Title
Valuing Time and Reliability: Commuters' Mode Choice from a Real Time Congestion Pricing Experiment

Author
Ghosh, Arindam

Publication Date
2001

UNIVERSITY OF CALIFORNIA,

IRVINE

VALUING TIME AND RELIABILITY: COMMUTERS' MODE CHOICE FROM A REAL TIME CONGESTION PRICING EXPERIMENT

DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

In Economics

By

Arindam Ghosh

Dissertation Committee
Professor David Brownstone, Chair
Professor Kenneth A. Small
Professor Duncan Black

2001

The dissertation of Arindam Ghosh is approved and is acceptable in quality and form for publication on microfilm:

Committee Chair

University of California, Irvine
2001

DEDICATION

To

My Parents

For their support and unwavering faith and for their sacrifices which made this dissertation a reality

My Wife

For everything

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS

Greek Symbol	
α	Alpha
β	Beta
γ	Gamma
δ	Delta
ε	Epsilon
ϕ	Phi
η	Eta
φ	Psi
λ	Lambda
π	Pi
θ	Theta
ρ	Rho
σ	Sigma
Φ	Theta (Capital)

ACKNOWLEDGEMENTS

I would like to express my deepest appreciation to my committee chair, Professor David Brownstone, from whom I learned the art of disaggregate modeling. It was from him that I learned how to work with data and create meaningful models and I am also thankful for all those “cute little tricks” with data that made discrete choice modeling fun. Without his guidance, persistent help, constant encouragement and advice this dissertation would not have been possible.

I would also like to thank Professor Kenneth Small for teaching me transportation economics. More importantly he taught me the value of dedication and the determination not to be satisfied with anything less than perfect. Professor Pete Fielding helped me in integrating my research to a broader context and helped me improve my writing style. Professor Tom Golob showed me how to be a good researcher and showed me how to look between the data.

Thanks also go to the members of the San Diego I-15 Congestion Pricing Project evaluation team: Janusz Supernak and Christine Kaschade of San Diego State University, Jackie Golob of Jacqueline Golob Associates, Dirk van Amelsfort of Bureau Goudappel Coffeng, the Netherlands (formerly of Delft University) and Kim Kawada of San Diego Association of Governments (SANDAG). Kathy Happersett of the Social Science Research Laboratory of San Diego State University led the team that collected the panel survey data. This research was partly financially supported by the University of California Transportation Center (UCTC) through their dissertation fellowship. I would also like to thank the University of California, Irvine for providing the Regents Dissertation Fellowship and Professor David Brownstone and Professor Tom Golob for supporting a large part of the research from their grant from the project "Consumer Response to Congestion Pricing on High-Occupancy/ Toll Lanes" sponsored by the University of California Transportation Center

In the course of my graduate studies, I have been very fortunate in having wonderful classmates who over the years have become very good friends. But for them, the journey would have been much more arduous. My heartfelt appreciation also goes to all the staff in the Economics department and the Social Science graduate office, for easing a lot of the official red tape. I would specially like to thank Kathy Alberti, Ivonne Maldonado, Alicia Sanchez and Iris Adams for making me feel at home in the US and for providing constant support and encouragement.

CURRICULUM VITAE

Arindam Ghosh

Citizenship	Indian	**Visa**	F-1

Fields Of Specialization

Applied Microeconomics, Applied Econometrics, Transportation, Industrial Organization

Education

- University of California, Irvine (1997 to 2001) Ph.D. in Economics
  Dissertation Title: Value of Time and Mode Choice Under Road Congestion Pricing
- Delhi School Of Economics, Delhi, India (1994 to 1996) M. A. in Economics
- Jadavpur University , Calcutta, India (1991 to 1994) B. A. (Hons) in Economics

Awards

A. Kimball Romney Award for Outstanding Graduate Paper (2001) , School of Social Sciences, University of California, Irvine.

Regents' Dissertation Fellowship (Fall 2000) University of California, Irvine.

Dissertation Fellowship (2000-2001) University of California Transportation Center , University of California, Berkeley

Regents' Fellowship (1997-1998) University Of California, Irvine.

Summer Fellowship (1998) Institute of Mathematical and Behavioral Sciences, University of California, Irvine.

Working Papers

Heterogeneity in Value-of-Time: Revealed and Stated Preference Estimates from the I-15 Congestion Pricing Project (2000)

To Pay Or Not To Pay: Commuters' Mode Choice Decisions Under Real Time Congestion Pricing
Presented At the 1st IGERT Graduate Student Research Conference, April (2000)
Presented at the 80th Annual TRB Meeting, January (2001)

The Size Distribution of International Box Office Revenues and the

Stable Paretian Hypothesis (2000)

Willingness-to-Pay to Reduce Commute Time and Its Variance: Evidence from the San Diego I-15 Congestion Pricing Project.
Joint with Camilla Kazimi, David Brownstone, Thomas. F. Golob and Dirk van Amelsfort
Presented to the 79th Annual Meeting, Transportation Research Board. On Preprint CD-ROM. Washington, D.C.: Transportation Research Board (Jan. 2000)

Teaching and Research Experience

Research assistant for the project **"Consumer Response to Congestion Pricing on High-Occupancy/ Toll Lanes"** conducted by Dr. David Brownstone & Thomas Golob from 1998 to 2000. Project sponsored by the University of California Transportation Center. Was responsible for organizing the survey data, estimating econometric models and developing a code in GAUSS to estimate an efficient Method of Moments Estimator for discrete choice models under choice based sampling.

Served as a teaching assistant for several intermediate and advanced undergraduate Economics courses. (Basic Economics, Financial Engineering and Econometrics). Led discussion sections and lab sessions. Conducted review sessions.

Research Associate at **Center for Development Economics, Delhi School of Economics, India** from August 96 to June 97. Was responsible for monitoring and analyzing different development indicators for India. Helped create and manage a database of leading indicators of the Indian stock market

Current Research

My current research analyzes the mode and scheduling choice for commuters on I-15 of San Diego, California. In particular, I study the value-of-time for this corridor from revealed and stated preference estimates and examine its policy implications.

Work-in-Progress

Transit Maintenance: A Survey

Principal-Agent Relationship with an Informed Principal: The Managerial Contract Model

Corruption: A review of Issues

Language Proficiency

English, Hindi, Bengali (Read, Write and Speak fluently).

Computer Skills

Stata, Gauss, EVIEWS, LIMDEP, Shazam, Matlab, Excel and other software packages.

Professional Affiliation American Economic Association

References

Prof. David Brownstone Professor of Economics	(949) 824 – 6231 dbrownst@uci.edu	3151 Social Science Plaza University of California, Irvine CA 92697-5100 Fax: (949)824-2182
Prof. Kenneth Small Professor Of Economics,	(949)824-5658, 7642 ksmall@uci.edu	
Prof. G.J. (Pete) Fielding Professor Emeritus	(949) 824-5448 gjfieldi@uci.edu	

ABSTRACT OF THE DISSERTATION

Valuing Time and Reliability: Commuters' Mode Choice from a Real Time Congestion Pricing Experiment

By Arindam Ghosh

Doctor of Philosophy in Economics

University of California, Irvine, 2001

Prof. Davis Brownstone, Chair

The valuation of travel time savings has been an important theme in transportation research because it is the single largest contributor to the benefits of many transportation projects. It also plays a central role in deciding about the size and scope of public investment and has important welfare implications. It can shed important light as to whether any congestion pricing scheme will have increase social welfare or not. And help us understand how commuters make their travel decisions. The San Diego I-15 Congestion Pricing Project (SDCPP) is a demonstration project where an existing High Occupancy Vehicle lane was converted to HOT (High Occupancy/Toll) lane. Beginning in 1996 these lanes were made available to solo drivers who pay a toll. The toll adjusts every six minutes to maintain free flowing traffic on the HOT lane. Carpoolers get to use the lane for free as before. This presents us with a unique opportunity to study commuters' choice between a tolled and a free alternative based on not only

what the commuters say they would do (SP), but also on what they actually did (RP).

The general result is that this tolled facility is used by high income, middle aged, homeowners, female commuters. An interesting result that comes out of this analysis is the dual effects of toll. If the actual toll rises above the mean toll then the commuter is more likely to take the FasTrak lane. Another interesting implication that the effect of toll is conditional on the level of uncertainty of travel time and conversely uncertainty in travel time encourages use of FasTrak lane only if toll rises above a threshold value. Commuters are more sensitive to variations in travel time in the morning, specially during the peak, than in the afternoon.

Another salient result is that the Value of Time estimates from Stated Preference models are significantly lower than the Revealed Preference estimates. The difference is consistent and persistent across the different models which lead to the conclusion that these differences are real. Probably it reflects the difference in responses of individuals to actual and hypothetical situations.

INTRODUCTION

The valuation of travel time savings has been an important theme in transportation research because it is the single largest contributor to the benefits of many transportation projects. Mackie et al (2001) estimates that for the U.K., time savings consists of roughly 80% of all benefits for most road improvement projects. The magnitude of benefits is very similar for the U.S. Thus the monetary valuation of time savings or Value of Time (VOT) is critical to the appraisal of a project if it is based on a cost-benefit analysis. By the same token VOT also plays a central role in deciding about the size and scope of public investment. Apart from its importance in investment appraisal and policy analysis, VOT has important welfare implications. It can shed important light on whether any congestion pricing scheme will increase social welfare or not. Since VOT is generally derived from a disagregate model of travel behavior, the model itself provides important insights into travel decisions. Understanding how commuters make their decisions and respond to changes in different travel attributes help us in formulating policies that are more effective.

Though VOT has been extensively studied over the years, there is surprisingly little consensus about its actual value. There has been a wide range of estimates and several surveys of the literature have put the valuation at 50% of wage rate (Small (1992), Walters (1992)). Most of the studies reviewed were based on revealed preference (RP), or on the actual choices made by the commuter. Typically these studies compare car driving against some form of public transit without controlling for factors like convenience and comfort. These omissions make these estimates biased upwards. Another major drawback of the RP approach is that in most cases the high degree of collinearity between time savings and cost of travel makes it very difficult to get precise estimate of VOT. To solve this problem, stated preference (SP) experiments were introduced where the commuter is asked hypothetical questions and their responses are used to infer VOT. It was expected that since SP methodology, by design, can control for different levels of travel attributes, it will yield precise estimates.

Calfee and Winston (1998) pointed out that most of the RP VOT studies estimated VOT by comparing different modes of travel (like whether to ride a bus or drive a car) and are likely biased due to unobserved attributes which vary across modes. However, to decide on the viability of public investment in roadways we need VOT that has been derived by observing a motorist's trade-off between a free-congested and a paid-uncongested alternative. This observation was not new to the literature but since none of the congestion pricing projects had this feature, RP studies were not feasible. So they (Calfee and Winston) conducted a stated preference study where respondents were queried about hypothetical choice situations regarding solo driving in a congested condition and paying a toll to avoid congestion. Their estimates of value of time were well below the previous estimates (roughly 10% of wage rate). In a similar study Hensher (2000) estimated it to be $8 which is again quite low compared to the revealed preference studies. Such low level of VOT implies that many of the public investment in roadway improvements are probably not beneficial.

However, in recent times there has been a number of congestion pricing projects that allow us to measure VOT without the bias factor mentioned before. Two such projects are in California – the SR-91 project and the San Diego I-15 Congestion Pricing Project (SDCPP). The former one has been extensively studied and a comprehensive review of VOT estimates can be found in Lam and Small (2001). This thesis studies the San Diego I-15 Congestion Pricing Project. SDCPP is a demonstration project where solo drivers are allowed to use the carpool lane by paying a toll. The toll is varied every six minutes to keep the lane freely flowing. Thus the toll varies in response to real time traffic volume. This is a rare opportunity to observe and analyze commuters' behavior in response to real time congestion pricing. Since the decision problem for the commuter is to choose between a tolled and uncongested alternative versus a free and congested alternative, this provides a good natural experiment to infer about VOT from the commuters' travel time and travel cost trade-off.

Brownstone et al (2000) and Brownstone et al (2001) analyze the morning commute for this corridor (using RP data) and derived a value of time of roughly $25-$30 which is well above the recent estimates from SP studies. Lam and Small (2000) have analyzed the SR-91 corridor in Orange County, California with a time-of-day pricing scheme. Using RP data, they estimate the value of time savings to be 58% of wage. Thus comparing the SP and RP studies we find that there is a significant difference in VOT estimates. Since most transportation projects are decided on the basis of cost-benefit analysis, such a significant difference will matter in the evaluation of a project.

Heterogeneity in VOT has received considerable attention in recent years. It has been pointed out in the literature that unobserved heterogeneity is the reason for such significant difference between revealed and stated preference estimates. Typically RP models don’t control for unobserved heterogeneity thus resulting in VOT estimates which are biased. The issue of heterogeneity in VOT has also been raised in the debate over the welfare implication of such congestion pricing schemes. Under the first best solution everyone should be charged according to his or her marginal benefit. But most of the congestion pricing schemes implemented charges one of the routes while leaving the other free. Thus in this second best world the natural question to ask is under what conditions these pricing schemes will increase consumer welfare. It has generally been agreed in the literature that if commuters have identical value of time then welfare is almost certain to go down (Weitzman (1974), Evans (1992)). But if there is heterogeneity in the value of time then congestion tolls may lead to an increase in welfare (Glazer (1981), Small (1992), Verhoef and Small (1999), Glazer & Niskanen (2000)) even when levied on only part of the network.

The research agenda of this thesis is to study behavioral models using data from I-15 Congestion Pricing Project and interpret the models to understand how commuters decide between a tolled and a free alternative. These models are also used to calculate VOT while controlling for heterogeneity.

The issue of uncertainty in travel time has received increasing attention in transportation research (Small (2000)). A commuter using the toll lane is not only paying for a reduction in travel time, but also for a reduction in the uncertainty in travel time. The value of variability (VOV) measures how much commuters are willing to pay for a marginal reduction in uncertainty. Uncertainty in travel time is incorporated in all the models and estimates of its valuation (VOV) are presented.

Chapter 1 of this thesis summarizes the literature on VOT and congestion pricing and describes in detail the I-15 Congestion Pricing Project. Chapter 2 describes the data that is used to evaluate the project. Five waves of surveys were conducted from fall 97 through fall 99. These survey data were supplemented with network data for this corridor. Chapter 2 also explains how these two data sources were merged to estimate the behavioral models,

Chapter 3 presents disagregate RP models of mode choice with survey data from Wave5 (fall 99). Mode choice[1] models are estimated using conditional logit modeling and allowing for observed heterogeneity. Observed heterogeneity is typically introduced by making the VOT a function of some observable attribute: either as a function of demographic characteristics or as a function of travel attributes. One of the criticisms leveled against this type of modeling is that it ignores individual heterogeneity (or the random component of traveler's decision process which cannot be observed) and thus results in biased estimates (Hensher (2000)). So results from models incorporating individual heterogeneity are also presented.

The two questions: 'Is there heterogeneity in VOT and how are these sensitive to model specification' are explored chapter 3. The conditional logit model answers our first question and shows that there can be heterogeneity depending on gender and road conditions. As to the second question several models are estimated which allow for individual heterogeneity and it is shown that the estimates are pretty stable. In part 2 of this chapter, which uses RP data from wave 3 (fall 98), this question is approached by estimating the model for a different time period and also adding another dimension to the choice problem. Here the morning and afternoon commute is modeled as a joint decision. The purpose of this approach is to control for unobserved heterogeneity as these two choices are made at two different points in time. These estimates again confirm the existence of heterogeneity and are very similar to those presented in part 1. Comparing the different models in part 1 and 2 we conclude that the mode choice models presented here are robust to the specification of unobserved heterogeneity.

The general results from these models are that this tolled facility (which will interchangeably be called the FasTrak lane or the HOT lane or the Carpool lane) is used by high income, middle-aged, homeowners and female commuters. An interesting result that comes out of this analysis is the dual effects of toll. Generally the toll a driver pays is representative of the cost of travel and a rising toll is associated with reduced FasTrak use. In this case however the toll can also act as a signal of congestion. Since the toll on the FasTrak lane adjusts to the level of traffic every six minutes, the commuter can extract information from the level of toll. The models estimated show that if the actual toll rises above the

---

[1] The word 'mode choice' will be used in this paper to denote the 3 choices commuters have in this case - solo driving in regular lane or paying a toll and driving in the express lane or carpool

mean toll (or what the commuter expects to pay) then the commuter is more likely to take the FasTrak lane.

The effect of toll also depends on the level of uncertainty of travel time. If the uncertainty rises above a particular value commuters are willing to pay very high prices to use the HOT lane. Conversely uncertainty in travel time encourages use of FasTrak lane only if toll rises above a threshold value.

In chapter 4 of this thesis, models of mode choice are estimated using SP data collected for wave 5 of the survey. The behavioral implications from these SP models, which were controlled for both observed and unobserved heterogeneity, were found to be same as the RP model. However, the VOT estimates from these models are significantly lower than the RP estimates and reflect a consistent pattern that has been observed in the literature. Chapter 5 is devoted to explaining the difference and several approaches are pursued using both RP and SP models to reconcile the difference. But it is consistent and persistent across the different models, which leads to the conclusion that these differences are real. Probably it reflects the difference in responses of individuals to actual and hypothetical situations.

The value of variability (VOV) measures how much commuters are willing to pay for a marginal reduction in uncertainty. Uncertainty is measured by taking the difference between 90th and median time savings. The models presented here estimates roughly $30 to be the amount commuters are willing to pay to reduce one hour of uncertainty in travel time. The general conclusion is that though commuters are willing to pay a high amount for a reduction in variability, it is dependent on the time of travel. Commuters are more sensitive to variations in travel time in the morning, specially during the peak, than in the afternoon.

CHAPTER 1

LITERATURE REVIEW AND DESCRIPTION OF THE I-15 CONGESTION PRICING PROJECT

1.1: LITERATURE REVIEW

The basic model for analyzing congestion was developed by Vickery (1969). The model considers the case of a single road which connects destination A to B. If the capacity of the road is below the optimal capacity for free flow it would result in a bottleneck and forming of queues. The private cost to the commuter is the waiting time in the queue (travel cost) and the penalty for arriving late or early (schedule delay cost). Assuming a uniform distribution of arrival of commuters, he computed the length of the queue, the maximum and minimum waiting time in the queue. He then examines the effect of a varying toll, which would eliminate the queue. He shows that in the new equilibrium a commuter would be indifferent between waiting at the queue and paying the toll and passing the bottleneck same time as he would have crossed in case of a queue. The toll would leave the commuters as well off as before and the revenue from the toll would be a "gain". Since the toll does not affect the arrival time of the commuters at the city center, this toll do not pose a burden. Since this seminal work there have been numerous development in the travel demand literature. The review is organized into four different sections. The first section summarizes the literature on congestion pricing and its impact on commuters. In the second section I summarize the results from mode choice literature. In the third section some of the literature on joint mode choice and departure scheduling literature are reviewed. The last section summarizes the literature on VOT and VOV. A comprehensive summary of congestion pricing projects can be found in Small & Gomez-Ilbanez (94).

1.1.1: LITERATURE ON CONGESTION PRICING

Starting with Vickery's model, the literature on congestion tolls has branched out into several directions and has become increasingly sophisticated over the years. The models differ in terms of heterogeneity in commuters, examining the impact of different pricing schemes and examining the welfare gain and losses in the second best world.

One branch of the literature studies the effect of Congestion Toll on Heterogeneous Commuters. Henderson (1974) introduced heterogeneity in the model by assuming different ratios of travel time cost to schedule delay cost. He found that congestion tolls are regressive. Glazer (81) shows that if consumers value quality of service equally, then any increase in the toll would reduce consumer welfare. But if the valuation is heterogeneous then the result is ambiguous. Cohen (1987) examines the distributional effect of congestion pricing arising from the different valuation of time of different commuters. He considers two groups with different value of time. He shows that the group with higher value of time will be better off and the other group (with lower value of time) would be worse off. Arnott et al (1994) examines optimal tolling schemes that would benefit the largest group of commuters possible. They examine the effect on commuters under three schemes. The first case considers a scheme under which the commuters are charged optimal tolls but are not paid back. The second scheme is where the toll revenues are paid as a lump sum amount to all commuters and the third is where the toll revenue is used to expand the capacity of the road. Relaxing the assumption of identical arrival times and assuming equal schedule delay costs for all commuters they found that in the case of optimal time varying toll, a toll simply replaces queuing costs with monetary costs. In this case tolls are welfare neutral and lump sum transfers make both the groups better off (than without a transfer). The effect of capacity expansion with the toll is not so clear. It would depend on the effect the expansion has on the peak.

The other strand of literature examines the effect of different types of Congestion Tolls. Arnott et al (1990) extends the rush-hour traffic flow models with endogenous departure time decision and examines different toll regimes and compares the efficiency of each regime. Application of the optimal fine toll (toll varying by level of congestion) does not change total schedule delay costs but eliminates travel time, resulting in a social saving on total travel cost. The nature of the equilibrium is dependent on the relative magnitude of value of time and value of being late at work. The authors (based on an empirical estimate by Small) assume that the latter is larger than the former. The effect of the fine toll would be to eliminate queue and ensure a smooth flow of traffic. Arnott et al (1993) introduced a demand function in the bottleneck model where the demand for trips is a function of the price. The supply curve is derived as a reduced form equation from the bottleneck model under different tolling regime. The equilibrium trip price and the number of trips are derived as a function of the capacity. They found that the fine toll is the most efficient one with the total price and number of trips equals the no toll case.

Chen-Hsiu (1994) considers the case of a multi-step toll. He finds that as the number of steps goes up to n, the proportion of queuing time is reduced by n/n+1. Small & Anderson (1998) examine the effect of tolls on revenue in case of an untolled alternative. The relationship between efficiency and profit maximization is not very clear. In the short run, the toll operator inflicts some negative externality on the free road which is not reflected in the price while in the long run from a capacity expansion, the free road has a positive externality which the monopolist cannot internalize. The basic conclusion is that the presence of transaction costs, congestion pricing is still possible by simplified tolling system and the private operators don't lose much revenue in the process.

There are numerous papers which look at the case of second best pricing or the existence of an untolled alternative Marchand (1968) analyzed second best tolls in the presence of an untolled alternative. He analyzes the equity impact of the pricing and concludes that the toll affects people whose marginal utility from using the network (which is underpriced relative to socially efficient level) is higher relative to a composite good. Verhoef et al (1996) considers a network model with two routes, one tolled and another untolled. The regulator has to decide on a toll that maximizes efficiency subject to the constraint that one of the routes cannot be tolled. There are two effects that they talk of: an overall demand effect or where the commuters efficiently decide not to use the route at all due to congestion pricing (captured by the elasticity of demand) and a modal split which means that the remaining commuters efficiently decides which route to take (depends on the marginal external cost of the two routes). They show that for most of the cases the second best would be inefficient compared to the first best case and the level of inefficiency would depend on the demand and cost parameters. The authors also show that second best tolling can be more efficient than the first best tolling if the cost of regulation is factored in. The probable cases are of that of an inelastic demand in combination of free flow cost differential (cost parameters on the untolled alternative exceeds tolled alternative) especially for fixed (route specific) costs of regulation.

