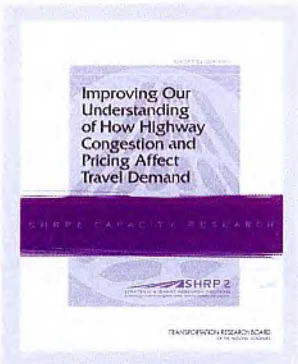

Improving Our Understanding of How Highway Congestion and Pricing Affect Travel Demand

DETAILS

0 pages | 8.5 x 11 | PAPERBACK
ISBN 978-0-309-43480-5 | DOI 10.17226/22689

CONTRIBUTORS

Parsons Brinckerhoff; Strategic Highway Research Program Capacity Focus Area; Transportation Research Board; National Academies of Sciences, Engineering, and Medicine

The Second
STRATEGIC HIGHWAY RESEARCH PROGRAM

SHRP 2 REPORT S2-C04-RW-1

Improving Our Understanding of How Highway Congestion and Pricing Affect Travel Demand

PARSONS BRINCKERHOFF
NORTHWESTERN UNIVERSITY
MARK BRADLEY RESEARCH & CONSULTING
UNIVERSITY OF CALIFORNIA AT IRVINE
RESOURCE SYSTEM GROUP
UNIVERSITY OF TEXAS AT AUSTIN
FRANK KOPPELMAN, AND
GEOSTATS

TRANSPORTATION RESEARCH BOARD
WASHINGTON, D.C.
2013
www.TRB.org

Subscriber Categories

Highways
Planning and Forecasting

The Second Strategic Highway Research Program

America's highway system is critical to meeting the mobility and economic needs of local communities, regions, and the nation. Developments in research and technology—such as advanced materials, communications technology, new data collection technologies, and human factors science—offer a new opportunity to improve the safety and reliability of this important national resource. Breakthrough resolution of significant transportation problems, however, requires concentrated resources over a short time frame. Reflecting this need, the second Strategic Highway Research Program (SHRP 2) has an intense, large-scale focus, integrates multiple fields of research and technology, and is fundamentally different from the broad, mission-oriented, discipline-based research programs that have been the mainstay of the highway research industry for half a century.

The need for SHRP 2 was identified in *TRB Special Report 260: Strategic Highway Research: Saving Lives, Reducing Congestion, Improving Quality of Life,* published in 2001 and based on a study sponsored by Congress through the Transportation Equity Act for the 21st Century (TEA-21). SHRP 2, modeled after the first Strategic Highway Research Program, is a focused, time-constrained, management-driven program designed to complement existing highway research programs. SHRP 2 focuses on applied research in four areas: Safety, to prevent or reduce the severity of highway crashes by understanding driver behavior; Renewal, to address the aging infrastructure through rapid design and construction methods that cause minimal disruptions and produce lasting facilities; Reliability, to reduce congestion through incident reduction, management, response, and mitigation; and Capacity, to integrate mobility, economic, environmental, and community needs in the planning and designing of new transportation capacity.

SHRP 2 was authorized in August 2005 as part of the Safe, Accountable, Flexible, Efficient Transportation Equity Act: A Legacy for Users (SAFETEA-LU). The program is managed by the Transportation Research Board (TRB) on behalf of the National Research Council (NRC). SHRP 2 is conducted under a memorandum of understanding among the American Association of State Highway and Transportation Officials (AASHTO), the Federal Highway Administration (FHWA), and the National Academy of Sciences, parent organization of TRB and NRC. The program provides for competitive, merit-based selection of research contractors; independent research project oversight; and dissemination of research results.

SHRP 2 Report S2-C04-RW-1

ISBN: 978-0-309-12969-5

Copyright Information

Note: SHRP 2 report numbers convey the program, focus area, project number, and publication format. Report numbers ending in "w" are published as web documents only.

Notice

SHRP 2 Reports

Available by subscription and through the TRB online bookstore:
www.TRB.org/bookstore

Contact the TRB Business Office:
202-334-3213

More information about SHRP 2:
www.TRB.org/SHRP2

THE NATIONAL ACADEMIES

Advisers to the Nation on Science, Engineering, and Medicine

The **National Academy of Sciences** is a private, nonprofit, self-perpetuating society of distinguished scholars engaged in scientific and engineering research, dedicated to the furtherance of science and technology and to their use for the general welfare. On the authority of the charter granted to it by Congress in 1863, the Academy has a mandate that requires it to advise the federal government on scientific and technical matters. Dr. Ralph J. Cicerone is president of the National Academy of Sciences.

The **National Academy of Engineering** was established in 1964, under the charter of the National Academy of Sciences, as a parallel organization of outstanding engineers. It is autonomous in its administration and in the selection of its members, sharing with the National Academy of Sciences the responsibility for advising the federal government. The National Academy of Engineering also sponsors engineering programs aimed at meeting national needs, encourages education and research, and recognizes the superior achievements of engineers. Dr. C. D. (Dan) Mote, Jr., is president of the National Academy of Engineering.

The **Institute of Medicine** was established in 1970 by the National Academy of Sciences to secure the services of eminent members of appropriate professions in the examination of policy matters pertaining to the health of the public. The Institute acts under the responsibility given to the National Academy of Sciences by its congressional charter to be an adviser to the federal government and, on its own initiative, to identify issues of medical care, research, and education. Dr. Harvey V. Fineberg is president of the Institute of Medicine.

The **National Research Council** was organized by the National Academy of Sciences in 1916 to associate the broad community of science and technology with the Academy's purposes of furthering knowledge and advising the federal government. Functioning in accordance with general policies determined by the Academy, the Council has become the principal operating agency of both the National Academy of Sciences and the National Academy of Engineering in providing services to the government, the public, and the scientific and engineering communities. The Council is administered jointly by both Academies and the Institute of Medicine. Dr. Ralph J. Cicerone and Dr. C. D. (Dan) Mote, Jr., are chair and vice chair, respectively, of the National Research Council.

The **Transportation Research Board** is one of six major divisions of the National Research Council. The mission of the Transportation Research Board is to provide leadership in transportation innovation and progress through research and information exchange, conducted within a setting that is objective, interdisciplinary, and multimodal. The Board's varied activities annually engage about 7,000 engineers, scientists, and other transportation researchers and practitioners from the public and private sectors and academia, all of whom contribute their expertise in the public interest. The program is supported by state transportation departments, federal agencies including the component administrations of the U.S. Department of Transportation, and other organizations and individuals interested in the development of transportation. **www.TRB.org**

www.national-academies.org

ACKNOWLEDGMENTS

This work was sponsored by the Federal Highway Administration in cooperation with the American Association of State Highway and Transportation Officials. It was conducted in the second Strategic Highway Research Program, which is administered by the Transportation Research Board of the National Academies. The project was managed by Steve Andrle, Deputy Director, SHRP 2.

The Parsons Brinckerhoff research team acknowledges Peter Vovsha (Principal Investigator 1) and Bob Donnelly (Project Manager) of Parsons Brinckerhoff; Mark Bradley (Principal Investigator 2) and John Bowman of Mark Bradley Research & Consulting; Hani Mahmassani (Principal Investigator 3) of Northwestern University; Tom Adler of Resource System Group; Kenneth Small and David Brownstone of the University of California at Irvine; Kara Kockelman of the University of Texas at Austin; Jean Wolf of GeoStats; and Frank Koppelman.

FOREWORD

Stephen J. Andrle, *SHRP 2 Deputy Director*

Driver response to congestion and road pricing is an essential element to forecasting the future use of roadway systems and estimating the effect that pricing has on demand and route choice. Though many studies have been conducted in the past and revenue studies are routinely done for proposed toll roads, there is still a need for improving the behavioral basis for forecast. The objective of this project was to develop mathematical descriptions of the full range of highway user behavioral responses to congestion, travel time reliability, and pricing. These descriptions were achieved by mining existing data sets. The report estimates a series of nine utility equations, progressively adding variables of interest. This research explores the effect on demand and route choice of demographic characteristics, car occupancy, value of travel time, value of travel time reliability, situational variability, and an observed toll aversion bias. The primary audience for this research is professionals who develop travel demand and traffic forecasts. Policy makers may also have an interest in the behavioral findings that could have policy implications. Equations for commercial drivers were not developed since their routes are normally determined, in part, by contracts and company policies.

The researchers for this study identified both revealed and stated preference data sets that could be mined to estimate equations on driver responses to congestion and tolls. The primary data sets were from Seattle and New York. Supporting data sets, used for testing transferability of the equations, included San Francisco, Minneapolis, Chicago, San Diego, Orange County (CA), and Baltimore. A hierarchical choice framework was used. The choice framework considers first residential location and activities, followed by primary destination and intermediate stops, mode of travel, occupancy (when applicable), time of day, departure window, and finally route choice.

