MID-AMERICA
TRANSPORTATION CENTER

Report # MATC-UI: 478

Final Report
25-1121-0001-478

University of Nebraska Lincoln

K-STATE
Kansas State University

KU
THE UNIVERSITY OF
KANSAS

MISSOURI
S&T
University of Science & Technology

THE UNIVERSITY OF IOWA

LINCOLN University

PRAIRIE VIEW A&M UNIVERSITY
1876

Advanced Decision Modeling for Real Time Variable Tolling – Development and Testing of a Data Collection Platform

Paul Hanley, Ph.D.
Associate Professor
Department of Civil and Environmental Engineering
Public Policy Center
University of Iowa

Hai Yu, Ph.D. Candidate
Graduate Research Assistant
Nikhil Sikka, Ph.D.
Senior Analyst, Environmental Resource Group, Inc.
Suyun Ma, Ph.D. Candidate
Graduate Research Assistant

THE UNIVERSITY OF IOWA

2012

A Cooperative Research Project sponsored by the U.S. Department of Transportation Research and Innovative Technology Administration

MATC

The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the information presented herein. This document is disseminated under the sponsorship of the Department of Transportation University Transportation Centers Program, in the interest of information exchange. The U.S. Government assumes no liability for the contents or use thereof.

Advanced Decision Modeling for Real Time Variable Tolling – Development and Testing of a Data Collection Platform

Paul Hanley, Ph.D.
Associate Professor
Department of Civil and Environmental Engineering
Public Policy Center
University of Iowa

Hai Yu, Ph.D. Candidate
Graduate Research Assistant
Department of Civil and Environmental Engineering
Public Policy Center
University of Iowa

Nikhil Sikka, Ph.D.
Senior Analyst
Environmental Resource Group, Inc.

Suyun Ma, Ph.D. Candidate
Graduate Research Assistant
Department of Civil and Environmental Engineering
Public Policy Center
University of Iowa

A Report on Research Sponsored by

Mid-America Transportation Center

University of Nebraska-Lincoln

2012

Technical Report Documentation Page

1. Report No.
25-1121-0001-478

2. Government Accession No.

3. Recipient's Catalog No.

4. Title and Subtitle
Advanced Decision Modeling for Real Time Variable Tolling – Development and Testing of a Data Collection Platform

5. Report Date
2012

6. Performing Organization Code

7. Author(s)
Paul Hanley, Nikhil Sikka, Hai Yu, Suyun Ma

8. Performing Organization Report No.
25-1121-0001-478

9. Performing Organization Name and Address
Mid-America Transportation Center
2200 Vine St.
PO Box 830851
Lincoln, NE 68583-0851

10. Work Unit No. (TRAIS)

11. Contract or Grant No.

12. Sponsoring Agency Name and Address
Research and Innovative Technology Administration
1200 New Jersey Ave., SE
Washington, D.C. 20590

13. Type of Report and Period Covered
Final Report,
January 2012-June 2012

14. Sponsoring Agency Code

15. Supplementary Notes

16. Abstract

Our current ability to forecast demand on tolled facilities has not kept pace with advances in decision sciences and technological innovation. The current forecasting methods suffer from lack of descriptive power of actual behavior because of the simplifications used in current economic decision models. These simplifications are in part due to the historical limitations on data collection. Today, we are seeing advances in the data collection technology that captures naturalistic behavior and this study is seeking to develop and test such technology. This will be the first phase of a naturalistic driving study on the topic of variable road tolling and decision making.

This investigation extends the state of knowledge of decision modeling under risk and ambiguity by developing a mobile data collection platform for the capturing of naturalistic choice outcomes, associated environmental states, decision makers' self-articulated perceptions of risk and assessments of ambiguity, and socio-economic attributes. This software/hardware development is the first step in advancing our ability to forecast future revenue sources in transportation. The data collection platform will be used in an extending study and will allow the study team to obtain insights for developing, testing and implementing new behavioral models that explicitly describe how we use imprecise information—in this case, ambiguous information signaled by variable toll lane charges—that are superior to those obtained to date.

To support the empirical observations of driver choice of diverting to HOT or non-HOT lanes in real-time, this study documents the development and testing of an in-vehicle perception acquisition device. The device is based on smart-phone technology and its software is designed using the latest IOS release. Drivers will be able to upload the software application onto their existing smartphone or tablet. The widespread acceptance and use of smartphones and tablets allows the collection and transmission of decision data collected in real-time without installing or altering the drivers' vehicles in any manner. This both increases the convenience—and perhaps the use of the application—and reduces the cost of data collection.

17. Key Words
Add key words here

18. Distribution Statement

19. Security Classif. (of this report)
Unclassified

20. Security Classif. (of this page)
Unclassified

21. No. of Pages
35

22. Price

Table of Contents

List of Figures

List of Tables

List of Abbreviations

Mid-America Transportation Center (MATC)
Nebraska Transportation Center (NTC)
High Occupancy Toll (HOT)

Disclaimer

The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the information presented herein. This document is disseminated under the sponsorship of the U.S. Department of Transportation’s University Transportation Centers Program, in the interest of information exchange. The U.S. Government assumes no liability for the contents or use thereof.

Abstract

Our current ability to forecast demand on tolled facilities has not kept pace with advances in decision sciences and technological innovation. The current forecasting methods suffer from lack of descriptive power of actual behavior because of the simplifications used in current economic decision models. These simplifications are in part due to the historical limitations on data collection. Today, we are seeing advances in the data collection technology that captures naturalistic behavior and this study is seeking to develop and test such technology. This will be the first phase of a naturalistic driving study on the topic of variable road tolling and decision making.

This investigation extends the state of knowledge of decision modeling under risk and ambiguity by developing a mobile data collection platform for the capturing of naturalistic choice outcomes, associated environmental states, decision makers' self-articulated perceptions of risk and assessments of ambiguity, and socio-economic attributes. This software/hardware development is the first step in advancing our ability to forecast future revenue sources in transportation. The data collection platform will be used in an extending study and will allow the study team to obtain insights for developing, testing and implementing new behavioral models that explicitly describe how we use imprecise information—in this case, ambiguous information signaled by variable toll lane charges—that are superior to those obtained to date.

To support the empirical observations of driver choice of diverting to HOT or non-HOT lanes in real-time, this study documents the development and testing of an in-vehicle perception acquisition device. The device is based on smart-phone technology and its software is designed using the latest IOS release. Drivers will be able to upload the software application onto their existing smartphone or tablet. The widespread acceptance and use of smartphones and tablets

allows the collection and transmission of decision data collected in real-time without installing or altering the drivers’ vehicles in any manner. This both increases the convenience—and perhaps the use of the application—and reduces the cost of data collection.

Chapter 1 Decision Science and Transportation Management and Funding through High Occupancy Tolled Facilities

High occupancy tolled (HOT) lanes are designed to affect overall traffic congestion by providing an alternative in which travel speeds are substantially faster than the general conditions. The alternative is provided as separated—either by a physical barrier or pavement marking—travel lanes from general non-HOT lanes. The means by which the higher speeds are maintained is through rationing: rationing as determined by directly charging drivers and supported by neoclassical economic theory for pricing a scarce resource. The charge, called a toll, is set dynamically based on traffic volume in the HOT as well as on the non-HOT lanes. Therefore, the foundation of HOT pricing is the collective decisions of individual drivers.