Braid (1996) examines the same situation as before with a bottleneck model. He compares three optimal pricing schemes and compares it with the base case of no–toll equilibrium. The cases are:

- First Best: Congestion toll on both roads.
- Second Best: A flat or no toll on one of the route. The optimal toll on the second route has two components – a positive time varying component which prevents queuing on the second route and uniform component that attracts traffic from the untolled route.
- Third Best: No toll on one of the route and a congestion toll on the second route.

An interesting parameter that was introduced in this model is the different capacity of the routes. The author shows that the efficiency gain from second best pricing is always greater than the third best and equals 2/3 of that of the first best pricing if capacities are same. If the capacity of the tolled alternative is less than the untolled one then the gain in efficiency relative to the third best is even larger but it also leads to a lesser gain in efficiency relative to first best pricing.

Using a similar network model and incorporating peak and off peak demand, Liu and McDonald (1999) isolate three effects due to the imposition of congestion pricing. The second best toll regime yields unequal traffic volume on the two routes with some of the commuters from the priced routes shifting to non-priced route. For the other 2 cases, i.e. first best and no toll situation, traffic volumes are identical in both routes. The second effect involves shifting peak period traffic to pre peak period. The presence of tolls reduces peak period travel and increases off peak period travel. The impact of first best tolls is larger than second best tolls. Lastly, both toll schemes reduce traffic volume relative to the no toll case and, as expected, the reduction is more for the first best tolls than the second best tolls. The authors then compare the welfare properties of the three cases and find that welfare is larger under all three scenarios if the routes are similar in size. The more dissimilar they are in size the larger is the welfare gain from congestion pricing.

1.1.2: LITERATURE ON MODE CHOICE

Empirical analysis of the above models is primarily done by what is commonly known as desegregate or behavioral travel demand modeling. It involves estimating discrete choice models using micro data. The typical approach in these empirical studies is to estimate discrete choice models of solo driving against some other mode, mainly public transit. These models are then used to estimate elasticities and value of time.

Bajic (1984) estimated a binary choice model between automobile and public transit system for Toronto metropolitan area. His estimated model seems to indicate that mode choice decisions are sensitive to the performance of the transportation system. Though the cost variables have the expected sign the low significance level of a number of parameters weakens the reliability of the value of time calculations. In general, the author found that walking and waiting time are more onerous than in-vehicle time and value of time is roughly 43% of hourly wage, which seem to conform to the existing estimates. He also estimated elasticities for the different mode choice and found that vehicle use is inelastic to fuel cost change and change in public transit fares from which the author concludes that public transit will not affect auto use in any significant way.

Southworth (1981) estimated separate mode choice and destination choice models for the West Yorkshire region of England. Though he found several significant demographic variables influencing mode and destination choice, the travel cost variable coefficients estimate were imprecise. The author found significant differences in value of time calculations between work and non-work trips. He also found that the value of time estimate is much lower for the destination choice model than the mode choice model from which the author concluded that association between cost and mode choice is stronger than cost and destination choice. Like previous studies he found that walking and waiting is more onerous than in-vehicle time. His estimate of in-vehicle value of time is on the lower end of the band at 11% of hourly wage.

Madan & Groenhout (87) estimated mode choice between public and private transport for Sydney work trips. The authors introduced a measure of employment density at work to explain mode choice. They showed that income and employment density are important explanators of mode choice and omitting them tend to overstate the effect of travel cost variables and travel characteristics. From the prediction success table it seems that though the model seems to do a fairly good job of predicting mode choice for the entire sample, it fails for a subsection of the sample for which the employment density is high. Their model seems to over-predict transit use. The elasticity measures tend to imply that though access time is not that different from in-vehicle time wait time is significantly more onerous. One notable result they derived was that the value of time might vary significantly across individuals.

Bhat (1997) introduced a new dimension in mode choice analysis – stops made for non-work purposes on way from home to destination and vice versa. Using data from Boston Metropolitan area the author estimated a joint unordered mode choice model and an ordered stop model. He finds that income, number of vehicles per driver, work duration and employment density are important factors influencing mode choice. Demographic characteristics, activity level, in-vehicle and out-of-vehicle time are important determinants of stops made. The elasticity estimates by the author seem to suggest that improvement in public transit and disincentives to solo driving does little to change solo divers who make stops. Thus it does not affect the congestion level since these are the people contributing to congestion by making more trips. On the other hand it affects solo drivers who did not make a stop and did not contribute much to pollution in the first place.

1.1.3: LITERATURE ON MODE AND SCHEDULING CHOICE

Typically the main assumption in a route/mode choice framework is the assumption of exogenous scheduling. Small (82) has shown that such assumptions would lead to a bias in policy prediction. Since then there have been quite a number of papers trying to model scheduling choice as an endogenous variable. Bhat (1998) formulates a joint mode and departure time choice for shopping trips using a nested structure in which mode choice is nested at a higher level and departure time choice at the lower level. He uses a multinomial logit model (MNL) for mode choice and standard ordered generalized extreme value (OGEV) model for departure time choice. They used survey data from the San Francisco Bay Area Household Survey which had single-weekday travel diary with information on shopping trips. They estimate a MNL, Nested Logit and MNL-OGEV models and compared their relative efficiency. They found that the MNL-OGEV performs best on grounds of data fit and failure to take into account correlation between adjacent departure times using the same mode will seriously bias the estimates.

McCafferty & Hall (1982) estimated a multinomial logit model of three period departure time choice and then tested it's stability by re-estimating the model after an exogenous effect of road closure which is expected to affect schedule choice. Their model fails to isolate any demographic factors influencing choice and the travel time coefficient was not significant. Though they estimated their model with only people with flexible working hours, their results showed a strong bias towards travel during peak hour. The conclusion the author reached was that this is evidence that flexible working hour may not be an important factor in schedule choice and probably it is guided by some other socio-economic factors. Hendrickson & Plank (1984) also argue that mode choice should consider scheduling choice as characteristic of travel varies by time of day. The authors used data collected in Pittsburgh, PA consisting of 1800 worker in the Central Business District and independent measurement of travel times and transit wait times. They estimated a logit model of simultaneous mode and schedule choice. Their value of time calculations show that access time to transit has the highest value followed by the waiting time. They also showed that lateness is more onerous than early arrival. They also computed elasticities with respect to several policy changes and found that reduction of wait time resulted in highest consumer surplus.

S Pellis (1987) argues that people decide on a safety cushion (time allocated for unexpected delays) based on two opposite forces. It is generally assumed that early arrival is more onerous than staying at home and they also value late arrival more than staying at home. Thus people maximize the time spent at home subject to the constraint of tolerable lateness. The author estimates the slack time substitution effect by a stated preference experiment. People were asked to select between pairs of travel time in which one had a reliable arrival time and the other was a cheaper option with some degree of risk. The value of slack time substitution was found to be 1.5 pence per min and value of lateness was found to be 5.6 pence, though it was sensitive to the frequency of lateness. People seem to leave home later as the reliability of service increases.

Bates et al (1990) estimated a departure time choice model by conducting an SP experiment with several departure schedules and cost attached to it. Their analysis brings out the importance of flexibility of schedule – people who are constrained by arrival time are less willing to move after peak hours than people with constrained departure times. In the afternoon, people with no departure constraints tend to leave much earlier than the peak. Overall they showed that people who travel during the peak hours are less willing to change their departure patterns and there is significant difference commuters who are arrival constrained from those who are departure constrained.

Abu-Eisheh and Mannering (87) also points out that the models which try to incorporate departure time explicitly in the model generally assumes that the departure time choices are discrete. This assumption is relaxed in this paper. The choice of departure time is a function of work start time, walk access time, travel time and cushion for delay. The first two are taken to be exogenous. Unlike other papers this paper relaxes the assumption that commuters don't have any influence over the travel time – the authors argue that this is true only in case of extreme congestion and the travel time is a function of the route specific characteristics (flow, speed etc), socio economic characteristics of the commuters and also vehicle characteristics. They used a single origin destination pair in Pennsylvania metropolitan area. The pair was connected by 3 distinct routes. They estimated models for morning commute using a survey data and their route choice model indicates that travel time affects choice in a negative way but only when interacted with income. The problem with this paper is that the estimates are not statistically significant.

Cascetta et al (1992) estimated a joint route-schedule choice model for morning commute using RP data from a single destination connected by several primary and secondary routes. They found that the difference between primary and secondary routes matter and travel times differentiated into the above categories performs better in explaining than travel time taken together. Safety perception and other comfort variables (no of left turns etc) seem to matter. They also find that there is difference between being early and being late. This perception was also found to vary across demographics. The authors performed a two-way anova analysis for departure time choices based on demographics and found them to significantly affecting route choice.

Mannering et al (1990) estimated afternoon commute decision using RP data from the Seattle area. In particular they estimate a logit model with three choices. The commuter can decide to never delay, delay and work, delay and spend time on other social activities. They found that the more trips the commuters make during afternoon peak hours the more they are likely to delay their departure. They also estimated a right truncated Poisson distribution to check the number of times the commuter would delay departure to avoid congestion. This also shows that people who have to travel frequently during peak hour would try to postpone their departure at a greater frequency and so would people with longer commute. To estimate the duration of delay the authors also estimated a hazard function which they found increasing, implying that the more the duration of delay the higher the probability that the delay would end. They found a 1% increase in travel time would lead to a 2.5% increase in delay.

Calpice (92) analyzed the impact of desired arrival time, availability of information on route and schedule switching for morning and afternoon commute based on a SP questionnaire for travelers in Austin, Texas. The author estimated a Poisson model of preferred arrival time based on socioeconomic variables. He found that these affect two groups of people – people with lateness tolerance (flexible arrival time) and people with no lateness tolerance – in very different manners. Commuters' response to information was estimated by a binary logit model with the choices being listening to radio or not. They found more aged people and females are more likely to listen to radios. Also if the commuter is travelling during peak and has a preferred arrival time, it influence the choice positively. Lastly they estimated a route and departure-time switch decision by multinomial logit. Separate models were estimated on route or time switching for AM and PM. Each case has four choices do nothing, switch route, switch time, do both. Commuters seemed to be influenced by network variables while making a route switch and socioeconomic variables and work office conditions seem to affect departure time choice.

Polak et al (93) estimated a schedule choice model for different activities using SP data. A theoretical model was developed and then estimated. They concluded that commuters seem to react by a shift in timing of the trip to later depending on the amount of time spent at work in a nonlinear manner. For leisure and shopping trips the effect was weakly nonlinear and sometimes linear. Lateness penalties seem to affect commuters more than shoppers or leisure activities. Shoppers and leisure seem to be more responsive to a delay in participation time (time spent at destination) and not so much to an increase. For shopping and leisure activities there seem to be evidence of a hierarchical structure of time and mode choice but for commuters they are simultaneous.

DePalma et al (1996) calibrated OLS and Tobit models to analyze departure schedules (both morning and afternoon) for commuters in Brussels. They found that the scheduling patterns for commuters with flexible and rigid arrival times are not that different which tend to imply that commuters with flexible arrival timing are not shifting their travel time to avoid the peak period delay. Automobile travelers with longer travel time tend to change their schedule more often than transit users which tend to suggest that cost of travel is nonlinear and dependent on mode choice. The authors also concluded an inertia to change on the commuters part in the sense that 50% of travelers were unwilling to change their schedule to save 10 minutes and whenever they changed their schedule it was mainly due to some personal engagement, rather than time savings considerations. They also reported that females are more likely to leave later and managers tend to leave earlier which indicates a strong influence of demographics on scheduling choice. Role-related constraints (like dropping off kids to day care center) seem to influence the decision to leave home and work related constraints seem to influence the decision to leave workplace.

Jou & Mahmassani (96) focused on the daily variability of departure time and route decisions for morning and afternoon work-related trips based on a two-week travel diary. They found that commuters are more likely to switch routes and or departure schedule in the morning than in the evening which they interpreted as an evidence of arrival time constraint at workplace. They also found that switching pattern for Dallas was quite similar to that of a similar study in Austin, specially in the evenings. The morning commute patterns were quite different in the two cities. They also found evidence of interdependence of schedule and route switching with the former occurring at a higher frequency than the later.

Koskenoja & Khattak (97) examine the factors that influence commuters' propensity to change afternoon departure times. They also noted that that travel characteristics have insignificant effect on schedule changes and it is more affected by demographic variables like occupation and workplace flexibility. They also examined whether these impacts are symmetrical on late and early departures and reported that there are some differences based on occupation (executives tend to leave later), work time flexibility and income but no major difference was again noticed from travel characteristics.

Emmerink & van Beek (97) examine the impact of factors influencing work-start time and how it influences road pricing and driver information systems. From a survey conducted in Netherlands it was found that income significantly influences work-start time and people with high income and education are more likely to have flexible work-hours. They interpret this as an indication that pricing is regressive since commuters with high income can avoid paying the toll. They also found that employees don't always fully exploit flexible schedule provided by their employer. The reason for this is social and family constraints.

1.1.4: LITERATURE ON VOT AND VOV

One of the earliest attempts at estimating value of time (VOT) was done by Beesly (1965). He used a graphical approach to solve for the trade-off between cost and time of travel. His estimate of VOT was roughly equal to 35% of average wage rate. Since then remarkable progress in discrete choice analysis has made it the predominant methodology in calculating VOT. Some of the earlier studies using discrete choice analysis are Lave (1968), Lisco (1967) and Hensher (1974). The estimates obtained by Lave and Lisco roughly equals 42% of wage rate whereas Hensher derived the VOT at 20% of wage rate. These earlier studies have been comprehensively reviewed by Hensher (1978). He criticizes the use of revealed preference data on a number of accounts. A key criticism is that the travel attributes that are measured by the researcher may fail to measure accurately the 'reported' values by the commuters. Thus he advances the case for stated preference analysis and cites some earlier studies like Lee and Dalvi (1969), Heggie (1976).

In a recent survey, Small et al (1999) points out another methodological problem (citing Stopher (1976)) of estimating value of time. Most of the earlier studies did not control for different attributes of travel. The difference between in-vehicle and out-of-vehicle VOT or those between congested and uncongested travel times are some of the examples. In other words the estimation of a homogenous VOT was found to be too restrictive. Mohring et al (1987), Train (1976) and Bradley et al (1986) are some of the studies which took such differences into account and found that valuation of congested travel time is considerably higher than uncongested travel times. However, Hensher et al (1989) did not find any significant difference between congested and uncongested travel times. An interesting difference between them is that the study by Hensher used stated preference data and the other three used revealed preference data.

There are several exhaustive surveys which summarize more recent findings. Among them are Small (1992), Waters (1992), Wardman (1998) and Small et al (1999). All of these surveys find an enormous variation in the value of time estimates ranging from a low of 20% (Bruzelius, 1979) to a high of 72% (Cambridge Systematics, 1977). Small (1992) and Walters (1992) conclude that 50 percent of gross wage rate is a reasonable estimate. Small (1992) summarizes the general consensus from these studies.

The value of in-vehicle time for non-business travel is usually found to be less than the gross wage rate and it rises (though not proportionally) with income.

Walking and waiting are more burdensome than car travel and thus have a much higher valuation of time.

Business travel seems to have a higher value of time than commuting.

Another problem associated with most of the earlier studies is that they compare auto travel with some other mode of travel such as bus or train which tends to make the estimate biased. Calfee and Winston (1998) argue that a correct value of time for road users should be estimated by the choice of commuters between a tolled and free alternative. Using a stated preference data, they put the estimate at a much lower level (an average of $ 3.88, or 19% of wage rate). In a recent stated preference analysis Hensher (2000) showed that using a multinomial logit model can seriously underestimate the value of time because it fails to take into account unobserved heterogeneity. Using a mixed logit framework he derived the estimate at roughly U.S. $8 (it can go up to $14 depending on the standard deviation). Though he does not report it as a proportion of wage rate it is clearly on the lower range of the estimates discussed above.

Revealed preference analyses in this area are very few because there are only three projects in US which gives us comparable estimates. They are: the Kate Freeway HOV-lane Pricing Project in Texas, the SR-91 Project in Riverside and the I-15 Congestion Pricing Project in San Diego (both in California). Using data from the I-15 congestion pricing project, Brownstone et al (2000) estimated VOT to be ranging anywhere between $18 to $30. Using data from SR-91 project, Lam and Small (2000) estimates VOT to be anywhere between $4.74 and $24.52, depending on the model one is looking at. Using the most complete model they put the estimate it at $ 22.87 or 72% of wage rate. Thus the revealed preference analyses have shown consistently higher values than the stated preference experiments.

One other area that has received considerable attention in recent time is the uncertainty in travel time. Small et al (1999) provides a comprehensive review of the issues, the theoretical and the empirical literature. The difficulty in identifying uncertainty and the lack of data has limited the number of studies measuring the value-of variability (VOV). Sena (1994), using a stated preference experiment estimate the value to be equal to $ 1 or $ 2 depending whether they are constrained by arrival time or not. In particular the author found that the ratio of VOV to VOT can vary from a low of 33% to a high of 150% depending on the model. Hensher (2000), using a stated preference experiment, estimates the value to be $5 per hour. These studies used standard deviation as a measure of uncertainty. Brownstone et al (2000) using a different definition of uncertainty (difference between 90th and median time savings) estimated that people are willing to pay $20 to reduce variability by an hour. Lam and Small (2000), using the same measure estimate it to be $ 15.12 per hour for males and $ 31.91 for females. Thus again we see a significant difference in estimates between stated and revealed preference data.

1.2: THE I-15 CONGESTION PRICING PROJECT

The Interstate 15 (I-15) Congestion Pricing Project is a federally funded demonstration project which allows the use of High Occupancy Lanes (HOV) to solo drivers for a price. The Express Lanes are dual reversible lanes in the median of I-15, extending approximately eight miles north from the I-15/State Route 163 junction with no intermediate entry or exit points. Prior to 1996 these lanes were open to car-poolers, vanpoolers, buses and emergency vehicles. These lanes were underutilized with peak-hour volumes less than 1000 vehicles per lane, while the regular lanes faced high levels of congestion. In 1991, as a part of planning for improvement in air quality in the region, the I-15 Congestion Pricing Project was identified. The aim was to use HOV lanes more efficiently, improve transit along the corridor, and relieve congestion on I-15.

Since it required special legislation to allow single occupancy vehicles (SOV) to use HOV lanes, the project faced substantial political obstacles. Another concern at this time was that excess use of the lane use by the SOVs would adversely affect the HOVs. Meanwhile a local mayor was pressuring the local transportation agency to find a way of funding public transit to his suburban area located on the same route. These factors resulted in a scheme to sell off the excess capacity of the lane to solo drivers and to use the revenue generated to finance public transit. In effect, the HOV lanes become both HOV and toll lanes, generally referred to as HOT lanes. The corridor is northeast of the main employment centers in San Diego with pronounced unidirectional commute patterns (southbound in the morning, and northbound in the evening). The HOT lanes are separated from the main lanes and operate in only one direction depending on whether it is morning or evening. Entry occurs at one point, and the entire length must be traversed before exiting. The lanes are operated in the southbound direction from approximately 5:30AM to 10:00AM and in the northbound direction from 2:30PM to 7:30PM. The location of the Express Lanes is shown on the map of Figure 1.

The demonstration project began in December 1996. In the first phase of the I-15 Congestion Pricing Project (December 1996 through March 1998) or the ExpressPass (EPR), a limited number of solo drivers (approximately 500) were allowed to use the Express Lanes at a fixed fee. The monthly fee was $50 to start with and was later revised to $70 and it allowed unlimited access to the solo drivers.

For the next phase of the project, started in March 1998, per-trip pricing was implemented. During the second phase (March 1998 through December 1999), subscribers were issued windshield-mounted transponders used for automatic vehicle identification. The second phase is referred to as FasTrak™. FasTrak user accounts are automatically debited a per-trip fee when they use their transponder. The fee is posted on changeable message signs upstream from the entrance to the lanes, and can be varied every six minutes. There is no limit on the number of subscribers to the FasTrak system. Instead, the fee is adjusted to maintain free-flowing traffic conditions in the HOT lanes[2]. The trip

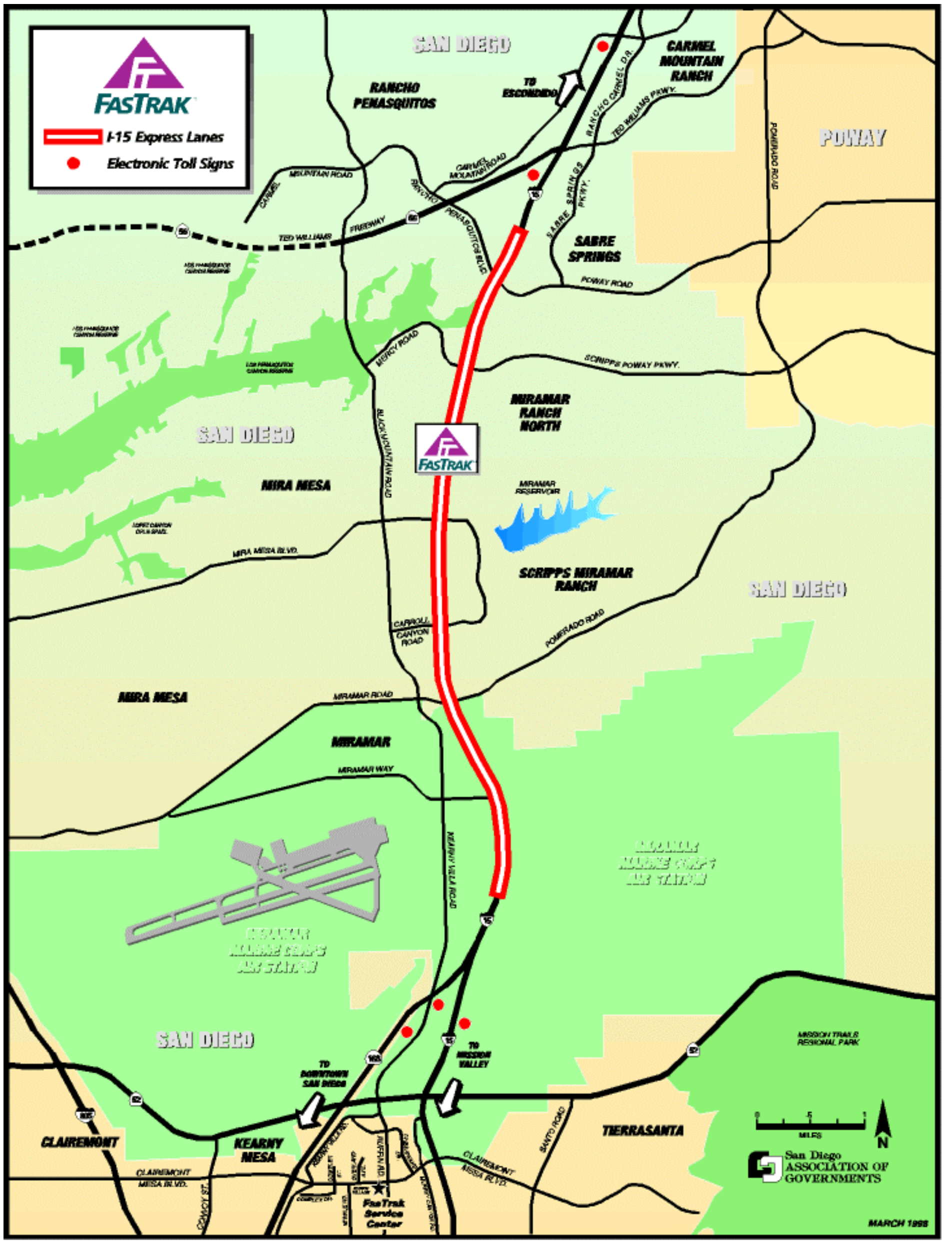
---

[2] The toll is set such that the capacity goal is met. The capacity goal is 1,300 vehicles per half-hour in the AM peak, and 1,440 vehicles per half-hour in the PM peak. This corresponds to level of service rating C (LOS C).

price is adjusted depending on the traffic conditions of the lanes, in order to maintain a satisfactory level of services for HOVs and can range from $.50 to $8.00. As before, HOVs will continue to use the Lanes at no cost. An electronic toll collection and violation enforcement system has been installed, and transponders were distributed to previous Express Pass holders and anyone else who applies. The program is advertised on signs along I-15 and in local newspapers. The revenues collected from the program are used to improve transit and carpool services in the I-15 corridor.