The basic utility equation features travel time and cost with coefficients estimated from the data sets. Additional levels of disaggregation may be used depending on the availability of data. In the next level, the equation specifies time to mean "free flow" and "congested" time. The data analysis indicates that drivers perceive every minute driving in congested conditions at 1.5 to 2.0 times longer than free flow travel time. In the next level, which adjusts the cost term for income, research shows that the value of travel time increases with income, but not linearly. The cost term is subsequently disaggregated by auto occupancy, followed by personal characteristics such as trip purpose, age, and gender. Sensitivity testing shows that segmentation by trip purpose is significant, but other personal characteristics are not extremely significant. Travel time reliability, considered in the next level, is the standard deviation of travel time adjusted for distance. This equation recognizes that the value of travel time reliability for short trips (e.g., 5 miles), especially trips to and from work, is greater. The next variable revealed from the data is a toll aversion bias, representing a psychological perception over and above time-cost trade-offs. The toll aversion bias is equivalent to 15–20 minutes of travel time even in areas with a long history of toll roads. The final term in the complete equation represents unobserved heterogeneity. This variable is significant because it represents what may be called "trip pressure" or other situational factors in which there is a penalty for lateness (e.g., trips to the airport or to pick up children). People making such trips are often willing to pay a toll rate higher than demographic or trip purpose characteristics would indicate.

This research reveals a number of policy implications. Drivers place a value on travel time across a wide range from $5 to over $50 per hour and approaching $100 per hour when trip pressure is high. Therefore, toll levels have to be significant to influence congestion. Travelers' responses to congestion and pricing are also dependent on the options available. Driver response to congestion and pricing usually escalates from changing a route or departure time, to switching to transit if available, to rescheduling trips, and finally moving or changing jobs. Providing travel options is an important complement to a road pricing strategy that is aimed at reducing congestion. Finally, improvements to travel time reliability are as important as improvements to average travel time. This implies that operational improvements and information provided to travelers may be as valuable as increases in speed.

The report contains extensive documentation on the estimation of these models and the policy implications. It also contains insights on the value of travel time reliability and the use of reliability in travel demand and simulation models.

CONTENTS

Executive Summary

Organization

Project SHRP 2 C04, Improving Our Understanding of How Highway Congestion and Pricing Affect Travel Demand, reviewed and advanced the state of the practice in modeling the effects of highway congestion and highway pricing on travelers' decisions, including choices of facility, route, mode, and time of day (TOD). This Executive Summary is intended for those who do not have extensive experience in travel demand modeling, but do wish to learn about and apply the results of the modeling research. The Executive Summary summarizes the objectives, data, methods, and key findings of the research. Each finding is accompanied by a discussion of the behavioral modeling issues, analysis results, and implications for transportation policy.

Reader Navigation

The C04 research aimed at breakthrough advances in travel demand modeling, and simulation was necessarily conducted within a highly specialized conceptual, mathematical, and technical context. Many important aspects of the work can only fully and meaningfully be described with theoretical constructs and important technical details that are difficult to comprehend by most general readers and even many model practitioners. The project team addressed this dilemma by maintaining continuity in the discussion, which can include both the more accessible content and the highly technical content, and offering navigation guides to the reader. Each section of the report typically begins with a discussion aimed at all readers; highly technical details are placed in Appendix A or are flagged in Chapter 6 with the header Related Technical Detail.

Technical Report

The Technical Report is intended for practitioners who perform or direct modeling of highway congestion and pricing policies and for planners and decision makers who wish to gain a deeper technical appreciation and understanding of particular issues. The discussion is organized around specific modeling questions and hypotheses that were tested, rather than around specific data sets and choice models.

Chapter 1 describes the research objectives and methodology. Chapter 2 describes the review and selection of data sources for tests of the travel demand models. The main research results described in Chapter 3 are organized by model types and features, and the model components are compared back-to-back. Chapter 4 covers the network simulation procedures for congestion and pricing studies, and Chapter 5 describes incorporation of the results into practice. Conclusions, main findings, and recommendations for future research are summarized in Chapter 6.

Technical Appendix and Supplemental Material

Appendix A comprises all technical details related to the specification and estimation of C04 models. An unabridged, unedited version of Chapter 3 is available online at: www.trb.org/Main/Blurbs/168141.aspx. Appendix A provides the full statistical results for the key models estimated and discussed in Chapter 3. It is organized according to the types of models estimated and the data sets used. This appendix is intended for those with experience in estimating discrete choice models and will be useful to those who wish to estimate choice models on their own data sets. Full model specifications are included in tabular form, and models are numbered in a logical way corresponding to the discussion in Chapter 3.

Appendix A also consists of three short technical memoranda that document and detail the processing and development of data sets and network simulation formulations, algorithms, and calibration.

Summary of Objectives and Methodological Principles

Primary Objectives and Focus of Research

The C04 project was designed to (1) synthesize research findings from the past 30 years on travelers' responses to changes in both traffic congestion and the price of travel, (2) select the most important and well-founded behavioral hypotheses, (3) test those hypotheses statistically on the most suitable data sets available in the United States, and (4) identify ways in which the developed functions could be incorporated in operational models of travel demand and network simulations. The scope for the project and the range and quantity of statistical analysis performed were extensive. This summary report is a distillation of the issues studied and the most important findings.

The focus of the research was to identify the most important contextual influences on individual behavioral sensitivity to highway congestion and pricing and to provide guidance on the relative magnitudes of those influences. In practice, behavioral sensitivities are often expressed as elasticity—the percentage change in a behavioral outcome divided by the percentage change in travel time or cost. For example, the elasticity of vehicle miles traveled (VMT) in response to gasoline price has sometimes been measured at –0.2, meaning that a 10% increase in fuel price will lead to a 2% decrease in VMT. Although it is tempting to think that behavioral responses can be predicted by using such a simple measure, the reality is much more complex. For example, the elasticity of VMT with respect to price is also influenced by the income level of the traveler, the presence or absence of good transit alternatives, and opportunities such as buying a more efficient vehicle or finding a job closer to home.

The following sections summarize the methodological principles of the research.

Highway Utility (Generalized Cost) Formulation

The most common approach to dealing with context specificity is to use disaggregate discrete choice models to predict the choices that a given type of traveler will make for a given type of trip under specific circumstances. A discrete choice model assumes that when all important aspects of a choice are weighed, each choice alternative will have a resulting utility, or attractiveness, for each specific traveler. The probability that the traveler will choose a specific alternative is a function of that alternative's utility relative to the utility of all other available alternatives.

A typical formulation of the utility (U) of a highway alternative can be written as a linear function of trip time and cost (including all forms of pricing) in the following way:

$$U = \Delta + a \times Time + b \times Cost + \varepsilon \tag{ES.1}$$

where
Δ = alternative-specific constant;
a = auto time coefficient;
b = auto cost coefficient; and
ε = term that captures the residual effects of all variables that are not explicitly represented.

Note that Δ, a, b, and ε may be either single terms or functions of other variables (e.g., income or trip purpose). In certain simple choice contexts like route choice, in which all alternatives are qualitatively similar, the alternative-specific constants can all be assumed to be equal to zero and dropped from the equation. However, in more general choice contexts that involve different modes of travel or different TOD periods, it is essential to account for qualitative differences between alternatives. In these choice contexts, Δ may include additional explanatory variables beyond travel time and cost.

Because utility has no physical dimensions, the coefficients are generally interpreted as relative to each other, as ratios. For example, if the estimated coefficient for auto travel time is –0.02/minute, and the estimated coefficient for auto cost is –0.10/$1, then the ratio of the two coefficients (a/b) is equal to $0.20/minute, or $12/hour. This particular ratio (here designated as value of time [VOT]) can be interpreted as the additional price that a traveler would be willing to pay for a marginal decrease in travel time (or to avoid a marginal increase in travel time).

The willingness to pay for reduced travel time varies substantially depending on the characteristics of the traveler and the context of his or her particular trip. Typical U.S. modeling analyses have only accounted for a limited amount of this variation, usually segmenting by trip purpose and sometimes by income segments.

Key Behavioral Hypotheses

Most of the tested hypotheses were related to how travel time and cost enter the highway utility formulation. The simplistic linear form of Equation ES.1 was questioned in many respects. Its primary drawbacks relate to the unrealistic assumption of a constant VOT.

Previous research identifies three major aspects of highway driving time that influence behavior and are perceived as important components of highway level of service (LOS): quantity (duration of time in the vehicle); quality (amount of stress or pleasure caused by the particular driving conditions); and reliability (level of uncertainty with respect to travel time and congestion levels). Each of these three travel time aspects and travel cost are likely to influence travel choices differently, although the corresponding effects are often intertwined.

Travel Time Quantity

The duration of travel time influences travelers' schedules and the alternative uses of the time they spend traveling. People with busy schedules are usually more likely to seek an alternative that will gain them a shorter-duration trip. For frequent trips (e.g., commuting), people may be more aware of the time duration difference between different travel alternatives and also more able to change their activity schedules to make optimal use of the time saved. Also, the longer is the block of time saved by choosing a specific travel option, the more likely it is that a traveler will perceive the time saved. The project team hypothesized that the utility of travel time duration is likely to vary according to the time constraints and activity schedule of the traveler and the frequency, TOD, and distance of a specific trip.

In practice, engineers often do not have accurate data on travelers' time constraints or trip frequency, particularly for forecasts. Trip purpose serves as a proxy, as regular trips to work or school tend to be the most frequent trips, and are often made by those with the busiest schedules.