This study begins to address the limitation of the neoclassical economic analysis for forecasting driver decisions, which adversely impacts the forecasting of revenues, congestion relief, and social surplus of travel. The forecasting of revenues is a concern to the facility owner and operator because of the desire to generate enough revenue to cover capital, operation and maintenance costs, as well as generating a surplus that can be used to subsidize other ancillary expenses such as mass transit. The second two, congestion relief and social surplus of travel, are less tangible in that they impact the aggregate economy in which the individual drivers perceive direct and substantial benefits. That is, a one minute reduction in travel time due to congestion would net the typical driver in the Chicago Metro $0.25 whereas the Chicago Metro as a whole would see a savings of $14,000. Perhaps the increment of saving is too small for the driver to perceive as a benefit—the cost of congestion in the Chicago Metro area is based on a study sponsored by the Metropolitan Planning Council (HDR 2008). Neoclassical economic theory is silent on the discontinuities generated by perceived benefits as thresholds that play out in applications. Additionally, the rational theory that underlies HOT pricing would treat the drivers' decision in absence of time constraints and with a time invariant discount rate.

Making a beneficial decision in a dynamic and hard to predict environment, such as that experienced on congested highways, is an enormous challenge. Not only does it require learning from past experience, but it also requires anticipating what might happen under circumstances not previously encountered.

1.1 Decision Science Theories that are Applicable to Driver Choice of HOT Facilities

1.1.1 Expected Utility Theory

Expected Utility Theory (EUT) has been widely used to understand drivers' behavioral responses to travel time uncertainty. EUT states that in situations involving uncertainty, the decision maker chooses outcomes on the basis of their expected utility values, i.e., the weighted sums of the utility values of outcomes multiplied by their respective probabilities. As such, the decision maker selects the alternative with the maximum utility (von Neumann and Morgenstern 1944; Einhorn and Hogarth 1981). However, after an extensive application of the conventional expected utility models in modeling uncertain behaviors, a conclusion was reached that decision makers often do not make choices in a way consistent with the utility models.

1.1.2 Random Utility Theory

Conventional Random Utility Model assumes that travelers have perfect information and show rational behavior to maximize their utility and is concerned with the valuation of certain and riskless outcomes. Rationality of travelers has been challenged in many studies (Fujii and Kitamura 2000; Avineri and Prashker 2004; Bogers et al. 2007). But developments in the random utility model framework have incorporated expected utility principles to model individual travel choice under risk (Noland and Small 1995; Polak et al. 2008). Maximum Expected Utility

Theory was introduced as one of the standard approaches to account for travel time uncertainty in terms of risk (Small et al. 1999; Bates et al. 2001). Although expected utility has been extensively used within random utility model and a linear utility specification has been the dominant approach to account for risk in travel time occurrences, recent investigation on individual heterogeneity in value of travel time by using a mixed multinomial model is showing non-linearity in both utility specification and probability weighting (Hensher et al. 2011). Furthermore, Sikka and Hanley 2012 add behavioral realism in an existing RUM framework by considering behavioral models available not only for risk analysis but also for ambiguity. In their research, they used a stated preference experimental protocol to simultaneously elicit people's attitude toward both risk and ambiguous routes. Especially, the specific hypothesis is that drivers do not always make rational decisions in route-choice situations and factors such as uncertainty in travel time and monetary cost play a significant role in route selection, and they further derived Willingness to Pay measures that take into account drivers' attitudes toward uncertainty and travel time variability (Sikka and Hanley 2012).

1.1.3 Cumulative Prospect Theory

Cumulative Prospect Theory has become the leading contender in an increasing trend in transportation studies toward the adaption of behavioral economics, cognitive sciences and psychology to travelers' decision making modeling. Behavioral and cognitive science has advanced knowledge regarding how people think about and make choices and how to mathematically represent these behaviors. The roots of behavioral theories lie in psychology, cognitive science, and neuroscience, and have rapidly grown with maturing technology such as functional magnetic resonance imaging, which allows for the testing and quantification of these theories (Sikka and Hanley 2012). Cumulative prospect theory and its predecessor, Prospect

Theory, were developed by Kahneman and Tversky as an alternative to expected and random utility theory for describing actual behavior (Kahneman and Tversky 1979; Tversky and Kahneman 1992). The basis of the two theories is that people are not strictly rational and frequently violate the axioms of rational theory due to cognitive limitations. The two principles, diminishing sensitivity and loss aversion, are invoked to explain the characteristic curvature of the value function and the weighting functions. Prospect theory departs from the tradition that assumes the rationality of economic agents; it is proposed as a descriptive, not a normative, theory. The idealized assumption of rationality in economic theory is commonly justified on two grounds: conviction that only rational behavior can survive in a competitive environment, and the fear that any treatment that abandons rationality will be chaotic and intractable. Both arguments are questionable. First, the evidence indicates that people can spend a lifetime in a competitive environment without acquiring a general ability to avoid framing effects or to apply linear decision weights. Second, perhaps more important, the evidence indicates that human choices are orderly, although not always rational in the traditional sense of this word. Cumulative prospect theory extends the theory in several respects. Difference between cumulative prospect and prospect theory is that weighting is applied to the cumulative probability distribution functions, as in rank-dependent expected utility theory but not applied to the probabilities of individual outcomes. This leads to overweighting of extreme events which occur with small probability, rather than to an overweighting of all small probability events. The main modification helps to avoid a violation of first order stochastic dominance and makes the generalization to arbitrary outcome distributions easier. Thus, concerning economic theory, Cumulative Prospect Theory is an improvement over Prospect Theory.

Cumulative prospect theory has received attention and been frequently applied in travel behavior research. A traveler is not assumed to be indifferent between a choice that may produce a positive outcome, called a gain, and one that may have a negative outcome, called a loss. This theory allows for modeling travelers' tendency to select a route with more consistent and known travel times over a route that may be faster but has the potential to be much slower under certain conditions. Kahneman and Tversky 1992 define loss aversion as people's tendency to value a loss more than they value an equivalent gain. In addition, based on empirical findings, when people are presented with an outcome that is highly likely to occur they tend to perceive the probability as being smaller, thereby reducing the likelihood they select the outcome compared to what a rational theory would suggest. A similar tendency is present with very low probabilities – people will perceive the chance of obtaining the outcome much greater than the theoretical probability would predict. This flexibility is needed to better forecast route-choice in traffic networks, which are inherently uncertain not only due to random disruptions like accidents, vehicle breakdowns, and weather closures, but also due to planned interruptions, construction and maintenance activities, and community and social events (Sikka and Hanley 2012). Furthermore, Sikka and Hanley 2012 compared drivers' choices between a general purpose and a tolled lane based on cumulative prospect theory and expected utility theory. They found that expected utility method overestimates the probability of a driver switching to a tolled lane, and the striking difference between applications of the two theories is a result of drivers overweighting the extreme probabilities for losses. The overestimation is the result of the traditional models that reveal how a rational driver ought to act. Cumulative prospect theory reflects drivers' hesitancy toward changing from a known state – general purpose lane – even if they are likely to sustain a loss by not switching. Finally, Sikka and Hanley 2012 proposed that cumulative prospect theory be the analytic framework that is applied to travelers' route-choice addresses to correct errors of applying traditional prescriptive models.