The project is described in much greater depth in Golob and Golob (2000), Golob, *et al.* (1998) and Supernak, *et al.* (1999). The San Diego Association of Governments maintains a web site (http://www.sandag.cog.ca.us/data_services/fastrak/) with the various evaluation reports. SANDAG is responsible for overall project coordination and management. The California Department of Transportation (Caltrans) is a primary project partner responsible for overseeing design specifications, physical improvements and operational changes on the facility. The California Highway Petrol (CHP) provides enforcement services and the San Diego Metropolitan Transit Board (MTDB) assists in planning and implementing transit services improvements funded by the project.

Figure 1. Location of the I-15 HOT (or Express) Lanes[3]



---

3 (Source: www.sandag.cog.ca.us/data_services/fastrak)

CHAPTER 2

DATA DESCRIPTION

San Diego State University has conducted a survey of commuters using this corridor in order to evaluate the project. This thesis uses data collected from that panel survey as well as loop detectors embedded in the roadway. Section 2.1 describes the panel survey and summarizes the demographics. Section 2.2 describes the collection, and estimation of time savings and toll data and the way it was matched to the survey data.

2.1: THE PANEL SURVEY

The panel survey consists of three samples of approximately equal size: a) FasTrak subscribers and former subscribers, b) other I-15 users and c) users of Interstate 8 (I-8, which is another north-south route with no HOT lanes) in San Diego as a control group. This analysis will use the data collected during the fall of 1998 and 1999 (October through November). During this time period, dynamic per-trip congestion pricing was well established. The target population is southbound morning commuters. FasTrak subscribers were picked at random from a list maintained by the billing agency, and the remaining respondents were recruited using random digit dialing of residential areas along the respective corridors. A partial quota sampling procedure was used to increase the number of carpoolers in non-subscriber parts of the sample. This procedure, known as choice based sampling, allows us to get enough observations for each mode and helps in improving the precision of the estimates. But the caveat is that any prediction or inference about the population has to be weighted by appropriate sampling weights. General results from analyses of the first wave of the evaluation panel, which was conducted in fall 1997, are summarized in J. Golob, *et al.* (1998).

Survey respondents were queried for detailed information about their most recent inbound and return trip along I-15 if that trip was made during the hours of operation of the HOT facility and covered the portion of I-15 corresponding to the facility.

Summary of the individual and household demographic data is presented in Table 1. For Wave 5 demographic characteristics for the entire survey sample are summarized because the estimation sample size varies across different models, but for wave 3 the demographics are summarized for the estimation sample only.

TABLE 1: DEMOGRAPHIC CHARACTERISTICS
(FIGURES ARE IN PERCENTAGES)

	FTP List & I-15 users		I-8 Users
	Wave 3	Wave 5	Wave 5
Income			
Less Than 40,000	3.91	8.88	25.43
Between 40,000 and 80,000	34.35	34.74	49.63
Above 80,000	60.34	55.18	23.82
Don’t Know /Refused	1.4	1.2	1.12
Total	460	1040	401
Home Ownership			
Owns Home	85.47	82.22	72.44
Rents/Lease/Other	14.53	17.78	27.56
Total	461	1046	450
Age			
18 to 34	14.78	15.2	18.27
35 – 44	41.30	34.71	32.96
45 – 54	32.61	30.87	28.73
55 or over	11.30	19.23	20.04
Total	460	1040	449
Gender			
Female	41.43	43.06	47.78
Total	461	1052	450
Education			
Less than Bachelors Degree	9.80	39.31	65.78
Bachelors Degree	60.35	29.68	17.33
Graduate School or Higher	29.85	31.01	16.89
Total	459	1048	450
Number of People Working Outside House			
None	0.87	4.13	5.79
One	31.24	32.25	32.96
Two	56.62	52.3	47.44
Three or more	11.28	11.33	13.81
Total	461	1042	449
Number of Licensed Drivers			
One	11.28	12.86	16.52
Two	68.76	66.7	60.94
Three or more	19.96	20.44	22.54
Total	461	1042	448

Number of Vehicles			
One	9.76	10.95	13.36
Two	52.71	49.28	38.98
Three or More	37.53	39.76	47.66
Total	461	1041	449
Number of People in Household			
One	6.52	7.02	11.16
Two	31.96	33.46	35.71
Three	47.17	22.69	17.19
Four or more	14.35	36.83	35.93
Total	460	1040	448

The sample of I-15 users has been fairly stable over the two waves. The estimation sample is quite similar to the entire sample and thus exclusion of observations (mainly due to missing information) has not brought in any systematic bias. Also note that the respondents who refused to disclose their income has been merged with the high-income group. Formal tests were conducted to test this and the null hypothesis (of similarity of high-income respondents and respondents refusing to disclose their income) could not be rejected. Since I-8 users are used for Wave 5, they are also being included separately.

The I-15 sample is wealthier than the I-8 sample in terms of income but quite similar in terms of other characteristics. The sample is considerably wealthy homeowners, primarily between 35 and 54 years of age. The typical household has three or more members, two licensed drivers and two workers owning two vehicles. Out of the 543 respondents in the FTP sample in wave 3, 301 were in wave 5 which means an attrition rate of approximately 45%. The attrition rate was marginally lower for other I-15 users at 40 %, where out of the 583 respondents in wave 3, 349 respondents were also in wave 5. Though this attrition rate might look large but such a rate is very common in transportation literature. But in spite of such significant replacement in the sample, the demographic composition has remained quite stable. Brownstone et al (1999) discusses the problem of attrition in this panel study. They find no evidence that attrition biases the result of discrete choice models similar to those in this thesis.

2.2 NETWORK DATA

Along with using the survey data, the choice models also rely on traffic flow data. The idea was to mimic the commuters' information-set on the day the choice was made. It has been noted in the literature that using travel attributes reported by commuters might seriously bias the results. Thus objective measures of travel-time, time savings and toll are constructed for the entire sample.

2.2.1:TIME SAVINGS AND SPEED INFORMATION

For mode choice modeling, we need to determine the time saved from travel on the HOT facility, which depends on the time of travel and the speed on the regular lanes. The speed on the regular lanes is a function of when the commuters arrive at the facility. To determine this arrival time, information from the panel survey and network data were used.

Respondents were queried as to which onramp was used for the morning commute and arrival time at that onramp. For the morning trip I have combined these responses with time-of-day speeds along the corridor to determine arrival at the HOT facility. They were also asked the time they reached the HOT lane (defined as the split of I-15 and 163) during their evening commute. The respondents provided the time they reached the HOT lane in evening.

Based upon arrival time, time saved from travel on the HOT facility is estimated using time-of-day speeds on the main lanes parallel to the facility and on the HOT lanes. Time saving is estimated for all respondents regardless of mode choice. Time-of-day speeds at several locations along I-15 were collected from California Department of Transportation (Caltrans) loop detectors embedded in the roadway. These point speeds were collected every six minutes of the morning and afternoon commute for the months of October and November 1998 and 1999. Speeds at loop detector locations are converted into speeds along the intervening segments (defined as the roadway between two loop detectors) using an algorithm developed as part of the DACCORD project.[4] The DACCORD algorithm basically assumes that the loop detector speed at the beginning of the segment applies to the first half of the segment and point speed at the end applies to the second half of the segment. Since loop detectors are placed near onramps, the I-15 corridor is effectively broken into segments traveling from onramp to onramp. Speed on the HOT facility is assumed to be 70 miles per hour based on several days of floating car experiments[5].

Median and 90th' percentile time savings were calculated using time-of-day speed data for two months. Median time saving from use of the HOT facility is estimated using the respondent's time of arrival at the facility. To capture the variability of time savings difference between 90th percentile and median time savings was calculated. The idea behind using these measures was to capture the effect that commuters, being experienced travelers, take into account the distribution of speed while making their mode choice decisions. Other measures

---

[4] DACCORD (Development and Application of Coordinated Control of Corridors) is a large European project with goals of designing, implementing and evaluating advanced dynamic traffic management systems.

[5] Speeds along the HOT lanes were measured by driving the lanes, recording start and end times, and then calculating average speed using the time differential and distance traveled. HOT lane speeds were measured every fifteen minutes of the morning peak period for five days. Speeds were generally close to 70 miles per hour with little variation across day and time.

of variability, such as standard deviation, were also tried but the difference between 90th and 50th percentile seems to give most precise results. Lam and Small (2000) found similar results with these measures. The following figures summarize median time savings for morning trips for wave 3 and 5. Figure 2 shows the median time savings by time-of-day for wave 3 and wave5. For each wave there are two median values for a given time of day because the median was taken over a month and the survey spanned over two months.

As is obvious from the graph, time savings dropped from wave 3 and wave 5. The drop is specially significant for Oct 99. Major construction work on Interstate 5 (I-5), which is another north-south freeway to downtown San Diego, diverted a lot of traffic to the I-15 around fall 98 (when wave 3 was conducted). The I-5 lanes re-opened around the time when wave 5 was being done and this resulted in a drop in time savings. The figure also shows a slight shift in the peak. Due to this drop in traffic some of the early morning time savings turned out to be negative mainly due to the assumption of constant 70 mph on the HOT lane. Since it is very unlikely that the regular lanes will be going faster than the HOT lane, these were set to zero. The time savings went back up again for Nov 99 and the traffic count data shows an increase in traffic volume. Though time savings went up for Nov 99, they are still below that of wave 3 estimates.

Figure 3 plots the variability for the same time period. The interesting thing to note is that though there has been a drop in time savings, variability has not changed much and the period with least time savings shows one of the highest variability. Thus though the lanes freed up, uncertainty remained the same throughout this period.

Figure 2: Median Time Savings For Wave 3 and 5

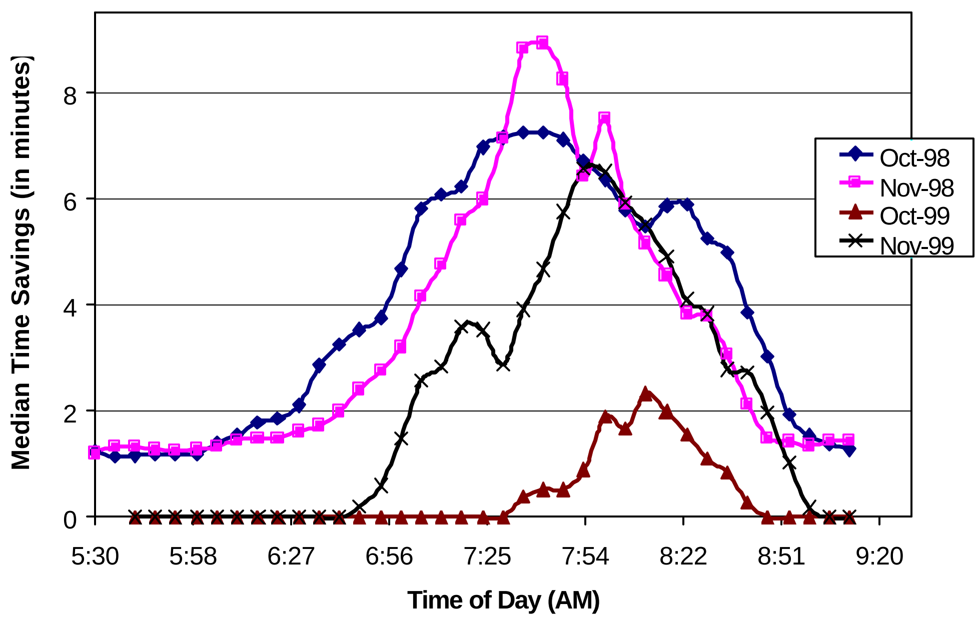
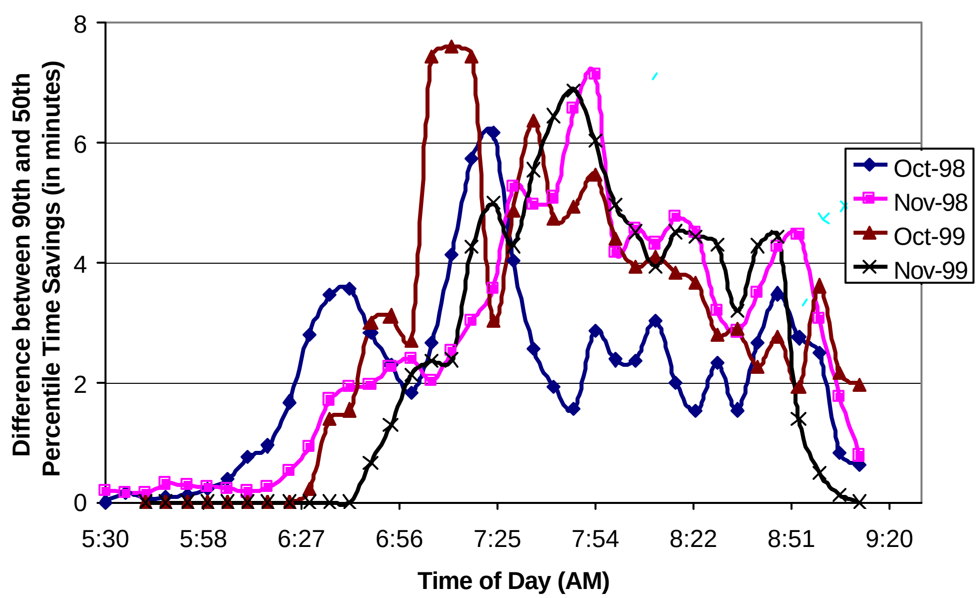


Figure 3: Variability Accross Waves



2.2.2: TIME SAVINGS AT ENTRANCE RAMP

The commuters in the sample use 20 different on-ramps to get on to I-15 but one of them deserves special attention. The Ted Williams Parkway is important on three counts. First it marks the northern end of the HOT Lanes. Secondly, a considerable number (36%) of the respondents use this onramp. Third, and most important, this particular on-ramp has a special connector that links it directly to the HOT lanes. Commuters accessing the HOT lane through this on-ramp bypass the queue at the ramp meter signal. There is considerable congestion in this area, which typically means substantial wait at the ramp. A visual inspection was conducted for 11 days in Fall 98 and Fall 99 from 5:30 to 9:12 AM and the waiting time was recorded for 15 minute interval by noting the license plate number of cars passing the onramp. The average wait by time of day was computed by taking mean over the 11 observations. This added time savings add an independent variation to the time savings from the regular lanes and helps us in controlling for collinearity. In addition to time savings, the commuters also reduce uncertainty by using this onramp. Though the data for Ted Williams Parkway is adequate for estimating average time savings but it is not adequate to form an estimate of variability. Thus the measure of variability used will be biased downwards. Figure 4 presents the data by time of day.

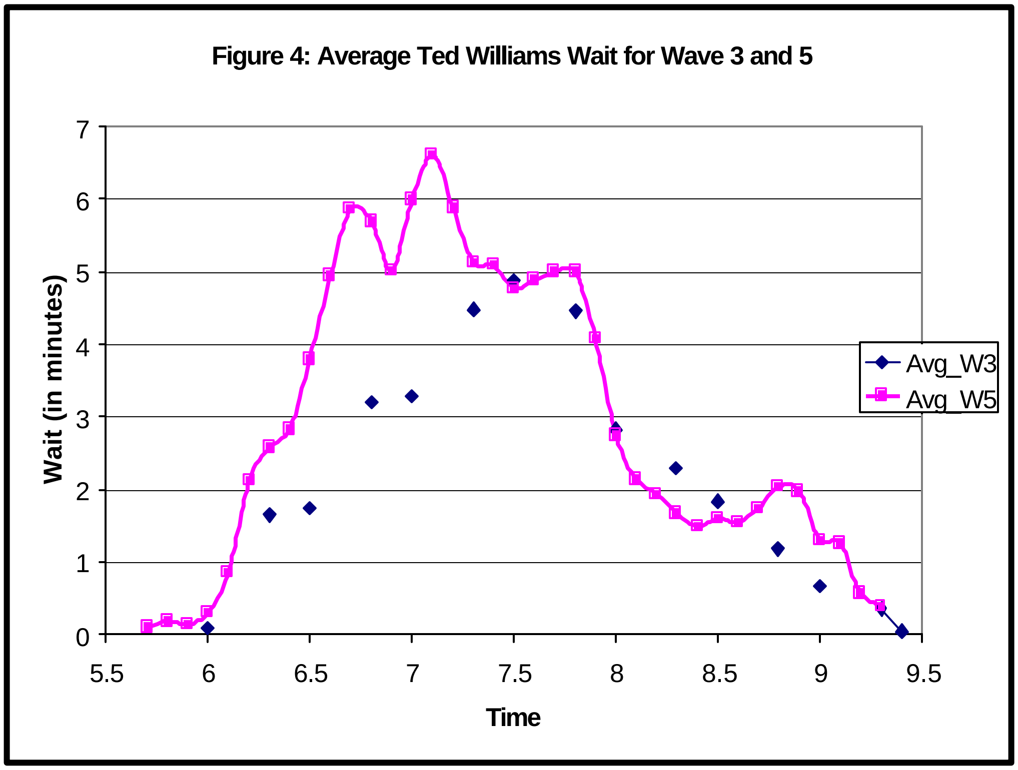


Figure 4: Average Ted Williams Wait for Wave 3 and 5
Wait (in minutes)
0
1
2
3
4
5
6
7
5.5
6
6.5
7
7.5
8
8.5
9
9.5
Time
Avg_W3
Avg_W5

2.2.3: PER-TRIP DYNAMIC TOLLS

Solo drivers who use FasTrak face a toll that is a function of arrival time at the HOT facility and the toll can change every six minutes. The level of congestion in the regular lanes determines the toll (e.g. tolls increase to avoid exceeding preset capacity constraints).[6] *While FasTrak subscribers are provided with a profile of maximum tolls that vary by time-of-day, actual tolls may be less than the maximum tolls depending upon usage of the facility. In extreme conditions, tolls may exceed the advertised maximum tolls although this is expressly advertised as a rare occurrence.*[7] *Typically during the early hours (5:30 AM in the morning and 3:00 PM in the afternoon) the toll is $ .50 and it rises with a $ .50 increment and reaches as high as $4.50 during the peak. The toll can go as high as $8, but in practice it rarely rises above $5. Toll data are available for every six minutes for October and November 1998 and 1999. From the travel time estimation described above, toll data was matched to each respondent's time of arrival at the HOT lane.*

Figure 5 shows the average toll by time-of-day for the months of October and November 1998 and 1999 (excluding Thanksgiving weekend). The tolls look quite similar across dates except for a slight shift in peak.

---

[6] The capacity goal is 1,300 vehicles per half-hour in the AM peak, and 1,440 vehicles per half-hour in the PM peak. This corresponds to level of service rating C (LOS C).

[7] See http://www.sandag.cog.uc/i-15fastrak/schedule.html for additional details.

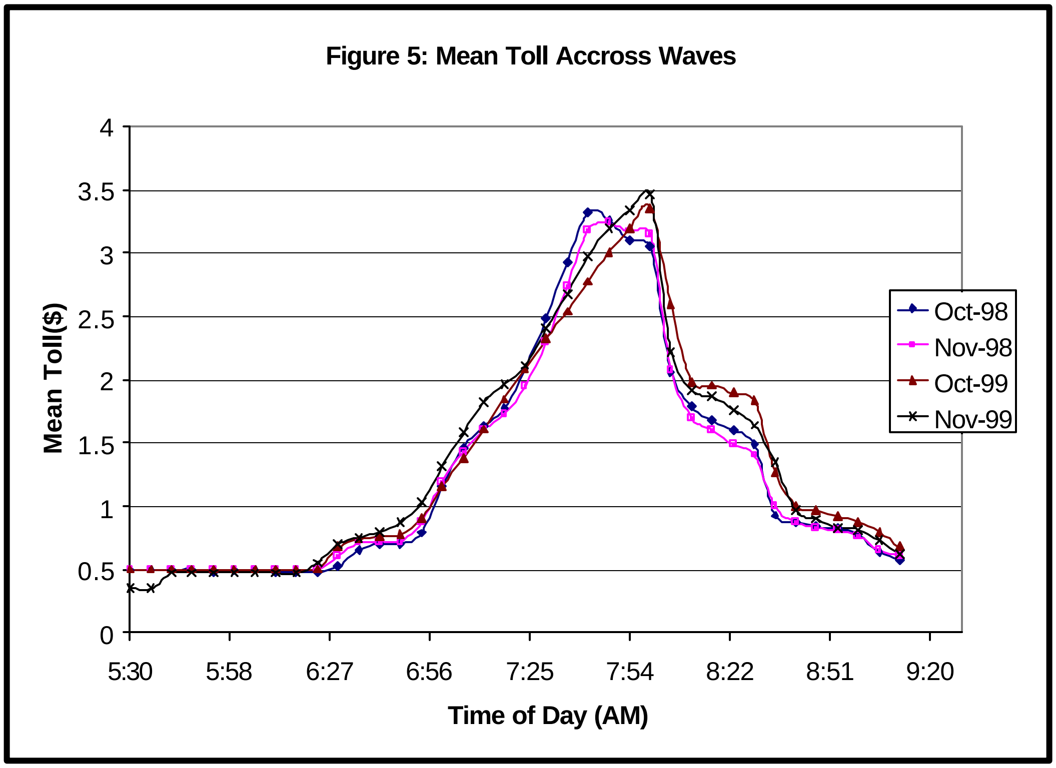
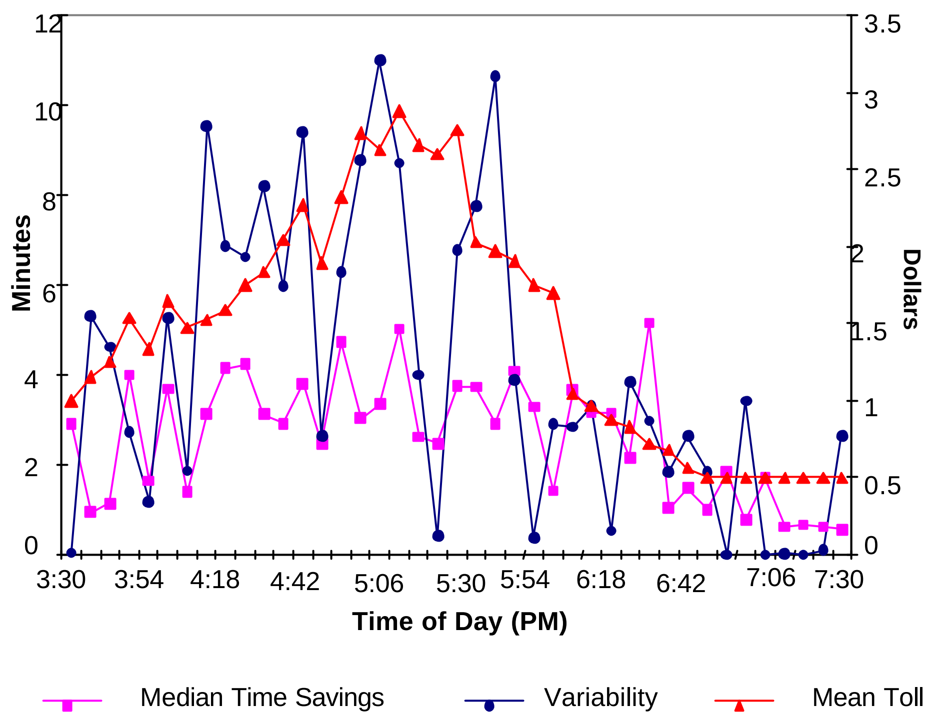


Figure 5: Mean Toll Accross Waves
Mean Toll($)
4
3.5
3
2.5
2
1.5
1
0.5
0
5:30
5:58
6:27
6:56
7:25
7:54
8:22
8:51
9:20
Time of Day (AM)
Oct-98
Nov-98
Oct-99
Nov-99

2.2.4: TIME SAVINGS, TOLL AND VARIABILITY FOR THE AFTERNOON COMMUTE

Since time savings and toll are also used for the afternoon commute for wave 3, Figure 6 presents median time savings, toll and variability for the afternoon commute. The median time savings is presented over the entire two-month period (Oct – Nov 1998). The picture is not at all well defined like the morning trip, and it shows a lack of well-defined peak in the afternoon.

