An Analysis of Privacy in Intelligent Transportation Systems (ITS) and Location Based Services (LBS)

BY

CAITLIN D. COTTRILL
B.A., Vanderbilt University, 1999
M.S., University of Tennessee, Knoxville, 2003

THESIS

Submitted as partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Urban Planning and Public Affairs
in the Graduate College of the
University of Illinois at Chicago, 2011

Chicago, Illinois

Defense Committee:

Piyushimita (Vonu) Thakuriah, Chair and Advisor
Kazuya Kawamura, Department of Urban Planning and Policy
Robert Sloan, Department of Computer Science
Siim Sööt, Emeritus Associate Professor, Urban Transportation Center
Richard Warner, Chicago-Kent College of Law

This thesis is dedicated to my family, whose love and support have provided the foundation upon which I have built my life. Most especially, it is dedicated to my father, Donald Cottrill, without whom none of this would have been possible.

ACKNOWLEDGEMENTS

I would like to thank my committee members, Piyushimita (Vonu) Thakuriah, Robert Sloan, Kazuya Kawamura, Siim Sööt and Richard Warner, for their support and assistance through the process of developing and completing the dissertation. Their input and guidance have been invaluable. Vonu, in particular, has been an amazing resource in furthering the goals of both the dissertation and my career, with the provision of input, care, support, a listening ear, a wise mind, and a kind heart in all endeavors. The success of this dissertation process is as much a credit to her guidance and wisdom as it is to my efforts.

I would also like to acknowledge the faculty and staff of the Urban Transportation Center, the College of Urban Planning and Public Affairs, and the Computational Transportation Science program for their support and assistance over the past five years. The process has been greatly enriched by their efforts.

TABLE OF CONTENTS

Table of Contents (Continued)

Table of Contents (Continued)

Table of Contents (Continued)

LIST OF TABLES

List of Tables (Continued)

List of Tables (Continued)

List of Figures

SUMMARY

This dissertation responds to the following three overarching questions of interest:

1. To what extent is privacy currently being protected in the locational environment?
2. What types of relationships exist between people's desire for location and mobility benefits and their willingness to trade private data to receive these benefits?
3. What methods may be used to balance the public's desires and expectations related to privacy in the locational environment and their desire to receive transportation benefits?

These questions were examined in three ways: (1) a qualitative assessment of current United States-based legal and technological solutions to protection of locational privacy, by combining a review of existing literature and privacy practices; (2) a content analysis of online or mobile device-based privacy policies and other aspects of the terms of agreement that transportation agencies and private companies require consumers to enter into for use of specific technologies; and (3) statistical analysis and econometric modeling of data from a primary survey of privacy preferences and expectations that was undertaken for the purposes of this dissertation.

For the content analysis of privacy policies, 101 policies from public and private transportation service providers were evaluated and compared for both readability and content (in particular, whether they are reflective of the components of privacy identified in the Federal Trade Commission's (FTC) Fair Information and Privacy Policies, including notice/awareness, choice/consent, access/participation, integrity/security, and enforcement/redress). The consumer survey gathered information related to consumer demographics, current use of technology, expectations relevant to treatment of data gathered in the mobile environment, and incentives and benefits that might alter those preferences. A general analysis of

demographics and other data was undertaken, followed by analysis of willingness to trade private data in return for transportation benefits related to safety, efficiency, and economics.

The content analysis of privacy policies revealed that the practices of both public and private providers of ITS services and LBS are inadequate in addressing privacy components identified by the FTC. In particular, consumers are given inadequate information related to the use and sharing of their private data, and they are not provided with sufficient information related to how they may view, correct and contest inaccurate information. Such a finding indicates that the public's expectations of privacy in the mobile environment may not be adequately reflected in the practices currently followed by service providers. In addition, a readability analysis of the evaluated policies indicates that these policies are generally written at a level above that which may be easily understood by the traveling public. Here, again, concerns are raised regarding the adequacy of information currently being provided to consumers in the context of the treatment of their personal data.

The survey analysis revealed that, while consumers do not exhibit signs of being overtly concerned about privacy in the mobile environment (as shown in low levels of noticing and reading privacy policies), when queried regarding their perceptions of privacy risk regarding the sharing of personal and travel data, they report fairly high levels of concern. Such findings indicate that consumer awareness of privacy issues associated with mobile technology use is lacking in regard to the implications of sharing data for ITS and LBS purposes. Participants were also asked to indicate the degree of compensation they would require to share certain data (including name, address, trip origins and destinations, and route details). A number of factors related to willingness to trade private data to receive safety, efficiency, and economic benefits

were tested in conjunction with compensation desires, resulting in findings that indicate that a portion of the population will be unwilling to share certain data (particularly name and address information) under any circumstances, while safety and efficiency benefits were generally seen to generate the greatest degree of incentives to reduce compensation desires.

Overall, it was determined that current practices regarding the protection of private consumer data in the mobile environment are inadequate to fully address the privacy expectations of the traveling public. A lack of consistency and comprehensiveness in current policies reveal a gap between the expectations of the traveling public and the actual practices of service providers. In addition, survey findings reveal that consumer awareness of privacy issues related to transportation technologies is low, despite an overall concern for risks associated with the sharing of data in these contexts.

The findings here indicate that, as mobile technologies are currently growing in scope and application, it is critical at this point to establish effective means of incorporating privacy protection into developing systems, over and above the policy and technology-based privacy solutions that currently exist. Findings here indicate that this may best be accomplished via a combination of the following elements:

- Policy methods:
    - Provision of an umbrella privacy policy for mobile services, which would allow for a “reasonable expectation” of privacy to be developed for mobile applications and services;
    - Enhanced clarity and transparency of privacy policies for individual services and applications.
- Technical methods:
    - Privacy by design
    - Enhanced use of anonymization techniques and other methods which may minimize risk of data disclosure and mining
- Consumer awareness:
    - Increase consumer-driven awareness campaigns

- Involve educational campaigns to increase consumer knowledge of potential risks and benefits of sharing data in the mobile environment

Application of these methods will both encourage adoption of and ensure that the benefits of ITS and LBS services are maximized for current and future applications.

CHAPTER 1: INTRODUCTION AND OBJECTIVES

1.1 Introduction

The 21st Century has been defined in large part by the rapid and extensive growth of information technology touching nearly every part of our daily lives. From growth in computing speed and efficiency, to increasingly powerful communications technology, to evolution of systems and networks increasing efficiency and timeliness, our reliance on technology has changed the face of societal interactions and expectations. This change is, perhaps, at no place more evident than in the realm of transportation.

While the 20th Century has been called the Age of the Automobile, the 21st century may be called the Age of Interactive Mobility. Planned and implemented systems allowing for the interconnectivity of transportation data from a variety of sources for a variety of benefits have become commonplace in our transportation network. Ubiquitous mobile peer-to-peer and vehicle-to-network mobile networking systems are currently in the planning stages, and such systems and technologies as Intelligent Transportation Systems (ITS) and Location-Based Services (LBS) have been growing rapidly in terms of scope, application, and sophistication. Since the establishment of the national Intelligent Vehicle Highway Systems (IVHS) program under the 1991 Intermodal Surface Transportation Efficiency Act (ISTEA), the funding allocated to national, regional, and local ITS programs has grown and changed significantly. Concurrent investments in and the rapid evolution of associated technologies (such as Global Positioning Systems and radio-frequency identification (RFID) tags) have allowed for widespread adoption of ubiquitous transportation technologies and systems such as ITS and LBS as cost-efficient

methods of increasing safety, efficiency, and expediency. Unlike traditional transportation investments, which focus on the physical space of roadways, sidewalks, and signal systems, ITS and LBS projects concentrate on the flow of information between travelers, vehicles, and the transportation network to encourage efficiency and safety.

One critical part of this shift in transportation investment is the need for expanded surveillance and dataveillance[1] to support these ITS and LBS technologies and applications. Palen (1997) noted that:

> An Intelligent Transportation System (ITS), by definition, involves the use of intelligence to enhance the operation of the transportation system. Intelligence, by definition, requires information. Information, by definition, is data formulated in a formation. Data is generated by surveillance. Therefore, surveillance forms the basis for the formation of information for an ITS. You can't have a usable ITS without surveillance.

Surveillance also forms a necessary component of LBS systems. Jiang and Yao (2006) state that, "From a societal aspect, LBS are a key instrument for the improvement of the quality of life and personal productivity. On the other hand, societal impacts of LBS also include surveillance and invasion of personal privacy, and changes in human spatial behavior." The increased use of surveillance technologies (such as sensors, Global Positioning Systems (GPS), Automated Number Plate Reading (ANPR), and Radio Frequency Identification (RFID)) in the transportation environment has resulted in large quantities of data being collected on travelers within the network. Unfortunately, while the benefits of this data to users have been well documented, the impacts of the concurrent loss of privacy on the traveling public have not been granted as much attention. Locational privacy, or the ability of a person to travel in public spaces with the

[1] *Dataveillance* has been described by Roger Clarke as, "…the systematic use of personal data systems in the investigation or monitoring of the actions or communications of one or more persons." (Clarke, 1988)

expectation that, in general, his location will not be tracked and recorded, is, however, poised to emerge as an issue of interest to consumers and service providers.

As the literature review will show, a number of studies have been undertaken regarding general issues of privacy in policy and legal frameworks, and the issue of privacy in technology-based mobile applications has been garnering increased attention. This dissertation will attempt to meld questions and issues from both realms in order to provide a better understanding of the links between the two, and how users may be impacted or influenced. Such an approach, as outlined below, will require reviewing literature from policy, theory and technology, as well as examining more closely approaches and issues related to privacy preservation from the viewpoint of industry experts, existing laws and regulations, and user preferences. The following section will address the research questions associated with both technological and policy-based privacy concerns of users in the mobile environment and ways that privacy may be protected.

1.2 Research Questions and Objectives

The research undertaken for this dissertation is designed around the following three primary questions:

1. To what extent and by what policy methods is privacy in ITS and LBS currently protected in the United States?
    a. What are the applicable rules, regulations, policies and guidelines that impact privacy in the mobile environment?
    b. What privacy-preserving technologies have been used in ITS and LBS? What would be a policy-based framework to effectively preserve privacy in given technological systems and networks?

   c. Are there consistent laws and regulations guiding the collection and use of private mobility data by both public and private transportation service and application providers? If not, how do relevant laws differ, and what gaps may be identified?

   d. Does a relationship exist between existing privacy policies and regulations and the potential for violation of locational privacy?

2. Does a relationship exist between individuals' privacy preferences and their willingness to trade private information for use in ITS and LBS technologies and applications?

   e. What expectations do travelers currently have regarding privacy protections afforded to them by transportation agencies and service providers?

   f. Do travelers currently demonstrate privacy concerns in the mobile environment?

   g. What are the component parts of locational privacy preferences, such as personal information (including name, address, or vehicle information) and travel information (such as origins, destinations, and other travel details)?

   h. How much influence, if any, does the desire to protect private information have on the willingness of persons to trade this information for various transportation benefits, including cost, efficiency, or safety benefits? Do any of the component parts of locational privacy identified in the above step have greater or lesser values in relation to these benefits?

   i. What types of expected benefits may most impact travelers' willingness to trade privacy components? Can these characteristics allow us to cluster potential users and determine ways to balance privacy preferences and application efficiency?

3. In light of the first two questions, what methods may be used to balance the public's desires and expectations related to privacy in the locational environment and their desire to receive transportation benefits?

The objectives behind answering these questions are as follows:

- To provide information on existing policy and technological tools for privacy preservation in ITS and LBS in the United States and under conditions similar to those experienced in the United States;
- To model the potential impacts of varying degrees and methods of privacy protection in ITS and LBS adoption and efficiency at both system (ITS in general) and application (individual LBS technologies) levels; and
- To utilize the results of the above to recommend methods of effectively protecting (or allowing for the protection of) traveler privacy while maximizing the usefulness of associated ITS and LBS technologies.

The effectiveness of many emerging transportation technologies will depend in large part on the extent of their deployment and use. Many proposed ITS and LBS technologies, particularly those associated with safety, traffic flow, and social networking, depend upon the ability to collect and share large amounts of data from vehicle to vehicle, from the vehicle (or traveler) to a network of traffic sensors connected to a central server, or from user to user. To encourage widespread acceptance of these technologies will require ensuring that potential users are aware of and comfortable with associated consequences of the sharing and use of this data. One of the primary aims of this study is to enhance the ability of transportation professionals to ensure that travelers are comfortable with the level of privacy afforded them by the implemented technologies by gaining a greater understanding of the public's privacy expectations. By concentrating on the overlap between technology and policy from the vantage point of the traveler, the study will address the issue of privacy in a manner relevant to transportation planners, technicians, and policy-makers.

