normal variates each with mean zero and standard deviation σ^k to be estimated. (For parsimony, we impose σ^{H2}=σ^{H3}≡σ^{HOV}.) The remaining errors η_{jn} are assumed to be iid extreme value, with variances depending on the data structure indicated in the next subsection.

We specify two components of μ_n in equation (2) with non-zero variances: one (μ_n^{Time}) for the coefficient of travel time, the other (μ_n^{Rel}) for the coefficient of (un)reliability. They have standard deviations σ^{Time} and σ^{Rel}. In sum, we specify six independent normal random terms (ν's and μ's) with five unknown standard deviations to be estimated.

We define the value of travel time (VOT) and value of reliability (VOR) as the ratios of marginal utilities of travel time and reliability to the marginal utility of money cost. Given our specification of utility, the measures are expressed as

$$VOT_n = \frac{\beta_n^{Time}}{\beta_n^{Cost}} \tag{4}$$

$$VOR_n = \frac{\beta_n^{\mathrm{Re}l}}{\beta_n^{Cost}} \ , \tag{5}$$

where β_n^{Time}, $\beta_n^{\mathrm{Re}l}$, and β_n^{Cost} are the coefficients of travel time, reliability of travel time, and toll in equation (1). These values depend on observables W_n and random components μ_n through equation (2).

Data Set

We combine survey data from three samples of people traveling between 4:00 am and 10:00 am on the California State Route 91 corridor westbound who have the option of using the express lanes. The surveys were taken over a 10-month period in 1999 and 2000. The first survey was a telephone survey generating 435 observations pertaining to actual travel on a particular day, conducted by researchers at California Polytechnic State University at San Luis Obispo (CalPoly) with our participation (Sullivan *et al.* 2000). Thus it consists of revealed preference (RP) data.

The second and third samples are from a two-stage mail survey collected by us through the Brookings Institution. The initial stage collected RP data from 79 respondents on actual trips taken during a week of travel, while a follow-up stage presented to each respondent eight stated preference (SP) scenarios.[4] In each SP scenario, the respondent was asked to choose between two otherwise identical routes with specified hypothetical tolls, travel times, and probabilities of delay; an illustrative scenario is presented in the appendix. The SP sample contains 78 respondents, who generated 610 observations; 54 of these people also answered the RP questions. Detailed descriptions of the samples are presented in Small, Winston, and Yan (2005b).

[4] The Brookings RP sample actually contains information for all commuting trips made within the survey week, which could be treated as separate observations. However, 87 percent of the respondents made the same choice every day and nearly all of the others varied on only one day. So we simplify, with little information loss, by creating a binary response variable equal to one if the respondent chose the express lanes for half or more of the days reported. We tried variants of this response variable with virtually no changes in results.

By constructing a sample that contains both RP and SP observations, we can overcome the main drawbacks of each type of data. The use of RP data is often hindered by strong correlations among travel cost, time, and reliability; whereas SP data raise concerns about whether the behavior exhibited in hypothetical situations applies to actual choices. By specifying some parameters to be identical and others different in the utility functions generating RP and SP choices, we can improve the precision in estimating common parameters (due to low correlations designed into the SP questions) while allowing for suspected behavioral differences in other parameters.

Table 1 presents some statistics on socioeconomic variables and trip distance. The Brookings RP sample appears to represent well the population characteristics of the SR91 catchment area, tracking census information for the two relevant counties except for household income—which, naturally, is higher for our respondents because most of them are commuters.[5] We estimate the average wage rate to be \$23/hour.[6] The CalPoly sample has higher household incomes and shorter trip distances than the Brookings samples, evidently being drawn from a smaller and more affluent geographical area. Our model includes income and trip distance as exogenous variables, so these sampling differences should not affect our parameter estimates.

Table 2 presents the choice shares of the nine alternatives associated with each RP sample. We observe a difference among carpooling propensities between the CalPoly and Brookings samples, with many fewer carpools in the latter. To better understand the difference,

---

[5] Our sample's median income is \$46,250, whereas the average incomes in the two counties where our respondents lived were \$36,189 and \$39,729 in 1995, as estimated by the Population Research Unit of the California Department of Finance.

[6] 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, and assign to each person in our sample the average BLS wage rate for that person's 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 CalPoly sample is disaggregated into four subsamples representing different ways of finding respondents (Sullivan *et al.* 2000). The “Random” subsample was obtained by telephone interviews drawn randomly from lists of telephone exchanges in the relevant area; the other three CalPoly subsamples came ultimately from license plates observed on the highway and are therefore choice-based (some were purposely carpool-enriched). Even in the CalPoly “Random” subsample, however, the combined carpool shares are considerably higher (24%) than in the Brookings RP sample (6%), despite both being obtained from random telephone calls. The CalPoly Random shares are much closer to the observed peak-period carpool shares on the SR91 roadway (Sullivan *et al.* 2000), so we conclude that the Brookings sample undersampled people who carpool — possibly because the telephone screening questions to determine eligibility for the survey were originally designed to only find solo drivers and subsequently modified. Thus, we use the CalPoly Random subsample as our measure of the population choice shares, and correct for choice-based sampling in our estimates by applying carpool-share weights to the other subsamples (Manski and Lerman 1977).

Specification and Estimation

We posit that motorists’ joint choices are influenced by their socioeconomic characteristics and the characteristics of their journey, including the total trip distance and the toll, travel time, and (un)reliability of travel time on the portion of the journey where a lane choice exists.

The express lane toll for a given trip is the published toll for the time of day the motorist reported passing the sign that indicates the toll level. It is discounted by 50% if the trip was in a carpool of three or more. (We asked respondents, even in the SP survey, to indicate their vehicle occupancy for actual trips.) We obtained estimates of travel time and reliability based on field measurements made by students who drove at many different times on eleven different days that correspond to the travel periods covered by our surveys. Thus we were able to construct the distribution of travel times across days as a function of the time of day. The most satisfactory fits of our model were obtained by specifying travel time as its median value across days, and unreliability as the difference between the 80th and 50th percentiles of the distribution of travel times across days. A rationale for using the 80th-50th percentile difference to measure unreliability is that people are more concerned with unexpected late arrivals than early arrivals; this measure also is less correlated with median travel time than a symmetric measure such as the variance. Small, Winston, and Yan (2005b) discuss the procedures used to estimate these measures and to validate their accuracy.

The socioeconomic variables include age, sex, household size, and per-capita income. Two other variables are constant across alternatives: trip distance and trip purpose (a dummy for work trip). We explored a number of other variables concerning arrival-time flexibility, occupation, education, and size of the workplace, but found that they have little explanatory power and omitting them did not materially influence the other parameter estimates.

All RP and SP variables were defined in the same way except that we measured reliability in the SP scenarios as the frequency of being delayed 10 minutes or more, which we convert into a probability for analysis, rather than trying to explain probability measures to survey respondents.