Figure 6 :Median Time Savings, Variability and Toll for Afternoon



CHAPTER 3

PASS AND MODE CHOICE MODELS FROM REVEALED PREFERENCE DATA

Each traveler on the I-15 has three alternatives available during peak hours: 1) solo driving on the main lanes, 2) solo driving using FasTrak, and 3) driving with others in a carpool. I model travelers' choices with a conditional logit model where all three alternatives are available to each traveler. Note that I have treated departure time from home as fixed. It is possible that commuters change their departure times in response to changes in tolls and congestion levels, but I do not have sufficient data to jointly model departure time and mode choice. Incorrectly assuming fixed departure times should not bias the value of time estimates since the congestion pricing scheme insures that the toll per minute of time saved is roughly constant over the morning commute period.

Of course a traveler can only legally use FasTrak if they first obtain a transponder or pass. So pass choice is included as an extra dimension of choice. In chapter 3.1 mode and pass choice models for morning commute are estimated from Wave 5 data. In 3.1.2 a conditional logit model of mode and pass choice is presented using corrected choice-based weights. Construction of these weights is outlined in 3.1.1. A common criticism against the multinomial logit model is the restrictive assumption of Independence of Irrelevant Alternative (IIA). Hensher (2000) argues that the estimated coefficients and value of time estimates are sensitive to model specification, specially to relaxation of the IIA assumption. So to check the stability of the model several non-IIA models are estimated. The results from these models are presented in 3.1.3. In 3.2 a different approach is taken to check the stability of the estimates. A similar mode and pass choice model is estimated for wave 3 which was conducted a year earlier than wave 5. Thus this tests whether the estimates are stable over time. An additional feature of this model is that it models morning and afternoon commute as a joint decision process. This is a unique modeling approach which allows us to observe similarities and differences between morning and afternoon mode choice decisions and allows us to controls for unobserved heterogeneity.

From the different models it can be concluded that the estimates are pretty stable across models. High income, middle aged and female commuters are more likely to own a pass and use it. Carpooling is probably done with family members. Time Savings has a positive and Toll has a negative impact on FasTrak use. The effect of variability depends on the actual level of toll prevailing on that particular day. It is positive if the toll on that particular day is higher than a threshold level. The reason for such dependence is because toll also acts a signal for congestion and commuters are more likely to take the FasTrak lane if the toll exceeds their expected amount. Failure to control for the signaling effect of toll can seriously bias the VOT estimates.

Morning and afternoon commute decisions are very similar except with respect to Variability. Morning commuters are sensitive to a reduction in uncertainty of travel time depending on the time of travel (peak travel is more sensitive than non-peak travel) and the level of toll (it is positive only when the toll is rising). Afternoon commuters are encouraged to take FasTrak if there is reduction in uncertainty but the effect is positive only up to a certain level of toll. This shows more tolerance for uncertainty in afternoon than in the morning.

3.1: MODE CHOICE MODELS FOR MORNING COMMUTE: USING RP DATA FROM WAVE 5

For this part I will be using survey data from Wave 5 (fall 99). The survey data will be complemented by network data as described in the previous chapter. The first model is a conditional logit of pass and mode choice. Results from both unweighted and weighted (correction for choice-based sampling) are presented. To test the stability of the estimates, the IIA assumption is relaxed and several models with flexible error structure is presented.

Table 2 describes the generation of the dependent variable and summarizes it for the sample to be used in estimation.

TABLE 2: THE DEPENDENT VARIABLE

Choice	**Definition**	**Frequency**	**Percent**
No Pass-Solo	Not a registered user and drove alone in regular lane	216	34.89
No Pass-Carpool	Not a registered user and carpooled	86	13.89
Pass-Solo	Registered User and drove solo in regular lane	68	10.99
Pass-FTP	Registered User and drove solo in FasTrak lane	221	35.7
Pass-Carpool	Registered User and Carpooled	28	4.52
	Total	619	100

I assume that the commuter has 5 mode choices (specified above) and she maximizes the following utility function by choosing particular mode:

$$U_{in} = V_i(X_n, Z_{in}) + e_{in}$$

Where i denotes model choice (i=1,..,5) and n denotes individual (n=1,…,619)

X is a vector of demographic characteristics like income, age which are assumed to be exogenous. Note that these variables vary across individuals but not across choices.

ε_{in} is a random variable which follows an extreme value distribution.

Z is vector of travel attributes which varies across both individuals and choices. The travel characteristics I look at for this model are:

1. **Median Time Savings**: Section 2 has described in detail how this measure was generated. The rationale for including this variable is that since commuters in this corridor are experienced travelers, they are aware of the distribution and they base their decisions on what they expect on average. Time savings is positive for the alternative of taking the FasTrak lane (solo or carpool) and zero otherwise.

2. **Toll:** The toll is positive for the alternative 'commuter drove solo in the FasTrak lane' and zero otherwise.

3. **Variability:** This measure was introduced to capture the uncertainty in traveling on the regular lane. This is defined as difference between 90^{th} and median time savings for the FasTrak alternative and zero otherwise. I expect the coefficient of this variable to be positive implying that the commuters dislike extra variability in their travel time. Other measures of variability, like standard deviation, were also used but this measure yielded more precise results.

4. **Toll*Variability**: This interaction term was introduced following the modeling approach in Brownstone et al (2001) and Brownstone et al (2000). One of the goals behind selection of the generic variables was to mimic the information set the traveler might have used when making the mode choice decisions. The percentile measures of time savings captures the distributional aspect and toll captures current road conditions. It is however reasonable to argue that their effects are not independent of each other. Suppose the commuter has some estimate of the distribution of travel time and on a particular day

faces severe congestion which would prompt her to update her estimate of 'bad delays' or what I have tried to capture by the variability term. But the toll by construction will also reflect such a bad day and hence will influence her updating process. It is this feedback process that I have tried to capture by the interaction term. This makes the effect of one conditional on another and helps in controlling the collinearity that exists between all these variables. It also helps introduces observed heterogeneity by making VOT a function of the level of variability and VOV a function of toll.

5. **Deviation of Actual Toll from Mean Toll**: The computation of toll makes it highly correlated with the level of congestion in the regular lanes. Commuters cannot see the level of congestion ahead but they can see the toll well before the lane starts. Being experienced travelers they will correctly interpret a high toll as a signal of congestion ahead and may therefore decide to use the FasTrak lane. Like toll, this variable is defined for FasTrak alternative and zero otherwise. Although we expect the effect of toll to be negative that is fully captured by the variable 'toll'; therefore this term captures the signaling effect of toll and is expected to be positive.

3.1.1: Choice- Based Sampling Weights

Though the Choice-based sampling technique helps us in estimating the parameters precisely, it also means that any inference has to be weighted by the correct weights. The standard maximum likelihood estimator is based on the assumption that the sampling method is random. Maximizing the likelihood function with choice based sampling lead to inconsistent estimators. McFadden (proof in Manski and Lerman (1977)) shows that in the case of a conditional logit model with full set of alternative specific constants, only the constants are inconsistent.

As a possible solution Manski & Lerman (1977) suggested estimating a weighted likelihood function. Suppose Q_n (i) is the fraction of the population selecting choice i and H_n (i) is the proportion for the sample. Then they estimated the likelihood function by the weighting each observation by the following weight W_n (i) = Q_n (i)/H_n (i). Their Weighted Exogenous Sample Maximum Likelihood Estimator (WESMLE) is the maximand of the weighted likelihood function.

Though WESMLE is consistent, it is not efficient. Imbens (1992) proposed an efficient method of moment estimator for such problems that is computationally less burdensome than previous efficient estimators like that of Cosslett (1981). I also implemented Imbens' estimator and was able to replicate his Monte Carlo study. But when it was applied to higher dimension model like in table 4 several problems arose. First, there were several numerical problems with convergence, which were detected due to the fact that one of the key assumptions was violated. In those cases where the numerical problem was solved and convergence was reached the standard errors were larger than the WESMLE case. Therefore, while the Imbens estimator might be asymptotically efficient, it is not better than the WESMLE for these data and models.

I do not have the population count by the five mode choices I have specified. The data that is available is a count of population by broad mode choices for morning commute (i.e. Solo, FTP and Carpool). The data was collected for 5 days in fall 98 and 99 where the number of people driving solo, using FasTrak or carpooling was measured. I will use the count data for fall 99 to illustrate the derivation of the weights.

The population share is given for:

Mode	Population Proportion
PASS-FTP	.097
Solo	.793
Carpool	.11

Let Q_i be the population for the i^{th} mode. (i=1,..,5)

Thus we know

$Q_4 = .097$ ..(1)

$Q_1+Q_3=.793$ ..(2)

$Q_2+Q_5 =.11$ ..(3)

Also note that the sample generated for pass holders and non-pass holders are random. This implies that within the pass holders the proportion of people driving solo or FTP or carpool is consistently estimated. The same logic applies for people without pass.

From (1) we have

$$Q_4 = \quad \text{Prob.(FTP|Pass)*Prob.(Pass)=.097}$$

Since Prob. of FTP given pass from the FTP sample, replacing the value

in the above equation we get

$$\text{Prob.(Pass) = .14} \qquad \text{(A)}$$

Similarly $\quad Q_3 = \text{Prob.(Solo|Pass)* Prob(Pass)}$

Since we know the proportion of Solo drivers within the FTP sample and the probability of holding a pass, replacing the values in the above equation we get.

$$Q_3 = (.215)*(.14) = .0301 \qquad \text{(B)}$$

Replacing the value from (B) in (2) we get

$$Q_1 = .763$$

Similarly we can solve for Q_2 and Q_5.

After solving, the values are summarized in the following table.

TABLE 3: PURE CHOICE-BASED WEIGHTS

Mode	Population Proportion (Q)	Sample Proportions (H)	Weights W=Q/H
No Pass-Solo	.763	.35	2.18
No Pass-	.098	.139	.707

Carpool			
Pass-Solo	.031	.11	.283
Pass-FTP	.097	.357	.272
Pass-Carpool	.0123	.045	.272

We will refer to this as the pure choice-based weights. These weights actually capture the inverse of the probability that a respondent is included in the sample. Thus, a respondent who is using FasTrak for all 5 days and someone who uses it for, say 2 days and drives solo for the other 3 days should not have the same probability of being included in our sample. The latter commuter should be more likely to be in our sample. However, since the current design does not distinguish between the two, the weights are modified. The survey asks the respondents about their mode choice for an entire week. Using these responses new weights are computed as follows:

$$(W\_adj)_n = \sum_i N_{ni} W_i \Big/ N_n$$

Where the subscript n stands for the individual and i stands for choice.

N_{ni}: Number of time n^{th} individual uses i^{th} mode.

N_n: Total number of trips made by the individual for the entire week.

The weights were adjusted so that their sum equals the sample size and I will refer to these weights as the adjusted choice-based weights.

3.1.2: ESTIMATION RESULTS FROM CONDITIONAL LOGIT MODELS

Table 4 presents the result from the Conditional Logit Model (both weighted and unweighted). For the unweighted model, the constants are inconsistent but all other parameters are consistently estimated. Most of the parameters are precisely estimated. Results from the unweighted model are discussed first. Carpooling is done by households with more than two workers, higher worker per vehicle and home owners and commuters who are between the age 35 and 45. Females, college educated, and high income commuters are more likely to have a pass and carpool. This indicates carpooling is probably done with household members. A medium commute distance is most favorable for carpooling, a result also found by Brownstone and Golob (1992). Households with higher worker per vehicle, home owners and those between 35 and 45 are more likely to hold a pass but its use is influenced by high income, the age group 35 to 55 and homeownership.

The coefficients of the generic variables in Table 4 are somewhat complicated to interpret due to the interaction term between reduction in variability (measured as the difference between the 90th percentile and median of the time savings distribution) and toll. This interaction term cannot be deleted from the model without significantly reducing the log likelihood and significantly increasing the standard errors of the toll and time savings coefficients. Similar results were obtained using different definitions of variability such as standard deviation. I also tried interacting the generic variables with the demographic variables, but found no significant interactions. In general, increased time saving is associated with a higher probability of choosing the HOT lane.

The interaction between toll and variability is positive. This implies that the effect of toll is negative only if variability is below 7.8 minutes. Variability will be above 7.8 minutes when the regular lanes are completely congested and commuters are willing to pay very high toll to use the lane. Variability exceeds this level for roughly 3% of my sample.

The opposite interpretation is true for variability. It has a positive influence on FasTrak use only if the toll rises above a certain level (approximately $1.20). The toll is above $1.20 for 60% of my sample. For any tolls below this level, Variability does not encourage commuters to use the FasTrak lane. A priori one would have expected that any level of variability would encourage use of the HOT lane. But to understand this apparent counterintuitive result one has to understand the dynamic toll in this case a reflection of the true congestion level on a particular day which is not known to a commuter. She knows the distribution of time savings and how variable or uncertain it is. She also knows that if a day she is travelling on is a particularly bad day (a draw from the 90th percentile) then that would be reflected by the toll. Thus when the toll is below a threshold level ($1.20 in this case) the probability that it is a 'bad' day is small and thus does not positively influence FasTrak use. This is further supported by the signaling effect of toll and is explained further in the next paragraph.

The positive and significant sign of the 'deviation of toll from its mean value' confirms the a priori expectation that commuters are interpreting toll as a signal for congestion. Thus if the toll rises above the level the commuter 'expects' (measured as the mean toll existing at that 'time of day') then it positively influences FasTrak use. Thus it further reinforces the idea that commuters are using toll to extract information about current road conditions. Apart from the interesting behavioral implication it is important to control for this effect to get an unbiased estimate of VOT. If it is not controlled for then the VOT estimates would be biased upwards which will be apparent when the VOT estimates are presented.

The weighted estimation (WESMLE) was performed using adjusted choice-based weights. The first noticeable feature is the uniform increase in standard errors of the estimates. This is an obvious outcome of the weighting scheme. As a result, the level of significance has dropped for almost all the estimates. All the generic variables, except variability, are still significant. Though these estimates are consistent, they are not efficient and since our interest is not in the alternative specific constants, there is little gain from estimating the WESMLE.

TABLE 4: ESTIMATION RESULT FROM THE MODE & PASS CHOICE MODEL

Base-No Pass-Solo	Unweighted		Weighted (WESMLE)	
	Coefficient	Std. Err.	Coefficient	Std. Err.
No Pass-Carpool				
Constant	-3.919	1.088	-4.235***	1.453
High Income & Don't Know/ Refused	-0.248	0.302	-0.508	0.392
Low Income	0.612	0.457	0.785	0.578
Distance	0.132**	0.052	0.112	0.072
Distance squared	-0.002**	0.001	-0.001	0.001
Worker Per Vehicle	1.062**	0.429	1.012*	0.557
Single Worker Household	-1.392***	0.385	-1.630***	0.527
Two Worker Household	-1.003***	0.323	-1.132***	0.427
Home Owner	0.665*	0.373	0.760	0.505
Age4555	-0.655**	0.311	-0.559	0.420
Pass-Solo				
Constant	-3.190	0.592	-5.201***	1.389
High Income	0.418	0.294	0.421	0.627
Low Income	-1.584	1.059	-1.509	2.641
Worker Per Vehicle	1.158***	0.440	1.122	0.931
Home Owner	0.866**	0.443	0.845	1.060
Age between 45 & 55	0.794***	0.270	0.908	0.614
Pass-FTP				
Constant	-1.987	0.418	-3.889***	0.898
High Income	0.968***	0.217	0.893**	0.383
Low Income	-1.014	0.664	-1.144	1.621
Home Owner	1.109***	0.333	1.163	0.741
Age between 35 & 55	0.633***	0.216	0.694	0.451

Pass-Carpool				
Constant	-3.818	0.501	-5.841***	1.169
High Income	0.900***	0.431	0.721	0.975
Low Income	-0.715	1.089	-0.806	2.758
Female	1.156***	0.422	1.212	1.009
College	0.809**	0.399	0.938	0.940
Generic Variables				
Median Time Savings	0.230***	0.036	0.214***	0.050
Toll	-0.534***	0.201	-0.674*	0.378
Variability	-0.07735**	0.042	-0.054	0.061
Toll*Variability	0.0773***	0.030	0.084*	0.052
Actual Toll-Mean Toll	0.794**	0.313	1.216**	0.607
Number of obs	619		619	
LR chi2(31)	486.15		1042.71	
Prob > chi2	0.00		0.00	
Pseudo R2	0.244		0.5243	
Log likelihood	-753.167		-472.968	

* indicates significance level (*=10%, **=5%, *=1%)

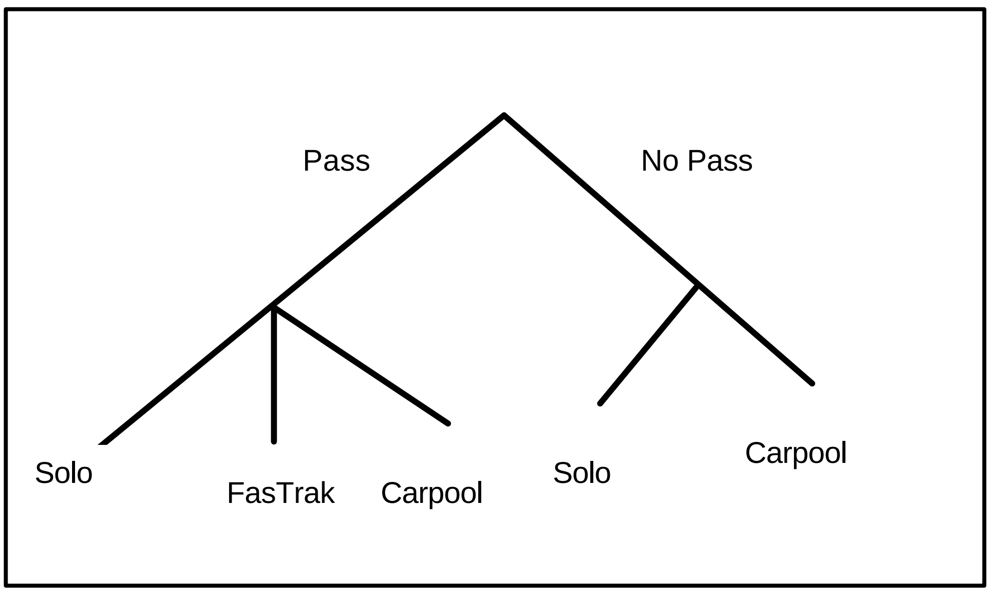
3.1.3. ESTIMATION OF NON-IIA MODELS

One of the compelling limitations of the conditional logit model is the "independence of irrelevant alternatives" property, which is implied by assuming that the error term is identical and independent across all choices. While there is a formal test of this assumption (due to Hausman), it is valid only under the condition that the coefficient estimates of the restrictive model are efficient. Since none of the above models are efficient, I cannot implement the test. Instead what I will do is estimate several non-IIA models and calculate the value of time estimate from it and check whether the calculation from the conditional logit is sensitive to these specification. In particular I will be estimating a Nested Logit Model, a Heteroscedastic Extreme Value Model and a Mixed Logit Model.

For any model other than the conditional logit, the parameter estimates from MLE are inconsistent due to the choice based sampling scheme. So the WESMLE is applied and consistent estimates are obtained. I will now briefly describe each of these models and discuss the overall results. The value of time calculations will, however, be presented in details in the next sub-section.

3.1.3.A: **NESTED LOGIT MODEL:** The conditional logit model is modified by recognizing the fact that there are common error components across choices. The model structure is presented in Figure 7. The model was estimated but it resulted in unacceptable estimate of the inclusive value. If the underlying random utility model is true then the inclusive value should lie between 0 and 1. But the inclusive value for the Pass case was significantly greater than one. Thus this indicates that although this nest seems logical and rational, it is not supported empirically. Therefore, no value of time estimate from this model will be presented.

FIGURE 7: TREE STRUCTURE FOR THE NESTED LOGIT MODEL



3.1.3.B: HETEROSCEDASTIC EXTREME VALUE (HEV) MODEL

The key assumption of conditional logit that is relaxed in the HEV model (Bhat (1995)) is the assumption of identical variance of the error terms. Though they are still assumed to be independent, error terms for different alternatives are allowed to have different variances. The utility for individual i for alternative j has a heteroscedastic random component.

$$U_{ij} = V_{ij} + e_{ij}$$

Where ε_{ij} has an extreme value distribution with cumulative distribution function (CDF) as:

$F(e_{ij})=exp(-exp(-e_{ij}/q_j)$

$$\text{and } Var(\varepsilon_{ij})=1/6\pi^2\theta_j^2$$

θ_j is the scale parameter for j^{th} choice. Since the scale parameter varies by choice the relative probabilities will not be constant thus relaxing the IIA assumption. For the purpose of identification one of the scale parameters is set to one. I assume that the variances differ only across pass holders and non pass holders. The scale parameter for pass holders is set to one. This is basically similar in spirit to the nested logit model but the advantage of this model is that it does not impose any a priori restriction on the inclusive value. Result from the estimation in presented in table 5a.