Travel Time Quality

Many drivers find driving in stop-and-go traffic more stressful than driving in free-flow conditions and would be willing to pay more to avoid time spent driving in heavy traffic. For example, many drivers will drive a longer distance to avoid congestion bottlenecks; they may spend the same amount of time or more in traveling, but in less stressful conditions. These differences in congestion levels tend to be related to differences in travel time reliability (discussed below), but those reliability effects are greater than the effects referred to here, which relate only to the physical and mental stress of time spent in the vehicle, and which would exist even if the duration of travel time was completely predictable.

Although the stress caused by driving in different conditions may vary from driver to driver, the analyst can expect to find some systematic effects. In general, people place a higher value on time savings that arise from reductions in congestion levels than on time savings from other types of system changes, such as the introduction of shorter-distance routes or closer destinations. This further implies that time savings will be valued more for specific facilities and times of day when congestion levels are the highest. People often report that driving becomes more stressful as the duration of the commute trip increases beyond 30 or 40 minutes. Thus, one might expect the value of time savings to increase with trip distance or duration, but not in a linear fashion. A possible explanation of the trip-length effects lies in the structure of the entire daily activity pattern rather than in the commute trip itself. When commuting time grows beyond 2.5–3 hours per day and is combined with 8.5–9 hours at work, the total work–commute time makes it difficult to incorporate other activities of a significant duration.

Travel Time Reliability

Even when average travel times for two highway alternatives are the same, drivers will generally prefer the more reliable alternative (least day-to-day variability in travel time) or the lower risk that the travel time will be significantly longer than average. Qualitative and quantitative research has indicated at least three following reasons for this preference:

1. Negative Consequences of Arriving Late at One's Destination. These can include missing an appointment, missing a flight, or losing pay for work time. This consideration gave rise to the *schedule delay* concept in measuring travel time reliability, which is discussed in the Technical Report.
2. Need for Buffer Time to Avoid Arriving Late. Travelers concerned about a late arrival must begin their trip earlier than they would if the travel time were more reliable. This behavior will avoid most instances of arriving late, but at the expense of departing earlier and sometimes arriving too early. This consideration gave rise to a concept in measuring travel time reliability that operates with estimates of buffer time based on the travel time distribution shape.
3. Discomfort Related to the Uncertainty of How Long the Trip Will Take on any Given Day. This approach operates with the simplest quantitative measures (e.g., standard deviation) of travel time variability and probably has the best chance of being incorporated in operational models.

Quantitative research into VOT variability and value of reliability (VOR) has lagged because of the lack of data on the day-to-day travel time variability that drivers face for particular trips. Measuring variability requires an estimate of the travel time distribution, which can be translated into measures such as standard deviation or buffer time represented by the 80th or 90th percentile of travel time versus the median. Estimates of the distribution for an entire trip distance from origin to destination (O-D) are also needed. However, because the correlation between travel times on different links in a network is a complex function of the network configuration and structure of the traffic flows, the distribution for an O-D time is not simply the sum of the distributions across highway links. As a result, O-D–level estimates are quite rare, unless one has extensive global positioning system (GPS)–based trace data over time, or some other means of estimating O-D variability (such as one specifically applied in the current study). Most previous quantitative research into VOR has been based solely on stated preference (SP) data, particularly in Europe. This project makes one of the most thorough efforts to date to overcome these data limitations and obtain estimates using revealed preference (RP) data from different regions of the United States.

Travel Cost

Various auto-related costs can influence the utility of a particular trip alternative (in terms of route, departure time, carpooling, and so forth). Although costs such as vehicle maintenance and insurance can vary with the mileage driven, the travel decisions made for a given trip are more strongly influenced by the direct costs of parking, tolls, and fuel. This research focuses on toll costs because they are most important in terms of future road-pricing policy and because they tend to provide the clearest and most statistically advantageous contexts for measuring the importance of price in choosing between highway travel options.

The sensitivity of travel behavior to a specific type of travel cost depends primarily on (1) how much of the cost a traveler actually has to pay (e.g., cotravelers may share the cost of fuel or tolls) and (2) how affordable that cost is for the particular traveler. The team hypothesized that the most important contextual differences determining the sensitivity of travel behavior to price are related to income, vehicle occupancy, and travel purpose. This behavioral mechanism should be considered in the regional network context. In particular, the presence of a reasonable transit alternative plays a major role in determining the final outcome of congestion and pricing.

Data and Methods to Test Behavioral Hypotheses

Only a few of the more than 100 surveys identified were adopted. These were chosen for the probability that they could support various travel time reliability measures, as well as toll and nontoll routes for the same trip. The team relied on the three types of travel survey data described below.

Revealed Preference Data

RP data are observed data on actual choices made by travelers. Although RP data are always preferable, the existing RP data sets have many limitations. RP data are collected from travel diaries in which survey respondents report all of the travel and activities they undertake in the course of a representative weekday. Because the respondents typically do not report either all of the available (but nonchosen) trip alternatives, or all of the travel times and costs related to the chosen and nonchosen alternatives, those supply-side LOS measures must be inferred from representations of the highway and transit networks. This task can be expensive and time-consuming. Only those RP data sets that can be supported by a well-calibrated regional travel demand model with network simulations implemented for multiple periods of a day are usable. The RP-type Household Surveys in New York City and Seattle, Washington, were adopted for many statistical tests because they could be supported by the necessary LOS variables.

Stated Preference Data

SP data are responses by survey respondents to questions about hypothetical travel situations. These data are collected in SP choice experiments that are customized around the context of an actual reported trip that a respondent has recently made. The SP approach can be used to study the demand for an alternative that does not actually exist, such as a new tolled highway facility. The choice set of available travel alternatives and the related times and costs are explicitly specified for respondents, which avoids the difficulty and expense of inferring those supply variables after the fact. The obvious disadvantage of SP data is that the choices are hypothetical, so there is less confidence that the analysis results will reflect true behavioral relationships. This concern is perhaps strongest in the case of complex variables such as travel time variability, which are difficult to portray clearly to SP respondents. In the current project, SP data were a complementary source that helped explore situations that were not observed in RP surveys.

Experimental Revealed Preference Data

The approach of experimental revealed preference data merges the best characteristics of RP and SP approaches by measuring actual choices from experimental contexts created on actual travel networks. An example is a system of tolls introduced temporarily in a road network and charged via electronic tolling, but only to specific drivers who are participating in the experiment. Experimental RP data from the recent Traffic Choices Study in Seattle were used in the current research.

Data Analysis Methods and Key Data Sets

After a careful assessment of RP and SP data sets from across the United States, the team selected a handful of data sets that would best support the planned range of analyses (Table ES.1). The New York and Seattle metropolitan regions were selected as the primary regions for RP data analysis because the Household Travel Survey data sets from these areas, together with associated highway and transit network supply data, could support detailed disaggregate model estimation. In addition, the Puget Sound Regional Council had carried out an innovative mileage-based tolling experiment that provided a unique source of RP data that complemented the other Seattle region RP and SP data sets. SP data sets from San Francisco and Los Angeles, California, were also analyzed.

Table ES.1. Data Sets Used for Analysis

Geographic Area	Planning Agency	RP Data	SP Data	Experimental RP Data
New York metropolitan region	New York Metropolitan Transportation Council	1997 Household Travel Survey (1-day diary)		
Seattle metropolitan region	Puget Sound Regional Council	2006 Household Travel Survey (2-day diary)	2006 highway toll SP experiment (follow-up with 2006 survey respondents)	2006 Travel Choices tolling experiment (GPS-based tolling simulation)
San Francisco County	San Francisco County Transportation Authority		2007 downtown area pricing SP experiment	
Los Angeles County	Los Angeles Metropolitan Transit Authority		2008 managed toll-lane SP experiment	

Table ES.2. Types of Choice Models Supported by the Data Sets

Data Set	Toll versus Nontoll Route Choice	Departure Time Choice	Mode and Occupancy Choice	Other Choice Dimensions
New York region 1997 RP household survey	X	X	X	X
Seattle region 2006 RP household survey		X	X	X
Seattle region 2006 Traffic Choices pricing experiment	X	X		
Seattle 2006 SP toll experiment	X	X		
San Francisco 2007 SP area pricing experiment		X	X	
Los Angeles 2008 SP managed-lane experiment	X	X	X	

Table ES.2 summarizes the aspects of choice represented in each of the primary data sets. The New York regional household travel survey RP data set supports the widest range of modeling. Because this region has tolled bridges and tunnels and has collected data on whether those tolls were paid on the respondents' actual trips, the data can be used to model the choice between tolled and free routes. The data also support models of departure time choice (analyzing the times of day that respondents chose to make their trips as a function of congestion levels) and mode choice (the decision to drive alone, carpool, or use bus or rail as a function of the travel times and costs of the different modes). New York, a transit-rich area with an extreme level of congestion and well-established toll facilities, provides a variety of data for exploring trade-offs between travelers' travel time and costs.

Because travel diary data sets provide a complete picture of a representative day's travel, they can be used to model other dimensions of travel choice, such as number of trips made for different purposes, as well as longer-term decisions, such as the number of automobiles to own.