As noted in the introduction, and expanded upon in the literature review, there has been comparatively little work done on the intersection of personal privacy preferences and the willingness of travelers to trade the personal information necessary for acceptance of ITS and LBS technologies. Additionally, there is little direct research on the potential for bi-directional influence of privacy-preserving technologies and policies in the locational environment. This dissertation is intended to help address those gaps in order to provide a better template for development of ITS and LBS technologies that will encourage adoption and use by directly addressing the privacy preferences of the traveling public. It is believed that the findings from the research may be able to provide a valuable voice in the emerging discussion

of locational privacy and its impacts on the behavior of the traveling public and the safety and efficiency of the transportation network.

An analysis of secondary data and a two part data collection effort are used to answer the research questions posed above. The following list outlines the approach that will be taken.

- First, a *review of related literature* was undertaken to define the context of the research questions. This review will provide an overview of literature related to privacy theory and privacy within the context of mobile applications, as well as literature related to the data collection and analysis methodologies to be used as part of the research. Current legal approaches to privacy are identified, and analyzed to determine how they apply to the mobile environment.
- Second, a *content analysis* of representative privacy policies was conducted in order to determine how well concerns and issues identified in current research and voiced by courts and law makers are reflected in the privacy policies of ITS and LBS service providers. This content analysis builds on the findings of the literature review, particularly in terms of items of interest to be addressed in privacy policies to reflect identified components of locational privacy.
- Finally, a *general survey* was designed and administered to understand locational privacy preferences and ways in which users conceptualize privacy in the context of benefits that they receive, risks that they avoid and potential compensation for allowing their data to be used for other purposes. This survey will establish a framework for determining general privacy preferences within a small population that may be evaluated in relation to current privacy practices in the mobile environment, as well as identifying preferences related to concerns about the sharing of data with public and private providers. These methodologies will be further discussed in Chapter 4.

The overall dissertation will focus on the relationship between privacy policies, ITS and LBS, and use of technology. Chapter 1 will describe the overall research questions and objectives and hypothesized impacts of the research undertaken. Chapter 2 will present the theoretical underpinnings for the current research, particularly as reflected in socio-technical, contextual integrity, and adoption theories and theories of justice. Chapter 3 will provide an overview of the pertinent literature, including the definition of and valuations of privacy and existing privacy protections. Chapter 4 will address the methodological underpinning of the research, including

literature related to content analysis and the determination and valuation of privacy preferences. Chapter 5 will present the results from a content analysis of existing privacy policies. Chapter 6 will present the results obtained by analyzing the survey results. Chapter 7 will review overall conclusions and policy implications, while Chapter 8 will detail limitations of the study and proposals for future research.

The dissertation, on the whole, is intended to assist in providing guidance for future information technology use in transportation. While we are currently experiencing rapid growth and evolution of transportation technologies, the potential for future systems incorporating even more ubiquitous, data-hungry mobility systems is enormous. One key point that should be noted is the need to plan for these future systems instead of simply reacting to the current environment. By establishing preferences and guidance for matters related to the privacy of personal information balanced with the desire for transportation enhancements, it will be possible to create policies for future systems designed with attention to this balance. By framing data collection within the current environment, a baseline of existing expectations related to actual practices can be established, while discussion of how findings related to such expectations may be translated to future systems will provide guidance for future policy. It is hoped that this dissertation will not only address current concerns, but also provide a resource for future policy needs.

CHAPTER 2: THEORETICAL UNDERPINNINGS OF PRIVACY IN THE LOCATIONAL CONTEXT

2.1 Introduction

While it is evident that the issue of privacy in the locational environment has become a topic of concern in recent years, the rationale behind this concern is perhaps less evident. Inured as we are to the expectations for sharing personal data such as phone numbers or email addresses in order to sign up for services, make purchases, or record preferences, it is perhaps conceptually easy to assume that provision of personal data is fast becoming a requisite for obtaining efficient and effective transportation information. What may be ignored here, however, are systematic changes in the provision of transportation services that may make this situation seem commonplace. This chapter will first provide an overview of socio-technical systems theory as it relates to the availability and provision of transportation services in order to better define the landscape under which locational privacy decisions begin to be made. By setting the context in this manner, it may be argued that the influences that emerging transportation technologies have on the greater society will necessitate that attention be paid to privacy rights of individuals, consistent with the regulatory protections afforded to data in realms such as health care and credit reporting. Next, privacy as contextual integrity will be examined in order to better set the stage for examining privacy within the context of mobility. Finally, a brief overview of justice theory and its application to the subject at hand will be presented.

2.2 Socio-Technical Systems Theory

2.2.1 Background

Advances in transportation technology are often predicated on the ability of the consumer to interface with emergent technologies such as real-time traffic information, location-based social networking, travel navigation and trip-planning. The degree of adoption of these technologies, and the amount to which the American public has begun to rely on them for transportation information and planning, indicates that a sea change is emerging in our experience of the mobile environment relative to available and developing technologies. Such an environment requires an analysis of the socio-technical transformation we are currently experiencing in the realm of transportation.

In the management sciences, socio-technical systems may be conceptualized as described by Bostrom and Heinen (1977):

> … a work system is made up of two jointly independent, but correlative interacting systems - the social and the technical. The technical system is concerned with the processes, tasks, and technology needed to transform inputs to outputs. The social system is concerned with the attributes of people (e.g., attitudes, skills, values), the relationships among people, reward systems, and authority structures. It is assumed that the outputs of the work system are the result of joint interactions between these two systems.

Much of the technological advancement currently being seen in the area of transportation is reliant upon a structure of socio-technical interaction where data generated by travelers within the social system are fed into technological systems to improve their effectiveness and reliability and, in turn, those data are then transformed into systems that impact how, when, where, and why we travel. This bi-directional influence of transportation technology and the society that uses it will be the subject of this section, as we work to ground the discussion of

the need for privacy policies relative to the expectations of the society to which such regulations would apply.

2.2.2 Description of Socio-Technology Theory and Application to Transportation

Erickson (2009) of IBM's Watson Lab proposes that socio-technical design includes the design of things that participate in complex and potentially conflicting systems distributed across time and space. Such a conceptualization of socio-technology as having both spatial and temporal components makes it particularly relevant to the area of transportation technologies, as these are also best evaluated over space and time. In addition, the recognition of complexity and conflict reflects that transportation technologies do not represent a simple system with linear influence, but rather a multi-dimensional network with feedback loops, multiple inputs, and numerous actors with potentially conflicting demands.

Such a conceptualization reveals the importance that comes to bear on policy considerations relevant to transportation technology, as these will help to guide and define the emerging societal structures in which adoption and adaptation will develop. The protection of personal information via privacy policy initiatives will be an important element in this development, as the bi-directional nature of technology and society leaves open the potential for societal or individual data to be included in technological systems, which then, in turn, leaves this data open to re-emergence in society. Conflicting desires of agents in the system may provide scope for intentional or unintentional misuse of these data, which heightens the need to have in place appropriate measures for protection and/or restitution in these cases. Ottens, *et al.* (2006) support this concept in their evaluation of socio-technical implications of ITS, arguing that social elements such as laws and regulations are necessary components of the

transportation system itself, as they impact and influence system functionality and behavior of actors within the society.

Here, issues relevant to law and regulations begin to clearly emerge. In the context of this dissertation, the primary area for concern is that of privacy, and how emerging location technologies and their application within both the transportation system and the overall society will be reflected in and influenced by laws and regulations pertaining to the protection of the private data needed for them to be effective. Shneiderman (2008) argues that socio-technical systems will require additional oversight from government regulators and others, noting that, "Attaining universal usability will make clear the need to also pursue 'universal sociability', that is, technology that supports social principles common to all communities, like civil liberties, privacy, or fairness…" By construing privacy as a common social principle, Shneiderman highlights the need for this facet to be addressed in both the design and adoption of systems, a need that may be best addressed via regulation and standards consistency. Linking this finding to the Fourth Amendment conception of "reasonable expectations," outlined in Chapter 3, strengthens the need to evaluate privacy in transportation technology and policy from the standpoint of a socio-technical system, as societal expectations may have large impacts on the development and adoption of ITS and LBS applications, particularly as people become more cognizant of the amount of data being provided to application developers and users, as well as the relatively lax standards currently guiding their storage, sharing and use.

Recent revelations regarding privacy and the sharing of data, particularly in reference to such social networking applications as Facebook, have begun to increase consumer awareness of privacy in a networked environment such as will be needed for effective deployment of ITS

and LBS applications. In the mobile environment, consumers are often unaware of the degree of data that may be collected, mined, and shared, and companies are often under little legal obligation to protect the privacy of data which have been shared. Here, lack of awareness may break down the socio-technical system as consumers are unaware of the role they are playing in the network, and thus cannot make effective decisions to guide the development of technology and technology policy. The Federal Trade Commission, in its 2010 report "Protecting Consumer Privacy in an Era of Rapid Change," suggested the utilization of a three-part model to address, in part, some of these concerns. The model includes the following components:

- Companies should adopt a "privacy by design" approach which will build privacy practices into everyday business practices, including the provision of reasonable security, minimal collection and retention of data, and promotion of accuracy.
- Companies should provide simpler, more streamlined consumer choices regarding data practices.
- Companies should take certain measures to make their data practices more transparent to consumers by providing clear privacy policies, allowing reasonable access to personal data, and requiring affirmative consent for data collection and use.

These measures, properly implemented, may begin to redress the current socio-technical imbalance, as they will allow for more consumer awareness to mediate technical considerations. Such recommendations will be further examined in the evaluation of privacy policies and the consumer survey.

2.2.3 Conclusions

Socio-technical theory provides a useful construct for evaluating the issue of privacy in the locational context, as it allows the researcher to investigate the bi-directional influences of society and technologies within the context of a widespread and multi-actor system such as are

seen in ITS and LBS. Issues of consumer awareness and the role that privacy policies and practices play in its development are relevant here, as effective socio-technical systems rest on some degree of knowledge regarding interactive action. The next chapter will look more thoroughly at the conceptualization of privacy within this type of system, including the provision of definitions of privacy, and overviews of the contributing components. The underlying legal justification for privacy regulation will be addressed, as will additional information on how such issues may be addressed via the examples of the Health Insurance Portability and Accountability Act of 1996 (HIPAA) and the Fair Credit Reporting Act (FCRA) of 1970.

2.3 Privacy as Contextual Integrity

In conjunction with socio-technical theories, several authors have posited that contextual integrity is a contributing factor to considerations of privacy in the public sphere. As Nissenbaum (2010) states,

> …a right to privacy is neither a right to secrecy nor a right to control but a right to *appropriate* flow of personal information…Privacy may still be posited as an important human right or value worth protecting through law and other means, but what this amounts to is a right to contextual integrity and what *this* amounts to varies from context to context.

The theory of contextual integrity, described in greater detail below, is particularly relevant to discussions of privacy in the transportation realm, as the contextual uses of ITS and LBS technologies vary with great rapidity, whether used on the public roadway, in a place of business, or in one's own home. In each case, the context of use may influence and impact the expectations that a user may have in relation to the data being shared, collected, and used. This

section will present an overview of the theory of contextual integrity and show how it relates to the current discussion of privacy in the locational environment.

2.3.1 Background and Overview

Barth, *et al.* (2006) state the theory of contextual integrity in the following way:

> Contextual integrity is a philosophical account of privacy in terms of the transfer of personal information. It is not proposed as a full definition of privacy, but as a normative model, or framework, for evaluating the flow of information between agents (individuals and other entities), with a particular emphasis on explaining why certain patterns of flow provoke public outcry in the name of privacy (and why some do not).

The ideas put forth in theories of contextual integrity draw upon Michael Walzer's theory of distributive justice, which speculates that societies are composed of a variety of distributive spheres defined by internal social goods (such as wealth, influence, education, and security). Distribution of these goods takes place according to internal norms or principles within each sphere, and is relatively autonomous.

Nissenbaum (2004) establishes the relationship between distributive justice and contextual integrity by calling upon complex equality, which posits that this, "is achieved when social goods are distributed according to different standards of distribution in different spheres and the spheres are relatively autonomous." In relation to privacy, Nissenbaum (2004) states that, "What matters is not only whether information is appropriate or inappropriate for a given context, but whether its distribution, or flow, respects contextual norms of information flow." Thus, the sphere within which data and information are shared will be subject to the norms and rules of that contextual sphere, and will vary dependent upon the current sphere of influence. For example, an individual may have differing expectations regarding privacy of location

information depending upon whether the current sphere of influence consists of work, family, friends, or an insurance company. In each case, generated data remain the same, but the norms associated with the flow and distribution of those data will be subject to expectations relating to the degree of comfort and the context in which they will or will not be shared.