3.1.3C: MIXED LOGIT MODEL

The mixed logit model (Train (1996)) allows us to incorporate unobserved individual heterogeneity by allowing some coefficients to be randomly distributed across individuals. Utility from alternative i is denoted by:

$$U_i = \beta`x_i + [\eta_i+\varepsilon_i]$$

where x_i is the vector of observed variables, β is the vector of parameters to be estimated, η_i is the random term with zero mean whose distribution over people and alternative depends on the underlying parameters and observed data relating to alternative i, ε_i is a random term with zero mean that is iid over alternatives and does not depend on underlying parameters and data and is normalized to the scale of utility. Stacking the utilities we have:

$$U = b`X + [h+e]$$

where $V(\varepsilon) = \alpha I$ with known (i.e. normalized) α and V whereas $V(\eta)$ is general and can depend on underlying parameter and data. For Mixed Logits each element of ε is iid extreme value and unlike standard Logit (which assumes $\eta = 0$ $\Rightarrow$ IIA) allows any distribution for η.

Denote the density of η by $f(\eta|\Omega)$ where Ω are the fixed parameters of the distribution. Given the value of η the conditional choice probability is simply logit $L(\eta) = \exp(\beta`x_i+\eta_i) / \sum_i \exp(\beta`x_i+\eta_i)$. Since η is not given the (unconditional) choice probability is given by:

$$P_i = L(\eta)f(\eta|\Omega)d\eta$$

Models of this form are called “mixed logit” since the choice probability is a mixture of logits with f as the mixing distribution. The probabilities do not exhibit IIA and different substitution patterns are attained by appropriate specification of f.

The likelihood function was maximized using simulation. The simulated likelihood function is maximized given a preset number of draws (1000 in this case). All the value of time coefficients, namely time savings, toll, variability and toll interacted with variability, were assumed to be independent and normally distributed over individuals. The estimation result returns an estimate of the mean and standard deviation of the random coefficients. This enables us to estimate a distribution of value of time. For the purpose of this model, value of time is estimated at the mean value and is compared against the conditional logit model. Result from this model is presented in table 5b.

Based on the log-likelihood value the HEV model shows significant improvement over conditional logit whereas the introduction of unobserved heterogeneity does not significantly improve the fit of the model.

TABLE 5A: ESTIMATION RESULT FROM H. E. V. MODEL

Base-No Pass-Solo	Coeff.	Std.Err.	t-ratio
No Pass-Carpool			
Constant	-7.484	3.230	-2.317
High Income	-0.968	0.740	-1.308
Low Income	1.479	0.940	1.574
Distance	0.186	0.110	1.688
Distance Squared	-0.002	0.001	-1.597
Worker Per Vehicle	1.946	1.104	1.763
Single Worker	-3.058	1.430	-2.138
Two Worker	-2.103	1.033	-2.037
Home Owner	1.368	0.824	1.661
Age4555	-1.170	0.867	-1.349
Pass-Solo			
Constant	-5.529	1.264	-4.374
High Income	0.566	0.526	1.075
Low Income	-0.868	1.199	-0.724
Worker Per Vehicle	1.344	0.869	1.547
Home Owner	0.870	0.771	1.129
Age4555	0.791	0.458	1.727
Pass-FTP			
Constant	-4.296	0.861	-4.988
High Income	1.041	0.410	2.538
Low Income	-1.051	0.964	-1.090
Home Owner	1.192	0.559	2.132
Age3555	0.704	0.393	1.792
Pass-Carpool			
Constant	-5.714	0.958	-5.967
High Income	0.788	0.686	1.150
Low Income	-0.934	1.287	-0.726
Female	0.603	0.625	0.966
College	0.872	0.591	1.477
Generic Variables			

Median Time Savings	0.319	0.097	3.289
Toll	-0.766	0.364	-2.104
Variability	-0.078	0.064	-1.220
Toll*Variability	0.087	0.051	1.715
Actual Toll-Mean Toll	1.180	0.556	2.123
Scale Parameters of Extreme Value Distns.			
No Pass- Solo	0.508	0.209	2.429
No Pass-Carpool	0.508	0.209	2.429
Pass-Solo	1.000	0.000	
Pass-FTP	1.000	0.000	
Pass-Carpool	1.000	0.000	
Std. Dev. of Extreme Value Distribution			
No Pass- Solo	2.526	1.040	2.429
No Pass-Carpool	2.526	1.040	2.429
Pass-Solo	1.283	0.000	
Pass-FTP	1.283	0.000	
Pass-Carpool	1.283	0.000	
Number of observations	619		
Log likelihood function	-477.921		
Chi-squared	1032.807		
Degrees of freedom	32		
R-sqrd	0.51935		

TABLE 5B: ESTIMATION RESULT FROM MIXED LOGIT MODEL

	Mixed Logit Model		
Base-No Pass-Solo	**Coeff.**	**Std.Err.**	**t-ratio**
No Pass-Carpool			
Constant	-5.029	1.342	-3.749
High Income	-0.379	0.365	-1.038
Low Income	0.739	0.524	1.410
Distance	0.139	0.062	2.248
Distance Squared	-0.002	0.001	-1.875
Worker Per Vehicle	1.279	0.563	2.274
Single Worker	-1.683	0.470	-3.585
Two Worker	-1.119	0.397	-2.818
Home Owner	0.938	0.457	2.054
Age4555	-0.714	0.394	-1.810
Pass-Solo			
Constant	-5.266	1.072	-4.911
High Income	0.440	0.485	0.905
Low Income	-1.538	1.934	-0.795
Worker Per Vehicle	1.204	0.825	1.459
Home Owner	0.825	0.786	1.050
Age4555	0.938	0.475	1.976
Pass-FTP			
Constant	-4.161	0.845	-4.923
High Income	1.147	0.401	2.861
Low Income	-1.419	1.370	-1.036
Home Owner	1.447	0.689	2.100
Age3555	0.773	0.409	1.891
Pass-Carpool			
Constant	-5.840	0.914	-6.392
High Income	0.904	0.805	1.123
Low Income	-1.137	2.088	-0.545
Female	0.922	0.760	1.214
College	1.153	0.800	1.441
Generic Variables			

Actual Toll-Mean Toll	1.345	0.671	2.004
Random Parameters			
Median Time Savings	0.285	0.079	3.623
SD	0.114	0.191	0.596
Toll	-0.950	0.667	-1.423
SD	0.468	0.540	0.867
Variability	-0.264	0.206	-1.282
SD	0.347	0.230	1.508
Toll*Variability	0.108	0.059	1.844
SD	0.026	0.077	0.333
Number of observations	619		
Log likelihood function	-471.498		
Chi-squared	1045.654		
Degrees of freedom	35		
R-sqrd	0.52581		

3.1.4: VALUE OF TIME (VOT) AND VALUE OF VARIABILITY (VOV)

The implied value of time for these models is a function of variability due to the interaction term between toll and variability. The value of time for these models is calculated as:

$$VOT = \left(b_{TIMESAVINGS} \middle/ b_{TOLL} + b_{TOLL*VARIABILITY} * Variability \right) * (-60)$$

An implicit assumption in the calculation of both VOT & VOV is that the difference between mean toll and actual toll remains constant when toll changes. So, change in toll in this case actually means a change in the toll schedule. All the estimates have been summarized after weighting by the adjusted choice-based weights to match it to the population.

Estimates of the VOT from all the different models are presented in Table 6. It has three parts. In the first part VOT estimates from the models are summarized. This represents the observed heterogeneity. Part 2 of the table represents estimation error of the parameters. Coefficients from the conditional logit model were bootstrapped 1000 times and the median is summarized. This meant repeatedly drawing the value of time coefficients from a joint normal distribution and storing the median value for each draw. For the heteroscedastic model and mixed logit model only the point coefficients are presented in part 1. The third part shows the variation due to the introduction of unobserved heterogeneity from the Mixed Logit model. 1000 draws were made from a normal distribution with mean as the estimated mean coefficients and standard deviation as the estimated standard deviation. For each of the draws the 25th, median and 75th percentile values were stored. Column 1 of part 3 presents the median of these values and column 2 presents the standard deviation of these values over 1000 draws. Thus columns 1 presents the observed heterogeneity and column 2 gives us an idea how large is the unobserved heterogeneity relative to the observed heterogeneity.

The estimates do not seem to vary much between these models. If we look at the range of the median values from the unweighted model then we will see that it reasonably covers all the median values from the other models. Generally the ranges almost coincide with each other. Having said that, the range of values is very interesting. It shows that there is a lot of heterogeneity in value of time. With an average wage rate of $50, the VOT as a proportion of wage rate can vary from a low of 34% to a high of 145% (approximately). At the median values it varies from 50% to 75%. Lam and Small (2000) estimated it to be 58%, which is well within the range of my estimate. So two separate studies for two different corridors are giving us consistent results.

The estimates from the Mixed logit model (at mean coefficient), presented in part 3 of table 6, are uniformly lower than the estimates from the other three models and this is a clear indication that there is indeed a bias factor affecting the estimates from multinomial logit models. Interestingly the bias factor seems to be in the opposite direction of what has been suggested by Hensher (2000).

A point to note is that the percentile values in part 3 are different from the ones presented in part 1 (column 4) of table 6. The correct way to compare the two estimates is the following. Consider the median estimate. Column 4 in part 1 presents the median of VOT estimated at the mean coefficients. Column 1 of part 3 represents the median of all the median values over 1000 draws. The non-linearity results in the difference in values. Note that VOT estimates from the mean coefficients are not that different from a MNL model. However when we explicitly vary the coefficients by taking into account the standard deviation the range of estimates is quite large. However, note that the estimates of standard deviations are not precisely estimated. It is perhaps an indication that after controlling for observed heterogeneity with all these travel attributes and demographics, the component of unobserved heterogeneity has been well controlled for.

TABLE 6: VALUE OF TIME ($ PER HOUR)

Part 1: VOT estimates from the coefficients of the model or observed heterogeneity				
Percentiles	MNL Model - Unweighted	MNL Model - Weighted	Heteroscedastic Extreme Value Model	Mixed Logit Model (At Mean Coefficient)
25%	25.67	19.70	24.75	18.00
50%	37.88	26.99	33.30	24.23
75%	73.15	43.16	51.25	37.31
Mean	56.58	26.09	27.77	20.27
Std. Dev.	64.01	35.98	51.86	37.57

	Part 2:Bootstrapping the coefficients - Estimation Error	
Percentiles	MNL Model -Unweighted	MNL Model -Weighted
25%	29.31	16.82
50%	35.78	23.60
75%	45.51	33.93
Mean	39.16	23.09
Std. Dev.	35.91	57.96

Part 3: Unobserved Heterogeneity - Mixed Logit Model		
Percentiles	Median Values for 1000 draws (observed heterogeneity)	Standard Deviation of drawn values (Unobserved heterogeneity)
25%	12.74	36.34
50%	21.22	22.56
75%	34.36	35.33

The dynamic toll in this case acts as a signal for congestion. How important is it to control for? To get an idea I have presented VOT estimates with and without controlling for the effect in Table 6A. The first two columns presents estimates from weighted and unweighted MNL estimates with the term ‘actual minus mean toll’. The third column represents estimate from a model without this term and the fourth column summarizes the median value from bootstrapping the coefficients for 1000 draws.

The first thing to note is that if we do not control for the signaling effect of toll then the VOT estimates will be biased upwards. In general these values are 30-40% higher than the models which control for this effect. Thus it is fairly important that this effect is well controlled for.

TABLE 6A: VALUE OF TIME ($ PER HOUR) WITH AND WITHOUT CONTROLLING FOR SIGNALLING EFFECT OF TOLL.

	Controlling for signaling effect.		**Not Controlling for the signaling effect of Toll**	
Percentiles	MNL Model - Unweighted	MNL Model - Weighted	At Parameter Estimates	Bootstrapping the coefficient and summarizing the median
25%	25.67	19.70	39.04	33.62
50%	37.88	26.99	49.32	44.81
75%	73.15	43.16	87.60	61.93

Value of Variability is the amount of money a commuter is willing to pay for a reduction in uncertainty by a marginal amount. It is expressed as dollar per

$$VOV = \left(b_{VARIABILITY} + b_{TOLL*VARIABILITY} * Toll / \quad \right) * (\ 60)$$

hour for convenience because variability, like travel time, is measured in time units. VOV is computed by taking the ratio of the derivative of the utility function with respect to toll and variability. Thus the interaction term is included in the calculations and it is multiplied by 60 to convert it from dollars per minute to dollars per hour.

The VOV is computed only for the unweighted conditional logit model since I have already shown that the VOT are pretty stable over different models. The values are weighted by appropriate sampling weights. Table 9 summarizes the result. As before, to check for sensitivity, the coefficients were bootstrapped and the median value has been summarized. The median VOV has quite a tight band and it roughly equals 70% of wage rate on an average. Note that these values are close to the estimate of approximately $32 by Lam and Small (2000) (for females). The ratio of VOV to VOT almost equals 100% which is again within the range estimated by Sena (94).

TABLE 7: VALUE OF VARIABILITY ($ PER HOUR)

	Model	Bootstrap (median values summarized)
10th Percentile	0.00	24.60
25th Percentile	25.67	29.31
Median	34.43	35.78
75th Percentile	55.48	45.51
Mean	49.88	39.16
Std. Dev.	60.92	35.91

Note

Model = Estimated at point estimates

Bootstrap = Estimation error of median VOV

Thus this analysis shows that non-IIA models are indeed useful for capturing the unobserved heterogeneity. But it will be wrong to assume that a simpler logit model would be less efficient at capturing this variation. Carefully constructed measures of travel characteristics will help us in observing much of the heterogeneity and may thus reduce the unobserved heterogeneity to a smaller proportion. Interacting the travel characteristics with demographics will enable us to introduce more variation into these estimates. It is this modeling approach that we consider in the next section.

3.2. MORNING AND AFTERNOON MODE CHOICE MODEL FROM WAVE 3 RP

DATA

In the previous section I have shown that conditional logit models with carefully constructed travel attributes and observed heterogeneity can produce reasonable heterogeneity in VOT and VOV estimates. In this section I will explore the issue of heterogeneity further and estimate a mode choice model of morning and afternoon commute using survey data from wave 3 (fall 98). This will enable us to examine the stability of these VOT and VOV estimates. Since the survey corresponds to the same time (October and November) seasonality is not an issue.

Morning and afternoon commute are modeled as a joint decision process. Most of the earlier studies focus either on the morning commute (Abu-Eisheh and Mannering 1987, Cascetta et al 1992, Southworth 1981, Swait and Ben-Akiva 1987) or the afternoon commute (Mannering and Hamed 1990). Only a handful of studies consider both morning and afternoon commute (Bates 1990, Calpice & Mahmassani 1992). Even in those cases the mode choices are kept separate. As a departure from earlier studies I will estimate a model which examines the process of joint determination. The motivation for such a modeling approach is twofold. The tolls charged for using this 8 mile long road can go up to $4 (it was later revised to $5) per use. Thus for commuters using the toll road on a regular basis this might add up to quite a considerable expenditure. So a budget constraint may be driving the dependence. Secondly, the morning and afternoon commute gives us an opportunity to examine choices made by the commuter at two different points in time. Modeling them together will help us in controlling for some of the unobserved heterogeneity. Some initial testing also suggests that the two choices are related.

3.2.1: GENERATION OF THE MODE CHOICE VARIABLE

Initially models of inbound (or morning) and return (or afternoon) mode choice were estimated for all users assuming the choice of FasTrak Pass (or the transponder) is exogenous to the model: those who would use FasTrak have already obtained the pass. But some specification testing suggested that it is endogenous. So like in the previous section the pass choice was combined with inbound and return trip information to generate the dependent variable. The dependent variable was generated by combining three responses from the commuters. The first question relates to the status of their FasTrak membership. The second and third questions relate to their mode choice for inbound and return commute. For the inbound trip the commuters were queried in details as to their mode choice. The question on afternoon mode choice simply asked them whether they used FasTrak for their afternoon commute or not. Table 8 summarizes the dependent variable.

Commuters who answered that they drove in the HOT lane without a FasTrak pass (which is illegal) were left out of analysis (approximately 7%). Also the few carpoolers who drove in the regular lanes were excluded. Due to lack of information for carpoolers for the afternoon trip, I assume that commuters who carpooled for inbound trip did the same for outbound trips.

The initial sample size was 1130 respondents out of which 200 said that they did not make any trips for the past two weeks. From the remaining 933 majority of the exclusion came from two sources. The respondents were asked the time when they reached the beginning of the HOT lane in the afternoon. Though most of the respondents answered it, 253 did not recall the time. Since this information was crucial in estimating the time savings and toll for each respondents, they were left out of the analysis. Travel time data are also missing for a few dates in December 98. After taking into account all these factors we have 458 observations with complete information which is 45% of the original sample size.

Table 8: Explanation of the Mode Choice Variable

Mode Choice	Frequency	Percent	Description
No Pass-Solo-Solo	133	28.85	Not registered user and driving solo both ways
No Pass-Carpool-Carpool	47	10.2	Not registered user and carpooling both ways
Pass-Solo-Solo	39	8.46	Registered user and solo driver both ways
Pass-Solo-FTP	26	5.64	Registered User: solo driver for inbound trip and FasTrak user for return trip
Pass-FTP-Solo	53	11.5	Registered User: FasTrak user for inbound trip and solo driver for return trip
Pass-FTP-FTP	154	33.41	Registered User and FasTrak user both ways
Pass-Carpool-Carpool	9	1.95	Registered User and carpooling both ways
Total	461	100	

3.2.2: MODEL AND ESTIMATION RESULTS

Results from the estimation of conditional logit model are presented in this section. Several specifications were tried and the one with highest explanatory power is presented. Section 3.2.2.1 describes the independent variables. Section 3.2.3 describes the generation of choice based weights. Section 3.2.4 discusses the results.

3.2.2.1: INDEPENDENT VARIABLES

The independent variables can be categorized into 3 different sets. The first set is the demographic variables. The second set consists of travel dummies which control for the time of travel. The third set variables are generic and they in turn are of two broad classes. The first type of generic variables (for e.g. combined toll and time savings) are meant to capture the common links between the morning and afternoon commute. The second set of generic variables pertains in particular to either morning or afternoon commute (for e.g. toll interacted with variability or deviation of toll from its mean value).

1. **Combined Toll:** I have already discussed in Section 2 how the morning and afternoon toll was collected and matched up to the respondents. The combined toll was constructed as:

$$\text{Combined Toll} = I_{FTP}\,(\text{Median Time Savings for the Morning}) + R_{FTP}\,(\text{Median Time Savings For Afternoon})$$

Where I_{FTP} is a dummy which takes a value of one for using FasTrak in the morning' alternative and 0 otherwise. R_{FTP} is also a dummy and it takes a value of 1 for the alternative using FasTrak lane in the afternoon. Thus the combined toll reflects the total amount the commuter pays for the day's travel. Note that I_{FTP} and R_{FTP} are alternative specific constants and do not denote actual choice.

2. **Combined Median Time Savings*:*** Like the combined toll, the combined time savings is defined as:

$$I_{FTP}(\text{Median Time Savings from Morning}) + R_{FTP}(\text{Median Time Savings from afternoon})$$

Where I_{FTP} is a dummy which takes a value of one for the 'using FasTrak in the morning' alternative and 0 otherwise. R_{FTP} is also a dummy and it takes a value of 1 for the alternative 'using FasTrak lane in the afternoon'. The time savings variable is positive for FasTrak users and carpoolers and zero for solo drivers.

3. Travel Dummies: The time period for both the morning and afternoon commute was divided into broad times-'before peak'; which is before 7:30 AM and before 5:00 PM, and 'peak', which is between 7:30 and 8:30 AM and 5 to 6 PM, and lastly 'after peak'. which is after 8:30 AM and after 6:00 PM. These dummies take the value 1 if the respondent travels during the peak or before peak and zero otherwise.

4. Variability: This measure was introduced to capture the uncertainty in traveling on the regular lane. As before this is defined as Difference between 90^th^

and median time savings for the FasTrak alternative and zero otherwise. This variable is kept separate for morning and afternoon.

5. Toll*Variability: This interaction term was introduced following the modeling approach as in the previous section.

6. Deviation of Actual Toll from Mean Toll: This was introduced as per similar arguments provided for the previous section.

3.2.3: Choice- Based Sampling Weights

I do not have the population count by the seven mode choices I have specified. The data that is available is a count of population by broad mode choices for morning commute (i.e. Solo, FTP and Carpool). The data was collected for 5 days in fall 98 and 99 where the number of people driving solo, using FasTrak or carpooling was measured. I will use the count data for fall 99 to illustrate the derivation of the weights.

The population shares from the traffic count are:

Mode	Population Proportion
Pass-FasTrak	.035
Solo	.808
Carpool	.157

Let Q_i be the population for the i^{th} mode. (i=1,..,7)

Thus we know

$Q_5+Q_6 = .035$ ..(1)

$Q_1+Q_3+Q_4=.808$ ..(2)

$Q_2+Q_7 =.157$ ..(3)

Note that the sample generated for pass holders and non-pass holders are random. This implies that within the pass holders the proportion of people driving solo or FTP or carpool is consistently estimated. The same logic applies for people without a pass.

Thus from (1) we have

$Q_5 +Q_6=$ Prob.(FasTrak-Solo|Pass)*Prob.(Pass)+

Prob.(FasTrak-FasTrak|Pass)*Prob.(Pass) =.035

Or Prob(Pass)(.12+.34)=.035 ….(A)

(Prob. of 'FasTrak-Solo and FasTrak-FasTrak given pass' from the FTP sample)

Thus Prob.(Pass)=.08

Now Q_3 = Prob.(Solo-Solo|Pass)* Prob(Pass)

= (.08)*(.08) = .0301 ....(B)

(.08 is the proportion of Solo drivers both ways within the sample with Pass)

Similarly we can solve for Q_4 and then solve for Q_1 from (2)

Following the same procedure we can solve for Q_2 and Q_7.

After solving, the values are summarized in the following table.

TABLE 9: PURE CHOICE-BASED WEIGHTS

Mode	Population Proportion (Q)	Sample Proportions (H)	Weights W=Q/H
No Pass-Solo-Solo	0.80	0.29	2.76
No Pass-Carpool-Carpool	0.16	0.10	1.55
Pass-Solo-Solo	0.01	0.08	0.08
Pass-Solo-FasTrak	0.00	0.06	0.08
Pass-FasTrak-Solo	0.01	0.12	0.08
Pass-FasTrak-FasTrak	0.03	0.34	0.08
Pass-Carpool-Carpool	0.00	0.02	0.08

I will refer to these as the pure choice-based weights. These weights capture the inverse of the probability that a respondent is included in the sample. Then, a respondent who is using FasTrak for all 5 days and someone who uses it for, say 2 days and drives solo for the other 3 days should not have the same probability of being included in our sample. The latter commuter should be more probable to be in our sample. But the current design does not distinguish between the two. So the weights are modified. The survey asks the respondents about their mode choice for an entire week. Using that response the new weights are computed as follows

$$(W\_adj)_n = \sum_i N_{ni} W_i \Big/ N_n$$

Where the subscript n stands for the individual and i stands for choice.