The Seattle region 2006 household RP survey data are comparable to those from the New York region, with the exception that the Seattle region had few tolled facilities at that time. Thus, the Seattle data will not support models of toll versus nontoll route choice, although they can be used to model the influence of other travel costs such as fuel and transit fares on mode choice. In contrast, the Seattle Traffic Choices experimental RP data offered tolls that varied by distance, facility type, and TOD, so these data are suitable for modeling route and departure time choices. The GPS-based data collection method did not provide information on vehicle occupancy or the use of non-auto modes, so its data are not useful for modeling mode or auto occupancy choice or other choice dimensions.

Each of the analyzed SP data sets focused on specific choice dimensions with respect to tolls. All of them offered respondents toll levels that varied by TOD, so the data sets are useful for modeling the effects of pricing (and congestion) on departure time. Some also offered the options of free alternative routes or competing transit alternatives, so those data sets can be used for route choice or mode choice analysis, respectively. There were at least two RP data sets and two SP data sets to model each of the main travel choice dimensions.

In order to focus on the national goals established for SHRP 2 and because of the necessity of managing the data sets and supporting them by the regional travel models and network simulations available to the project team, data from other countries were not considered. The team believes most of the results can be extended (at least qualitatively or conceptually) to areas outside the United States. But because regional conditions play a significant role in shaping travel behavior, a direct transfer of model structures and coefficient values to areas outside the United States is not recommended.

Analysis Approach for Improved Demand Modeling

The team tested the same types of variables and functional specifications of the generalized cost on multiple data sets and choice contexts and looked for consistencies that suggested the most reliable practices and productive paths for modeling the effects of highway congestion and pricing. This systematic approach involved the following steps:

- The first models used the most basic specifications for each data set and each choice dimension (e.g., route or mode choice) separately, using only simple time and cost variables included in Equation ES.1;
- Additional variables (e.g., segmentation variables for income and car occupancy) were then systematically added to the models;
- Model complexity was gradually increased when possible to simultaneously model different travel decisions using nested hierarchical choice structures, such as a joint model of departure time choice and mode choice; and
- The probabilistic distribution (spread) of the key travel time coefficients was estimated using advanced mixed logit specifications.

This stepwise analysis approach across several data sets revealed the stability and generality of the modeling results and allowed a recommendation for a generalized cost specification.

The team recommends using a highway utility function of the general form (Equation ES.2), which is an extension of the simplified form (Equation ES.1):

$$U = \Delta + a_1 \times Time \times (1 + a_2 \times D + a_3 \times D^2) + b \times [Cost/(I^e \times O^f)] + c \times STD/D \qquad \text{(ES.2)}$$

where

Δ = alternative-specific bias constant for tolled facilities;
a_1 = basic travel time coefficient (ideally estimated as a random coefficient to capture unobserved user heterogeneity);
$Time$ = average travel time;
a_2, a_3 = coefficients reflecting the impact of travel distance on the perception of travel time;
D = travel distance;
$Cost$ = monetary cost (e.g., tolls, parking, and fuel);
I = (household) income of the traveler;
O = vehicle occupancy;
e, f = coefficients reflecting effect of income and occupancy on the perception of cost;
STD = day-to-day standard deviation of the travel time; and
c = coefficients reflecting the impact of travel time (un)reliability.

Equation ES.2 includes travel cost explicitly scaled by income and vehicle occupancy; travel time reliability (variability) is included separately from the typical, or median, travel time. The team found that the standard deviation of travel time divided by distance gave the strongest and most consistent results for the reliability effect. Conceptually, this variable represents the day-to-day variability in highway travel speed (time divided by distance is the inverse of speed).

The distance-based term by which travel time is multiplied expresses an important distance effect on travel time perception and VOT. This term represents a polynomial function of distance that scales travel time in the following way:

$$Time \times (1 + a_2 \times D + a_3 \times D^2) \qquad \text{(ES.3)}$$

For travel segments with a short average distance, the distance-related effects are insignificant, and the entire multiplier can be dropped. For longer segments, such as commuting to work in large metropolitan areas, the distance effects are significant. VOT can grow or decline with distance depending on the sign and magnitude of coefficients a_2, a_3.

The suggested form for accounting for travelers' perceptions of travel costs corresponds to the monetary cost scaled by power functions of both income and vehicle occupancy:

$$Cost/(I^e \times O^f) \tag{ES.4}$$

Coefficients e, f lie in the unit interval. If the coefficient is close to zero, the corresponding scaling effect becomes insignificant and can be dropped. If the coefficient is close to 1.0, the scaling effect reaches the maximum. The most statistically significant values were found by exploring the entire range of possible values.

The recommended and most statistically significant main measure of travel time reliability is the day-to-day standard deviation of travel time by auto, divided by distance:

$$STD/D \tag{ES.5}$$

This reliability measure is especially practical for RP-based models because it obviates one of the most problematic features of most RP data sets: correlation between travel time, travel cost, and practically any travel reliability measure, including standard deviation or buffer time. This measure has a plausible behavioral interpretation: Travelers may perceive travel time variability as a relative (qualitative) measure rather than an absolute (quantitative) measure. This behavioral assumption is appealing in the context of the entire highway utility (Equation ES.2), in which travel time and cost are included in an absolute fashion; thus the reliability term plays a complementary role and explains what has not yet been explained by the time and cost terms.

VOT can be calculated as

$$VOT = \frac{a_1}{b} \times (1 + a_2 \times D + a_3 \times D^2) \times (I^e \times O^f) \tag{ES.6}$$

In a general case, VOT is a function of travel distance, income, and car occupancy for each travel segment. If the model is explicitly segmented by these variables, then the formula for VOT can be simplified and made specific to each segment by differentiation of the basic time and cost coefficients.

VOR can be calculated as follows:

$$VOR = \frac{c}{b} \times \frac{(I^e \times O^f)}{D} \tag{ES.7}$$

Like VOT, VOR is a function of travel distance, income, and car occupancy for each travel segment unless a more detailed explicit segmentation is applied. VOR is inversely proportional to distance (i.e., the longer the distance, the greater the magnitude of the reliability measure), although a longer distance tends to dampen travel time variation. The portion of travel time variability that is proportional to the average travel time is accounted for in the loaded travel time coefficient a_1. Thus, only the residual variation of travel time expressed as standard deviation per unit distance is accounted for in the reliability term.

Finally, the reliability ratio (RR) can be calculated as a measure of the relative importance of reduction of (un)reliability versus average travel time savings:

$$RR = \frac{VOR}{VOT} = \frac{c}{a_1} \times \frac{1}{(1 + a_2 \times D + a_3 \times D^2) \times D} \tag{ES.8}$$

Logically, the reliability ratio is a function of travel distance rather than a fixed value. The reliability ratio usually declines with distance, a fact that cannot be taken out of the functional form (Equation ES.2), in which the terms for (loaded) travel time and travel time reliability complement each other.

This form of highway utility (Equation ES.2), with its choices for route type, mode, car occupancy, and TOD, can be incorporated in operational travel demand models in the near future.

Through these primary choice dimensions, the impacts of congestion and pricing can be further propagated through the model system chain to affect destination choice, trip frequency, and other choice dimensions.

Impacts of Congestion and Pricing on Travel Demand: Behavioral Insights and Implications for Policy and Modeling

Key study findings are related to the estimated values of the parameters in Equation ES.2. These findings are presented as a series of 11 key behavioral insights, along with the implications of these insights for road pricing congestion management and improved travel modeling:

- Variation in VOT across highway users;
- Income and willingness to pay;
- Auto occupancy or group travel and willingness to pay;
- Constraints on TOD shifting (carpools and single-occupant vehicles);
- Importance of VOR and its relationship to VOT;
- Effect of travel distance on VOT and VOR;
- Evidence of negative toll bias;
- Hierarchy of likely responses to change in tolls and congestion;
- Summary of user segmentation factors;
- Avoiding simplistic approaches to forecasting; and
- Data limitations and GPS-based data collection methods.

Variation in Value of Time Across Highway Users

Key Finding

VOT varies widely, from $5 to $50/hour across income groups, vehicle occupancies, and travel purposes. There is significant situational variation (unobserved heterogeneity), with some people willing to pay almost nothing to save time, and with others willing to pay more than $100/hour.

Implications for Policy

The wide distribution of willingness to pay confirms that pricing can effectively serve the important function of market discrimination and demand management. Because most travelers have a relatively low willingness to pay, any price that affects all travelers, such as a general toll for all lanes of a highway, may influence demand at fairly modest levels. In contrast, prices for high-occupancy toll (HOT) and express lanes can be set at fairly high levels and adjusted to attract a relatively small percentage of travelers with the highest willingness to pay. Pricing policies should be applied after a careful analysis of possible negative implications for low-income users.