A number of specification issues arise when we combine the RP and SP data sets. The RP analysis is described by equations (1)-(3), to which we append superscript *RP* to distinguish those observations. The SP analysis, however, is different because we asked each respondent to express only a binary choice between express or regular lanes; and we asked for this choice in eight different scenarios (each with different hypothetical values of travel variables). Because the SP choice is binary, it is convenient in the case of SP respondents to replace (1) by the utility *difference* between the express and regular lane. Thus in each choice scenario t, the respondent n chooses the express lane if and only if

$$U_{nt}^{SP} \equiv X_{nt}^{SP} \beta_n^{SP} + \varepsilon_{nt}^{SP} \geq 0, \tag{6}$$

with (2) taking the form:

$$\beta_n^{SP} = W_n^{SP} \gamma^{SP} + \mu_n^{SP}. \tag{7}$$

We account for two additional types of error correlation that may arise due to the nature of the SP sample and to its being combined with the RP samples. First, we expect ε_{nt}^{SP} to exhibit a typical panel structure whereby it contains one random term, ξ_n, common to all the choice scenarios considered by individual n. Second, in the 55 cases where the same individual answered both the RP and SP questions, we expect a part of the error to be common between them: specifically, we expect the SP error to contain a term proportional to ν_n^X from equation (3), representing the RP preference for travel in the express lane. Thus we specify the SP error as:

$$\varepsilon_{nt}^{SP} = \xi_n + \theta \nu_n^X + \eta_{nt}^{SP} \tag{8}$$

where ξ_n has zero mean and variance normalized to one; θ is a parameter to be estimated; and η_{nt}^{SP} has a logistic distribution with standard deviation σ^{SP}.

Finally, we follow standard practice in combining RP and SP data by allowing for differences between revealed and stated choices in the variance of random preferences; for similar reasons we allow for a difference between the two RP data sets, namely Brookings RP

(BR) and CalPoly (C). We can normalize one variance, which we choose by setting $\sigma_{SP}=\pi/\sqrt{3}$ (as in binary logit). We parameterize the other variances as ratios

$$\tau^{BR} = \frac{\sigma^{SP}}{\sigma^{BR}} \tag{9}$$

$$\tau^{C} = \frac{\sigma^{SP}}{\sigma^{C}} \tag{10}$$

which are described in our estimation results as “scale parameters”.

We express the log-likelihood function for our sample as the summation of choice probabilities for RP observations (choice among nine alternatives) and for SP observations (binary choice), with the common error term ν_n^X entering both RP and SP choices for those people who are members of both Brookings samples. Each choice probability is a mixed logit expressed as an integral of a multinomial or binary logit probability, conditional on normal random variates, over the distribution of those variates. We obtain parameter estimates by using the *maximum simulated likelihood estimator* (MSLE) described by Lee (1992) and Train and McFadden (2000). The MSLE performs the integration by Monte-Carlo simulation.

Identification

Our model is identified by assuming that any unobserved influences on transponder, vehicle occupancy, and route choices do not vary systematically by time of day; if they did, they would be correlated with the cost, time, and reliability of travel and would therefore bias those coefficients. The validity of this assumption depends to a large extent on how well our observed variables capture taste variation across time of day. Fortunately, it appears that such variation is reflected in several of our variables. For example, a motorists’ sex is correlated with the time of day of travel: females constitute only 15% of those people traveling during the interval 4:00-5:00

am, but 39% of the 7:00-8:00 am group. Similarly, the proportion of respondents whose trips are work trips varies from 100% at the earliest time to 58% at the latest time.

In Small, Winston, and Yan (2005a), we conducted a formal test of whether unobserved taste variation by time of day affects cost and travel-time parameters. We did this by estimating models of lane choice that included five time-of-day dummy variables. The findings indicated that values of time and reliability were not affected very much, which is likely to be true in the model here as well. We do not include the time-of-day dummies in the current model because in the previous work they reduced the precision of the estimates.

Estimation Results

Estimation results are presented in Table 3. We group the RP parameters as those for generic variables that influence all three choice dimensions (transponder, lane, and vehicle occupancy) simultaneously, and those that influence just one of those dimensions. We also group separately those parameters influencing only the SP lane choice and those having a common effect on RP and SP choices.[7] Most influences are statistically significant and have the expected sign. As indicated by the generic RP coefficients and the SP coefficients, motorists pay attention to the toll, travel time, and reliability when choosing among the available alternatives.

Observed heterogeneity in preferences is indicated by interactions, for instance, between cost and income and between travel time and various functions of trip distance. As expected, motorists with higher incomes are less responsive to the toll, a statistically significant effect for

[7] We conducted an extensive exploration of alternative specifications and functional forms for the explanatory variables, including removing the equality constraints between certain RP and SP parameters reflected in the "combined estimates" in the table. The model presented here is generally robust to such variations and in most cases is not rejected by statistical tests against more general models, although frequently the generalizations achieve notably worse precision.

RP respondents. The deterrent effect of travel time varies with distance in an inverted U pattern, initially rising but then falling for trips greater than 32 miles. Following Calfee and Winston (1998), we conjecture that this pattern results from two opposing forces: the increasing scarcity of leisure time as commuting becomes longer, and the self-selection of people with lower values of time into farther out residences. For SP, we allow the coefficient on travel time to differ between people with long and short commutes, but the difference is negligible.

We also find observed heterogeneity in alternative-specific preferences. Middle-aged females and commuters are more inclined than other motorists to acquire a transponder. Middle-aged females with large families are more likely than other motorists to carpool, perhaps because they are more likely to make trips where family members ride together. Finally, as indicated by the combined estimates, women, middle-aged motorists, and motorists in smaller households are more likely than others to choose the toll lanes, even given transponder acquisition and car occupancy.[8]

Substantial unobserved heterogeneity is indicated by the standard deviations of the random normal variates indicating preferences over travel time and reliability, as well as those indicating absolute preferences for the express lanes and for carpooling. (We tried also to estimate a random coefficient for the toll, but we were unable to obtain stable results.) The standard deviations are estimated with good precision and are substantial in magnitude, ranging from roughly one-fourth of the corresponding mean coefficient to a multiple of it.[9] The scale

[8] To better understand why women are more likely to use the toll lanes, we tried including an interaction of four variables: gender, age, household size, and either the express-lane dummy or the travel-time-uncertainty variable. This interaction sought to test whether working mothers with children might prefer the toll lanes or be more averse to unreliability due to tighter schedules. However, we could not find a measurable effect.

[9] The ratio of standard deviation to the mean coefficient is directly estimated for (un)reliability at 1.32. In the case of travel time, the estimated standard deviation of 0.39 may be compared with the SP coefficient of travel time of about -0.36 and with the derivative of utility with respect to RP median travel time, which is -0.76 at the mean trip distance of the Brookings RP sample.

and correlation parameters that describe the error structure are also estimated precisely and show, as expected, that the RP and SP responses from a single individual are strongly correlated.