N_{ni}: Number of time n^{th} individual uses i^{th} mode.

N_n: Total number of trips made by the individual for the entire week.

The weights were adjusted so that their sum equals the sample size and I will refer to these weights as the adjusted choice-based weights.

3.2.4: ESTIMATION RESULTS

The result from estimating the conditional logit model is presented in Table 10. The base case against which all other cases are compared is No Pass-Solo-Solo. Several different specifications were tried and this model fits best. The demographic variables are constructed to capture the effect on pass holders, FasTrak use and carpooling.

The signs of the coefficients match our a priori expectations and most of them are precisely estimated. Owning a pass is positively associated with high income, a single worker household, homeownership, and using the Ted Williams onramp. Women are more likely to hold pass and higher vehicle per driver has a negative effect on pass holding. This latter is a counter intuitive result. Vehicle per driver was introduced as a proxy for wealth and was expected to be positively related to pass holding. A possibility is that this measure is capturing a very high wealth effect. Some of this commuters report almost six or seven vehicle per driver and probably such high level of wealth is associated with flexible work schedule thus reducing the necessity of a pass. The use of FasTrak is influenced positively by age (the middle age group 35 to 55) and a two worker household. The travel dummies suggest that commuters reaching the lane after 6 PM then prefer driving in the regular lane.

Households with higher ratio of worker to vehicle and homeowners are more likely to carpool. Women with children are more likely to carpool. This implies that carpooling is probably done with family members. A medium range distance is most favorable for carpooling, a result also found by Brownstone and Golob (1992). Carpoolers also dislike driving between the morning peak.

The effect of time savings is positive for commuters, as expected. The effect of toll is negative. The effect of variability, or the difference between 90th percentile and median time savings, brings out the difference between morning and afternoon commute. For the morning commute, the effect of variability is positive only if the arrival time at the HOT lane is after 7:30 AM but not before that. This suggests that commuters traveling before 7:30 AM are less sensitive to uncertainty in travel time than commuters travelling after 7:30 AM. This makes sense since the peak starts at 7:30 AM, and people travelling during the peak are more sensitive to variability as they have less cushion time (assuming they are constrained by their arrival time). The term interaction of toll and variability was found to be insignificant for this model. However, note that the period for which variability does not encourage the use of HOT lane is also the time with least amount of toll and it is only after 7:30 AM that the toll starts going up steeply. So the insignificance of the interaction term may be due to high collinearity between toll and time savings. It was better controlled for in Wave 5 due to a better measurement of Ted Williams time savings. For the afternoon commute the effect of variability is positive only up to a threshold level of toll. In this case it is approximately $4.25 which is never reached for this sample (the tolls are capped at $4 for this period). But it shows that the afternoon commute is near its cap. Thus the effect of Variability varies between morning and afternoon commute. For the morning commute variability will encourage use of HOT lane only if either the toll reaches above a certain level or they are travelling in the peak. But for the afternoon commute it encourages use of HOT lane but only after a certain level of toll.

The signaling effect of the toll is captured by the term Actual Toll prevailing on that particular day minus the mean toll for that time-of-day. It is positive, which implies that if the actual toll exceeds the usual toll then people are more likely to use FasTrak. This is a unique feature of the dynamic pricing. Although the commuter cannot see the actual congestion ahead she can infer it from the toll level. Apart from the interesting behavioral implication it is important to control for this effect to get an unbiased estimate of VOT. If it is not controlled for, then the VOT estimates would be biased upwards as shown in the previous section

Note that this model is not corrected for choice-based sampling. This implies that the constants are inconsistent but all other terms are consistently estimated. The weighted estimation results are not presented, but it does not alter any of the key results. An obvious impact of the significant difference in the weights is a substantial increase in the standard errors thus making most of the coefficients insignificant. However, none of these estimates are efficient and there is little gain going from the unweighted to the weighted estimates since we are not interested in constants.

Table 10: ESTIMATION RESULT

	Coefficient	Std. Err.	z
No Pass Carpool-Carpool			
Constant	-7.963	1.576	-5.053
Traveling Between 7:30 and 8:30 AM	-2.258	0.464	-4.862
Pass-Solo-Solo			
Constant	-3.042	0.526	-5.781
Traveling after 6 PM	1.187	0.395	3.004
Constant for **Pass Solo-FTP**	-4.213	0.607	-6.939
Pass-FTP-Solo			
Constant	-3.685	0.602	-6.118
Traveling after 6 PM	1.346	0.349	3.852
Constant for Pass **FTP-FTP**	-2.676	0.623	-4.297
Constant for Pass **Carpool Carpool**	-12.341	1.699	-7.265
Variables affecting Pass choice			
High or Don't Know Income	1.496	0.238	6.281
Female	0.479	0.241	1.988
Vehicle Per Driver	-0.665	0.301	-2.208
Single Worker Household	0.601	0.304	1.979
Home Owner	1.005	0.343	2.931
Uses Ted Williams	0.709	0.254	2.792
Variables affecting FTP choice			
Age between 35 & 55	0.569	0.244	2.331
Two Worker Household	0.500	0.264	1.890
Variables affecting Carpool choice			
Female * Have Children	0.663	0.370	1.794
Distance	0.149	0.079	1.892
Distance Squared	-0.002	0.001	-1.811
Worker Per Vehicle	1.567	0.620	2.527
Home Owner	2.553	0.777	3.286
Combined Generic Variables			

Median Time Savings	0.139	0.040	3.447
Toll	-0.373	0.099	-3.750
Generic Variables for Morning Commute			
Reduction of Variability	0.159	0.076	2.083
Reduction of Variability before 7:30 AM	-0.361	0.077	-4.709
Deviation of Actual Toll from Mean Toll	0.922	0.285	3.238
Generic Variables for Afternoon Commute			
Reduction of Variability	0.195	0.062	3.147
Toll*Reduction in Variability	-0.046	0.025	-1.868
Deviation of Actual Toll from Mean Toll	0.654	0.325	2.009
Number of obs	458		
LR chi2(30)	491.02		
Prob > chi2	0		
Pseudo R2	0.2755		
Log likelihood	-645.7171		

3.2.5: VALUE OF TIME AND VALUE OF VARIABILITY ESTIMATES

The model estimated in the previous sub-section will be used for computing the value of time. The estimates are separated out in terms of morning and afternoon commute. The unweighted model in Section 3.3 is used to make the value of time calculations. To make it compatible to the population of morning commuters, the value of time is weighted by the adjusted choice based weights.

TABLE 11: VALUE OF TIME ESTIMATES ($/HR) FOR MORNING AND AFTERNOON COMMUTE

	Morning		Afternoon	
Percentile	**Model**	**Bootstrap**	**Model**	**Bootstrap**
5th Percentile		15.15	9.35	8.75
25th Percentile		18.45	11.00	11.06
Median	22.36	22.75	13.74	14.06
75th Percentile		28.41	19.21	17.14
90th Percentile		33.89	22.36	20.70
Mean		24.01	14.83	14.46
Std. Dev.		9.01	4.75	4.95

Note

Model = Estimates from the coefficients of the model

Bootstrap = Bootstrapping the estimation error of the median VOT

The table lists each set of estimates under two separate heading. The first one or 'Model' stands for the estimation with the point estimates. For the morning commute there is just one value. To check the sensitivity of these estimates the coefficients were bootstrapped (1000 draws) and the resulting distribution is presented. The range of values represents the estimation error and puts a confidence band around the estimated median VOT. The respondents were not asked about their wage rate in this wave. Table 6 presents an approximation[8] of the hourly wage. Using the mean wage morning VOT can range between 70% to 101% of hourly wage.

TABLE 12: SUMMARY OF HOURLY WAGE

	N	Median	Mean	Std. Dev.
Hourly Wage	951	29	33.5	15.11

The afternoon commute has uniformly lower VOT estimates. Since the afternoon VOT depends on the level of variability (through the interaction term) the model results summarizes the VOT for the sample and the bootstrap presents estimation error of the median VOT. There is a drop of $8 from the morning on an average. The estimate is roughly 40% of wage rate. The morning VOTs are lower than the VOTs presented in table 6A. One of the reason VOTs are lower for this model is that time savings for Ted Williams onramp is not as precisely estimated for this wave (wave 3) as it is for wave 5. Another reason for such lower estimate may also be from due to the fact that the time savings and toll coefficients are constrained to be equal for morning and afternoon commute. Though statistically their equality cannot be rejected, the not-so-satisfactory measure of Ted Williams time savings and the noisy distribution of time savings in the afternoon (figure 6) may have resulted in the lower estimates.

[8] Wage =(Household Income/Number of workers*2000)

Value of Variability is the amount of money the commuter is willing to pay for a reduction in uncertainty by a marginal amount. It is expressed as dollar per hour for convenience because variability, like travel time, is measured in time units. It is computed by taking the ratio of the derivative of the utility function with respect to toll and variability. Thus the interaction term is included in the calculations and it is multiplied by 60 to convert it from dollars per minute to dollars per hour. Table 13 presents the result from the model and the bootstrapped values. The estimates from the coefficients of the model are presented under 'model' and the bootstrapped values put an error band around the estimate. For the afternoon commute the summarized values are bootstrapped median values. Comparing the median values we can say that the value of variability for morning commute is slightly higher than in the afternoon, though the difference is negligible. It is quite high which seem to suggest that commuters are willing to pay quite a high amount to reduce uncertainty.

TABLE 13: VALUE OF VARIABILITY FOR MORNING COMMUTE AFTER 7:30 AM & AFTERNOON COMMUTE

	Morning		Afternoon	
Percentiles	**Model**	**Bootstrap**	**Model**	**Bootstrap**
10%		11.88	5.26	15.33
25%		18.20	13.72	22.93
50%	25.73	25.67	20.20	25.65
75%		34.39	24.54	40.84
90%		44.16	28.19	62.86
Mean		27.03	18.33	36.43
Std. Dev.			8.86	

Note

Model = Estimates from the coefficients of the model

Bootstrap = Estimation error of VOV

CHAPTER 4

VALUE OF TIME FROM STATED PREFERENCE AND REVEALED PREFERENCE ROUTE-CHOICE AND MODE CHOICE MODELS

4.1: STATED PREFERENCE AND REVEALED PREFERENCE ROUTE-CHOICE MODELS

One of the main criticisms of the VOT calculations from earlier studies has been that typically car driving is compared to some other mode, mainly a bus or some form of public transit. This lead to biased estimates due to unobserved factors like comfort and convenience. Stated preference (SP) experiments do not have this problem and are expected to give more precise results. SP analyses like Calfee and Wisnton (1998) and Hensher (2000) have produced estimates that are quite low. Recent congestion pricing projects like the I-15 and SR-91 do not have the same drawback as the earlier RP studies since the choice is between choosing a free and tolled alternative. Studies from these two projects by Brownstone et al (I-15) (2001), Lam and Small (SR-91) (2000), Brownstone et al (I-15) (2000) have all produced estimates significantly higher than the stated preference analyses. The debate over this difference in estimate is an important one. If the lower estimates are true then it will imply that the benefit of such projects is very small and if the higher estimates are right then it will suggest that there are considerable benefits to be gained.

As a part of the wave 5 survey we conducted a simple stated preference experiment. Each respondent was given a hypothetical toll and time savings and asked whether they will use the FasTrak lane or not. With the regular survey questions the following question was asked to all I-15 users (pass holders and non-pass holders).

> **W5q86 [I-15 ONLY:] Suppose** that on a weekday morning trip **you drove alone on I-15**, and the toll for driving alone in the carpool lanes was $______. If paying this toll would save

you ______ minutes, would you pay and use the carpool lanes?

For the I-8 users the same question was asked but in a slightly different manner

W5q87 [I-8 ONLY:] Suppose there was a carpool lane on I-8 about 8 miles long. And suppose that on a weekday morning trip **you drove alone on I-8**, and the toll for driving alone in the carpool lane was $_____. If paying this toll would save you _____ minutes, would you pay and use the carpool lane?

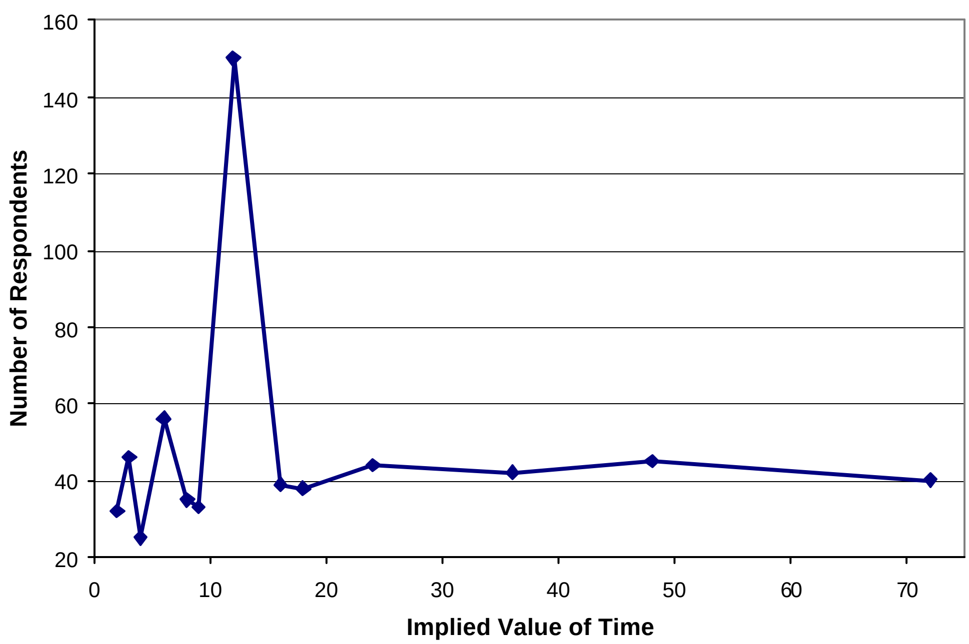
The tolls and time savings were filled by a standard independent orthogonal design with the toll varying between $1 and $6 and the time savings varying between 5 to 30 min. This is a very simplified study compared to those by Calfee & Winston and Hensher. Both these studies were far more elaborate and sophisticated in their approach. Calfee and Winston conducted a mail-in survey for commuters in a US metropolitan area and the different scenarios control for variability. The study by Hensher was a laptop computer based face to face interview with several choice scenarios and measures for variability. The stated preference study used in this section was done over the phone along with the regular survey and does not control for variability.

One of the problems sometimes noted in SP VOT survey design is that the survey by design does not allow for very high VOT. Even when the possibility is there, not enough people are given the option. To ensure that this particular survey does not suffer from this flaw, I have tabulated the hypothetical toll and time savings combination for the I-15 sample and how many people were actually given the option. Figure shows the implied VOT from these hypothetical tolls and time savings on the X-axis and number of respondents given that option on Y axis. From the following table and figure it is obvious that the number of respondents are quite uniformly distributed across different combinations. The highest VOT that can de demonstrated is $ 72 (5 minutes of time savings and $ 6 toll) and quite a fair number of respondents were given that option. The I-8 sample was given similar combinations.

TABLE 14: HYPOTHETICAL TOLL AND TIME SAVINGS FOR THE I-15 SAMPLE

	Hypothetical Time Savings				
Hypothetical Toll	5	15	20	30	Total
1	39	25	46	32	142
3	42	37	33	56	168
4	45	39	34	35	153
6	40	44	38	40	162
Total	166	145	151	163	625

Figure 8: Number of Respondents by Implied VOT from SP



Thus we have three different groups of people with different degrees of familiarity with the congestion-pricing project. The first group is the FasTrak pass holders. Their actual decision of using the toll lane or not enables us to estimate a binary route choice model based on revealed preference data (RP). We also have stated preference (SP) data on the same FasTrak sample and thus can estimate the same route choice model with the SP data and compare the results. The second group are other I-15 users who are familiar with the project but have decided not to use it. For the third group, or the I-8 users, this is a purely hypothetical question. The results across these three different groups will help us in controlling for sample selection bias i.e. where the researcher observes only that group which has decided to take a particular mode and not those who have decided not to use it. This generally tends to bias the results in favor of that particular mode. It has also been one of the explanations for differences in SP and RP estimates.

A binary route choice model is estimated for all the four cases and the results are summarized in Table 15. The same model is estimated for all four cases so as to maintain consistency, though some of the demographic and generic variables are not significant. To compare the VOT estimates across different models it is essential to control for any non-linear effect of time savings or toll. The estimation result presented in Table 16 clearly shows that the effect is non-linear.

Though the SP response is used to estimate the coefficients, to compare VOT across models the RP toll and time savings is used for all samples. For the

FasTrak pass holders and other I-15 users I use the actual toll and time savings they faced. For The I-8 users these were imputed. From the combined I-15 sample the average time to reach the carpool lane was computed by time-of-day. Then given the time the respondents (for I-8) said they reached the freeway, the average travel time was added to compute a hypothetical arrival time at the FasTrak lane. Then the toll and time savings were merged by the hypothetical arrival time. This restricts the range to be same for all samples and make the comparison consistent.

The value of time (VOT) for this model is defined as:

$$VOT = \left(\frac{b_{TIMESAVINGS} + b_{TIMESAVINGS^2} * 2TimeSavings}{b_{TOLL} + b_{Toll^2} * 2Toll} \right) * (-60)$$

The VOTs are summarized in Table 3. Since VOT is a function of toll and time savings these models controls for observed heterogeneity. For each model the VOT estimates are presented under two different headings. The first set of estimates under 'model' is calculated at the coefficient point estimate. The second set summarizes the estimation error in median value of time. This is calculated by bootstrapping the coefficients. VOT from both methods is much higher for the RP model than for the SP models. The question remains as to whether these differences are significant or not. To test the sensitivity of these results the coefficient of time savings, toll and their squared values were bootstrapped. This captures the estimation error of the parameter and helps us in putting a confidence band. The value was stored for 1000 draws and the median value is summarized in Table 16.

It is clear that the FTP sample has slightly higher VOT estimates than the other two groups (in the SP case). This can be interpreted as an evidence of sample selection bias. But the bootstrapped confidence bands for all three groups are pretty close which means statistically the difference is insignificant. The estimates from the RP model are uniformly and significantly higher than the SP values. To get an idea as to what is the ratio of these VOT values to the wage rate, which is a standard convention, I have presented the mean wage rates (as reported by the respondents) for the three different groups in Table 17. Note that these wage rates are from wave 5 and slightly different from wave 3 (which were imputed) presented earlier. To fix some bounds, the lowest estimate from the RP model (using the 25th percentile value) is approximately 37% of wage rate. The lowest estimates from the SP model are 26%, 29% and 40% for the FasTrak, Other I-15 and I-8 samples respectively. The highest value in case of RP can go up to 90% (using 75th percentile value). The maximum values for SP are 31%, 39%, and 48% for the three groups respectively. Comparison of median values as a percentage of wage rate yields roughly 50% higher in the RP case than the SP estimates. Thus the difference between the SP and RP results are real and significant. An interesting point to note is that the SP values we have derived are very similar to the ones derived by Calfee and Winston (1998) and Hensher (2000). The bootstrapped values also show a significantly higher value for the RP case. On average the median values from RP are 50 to 75% higher than the SP models. The only conclusion that can be reached from this analysis is that there is a significant difference between SP and RP experiments, which cannot be attributed to any methodological difference. People respond differently to experimental and actual situations.

TABLE 15: ESTIMATION RESULT FOR THE ROUTE CHOICE MODELS FOR THE REVEALED PREFERENCE & STATED PREFERENCE EXPERIMENTS

(Dependent Variable is 1 if used FasTrak and 0 otherwise)

Positive Coefficient Favors Outcome	Revealed Preference	Stated Preference		
	FasTrak Users	FasTrak Users	Other I-15 Users	I-8 users
Constant	-0.883	1.230	-0.817	-0.150
	(0.935)	(1.195)	(0.936)	(0.836)
High Income Dummy	0.955**	1.023**	0.421	0.316
	(0.353)	(0.353)	(0.299)	(0.342)*
Low Income	-0.193		-0.302	-0.678
	(0.955)		(0.495)	(0.390)
Female	0.559	0.144	0.014	0.968
	(0.358)	(0.343)	(0.302)	(0.287)
Graduate School Dummy	0.367	-0.784**	0.677**	-0.966**
	(0.374)	(0.352)	(0.311)	(0.394)
1 kid under 16	-0.600	-0.158	0.312	-0.414
	(0.409)	(0.405)	(0.345)	(0.387)
Two Worker Household	0.599*	-0.411	-0.206	-0.050
	(0.361)	(0.355)	(0.288)	(0.303)
Age between 35 and 45	0.188	-0.311	-0.232	0.039
	(0.354)	(0.336)	(0.290)	(0.301)
Distance	0.023	-0.012	0.001*	-0.001
	(0.016)	(0.016)	(0.001)	(0.001)
Worker Per Vehicle	-1.969**	-0.977	0.259	0.364
	(0.660)	(0.677)	(0.423)	(0.451)
Home Owner	0.971**	0.301	-0.662*	-0.744**
	(0.501)	(0.548)	(0.368)	(0.337)
Time Savings	0.770**	0.261**	0.330**	0.179**
	(0.215)	(0.080)	(0.081)	(0.070)
Time Savings Squared	-0.038*	-0.003	-0.006**	-0.003

Squared				
	(0.021)	(0.002)	(0.002)	(0.002)
Toll	-1.289**	-0.908**	-1.773**	-0.925**
	(0.621)	(0.443)	(0.343)	(0.327)
Toll Squared	0.209	0.010	0.168**	0.063
	(0.161)	(0.057)	(0.048)	(0.046)
Number of obs	266	306	379	322
LR chi2(14)	66.21	164.98	143.53	95.75
Prob > chi2	0	0	0	0
Pseudo R2	0.2092	0.4001	0.3004	0.2246
Log likelihood	-125.11085	-123.692	-167.097	-165.2778

Standard Errors are in Parenthesis
* denotes level of significance (* =10%, ** = 5%,)

TABLE 16: VALUE OF TIME ESTIMATES FROM THE ROUTE CHOICE MODELS

	Revealed Preference		Stated Preference					
	FasTrak Users		FasTrak Users		Other I-15 Users		I-8 users	
Percentiles	Model	Bootstrap	Model	Bootstrap	Model	Bootstrap	Model	Bootstrap
25%	20.14	29.12	14.84	12.76	12.37	10.94	12.46	9.30
50%	40.58	37.15	16.12	16.08	13.28	12.99	13.17	12.98
75%	51.15	46.77	17.12	20.94	15.55	15.68	14.85	17.52

Note:

- VOT is in terms of $/hr
- Model = VOT estimated at the coefficients
- Bootstrap = Bootstrapping the estimation error in median value of time, at a given value of variable.