Implications for Modeling

Most models used for travel demand forecasting have assumed a single VOT. Only occasionally have different cost coefficients been used for different income groups and vehicle occupancy levels. Differentiation of VOT is even less typical in network simulation procedures. These practices result in significant aggregation biases that affect the accuracy of traffic and revenue forecasts. Whenever possible, random coefficients should be used to estimate the distribution of VOT across the population. For general use, newer activity-based forecasting models that use a microsimulation approach can simulate a different VOT for each person and trip, providing the most disaggregate treatment of VOT, and thus avoiding one important source of possible errors and biases in the forecasts.

Income and Willingness to Pay

Key Finding

Household and personal income has a strong relationship with VOT and willingness to pay, but the relationship appears to be less then linear. To account for the income effect, cost variables in travel models (including tolls) should be divided by household income, raised to a power in the range 0.6 to 0.8 depending on the trip purpose (e.g., for a power of 0.7, doubling income increases VOT by 62%; halving income decreases VOT by 38%).

Implications for Policy

The income effect is strong, so that many of the benefits of pricing are purchased by those who can most afford them, and equity considerations cannot be discounted. Lower-income travelers also derive benefits in the form of increased options, as well as improvements in traffic conditions if total capacity can be increased through priced facilities. The parallel effect of car occupancy mitigates the income effect. Low-income commuters have more opportunities to carpool and share commuting costs than do high-income commuters. High-occupancy vehicle (HOV) and HOT lanes, as well as transit, represent viable alternatives for low-income travelers.

Implications for Modeling

Forecasting models typically use income either in a simplified linear form to scale travel costs or as a segmentation variable, with different cost coefficients in different income ranges. Neither approach seems entirely appropriate. The assumption of linearity with income seems too strong, particularly in higher income ranges, and the piecewise linear approach often results in strong nonlinearities or discontinuities in the effect of income that do not have a strong statistical or behavioral basis. The recommended approach is empirically justified across a wide body of evidence and provides a smooth response surface for forecasting.

Auto Occupancy or Group Travel and Willingness to Pay

Key Finding

Auto occupancy has a strong estimated relationship with VOT and willingness to pay. The relationship appears to be slightly less than linear. To account for occupancy effects, cost and toll variables in travel models should be divided by occupancy raised to a power in the range 0.7 to 0.8.

Implications for Policy

Because a group of vehicle occupants is generally willing to pay more than a solo driver, a tolled facility is likely to attract a higher percentage of multioccupant vehicles than will a free facility, even if no special carpool discount is offered; in effect, a carpool discount is being offered to a group that tends to value it the least. However, ridesharing increases system capacity, and the conversion of HOV lanes to HOT lanes may potentially discourage carpooling among individuals with higher VOT by offering solo drivers the same travel time advantage without the added inconvenience of ridesharing. Thus, free or discounted toll lanes for carpools will encourage carpoolers, even if it does not attract much additional ridesharing.

In general, low-income commuters have a higher probability of forming a carpool than do high-income commuters. Low-income workers normally have a fixed work schedule, which simplifies carpooling logistics, and they tend to live in dense residential clusters where the process of collecting and distributing passengers requires minimal extra time. In addition, low-income jobs tend to form clusters of multiple jobs. This factor may vary with the structure of the metropolitan area, however, as large clusters of high-income jobs may be present in the central business district (CBD).

The higher opportunity for carpooling among low-income workers mitigates the equity concerns regarding pricing, because costs can be shared within the carpool. Faced with significant tolls, high-income workers can only switch to transit, but low-income workers can use transit or HOV lanes. This consideration is frequently missed in policy analysis of pricing projects, which may result in exaggerated equity concerns.

Implications for Modeling

Dividing travel cost by vehicle occupancy is a fairly standard practice in applied modeling. The team's main recommendation is to divide costs by a function of occupancy that is somewhat less than linear. Income-specific components in the car occupancy choice are needed to reflect differential opportunities to carpool by income. Simplified approaches based on average occupancy coefficients tend to mask these important effects and portray pricing projects in an extreme way with respect to different income groups.

Constraints on Time-of-Day Shifting: Carpools and Single-Occupant Vehicles

Key Finding

Although commute carpools generally have a higher VOT, they also tend to have tighter scheduling constraints and tend to be less flexible in their capacity to shift departure time away from the peak period and hour.

The team consistently found that ridesharing commuters are more likely to travel in the heart of the peak periods than those who drive alone. Carpool commuters must coordinate their commute schedules with cotravelers, so it is less likely that they can adjust their departure times earlier or later to avoid peak congestion or pricing.

Implications for Policy

Because carpooling commuters generally have little opportunity to retime their trips to avoid peak congestion times, TOD pricing and other peak-spreading policies will tend to be less successful in influencing their behavior. To avoid inadvertently discouraging ridesharing, policies should be designed to offer some level of advantage of travel time or price (or both) for HOVs. Congestion pricing policies could be more effective if they were accompanied by policies encouraging employers in the CBD (or other relevant congestion pricing zones) to shift working hours and introduce flexible or compressed work weeks.

Implications for Modeling

Modeling studies intended to predict peak spreading behavior and response to TOD pricing should include different sensitivities for different car-occupancy levels. In general, the propensity to switch from the peak hour to a different hour is inversely proportional to vehicle occupancy.

Importance of Value of Reliability and Relationship to Value of Time

Key Finding

Travelers tend to value variation in travel time reliability (day-to-day variability) at least as highly as they value variations in the usual travel time. Various ways of specifying the variability variable were tested, but the measure that produced the most consistent results was the standard deviation of travel time divided by journey distance.

Evaluating the estimation results to impute the reliability ratio (the value of reducing the standard deviation of travel time by 1 minute divided by the value of reducing the average travel time by 1 minute, or VOR/VOT for an average trip distance) obtained ratios in the range 0.7 to 1.5 for various model specifications. SP studies from Europe give typical values in that same range for auto travel, with higher ranges up to 2.5 for rail and transit travel. The SP results, however, indicate that the estimates may vary depending on how the reliability concept is presented to respondents. Thus, it is crucial to obtain new estimates based on actual choices at the trip level.

Implications for Policy

Highway investments that can improve travel time reliability will tend to be just as beneficial for travelers as investments to reduce typical travel times. This finding underlines the importance of addressing key bottleneck points and using transportation systems management and intelligent transportation systems to monitor and adapt to congestion levels on the network, as well as systems to avoid nonrecurrent congestion and to recover from it as quickly as possible. For managed lanes and other priced facilities, the "guarantee" of a reliable travel time may be of great value. This makes variable pricing, and especially that of dynamically priced lanes, one of the more effective pricing forms that are attractive for the user. These findings also emphasize the importance of effective accident management, as the consequences of traffic accidents constitute a significant share of long delays.

Implications for Modeling

Although models can be estimated using measures of day-to-day travel time variables from real and simulated highway networks, further progress is needed before this method is feasible for most travel demand forecasts, particularly in terms of widespread collection of data for actual levels of travel time variability at the O-D level. Certain technical issues must also be resolved on the network simulation side, specifically the incorporation of travel time reliability in route choice and the generation of O-D travel time distributions instead of average travel times. In the near term, this method may be most applicable to corridor- and facility-level forecasts. Some simplified implicit measures of reliability (such as perceived highway time by congestion levels, as explained below) can be applied with the existing model structures and network simulation procedures.

Effect of Travel Distance on Value of Time and Value of Reliability

Key Finding

Savings on average or typical travel time (VOT) are valued more highly for longer trips than for short trips, except for a special effect on commuting trips over 40 miles. For VOR, there is a relative damping effect for longer trips. These findings suggest the efficacy of using higher-priced managed lanes to address key bottlenecks and lower distance-based tolls on the wider highway network.

Implications for Policy

Traffic bottlenecks tend to increase the variability (unreliability) of all trips that pass through them, regardless of total trip distance, and the results indicate that all travelers will derive considerable benefit from making the system more reliable. In contrast, improvements that increase average speeds or reduce travel distances without substantially improving reliability will not be valued very highly by those who only use the facility for a short distance. Thus, distance-based tolls are appropriate in general, but higher prices that are not based on distance may be more appropriate for addressing key bottlenecks.

Implications for Modeling

Because VOT and VOR tend to vary with O-D trip distance, using a constant VOT and VOR for a wide range of trip lengths is an unreasonable simplification pertinent to most travel models. For the most accurate predictions, this distinction should be used in demand-forecasting models.

Evidence of Negative Toll Bias

Key Finding

There is a significant negative bias against paying a toll, regardless of the toll amount. This preference is generally supported across travel purposes by RP and SP data, as well as by research in behavioral economics. The estimated toll penalty effect for auto trips is generally equivalent to as much as 15–20 minutes of travel time.

Implications for Policy

The resistance to paying a toll appears to present an obstacle to the effective widespread introduction of congestion pricing policies. However, a pricing policy can be effective even if only a limited proportion of drivers pay the toll, and as with VOT, the resistance to paying any toll at all may vary widely across the population. In that sense, toll bias becomes another dimension of market discrimination, similar to VOT. Resistance can be overcome by a guaranteed superior LOS in terms of travel time savings and improvements in reliability. Tolling existing facilities only to collect revenue, but without a substantial LOS improvement, would generally be perceived negatively by highway users.