2.3.2 Components of Contextual Integrity

Barth, *et al.* (2006) describe contextual integrity in terms of contexts, roles, and types of information of interest. They propose a temporal framework, stating that, "Temporal logic with past and future operators is used to say, for example, that certain information may be disclosed only if the subject mentioned has previously given permission or that if certain information is made public, notification must be sent to the concerned party. (Barth, *et al.* 2006)" Two types of norms are identified based on their temporal context, with positive norms permitting communication in the case of satisfaction of its temporal condition, and negative norms permitting communication *only* if the temporal condition is fulfilled. The act of communication here described consists of sending data from a subject to a recipient, and the data model set forth by the authors also acknowledges that these data may be combined with additional messages (or "mined") to provide additional knowledge about the subject.

This conceptualization is particularly relevant in a transportation setting, where, as noted, contexts may change rapidly due to changing environments. A traveler, for example, may be willing to let a sphere of "friends" know that his travel path includes a coffee shop, a drugstore, and a bar, but may be unwilling to let the sphere of "colleagues" have access to this path. By ensuring past, present, and future compliance with the contextual preferences of the individual, it may be possible to ensure consistency of expectations, if not complete comfort,

within locational privacy. Acknowledging the presence of mining also shifts expectations, as travelers will need to be aware that data they have shared in one context may be shared, with or without their knowledge, in a different context. One question relevant to privacy policies here is whether they adequately describe the past, present and future contexts within which shared data may be used. This dissertation will, in part, work to address whether these contexts are adequately presented to the traveler.

Barth, *et al.* (2006) also identify four key constructs used in defining contextual integrity, namely:

- Informational norms: communication of personal information between parties;
- Appropriateness: whether the data in question is in conformity with relevant informational norms;
- Roles: the capacity an agent (whether data provider, receiver, or subject) is serving in respect to a certain data transmission; and
- Principles of transmission: constraints that regulate the flow of information between entities subject to informational norms. These principles include:
    - Confidentiality: Prohibits agents that receive information from sharing it with other agents in the future;
    - Reciprocity: Principle that guides whether information flow is one-way (as from patient to physician) or bi-directional (as between friends); and
    - Dessert: Norms guiding whether information is "deserved" (such as physicians "deserving" to know certain information about their patients in order to make informed diagnoses).

If these principles are not followed according to ascribed norms, contextual integrity may be said to be violated. In relation to locational privacy, privacy policies of service providers are generally used to inform consumers of how these norms are treated, but (as will be seen in Chapter 5), this may be an inadequate method of ensuring consumer awareness.

2.3.3 Application of Contextual Integrity to Privacy in ITS and LBS

ITS and LBS applications, as described above, constitute a multi-directional socio-technological system, with roles of data generators, receivers, and users changing rapidly dependent upon one's use of the system. Such a system will impact the contextual integrity norms described above, particularly in relation to the expectations of persons in regard to the privacy they may feel within their roles as travelers. Zimmer (2008) has examined contextual integrity in the realm of privacy on the roadway using the case study of Vehicle Safety Communication (VSC) technologies, stating that, "…the design of VSC technologies might significantly alter the flow of personal information in the context of highway travel, contributing to the growing ubiquity of public surveillance, and threatening the value of privacy in public." Zimmer (2008) argues that privacy in public has been an area ignored in much privacy research due to two factors:

> *Conceptually*, the idea that privacy might somehow be violated in public space is often considered paradoxical. For many, the value of privacy applies to an individual's private sphere alone…The second explanation why privacy in public is often overlooked recognizes that the *empirical* status of privacy in public has failed to garner proper attention by privacy theorists. Simply put, prior to recent advances in information technology, the problem of privacy in public was not experienced in one's everyday life to the extent it is today.

In Zimmer's argument, recent advances in and uses of technology on the roadway have brought the notion of privacy in public to the forefront of concern, as it has changed norms and expectations of data flow by travelers due to increasing amounts of surveillance and the capability to record, store and use previously unthought-of amounts of data. The shifting of these norms via such technologies as VSC has enlarged the potential landscape for violations of contextual integrity.

Zimmer describes two primary areas of concern for norm violation. First, he notes that current expectations regarding the norms of appropriateness on the highway include the expectation that visual information, such as vehicle type and license plate information, is open to observation and sharing. With the advent of VSC technologies, however, such observations may become more detailed and more precise, and may be recorded for future use. Such technologies, he argues, disrupt the current norms of appropriateness by opening up the availability of data related to precise information about the vehicle (such as longitudinal location or braking ability) to a variety of agents acting in differing roles (such as law enforcement agencies or fellow travelers). The second area of concern relates to norms of distribution (discussed above as norms of transmission), with Zimmer (2008) positing that VSC technologies may have, "… the potential to disrupt the natural barriers that previously limited the ability to track individual vehicles over space and time." Such a disruption in the expected norms of appropriateness and transmission would have the potential to violate contextual integrity, and leave the consumer open to unexpected consequences related to the potential for collecting, storing, and sharing of personal data in ways that may not be in concordance with anticipated societal behavior. Such concerns will, in turn, may impact potential adoption and use of certain ITS and LBS technologies, which will be discussed in the next section.

2.4 Adoption Theory

2.4.1 Introduction

The rapid evolution of ITS and LBS would not be possible without adoption by consumers and agencies. While technologies have advanced to allow for real-time information to be collected,

shared, and used, and applications (or "apps") can be used for social networking and way finding, without widespread adoption, the systems do not necessarily have the inputs to function effectively. This section will first outline the benefits of ITS and LBS technologies in order to present a rationale for adoption, followed by background information on adoption theory.

2.4.2 Benefits of ITS and LBS

Bertini, *et al.* (2005) documented the following benefits of deployment of various types of ITS in urban areas:

- Arterial management systems can potentially reduce delays between 5% and 40% with the implementation of advanced control systems and traveler information dissemination.
- Freeway management systems can reduce the occurrence of crashes by up to 40%, increase capacity, and decrease overall travel times by up to 60%.
- Freight management systems reduce costs to motor carriers by 35% with the implementation of the commercial vehicle information systems and networks.
- Transit management systems may reduce travel times by up to 50% and increased reliability by 35% with automatic vehicle location and transit signal priority implementation.
- Incident management systems potentially reduce incident duration by 40% and offer numerous other benefits, such as increased public support for DOT activities and goodwill.

Of note here is that the types of benefits obtained via implementation of ITS systems vary widely in terms of both types of measurable benefits and beneficiary type. The benefits identified here relate to the following six goals for transportation operations identified by the U.S. DOT: safety, mobility, efficiency, productivity, energy and environmental impacts, and customer satisfaction. In order to track both ITS deployment and accrued benefits, the

Intelligent Transportation Systems Joint Program Office (ITS JPO) of the U.S. DOT has developed a Benefits database that addresses the topic areas shown in Table I below.

Table I: Major Topics in ITS

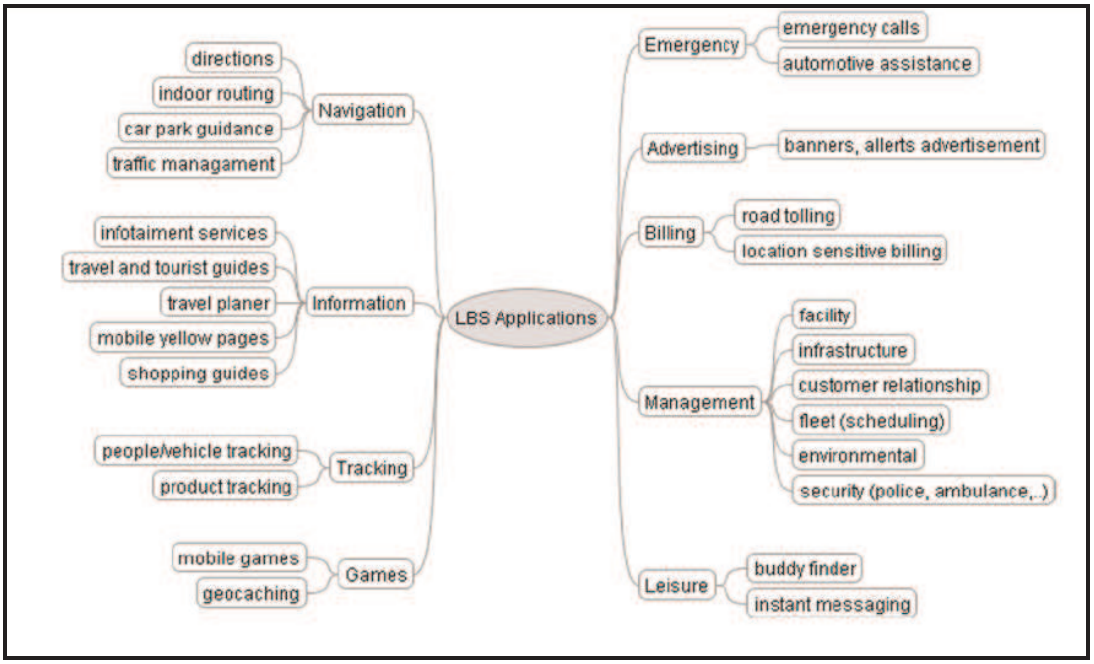
Intelligent Infrastructure	Arterial Management
	Freeway Management
	Crash Prevention & Safety
	Road Weather Management
	Roadway Operations & Maintenance
	Transit Management
	Transporation Management Centers
	Traffic Incident Management
	Emergency Management
	Electronic Payment & Pricing
	Traveler Information
	Information Management
	Commercial Vehicle Operations
	Intermodal Freight
Intelligent Vehicles	Collision Avoidance
	Collision Notification
	Driver Assistance

While the identified ITS systems are equally as ubiquitous as LBS, they may be less visible to the average consumer and are often adopted at the agency or organizational level rather than by individual consumers. Such project types as arterial and freeway management and roadway operations and maintenance may have clear benefits to the consumer in terms of travel efficiency and safety, but are likely not recognized by the consumer. For example, a traffic signal interconnect project, which maintains green lights along a segment of roadway in order to make travel more efficient, will provide benefits to the consumer, but will likely not be consciously acknowledged or recognized by the traveler. Thus, adoption decisions are made at the agency level based on estimated benefits, political and organizational structure, and anticipated consumer reaction. Conversely, such ITS systems as electronic payment and pricing

and intelligent vehicle systems do require conscious adoption by the consumer to function effectively, which increases the importance and impact of consumer reaction and willingness to adopt. For these projects, propensity to adopt should be viewed from both organizational and individual viewpoints, thus benefits should be clearly acknowledged from conception.

Unlike ITS projects, which are generally adopted at the systemic level by agencies and then passed down to the consumer, LBS adoption generally takes place at the level of the consumer; thus, their benefits must be clearly apparent to individual users. LBS applications vary widely, as shown in Figure 1 below, developed by Steiniger, *et al.* (2006).

Figure 1: LBS Adoption Categories



The range of services represented here, from social networking to wayfinding to emergency services, begins to demonstrate why LBS have grown rapidly in adoption in recent years (for

example, Stevens and Goasduff (2008) estimate that, “Worldwide subscribers to location-based communications services on mobile devices will increase by nearly 168 per cent in 2008 while revenue will grow by 169 per cent…”).

The emergence and seeming ubiquity of these location based services has revealed a plethora of data collection possibilities, particularly in combination with the wide spread of GPS enabled cellular phones. According to the CIA World Factbook (2011), 270 million U.S. citizens (89%) were mobile cellular telephone subscribers in 2008. In addition, the number of persons using mobile cellular devices for non-phone uses is also growing, particularly in conjunction with GPS enabling of these devices. For example, a 2009 Neilsen report found that 16% of teenagers with these devices use location-based services on their phone. According to Junglas and Watson (2008) from 2006 to 2010 the U.S. market for location based services is expected to grow from $150 million to $3.1 billion. In this environment, location-based applications (or “apps”) are also quickly gaining in popularity, with such services as Foursquare (with over half a million users in its first year (Parr, 2010)) and Google Latitude (which reports over 3 million active users (Siegler, 2010)) emerging as potential major players in the technological realm.