TABLE 17: REPORTED WAGE RATE FOR THE THREE DIFFERENT SAMPLES

Reported Wage ($ per hour)	Mean
FasTrak Sample	56.67
Other I-15 Users	40.78
I-8 Users	29.93

4.2: JOINT ESTIMATION OF RP AND SP DATA

The route choice model from SP data yielded VOT estimates which are significantly different from the RP route and mode choice models. The question that arises then why is there such a significant difference and can there be a modeling approach that reconciles these differences. The first approach taken here is to follow a common modeling strategy taken in the literature: that is to jointly model SP and RP or alternatively known as 'data enrichment'. But to do a joint estimation it is essential that a test is conducted as to the validity of 'data enrichment'.

To test the validity of data enrichment a joint model is estimated while controlling for scaling. The RP mode choice model in section 3.1.2 and the SP route choice model in section 4.1 are combined for the joint model. The equality constraint is imposed on toll and time savings coefficient whereas all the other coefficients are allowed to vary across the two datasets. The joint model is estimated using a nested logit specification where SP & RP choices are estimated as separate nests. One of the inclusive values is a set to 1 (RP in this case) for identification. Note that the inclusive value measures the scaling factor. The test is conducted following Hensher et al (2001). The null hypothesis is that the coefficients are homogeneous across the RP and SP dataset. The test statistic is:

$$-2*[(L^{rp}+L^{sp})-L^{joint}] \text{ follows } \chi^2 \text{ with b-1 d.f}$$

where b is the number of common parameters.

The value of χ^2 is 541.6337. With 1 degree of freedom the null hypothesis is rejected. Thus combining the two datasets is not a valid procedure in this case. Though we cannot combine the datasets it should be noted that supplementing the RP with SP data yields more precise coefficient of travel attributes. The effect of variability of savings still depends on the level to toll and it is positive for any toll above $ .40. Since the minimum level of toll is $ .50, the effect of variability is positive for the entire sample. The Value of Time estimates now cover the entire distribution of SP and RP estimates. The lower part of the distribution corresponds with the SP and the upper part corresponds to the RP estimates. Thus the joint model has resulted in a distribution of VOT that can accommodate both RP and SP estimate but as the test suggest that the underlying variation for the two datatset is different and joint modeling is not a viable option in this case. It further accentuates the point that there is a fundamental difference between RP and SP data.

Table 18 : Joint Estimation of SP-RP (Wave 5)

Base-No Pass-Solo	**Coeff.**	**Std.Err.**	**t-ratio**
Revealed Preference Parameters			
No Pass-Carpool			
Constant	-3.994	1.120	-3.568
High Income	-0.526	0.314	-1.673
Low Income	0.803	0.426	1.886
Distance	0.100	0.054	1.856
Distance Squared	-0.001	0.001	-1.550
Worker Per Vehicle	0.859	0.469	1.833
Single Worker	-1.632	0.409	-3.988
Two Worker	-1.133	0.341	-3.324
Home Owner	0.643	0.378	1.699
Age4555	0.117	0.204	0.573
Pass-Solo			
Constant	-5.076	1.058	-4.798
High Income	0.418	0.497	0.841
Low Income	-1.515	1.950	-0.777
Worker Per Vehicle	1.350	0.809	1.669
Home Owner	0.758	0.791	0.958
Pass-FTP			
Constant	-3.747	0.628	-5.970
High Income	0.928	0.306	3.038
Home Owner	1.204	0.549	2.193
Age3555	0.748	0.356	2.101
Pass-Carpool			
Constant	-5.901	0.925	-6.379
High Income	0.714	0.794	0.900
Low Income	-0.647	2.041	-0.317

Female	1.096	0.798	1.373
College	1.058	0.772	1.369
Generic Variables			
Equal Across SP& RP			
Median Time Savings	0.212	0.040	5.332
Toll	-0.995	0.207	-4.800
RP Only			
Variability	-0.049	0.048	-1.021
Toll*Variability	0.109	0.033	3.340
Actual Toll-Mean Toll	1.416	0.462	3.065
Stated Preference Estimates			
Constant	-1.378	0.732	-1.883
High Income	1.189	0.478	2.490
Low Income	-0.940	0.764	-1.230
Female	0.537	0.418	1.286
Graduate Degree	0.441	0.446	0.989
Age Between 35 & 45	-0.728	0.430	-1.695
Home Owner	-1.417	0.555	-2.553
Inclusive Value			
RP	1.000		0.000
SP	0.537	0.111	4.835
SP	0.537	0.111	4.835
Number of obs.	594		
Log likelihood	-741.4452		
Restricted log likelihood	-2276.775		
Chi-squared	3070.66		
Degrees of freedom	38		
R-sqrd	0.67434		
RsqAdj	0.67012		

TABLE 19: THE LIKELIHOOD RATIO TEST FOR DATA ENRICHMENT

Ho: Parameters (Time Savings & Toll) are homogeneous across SP & RP data	
Hi: They are not	
Test Statistic is $-2*[(L^{rp}+L^{sp})-L^{joint}]$ follows Chi-sq with b-1 d.f where b is the no. of common parameters.	
SP Likelihood	-294.92418
RP Likelihood	-717.33787
Joint Likelihood	-741.4452
Chi-sq	541.6337
Thus Null hypothesis is rejected	

TABLE 20: VALUE OF TIME AND VALUE–OF VARIABILITY FROM THE JOINT MODEL

Percentile	Value of Time	Value of Variability
25%	12.78	7.88
50%	16.99	15.72
75%	26.38	31.01
90%	34.42	50.22
95%	34.44	61.76
Mean	11.81	12.50
Std. Dev.	32.26	40.60

4.3: SP MODEL WITH UNOBSERVED HETEROGENEITY

The SP models presented in the previous sections controls for observed heterogeneity but not for unobserved heterogeneity unlike the RP models. As we have seen in case of the RP models, introduction of unobserved heterogeneity can bring in a lot of variation in the estimates. Thus a mixed logit route choice model was estimated using the SP data. As before, the idea is to vary the coefficients used to calculate VOT randomly across respondents and then estimate the distribution of VOT. From an initial model similar to the one in section 4.1 insignificant coefficients were eliminated till all the remaining coefficients are significant. As a result all the demographic variables were eliminated except the dummy for high income. The sample used for this model is all I-15 users (FasTrak users and other I-15 users).

Initially time savings, toll and their squared terms were included with different assumptions about the random parameters. The model that had the highest explanatory power is presented in Table 20A. It is a relatively simple model with time savings and toll. The coefficients of time savings and toll assumed to distributed normally across respondents. All the parameters are precisely estimated and the resultant VOT estimates are presented in Table 20B. The first column presents estimate of VOT at the mean coefficients and is approximately $10 which is consistent with what was derived before. The second and third column is presents the variation due to unobserved heterogeneity. Coefficients of time savings and toll were drawn from a normal distribution with mean as the mean coefficient and variance as the estimated variance. Thousand such values were drawn and the 25th, 50th, 75th percentile values were stored. This process was repeated for 1000 times and the second and third columns list the median and mean of these percentile measures. The fourth column presents the standard deviation of these values. The low value of the standard deviation imply that these percentiles are do not vary a lot. The range of the distribution (interquartile range) is significant but in general the distribution is lower than the RP estimates.

TABLE 20A: SP MODEL WITH UNOBSERVED HETEROGENEITY

	Coefficient	Std. Error	t-ratio
Constant	-0.409	0.366	-1.117
High Income	1.145	0.402	2.846
Random Parameters			
Time Savings	0.224	0.081	2.752
Standard Deviation of Time Savings	0.123	0.058	2.133
Toll	-1.346	0.500	-2.694
Standard Deviation of Toll	0.614	0.324	1.893
Number of Observations	1011		
Log likelihood function	-519.409		
Restricted log likelihood	-700.772		
Chi-squared	362.726		
Degrees of freedom	6		
Significance level	0.000		
R-sqrd	0.259		
RsqAdj	0.254		

TABLE 20B: VALUE OF TIME ESTIMATES

		Summary Values for 1000 draws		
		Median	Mean	Std. Dev
At Mean Coefficient	9.99			
25th Percentile		5.94	5.94	0.17
Median		9.92	9.93	0.19
75th Percentile		15.22	15.22	0.28

CHAPTER 5

EXPLAINING THE DIFFERENCE

What we have learnt from chapters 3 and 4 is that there is a significant difference between the RP and SP estimates. One of the main purposes of SP was to yield more precise estimates. Combined with RP data it can be used to control the high degree of collinearity among the measures of travel attributes. As we saw in section 4.2 such data enrichment process yields more precise coefficient estimates. The other use of SP was to yield estimates in cases where RP data is missing. But the preceding analysis has shown that there is a fundamental difference between how people answer these SP surveys and how they respond to real situations. This is shown by the rejection of the data enrichment process which means that the underlying variance is very different for these two datasets. In this particular case the significant difference between RP & SP VOT estimates may lead to very different conclusions about project evaluation and welfare impacts of congestion tolls.

Thus it is important to try and understand why they are dfferent. This chapter explores the reason for the difference between SP and RP estimates. Several possibilities are explored to try and reconcile the difference between the two. The general conclusion of this chapter is that the differences are irreconcilable and so more care should be taken in interpreting these SP estimates.

5.1: SP MODEL CONDITIONAL ON ACTUAL CHOICE

The first approach taken to reconcile the difference is conditioning the SP model on actual choices made. The idea is that commuters who actually took the toll lane are different and have a higher VOT than people who did not. If the preceding hypothesis is true then their SP answers would also indicate a higher VOT. The model used in this sub-section is the same route choice SP model as in chapter 4.1 but with dummy variables introduced representing actual choices made. Dummy variables for FasTrak and Solo (1 if actual choice was FasTrak and Solo respectively for last trip) are introduced as levels and interacted with toll and time savings to bring out the difference in VOT. The dummy variable for FasTrak use is positive which seem to suggest that commuters who actually used the HOT lane for the last trip are more likely to answer positively to the SP question. Conversely people who did not take the lane are less likely to answer positively to the SP question.

The interaction term between the FasTrak use dummy variable and travel attributes brings out the difference between FasTrak users and others in terms of sensitivity to time savings and toll. The results suggest that commuters who actually used the HOT lane are more sensitive to time savings and toll. Time savings and toll were also interacted with the solo driving dummy but they were insignificant. Though the coefficients confirm our a priori belief about the effect, the differences are not that significant in terms of VOT estimates. Actually contrary to expectation, VOT for people who did not use HOT lane are higher. But a formal test shows that the differences are statistically insignificant. In general the estimates are very similar to the SP estimates derived earlier. Thus conditioning SP choice on actual choices does not account for the difference in VOT.

The same model was estimated with ‘proportion of FasTrak use to total number of trip for the week’ and the results were qualitatively same as this one.

TABLE 21: SP MODEL CONDITIONAL ON ACTUAL CHOICE.

Dummy corresponds equals 1 if actual choice for last trip is FasTrak & Solo respectively.

Positive Favors FasTrak Use	Coefficient	Std. Err.	z
Constant	-1.586	0.563	-2.816
High Income	0.511	0.174	2.930
Low Income	-0.195	0.313	-0.621
Vehicle Per Driver	0.223	0.173	1.290
Female	1.168	0.409	2.853
Single Worker Household	0.325	0.179	1.820
FasTrak Dummy	1.626	0.646	2.518
Solo Dummy	-0.698	0.192	-3.635
Time Savings	0.278	0.046	6.062
Time Savings Squared	-0.004	0.001	-3.898
Time Savings for Female	-0.043	0.020	-2.153
Time Savings for FasTrak Users	0.061	0.028	2.197
Toll	-1.099	0.206	-5.343
Toll Squared	0.072	0.029	2.532
Toll for FasTrak Users	-0.368	0.152	-2.416
Number of obs		995	
LR chi2(14)		448.76	
Prob > chi2		0	
Pseudo R2		0.3291	
Log likelihood		-457.42683	

Table 22: Value of time ($/hr)

	Used FasTrak last trip		Did Not Use FasTrak last trip	
Percentiles	Female	Male	Female	Male
10%	8.44	10.50	7.81	10.74
25%	9.29	11.15	10.22	12.74
50%	10.99	13.18	11.89	14.81
75%	12.52	15.01	14.61	17.81
90%	14.05	16.23	15.91	19.61
Mean	11.19	13.27	12.20	15.31
Std. Dev.	2.31	2.24	3.27	3.55

5.2: THE RP MODEL WITH PERCEIVED TIME SAVINGS

The survey queried commuters on the amount of time they saved (or would have saved) on their last trip by taking the HOT lane. The answers generally clustered around 15 or 20 minutes, which exceeds the engineering estimates for the same time periods. Golob et al (2001) has shown that this perceived time savings is an important determinant of the 'use of FasTrak'. If commuters are taking their decisions based on their perception of time savings which is higher than the engineering estimates then the VOT from the RP model will be biased upwards. That's because their perceived per dollar time savings is higher than the engineering estimates.

Thus to control for this, difference between perceived and actual time savings is introduced in the model and is called 'Excess Time Savings'. I have modeled it as an individual specific variable and not a generic variable. So it is introduced by interacting it with alternative specific constants. The first table summarizes the variable which seem to suggest that most of the respondents overestimate time savings by roughly the same amount except people with transponder who decided not to use the HOT lane (Pass-Solo). They have the closest estimate to the engineering estimates.

The model suggests that the variable has a significant negative impact on Pass-Solo option and a significant positive impact on Pass-FTP option. So people with transponders are more likely to take the HOT lane if their perception exceeds the engineering estimates. The obvious problem of using such a framework is the endogeneity problem. The assumption of perceived time savings being exogenous is probably false. But the purpose of this exercise is to control for this effect and see whether the difference between SP and RP estimates can be reconciled.

The VOT estimates are presented in Table 25. The hope was that by controlling for these factors the RP estimates will be closer to the SP estimates but the estimates are clearly much higher than the SP estimates. A similar model was tried but replacing median time savings with perceived time savings but the results did not change in any fundamental manner.

TABLE 23: SUMMARY OF PERCEPTION ERROR BY MODE CHOICE

Choice	Mean	Std. Dev.	Freq.
No Pass-Solo	9.969	12.246	215
No Pass-Carpool	12.076	10.621	93
Pass-Solo	5.180	5.795	67
Pass-FTP	12.281	10.052	220
Pass-Carpool	11.886	9.516	28
Total	10.671	10.795	623

TABLE 24: RP MODE CHOICE MODEL WITH 'EXCESS SAVINGS'

Base-No Pass-Solo	Coef.	Std. Err.	Z
No Pass-Carpool			
Constant	-3.912	1.091	-3.587
High Income	-0.241	0.305	-0.791
Low Income	0.700	0.466	1.501
Distance	0.133	0.052	2.547
Distance Squared	-0.002	0.001	-2.021
Worker Per Vehicle	0.992	0.433	2.293
Single Worker	-1.408	0.393	-3.587
Two Worker	-1.033	0.325	-3.174
Home Owner	0.754	0.376	2.002
Age4555	-0.685	0.312	-2.193
Pass-Solo			
Constant	-2.638	0.643	-4.101
High Income	0.323	0.303	1.066
Low Income	-1.542	1.070	-1.441
Worker Per Vehicle	1.160	0.451	2.573
Home Owner	0.942	0.478	1.972
Age4555	0.775	0.280	2.769
Perceived –Actual Time Savings	-0.076	0.021	-3.678
Pass-FTP			
Constant	-2.263	0.449	-5.044
High Income	1.045	0.224	4.663
Low Income	-0.871	0.669	-1.301
Home Owner	1.229	0.346	3.557
Age3555	0.571	0.221	2.586
Perceived –Actual Time Savings	0.020	0.009	2.186

Pass-Carpool			
Constant	-3.700	0.496	-7.457
High Income	0.845	0.436	1.938
Low Income	-0.711	1.090	-0.653
Female	1.051	0.425	2.471
College	0.760	0.406	1.871
Generic Variables			
Median Time Savings	0.231	0.037	6.322
Toll	-0.563	0.208	-2.701
Variability	-0.075	0.044	-1.722
Toll*Variability	0.076	0.031	2.470
Actual Toll-Mean Toll	0.827	0.322	2.568
Number of obs	599		
LR chi2(33)	493.48		
Prob > chi2	0		
Pseudo R2	0.2559		
Log likelihood	-717.31185		

TABLE 25: VALUE OF TIME

Percentiles	$/hr
10%	24.62
25%	24.62
50%	35.43
75%	63.87
90%	109.24
Mean	45.13
Std. Dev.	47.96

5.3: A RANDOM EFFECTS DYNAMIC MODEL OF ROUTE CHOICE

This sub-section explores the panel structure of this dataset. Though I have compared similar mode choice models across two time periods and shown the stability of the estimates, the current section is an attempt to exploit the dynamic nature of choice. If past choices have significant impact on current choice then the expectation is to control for it by estimating a random effects logit model. The carpoolers were excluded from the sample and a binary dynamic route choice model is estimated.

The reason for excluding the carpoolers was to simplify the model structure since a random effects multinomial logit model is not that well developed in the literature. The question is whether excluding the carpoolers brings in any systematic bias in the estimates. To check that, a mode choice model, similar to the one in section 3.1.2, is re-estimated but excluding the carpoolers. It is clear from table 26 that the results are unchanged. The VOT estimates are also same as in the model with carpoolers. Thus excluding the carpoolers does not bring in systematic bias into the VOT estimates as is apparent from table 27.

Results from the random effects model is presented in table 28. The results confirms what we have learned from the cross sectional models. High income, middle aged, female commuters are more likely to use the HOT lane. The effect of toll and time savings are as expected. The sign of variability is positive unconditional of the level of toll. This is a different result from earlier models where the effect of variability was dependent on the level of toll and was positive only after the toll crossed a threshold value. But in the dynamic model the effect of variability is positive and the interaction term between toll and variability was not precisely estimated and thus dropped. The value of 'Rho' which measures the proportion of variation caused by the random component in the model is close to 1 suggesting that there is strong correlation across temporal choices for a given individual. Although people respond to current travelling condition, there is strong dependence of current choice on past choice.

The VOT estimate presented in table 28 differentiates between high income group (household income more than 80k) and other income groups. The estimate for the high income group is on the higher side and probably due to the fact I do not control for transponder choice. These estimates are very similar to the ones from the RP route choice model for wave 5 (section 4.1). Though the model cannot reconcile the difference it is interesting in the sense that it shows that even though the surveys were conducted a year apart, commuters are greatly influenced by past choices and thus 'habit' is probably an important determinant of mode choice.

TABLE 26 THE RP MODEL BUT EXCLUDING THE CARPOOLERS

Base: No Pass-Solo	Coef.	Std. Err.	z	P>z
Pass-Solo				
Constant	-3.483	0.622	-5.600	0.000
High Income	0.436	0.296	1.474	0.140
Low Income	-1.593	1.069	-1.490	0.136
Worker Per Vehicle	1.550	0.519	2.986	0.003
Home Owner	0.824	0.447	1.843	0.065
Age4555	0.856	0.276	3.104	0.002
Pass-FTP				
Constant	-1.622	0.491	-3.305	0.001
High Income	1.019	0.228	4.471	0.000
Low Income	-1.037	0.725	-1.430	0.153
Home Owner	1.029	0.357	2.881	0.004
Age3555	0.684	0.240	2.854	0.004
Generic Variables				
Median Time Savings	0.342	0.047	7.292	0.000
Toll	-1.110	0.295	-3.766	0.000
Variability	-0.177	0.089	-1.995	0.046
Toll*Variability	0.153	0.056	2.731	0.006
Actual Toll-Mean Toll	1.149	0.368	3.125	0.002
Number of obs	505			
LR chi2(16)	271.39			
Prob > chi2	0			
Pseudo R2	0.2446			
Log likelihood	-419.10234			

TABLE 27: VALUE OF TIME & VALUE OF VARIABILITY

Percentile	VOT	VOV
25%	18.49	15.77
50%	26.86	31.29
75%	49.63	67.10
95%	88.51	164.59
Mean	36.06	47.51
Std. Dev.	38.05	57.50

TABLE 28: A RANDOM EFFECTS PANEL MODEL OF RP ROUTE CHOICE (1 IF USED FASTRAK, 0 OTHERWISE)

	Coef.	Std. Err.	z
Constant	-8.371	1.767	-4.736
High Income	2.825	0.979	2.885
Low Income	0.843	0.870	0.969
Female	1.310	0.606	2.160
Home Owned	4.518	1.224	3.692
Graduate	0.738	0.592	1.247
Age between 35 & 45	1.099	0.570	1.929
Median Time Savings	0.789	0.168	4.701
Toll	-1.742	0.484	-3.595
Reduction in Variability	0.252	0.112	2.253
Toll*High Income	1.379	0.579	2.382
Time Savings*High Income	-0.562	0.191	-2.943
Actual - Mean Toll	1.325	0.698	1.898
Ln(Sigma-sq-u)	2.895	0.365	7.938
Sigma (u)	4.252	0.775	
Rho	0.948	0.018	

Log likelihood	-341.61911
Number of obs	659
Number of groups	445
Obs per group: min	1
avg	1.5
max	2
Wald chi2(12) (all coeff. 0)	35.35
LR chi2(12) (all except constants)	108.15

TABLE 28A: VALUE OF TIME

Value of Time	$/hr
High Income (above 80k)	37.60
Other Income Groups	27.19

5.4: SP Model with imputed Variability

One of the reasons an SP design may fail to produce close RP result is a failure to replicate the exact conditions under which the RP choices were made. In this case the SP question was relatively simple and omitted a crucial aspect of RP choice – Variability. Since the survey interview was conducted over a telephone it was not possible to conduct a complex SP question that will control for variability.

The respondents could have responded to the omission in either of the following manners: they may either assume variability to be zero or they may respond while considering some positive amount of variability. Since it is not possible to guess which assumption they might have used I will explore all the possibilities. One possibility is that they assumed Variability to be zero, which is the SP model I have already estimated.

The second approach is to estimate the RP mode choice model but imputing zero variability while calculating VOT. I will use 2 RP models to check this hypothesis: from table 4 (morning RP model) the implied VOT is $25.84 and from table 10 (Morning-Afternoon RP model) the implied VOT is 22.36. Both these values are well above the SP values.

The last approach is to form a guess of the variability the respondents assume and estimate the SP model with this imputed variability. Though it is very difficult to guess but one possibility is to assume that the respondent had in mind the variability associated with her general time of travel. The best estimate of this measure would therefore be the engineering estimate of Variability. Result from the estimation of the SP model with imputed variability measure is presented in table 29. This is the same SP route choice model with two additional variables. One is the variability (90^{th} – 50^{th} percentile time savings) and the other is the interaction between hypothetical toll and variability (to give it the same flavor as the RP model). Though the coefficients are barely significant, the model suggests a different behavior than the RP model. The effect of variability is positive (as opposed to negative) and it becomes negative (as opposed to positive) only after toll crosses a certain threshold level. In this case it is approximately $ 4.60, which in this context is a very unlikely event (the maximum charged is generally $4.50). Though this is an interesting result, however in terms of VOT this model is not that different, As shown in table 30 the values are slightly higher but still way below the RP estimates.