Resistance to paying a toll is likely to fade as road pricing becomes more ubiquitous and more convenient. In the past, drivers had to wait in lines to pay tolls, which in itself could explain a good deal of resistance to tolls. Now, electronic tolling has made paying the toll both faster and less noticeable in terms of the amount of money being spent. The more widespread becomes electronic road pricing, the more it can be expected that antitoll bias will reduce.

Implications for Modeling

Antitoll thresholds are avoided in forecasting on the basis that they are not rational in economic terms. Empirically, however, they do appear to be real, so they should be included to obtain the most accurate results, at least for short-term forecasts. This bias will result in a more conservative traffic and revenue forecast if travel time savings are insignificant, but it also may result in a more optimistic forecast for pricing projects that improve travel time significantly. For longer-term forecasts, it may be appropriate to explore scenarios with reduced or eliminated antitoll bias terms.

Hierarchy of Likely Responses to Changes in Tolls and Congestion

Key Finding

Traveler responses to congestion and pricing depend on the range and attractiveness of available alternatives. The models estimated for this project covered a range of travel choices. When possible, nested hierarchical models were estimated to determine which types of choices are most sensitive to travel time and cost changes. The highest propensity for change appears to be between tolled and nontolled routes. A change of route requires little or no adjustment in travel schedule, and the choice can even be made en route. Travelers also show a fairly high propensity for making minor shifts in departure time of an hour or less, since the smaller is the shift, the less rescheduling of activities is required, and the more familiar the traveler is likely to be with the typical traffic conditions over time.

Somewhat less likely are changes in travel mode or car occupancy, which may include switching between auto and transit or between driving alone and ridesharing. Mode shifting is most prevalent for commute trips and other frequent trips for which information about transit services or possible carpoolers is most available or worth investigating.

Less likely responses to changes in congestion or pricing are changes in the choice of destination locations, the rescheduling of trips to very different times of day, and changes in the frequency of making trips from home. These types of changes are the least likely for activities that are most constrained, such as work and school trips or medical appointments. For more flexible types of trips, these types of shifts may actually be more likely than changing the mode of travel.

In the longer term, people may make greater changes as opportunities arise and life-cycle transitions occur. These shifts include changing the number or type of vehicles owned (or both) and the location of home, work, school, and other key travel anchor points relative to one another. The present project outlines an approach to modeling longer-term responses to congestion and pricing by means of accessibility measures that are derived from the estimated primary choice of route, mode, and TOD.

Implications for Policy

Decisions influencing traffic congestion and the cost of driving can affect travel behavior, and the relationships are often complex and can shift over time. This aspect of travel behavior argues for using advanced demand simulation models to guide policy, rather than relying on mental models and experience. The most predictable effects tend to be those that require only minor adjustments, such as choosing to travel at a slightly different TOD. To make pricing policies more effective in tackling congestion, the presence of competitive alternative modes and destinations should be carefully considered. Pricing policies are most effective in combination with transit improvement and smart land use development.

Implications for Modeling

Modeling systems should be able to represent the influences of travel time and cost on all of the types of decisions listed above, and the models should be integrated so that appropriate relative sensitivities are reflected at the different hierarchical levels. These relative sensitivities should also allow for variation in travel segments and travelers. Such modeling requires an activity-based microsimulation model, ideally used in combination with accurate dynamic simulation of traffic congestion.

Summary of User Segmentation Factors

Key Finding

Many potential factors can affect VOT, VOR, or traveler responses to congestion and pricing, including person, household, land use, and travel characteristics. It will never be possible in regional travel models designed for long-term forecasting to account for all the details of user characteristics. However, it is possible to account explicitly for the most important and systematic effects and to apply reasonable assumptions about the probabilistic distributions of VOT and VOR in order to account for the residual heterogeneity.

Implications for Policy

Most of the important factors that affect traveler responses to congestion and pricing are highly differentiated by highway user groups. In calculation of user benefits, the analysis must be implemented with a user segmentation that at a minimum includes trip purpose (work and nonwork); income group and car occupancy (three to four categories each); and commuting distance and household size (two to three categories each). It is highly desirable to account for significant unobserved user heterogeneity and situational variability by applying probabilistic VOT/VOR rather than deterministic VOT/VOR. Simplified methods that operate with an average VOT/VOR are subject to significant aggregation biases and will not adequately portray a pricing project.

Implications for Modeling

Segmentation is crucial for policy evaluation, and modeling systems should be segmented according to the main effects described above. Traditional four-step demand models and static traffic assignments, still the most common tools in practice, are of little use because limited segmentation is one of their major constraints. In addition, it is practically impossible to incorporate distributed parameters in these aggregate constructs. Activity-based models (ABMs) on the demand side and dynamic traffic assignment (DTA) on the network simulation side offer the potential for significantly better platforms for modeling highway congestion and pricing because they are both based on the concept of individual microsimulation.

Avoiding Simplistic Approaches to Forecasting

Key Finding

Although many key effects and tendencies related to the highway utility function are similar across data sets and regions in the United States, many additional effects associated with person types, household composition, transit availability, and land use are specific to each region. Therefore, any simplified surrogate equations or elasticity calculations need to be interpreted and applied with caution.

Implications for Policy

Interregional comparisons and analogies and general rules with respect to expected demand elasticity in relation to congestion and pricing must be applied cautiously. In general, they should not be used for evaluating pricing projects and policies or comparing different pricing alternatives. Properly portraying congestion and pricing effects, as well as the large magnitude of possible impacts (positive or negative), fully justifies a serious modeling approach with a corresponding data collection effort.

Implications for Modeling

The functional forms for the highway utility function developed in the present research should be applied within a framework of regional travel models in which all needed structural inputs and market segments can be supported. Such travel models can fully address regional specifics and take advantage of available data. The best framework is a complete regional travel model system in which an advanced travel demand model (preferably of the activity-based microsimulation type) is integrated with an advanced network simulation tool (preferably DTA with microsimulation of individual vehicles).

Data Limitations and Global Positioning System–Based Data Collection Methods

Key Finding

The availability of data sets adequate to support the analyses undertaken in this study was extremely limited, especially for travel time reliability. This kind of difficulty should decrease, because the use of GPS and probe vehicles and other distributed wireless technologies to collect data on actual travel times and speeds is growing rapidly.

Implications for Policy

With more comprehensive and credible data on travel times and speeds, including measures of travel time reliability, policy makers will have a significantly better basis for advocating new projects and policies, including pricing. The entire issue of improving travel time reliability can finally shift from qualitative analysis to quantitative analysis.

Implications for Modeling

New sources of information are essential for estimation and calibration of travel demand models and network simulation tools. Crude LOS variables created by static assignment procedures have always formed one of the weakest components in travel modeling, frequently manifested in illogical values of model coefficients that must be constrained to ensure reasonable model sensitivities. All travel demand and network simulation models would benefit from better estimates of O-D travel times by TOD. Special benefits would be provided to and could be exploited by advanced models that incorporate travel time reliability measures.

Network Simulation Models to Support Congestion and Pricing Studies

Salient points of the C04 research with respect to network simulation tools include the following:

- Need for Microsimulation. Capturing user responses to pricing and reliability is best accomplished through microsimulation of individual traveler decisions in a network platform; a time-dependent analysis tool is required because the time dimension is essential to evaluating the impact of congestion pricing and related measures. Microsimulation of individual traveler choices provides the most general and scalable approach to evaluate the measures of interest in this study;
- Need for More Robust DTA. The current generation of available simulation-based DTA models only considers fixed, albeit time-varying, O-D trip patterns. Greater use and utility will result from integrating DTA with an activity-based demand model and incorporating user attributes, including systematic and random heterogeneity of user preferences;
- Improved Algorithms for Regional Scale Modeling. Finding equilibrium time-varying flows has been based on the relatively inefficient method of successive averages, a method that does not scale well for application to large metropolitan networks. New implementations of the method of successive averages and other algorithms that exploit the vehicle-based approach of simulation-based DTA have been demonstrated on large actual networks in this research effort;
- Traveler Heterogeneity. Incorporating heterogeneity of user preferences is an essential requirement for modeling user responses to pricing in both travel demand models and network simulation tools. New algorithms that exploit nonparametric multicriteria shortest-path procedures allow VOT to be continuously distributed across users. Efficient implementations of these algorithms have been demonstrated for large network application as part of this study; and
- Network Reliability Measures. In a network simulation model, (1) route choice must include the reliability measures in a way consistent with mode and other choices, and (2) network path–building algorithms must generate O-D measures, along with average travel time and cost, to feed back to the demand model. Two practical approaches are proposed to estimate variability measures of travel time in the context of network assignment tools. The first exploits trajectory information in micro- and mesosimulation tools; the second employs a robust relation established between the first and second moments of the travel time per unit distance. These methods are fully compatible with the adopted functional form of the highway utility and reliability measures like standard deviation of travel time per unit distance.

The proposed integrated model framework is a demonstration of a trip-based integration of a well-calibrated mode choice model in practice and a simulation-based dynamic traffic micro-assignment model. This framework is sufficiently flexible to incorporate other dimensions (e.g., destination and departure time choices) in addition to the mode choice dimension from the demand side, and it can be readily extended to an activity-based integration of demand models and an activity-based dynamic traffic microassignment model.