Benefits that can be gained from use of LBS applications vary widely depending on the type of LBS used. Mobile social networking applications, also known as Mobile Social Software (MoSoSo), have been described by Lugano (2007) as, “a class of mobile applications whose scope is to support social interaction among interconnected individuals [...] exploiting the media convergence process and the increasing power of mobile devices to offer a variety of services.” These services, such as Google Latitude, Gowalla, Foursquare, and Facebook Places, allow participants to engage in social networking from a mobile environment, “checking in” to

current locations and distributing that information to friends and social contacts. Benefits to the user of these applications may come in the guise of maintaining contact with parties of interest, economic incentives from retailers or other businesses (such as discounts offered by retailers to “Mayors” in Foursquare), or increased knowledge of the whereabouts of persons in one’s social network. Mobile navigation services such as Google Maps, Waze, and MapQuest allow a user to query location and direction information from a static point or while in transit, thus allowing for more effective decision-making regarding route planning. Recently, some of these services have been partnering with transit services, state DOTs, and other transportation data providers to provide more comprehensive and timely transportation information to, for example, allow travelers to adjust their route-planning decisions based on real-time information regarding travel times and traffic conditions. These services provide a number of benefits for the consumer, including time and cost savings, as well as security regarding the accuracy of their travel decisions. Mobile automotive assistance and emergency service applications, such as OnStar, which provides a variety of services ranging from navigation and direction assistance to emergency services and diagnostics, have also seen increased adoption as their safety and security benefits are acknowledged by consumers. Balancing the benefits of ITS and LBS adoption are concerns that individuals and organizations may have regarding adoption, including privacy concerns and general unease with new technologies. Adoption theory, presented below, may provide some indications of factors that may influence individual and organizational decisions regarding adoption.

2.4.3 Background on Adoption Theory

The topic of adoption of innovative technologies is one that has seen increasing interest in recent years. As we increasingly rely on technology for the performance of both personal and professional tasks, adoption of innovative technologies at both individual and organizational levels has become of interest both in system design and in personal interactions, as outlined in Section 2.2. Research related to characteristics of early adopters and the diffusion of innovations across society has provided information that may be useful for a variety of actors, including consumers (who may be interested in determining at what point they may be interested in adopting new technologies), developers, marketers, and organizations involved in the distribution and application of technological innovations. This section will present an overview of research related to adoption and diffusion of technologies, including how it may relate to adoption of ITS and LBS technologies described above.

Research on diffusion of innovations has been growing steadily since the 1940s. According to Rogers (1976), "The main elements in the 'classical model' of the diffusion of new ideas that emerged are (1) the *innovation*, defined as an idea, practice, or object perceived as new by an individual or other relevant unit of adoption, (2) which is *communicated* through certain *channels* (3) over *time* (4) among the members of a *social system*." Rogers posits that these elements should be evaluated in the context of relationships between individuals, and between social and societal networks. In the context of the current research, it will be necessary to determine the types of relationships extant in the traveler's current sphere of influence, including between friends (as with social networking applications), with law enforcement agencies, or with transportation service providers. Here, these relationships will

be examined both by establishing those services which have been adopted by the user, as well as determining what factors may impact likelihood of adoption

Rogers' innovation-adoption curve, shown in Figure 1, provides a graphical representation of how innovations disperse over time, along with the expected market share in relation to that adoption. The bell-shaped curve indicates a generic expectation of the percentages of populations that would be expected to adopt a certain innovation over time, with Innovators and Early Adopters acting as "proving grounds" for new technologies, and Laggards adopting innovations somewhat reluctantly. This breakdown provides a useful overview of what an application or service developer may expect in terms of adoption, though timeframes may vary based upon the type of application or service developed, as well as its' ubiquity in relation to associated technologies. For example, an electronic toll pass may see widespread adoption, but only by those drivers who routinely travel on toll roads. On the other hand, applications such as Google Maps or Foursquare may see wider adoption over a broader segment of the population due to their ability to be used on smartphones, which have been gaining broader adoption. In addition to describing the temporal aspects of adoption, another area of interest is that of characteristics of persons in each of the identified categories. Rogers and Shoemaker (1971) provided the outline of characteristics of adopter categories shown in Table II.

Figure 2: Rogers' Innovation-Adoption Curve

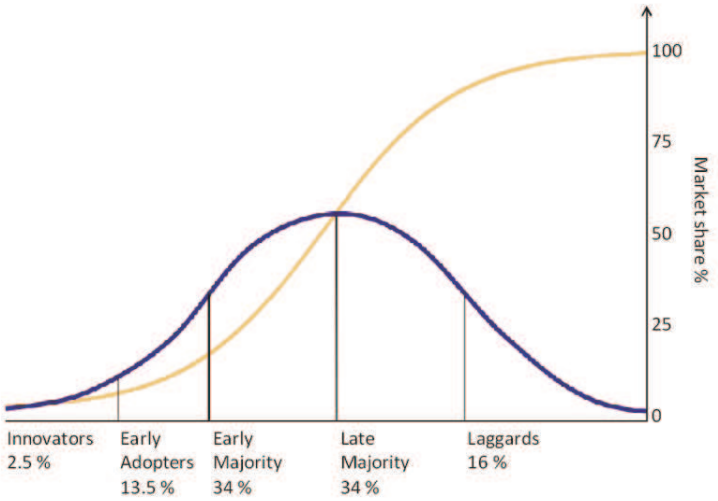


Table II: Characteristics of Adopter Categories

Adopter Category	Salient Values	Personal Characteristics	Communication Behavior	Social Relationships
Innovators	"Venturesome"; willing to accept risks	Youngest age; highest social status; largest and most specialized operations; wealthy	Closest contact with scientific information sources; interaction with other innovators; relatively greatest user of impersonal sources	Some opinion leadership; cosmospolite
Early Adopters	"Respect"; regarded by many others in the social systems as a role-model.	High social status; large and specialized operation.	Greatest contact with local change agents	Greatest opinion leadership of any category in most social systems; localite
Early Majority	"Deliberate"; willing to consider innovations only after peers have adopted.	Above average social status; average-sized operation.	Considerable contact with change agents	Some opinion leadership
Late Majority	"Skeptical"' overhelming pressure from peers needed before adoption occurs.	Below average social status; small specialization; small income	Secure ideas from peers who are mainly late majority or early majority; less use of mass media	Little opinion leadership
Laggards	"Traditional"; oriented to the past	Little specialization; lowest social status; smallest operation; lowest income; oldest	Neighbors, friends, relatives who have similar values are main source of information.	Very little opinion leadership; semiisolates

Rogers and Shoemaker, 1971

These general characteristics provide some scope for understanding the personalities of individuals and organizations involved in the adoption of new technologies from a social standpoint. Also of note, however, are the characteristics of the technologies themselves.

Fichman (2000) cites Rogers (1995) in noting the following five characteristics of technologies that may have systematic effects on assimilation and diffusion:

- Relative advantage: The degree to which an innovation is perceived as being superior to the idea it supersedes. This characteristic is generally seen as being a positive attribute relative to adoption.
- Compatibility: The degree to which an innovation is seen as consistent with existing values, past experiences, and current needs of potential adopters. This attribute is seen as positively impacting potential for adoption.
- Trialability: The degree to which an innovation may be experimented with or practiced on. This attribute generally positively impacts potential for adoption.
- Observability: The degree to which the results of an innovation are visible to users and others. This characteristic generally has a positive impact on adoption.
- Complexity: The degree to which an innovation is perceived as relatively complicated to comprehend and use. This characteristic is generally regarded as negatively impacting adoption.

Of note is that these characteristics of innovations may have impacts on both an individual and an organizational level. In particular, the networks that influence a potential adopter's perception of the degree of each of the five characteristics represented by an innovation may be significantly different on an individual versus an organizational level. Additionally, perceptions of risk may differ when envisioning adoption of an innovation for personal use versus organizational needs. While ITS and LBS technologies described above are quite reflective of these characteristics, other factors may also come into play when determining the likelihood of encouraging individual adoption.

A number of demographic characteristics have been shown that may influence an individual's adoption of innovative technologies for personal use. Such characteristics may include such components as age, education, and prior experience with technology (Munnukka, 2007). Additional work by Junglas, *et al.* (2008) indicate that the characteristics of agreeableness, conscientiousness, and openness to experience may also have impacts on concern for privacy. Finally, while males have traditionally dominated the categories of innovators and early adopters (Caruso and Salaway, 2007), there is evidence that women are becoming more willing to participate earlier in technological innovations. For example, in a study of online social networking sites, a majority of users of such sites as Facebook (63%), MySpace (63%), and Friendster (58%) were found to be female (Rapleaf, 2007). According to the Business Week analysis of the study, males tend to gravitate towards more transactional-based sites such as those targeted towards news, sports and financial information, while women's online behavior is more geared towards relationship-driven sites (Hoffman, 2008). The nature of shared information on these sites is interesting given findings in an earlier study of Internet users that indicated that males believe "censorship" to be the greatest threat to the Internet, while females cited their greatest concern as "privacy" (Herring, 2001). These characteristics of behavior online may have implications for understanding how users will respond to new technologies in the mobile environment. Additionally, by identifying characteristics that may make an individual more or less likely to adopt mobile technologies it may be possible to also better define some of the measures that may be taken to encourage adoption through the mitigation of probable concerns.

2.4.4 Conclusions

Pedersen (2005) has argued that, “For researchers, an important issue is how mobile end-user services differ from traditional [Internet and Communication Technology] ICT-services in ways that affect their adoption. For example, the personalization, location specificity and ubiquity of these services are suggested as important characteristics making their adoption different from other ICT-services.” In essence, while traditional ICT services provide specific benefits for use, evaluation of these benefits may not adequately represent a user’s experience with mobile technologies as used in ITS and LBS. Clarification of the benefits that may be gained from adoption of ITS and LBS technologies, shown above, provides further support for encouraging the adoption of these technologies via clearer privacy policies. If these benefits are presented to consumers in an understandable way in conjunction with information that may allay concerns such as privacy loss or distrust, the likelihood of adoption and use may be increased. Current shortfalls in the presentation of benefits and risks will be further examined in Chapters 5 and 6.

2.5 Justice Theory

2.5.1 Introduction

The privacy theories presented here are also closely tied to theories of justice and civil liberties. This section will provide a brief overview of such theories, and indicate how they relate to the current study.

2.5.2 Background

Rawls' book, A Theory of Justice (1971, revised in 1975 and 1999), has formed the basis for much justice theory. Here, Rawls proposes an "Original Position," in which principles are chosen by individuals from behind a "veil of ignorance," described in the following way: "…no one knows his place in society, his class position or social status, nor does anyone know his fortune in the distribution of natural assets and abilities, his intelligence, strength, and the like. I shall even assume that the parties do not know their conceptions of the good or their special psychological propensities (Rawls, 1999)." In relation to the current discussion of privacy, such justice theory applies in a number of ways. Garrett (2005) argues that, "John Rawls could defend the right to privacy by pointing out that our representatives in the Original Position would include privacy rights among the adequate scheme of Equal Liberties guaranteed by his First Principle. Privacy enables us to pursue our personal conceptions of the good, with those we wish to associate, so long as we do not violate the rights and liberties of others." Such a conception would indicate that privacy is a basic liberty (a viewpoint supported by interpretation of the Fourth Amendment). By first placing the right to privacy within the Rawlsian framework, we may next evaluate how privacy as justice may be evaluated within the current environment.

2.5.3 Current Implications

The preponderance of surveillance and dataveillance in today's society has led to increasing attention being paid to the ethics and impacts of attendant privacy concerns. If we are to assume that privacy would be considered a basic right under Rawles' justice formulation, then invasions of privacy would engender concerns related to justice and rights. Such discussions

have recently been lively in the area of ubiquitous online marketing technologies, with Ashworth and Free (2006) arguing that, “an important component of consumers’ privacy concerns relates to fairness judgments, which in turn comprise of the two primary components of distributive and procedural justice.” In Ashworth and Free’s (2006) description, “Distributive justice relates to the perceived fairness of the allocation of outcomes and is assumed to reflect a concern for one’s material well-being…[P]rocedural justice refers to the fairness of the rules or policies that are used to allocate outcomes.” By invoking the conception of “fairness” here, Ashworth and Free indicate consumer concern for the following:

- That notice has been given that data collection will occur;
- That the consumer is treated as a valued and respected individual; and
- That the material outcome is comparable to the information provided.

Consumer determinance of whether these components have been met is subject to normative standards such as openness, information access, permission, and honesty. Such norms are in keeping with the Federal Trade Commission’s Fair Information and Privacy Principles, discussed further in Chapter 3.

The ubiquity of such marketing is a concern translatable to the mobile environment. In essence, the fairness of collection of data on consumers by ubiquitous technologies and techniques may be called into question, particularly if adequate notice is not given. If consumers are assumed to “opt-in” to use of such techniques by simple use of a service, they may feel that the norms of openness, permission and honesty have been violated. Control and choice also come into play, as their withholding may indicate to the consumer that he or she is not a valued and/or respected individual. Finally, lack of information pertaining to collection, use and access to collected data by companies or organizations may cause consumers to

question whether the inputs (data) and outcomes (benefits) are balanced, thus heightening concern for fairness. In the mobile environment, rapidity of change and a plethora of involved agents (including mobile application developers and local, state, and federal transportation agencies) may increase a consumer's likelihood of assuming that fairness has not been met, thus calling justice into question.