In Table 4, we use our parameter estimates to compute properties of the distributions across individuals of motorists' implied value of time (VOT) and reliability (VOR). We do this both for all road users combined and for users of the express lanes and the free lanes separately. We use the Brookings RP sample for enumeration because it best represents the population, as argued previously. (We do not have to be concerned about possible bias from choice-based sampling that understated carpool frequency because we do not use information about the Brookings survey respondents' choices here.) In the table, we characterize heterogeneity in VOT and VOR by the interquartile range (i.e., the difference between the 75th and 25th percentile values) across individuals, a measure that is relatively robust to the high upper tails typically found in distributions of ratios of random variables. The results are obtained using simulation procedures described in Small, Winston, and Yan (2005a).

As shown by the 90% confidence intervals in the second column, all of the reported estimates are statistically different from zero using a one-sided test at a five percent significance level. We find that the median value of time is $19.63/hour, which is about 85% of the average wage rate and thus near the upper end of the range expected from previous work (Small 1992). The median value of reliability is slightly higher at $20.76/hour. Motorists also exhibit a wide range of preferences for speedy and reliable travel as total heterogeneity in VOT and VOR is nearly equal to, or greater than, the corresponding median value. As expected, express-lane users have on average higher values of travel time and reliability than do users of the free lanes — but there is a wide and overlapping range within each group.

Simulating Highway Policies

We explore the policy implications of the preference heterogeneity we have found by developing a simulation model that uses our econometric results. It allows us to examine current HOT and HOV policies and alternative pricing policies. We begin with a situation closely resembling the SR91 road-pricing experiment: two 10-mile roadways, Express and Regular, are assumed to connect the same origin and destination and to have the same free-flow travel time of 8.0 minutes—the estimated travel time at 4:00 a.m., corresponding to a speed of 75 miles per hour. We model a four-hour peak period and find equilibria by iterating between the supply and demand sides of the model.

The supply side is a standard static congestion model in which travel delays are proportional to the fourth power of the volume-capacity ratio (US Bureau of Public Roads 1964). Capacity is 2,000 vehicles per hour per lane. Unreliability is assumed to be a constant fraction 0.3785 of travel delay (travel time minus free-flow travel time) — the fraction observed in our data on the free lanes averaged over 5-9 a.m.

Our demand function is obtained from the estimated demand model by sample enumeration, using the Brookings RP sample which, as noted, is random and mostly representative of the population. However, the estimated model is conditional on travel in this corridor. We want to include the possibility of individuals altering the decision to travel in the corridor in response to policies we simulate, because other studies have shown that the overall elasticity of corridor demand can strongly affect the relative benefits of alternative pricing strategies (Verhoef, Nijkamp, and Rietveld 1996). Therefore we extend our mixed logit model to accommodate the possibility of non-travel by adding an “outside choice” representing non-travel (on the corridor).

Let Ω={-1,0,1,…,8} denote the choice set for a potential road user, where alternative -1 is the outside choice and alternatives 0–8 represent the different combinations of routes, transponder acquisition, and car occupancy defined previously. It is convenient to let $\widetilde{\Omega}$={0,1,…,8} denote the subset of choices involving travel on the corridor.

The utility of individual *n* choosing alternative *j* is:

$$U_{-1n} = \bar{\delta}_{-1} + \eta_{-1n} \tag{11a}$$

$$U_{nj} = X_j^{RP} \beta_n^{RP} + \varepsilon_{jn}^{RP}, \ j \geq 0 \tag{11b}$$

with β_n^{RP} and ε_{jn}^{RP} as given by equations (2)-(3). Thus each traveler's utility for non-travel is divided into a mean $\bar{\delta}_{-1}$, which is constant for all commuters, and random deviation ε_{-1n}; whereas the utility for each alternative involving travel is the same as in the RP part of our estimated model (1)-(3). We henceforth omit the superscript *RP* for simplicity.

The random preferences for individual *n* are therefore represented by the vector $\Psi_n = (\eta_n, \nu_n, \mu_n)$, where $\eta_n = (\eta_{-1n}, \eta_{0n}, ..., \eta_{8n})$, $\nu_n = (\nu_n^T, \nu_n^X, \nu_n^{H2}, \nu_n^{H3})$, and $\mu_n = (\mu_n^{Time}, \mu_n^{\mathrm{Re}l})$. The density function of Ψ_n is specified as $\rho(\Psi_n) = \rho_\eta(\eta_n) \cdot \rho_{\nu\mu}(\nu_n, \mu_n)$; here $\rho_{\nu\mu}(\cdot)$ is a product of independent normals with standard deviations as estimated, while $\rho_n(\cdot)$ takes the nested-logit form in which the outside alternative -1 is one nest and the travel alternatives $\widetilde{\Omega}$ are another nest with similarity parameter λ. This specification captures the idea that the substitution pattern between any two travel choices may be different from that between non-travel and travel. The market share of alternative $j \in \widetilde{\Omega}$, within the sub-market represented by people with characteristics of traveler *n* in our enumeration sample, is found by integrating the nested-logit probability formula, conditional on random parameters ν_n and μ_n, over the distribution function of those random parameters:

$$S_{jn} = \int_{(\nu_n,\mu_n)} S_{jn}^{(\nu_n,\mu_n)} \cdot \rho_{\nu\mu}(\nu_n,\mu_n) d(\nu_n,\mu_n) \tag{12a}$$

where

$$S_{jn}^{(\nu_n,\mu_n)} = \frac{\exp(X_{jn}\beta_n/\lambda)}{\exp(I_n)} \cdot \frac{\exp(\lambda I_n)}{\exp(\bar{\delta}_{-1}) + \exp(\lambda I_n)} \tag{12b}$$

is the share conditional on values of the normal random variates, and

$$I_n = \ln \sum_{j=0}^{8} \exp(X_{jn}\beta_n/\lambda) \tag{12c}$$

is the inclusive value of travel choices. The non-travel share is

$$S_{-1n} = \frac{\exp(\bar{\delta}_{-1})}{\exp(\bar{\delta}_{-1}) + \exp(\lambda I_n)}. \tag{12d}$$

The total demand for an alternative j is therefore

$$D_j = \sum_n p_n S_{jn}, \tag{13}$$

where p_n is the number of people represented by motorist n. This number is just $p_n = N/79$, where N is the total population size, since our enumeration sample consists of 79 equally-weighted individuals. The traffic volume arising from those individuals who choose travel alternative $j \geq 0$ is $V_j \equiv D_j/O_j$, where O_j is the vehicle occupancy associated with alternative j.

Calibrating the Expanded Demand Model

To use our model for simulation, we need to calibrate three parameters: the alternative-specific constant for the outside choice ($\bar{\delta}_{-1}$), the coefficient of the inclusive value of travel (λ), and the population size (N). Because we expect the travel alternatives to be much closer substitutes for each other than for non-travel, we choose λ as small as possible without causing numerical instability: namely, λ=0.2. This choice does not seem to have much effect on the nature of the results. We calibrate the other two parameters ($\bar{\delta}_{-1}$ and N) to replicate observed traffic conditions during the morning peak on SR91 in the summer of 2000, which took place with an express-lane toll of $3.30 with 50% discount for HOV3. These conditions are: travel-time difference between the express and the free lanes of 3.4 minutes (according to our field measurements), and non-travel share of 10%.[10] The parameters that achieve these results are shown in the first column of numbers in Table 5.