Dynamic mode share and toll road usage results of the proposed integrated model are demonstrated on the large-scale New York metropolitan network. The convergence of the proposed algorithms is also examined. The proposed model uniquely addresses the needs of metropolitan agencies for prediction of mode and path choices and the resulting network flow patterns, and it can evaluate a wide range of road-pricing scenarios on large-scale networks.

Incorporation of Results in Applied Travel Models

Different model structures offer different options for the inclusion of advanced forms of the highway utility function. Although certain components can be incorporated in any properly segmented model, others, such as travel time reliability measures or probabilistically distributed VOT, impose strict constraints on the model structure. The main related issues of incorporation of the proposed form of the utility function are addressed in the following sections.

Transferability of Model Structures and Parameters Between Regions, Choice Contexts, and Studies

The study results have three levels of generalization: (1) understanding of general rules of travel behavior and identification of major impacts and mechanisms leading to conceptual model structures, (2) mathematical structures of associated choice models and associated forms of the highway utility function, and (3) estimated choice models with the obtained values of coefficients and significance of particular variables.

The first two levels of transferability—model approaches and structures—can be effectively generalized. Most of the functional forms for highway utility were statistically significant in such different regions as New York City and Seattle, and there was agreement between major findings based on RP and SP types of data. However, a direct transfer of model coefficient values from region to region, or from choice context to choice context, is not recommended. For different areas, even similar choice contexts such as trip mode choice versus tour mode choice may require a significant rescaling of parameters. It practice, it also may be difficult to ensure exactly the same level of model segmentation and variable definition.

The best way to transfer a model structure from region to region is to reestimate it based on local data using the model specification in the current study as the prototype. In transferability tests (e.g., from New York to Seattle), the absolute majority of model coefficients that were significant for one region were significant for the other region, although the values varied.

A second-best approach is to recalibrate the model on aggregate local data rather than fully reestimating it in a disaggregate fashion. Recalibration can be done after the model has been implemented and the results have been compared to the aggregate targets externally established for each choice dimension. Recalibration and full reestimation differ in that only a subset of parameters (bias constants that do not interact with any person, household, land use, or LOS variables) are allowed to change.

Using Study Results in Applied Forecasting Models

An applied forecasting model must meet certain requirements that in turn impose objective limitations on the functional forms of highway utility, specifically, travel time reliability measures. The research results of this study are grounded in one or more of four applied modeling contexts:

- Aggregate (Four-Step) Demand Models. Although these models offer a limited framework for incorporating congestion and pricing effects, some of the main features of the highway utility function can be incorporated by including the suggested generalized cost components in the mode choice utilities for highway modes. The mode choice model has to differentiate highway modes by three to four occupancy categories and toll or nontoll route, which would result in six to eight highway modes. After adequate segmentation by trip purpose, income groups, and TOD, several hundred trip tables may be generated. However, any additional segmentation by using person, household, or land use characteristics or adding additional choice models would be impossible.

- ABMs Implemented in a Microsimulation Fashion. These models are characterized by a fully disaggregate structure and rely on individual microsimulation of households and persons. They take full advantage of a detailed level of segmentation by household and person characteristics and can include complicated decision-making chains and behavioral mechanisms. The suggested form of the highway utility can be fully implemented, including route-type, mode, and TOD choices. Variables such as income and parameters like VOT can be continuously distributed to account for unobserved heterogeneity (situational variation).
- Static Traffic Assignment. It is probably impossible to incorporate travel time reliability measures in this framework except by use of simplified proxies. Several simplified approaches can be implemented with these models that are still used by many metropolitan planning organizations and departments of transportation. For example, the perceived highway time concept can be readily incorporated on both the demand and network simulation sides. Some improvements to the current state of the practice can be achieved with a multiclass assignment in which vehicle classes are defined by occupancy, route type, and (possibly) VOT-based groups. This practice, however, may result in more than 20 vehicle classes and long run times for large regional networks.
- DTA with Microsimulation of Individual Vehicles. These models are characterized by a fully disaggregate structure and rely on individual microsimulation of vehicles. Similar to ABMs, they can take full advantage of a detailed level of segmentation by household and person characteristics linked to each vehicle, and they can also incorporate probabilistically distributed VOT to account for unobserved user heterogeneity. With the new technical features described in this study, these models can incorporate the suggested O-D measures of travel time reliability in route choice, as well as generate reliability skims to feed back to the demand model.

The major applications framework for the proposed models primarily takes into account the full regional model framework, although facility- and corridor-level models are also taken into account. For deep understanding and proper modeling of congestion and pricing impacts, a full framework, with chosen and nonchosen alternatives, should be available to both users and nonusers, for which full regional travel data set and model are needed. At both the model estimation stage and the application stage, it is essential to know LOS variables such as travel time, cost, and reliability for nonchoice routes, modes, TOD periods, and destinations.

The most promising long-term direction for DTA modeling is the integration of an activity-based demand model with DTA, in which both models are implemented in a fully disaggregate microsimulation fashion with enhanced typological, temporal, and spatial resolution.

Incorporation of Travel Time Reliability in Operational Models

The incorporation of travel time reliability measures in demand models, and especially in network simulations models, still represents a major challenge, especially if the modeling system is to be practical in terms of run time and data support. In general, there are four possible approaches to quantifying reliability:

- Indirect Measure: This concept is based on statistical evidence that in congestion conditions, travelers perceive each minute with a certain weight. Perceived highway time is not a direct measure of reliability, but it can serve as a proxy for reliability because the perceived weight of each minute spent in congestion is a consequence of associated unreliability.

- First Direct Measure: Time Variability (Distribution) Measures. This direct approach assumes that several independent measurements of travel time are known that allow for forming the travel time distribution and calculation of derived measures, such as buffer time.
- Second Direct Measure: Schedule Delay Cost. According to this concept, the direct impact of travel time unreliability is measured through cost functions (penalties expressed in monetary terms) of being late (or early) compared with the planned schedule of the activity. This approach assumes that the desired schedule is known for each person and activity undertaken in the course of the modeled period.
- Third Direct Measure: Loss of Activity Participation Utility. This approach assumes that each activity has a certain temporal utility profile and that individuals plan their schedules to achieve maximum total utility over the modeled period, taking into account expected travel times. An extended travel time due to unreliability can be associated with a loss of a participation in the corresponding activity. Similar to the schedule delay concept, this approach suffers from data requirements that are difficult to meet in practice.

Current possibilities for incorporating each approach within the specific frameworks of both demand modeling and network simulation and supporting it with the necessary input data are summarized in Table ES.3.

Summary of Recommended Model Parameters

A summary of the recommended (default) values for all coefficients applied in the highway utility function (Equation ES.2) is given in Table ES.4. These parameters are recommended for use in operational models only if a full disaggregate estimation of regional data cannot be implemented. In that case, careful aggregate validation and calibration of the entire model system, including route-type, mode, and TOD choices, will be needed.

Table ES.3. Incorporation of Travel Time Reliability Approaches in Operational Models

Method	Demand Model	Network Simulation
Perceived highway time	Straightforward and does not require structural changes	Straightforward and does not require structural changes
Time distribution (mean variance)	Straightforward and does not require structural changes	Network route choice has to incorporate reliability measures that are not additive by links; this requires explicit route enumeration. O-D reliability measures need to be generated
Schedule delay cost	Preferred arrival time has to be externally specified for each trip	Network route choice has to incorporate reliability measures that are not additive by links; this requires explicit route enumeration. O-D travel time distributions should be generated either analytically or through multiple simulations
Loss of participation in activities	Temporal utility profiles have to be specified for each activity; entire-day schedule consolidation model has to be applied	Network route choice needs to incorporate reliability measures that are not additive by links; this requires explicit route enumeration. O-D travel time distributions have to be generated either analytically or through multiple simulations