Another issue that is important to note in terms of justice in relation to privacy is that of sensitivity of information. Ashworth and Free (2006) describe the issue of information sensitivity in terms of distributive justice by stating the following:

> First, the collection of sensitive information is likely to reduce consumers' outcome of the exchange because the potential consequences associated with the collection of sensitive information are more severe than the consequences associated with less sensitive information. Second, sensitive information may well increase consumers' evaluation of their input to the exchange as they are now providing information they perceive to be more valuable.

In this manner, by collecting "sensitive" information (such as name, address, or financial information), companies may upset the valuation of equity by the consumer, thus leading to a feeling of injustice. This issue may be compounded if, as noted above, adequate notification is not given.

In terms of determination of justice relating to privacy in the mobile environment, outcomes (or benefits) to the consumer may not be immediately noticeable or understandable. While professionals in the field may understand that having data related to travel patterns, social networks, or mode preferences may increase the efficiency and cost effectiveness of transportation planning and policy, the volume of data needed to make such improvements may be beyond the ken of the average consumer. While explanations may be attempted, a lack

of immediate, easily discernable benefits may upset the consumer's beliefs towards distributive justice, and the lack of explanation may hamper feelings of procedural justice.

The issue of information and data privacy, defined by Culnan and Bies (2003) as, "the ability of individuals to control the terms under which their personal information is acquired and used," within a technological environment raises the concerns addressed above. In addition, concerns may be raised about the use of such data, particularly in regards to sharing with third parties, and the mining of data. If consumers are not informed that their data will be shared with third parties, particularly if such sharing is for economic benefit on behalf of the collecting company, they may feel that the norms of honesty, information access, and permission have been violated. If data are combined and mined, sensitive information may be revealed, which the consumer would have preferred been kept private. Such a combination of factors are especially troublesome in the mobile environment, as mobility and location information may be used to determine, as noted by the courts in Chapter 3, habits, preferences, political leanings, and other information that may be valuable to marketers and application developers, but that may also be regarded as highly personal and private by consumers. Equity may also become an issue here, as differing educational levels and economic access to services may create disparities in understanding of privacy policies as well as in benefits and protections obtained. Though this issue will not be addressed directly by the current dissertation research, it will be an issue that demands attention in coming years.

2.5.4 Conclusion

The issue of justice in relation to privacy in the mobile environment is one that builds on previous conceptions of rights as they apply to the population in general, and to practices

surrounding ubiquitous data and information collection. The use of policy strategies, such as the FCC's Fair Information policy, to address these issues has been the most common approach to ensure retention of rights and attention to justice; however, the rapid evolution of technologies and associated changes in expectations will require that additional attention and management take place.

2.6 Conclusions

The theory of contextual integrity is highly applicable to the conceptualization of ITS and LBS technologies as components of a socio-technical system. By framing the dissertation subject within the realm of bi-directional influences of societal, personal, and technological norms, we may better understand the impacts and expectations that may be expected from the need for personal private data to interact within these systems, and the desire on the part of societal actors to have these data protected while not forfeiting the concomitant benefits anticipated by the implementation of these systems. These concerns and the trade-offs between benefits and costs, particularly in relation to theories of justice, will also impact the potential for adoption as discussed in Section2.4.

The current lack of consumer awareness, explored more fully in Chapter 6, reflects concerns related to the effective implementation of ITS and LBS. As shown here, an effective socio-technical system requires the recognition of contextual changes in spheres of influence, and appropriate measures taken to reflect these contexts. Currently, lack of knowledge on the part of consumers regarding the collection and treatment of data makes it difficult for these spheres to be adequately reflected. Such difficulty, which also indicates inadequate attention to

justice in the mobile environment, has the potential to negatively impact potential for adoption, as well as setting the stage for unrealistic expectations of privacy. The next chapter will provide more detailed information related to the conceptual and definitional components of privacy within the framework discussed above.

CHAPTER 3: THE STATE OF THE ART IN LOCATIONAL PRIVACY: REVIEW AND ANALYSIS OF EXISTING POLICY, LEGAL AND TECHNICAL APPROACHES

3.1 Introduction

The following review and analysis of relevant literature, law and technology focuses on the underpinnings of privacy in location services, as well as foundational studies and overviews related to privacy in Intelligent Transportation Systems and Location Based Services. Of note is that the review will be fairly wide-ranging in scope, as contributing elements to privacy theory and concepts hinge on a number of factors, particularly in the context of location and mobility studies. Because of the variety of actors and agents concerned with the provision, collection, use, and sharing of data in the mobile environment (including individuals, public agencies, and private organizations), current legal, technical and policy approaches to privacy must be evaluated from a number of viewpoints. This chapter will attempt to provide the reader with a thorough overview of both the context of privacy, as well as analysis of how privacy is currently treated within the regulatory framework. Such an approach will better establish the concerns to be addressed in the empirical portion of the dissertation.

One issue of note is that while technical studies reviewed here are applicable to the questions at hand, many are somewhat limited in scope, as they primarily focus on only one aspect of privacy in ITS and/or LBS. It is hoped, however, that by evaluating these studies the complexities of issues apparent in locational privacy will become more evident to the reader, and will identify the broad spectrum of concerns that must be addressed when defining relevant issues. In order to accomplish this task, definitions of privacy will first be presented to

establish the context from which the study will progress. Next, legal issues related to privacy in the public sphere will be addressed, followed by a more policy-oriented overview of privacy considerations. Following this, prior examples of how privacy has been treated in the Health Insurance Portability and Accountability Act (HIPAA) and the Fair Credit Reporting Act (FCRA) will be presented in order to provide a context for governmental approaches that may be taken to provide protection of private information. Finally, policy documents establishing a framework for privacy preservation in contexts related to location services will be synthesized in order to identify gaps and areas of concern, as well as providing a basis for better understanding the content analysis of privacy policies.

3.2 Defining Privacy

Many definitions of privacy have been proposed, but most tend to have issues of control of information and its flow as their foundations. Westin (2003) has defined privacy as, "the claim of an individual to determine what information about himself or herself should be known to others." This broad definition contains within itself a wealth of further claims related to different states of privacy, and to the context of the person and his or her information. By approaching the privacy claim from the viewpoint of context, as reviewed in Section 2.3, the emerging literature on the social, political, and economic variations inherent in the experience of privacy reveal a range of expectations dependent upon the person's individual understanding. The concept of privacy as based upon a subjective or contextually-based understanding, as described in Section 2.3, is also consistent with the legal understanding of the subject – for example, the Fourth Amendment, central to legal justifications for privacy

protection, has been understood by the courts to be centered on “reasonable expectations of privacy” (Slobogin, 2002). What should be addressed, then, is how the central definition of control and expectation as the foundation of privacy claims is manifested in various arenas. This section will outline the components of privacy as defined within the domains of government, economics, and social interactions, and will then examine the claims set forth in the context of locational privacy.

3.2.1 Privacy in Policy

Federal, State and Local agencies are bound by numerous privacy requirements, including, though not limited to, the Fourth Amendment, the 1974 Privacy Act, and the Electronic Communications Privacy Act (ECPA), as well as codes and regulations specifically applicable to individual agencies, such as Title 13 of the U.S. Code, which pertains to privacy requirements for the U.S. Census. While this thesis will not attempt to provide an exhaustive overview of privacy requirements pertaining to governmental agencies, it will be necessary to analyze the most fundamental principles insofar as they pertain to citizens and other residents/visitors. The Federal Trade Commission’s (FTC) “Fair Information Practice Principles” provides a fairly clear example of the elements that comprise the government’s approach to privacy from the standpoint of the consumer. The Principles identify five “core principles” relevant to privacy policy (FTC, 2007), namely:

(1) Notice/Awareness;
(2) Choice/Consent;
(3) Access/Participation;
(4) Integrity/Security; and
(5) Enforcement/Redress”.

These principles form the government's definition of privacy as it may be reasonably expected by the consumer.

The first FTC principle, notice/awareness, may be considered as the most fundamental, as it sets the context for the remaining four. According to the FTC (2007), "Consumers should be given notice of an entity's information practices before any personal information is collected from them. Without notice, a consumer cannot make an informed decision as to whether and to what extent to disclose personal information." The second principle, choice/consent, rests upon the belief that consumers should have options regarding if information will be collected and how that information will be used. This principle is particularly relevant in relation to secondary usage of data, in which information collected for one purpose may be used for a different and potentially unrelated purpose. The access/participation principle "refers to an individual's ability both to access data about him or herself…and to contest that data's accuracy and completeness" (FTC, 2007). Integrity/security is concerned with the responsibility of collectors to ensure the integrity of collected data via such means as cross-referencing against reputable data sets, as well as with ensuring that collected data is protected from loss and unauthorized access via both managerial and technical means. Finally, enforcement/redress is intended to ensure the efficacy of the preceding principles by providing a mechanism for enforcement. From a definitional standpoint, these principles expand the notion of "control" that rests at the heart of the broad definition of privacy by introducing the associated concepts of notice, choice, access and security. These concepts are particularly relevant from the point of view of government and commerce, insofar as they are related to information privacy. These

principles will be discussed in more detail below in order to provide a framework for discussion of applicable industry privacy guidelines.

3.2.2 Economic Approach to Privacy

From the viewpoint of economics, the concept of control in relation to data ownership becomes a primary concern. According to Stigler (1980), "'Privacy' connotes the restriction of the collection or use of information about a person or corporation; the information in question 'belongs' to the individual." According to Zahn and Rajamani (2008), Stigler further defines privacy as, "the concealment of useful information assuming an economic value in transactions." Stigler's definition assumes that control of data has an economic grounding – our ownership and control of data is defined in relation to our willingness to trade it for an economic benefit. In these findings, privacy is defined as a marketable commodity, subject to economic evaluation and individual willingness-to-trade. Under this definition, the desire of the individual for privacy is balanced against the market's willingness to "buy" information, whether through economic incentives (such as offering discounts to persons who use frequent shopper cards that track purchasing habits) or through incentives of convenience (such as the Department of Homeland Security's "Trusted Traveler" programs). "Control" under this definition is thus primarily understood as ownership of data at the point of contact, and may be subject to limitations over control of secondary use once the data has been "bought" by an outside entity. The discussion below will further expand upon the economics of privacy, particularly in relation to the transportation context.

Some researchers have argued that privacy should be regarded as a marketable commodity that may be traded in return for economic or other benefits. In this conception, the

market for personal data will self-regulate as customers establish the price at which they will be willing to sell personal data, while the market will establish the price which it is willing to buy the data. De Boni and Prigmore (2001) offer perhaps the most succinct summation of this viewpoint when they state that, “customers who value their privacy will be unwilling to enter into transactions with businesses that do not respect their customer’s privacy. Thus, if privacy truly is a concern for most consumers, market forces will ensure that businesses that respect customer privacy will succeed; those that do not will fail.” In an economy based on the exchange of physical goods, this explanation works reasonably well, considering DeLong and Froomkin’s (2000) argument that in the traditional market system the three features of excludability, rivalry, and transparency dominate property rights and exchange. In the information economy, however, these three features are subject to debate, as data and information are subject to fewer physical restrictions. In a primary market with one seller and one buyer, it is reasonable to assume that market forces will work reasonably well. In a cross-market system, however, where information purchased in one market may then be sold for use in a second market without diminishing the ability of the initial purchaser to use the information, the situation becomes decidedly murkier. The following section will outline the general argument for a market-based approach to location privacy, and will then review the primary arguments against such a conception.

Stigler (1980) has defined privacy as, “the restriction of the collection or use of information about a person or corporation; the information in question ‘belongs’ to the individual.” While the understanding of private information as owned by the individual is fairly common, De Boni and Prigmore (2001) note that, “[this conception] is substantially

undermined by the commonly held view that, '…a market for personal information would not have as its first objective the regulation of privacy abuses or the enhancement of privacy, but rather, an economically more efficient flow of personal information…' The right to information privacy is seen as subsidiary to the right of businesses to do business." Such a market-oriented approach is understandable if one takes the view of the value-added nature of the use of information in market efficiency. Zahn and Rajamani (2008), for example, argue that the collection and analysis of information by businesses allows them to achieve greater efficiency by, "better understanding the needs of existing and prospective consumers and effectively assigning services and products to cater to their requirements in a cost and time-efficient way." In this case, the value that an individual would place on certain elements of personal information is less than the valuation of the same information as it may be used by a buyer, thus creating a benefit for the buyer.

One question that the concept of personal information as property subject to value-adding processes in the marketplace raises is that of property as intentional versus incidental. In this formulation, "intentional" information may be conceived of as private data that is intrinsic to maintaining the actor's control over how he or she is perceived by others (reputation) and when he or she wishes to interact with others. The individual will likely place a fairly high value on this information, as loss of control over its access and dissemination may have immediate social or financial repercussions. "Incidental" information, on the other hand, is data created secondary to intentional information and is likely not valued as highly. Under this understanding, location and travel data along a route would be incidental to the intentional information of origin and destination. As such, if market mechanisms are put into place that

would allow for the collection of data pertaining to route choice at a fairly high resolution, the relative price would likely be lower than if low resolution data is collected that contains both intentional and incidental information. In this case, the incidental route choice information may be very valuable to transportation planners and others once it has been incorporated into value-adding processes such as analysis of traffic growth on a corridor over time. In this manner, a market-oriented approach to the dissemination travel and location data seems quite pragmatic.