1.1.4 Expectancy-Valence Theory

Expectancy-Valence theory was proposed by Busemeyer and Stout 2002 to capture the outcomes for repeated decisions. The outcome is a function of the person's current evaluation of experience and their previous outcome from the same decision. It is not assumed that the person making the decision has a priori probabilities of outcomes as the traditional neoclassic rational theories suggest. Rather, the person acquires the relevant probabilities over time with each new experience updating the person's expectation of outcomes. Similar to Bayesian updating, the person may begin with an arbitrary level of knowledge (probability estimate) and with each experience builds toward stable subjective expectations. The theory is robust to account for people's subjective weighting of loss more heavily than similar magnitude gains; the ability to variably weigh recent versus time distant experiences; and capturing non-linear utility. The advance from Cumulative Prospect theory is its ability to track the change in subjective probabilities for outcomes.

Chapter 2 The Context of High Occupancy Tolled Lanes for Application of Decision Science

A variable toll provides information that is used by drivers to reduce travel time uncertainty. The information signaled by the toll amount is either perceived as a gain or loss; a gain if the drivers believe the information costs less than the travel time savings and as a loss if they believe that using the tolled lane will not save them time. Additionally, in the case of a gain the driver must value the travel time savings as being in excess of the information cost, whereas, it would be viewed as a loss if the travel time savings does not exceed the cost of information.

Complicating the decision processes is the quality of the information, as perceived by the driver, conveyed by the toll in reducing the travel time uncertainty. A perfect signal would require the toll to flawlessly correlate with travel speed—assuming every driver values monetary amounts identically, the chosen increment that the toll increases would map precisely to the increase in travel speed, for example, every one dollar increase guarantees an increase of one mile per hour. Should the information perfectly signal travel time savings, that is reducing uncertainty to zero, and the drivers are rational, then expected or random utility theory could be applied using fixed risk averse, risk seeking or risk neutral utility functions. However, when the information is imperfect, containing uncertainty, the drivers will react differently based on the level of perceived imperfection.

Information uncertainty occurs when drivers either do not understand the exact relationship between the toll and speed, assuming it exists; do not trust that the relationship holds, perhaps based on experience using the tolled lanes; or do not have experience with the given toll lanes. Under this condition, traditional utility theory does not model actual choice behavior sufficiently. Cumulative prospect theory, the theory descriptive of actual choice behavior, posits that drivers will react to uncertainty differently based on their current condition and perceived likelihood of improving their situation. In a dynamic environment with the opportunity to repeat similar—if not identical—decision tasks, expectancy-valence theory may be a superior applied model for HOT lane evaluation.

Under cumulative prospect theory, there is a mid-range of uncertainty, taken as 25% to 75%, within which drivers will react with risk aversion to perceived improvements, whereas if the drivers view the toll as a loss, they will react with risk seeking behavior (Van Lint and van Zuylen 2005). Under extreme uncertainty in the information, the drivers would swap their reactions to risk. That is when the toll is signaling a travel time savings relative to the general purpose lanes with a chance above or below 25%, drivers will be risk seeking and if the toll signals slower travel times they will be risk averse. Table 2.1 summarizes the reactions of drivers to possible outcomes, both gains and losses, to tolled lanes for given ranges of probabilities. The following discussion expands the summary.

Table 2.1 Drivers' Reaction to Outcomes within Mid-range and Extreme Probabilities of Occurrence

		Outcome	
		Slower travel speed (loss)	**Faster travel speeds (gain)**
Probability of Outcome Occurring	**Mid-range of perceived information uncertainty (25 to 75%)**	Risk seeking – more likely to select the toll lane if they are currently in a general purpose lane and traveling at a slower than desired speed. They will gamble on switching lanes, paying the toll for the potential to improve their travel speed.	Risk averse – more likely to stay in the general purpose lane when traveling at a desired speed even if they could travel faster in the tolled lane. They will not gamble on switching lanes, paying the toll for the potential to improve their travel speed.
	Extreme range of perceived information uncertainty (above or below 25%)	Risk averse – more likely to stay in the general purpose lane even if they are traveling at a slower than desired speed. They will not gamble on switching lanes, paying the toll for the potential to improve their travel speed.	Risk seeking – more likely to select the toll lane if they are currently in a general purpose lane. They will gamble on switching lanes, paying the toll for the potential to improve their travel speed.

Why would drivers be risk averse under mid-range levels of uncertainty of improving their travel speed by switching to the tolled lane from the general purpose lane? When drivers are in the general purpose lanes, according to the cumulative prospect theory they will react in a risk averse manner toward switching into the tolled lane if they perceive they will be traveling near an acceptable speed, gauged by past experience or relative to the posted speed limit. Typically, consumers will tend not to gamble on improving their situation when the gamble has a moderate chance of producing a gain because they shy away from the chance of decreasing their positive situation. People prefer the known and are content when we are pleased.

Why would a driver be risk averse under extreme uncertainty of improving their travel speed by switching to the tolled lane? When drivers view that staying within the general purpose lane will be much slower than the tolled lane, as measured by travel speed, switching into the tolled lane would be appealing because they seek to reduce a large perceived loss if they stay in the general lane. They are willing to pay the toll price with the hope, although small, that switching into the toll lane will be faster and protect them from the larger loss of the general lane. This parallels consumers' choice of purchasing insurance to safeguard a large loss even though there is an extremely small probability that the loss will occur.

Why would a driver be risk seeking under mid-range levels of uncertainty of improving their travel speed by switching to the tolled lane? When drivers view the general purpose lanes as moving slower than they desire, which could be based on their experience or be based on the rate of speed related to the posted speed limit, they will be more likely to switch into the tolled lane if they believe there is a mid-range chance that the tolled lane will be faster. In this case, the drivers are saying that they are already losing by staying in the general lane and will continue to lose if they stay in the lane, so why not take a chance by paying the toll and switching lanes. People tend to look for a way to reduce known losses by taking more chances at improving their situation even though it is possible to increase their loss (optimism bias).

Why would a driver be risk seeking under extreme levels of uncertainty of improving their travel speed by switching to the tolled lane? Drivers will tend to select the tolled lane when they believe they will gain much higher travel speeds by switching into the tolled lane than staying within the general purpose lane. Similar to the tendency for consumers to ignore the purchase price of a lottery ticket even though the ticket has a very low probability of winning, the drivers will select the tolled lane hoping for the larger pay-off of fast speeds.