Table ES.4. Recommended Coefficient Values

	Model Coefficients								Examples of Population and Travel Characteristics			Derived Measures					
			Distance (mi)														
Travel Purpose	Toll Bias	Time (min)	Linear	Squared	Cost (cents)	SD per mi (min/mi)	Exponent for Income	Exponent for Car Occupancy	Household Income ($/year)	Car Occupancy	Distance (mi)	Time Coefficient with Distance Effect	Toll Bias Equivalent (min)	Cost Coefficient with Income and Occupancy Effects	VOT ($/h)	VOR ($/h)	Reliability Ratio
To Work and Business	–0.85	–0.0425	0.02024	–0.000266	–1.25	–0.625	0.6	0.8	30,000	1.0	5.0	–0.0465	18.3	–0.0026	10.8	29.1	2.69
									30,000	2.0	5.0	–0.0465	18.3	–0.0015	18.9	50.7	2.69
									30,000	3.0	5.0	–0.0465	18.3	–0.0011	26.1	70.2	2.69
									30,000	1.0	10.0	–0.0500	17.0	–0.0026	11.6	14.6	1.25
									30,000	2.0	10.0	–0.0500	17.0	–0.0015	20.3	25.4	1.25
									30,000	3.0	10.0	–0.0500	17.0	–0.0011	28.0	35.1	1.25
									30,000	1.0	20.0	–0.0552	15.4	–0.0026	12.9	7.3	0.57
									30,000	2.0	20.0	–0.0552	15.4	–0.0015	22.4	12.7	0.57
									30,000	3.0	20.0	–0.0552	15.4	–0.0011	31.0	17.5	0.57
									60,000	1.0	5.0	–0.0465	18.3	–0.0017	16.4	44.2	2.69
									60,000	2.0	5.0	–0.0465	18.3	–0.0010	28.6	76.9	2.69
									60,000	3.0	5.0	–0.0465	18.3	–0.0007	39.6	106.4	2.69
									60,000	1.0	10.0	–0.0500	17.0	–0.0017	17.7	22.1	1.25
									60,000	2.0	10.0	–0.0500	17.0	–0.0010	30.7	38.4	1.25
									60,000	3.0	10.0	–0.0500	17.0	–0.0007	42.5	53.2	1.25
									60,000	1.0	20.0	–0.0552	15.4	–0.0017	19.5	11.0	0.57
									60,000	2.0	20.0	–0.0552	15.4	–0.0010	33.9	19.2	0.57
									60,000	3.0	20.0	–0.0552	15.4	–0.0007	46.9	26.6	0.57
									100,000	1.0	5.0	–0.0465	18.3	–0.0013	22.3	60.0	2.69
									100,000	2.0	5.0	–0.0465	18.3	–0.0007	38.9	104.5	2.69
									100,000	3.0	5.0	–0.0465	18.3	–0.0005	53.8	144.5	2.69
									100,000	1.0	10.0	–0.0500	17.0	–0.0013	24.0	30.0	1.25
									100,000	2.0	10.0	–0.0500	17.0	–0.0007	41.8	52.2	1.25
									100,000	3.0	10.0	–0.0500	17.0	–0.0005	57.8	72.2	1.25
									100,000	1.0	20.0	–0.0552	15.4	–0.0013	26.5	15.0	0.57
									100,000	2.0	20.0	–0.0552	15.4	–0.0007	46.1	26.1	0.57
									100,000	3.0	20.0	–0.0552	15.4	–0.0005	63.8	36.1	0.57

(continued on next page)

Table ES.4. Recommended Coefficient Values (continued)

Travel Purpose	Model Coefficients: Toll Bias	Time (min)	Distance (mi): Linear	Distance (mi): Squared	Cost (cents)	SD per mi (min/mi)	Exponent for Income	Exponent for Car Occupancy	Examples of Population and Travel Characteristics: Household Income ($/year)	Car Occupancy	Distance (mi)	Derived Measures: Time Coefficient with Distance Effect	Toll Bias Equivalent (min)	Cost Coefficient with Income and Occupancy Effects	VOT ($/h)	VOR ($/h)	Reliability Ratio
From Work and Business	−0.95	−0.0425	0.02024	−0.000266	−1.44	−0.545	0.6	0.8	30,000	1.0	5.0	−0.0465	20.4	−0.0030	9.4	22.1	2.34
									30,000	2.0	5.0	−0.0465	20.4	−0.0017	16.4	38.4	2.34
									30,000	3.0	5.0	−0.0465	20.4	−0.0012	22.7	53.1	2.34
									30,000	1.0	10.0	−0.0500	19.0	−0.0030	10.1	11.0	1.09
									30,000	2.0	10.0	−0.0500	19.0	−0.0017	17.6	19.2	1.09
									30,000	3.0	10.0	−0.0500	19.0	−0.0012	24.3	26.6	1.09
									30,000	1.0	20.0	−0.0552	17.2	−0.0030	11.2	5.5	0.49
									30,000	2.0	20.0	−0.0552	17.2	−0.0017	19.4	9.6	0.49
									30,000	3.0	20.0	−0.0552	17.2	−0.0012	26.9	13.3	0.49
									60,000	1.0	5.0	−0.0465	20.4	−0.0020	14.3	33.4	2.34
									60,000	2.0	5.0	−0.0465	20.4	−0.0011	24.8	58.2	2.34
									60,000	3.0	5.0	−0.0465	20.4	−0.0008	34.4	80.5	2.34
									60,000	1.0	10.0	−0.0500	19.0	−0.0020	15.3	16.7	1.09
									60,000	2.0	10.0	−0.0500	19.0	−0.0011	26.7	29.1	1.09
									60,000	3.0	10.0	−0.0500	19.0	−0.0008	36.9	40.3	1.09
									60,000	1.0	20.0	−0.0552	17.2	−0.0020	16.9	8.4	0.49
									60,000	2.0	20.0	−0.0552	17.2	−0.0011	29.5	14.6	0.49
									60,000	3.0	20.0	−0.0552	17.2	−0.0008	40.8	20.1	0.49
									100,000	1.0	5.0	−0.0465	20.4	−0.0014	19.4	45.4	2.34
									100,000	2.0	5.0	−0.0465	20.4	−0.0008	33.7	79.1	2.34
									100,000	3.0	5.0	−0.0465	20.4	−0.0006	46.7	109.4	2.34
									100,000	1.0	10.0	−0.0500	19.0	−0.0014	20.8	22.7	1.09
									100,000	2.0	10.0	−0.0500	19.0	−0.0008	36.3	39.5	1.09
									100,000	3.0	10.0	−0.0500	19.0	−0.0006	50.1	54.7	1.09
									100,000	1.0	20.0	−0.0552	17.2	−0.0014	23.0	11.4	0.49
									100,000	2.0	20.0	−0.0552	17.2	−0.0008	40.0	19.8	0.49
									100,000	3.0	20.0	−0.0552	17.2	−0.0006	55.4	27.3	0.49

(continued on next page)

Table ES.4. Recommended Coefficient Values (continued)

	Model Coefficients								Examples of Population and Travel Characteristics			Derived Measures					
Travel Purpose	Toll Bias	Time (min)	Distance (mi) Linear	Distance (mi) Squared	Cost (cents)	SD per mi (min/mi)	Exponent for Income	Exponent for Car Occupancy	Household Income ($/year)	Car Occupancy	Distance (mi)	Time Coefficient with Distance Effect	Toll Bias Equivalent (min)	Cost Coefficient with Income and Occupancy Effects	VOT ($/h)	VOR ($/h)	Reliability Ratio
Nonwork	−1.2	−0.0335	0	0	−0.5228	−0.418	0.5	0.7	30,000	1.0	5.0	−0.0335	35.8	−0.0030	6.7	16.6	2.50
									30,000	2.0	5.0	−0.0335	35.8	−0.0019	10.8	27.0	2.50
									30,000	3.0	5.0	−0.0335	35.8	−0.0014	14.4	35.9	2.50
									30,000	1.0	10.0	−0.0335	35.8	−0.0030	6.7	8.3	1.25
									30,000	2.0	10.0	−0.0335	35.8	−0.0019	10.8	13.5	1.25
									30,000	3.0	10.0	−0.0335	35.8	−0.0014	14.4	17.9	1.25
									30,000	1.0	20.0	−0.0335	35.8	−0.0030	6.7	4.2	0.62
									30,000	2.0	20.0	−0.0335	35.8	−0.0019	10.8	6.7	0.62
									30,000	3.0	20.0	−0.0335	35.8	−0.0014	14.4	9.0	0.62
									60,000	1.0	5.0	−0.0335	35.8	−0.0021	9.4	23.5	2.50
									60,000	2.0	5.0	−0.0335	35.8	−0.0013	15.3	38.2	2.50
									60,000	3.0	5.0	−0.0335	35.8	−0.0010	20.3	50.7	2.50
									60,000	1.0	10.0	−0.0335	35.8	−0.0021	9.4	11.8	1.25
									60,000	2.0	10.0	−0.0335	35.8	−0.0013	15.3	19.1	1.25
									60,000	3.0	10.0	−0.0335	35.8	−0.0010	20.3	25.4	1.25
									60,000	1.0	20.0	−0.0335	35.8	−0.0021	9.4	5.9	0.62
									60,000	2.0	20.0	−0.0335	35.8	−0.0013	15.3	9.5	0.62
									60,000	3.0	20.0	−0.0335	35.8	−0.0010	20.3	12.7	0.62
									100,000	1.0	5.0	−0.0335	35.8	−0.0017	12.2	30.3	2.50
									100,000	2.0	5.0	−0.0335	35.8	−0.0010	19.8	49.3	2.50
									100,000	3.0	5.0	−0.0335	35.8	−0.0008	26.2	65.5	2.50
									100,000	1.0	10.0	−0.0335	35.8	−0.0017	12.2	15.2	1.25
									100,000	2.0	10.0	−0.0335	35.8	−0.0010	19.8	24.6	1.25
									100,000	3.0	10.0	−0.0335	35.8	−0.0008	26.2	32.7	1.25
									100,000	1.0	20.0	−0.0335	35.8	−0.0017	12.2	7.6	0.62
									100,000	2.0	20.0	−0.0335	35.8	−0.0010	19.8	12.3	0.62
									100,000	3.0	20.0	−0.0335	35.8	−0.0008	26.2	16.4	0.62

Note: SD = Standard deviation.