As noted above, however, such an outline considers only a one-to-one primary market structure. As Stigler (1980) points out, however, "The primary peculiarity of information as a property right is commonly held to be its public goods character: if *A* gives (sells) information to *B*, there is usually no efficient way to insure that *B* does not disseminate the information to *C* (while still retaining possession of the information)." In the case of location and other types of information property, the potential for dissemination beyond the initial transaction may be the point at which a market structure becomes more problematic. As noted in Danezis, *et al*. (2005), students who took part in a compensation auction of GPS data increased their bids when it was made known to them that there was commercial interest in the collected data. Here, it may be inferred that knowledge that the data was seen as valuable to a third party increased the relative value that respondents placed on their information. In the initial Danezis survey, participants were informed that collected location data would be retained and possibly used again for further research. Based on the wording of the survey, participants could infer that such further research would also be conducted by the University group. In this case, the

initial market would remain stable. The introduction of knowledge of the potential for cross-market sharing, however, increased the value of the data.

According to Hui and Png (2006), “In deciding how much personal information to reveal, consumers balance the benefit from consuming the primary item against direct privacy costs. The higher the rate at which consumers expect sellers to cross-sell personal information, the less information consumers would reveal.” The ability of the seller to accurately price his or her intentional and incidental data, however, is contingent upon knowledge of how that data will be used by the buyer. Here, because information property does not meet the requirements of excludability and rivalry, it may be possible that the seller will undervalue his or her actual private information. Hui and Png (2006) here note the potential for cross-selling to result in unsolicited promotions, which may intrude on the value of seclusion. On the other hand, the potential benefits of cross-selling in some instances (such as dissemination of location data from a state department of transportation to a city office of emergency management) may be beneficial enough that a potential seller will overvalue certain aspects of privacy. In this case, it is reasonable for some regulatory intervention to ensure that adequate knowledge about both primary and secondary markets is given to the individual.

A final consideration that should be addressed is that of personal valuation versus societal valuation from the viewpoint of government services. The greatest benefits from collection of travel and location data by government agencies are related to the potential to make transportation networks more efficient. In this case, while the individual traveler whose data are collected may not receive significant individual benefits, the society that uses the network will. Given that it is the aggregation of individual data records that would allow this

potential benefit to accrue, the individual must make his or her valuation based not only on his individual preferences, but also on potential societal benefits. In the case of government entities subject to legal requirements for use and dissemination of individual data, the valuation may be relatively simple. However, the transportation network is composed of both public and private entities, and the potential for sharing of collected data for uses beyond network efficiency and subject to legal restrictions is great. Given these conditions, the traveler must determine his or her value of location data based on individual and societal benefit versus potential cost of intrusion and dissemination beyond the primary market.

As shown here, the treatment of private data and information as marketable commodities is somewhat problematic given competing notions of valuation, ownership, and use. With good information and reasonable understanding of primary and secondary uses, the individual and societal benefits would likely make a market price for information acceptable to the seller. However, without this information, the potential for the seller to significantly over- or under-value his or her privacy is great. For the buyer, less information is likely beneficial, as studies have indicated that while persons indicate a high preference for privacy, they are willing to sell at a relatively low price. The market argument for privacy is compelling to a point; however, the barriers identified above should be taken into consideration. The relationship between economic incentives and willingness-to-trade will be further explored in the general survey discussed in Chapter 6.

3.2.3 Privacy as Contextually Defined

Finally, the definition of privacy within the social arena is subject to a number of contextual factors, as noted in Chapter 2. Westin (2003) provides what is perhaps the most basic

breakdown of the experience of privacy by the individual, identifying the following four states: Solitude, intimacy, anonymity and reserve. Margulis (2003) defines the four states as follows:

- Solitude: Being free from observation by others.
- Intimacy: Small group seclusion for members to achieve a close, relaxed, frank relationship.
- Anonymity: Freedom from identification and from surveillance in public place and for public acts.
- Reserve: Based on a desire to limit disclosures to others; it requires others to recognize and respect that desire.

In this formation, privacy is defined in relation to the environment in which an individual is based and provides scope for altering the degree of disclosure of private information based upon that environment. In this manner, the context of privacy broadens beyond a singular experience between an individual and a uniform "other" and enters into the space of individual and group relationships. By understanding privacy within the dynamic of both the individual's relationship to the broader world and the interaction between the individual and his or her identified group memberships, the degree of control over information that the individual expects is allowed to expand or contract within the context of a social contract while still affording the individual the ability to control the disclosure or use of that information beyond the bounds of the specified situation.

3.2.4 Locational Privacy

The above analysis has focused primarily on general concepts of privacy as understood in a variety of socio-political arenas. We turn now to evaluating the identified components in the specific realm of locational privacy. Because it is concerned primarily with privacy in the realm

of travel taking place in the public sphere, and due to the variety of commercial and governmental interests in tracking travel patterns and habits, the issues of control and context are particularly relevant to the understanding of locational privacy. The definitions cited above, however, are primarily concerned with the understanding of privacy as a static concept that allows time for reflection and decision-making. While many of the cited privacy components are relevant to locational privacy – such as knowledge of and control over collected data, the economic value of private data, and the definition of privacy as dependent upon individual and group relations – the dynamic aspect of locational privacy introduces a spatiotemporal element that has not yet been adequately developed within the literature.

From a general standpoint, traditional concepts of privacy take place at a point of reference, while locational privacy exists as a rapidly changing route. This element of change over time and space heightens the degree of awareness and information needed by the individual to make informed decisions regarding the collection and use of his data and increases the potential that aspects of control may be lost. For example, when an individual uses the Internet he or she is bound by the privacy policy of the associated ISP. If the individual wishes to switch ISPs, he or she has the ability to review and accept or decline that ISP's policy. In a dynamic environment, the rapidity of change related to context and authority may make such consideration impossible in terms of efficiency. This, in turn, heightens the importance of where and when decision-making takes place, indicating that the definition of privacy may need to be expanded to include a spatiotemporal aspect for the purposes of locational privacy. Additionally, collected data that include both spatial and temporal identifiers may increase the likelihood that personally identifying information may be gleaned, as it is possible to determine

not only where a traveler has been, but also at what times, thus introducing the potential to identify activities that may have taken place.

3.2.5 Conclusions on Definitions

The above review demonstrates the plurality of definitions of privacy that exist in view of the context in which they are referenced. The fundamental basis of control – over data, access, value, and social interactions – holds true for all, but how that control is understood manifests itself in a variety of ways. The interaction of the social, political and economic within the transportation experience highlights and enhances the degree of overlap between the components. Additionally, the introduction of the spatiotemporal dynamic in reference to locational privacy initiates a need to consider the potential for rapid change within the immediate experience of privacy by the individual. For purposes of this research, locational privacy will be studied in response to economic questions (what are the risks associated with providing data in the mobile environment and what compensation or benefits do travelers expect to receive in return), as well as in context (with whom and for what purposes are travelers willing to share private information). It is hoped that such an approach will allow for exploration of privacy matters relevant to travelers in the mobile environment.

3.3 Dimensions of Privacy

3.3.1 Introduction

As is evident from the above review, the concept of privacy is not one easily defined or understood. One method that has been used to address this issue is the use of

multidimensional constructs. Multidimensional constructs allow several related but distinct dimensions to be treated as a single theoretical construct (Edwards, 2001), thus providing the ability to address complex subjects in a relatively streamlined manner. The following section will review and evaluate different approaches that have been taken to describe privacy as a multidimensional construct.

3.3.2 Background

In "Privacy as a Concept and a Social Issue: A Multidimensional Development Theory", Laufer and Wolfe (1977) argue that in order to understand privacy as both a contemporary and a future social issue it must first be understood from a conceptual viewpoint. To establish the concept of privacy, the authors use a multidimensional construct approach that focuses on the environmental, interpersonal and self ego aspects of privacy, and argue that, "This multidimensional structure enables us to understand perceived privacy and privacy invasion as well as to predict the types of situations that can potentially create privacy or invasion experiences" (Laufer and Wolfe, 1977). In the years since this article was published, a number of authors, including Westin (2003) and Stewart and Segars (2002), have added additional elements to this fundamental concept, expanding it in terms of informational privacy and establishing a system that categorizes the public based on broad levels of privacy concern. The resulting constructs establish a good foundation for the examination of locational privacy concerns, and provide a set of considerations that should be addressed in its development.

Laufer and Wolfe's construct hinges on the conception of privacy as an individual as well as a social-historical concern, requiring the elucidation of privacy from both individual and normative perspectives. They argue that the two perspectives are interdependent, with time as

the dynamic basis for the interdependency. A bi-part conception of time as short-term (individual) and long-term (societal) becomes relevant as the authors expand upon their tri-part dimensions as follows:

- Self-ego dimension: A process of development that focuses on individuation and personal dignity, with key components of this concept including voluntary and enforced aloneness.
- Environmental dimension: Those elements (cultural, sociophysical and life cycle) that create boundaries to meaning and experience.
    - Cultural: The mores of a society that establish lines between actions understood as private and those that are public; these mores change over time.
    - Sociophysical: Related to the interplay of place and society and how their experience and demands construct both individual and societal concepts of private v. public behavior.
    - Life cycle: Understood from the viewpoint of the individual from birth to death, as different periods of time within the developmental process will leave him or her subject to different roles and changes in societal mores.
- Interpersonal dimension: Defined in relation to an individual's relationship to others via the management of information.

Within this framework, the authors posit that the issue of control/choice functions as a mediating variable, as the choice to relate or separate from others is experienced in any given privacy situation. The authors, however, feel that this issue, while influential, is conceptually separate from the construct of privacy. The construct as outlined generally conceives of privacy as being a function of the relationships between self and self, self and others, and self and the environment located within the potential for these relationships to change over time.

Stewart and Segars (2002) also take a multidimensional approach, using an instrument developed by Smith, *et al*. (2001) that reflects the following four factors of concern in information privacy: collection, errors, secondary use, and unauthorized access. This construct suggests that, "individuals with a high concern for information privacy perceive that: (1) too

much data are collected, (2) much of the data is inaccurate, (3) corporations use personal information for undisclosed purposes, and (4) corporations fail to protect access to personal information" (Stewart and Segars, 2002). The authors conducted a survey of 400 consumers to determine if the posited construct adequately reflected information privacy concern. The findings indicated that the four factors identified above likely function as second-order factors, subject to a larger theme such as consumer control of information or procedural fairness. These findings closely reflect the concerns addressed in the Fair Information Practice Principles outlined by the FTC in Section 3.2.1.

Finally, Westin's privacy indices classify the public into the three categories of high/fundamentalist, medium/pragmatist and low/unconcerned. In the earliest studies, conducted for Harris-Equifax in 1990 and 1991, Westin (2003) used the following four questions to classify individuals:

> (1) Whether they are very concerned about threats to their personal privacy today, (2) Whether they agree strongly that business organizations seek excessively personal information from consumers, (3) Whether they agree strongly that the Federal government since Watergate is still invading the citizen's privacy, and (4) Whether they agree that consumers have lost all control over circulation of their information.

Westin (2003) classified as privacy fundamentalists those persons who responded affirmatively to three or four questions (roughly 25%), pragmatists as those who responded affirmatively to two questions (roughly 57%), and unconcerned those who answered one or none affirmatively (roughly 18%). While Westin has conducted numerous additional studies on this topic, his use of the tri-part division of the citizenry has continued, and has been adopted by many in related literature. Westin's approach differs from the other two reviewed here, insofar as it aims to

classify consumers based on privacy concerns rather than establishing a methodology by which to define or measure privacy concerns themselves.

3.3.3 Privacy in Relation to Trust

In discussing privacy, a related concept that should be addressed is that of trust. As in the contextual understanding of privacy, relationships between data producers, collectors, and consumers are based on a number of factors, including the degree to which data producers trust that the collectors and consumers will respect the contextual norms associated with these data. The relationship between trust and privacy has been examined by a number of researchers, including Karvonen (2010), Liu, *et al.* (2004), and Metzger (2004). Generally, it is found that the concepts of privacy and trust are closely linked, with trust serving as an intermediate variable in consumer willingness to release private information to agencies and organizations (Liu, *et al*. (2004)).