2.1 Discussion and Proposed Extension of Research

The introduction of tolled lanes as a means to reduce congestion, thereby, functionally increasing their capacity and their ability to supplement the fuel tax revenue, highlights the need to move from proscriptive theories, such as expected and random utility theory, which detail how rational drivers should decide, to a behavioral model. The move should be toward a descriptive behavioral theory, such as cumulative prospect and expectancy-valance theory that will allow for improved lane usage and revenue forecasts. Today, the most popular descriptive theory of decision making under risk and uncertainty is cumulative prospect theory. Its advantages over standard expected utility and random utility theories are in its ability to forecast decisions based on observed outcomes that are grounded in observed psychological behavior. Cumulative prospect theory encompasses risk aversion and risk seeking, differing actions depending on perceived negative or positive outcome, and our misapplication of extreme probabilities—it accounts for framing effects and overweighting and underweighting of extreme probabilities. The cumulative prospect theory now has a firm axiomatic foundation and its formulation makes it tractable.

As illustrated in the simple scenarios presented in Sikka and Hanley 2012, ignoring the tendency to overweight low probabilities when negative outcomes are possible greatly impacts forecasts. The behavioral model examples show that when drivers perceive a delay to be equally as likely as a travel time savings, there is a likelihood to greatly overestimate the probability that drivers will switch from a general purpose to a tolled lane. The overestimation is the result of the traditional models that reveal how a rational driver ought to act. Congestion and revenue forecast would suffer in accuracy. Granted, the illustrations above are simplified in the sense that the model parameters are those accepted in the literature, but are not specifically derived for the context of lane tolling.

To overcome the limitations of using existing parameters, Sikka (2012 dissertation) estimated parameters within the context of variable tolling. Sikka completed a laboratory experiment using conjoint analysis, which used a survey instrument to assess behavioral responses and preferences. The purpose behind conducting these experiments was to determine the independent influences on observed behavior and their use of varying tolls as information regarding likely outcomes—either delays or travel time savings. The independent influences in a route-choice will include travel time, travel cost, travel time savings, and variability. More detailed discussion on conjoint experiments can be found in Louviere 1994, Hensher et al. 2005, and Bliemer and Rose 2006.

A recent challenger to cumulative prospect theory is expectancy-valance theory. Within this modeling framework, drivers learn from past experiences and update existing estimates for likely outcomes. Expectancy-valance theory is generalizable to incorporate much of the flexibility of cumulative prospect theory, such as unequal weighting of losses and gains and risk aversion. Therefore, advancement in applied analysis of drivers' choice of HOT facilities, which is a repeated decision, will feature in the exploration of behavioral experience-based theories.

Chapter 3 Experimental Approach for Evaluating Drivers' Decisions to Use HOT Facilities

3.1 Using Description-based versus Experience-based Experiments to Elicit Decisions

The psychology literature categorizes decision theories as static or dynamic. That is, delineating whether the decision maker enters and then leaves the decision process with the same subject outcome probability. If the estimate is stable after living the experience then the theory falls under a static paradigm. Cumulative prospect theory, even with its ability to capture cognitive biases, offers such a static model in that the person is facing a one-time decision and has a prior estimate of the likelihood of the outcome. When a theory describes a person who faces repeated decisions of exact or similar situations and the person learns from the experience by updating their estimates of likelihoods, the theory falls within a dynamic category. The theory of Expectancy-Valance belongs to this category. These categories are also labeled as description-based for static processes and as experience-based for dynamic processes.

The task of driving and choosing a route to a destination can follow either paradigm. If the person is traveling in an unknown area or perhaps is taking a very infrequent trip, the driver could fall under a static theory and, therefore, the decision process could be elicited by a description-based experiment. That is, using such procedures as stated preference surveys, the driver can be provided with scenarios, based in reality, which use actual travel times between known origins and destinations. The driver is provided with statistics that describe their trip (as in Sikka and Hanley 2012) and the model parameters can be estimated using statistical techniques.

However, the driving environment is dynamic. The parameters that determine travel times are time varying and stochastic—influenced by season, weather, crashes, and construction to list a few. A person is most likely repeating a route multiple times a week or month and therefore, it is posited that they will be updating their subject probabilities with experience. Under this scenario, in the onset, the appropriate model would be from the dynamic paradigm. This holding, the drivers' decision parameters must be elicited through repeated decisions using experience based techniques.

In particular, Kudryavtsev and Pavlodsky 2012 explored behavior in two classes of decision tasks: description-based tasks based on description of exact probabilities and magnitudes of outcomes, and experience-based tasks based on people's past experience without the benefit of such descriptions. They conducted experiments to compare the prediction power of a number of decision learning models in both kinds of tasks when they focused on individual, rather than aggregate, behavioral characteristics. Prospect theory, Expectancy-Value model (EVL) and three combinations of these well-established models are involved. Furthermore, they found that models involving linear weighting of gains and losses perform better in both kinds of tasks, from the point of view of generalizability and individual parameter consistency. And they conclude that overall, when both prospects are mixed, the assumption of diminishing sensitivity does not improve models' prediction power for individual decision makers. Finally, for some of the models' parameters, they document consistency at the individual level between description- and experience-based tasks (Kudryavsev and Pavlodsky 2012).

Chapter 4 Prototype Data Collection System

This investigation extends the state of knowledge of decision modeling under risk and ambiguity by developing a mobile data collection platform (in-vehicle Perception Acquisition Device) The main purpose of the platform is to capture naturalistic choice outcomes, associated environmental states, decision makers' self-articulated perceptions of risk and assessments of ambiguity, and socio-economic attributes. This software/hardware development is the first step in advancing our ability to forecast future revenue sources in transportation. The data collection platform will be used in an extending study in which insights, that are superior to those obtained to date, will be obtained for developing, testing and implementing new behavioral models. The study will explicitly describe how we use imprecise information, in this case, ambiguous information signaled by variable toll lane charges.

Using smart phones, drivers' real-time choices, with real-world consequences, data will be obtained that is related to route-choice, specifically the choice of switching into a tolled lane or not. The dynamic data will include: time and location of choice-making; price of the variable toll at time and location; and verbal response to prompts regarding why they did or did not switch to the toll lane. With audio prompts, the system requests the drivers' guesses at the expected delay—or expected time savings—and whether they think the current toll is a bargain for the conditions. In order to reduce driving distractions when other pressing matters need attention, drivers are asked, but not forced, to respond the prompts. The dynamic data is merged with the socio-economic attributes of the drivers along with other known attributes of the trip underway, such as trip type and number of passengers. Additional traffic data obtained via the web portal run by the Minnesota Department of Transportation, such as toll being charged, vehicle volume by location and time of day, will be collected as external and objective measures.

4.1 Data Collection Equipment Development and Testing

In this funded study, the customized smart phone and tablet application was developed for technology running IOS 5 or later. This prototype application runs on iPhones and iPads and requires a cellular data service. After the software development was completed, it was field tested on U.S. 218 in Johnson County, IA. A virtual HOT lane facility was created on U.S. 218 that replicates the I-394 HOT facility in the Minneapolis Metropolitan Area. The replicated HOT lane mimicked critical attributes of operating variable tolled lanes in a congested network.

A data collection center was designed and implemented for the study. The site is located at the Public Policy Center, University of Iowa. The data from the data collection device was transmitted to the collection center and analyzed to streamline the functionality of the mobile data collection device.