We recognize that most areas considering new pricing or express-lane policies have far more congestion than was observed on SR91 in 2000 —five years after a 50 percent increase in SR91 capacity. Indeed, traffic on SR91 itself has increased greatly since 2000. Thus, we simulate a situation where travel speed is 30 miles per hour under a no-toll scenario, which on our corridor means a travel time of 20 minutes. The adjusted travel conditions require assuming a larger population N. Rather than assuming constancy in the rather abstract parameter $\bar{\delta}_{-1}$, we assume constancy in the total travel elasticity with respect to full cost under a no-toll situation, which can be computed from our initial calibration and also compared with other studies.[11] The resulting elasticity value (-0.36) shown in the middle column of Table 5 may be compared with the value of -0.58 estimated by Yan, Small, and Sullivan (2002) for the full-cost elasticity of morning trips under the actual pricing scheme in effect in 2000, based on the CalPoly data and

---

[10] The 10% figure is a rough estimate based on the likelihood that a small portion of trips were foregone due to congestion, that an alternative route available for some travelers (SR241) had about 8%-9% share of the CalPoly sample, and that the share of public transit is less than 1%.

[11] We computed the full-cost elasticity by using our expanded demand model, with initial calibrated parameters just described, to simulate changes in total travel under a no-toll scenario and under a scenario where travel time and reliability are both increased 10 percent for all alternatives.

using a model with no heterogeneity. We also computed the money-price-elasticity of express-lane travel, starting with the toll scheme of summer 2000 (first column of Table 5), obtaining a value of -1.12. The elasticity is somewhat higher than the values of -0.7 to -0.8 reported in Yan, Small, and Sullivan, which are based on a model that did not account for preference heterogeneity. We believe the elasticities calculated here are realistic because preference heterogeneity creates a subset of people who are quite ready to shift into or out of the express lanes in response to tolls, but not very likely to shift from travel to non-travel.

The calibration exercise leads to the parameters shown in the last row of the table for use in our policy simulations. We perform sensitivity tests using different values of the total traffic elasticity, including zero.

Simulation of Policy Scenarios

Based on our equilibrium model of supply and demand, we simulate results for several alternative highway pricing and operational policies. For each, we calculate tolls, travel times, traffic volumes, revenues, changes in consumers' surplus, the distributional effects of the changes by income group, and the total change in social welfare. In our base-case or "no-toll" policy, the two roadways are not distinguished and have no toll. The change in consumer surplus for traveler *n*, relative to the base case, is determined by the log-sum rule for nested logit (Choi and Moon 1997):

$$\Delta CS_n = \frac{1}{m_n} \Delta \ln\left[\exp\left(\bar{\delta}_{-1}\right) + \exp\left(\lambda I_n\right)\right], \tag{18}$$

where Δ indicates the difference between a given scenario and the base scenario, I_n is given by (12c), and m_n is the individual's marginal utility of income, determined from the coefficient of the toll variable using Roy's identity.[12] The change in social welfare is the sum of expected changes in all individuals' consumer surplus and in toll revenues.

Besides the base scenario ("No toll"), the scenarios that we consider are:

- a conventional carpool lane open to carpools of two or more people (HOV);
- a conventional HOT lane open to carpools and to paying solo drivers (HOT);
- an express lane open to paying customers only ("One-route toll");
- full-fledged differential pricing of both sets of lanes ("Two-route toll");
- differential pricing of both sets of lanes but with carpools free ("Two-route HOT").

Where applicable, prices are chosen to maximize welfare subject to the equilibrium constraints and to any applicable constraints on who can travel on each set of lanes.

The policies differ in the number of options they offer travelers, a difference that by itself cause welfare changes in a discrete choice model. We define the options, summarized in Table 6, to make those welfare changes correspond to actual changes that would be perceived by users. The No-toll scenario offers just three alternatives (solo, HOV2, or HOV3, all on the general lane with no transponder) because no role exists for a transponder and motorists have no reason to distinguish between the different lanes.[13] The HOT lane policy allows all of the nine alternatives

[12] We report all results on a per-trip basis, so Roy's identity equates m_n to minus the coefficient of the toll. Based on the results of Table 3, $m_n = -(-2.4042+1.3869*H_n)$, where H_n represents this traveler's value of the high-income dummy.

[13] Analytically, the condition that users be indifferent between the lane types, *ceteris paribus*, requires one other change besides setting the transponder preference to zero: we also calibrate the average express-lane preference in the coefficient of "Express lane dummy" to equalize the volume-capacity ratio between the two sets of lanes when both have equal travel times. We also computed results, not shown here, for a different no-toll scenario offering separate Express and General lanes, both free but differing is some perceptible way — as occurs for example with the free bypass lanes offered on certain expressways in the Chicago area. Under these travel conditions, the no-toll scenario has six alternatives.

that we used to estimate the demand model. The HOV and the two-route policies each foreclose three of the nine alternatives.

Under the no-toll and HOV scenarios, we set the transponder preference (negative on average in the model shown in Table 3) to zero. However, when we compute consumer surplus in scenarios that do entail a transponder choice, we account for the preferences for and against acquiring a transponder that are represented by the coefficients of "Transponder dummy" in our demand model and by the random variate ν_n^T in equation (3). We interpret these preferences as reflecting the nuisance or cost of acquiring a transponder and the benefits of day-to-day flexibility in lane choice.

Simulation Results

Simulation results are presented in Table 7. Generally, the change in consumer surplus is affected more by a policy's impact on the toll, travel time, and reliability of travel time than by the availability of alternatives or our adjustment of alternative-specific preferences.

Turning to specific policies, the introduction of HOV lanes improves efficiency by encouraging carpooling—more than doubling its share from the base case. This policy significantly reduces travel time on the express lanes, but leaves the general lane congested. In all likelihood, the policy would be much less effective in a smaller metropolitan area: Dahlgren (1998) finds that HOV lanes are favorable (in terms of reducing total person delay) only when initial congestion is substantial (delays of 35 minutes or more) and when the initial modal share of carpools is sizable (20 percent or more).

Allowing solo motorists to use the express lane if they pay a toll (HOT) generates a small welfare improvement over the HOV policy by enabling a small share of travelers to switch lanes and drive faster. In the HOT-like policy without an HOV exemption (one-route toll), the general lane becomes even more congested and the welfare improvement over the initial HOV policy is negligible.

Notwithstanding their sacrifice of economic efficiency, variants of HOT and HOV policies have attained a certain degree of public acceptance, suggesting that their favorable distributional features are compelling. The HOT and HOV policies impose little or no direct loss in consumer surplus to any observable group. One-route tolling, in contrast, creates a small consumer surplus loss for the median traveler, although like all the policies it creates a notable gain for high-income travelers because some of them have high values of time and reliability.