TABLE 29: SP MODEL WITH IMPUTED VARIABILITY

	Coefficient	Std. Err.	z
Constant	-1.648	0.656	-2.512
High Income	0.815	0.209	3.907
Low Income	-0.234	0.428	-0.547
Vehicle Per Driver	0.057	0.199	0.284
Gender Dummy	0.210	0.207	1.014
Single Worker Household	0.300	0.211	1.425
Time Savings	0.276	0.053	5.203
Time Savings Squared	-0.004	0.001	-3.245
Toll for FasTrak Users	-1.009	0.264	-3.823
Toll Squared	0.071	0.035	2.045
Variability	0.140	0.089	1.579
Toll*Variability	-0.030	0.022	-1.351
Number of obs	625		
LR chi2(12)	250.47		
Prob > chi2	0		
Pseudo R2	0.2912		
Log likelihood =	-304.8		

TABLE 30: VALUE OF TIME

Percentiles	$/hr
10%	3.10
25%	6.90
50%	13.63
75%	23.09
90%	39.63
Mean	17.82
Std. Dev.	16.14

5.5 RP Model with 'changed departure time'

Why do commuters systematically overestimate time savings? Is it an error in judgement in their part or are the engineering estimates suffering from measurement error? The answer to this question is outside the scope of this thesis but let us consider the consequence of these two possibilities. If they are making a mistake in judgement and basing their mode choice decision on an inflated time savings then that would be consistent with the high RP estimate. It would also be consistent with the low SP estimates because the hypothetical amounts were based on the engineering estimates.

On the other hand consider the other possibility. It is possible that the respondents are considering some extra time savings which the engineering estimates fail to capture. In response to a question to FasTrak users as to 'when would they have left home had there been no FasTrak' almost 90% said that they have leave earlier than they do now. On asked further by how much earlier most of them answered 10 to 15 minutes. Coincidentally that is roughly the same margin by which the respondents overestimate time savings. So it is possible that their perceived time savings include this delayed departure time. But I lack the data required to estimate a simultaneous departure time and mode choice model.

Alternatively I will try to control for it by using the time they said they would have to pre-pone their departure time if there were no FasTrak. Table 30 presents an RP route choice model with this variable which has a positive and significant impact on FasTrak use. A higher value of this variable means that they have to leave home earlier in the absence of FasTrak. It is positive which is what we would expect a priori. It is very difficult to say anything further since as was with perceived time savings, the assumption of exogeneity is a suspect. The hope again was to control for this effect and compare the VOT. Table 31 presents the VOT from this model which is again unchanged and is very close to the RP estimates I have derived earlier.

Table 31: A RP Route Choice Model for FasTrak Users only with 'changed departure time due to FasTrak

	Coefficient	Std. Err.	z
Constant	0.961	1.069	0.898
High Income	0.680	0.397	1.711
Female	-0.041	0.393	-0.104
Graduate School	0.324	0.421	0.770
Age between 35 and 45	0.722	0.418	1.727
Distance	0.036	0.019	1.884
One kid Under 16	-0.780	0.453	-1.720
Worker Per Vehicle	-2.560	0.789	-3.245
Two Worker Household	1.079	0.440	2.454
Owns Home	0.016	0.636	0.025
Median Time Savings	0.472	0.095	4.982
Toll	-1.350	0.518	-2.607
Variability	-0.358	0.149	-2.398
Toll*Variability	0.241	0.099	2.437
Actual-Mean Toll	1.376	0.734	1.876
Leave Earleir if no FasTrak (in min.)	0.019	0.009	2.106
Number of obs	243		
LR chi2(15)	66.6		
Prob > chi2	0		
Pseudo R2	0.2472		
Log likelihood	-101.39		

TABLE 32: VALUE OF TIME

Percentiles	VOT ($/hr)
25%	20.98
50%	26.62
75%	63.91
Mean	26.21
Std. Dev.	79.74

CONCLUSIONS

The goal of this thesis was to study commuters' behavior under real-time congestion pricing and to infer about the amount they are willing to pay to reduce congestion (VOT) and uncertainty (VOV). Value of time (VOT) and value of variability (VOV) was measured from their actual choices. These estimates were compared to VOT estimates based on their responses to hypothetical situations. In addition to measuring VOT, this thesis also studied heterogeneity in VOT.

The various disagregate model estimates show that demographic characteristics have an important impact on mode choice decisions. Carpooling is more frequently done by households with more than two workers, higher workers per vehicle, home owners and those between the age of 35 and 45. Females, college educated, and high income commuters are more likely to own a transponder and carpool. This indicates that carpooling is probably done with household members. A medium commute distance is most favorable for carpooling. Households with higher workers per vehicle, home owners and those between the age of 35 and 45 are more likely to own a transponder. However transponder use is positively influenced by high income, the age group 35 to 55, and home ownership.

The generic variables also have significant impact on mode choice. Time savings has a positive impact and Toll has a negative impact on HOT lane use. The new result from this analysis is the interaction between toll and variability, which makes the effect of one conditional on the other. This implies that the effect of toll is negative only if variability is below a certain threshold level. A high variability means that the travel time in regular lanes are very uncertain and commuters are willing to pay very high toll to use the HOT lane. The opposite interpretation is true for variability. It has a positive influence on FasTrak use if the toll rises above a certain level. For any tolls below this level commuters are not encouraged to use the FasTrak lane by increased variability. A priori one would have expected that any level of variability would encourage use of the HOT lane. But this apparently counterintuitive result is explained by the fact that the toll in this case is dynamic and is a reflection of the true congestion level on a particular day. The actual congestion level is not known to a commuter, but they can observe the toll and draw conclusions from it. When the toll is below a threshold level ($1.20 in this case), the probability that it is a 'bad' day is small and thus does not positively influence FasTrak use.

The positive and significant sign of the 'deviation of toll from its mean value' confirms the a priori expectation that commuters are interpreting toll as a signal for congestion. Thus if the toll rises above the level the commuter 'expects' (measured as the mean toll existing at that 'time of day') then it positively influences FasTrak use. Thus it further reinforces the idea that commuters are using toll levels to extract information about current road conditions. Apart from the interesting behavioral implication, it is important to control for this effect to get an unbiased estimate of VOT. If it is not controlled for, then the VOT estimates would be biased upwards.

The afternoon commute is very similar to the morning commute except in terms of variability. The effect of variability is positive if the toll is below a certain level. This is opposite to what we got for the morning commute. This suggests that in the afternoon, though commuters are willing to pay to reduce the uncertainty, they have a ceiling. This is probably because the afternoon travel is not constrained by arrival time and thus makes the commuter more tolerant to variability.

From the different models I find that there is significant difference between revealed preference and stated preference estimates of value of time. The reason for this difference cannot be attributed to sample selection bias, but rather commuters responding differently to controlled experiments and actual choice situations. An attempt to merge the two datasets was statistically rejected which seems to imply that commuters are behaving fundamentally different. Several approaches were taken to reconcile the differences but they were consistent and persistent.

This thesis further shows that there is significant heterogeneity in the valuation of time depending on travel time, gender and other factors. These estimates are stable over different model specifications and time. The results also indicate a very interesting aspect of dynamic pricing. Value of Variability estimates show that people are willing to pay a high price for a reduction in uncertainty and it is dependent on the time of travel.

One of the main drawbacks of this modeling is the assumption that time of travel is exogenous. Small (1992a) points out that such an assumption will bias the estimates. Some of this effect has been controlled by interacting time of day with travel attributes. The problem with making time of travel completely endogenous is the unavailability of data. However, it is encouraging to note that my estimates are very close to those derived by Lam and Small (2000) who explicitly incorporated time-of-day choice.

The other natural extension pursued was to introduce a time dimension in the choice problem. A random effect panel model was estimated to capture the dynamic nature of choice over time. I found that commuters are greatly influenced by their past behavior and thus there is little change over time. The result from this model was not qualitatively different from the cross-sectional model.

A result that emerged from this analysis is that the commuters' perception of time savings is an important determinant of her mode choice and probably the key to explaining the difference between SP and RP. Commuters seem to systematically overestimate the time savings. Thus when we infer VOT from a RP model the engineering estimates, which are lower, pushes up the time savings coefficient. So when the SP questions are asked, which are loosely based on the engineering measures, yields lower VOT estimates. This is just a conjecture and the only way to formally test this hypothesis would be to devise a system of equations where mode choices and perceptions can both be modeled as endogenous processes. Perhaps that would be an interesting endeavor and help us in understanding commuters' travel behavior in a much better manner.

Bibliography

Abu-Eiseh S. and Mannering F. L (1987) "Discrete/Continuous Analysis of Commuters" Route and Departure Time Choices", Transportation Research Record1138, pp. 27-34

Ameniya Takeshi and Quang H. Vuong (1987) "A Comparison of Two Consistent Estimators in the Choice-Based Sampling Qualitative Response Model", Econometrica, Vol. 55 No. 3, 699-702

Arnott, R., A. de Palma, and R Lindsey: (1990) "Economics of a bottleneck" Journal of Urban Economics, 27, 111-130

Arnott, R., A. de Palma, and R Lindsey: (1993) "A Structural model of peak-period congestion: A traffic bottleneck with elastic demand" American Economic Review, 83(l), 161-179

Arnott, R., A. de Palma, and R Lindsey (1994) "The Welfare effects of Congestion Tolls with heterogeneous commuters", Journal of Transport Economics and Policy, 28(2), 139-161

Arnott, R., A. de Palma, and R Lindsey: "Recent Development in the bottleneck model"

Arnott R and K Small (1994): "The Economics of traffic Congestion", American Scientist, 82, Sept-Oct, 446-55

Bates J. J. et al (1990) "A Model of Departure Time Choice in the presence of Road Pricing Surcharges", Presented at PRTC 1990.

Beesley, M.E. (1965), "The Value of Time Spent in Travelling: Some New Evidence", Economica, May, pp. 174-85

Bhat C. R. (1998) "Analysis of Travel Mode and Departure Time Choice for Urban Shopping Trips", Transportation Research B, Vol. 32, No. 6, pp. 361-371

Bhat C. R. (1997) "Work Travel Mode Choice and Number of Non-work Commute Stops", Transportation Research B, Vol. 31, No. 1, pp. 41-54

Bhat, Chandra R. (1995), "A heteroscedastic extreme value model of intercity travel mode choice", Transportation Research, Part B, Vol. 29B, no. 6, pp. 471-83

Bradley Mark (1997), "A Practical Comparison of Modeling Approaches for

Panel Data", in Panels for Transportation Planning: Methods and Applications, edited by Thomas F. Golob, Ryuichi Kitamura and Lyn Long, Kluwer Academic Publisher, pp. 281-304

Bradley, M., Marks, P. and Wardman, M. (1986),"A Summary of Four Studies into the Value of Travel Time Savings", Transportation Planning Methods, Proceedings of Seminar M, PRTC, pp. 271-85

Braid, R. M. (1996) "Peak-Load Pricing of a Transportation Route with and Unpriced Substitute", Journal of Urban Economics, 40,pp. 179-197

Brownstone D., Ghosh A., Golob T., Kazimi C., van Amelsfort D., (2001) "Willingness-to-Pay to Reduce Commute Time and its Variance: Evidence from the San Diego I-15 Congestion Pricing Project, TRANSPORTATION RESEARCH B, forthcoming

Brownstone David, Kazimi C., Golob T. F., (1999) "Modeling Non-ignorable Attrition and Measurement Error in Panel Surveys: An Application to Travel Demand Modeling." Presented at the International Conference on Survey Nonresponse, October 28-31, Portland, OR. To appear in Survey Nonresponse, New York: Wiley.

Brownstone David and Golob T. F. (1992) "The Effectiveness of Ridesaring Incentives: Discrete Choice Models of Commuting in Southern California", Regional Science and Urban Economics 22, pp. 5-24

Calfee, J. and C. Winston (1998). "The value of automobile travel time: implications for congestion policy.", Journal of Public Economics, 69: 83-102.

Cambridge Systematics, Inc. (1977). "The Development of a Desegregate Behavioral Work Mode Choice Model", prepared for California Department of Transportation and Southern California Association of Governments. Cambridge, MA.

Caplice C. and Mahmassani H.S (1992) "Aspects of Commuting Behavior: Preferred Arrival Time: Use of Information and Switching Propensity", Transportation Research A, Vol. 26A, No 5, pp 409-418

Cascetta E. et al (1992), "Analysis and Modeling of Commuters" Departure Time and Route Choice in Urban Economics", Presented at 2nd International CAPRI Seminar on: Urban Traffic Network

Cosslett Stephen R. (1981), "Maximum Likelihood Estimator for Choice-Based Samples", Econometrica, Vol. 49 No. 5, 1289-1316

Cohen Y (1987): “Commuter Welfare Under Peak Period Congestion Tolls: Who gains and who looses?”, The International Journal of Transport Economics, 14, 239-66

De Palma A., Khattak A. J., Gupta D. (1997) "Commuters” Departure Time Decisions in Brussels", TRANSPORTATION RESEARCH BOARD Annual Meeting 1997

Emmerink R. H. M. and van Beek P. (1997) "Empirical Analysis of Work Flexibility: Implications for Road Pricing and Driver Information Systems, Urban Studies, Vol. 34, No. 2, pp. 217-234

Evans, Andrew W. (1992): “Road Congestion Pricing: When is it a good policy?“Journal of Transport Economics and Policy, September, pp. 213-44

Glazer, A and Niskanen, E. (2000), "Which Consumers Benefit from Congestion Tolls ?", Journal of Transport Economics and Policy, Volume 34, Part I, pp. 43-54

Glazer, A (1981): “Congestion Tolls and Consumer Welfare”, Public Finance, 36(1), pp. 77-83

Golob, J.M. and Golob, T.F. (2000). “Studying road pricing policy with panel data analysis: The San Diego I-15 HOT lanes”, paper presented at the 9th Conference of the International Association for Travel behavior Research, Gold Coast, Australia, July 2-7.

Golob, J. M., Supernak J., Golob, T.F. and Kawada, K. (1998). "An evaluation of a high occupancy toll (HOT) lane demonstration project in San Diego", paper presented at European Transport Conference, 14-18 September, Loughborough University, England.

Golob, T.F. (1999). "Joint models of attitudes and behavior in the evaluation of the San Diego I-15 Congestion Pricing Project". Report UCI-ITS-WP-99-07, Institute of Transportation Studies, University of California, Irvine.

Golob, T.F., Kitamura, R. and Supernak, J. (1997). "A panel-based evaluation of the San Diego I-15 Carpool Lanes Project", in Golob, T.F., Kitamura, R. and Long, L., eds. Panels for Transportation Planning: Methods and Applications, 97-128, Kluwer Academic Publishers, Boston.

Gonzolez, R. M. (1997). The value of time: a theoretical review. Transport Reviews, 17(3): 245-266.

Heckman James J. ((1981), “Statistical Models for Discrete Panel Data”, in Structural Analysis of Discrete Data with Econometric Application ed. By Charles F. Manski and Daniel McFadden, MIT Press, pp. 114-178

Heggie, I. G. (1976), "A Diagnostic Survey of Urban Journey to Work Behavior', in Heggie (ed) 'Modal Choice and Value of Travel Time', Oxford, Oxford University Press.

Henderson J. V (1974): “Road Congestion: A reconsideration of pricing theory” Journal of Urban Economics, 9, 349-64

Hendrickson C. and Plank E. (1984) “The Flexibility of Departure Times for Work Trips”, Transportation Research A, Vol. 18A, No. 1, pp. 25-36

Hensher, David A. (1974), "A Probabilistic Desegregate Model of Binary Mode Choice", in Hensher (ed): 'Urban Travel Choice and Demand Modelling', Special Report no. 12, Australian Road Research Board, Melbourne.

Hensher, David A. (1978), "Valuation in Journey Attributes: Some Existing Empirical Evidence", in Hensher and Dalvi (ed), 'Determinants of Travel Choice', New York:Praeger, pp. 203-65

Hensher, David A. Milthorpe, F. W., Smith, N. C. and Bernard, P. O. (1989), "Urban Tolled Roads and the Value of Time Savings", The Economic Record, Vol. 66, No. 193, pp. 146-56.

Hensher, David A. (2000) "Measurement of Valuation of Travel Time Savings", Working Paper (ITS-WP-00-13), Institute of Transportation Studies, The Australian Key Centre in Transport Management, The University of Sydney and Monash University

Imbens Guido W. (1992) “An Efficent Method of Moment Estimator for Discrete Choice Models with Choice-Based Sampling”, Econometrica, Vol. 60 No. 5, 1187-1214

Jou Rong-Chang and Mahmassani H. S. (1994) "Comparability and Transferability of Commuter Behavior Characteristics Between Cities: Departure Time and Route Switching Decisions, TRANSPORTATION RESEARCH BOARD Annual Meeting 1996

Koskenoja P. and Khattak A. J. (1997) "Analyzing Work Departure Time Variability in Brussels", TRANSPORTATION RESEARCH BOARD Annual Meeting, 1997

Kraus Marvin, Herbert Mohring and Thomas Pinfold (1976), “The Welfare

Costs of nonoptimum pricing and investment policies and investment policies for Freeway Transportation", The American Economic Review, 66, 532-547

Laih, C-H, (1994), "Queuing at the bottleneck with single and multi-step tolls", Transportation Research A, 28A (3) 197-208

Lam, C. Terrence and Small, K. A (2001), "The Value of Time and Reliability: Measurement from a Value Pricing Experiment', Transportation Research E. Vol 37.

Lave, C. A. (1969), "A Behavioral Approach to Modal Split Forecasting", Transportation Research, vol. 3.

Lee, N. and Dalvi M. Q. (1969), "Variation in the Value of Travel Time", The Manchester School, Vol. 37, No. 3.

Liu, L.N and J.F.McDonald (1999),"Economic Efficiency of second-best congestion pricing scheme in urban highway system", Transportation Research Part B pp. 157-188

Lisco, T.E. (1967),"The Value of Commuters Travel Time: a Study in Urban Transportation" Ph.D. these, University of Chicago.

Mackie P. J., Jara-Diaz S., Fowkes A. S. (2001) "The Value of Travel Time Savings in Evaluation", Transportation Research E Vol 37 pp. 91-106

Madan D. B. and Groenhout R. (1987) "Modeling Travel Mode Choices For Sydney Work Trip", Journal of Transport Economics and Policy, pp. 135-149

Manski, C.F. and Lerman. S. (1977) "The estimation of choice probabilities from choice-based samples". Econometrica, 45, 1977-1988.

Marchand Maurice (1968) "A Note on Optimal Tolls in an Imperfect Environment", Econometrica, 36, pp. 575-581

Mannering F. L and Hamed H. M (1990) "Occurrence, Frequency, and Duration of Commuters" Work to Home Departure Delay", Transportation Research B, Vol. 24B, No. 2, pp. 99-109

McCafferty D. and Hall F. "The Use of Multinomial Logit Analysis To Model The Choice of Time to Travel", Economic Geography, vol. 58, no. 3

Mohring, H. (1999) "Congestion" in Essays in Transportation Economic and Policy: A Handbook in honor of John R. Meyer: Brookings Institute Press, pp181-222

Mohring, H., Schroeter, J. and Wiboonchuitkula, P (1987), "The Values of Travel Time, Travel Time, and a Seat on a Bus", Rand Journal of Economics, Vol. 18, No. 1, pp. 40-56

Pells S (1987) "The Evaluation of Reduction in the Variability of Travel Times on the Journey to Work", PRTC Seminar C 7-11 Sept. 1987

Polak J., Jones P., Vythoulkas P.(1993), "Travelers" Choice of Time of Travel Under Road Pricing", 21st PRTC Summer Annual Meeting.

Small, K. A, Noland R, Chu Xuehao and Lewis David, (1999), "Valuation of Travel-Time Savings and Predictability in Congested Conditions for Highway User-Cost Estimation", Report 431, National Cooperative Highway Research Program, Transportation Research Board, National Academy Press, Washington, D. C.

Small, K.A and Anderson, David R. (1998): Revenue from Fine and CoarseToll Schedules, Working Paper, University of California, Irvine

Small, K. A. (1992A): Trip Scheduling in Urban Transportation Analysis, The American Economic Review, Vol. 82, No. 2, Papers and Proceedings of the Hundred and Fourth Annual Meeting of the American Economic Association, pp. 482-486.

Small, K.A. (1992B): "Urban Transportation Economics", Harwood Academic Publishers.

Small, K. A. and Clifford Winston (1986): "Efficient Pricing and Investment Solutions to Highway Infrastructure Needs", The American Economic Review Vol. 76 Issue 2, Papers and Proceedings of the Ninety-Eighth Annual Meeting of The American Economic Association, Pg. 165-169

Southworth F. (1981) " Calibration of Multinomial Logit Models of Mode and Destination Choice", Transportation Research B, Vol. 15A, pp. 315-325

Stopher, Peter, R. (1976), "Derivation of Values of Time for Travel Demand Models", Transportation Research Board, National Research Council, Washington, D. C., pp. 12-18.

Supernak, J. (1991). "Assessment of the Effectiveness of the Reversible Roadway for High Occupancy vehicles on Interstate Route 15", Final Report Prepared for the California Department of Transportation, Department of Civil Engineering, San Diego State University.

Supernak, J., J. M. Golob, K. Kawada and T.F. Golob (1999). "San

Diego"s I-15 Congestion Pricing Project: Preliminary findings". Presented at annual meeting of Transportation research Board, January, Washington, DC.

Swait J. and Ben-Akiva (1987) "Empirical Test of a Constrained Choice Discrete Choice Model: Mode Choice in Sao Paulo, Brazil", Transportation Research B, Vol. 21 B, No. 2. pp. 103-115

Train, K. (1976), "Work Trip Mode Split Models: An Empirical Exploration of Estimate Sensitivity to Model and Data Specification", Working Paper No. 7602, Urban Travel Demand Forecasting Project, Institute of Transportation Studies, University of California, Berkeley.

Train, K. (1998),"Recreational Demand Models with Taste Difference over People", Land Economics, 74(2), pp. 230-239

Verhoef, Erik T. and Small, K. A (1999), "Product Differentiation On Roads Second -Best Congestion Pricing with Heterogeneity under Public and Private Ownership" Working Paper (99-00-01), Department of Economics, University of California, Irvine

Verhoef, E.T, P. Nijkampand P. Rietveld (1996): "Second-Best Congestion Pricing: The case of an untolled alternative", Journal of Urban Economics, 40(3), 279-302

Vickery, W.S. (1969)"Congestion Theory and transport Investment", The American Economic Review (papers and proceedings), 59, pp. 251-60

Wardman, M. (1998). "The value of travel time: A review of British evidence". Journal of Transportation Economics and Policy, 32: 285-316.

Walters, II, W. G. (1992) "Values of Travel Time Savings and the Links with Income", Paper Presented at the Annual Meeting of the Canadian Transportation Research Forum, Banff, Alberta.

Weitzman, Martin (1974), "Free Access vs. Private Ownership as Alternative Systems for Managing Common Property", Journal of Economic Theory, 8, pp.

Yang H and H J Huang (1997):"Analysis of the time varying pricing of a bottleneck model with elastic demand using optimal control theory", Transportation Research B, 31(B)(6), 425-40

Yang Hai, and William H.K. Lam (1996): "Optimal Road Tolls Under conditions of Queuing and Congestion," Transportation Research, 30A, PP 319-332