4.1.1 System States of the Application

Figure 4.1 illustrates the states of the software application that runs on the data collection device, which is currently either an iPhone or iPad, with a future application for Android operating systems. The flow through the application begins prior to S_1 when the application is running in the background checking if the HOT lane facility is approaching. Specifically, the application utilizes the device's hardware to first orient within the cellular network and as the device approaches the facility it references the built-in GPS hardware for more precise spatial definition. In all cases, the device does not record the location of the device when it is outside the HOT facility to prevent invasion of privacy concerns. At S_1, the system begins to orient itself with the aid of the GPS and checks for its relative position with known HOT facility features, such as the dynamic signs that announce the current toll level. Once the device passes under a toll sign, it generates a time stamp for use in modeling the time a driver had to make the decision to use or not use the HOT lane. At this point, the driver hears an audio prompt asking for an estimate of how much time will be saved if the HOT lane is taken. The system then activates the microphone and records the driver's verbal response. The microphone only remains active for a given amount of time and when the driver does respond due to other demands for attention, the system will record a missing data point. See a following section that summarizes the safety of using audio prompts in in-vehicle devices. The system stays at S_2 until it detects that the device has past the entrance of the HOT lane. At that point, the system enters state S_3 where the driver hears an audio prompt asking if the driver did select to take the HOT lane. The second prompt is required because the horizontal accuracy of the hardware is not accurate enough to determine in which lane of travel the vehicle is located (as of the application development, the horizontal accuracy was five meters). The system waits a predetermined amount of time before turning off the microphone and enters state S_4. The data collected in S_2 and S_3 is stored on the device and prepared for export from the device. The flow through states S_1 to S_2 continues for the duration of the HOT facility and then the device turns itself off.

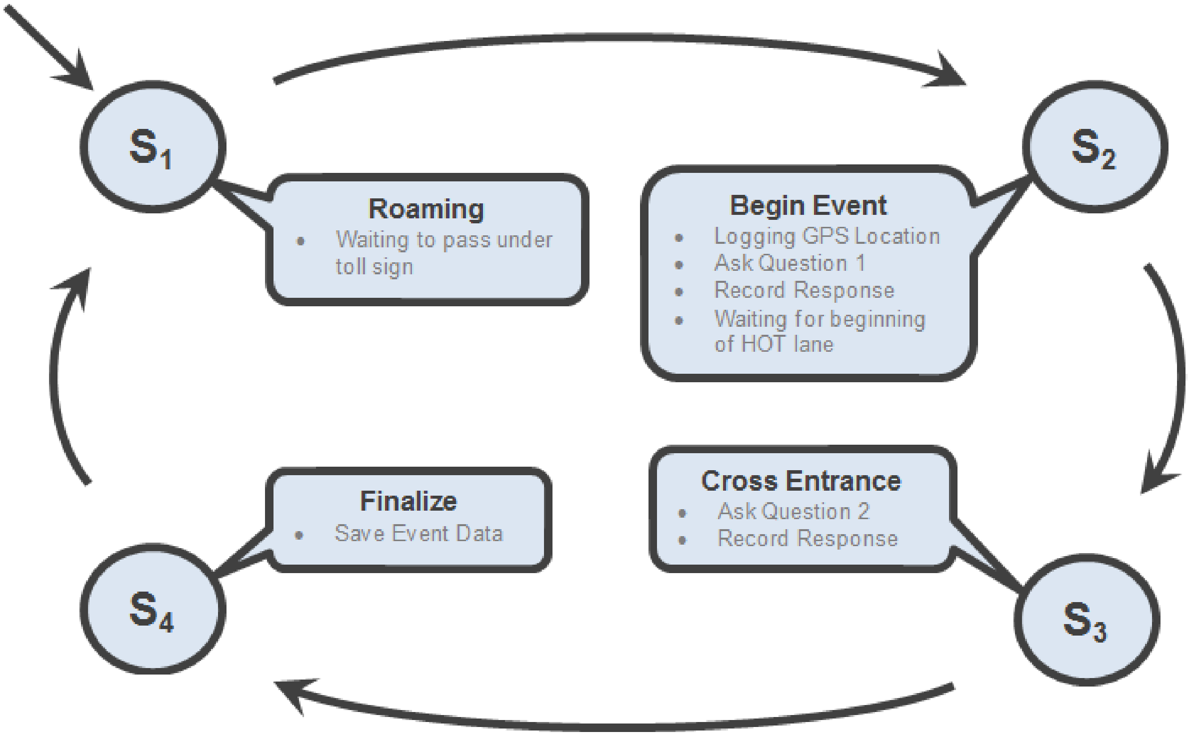

Figure 4.1 Application State Diagram

4.2 System Design and Security

Figure 4.2 contains an illustration of the data collection system in its totality. The system consists of a central web server that receives data from the in-vehicle collection device; retrieves traffic volumes and speeds by lane along I-394 from the MNDOT traffic management center and toll rates by HOT segment from the MNDOT; and stores the text files to a database server and the audio files to a cloud server.

The web server resides in a secured office at the University of Iowa's Public Policy Center with access limited to the project principal investigator and the systems administrator. The server is compliant with the university security protocols and located behind all security firewalls of the University of Iowa. The server is continuously monitored for security software updates.

The data collection device, shown as an iPad, has the custom developed software application described above. The data collection device has a store of events generated for each HOT segment the driver encounters. The device is currently configured for a manual upload through a secure web server.

Ruby and Sinatra are the scripting and web server software driving the data collection.

Once the data upload is confirmed by the data collection device operator—uploading can only occur if the device is not in motion—it is autonomously stored in a custom database. The database stores the event data with only an identification code that contains no personal identification information. The link between the data identification code and the demographic data is known only by the principal investigator and stored in a separate database on a different secured server. The names and addresses of the participants will be never be used in analyzing data or reporting findings, therefore it is stripped from all demographic data.

The uploaded audio event data from the collection device is removed from the text-based information and then stored as mp3 files remotely on Amazon’s S3 cloud service. The service is secure with access limited to the study team. To further maintain the privacy of the users, only an event identification code is stored that can link the audio file to the database. In order to provide an additional layer of protection, the event ID is the only common element between the audio files and text files.

On a daily basis, the traffic volume and speed data for each lane along I-394 is retrieved and passed into storage through the web server for later linkage with the driver event data. The volume and speed have a fine time resolution of thirty seconds, therefore, the objective traffic conditions the drivers faced while making their choice whether to use the HOT lane are known. The toll rates that were in effect when the driver selected to use the HOT lane are retrieved on a daily bases from the MNDOT MNPass site and passed into the database server.