The first-best policy of tolling both lanes (Two-route toll) produces a sizable gain in welfare over HOV and HOT policies, largely because it greatly reduces congestion on both lanes. However, it causes travelers to suffer high and disparate losses in consumer surplus, amounting to $2.36 per trip on average and over $5.36 per trip for one-fourth of all travelers. Furthermore, the largest losses are associated with the lowest income groups, who tend to have the lowest values of time and reliability. These features make efficient pricing unattractive politically.

However, we find that policymakers can achieve most of the gains from first-best pricing, while partly addressing distributional concerns, by adding a carpool exemption to the two-route pricing policy ("Two-route HOT").[14] The carpooling share, already high in two-route pricing because of its financial incentives, increases even more, while congestion on both routes is

[14] This policy, like others involving an express lane, incorporates the absolute preferences for the express lane from our demand model, which on average are slightly positive. For this reason, traffic is equilibrated when the express lane is slightly slower than the general lane, even though the express lane has a (moderately) higher price. While this may appear anomalous, it reflects other advantages of the express lanes on SR91, such as lack of trucks and intermediate entrances and exits, which we think would apply in many other express-lane applications so we have allowed the feature to appear.

comparable to the levels under two-route pricing. Remarkably, travelers on average obtain a substantial welfare gain from Two-route HOT; nevertheless, the policy is vulnerable to the concern that a substantial fraction of people incur sizable losses — those in the most disadvantaged quartile of users lose $1.91 per trip or more.

The preceding analysis has shown that by designing road pricing policies that account for motorists' varying preferences, it is possible to identify a range of potentially attractive policies beyond simple HOV and HOT. We show in the appendix that this basic finding does not change if we assume different values for the elasticity of total travel.

It is important to point out that if we assumed that travelers were homogeneous, our findings would change considerably, along the same lines as in Small and Yan (2001) and Verhoef and Small (2004). Given homogeneous preferences, the relative welfare gain from the one-route toll, whose efficiency relies mainly on creating differential services, would drop significantly; HOT would become nearly identical in effect to HOV (because no additional benefit results from separating users based on their preferences); the one-route toll is set much lower (because it cannot rely on attracting users just from the upper tail of the VOT and VOR distributions); and the toll differences under the two-route toll and the two-route HOT are smaller, for the same reason. The differences reflect an underlying feature of heterogeneity: diversity in users' preferences creates options to improve social welfare by providing differentiated services instead of by encouraging carpools and reducing traffic. Thus, by accounting for heterogeneity we have given a policymaker more flexibility.

This observation raises the question of whether it is possible to craft a policy that achieves even a better compromise between efficiency and political feasibility than the policies that we have explored. In particular, we seek a more efficient policy with the same attractive distributional features as the “One-route toll”. We choose the “One-route toll” as a benchmark for political feasibility because there are at least two cases in North America where a policy resembling it has been successfully implemented. One case is the SR91 itself: although carpools did not all pay full fare on SR91, those with only two occupants did and those with three or more occupants paid half the fare during most of the time when the original private toll road was in effect.[15] The other case is Highway 407 in the Toronto area. This is a publicly built highway through the suburbs that parallels, a few miles away, a very congested east-west route through the city known as Queen Elizabeth Way. Projects such as these, and several others being actively considered, suggest that the public is willing to tolerate a toll road without HOV exceptions if a free alternative is available. In our simulation of the One-route toll, the free alternative exists in the form of a roadway immediately adjacent to the priced one and so should be at least as acceptable as Highway 407.

We therefore quantify a benchmark for political viability as the 25th percentile of consumer-surplus change experienced by travelers under the “One-route toll” policy (-\$0.98 per trip). We then search for toll values in an alternate version of two-route pricing that maximizes welfare subject to meeting this distributional benchmark. The result is the “Limited two-route HOT” policy shown in the last column of Table 7. It uses a sharply differentiated toll that for the express lane is of a magnitude (\$9.65) comparable to that in the Two-route toll policy, but for the general lane is much lower (\$1.90) than either of the other two-route pricing policies. The policy achieves a general-lane speed (36 mi/hr) intermediate between those of the no-toll and the two-

[15] Although there were complaints about the high tolls and about charging HOV3+ vehicles, they did not appear to have undermined the stability of the arrangement, which had a strikingly high acceptance level in various polls. What did eventually undermine the toll levels and schedule was an unrelated issue: the franchise allowed the private operator to veto any capacity improvements in the corridor, which it did through a lawsuit in a highly publicized dispute with the California Department of Transportation.

route policies. It also achieves greater efficiency than all policies that leave the general lanes unpriced—bear in mind that the relatively small $0.16 gain per trip over the simple HOT policy applies to millions of trips annually.

From a distributional perspective, motorists on average achieve a consumer surplus *gain* of $1.36 compared with no tolls, exceeding the gain achieved by any other two-route pricing policy. Furthermore, travelers are treated much more evenly than in the other two-route pricing policies, with the inter-quartile range only modestly greater than with HOT or HOV. Thus, the Limited two-route HOT policy succeeds in improving efficiency more than most other policies, while maintaining the attractive distributional characteristics of our benchmark for political feasibility.

We stress that our policy simulations are based on an experiment concerning only a single ten-mile stretch of highway. Most significant congestion affects a much broader region. If the distributional advantages of differentiated pricing enable it to be broadly adopted, its welfare gains will be greatly magnified.

Conclusion

Methodological advances in microeconometrics have enriched our understanding of consumer behavior by recognizing that consumers are not homogeneous. Applications have shown that accounting for heterogeneity is important when assessing such diverse policies as economic deregulation, job training, and poverty programs.

Heterogeneity also plays an important role in policy toward highway transportation. In particular, accounting for heterogeneity creates the opportunity not only to introduce HOT lanes, as has been previously recognized, but to introduce more far-reaching pricing policies within the limits of distributional effects that appear to be politically acceptable in certain circumstances. We have been able to design a differentiated road-pricing scheme that fills in the gap between optimal but socially unpopular first-best pricing and pragmatic but less efficient policies like carpool or HOT lanes. Recent experiments have shown that policymakers are no longer unwilling to use the price mechanism to allocate scarce road capacity. The changing times give cause for optimism that more efficient policies, offering choices that appeal to diverse users, will become serious candidates for implementation.