Within the framework of ITS and LBS, trust may be displayed in two ways – first is trust in the system itself, for example, trust that accurate directions have been given or that correct information is being provided. A second factor relates to trust in the agencies providing information. Here, trust manifests as a belief that data will be treated appropriately within the framework of contextual norms. In the context of ubiquitous computing ("ubicomp") in the mobile environment, Karvonen (2010) states the following, "Ubiquitous systems gather information from their users and the user has to be able to trust the system to give out the needed information regarding him/her. Furthermore, the ideology of invisibility with ubicomp systems causes extra requirements for the development of user acceptance and trust." Ubiquitous computing in the mobile environment, whether in the form of current technologies

such as location-based applications and GPS-enabled mobile technologies, or of proposed technologies such as peer-to-peer safety information, may be largely invisible to those in the system, thus requiring both enhanced consumer data protection, as well as fairly transparent implementation. Recent privacy-related violations from such companies as Facebook (Helft, 2010), Apple (Zyskowski, 2011), and Google (Halliday, 2010) have put privacy concerns at the forefront of trust issues. Such a trend is clearly evident in findings from a recent survey conducted by Harris Polls for TRUSTe (2011) which showed that, "Privacy concerns rank #1: Most consumers expressed great concern about their data privacy both when using smartphones in general, and when using mobile apps in particular; this concern increases with the age of the user."

The connections between privacy and trust may be especially relevant given the limited ways that companies and organizations may communicate to consumers the methods by which their private data are protected. Privacy policies (detailed further in Chapter 5) may not adequately address all facets of user concern regarding privacy, and may, in fact, be unread by consumers. Negative publicity such as that cited above, or general distrust in government agencies or private corporations may enhance these concerns, and erode the trust necessary to help enhance likelihood of technology adoption on the part of consumers. No clear method of allaying these concerns exists, beyond ensuring that companies and organizations act in a manner that does not engender privacy concerns, or providing transparent and clear information regarding how collected data are to be used, managed, and shared. Allowing users to "opt-in" to services and providing them some measure of control over how their data are used may also help ease trust concerns.

3.3.4 Application to Location Privacy

The contributing factors to privacy examined here each add a layer to the general concept. We now turn to examining how they impact the understanding of locational privacy. As discussed in Section 3.3.2, perhaps the key element of locational privacy in relation to general privacy is that of the rapidity of change. While Laufer and Wolfe specifically address the temporal aspects of privacy, they do so in a manner that still assumes relatively slow motion (over a lifetime or longer). The dynamic nature of travel and location change demand a more instantaneous understanding, and will perhaps necessitate an expansion of the temporal understanding to reflect the necessity of traveling rapidly between societies and their mores (for example, the crossing of state boundaries where different privacy regulations exist). While Laufer and Wolfe discuss elements of change within a society, they are less clear on aspects of change between societies and environments, where cultural and environmental settings and expectations differ. Insofar as the self remains the same, the self-ego aspect may remain broadly stable, but the introduction of the spatiotemporal aspect inherent in locational privacy demands a more nuanced approach to the levels (such as local, state and federal) at which privacy is experienced.

Stewart and Segars' construct is relevant for locational privacy insofar as ITS networks necessitate the use of surveillance and the collection of varying amounts of data for maximum efficiency. By observing that the aspects of collection, errors, secondary use and unauthorized access may be subject to the greater concern for control, the authors implicitly highlight the need to closely attend to elements of notification, as without knowledge that data has been collected the element of control is lost and the other aspects rendered nearly meaningless.

Again referencing the spatiotemporal nature of transportation, this becomes particularly relevant in a dynamic environment where authority over data collection may change rapidly. Finally, Westin's categorization of the public into different levels of privacy concern accentuates the need to address privacy from a variety of vantage points. Persons within the category of "fundamentalists" will have greater issues associated with the collection and sharing of data relative to those persons within the pragmatist or unconcerned categories, and it will be necessary to address these concerns in such a manner that prevents the loss of benefits for the greater traveling public in order to gain the acceptance of a minority, while not denigrating the valid concerns that they bring to the table. The concept of trust contributes to all of these factors, as it will play an underlying or mediating role in how a consumer views privacy in relation to application and decision-making. In short, the constructs outline above provide the foundation for an understanding of the elements that define locational privacy, but may fall short in accounting for the realities of the environment in which travel takes place.

3.4 Legal Issues of Privacy in Public Places

3.4.1 Introduction

The legal issues associated with privacy in ITS and LBS are primarily based in the area of "privacy in public". Because much surveillance-based ITS technology (such as red-light cameras, electronic toll collection, and traffic cameras) is utilized on the public roadway, there has been some discussion as to how far the right to privacy may reasonably be expected to extend beyond the confines of the private domain. The following section will provide review and critical examination of the legal issues associated with this topic, including a review of the legal

foundation for privacy in public, and the findings from two recent court cases addressing the legality of the use of GPS technology to track a person of interest.

3.4.2 Framing the Argument for Privacy in Public

Slobogin (2002) frames the issue of privacy in the public sphere within the U.S. Supreme Court decision in United States v. Knotts, in which the Court found that the Fourth Amendment does not apply in the case of tracking a car's movement via the use of an electronic beeper. He cites the Court in its' decision that, "[a] person traveling in an automobile on public thoroughfares has no reasonable expectation of privacy in his movements from one place to another." The author also notes that the Court's decision was reached by citing the principle of "reasonable expectation", and asks the general question, "When one's every movement is readily observable by others, how can one expect constitutional protection of those movements?" This finding and Slobogin's question address the general issue of what amounts to a "reasonable expectation" of privacy in public.

To answer generally the degree of privacy expectation we may have in a public space, Slobogin cites the "bar example" originally presented by Supreme Court Justice William Rehnquist, in which he states, "[T]here would be an uneasiness, and I think a justified uneasiness, if those who patronized the bar felt that their names were being taken down and filed for future reference," in order to make the broader point that he feels the general public believes we have a right to public anonymity. Slobogin addresses this point by attempting to establish a constitutional right to anonymity in the following three ways:

1. By showing that, “indiscriminate technological public surveillance seriously undermines the way we would like our society to function, because of its effect on public anonymity.”
2. By arguing that, “a number of constitutional principles, while not explicitly recognizing a right to public anonymity, provide solid groundwork for it.”
3. By reporting the results of an empirical study “that suggests that American citizens feel public camera surveillance by the government is more intrusive than a variety of other police actions that the Supreme Court has labeled a ‘search’ or ‘seizure.’” (Slobogin, 2002)

In reference to the first point, Slobogin (2002) argues that excessive public surveillance will impact society’s citizens by instilling a degree of fear that will cause them to, “act less spontaneously, more deliberately, less individualistically, and more conventionally…” In short, Slobogin contends that the type of surveillance cited here will lead to a lapse in the individuality and an increase in the conformity of the citizens that it is hoped to protect. On the second point, Slobogin (2002) cites a number of Constitutional bases related to surveillance and its privacy implications, including the First Amendment right to free assembly, the Due Process Clause of the Fourteenth Amendment, and the general right to privacy found in “the penumbras of the First, Third, Fourth and Fifth Amendments, …or the Ninth Amendment’s reservation of rights to the states.” He holds the Fourth Amendment, however, as most applicable to the issue of privacy in public. In reporting on the Court’s finding in Katz v. United States, in which government agents bugged a phone booth, Slobogin quotes the decision that, “what a person seeks to preserve as private, even in an area accessible to the public, may be constitutionally protected” (Slobogin, 2002). He also notes, however, in relation to the Knotts case cited above, that the Court additionally found that, “what a person knowingly exposes to the public…is not a subject of Fourth Amendment protections” (Slobogin, 2002).

Slobogin addresses his third point via data from an empirical study, in which he focuses primarily on the concept of the "reasonable expectation" of privacy to develop the context in which privacy is understood by the Courts under the Fourth Amendment to the United States Constitution. To develop a method by which this "reasonable expectation" may be codified, Slobogin conducted a survey of 190 persons in Gainesville, Florida called for jury duty.[2] Survey participants were asked to evaluate 15 scenarios involving various forms of surveillance in which, "the police were looking for evidence of crime but that the target of the police action had not engaged in any criminal activity" (Slobogin, 2002). The survey results indicated that participants have a higher expectation of privacy than is typically recognized by the Court, insofar as some actions allowable under the Fourth Amendment were generally regarded as fairly intrusive (such as helicopter flights 400 feet above a backyard, being followed by an officer, and having garbage searched through on the curbside). Based on these results, Slobogin (2002) states that, "Put simply, the participants are better than the Court at identifying expectations of privacy society is prepared to recognize as reasonable."

Slobogin's findings here indicate that by placing the onus of privacy determinations on the expectations of the public, scope for imbalance emerges. Given that the conducted survey revealed that actions currently defined by the law as "legal" were viewed as intrusive by many respondents, it is possible to argue that a lack of awareness regarding allowable privacy practices creates an *unreasonable* expectation of privacy. This finding will be further reviewed below, in the content analysis of privacy policies and consumer survey.

[2] The generalizability of these results may be questionable, as jury pools are selected from registered voters, who may have different characteristics than the population as a whole. See Holder, 2006.

3.4.3 Current Legal Issues in Location Technologies

The extent to which privacy may be expected in public spaces in relation to emerging location technologies has not yet been determined, as evidenced by conflicting findings related to the use of GPS devices to track persons of interest. Two appeals courts recently addressed the issue and issued conflicting findings as to whether a warrant is required. In *State of Wisconsin v. Sveum* (2009) an appellate court addressed the issue of whether placing a GPS tracking device on a car implicates the Fourth Amendment. The holding stated that, "…neither a search nor a seizure occurs when the police use a GPS device to track a vehicle while it is visible to the general public." The court cited *United States v. Knotts* and *United States v. Karo* in its decision, and argued that the GPS device as used provided only such information as would normally be available through visual tracking in a public place (i.e., the location of the tracked vehicle). Of note, however, is the Court's following statement:

> We are more than a little troubled by the conclusion that no Fourth Amendment search or seizure occurs when police use a GPS or similar device as they have here. So far as we can tell, existing law does not limit the government's use of tracking devices to investigations of legitimate criminal suspects. If there is no Fourth Amendment search or seizure, police are seemingly free to secretly track anyone's public movements with a GPS device…
>
> We are also concerned about the private use of GPS surveillance devices. As the Seventh Circuit and a recent New York Times article indicate, GPS technology is available at low cost to the general public…Although there are obviously legitimate private uses, such as a trucking company monitoring the location of its trucks, there are also many private uses that most reasonable people would agree should be prohibited…
>
> Consequently, we urge the legislature to explore imposing limitations on the use of GPS and similar devices by both government and private actors. Such limitations would appear to be consistent with limitations the legislature has placed on electronic intercepts of communications. (Wisconsin v. Sveum, 2009)

In a conflicting finding that nonetheless reflects the concerns stated in the *Sveum* case, the New York Court of Appeals found in *The People v. Weaver* that the use of a GPS unit placed without warrant on the defendant's vehicle did constitute a violation of his Fourth Amendment rights. Citing the unprecedented ability to gather data via GPS technology, the Court stated,

> One need only consider what the police may learn, practically effortlessly, from planting a single [GPS] device. The whole of a person's progress through the world, into both public and private spatial spheres, can be charted and recorded over lengthy periods…Disclosed in the data retrieved from the transmitting unit, nearly instantaneously…will be trips the indisputably private nature of which takes little imagination to conjure: trips to the psychiatrist, the plastic surgeon, the abortion clinic, the AIDS treatment center, the strip club, the criminal defense attorney, the by-the-hour motel, the union meeting, the mosque, synagogue or church, the gay bar and on and on. What the technology yields and records with breathtaking quality and quantity, is a highly detailed profile, not simply of where we go, but by easy inference, of our associations -- political, religious, amicable and amorous, to name only a few -- and of the pattern of our professional and avocational pursuits (*People v. Weaver,* 2009).

The Court's decision appealed not only to current law enforcement use of GPS devices, but also cited potential future uses if no legal precedent was set at the current juncture. The differing findings, but similar concerns, cited in both the *Weaver* and *Sveum* lend credence to the idea that privacy is a relevant and timely issue both in the courts and on the roadway. Combined with Slobogin's findings, it is evident that the expectations of the public and the courts in regard to privacy in the public sphere are changing based on technological advances and the potential ramifications of these advances.

3.5 Privacy Issues in ITS

Once the relevance of the issue of privacy in public places has been determined, it is necessary to focus on the more targeted issue of privacy in relation to ITS. Unlike Slobogin, who focused

on the issue of privacy in relation to surveillance, Briggs and Walton (2000) focus most heavily on issues relevant to the retention and management of the data that result from ITS surveillance applications, identifying the following questions as being particularly relevant:

- Do the ITS applications have the ability to collect personal information, and what is the extent of this information?
- Does the traveler know that his or her data is being collected and does he or she have any control over the collection?
- How long and for what purposes will collected data be stored and used?