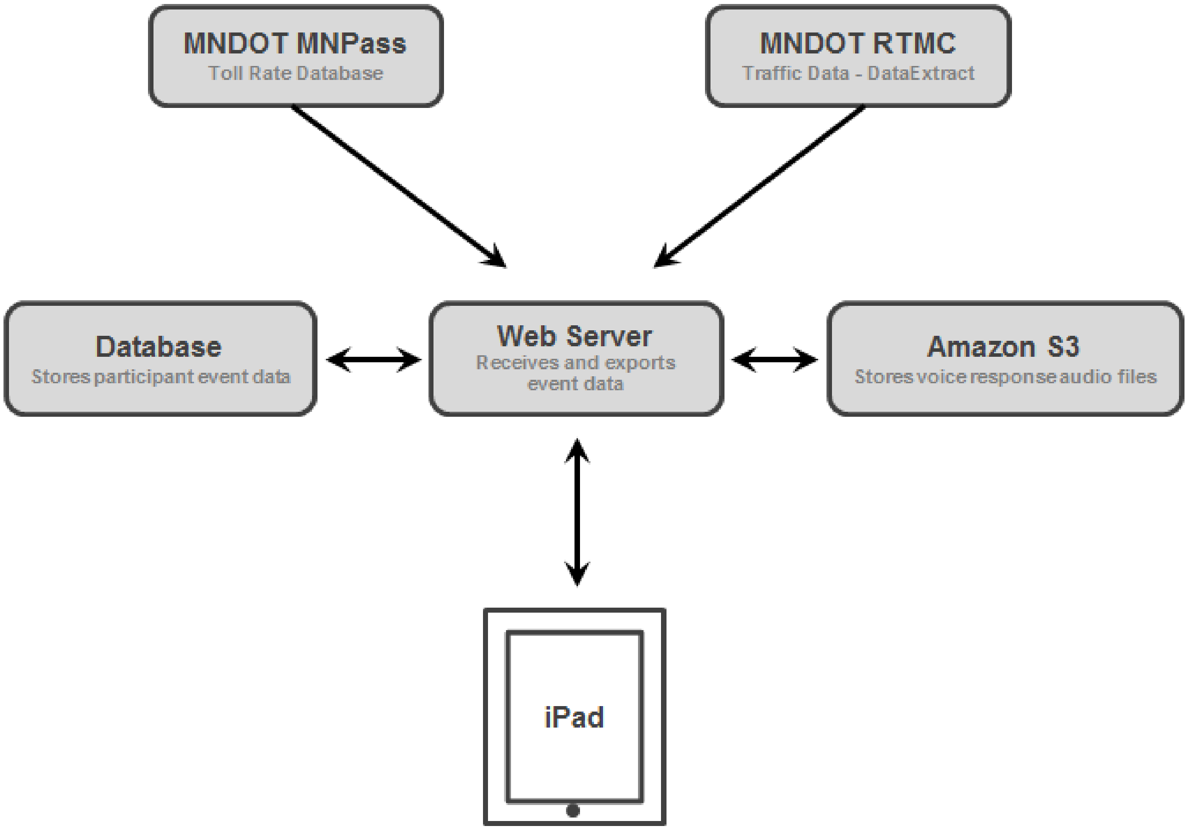

Figure 4.2 System Diagram for Data Collection System

4.3 Safety of Audio Prompting and Speech Interface in In-Vehicle Devices

Voice control devices are increasingly prevalent in vehicles. Evidence shows that the brain does not perform two tasks at the same time, rather it switches from one task to the other, and when overloaded, reaction time usually slows due to this switching and the brain selects to pay attention to incomplete information. Thus, driver distraction caused by interacting with technology while driving is a potential hazard.

The actual hazard level caused by driver distraction depends on characteristics of the secondary task and how the driver is interacting with the technology. It is the interaction with technology that mainly influences distraction, therefore reducing the complexity of interaction, such as using audio to replace visual and manual interaction, reduces the risk of distraction. As shown in the studies discussed below, voice interaction is consistently shown to be superior in terms of reduced driver distraction, increased driving performance and reduced crash risk when compared with manual control and graphical displays.

Baron and Green 2006 found that voice interaction allows drivers to visually concentrate on the driving task. More recently, Perez et al. 2011 evaluated voice control versus manual input navigation systems. They found voice control systems required less mental effort allowing for longer and more frequent glances to the forward roadway or driving-related locations. From their experimental results, the authors concluded that voice control outperformed visual-manual entry in terms of eye-glance requirements during destination entry; moreover, in many cases voice control systems have sizable advantages over visual-manual control devices (Perez et al. 2011).

As a result of the increased concentration on the driving task, Jensen and Skov 2010 found that audio prompting and response caused fewer longitudinal control errors (i.e. speeding violations) and fewer lateral control errors (i.e. lane excursions). Audio prompting and response contribute to improved driver performance, to decreased frequency and duration of eyes-off-the-road glances that is required for the visual and audio-visual in-vehicle technologies. Similarly, when Maciej and Vollrath 2009 evaluated in-vehicle technology interfaces they discovered speech interaction led to significant improvements in driving performance (measured via lanekeeping and reaction times) and in visual attention (measured via gaze time toward the driving task), as well as a significant reduction in the subjective distraction (rating scales) in their participants' lane change experiment. Given the great potentials for increased driver safety, the authors conclude that "speech control is a must in the car of the future" (Maciej and Vollrath 2009).

The data collection device developed in this study follows the recommendations of previous research by using audio and speech interfaces. The application uses audio prompts and does not require the driver to cast glances at the device to check their location or read what questions will be asked. They only receive two spoken prompts; "How much time do you believe you are saving?" and "Are you in the HOT lane?" The device then records their verbal response without intrusively repeating the question; instead, the audio recording terminates at a predetermined time with or without a driver's response or interaction. Based on current findings and experimental results, the proposed data collection device does not represent an unacceptable source of distraction because it does not require the driver's visual attention and requires less mental effort than existing visual-manual controlled in-vehicle navigation systems currently in use.

Chapter 5 Simulation of the I394 High Occupancy Tolled Facility

The testing phase of the in-vehicle perception acquisition device for the I394 HOT lane facility was simulated along U.S. 218 in Johnson County, IA. The critical roadside infrastructure was identified using schematic maps from the Minnesota Department of Transportation web site from Wayzata Blvd. to I-394, which are shown in figure 5.1. Locations of HOT access signs, dynamic toll signs, transponder readers, and the beginning and ending of each HOT tolled section were obtained. Using Google maps and related image data, the specific locations of the roadside infrastructure were obtained as latitude and longitude coordinates; location data for both the eastward and westward directions of I-394 were also obtained.

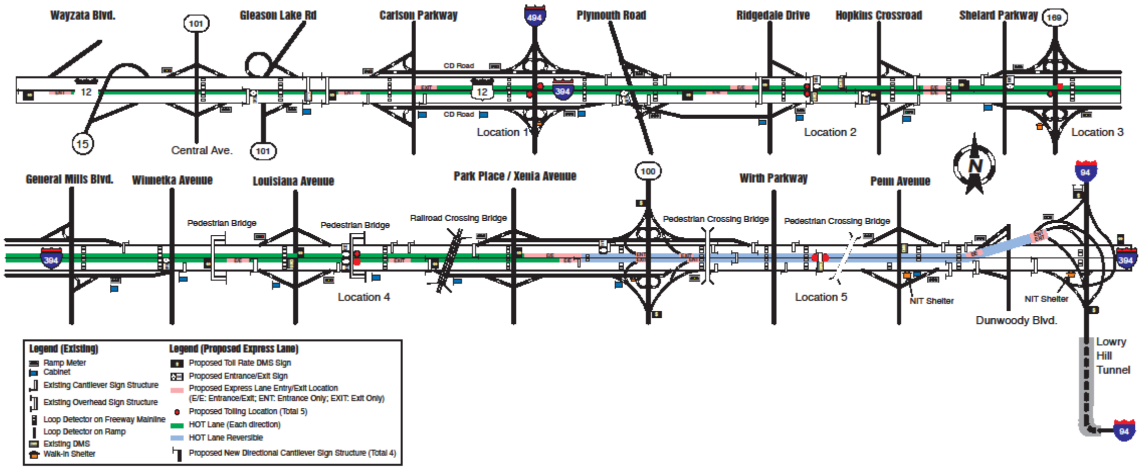

Figure 5.1 I-394 HOT Facility Schematic (http://www.mnpass.org/pdfs/394mnpass-schematic.pdf)