Table 1. Descriptive Statistics

	Value or Fraction of Sample		
	Cal Poly-RP	Brookings-RP	Brookings-SP
Age (years):			
<30	0.11	0.12	0.10
30-50	0.62	0.62	0.64
Household Income ($/year):			
<60,000	0.38	0.83	0.83
>100,000	0.22	0.02	0.04
Female Dummy	0.32	0.37	0.37
Mean Actual Trip Distance (Miles)	34.2	44.8	42.6
Number of Respondents	435	79	78
Number of Observations	435	369	610

Table 2. Choice Shares of Calpoly and Brookings RP Samples

Alter-native	Calpoly Sample				Brookings RP sample (%)
	Random (%)	New Plates (%)	Repeat (%)	UCI (%)	
*NF*1	41	28	17	11	51
*TF*1	16	26	33	39	23
*TX*1	19	15	16	22	20
*NF*2	7	9	3	0	0
*TF*2	3	8	16	6	2.5
*TX*2	8	5	7	22	1
*NF*3	3	3	0	0	0
*TF*3	1	3	3	0	0
*TX*3	2	3	5	0	2.5
All carpool	24	31	34	28	6
No. of Obs.	201	191	58	18	79

Legend:
Transponder acquisition: N=No, T=yes
Lane: F=Free lane, X=Express lane.
Car occupancy: 1=solo, 2=HOV2, 3=HOV3+

Table 3 Estimation Results

Variable	Coefficients (standard error)[a]
RP Estimates	
Generic variables:	
Toll ($)[b, c]	-2.4042 (0.3994)
Toll[b, c] × dummy for high household income (> $60K)	1.3869 (0.3395)
Median travel time (min.)[b] × trip distance (units of 10 miles)	-0.5753 (0.1751)
Median travel time[b] × trip distance squared	0.1128 (0.0394)
Median travel time[b] × trip distance cubed	-0.0050 (0.0020)
Travel-time uncertainty (80%-ile minus the median) (min.)[b]	-0.7489 (0.2668)
Transponder choice:	
Transponder dummy× Brookings dummy	-2.0101 (0.7472)
Transponder dummy × Calpoly dummy	-3.6342 (0.7374)
Female dummy × age30-50 dummy × transponder dummy	1.8535 (0.7979)
Commute dummy × transponder dummy	1.2502 (0.6967)
Standard deviation of transponder dummy (σ^{T})	0.3276 (0.9422)
Lane choice:	
Express lane dummy × Brookings dummy	0.2564 (1.1386)
Express lane dummy × Calpoly dummy	0.2264 (1.1691)
Standard deviation of express lane dummy ($\sigma^{X\text{-}RP}$)	3.7879 (0.8261)
Carpool choice:	
Carpool dummy × Brookings dummy	-11.5192 (1.0339)
Carpool dummy × Calpoly dummy	-11.6719 (0.8883)
Female × age30-50 × household size × carpool	1.4404 (0.3563)
HOV3 dummy × Brookings dummy	-9.2262 (0.9886)
HOV3 dummy × Calpoly dummy	-7.4263 (0.9909)
Common standard deviation of HOV2, HOV3 dummies (σ^{HOV})	10.3225 (0.7837)
SP Estimates	
Express lane dummy	-3.0651 (1.1953)
Standard deviation of express lane dummy ($\sigma^{X\text{-}SP}$)	1.0530 (0.5237)
Toll ($)[b, c]	-1.4165 (0.3028)
Toll[b, c] × dummy for high household income (> $60K)	-0.2492 (0.4808)
Travel time (min.)[b] × long commute dummy (> 45 minutes)	-0.3538 (0.0812)
Travel time[b] × (1 - long commute dummy)	-0.3843 (0.0616)
Travel-time uncertainty[d]	-7.1139 (1.4507)

(Table 3 – continued)

Combined Estimates

Female dummy × express lane dummy	2.2434 (0.8384)
Age30-50 dummy × express lane dummy	1.9277 (0.7955)
Household size (number of people) × express lane dummy	-0.7371 (0.2117)
Standard deviation of travel-time coefficient (σ^{Time})	0.3866 (0.0694)
Ratio of std. dev. to mean of coefficient of travel-time uncertainty ($\sigma^{Rel}/\gamma^{Rel}$)	1.3233 (0.3805)
Correlation parameter between RP and SP express lane choice (θ)	1.4808 (0.3209)
Parameters Associated with Scaling	
Scale parameter: Calpoly sample (τ^{C})	0.4143 (0.0902)
Scale parameter: Brookings RP sample (τ^{BR})	0.6064 (0.2029)
Number of observations	1124
Number of persons	538
Log-Likelihood	1059.63

[a] Standard errors reported are the "sandwich" estimate of standard errors from Lee (1995). That is, each is the square root of the corresponding diagonal element in the matrix $\hat{V} = (-H)^{-1} P (-H)^{-1}$, where H is the Hessian of the simulated log-likelihood function and P is the outer product of its gradient vector (both calculated numerically). This estimate accounts for the simulation error in the likelihood function.

[b] All cost, travel-time, and unreliability variables are entered as the difference between values for toll and 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. For SP, car occupancy is determined from a question asking whether the respondent answered as a solo driver or as part of a carpool, and if the latter what size carpool.

Table 4. Implied Values of Time and Reliability for the Brookings Sample

	Median Estimate	90% Confidence Interval [5%-ile, 95%-ile]
Value of time ($/hour)		
Median for:		
Entire sample	19.63	[8.75, 34.61]
Express lane users	25.51	[11.50, 39.99]
Free lane users	18.63	[7.76, 29.08]
Total heterogeneity for:		
Entire sample	19.02	[12.57, 30.96]
Express lane users	29.30	[14.65, 55.97]
Free lane users	17.73	[11.37, 28.05]
Value of reliability ($/hour)		
Median for:		
Entire sample	20.76	[8.37, 40.71]
Express lane users	23.78	[10.01, 48.29]
Free lane users	19.50	[5.73, 34.54]
Total heterogeneity for:		
Entire sample	35.51	[14.95, 66.71]
Express lane users	44.70	[18.27, 84.24]
Free lane users	32.95	[13.70, 62.01]

Table 5. Initial Calibration

	Summer 2000 conditions		More congested conditions
	HOT-lane scenario	No-toll scenario	No-toll scenario
Assumed toll:[a]			
Express lanes	\$3.30	\$0	\$0
Free lanes	\$0	\$0	\$0
Calibrated parameters:			
N	17,570	17,570	24,710
$\bar{\delta}_{-1}$	-12.65	-12.65	-23.41
Travel time (min.):			
Express lanes	9.83	12.03	20
Free lanes	13.23	12.03	20
Arc elasticities (based on 10% increase):			
Total corridor traffic volume, with respect to full cost	-0.40	-0.36	-0.36
Express-lane traffic volume, with respect to express toll	-1.12	N.A.	N.A.

[a]HOV3+ pays half of the toll

Table 6. Availability of Each Alternative By Scenario

Alternative				Scenario			
Number	Description			No Toll	HOV	HOT, One-route toll	Two-route toll, Two-route HOT
	Mode	Transponder?	Lane				
0	Solo	N	G	x	x	x	
1	Solo	T	G			x	x
2	Solo	T	X			x	x
3	HOV2	N	G	x	x	x	
4	HOV2	T	G			x	x
5	HOV2	T	X		x	x	x
6	HOV3	N	G	x	x	x	
7	HOV3	T	G			x	x
8	HOV3	T	X		x	x	x

Note: x means that the alternative is available in that scenario. G indicates the general (non-express) lanes; these are the same as the "free lanes" (F) in the HOT scenario from which the demand model was estimated and which are defined in the first column of Table 2.

Table 7. Simulation Results

	No Toll	HOV	HOT	One-route toll	Two-route toll	Two-route HOT	Limited two-route HOT
Toll on Express Lane	$0	$0	$9.23	$8.69	$10.14	$6.33	$9.65
Toll on General Lane	$0	$0	$0	$0	$8.16	$5.34	$1.90
Travel times (min.):							
Express Lane	20.00	11.75	12.35	11.56	11.62	13.12	12.51
General Lane	20.00	18.78	19.17	22.63	12.76	12.48	16.51
Aggregated choice shares (%):							
No travel on corridor	7.44	3.27	3.37	6.44	8.48	2.96	3.33
Solo on Express Lane	24.79	0	2.57	8.94	7.96	7.57	1.71
Solo on General Lane	49.58	51.99	52.55	54.47	27.86	24.98	46.34
HOV2 on Express Lane	5.13	32.56	30.00	15.35	16.33	23.10	31.68
HOV2 on General Lane	10.25	2.92	3.00	7.24	22.08	26.84	6.59
HOV3+ on Express Lane	0.94	8.80	8.11	6.70	8.52	6.89	8.91
HOV3+ on General Lane	1.88	0.56	0.40	0.86	8.77	7.65	1.43
All HOV3+	2.82	9.36	8.51	7.56	17.29	14.54	10.34
All HOV	18.20	44.84	41.51	30.15	55.70	64.48	48.61
Consumer surplus change ($/person)[a]:	0	2.11	2.01	0.50	-2.36	0.98	1.36
Classification by source							
Adding more alternatives	0	0.02	0.18	0.18	0.11	0.11	0.11
Transponder & lane preferences	0	0.27	0.80	0.80	0.68	0.68	0.68
Change in toll, time, reliability	0	1.82	1.03	-0.48	-3.15	0.19	0.57
Distribution in population							
75%-ile	0	2.92	2.71	0.65	0.00	3.51	2.80
50%-ile	0	0.77	0.62	-0.27	-2.68	0.52	0.33
25%-ile	0	0.26	0.26	-0.98	-5.36	-1.91	-0.98
Distribution by income group							
High income (≥ 60K)							
75%-ile	0	6.47	6.80	6.01	5.16	8.30	6.88
50%-ile	0	1.68	1.61	1.27	0.92	4.48	2.47
25%-ile	0	0.72	0.67	-0.92	-3.34	0.15	0.04
Low income (< 60K)							
75%-ile	0	2.50	2.10	0.29	-0.55	2.47	2.00
50%-ile	0	0.64	0.51	-0.37	-3.20	0.11	0.08
25%-ile	0	0.22	0.22	-0.98	-5.60	-2.19	-1.04
Toll revenue ($/person)	$0	$0	$0.24	$1.64	$5.35	$1.81	$1.05
Welfare change ($/person)[a]	$0	$2.11	$2.25	$2.14	$2.99	$2.79	$2.41

[a] Consumer surplus and social welfare are measured relative to the no-toll scenario and are divided by the number of users *in the no-toll scenario* in order to put them in per-vehicle terms, including all people who would use the corridor under the no-toll policy. Social welfare is the sum of consumer surplus plus revenue.

Appendix

Stated Preference Survey Questionnaire

Eight hypothetical commuting scenarios were constructed for respondents who travel on SR91. Respondents who indicated 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:

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 ☐

Sensitivity Analysis for Simulation Results

The simulation results presented in the text were based on a total traffic full cost elasticity of -0.36. Tables A.1 and A.2 present simulation results where we assume the full-cost elasticity of total corridor traffic is -0.60 and zero, respectively. For the case of zero elasticity, the model has no outside option and so there is no parameter $\bar{\delta}_{-1}$ to calibrate. In the other case, the two parameters N and $\bar{\delta}_{-1}$ are simultaneously calibrated, as in the main results, to achieve the desired elasticity and travel-time differential with no toll. We caution that numerical results are not necessarily comparable with different assumed elasticities because they imply different starting shares for HOV. Our main purpose for presenting them is to show that the welfare rankings of various policies, and the nature of their distributional impacts, do not depend on this assumed elasticity.

Table A.1 Simulation Results (elasticity=-0.60)

	No Toll	HOV	HOT	One-route toll	Two-route toll	Two-route HOT
Toll on Express Lane	$0	$0	$8.41	$8.53	$9.41	$6.02
Toll on General Lane	$0	$0	$0	$0	$7.01	$5.32
Travel times (min.):						
Express Lane	20.00	12.35	13.01	11.29	11.41	13.84
General Lane	20.00	19.82	20.02	22.88	13.00	12.79
Aggregated choice shares (%):						
Outside choice	16.50	9.79	10.79	15.86	19.69	9.80
Solo on Express Lane	22.18	0	2.83	7.38	7.19	7.86
Solo on General Lane	44.35	47.94	47.99	48.93	26.67	21.85
HOV2 on Express Lane	4.79	30.79	27.94	14.10	14.11	20.44
HOV2 on General Lane	9.59	2.71	2.99	6.92	18.80	26.71
HOV3 on Express Lane	0.86	8.45	6.91	5.95	7.21	5.99
HOV3 on General Lane	1.73	0.43	0.55	0.86	6.33	7.34
All HOV3	2.59	8.88	7.46	6.81	13.54	13.33
All HOV	16.97	42.38	38.39	27.83	46.45	60.48
Consumer surplus change ($/person relative to No Toll):[a]	0	1.50	1.47	0.46	1.77	0.49
Classification by source						
Adding more alternatives	0	0.02	0.16	0.16	0.10	0.10
Transponder & lane preferences	0	0.25	0.74	0.74	0.64	0.64
Change in toll, time, reliability	0	1.23	0.57	-0.44	-2.51	-0.25
Distribution in population						
75%-ile	0	2.04	1.94	0.56	0.03	2.73
50%-ile	0	0.10	0.03	-0.14	-1.76	0.08
25%-ile	0	0.03	0.01	-0.87	-4.61	-1.94
Distribution by income group						
High income (≥ 60K)						
75%-ile	0	5.04	5.04	5.16	4.85	6.67
50%-ile	0	0.23	1.08	1.02	1.47	3.50
25%-ile	0	0.08	0.02	-0.65	-2.28	0.18
Low income (< 60K)						
75%-ile	0	1.75	1.42	0.20	-0.02	1.86
50%-ile	0	0.09	0.02	-0.25	-2.26	0.00
25%-ile	0	0.03	0.01	-0.89	-4.83	-2.28
Toll revenue ($/person)	$0	$0	$0.24	$1.40	$4.24	$1.63
Welfare change ($/person)[a]	$0	$1.50	$1.71	$1.86	$2.47	$2.13

[a] See Table 7.