The distance between specific roadside infrastructure locations was calculated based on latitude and longitude coordinates using the Harversine formula. Starting at mile marker 89 and proceeding south—corresponding to the eastbound direction of I-394—the locations of each HOT access sign, dynamic toll sign, transponder reader, and the beginning and ending of the HOT lanes were mapped to the layout. For each piece of roadside infrastructure, the corresponding latitude/longitude coordinate was obtained along U.S. 218 and stored in the data collection device. The same procedure was applied to the northbound lanes of U.S. 218 to correspond to the westbound direction of I-394. See figure 5.2 for a graphic illustration along

U.S. 218. A bounding circle—a geofence—was created around the virtual HOT facility to trigger the application to activate the GPS hardware in the device.

The location of the dynamic toll sign relative to the HOT lane entrance is important to the study because this, along with travel speed, determines the time a driver has to make a decision whether to use or not use the facility. Additionally, the driver is given the key information—the price signal—that in theory will forecast their realized choice. Therefore, the data collection device utilizes geofences to trigger the first audio prompt to occur at or very near the toll sign location. The first question asks the driver for their subjective estimation of time savings if the HOT lane is taken. A second geofence is assigned to the entrance of the HOT lane, so when the vehicle crosses the imaginary line the data collection device prompts the driver with the audio question “Did you choose the HOT lane?” As explained above, the hardware limits the horizontal position accuracy such that it cannot reliably determine in which lane the vehicle is located. The programing of the geofences continues for each of the five tolled road segments for the simulated I-394 facility. The mapping between the eastbound I-394 HOT facility to the virtual U.S. 218 HOT facility is shown in the figure with only the first two HOT lane segments labeled.

Legend

- Circle ● : Access Sign for HOT lane
- Star ★Toll sign for HOT lane
- Diamond ♦ : Entrance to HOT lane
- Square ■ : End of HOT lane
- Triangle ▼: Electronic toll charger
- Blue circle ○The 1st question ‘how much time do you think you will save’ is triggered
- Orange circle ○The 2nd question ‘are you on HOT lane’ is triggered

Figure 5.2 Virtual HOT Facility on U.S. 218 with Greater Geofence - Southbound

Chapter 6 Performance of the In-vehicle Perception Acquisition Device

The in-vehicle Perception Acquisition Device was tested on the virtual HOT facility along U.S. 218 in Johnson County, IA. Every three seconds, the device logged the position (as latitude/longitude coordinates), speed and bearing of the vehicle along the virtual HOT facility. A critical performance measure is the distance traveled into the geofenced area before the device prompted the driver with the first audio question (see the figure for the locations where the questions were triggered). On average, the first question was asked approximately two seconds after crossing into the appropriate geofence. The mean distance past the perimeter was 103 feet (average vehicle speed was 67 mph). The figure illustrates one such geofence represented by the reddish shaded oval that has two vehicle tracks (red and blue dots). The virtual toll sign is shown as a black dot on the perimeter of the geofence.

Table 6.1 Descriptive Statistics for Distance Past Geofence - Question 1

	Feet
Minimum	86
Maximum	138
Mean	102.6
Median	101
St. Dev	17.4
N	20

Figure 6.1 Illustration of Geofence Prompting Question 1

The timing for triggering the second question was also established with geofences. A geofence was created for each of the simulated HOT entrances and once the device crossed it the second question was asked. When the question was asked the system began recording the driver's response. On average, the distance traveled before the audio recording terminated was 2,028 feet. At the average speed of 67 mph, this distance is covered in about 42 seconds. Although this might be an extremely long time for the recording of the second question, it provides the flexibility to reduce the open microphone time with additional testing. The figure illustrates the length of vehicle tracts for which the recording of the driver's response to the second question was active. The geofence that triggered the second question is shown in the reddish oval and the vehicle tracts are shown with red and blue dots (only two tracts are shown for clarity).

Table 6.2 Descriptive Statistics for Distance Past Geofence - Question 2

	Feet
Minimum	1951
Maximum	2094
Mean	2028.6
Median	2028
St. Dev.	45.0

Figure 6.2 Distance Traveled while Recording Question 2

Horizontal position of the device was not thoroughly tested in this prototype. However, on several trials the driver either maintained position in the left or right hand lane. As shown in the figure, the tract represented by the blue dots corresponded to the driver staying in the right-hand lane and red corresponds to the left-hand lane. The opposite designation is shown in the figure. Visual inspection reveals that the correct lane of travel was detected, suggesting that the device is sensitive enough to determine whether the vehicle is in the HOT lane or not, therefore eliminating the need to ask the second question. However, further testing of the horizontal accuracy and reliability is required in the urban environment and lane widths on I-394.

Chapter 7 Conclusions

This investigation extends the state of knowledge of decision modeling under risk and ambiguity by developing a mobile data collection platform for the capturing of naturalistic choice outcomes, associated environmental states, decision makers' self-articulated perceptions of risk and assessments of ambiguity, and socio-economic attributes. This software/hardware development is the first step in advancing our ability to forecast future revenue sources in transportation. The data collection platform will be used in an extending study and will allow the study team to obtain insights for developing, testing and implementing new behavioral models—that are superior to those obtained to date—that explicitly describe how drivers use imprecise information—in this case, ambiguous information signaled by variable toll lane charges.

The data collection device developed in this study follows the recommendations of previous research with audio and speech interfaces. The device uses audio prompts and does not require the driver to cast glances at the device to check their location or read what questions will be asked. They just receive two spoken prompts; "How much time do you believe you are saving?" and "Are you in the HOT lane?" The device then records their verbal response without intrusively repeating the question; instead, the audio recording terminates at a predetermined time with or without a driver's response or interaction. Based on current findings and experimental results, the proposed data collection device does not represent an unacceptable source of distraction because it does not require the driver's visual attention and requires less mental effort than existing visual-manual controlled in-vehicle navigation systems currently used.

The testing phase of the in-vehicle perception acquisition device the I-394 HOT lane facility was simulated along U.S. 218 in Johnson County, IA. The critical roadside infrastructure was identified using schematic maps from the Minnesota Department of Transportation web site from Wayzata Blvd. to I-394. The locations of HOT access signs, dynamic toll signs and transponder readers, and the beginning and ending of each HOT tolled section were mapped along U.S. 218 based on the relative distances on I-394.

The in-vehicle perception acquisition device was tested on the virtual HOT facility along U.S. 218. At a minimum of every three seconds, the device logged the position (latitude/longitude coordinates), speed and bearing of the vehicle along the virtual HOT facility. On average, the first question, “How much time do you believe you are saving?” was asked approximately two seconds after crossing into the appropriate geofence. The mean distance past the perimeter was 103 feet (average vehicle speed was 67 mph).