Table A.2 Simulation Results (elasticity=0: no outside choice)

	No Toll	HOV	HOT	One-route toll	Two-route toll	Two-route HOT
Toll on Express Lane	$0	$0	$8.47	$8.64	$10.46	$6.17
Toll on General Lane	$0	0	$0	$0	$9.29	$5.19
Travel times (min.):						
Express Lane	20.00	10.95	11.65	11.57	11.83	12.32
General Lane	20.00	18.09	18.28	22.14	12.39	12.28
Aggregated choice shares (%):						
Outside choice	0	0	0	0	0	0
Solo on Express Lane	26.60	0	3.49	10.39	9.68	8.75
Solo on General Lane	53.19	54.32	54.48	57.63	26.64	26.65
HOV2 on Express Lane	5.65	31.68	28.74	15.23	16.42	21.66
HOV2 on General Lane	11.30	4.24	4.49	8.50	26.39	28.07
HOV3 on Express Lane	1.09	8.92	7.93	6.95	8.94	6.63
HOV3 on General Lane	2.17	0.83	0.87	1.30	11.93	8.23
All HOV3	3.26	9.75	8.80	8.25	20.57	14.86
All HOV	20.21	45.67	42.03	31.98	63.38	64.60
Consumer surplus change ($/person relative to No Toll):[a]	0	2.94	3.09	1.17	-2.39	1.87
Classification by source						
Adding more alternatives	0	0.13	0.94	0.94	0.60	0.60
Transponder & lane preferences	0	0.24	0.48	0.48	0.42	0.42
Change in toll, time, reliability	0	2.57	1.67	-0.25	-3.41	0.85
Distribution in population						
75%-ile	0	4.07	4.00	1.15	0.12	4.45
50%-ile	0	1.47	1.64	0.14	-3.05	1.29
25%-ile	0	0.56	0.91	-0.44	-5.78	-1.48
Distribution by income group						
High income (≥ 60K)						
75%-ile	0	8.35	8.68	6.70	6.04	9.81
50%-ile	0	3.06	4.02	1.79	1.49	5.30
25%-ile	0	1.42	2.23	0.13	-3.10	0.93
Low income (< 60K)						
75%-ile	0	3.38	3.04	0.71	-0.94	3.46
50%-ile	0	1.22	1.40	0.01	-3.60	0.77
25%-ile	0	0.49	0.82	-0.50	-6.08	-1.73
Toll revenue ($/person)	$0	$0	$0.30	$1.76	$6.25	$1.92
Welfare change ($/person)[a]	$0	$2.94	$3.39	$2.93	$3.86	$3.79

[a] See Table 7.

References

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.

Choi, K.-H., and C.-G. Moon (1997), "Generalized Extreme Value Model and Additively Separable Generator Function," *Journal of Econometrics*, 76, pp. 129-140.

Dahlgren, J. (1998), "High Occupancy Vehicle Lanes: Not Always More Effective than General Purpose Lanes," *Transportation Research A*, 32, pp. 99-114.

De-Palma, A., and R. Lindsey (2004), "Congestion Pricing with Heterogeneous Travelers: A General-Equilibrium Welfare Analysis," *Networks and Spatial Economics*, 4, pp. 135-160.

Jiang, Meilan, and Takayuki Morikawa (2004), "Theoretical Analysis on the Variation of Value of Travel Time Savings," *Transportation Research A*, 38, pp. 551-571.

Hensher, David A. (2001), "The Valuation of Commuter Travel Time Savings for Car Drivers: Evaluating Alternative Model Specifications," *Transportation*, 28, pp. 101-118.

Hess, Stephane, Michel Bierlaire, and John W. Polak (2005), "Estimation of Value of Travel-Time Savings Using Mixed Logit Models," *Transportation Research A*, 39, pp. 221-236.

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.

Lee, L. (1995), "Asymptotic Bias in Simulated Maximum Likelihood Estimation of Discrete Choice Models," *Econometric Theory*, 11, pp. 437-483.

Manski, C. F., and S. R. Lerman (1977), "The Estimation of Choice Probabilities from Choice Based Samples," *Econometrica*, 45, pp. 1977-1988.

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

Mohring, H. (1999), "Congestion," in J. Gomez-Ibanez, W. Tye, and C.Winston, eds., *Essays in Transportation Economics and Policy: A Handbook in Honor of John R. Meyer*, Brookings Institution Press, Washington, DC.

Orski, C. K. (2001), "Carpool Lanes—An Idea Whose Time Has Come and Gone," *TR News*, 214, May-June, pp. 24-28.

Poole, R. W., Jr., and T. Balaker (2005), *Virtual Exclusive Busways: Improving Urban Transit While Relieving Congestion*, Reason Foundation Policy Study 337.

Santos, G., editor (2004), *Road Pricing: Theory and Evidence*, Elsevier Press, North Holland.

Schrank, David, and Tim Lomax (2005), *The 2005 Urban Mobility Report*, Texas Transportation Institute, Texas A&M University System, College Station, Texas. http://mobility.tamu.edu/ums/report.

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

Small, K.A., C. Winston, and C. Evans (1989), *Road Work: A New Highway Pricing and Investment Policy*, Brookings Institution, 1989.

Small, K.A., C. Winston, and J. Yan (2005a), "Uncovering the Distribution of Motorists' Preferences for Travel Time and Reliability," *Econometrica*, 73, pp. 1367-1382.

Small, K.A., C. Winston, and J. Yan (2005b), Supplement to "Uncovering the Distribution of Motorists' Preferences for Travel Time and Reliability," *Econometrica Supplementary Material*, 73, http://www.econometricsociety.org/suppmatlist.asp.

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.

Steimetz, S.S.C., and D. Brownstone (2005), "Estimating Commuters' 'Value of Time' with Noisy Data: A Multiple Imputation Approach," *Transportation Research B*, 39, pp. 865-889.

Sullivan, E., with K. Blakely, J. Daly, J. Gilpin, K. Mastako, K. Small, and J. Yan (2000), *Continuation Study to Evaluate the Impacts of the SR 91 Value-Priced Express Lanes: Final Report*. Dept. of Civil & Environmental Engineering, Calif. Polytechnic State University at San Luis Obispo, Dec. (http://ceenve.ceng.calpoly.edu/sullivan/SR91).

US Bureau of Public Roads (1964), *Traffic Assignment Manual*, US Bureau of Public Roads, Washington, D.C.

Verhoef, E.T., P. Nijkamp, and P. Rietveld (1996), "Second-best congestion pricing: the case of an untolled alternative," *Journal of Urban Economics*, 40, pp. 279-302.

Verhoef, E.T., and K.A. Small (2004), "Product Differentiation on Roads: Constrained Congestion Pricing with Heterogeneous Users," *Journal of Transport Economics and Policy*, 38, pp. 127-156.

Yan, J., K. Small, and E. Sullivan (2002), "Choice Models of Route, Occupancy, and Time-of-Day with Value Priced Tolls," *Transportation Research Record*, 1812, pp. 69-77.