The timing for triggering the second question, “Are you in the HOT lane?” occurred on average two seconds after the simulated HOT entrances were crossed. On average, the distance traveled before the audio recording terminated was 2,028 feet. At the average speed of 67 mph, this distance is covered in about 42 seconds. Although this might be an extremely long time for the recording of the second question, it provides the flexibility to reduce the open microphone time with additional testing.

Horizontal position of the device was not thoroughly tested in this prototype. However, on several trials the driver either maintained position in the left or right hand lane. The tract for each run was plotted on a spatially registered orthophotograph and was visually inspected. The visual inspection revealed the correct lane of travel was detected, suggesting that the device is sensitive enough to determine whether the vehicle is in the HOT lane or not, therefore eliminating the need to ask the second question. However, further testing of the horizontal accuracy and reliability is required in the urban environment and lane widths on I-394.

7.1 Future Extension

This investigation extends the state of knowledge of decision modeling under risk and ambiguity by developing a mobile data collection platform for the capturing of naturalistic choice outcomes, associated environmental states, decision makers' self-articulated perceptions of risk and assessments of ambiguity, and socio-economic attributes. This software/hardware development is the first step in advancing our ability to forecast future revenue sources in transportation. The data collection platform will be used in an extending study and will allow the study team to obtain insights for developing, testing and implementing new behavioral models—that are superior to those obtained to date—that explicitly describe how drivers use imprecise information—in this case, ambiguous information signaled by variable toll lane charges.

The next step in this research program is to field test the in-vehicle perception acquisition device alone the I-394 HOT facility as well as the I-35W facility in Minneapolis, MN. The field testing on the actual HOT facilities will allow further development of the collection system—both for the client and server. Upon completion of the modifications, a pool of drivers along the I-394 and I-35W corridor will be recruited and enrolled in a study to collect real-time information on the choice of HOT lanes. The data will lead to the testing of the two leading behavioral decision theories; cumulative prospect theory and expectancy-valance theory. In addition to distinguishing between the predictive powers of the existing theories, the development of refined learning-based decision models will be developed and tested.

References

Avineri, E., and J.N. Prashker. 2004. “Violations of expected utility theory in route-choice stated preferences: the certainty effect and inflating of small probabilities.” *Transportation Research Record: Journal of the Transportation Research Board* 1894: 222-229.

Baron, Adriana, and Paul Green. 2006. “Safety and usability of speech interfaces for in-vehicle tasks while driving: A brief literature review.” *Technical Report UMTRI-2006-5*, University of Michigan Transportation Research Institute, Ann Arbor.

Bates, J., J. Polak, P. Jones, and A. Cook. 2001. “The valuation of reliability for personal travel.” *Transportation Research Part E: Logistics and Transportation Review* 37, no. 2: 191–229.

Bliemer, M.C.J., and J.M. Rose. 2006. “Designing stated choice experiments: state-of-the-art.” Paper presented at the 11th International Conference on Travel Behavior Research, Kyoto, Japan.

Bogers, E.A.I., M. Bierlaire, and S.P. Hoogendoorn. 2007. “Modeling learning in route choice.” *Transportation Research Record: Journal of the Transportation Research Board* 2014: 1–8.

Busemeyer, Jerome R., and J.C. Stout. 2002. “A Contribution of Cognitive Decision Models to Clinical Assessment: Decomposing Performance on the Bechara Gaming Task.” *Psychological Assessment* 14, no. 3: 253-262.

Einhorn, Hillel J., and Robin M. Hogarth. 1981. “Behavioral Decision Theory: Processes of Judgment and Choice.” *Journal of Accounting Research* 19, no. 1: 1-31.

Fujii, S., and R. Kitamura. 2000. “Anticipated travel time, information acquisition, and actual experience: Hanshin expressway route closure, Osaka-sakai, Japan.” *Transportation Research Record: Journal of the Transportation Research Board* 1725: 79–85.

Hensher, D.A., J.M. Rose, and W.H. Greene. 2005. *Applied Choice Analysis: A Primer*. Cambridge University Press, UK.

Hensher, D.A., W.H. Greene, and Z. Li. 2011. “Embedding risk attitude and decision weights in non-linear logit to accommodate time variability in the value of expected travel time savings.” *Transportation Research Part B: Methodological* 45 no. 7: 954-972.

HDR Decision Economics. 2008. “Moving at the Speed of Congestion: The True Costs of Traffic in the Chicago Metropolitan Area.” Metropolitan Planning Council, Chicago.

Jensen, Brit Susan and Michael B. Skov. 2010. “Studying driver attention and behaviour for three configurations of GPS navigation in real traffic driving.” *CHI 2010 Proceedings of* *the 28th International Conference on Human Factors in Computing Systems*, New York, USA: ACM Press.

Kahneman, D., and A. Tversky. 1979. “Prospect theory: an analysis of decisions under risk.” *Econometrica* 47, no. 2: 263–291.

Kudryavtsev, A., and J. Pavlodsky. 2012. “Description-based and experience-based decisions: individual analysis.” *Judgment and Decision Making* 7, no. 3: 316-331.

Louviere, J. J. 1994. *Conjoint Analysis*. Blackwell Business, Cambridge.

Maciej, Jannette, and Mark Vollrath. 2009. “Comparison of manual vs. speech-based interaction with in-vehicle information systems.” *Accident Analysis & Prevention* 41, no. 5: 924-930.

Noland, R., and K. Small. 1995. “Travel-time uncertainty, departure time choice, and the cost of morning commutes.” *Transportation Research Record: Journal of the Transportation Research Board* 1493: 150–158.

Perez, Miguel A., Derek Viita, Jonathan M. Hankey, Sherri Voran-Nowak, and Steven C. Tengler. 2011. “Voice- and visual-manual-control navigation system evaluation based on user performance in destination entry and navigation tasks.” Transportation Research Board Annual Meeting, Washington D.C.

Polak, J., S. Hess, and X. Liu. 2008. “Characterizing heterogeneity in attitudes to risk in expected utility models of mode and departure time choice.” Paper presented at the annual meeting of the Transportation Research Board, Washington, D.C.

Small, K.A., R.B. Noland, X. Chu, and D. Lewis. 1999. “Valuation of travel-time savings and predictability in congested conditions for highway user-cost estimation.” *NCHRP Report 431*, Transportation Research Board, Washington, D.C.

Tversky, Amos, and Daniel Kahneman. 1992. “Advances in Prospect Theory: Cummulatve Representation of Uncertainty.” *Journal of Risk and Uncertainty* no. 5: 297-323.

Van Lint, J.W.C., and H.J. van Zuylen. 2005. “Monitoring and Predicting Freeway Travel Time Reliability: Using Width and Skew of Day-to-Day Travel Time Distribution.” *Transportation Research Record: Journal of the Transportation Research Board* no. 1917: 54-62.

Von Neumann, John, and Oskar Morgenstern. 1944. *Theory of Games and Economic Behavior.* New Jersey: Princeton University Press.