Briggs and Walton argue that these questions must be addressed within the context of a number of issues, including anonymous data collection, visual images, secondary uses of data, law enforcement access to data, litigation involving data, data creep and opt-in versus opt-out conditions. These issues may be subject to different degrees of acceptance by different stakeholders, including commercial freight carriers and shippers and the general public.

The authors provide the following two general characteristics for an ITS application to be subject to privacy concerns:

- It enables the identification of an individual vehicle or occupant.
- It collects and stores proprietary information about a vehicle or individual (Briggs and Walton, 2000).

The authors identify the following applications, among others, as meeting these conditions: border crossing systems for commercial vehicles, vehicle probe applications (which track individual vehicles along their travel trajectory to measure such things as traffic speed), video surveillance applications, smartcard applications, and incident or accident logs (Briggs and Walton, 2000). These applications and the technologies that underlie them, such as cellular

phone geolocation, automatic vehicle identification and video license plate reading, will open up a heretofore unavailable level of data on individual drivers and will require close attention to balancing the desire for the data and the need to protect the privacy of those on whom data is collected.

The authors, like Slobogin, provide an overview of laws and regulations relevant to the issue of privacy in ITS, in particular those related to telecommunications. In addition to the overarching applicability of the Fourth Amendment, the authors cite the Electronic Communications Privacy Act of 1986 and 1994 Communications Assistance for Law Enforcement Act, the Telecommunications Act of 1996, the Privacy Act of 1974 and Freedom of Information Act, and general Fair Information and Privacy. (Briggs and Walton, 2000) The authors point out that while each of these acts has some relevance to the issue of privacy in ITS, there is no single regulatory act that specifically addresses the issue, and that these acts address the issue only within the public sector.

Briggs and Walton use a survey of electronic toll collection (ETC) agencies to further explore how companies providing services in the mobile environment currently address privacy of customer data, including collection procedures and secondary uses. ETC operators, a mixture of public and private organizations, are generally subject to two overarching privacy concerns: the potential for malicious outside attackers to obtain data and use it for unintended purposes, and the potential for secondary use of data for purposes such as speed tracking and marketing (Briggs and Walton, 2000). To assess how the ETC operators address these concerns, Briggs and Walton first examined and compared the applications and customer contracts of 12 United

States ETC agencies. Of these 12 agencies, nine additionally participated in a survey conducted for the study. In the survey, Briggs and Walton asked participants to report the following:

- what information is collected from customers,
- what secondary uses of the data are allowed,
- the length of time for which user data is stored, and
- methods of information security employed.

The survey produced a number of interesting results regarding the mechanics of data collection, maintenance and use; however, a number of questions about the survey exist, including a relatively small sample size and the experiences of public versus private operators. Because of the variety and number of organizations involved with the implementation of ITS applications, future work in this area may need to provide a more extensive approach to surveying, ensuring that the experiences reported are representative of the types of operators involved. This study provides the groundwork for the more extensive content analysis of privacy policies explored in Chapter 5.

Keeping the above limitations in mind, a number of Walton and Briggs's findings prove relevant to the issue at hand. First, the authors found that while agencies do ensure that they are meeting relevant guidelines and following applicable regulations, they tend to use public perception as a guide for their practices. Additionally, they find that provision of customer choices and taking a voluntary (specifically an opt-in) approach to these programs are critical to gaining acceptance. Finally, they provide a number of recommendations for data collecting agencies and organizations, including recommendations regarding data collection, retention, sharing and use, and the need to build privacy protection into organizational structures and regulations. These recommendations, along with the depth of information covered over the

report, provide an excellent grounding for the issues that must be addressed to ensure the protection of traveler privacy within ITS, and more broadly, they more or less respond to the constitutional concerns Slobogin raised above. Because the analysis addresses both regulatory and technical issues within the context of ITS, it also provides a good bridge for planners who must work with both to ensure that enough useful data is collected to justify the implementation of ITS applications and that the privacy of the users is protected. The issue of user choice, in particular, provides an opening to discuss technological approaches to vehicle-based privacy that have been suggested, along with some of the limitations that have been identified.

The next section brings these recommendations into clearer focus by providing an overview of how regulatory guidance has been used to address privacy in the realms of health and credit reporting. The examination of the HIPAA and FCRA within the setting of the Privacy Act of 1964 will establish the groundwork for determining what elements of privacy may be reasonably addressed within policies directed towards protection of private information within the locational setting.

3.6 Approaches to Privacy Preservation in the Regulatory Context

3.6.1 Introduction

As technological changes allow for ever more intrusions into one's "private" life, the public's expectations related to privacy have also shifted. As these shifts occur, it has often been difficult for regulation to keep pace. At a 2008 conference exploring the question of whether law and ethics are able to keep pace with science and technology, it was discussed that, "As

developments in science and technology accelerate (the number of important scientific discoveries doubles every 20 years, and the number of patent applications filed increases 5 percent each year), laws that regulate them are being bogged down (Magruder, 2008)." The same concerns emerge with more overarching issues such as privacy. The next section of the dissertation will explore how the Health Insurance Portability and Accountability Act of 1996 (HIPAA) and the Fair Credit Reporting Act (FCRA) of 1970, which was amended in 2003 by the Fair and Accurate Credit Transactions Act, have been developed to keep pace with emerging technologies, in the hope that they may further describe current regulatory approaches toward privacy protection.

3.6.2 The Privacy Act of 1974

3.6.2.1 Background

Both HIPAA and the FCRA are framed on the groundwork of The Privacy Act of 1974, which addressed the following four policy objectives:

1. To restrict disclosure of personally identifiable records maintained by agencies.
2. To grant individuals increased rights of access to agency records maintained on themselves.
3. To grant individuals the right to seek amendment of agency records maintained on themselves upon a showing that the records are not accurate, relevant, timely, or complete.
4. To establish a code of "fair information practices," which requires agencies to comply with statutory norms for collection, maintenance, and dissemination of records.

The Act was amended by the Computer Matching and Privacy Act of 1988, and the Computer Matching and Privacy Protection Amendments of 1990 further clarified the due process

provisions. This guidance has provided the underlying structure of privacy as understood in the federal environment, and has had repercussions across other areas.

3.6.2.2 Health Insurance Portability and Accountability Act

The privacy of personal medical information has long been recognized as central to patient-doctor relationships. In a time of increasing amounts of data collection and record creation, managing the outward flow of such information has proven a critical issue in maintaining confidentiality and ensuring that proper privacy norms are respected. Current medical practice requires that the patient disclose a large amount of personal information for adequate and accurate care. In order that care may be continued, that patients and physicians may be compensated through insurance agencies, that legal action may be taken if necessary, and for research needs (among other reasons), careful and often detailed records of patient histories are created and may be maintained for lengthy periods of time. While these records are generally considered confidential, they may be shared for purposes of insurance payments, health of others, and various other reasons. In response to these and other issues, the Health Insurance Portability and Accountability Act of 1996 (HIPAA) was enacted.

According to the U.S. Department of Health and Human Services,

> The HIPAA Privacy Rule provides federal protections for personal health information held by covered entities and gives patients an array of rights with respect to that information. At the same time, the Privacy Rule is balanced so that it permits the disclosure of personal health information needed for patient care and other important purposes. (DHHS, accessed February 2011)

The Privacy Rule applies to health plans, health care clearinghouses, and to health care providers that transmit health information in electronic form (“covered entities”). Health care information protected by the Privacy Rule includes all, “individually identifiable health

information held or transmitted by a covered entity or its business associate, in any form or media, whether electronic, paper, or oral. The Privacy Rule calls this information "protected health information (PHI)." (HHS, 2003)" Covered information includes data that may, or may be able to, identify the individual relating to the following:

- the individual's past, present or future physical or mental health or condition,
- the provision of health care to the individual, or
- the past, present, or future payment for the provision of health care to the individual.

If health information has been de-identified (via formal determination by a qualified statistician, or by removal of information that may be used to identify the individual or his or her relatives or others), there are no restrictions on its' use or disclosure; however, a covered entity may not otherwise use or disclose health information covered under this rule except as permitted or required by this Rule, or as authorized by the subject in writing. Protected health information is required to be disclosed in the following situations: to a subject if he or she has requested access, or to HHS if it is undertaking a compliance investigation, review or enforcement action. Written authorization must be obtained for disclosure or use of protected information if it is to be used for reasons other than those authorized by the Privacy Rule.

A key principle of the Privacy Rule is that of "minimum necessary" use and disclosure, which indicates that covered entities must make reasonable efforts to request, use and disclose only the minimum amount of protected health information necessary for its' purposes. In addition, covered entities must restrict access to protected health information internally via policies and procedures that identify who may have access to what information and for what purposes, and must also establish policies and procedures relevant to disclosures of protected health information. The Privacy Rule also requires that covered entities must provide a notice

of their privacy practices, including the following: “The notice must state the covered entity’s duties to protect privacy, provide a notice of privacy practices, and abide by the terms of the current notice. The notice must describe individuals’ rights, including the right to complain to HHS and to the covered entity if they believe their privacy rights have been violated. The notice must include a point of contact for further information and for making complaints to the covered entity. Covered entities must act in accordance with their notices. (HHS, 2003)” This requirement addresses the issue of adequate protection of the privacy rights of the patient and, in addition, instituted practices related to the following patient rights:

- The right to review and obtain a copy of his or her own protected health information as gathered by the covered entity, with some exceptions;
- The right to amend these records if incorrect or incomplete;
- The right to obtain an accounting of disclosures of these records to covered entities or their business associates;
- The right to request restriction of use or disclosure of protected health information, though this request may be denied by the covered entity; and
- The right to request confidential communications via contact by alternative means or locations.

These policies significantly increased the rights of the patient relative to prior practices.

In addition to the Privacy Rule, a Security Rule was also developed that, “…specifies a series of administrative, physical, and technical safeguards for covered entities to use to assure the confidentiality, integrity, and availability of electronic protected health information. (HHS, accessed February, 2011)” While much of the rule is reflective of the Privacy Rule outlined above in terms of covered entities, administrative procedures, and other standard regulations, the focus on setting standards for technical safeguards of electronic data is in direct response to advances that have been made in electronic data collection, storage, use and sharing. The

Security Rule provides guidance related to risk determination and access to information, but does not specify what technological methods should be used. In this way, the Security Rule addresses concerns about safeguarding collected data in such a way that technological innovations in security may be implemented.

Broader applications of such protections may be seen as applicable to the transportation industry, particularly in regard to requirements for notice, access, sharing, and security of collected data. While some regulatory aspects of application differ from the location services industry (particularly in regard to the public-private nature of such industry agents), the overview of the provisions of the Privacy and Security Rules of HIPAA provide scope for evaluation of existing practices within transportation, as well as recommendations for necessary components of any proposed regulation.

3.6.2.3 Fair Credit Reporting

A second area that may inform analysis of the privacy and protection of consumer data in the location industry is that of credit reporting. The federal Fair Credit Reporting Act (FCRA) was originally passed in 1970, with amendments made in the Consumer Credit Reporting Reform Act (CCRRA) of 1996 and the Fair and Accurate Credit Transactions Act (FACTA) of 2003. According to EPIC.org (accessed February, 2011), “The FCRA was the first federal law to regulate the use of personal information by private businesses.” The fact that the FCRA regulated private businesses makes it especially applicable to the question of regulation of the privacy of personal data in the transportation industry, as transportation involves both public and private service providers.

According to McEneney and Kaufmann (2004), the FCRA requires consumer reporting agencies to maintain procedures to assure accuracy of information, exclude obsolete consumer information, limit disclosure of reports, and allow consumers to review and correct their information. These requirements are reflective of those identified in the above-reviewed HIPAA Privacy Rule. These rights, again, are reflective of the recognition for the need for stronger consumer protections, particularly with regard to notice, access, and enforcement – all hallmarks of the FTC Fair Information and Privacy Principles described in the next chapter.

While somewhat less far-reaching than the HIPAA Privacy and Security Rules, the FCRA provides scope for ascertaining what expectations consumers may have in interactions with private agencies that have access to their personal data. Understanding such regulation and the rights it affords to consumers will be a critical component in framing recommendations for the treatment of privately-collected location and mobility data.

3.6.2.4 Conclusion

This section has examined key United States privacy acts in the context of evaluating current standards for expectations of privacy by consumers. These acts, along with the FTC Fair Information and Privacy Principles, form the underlying structure upon which policies and procedures relevant to privacy in the locational environment may be built. Key issues of note include the focus on notice, access, control, and enforcement, as well as the regulation of both public and private service providers. Such existing policies as these provide scope for arguing for a broader-spectrum policies that would specifically address issues of concern directly targeted to ITS and LBS technologies and services, as they provide acknowledgement of targeted areas of concern and applicable methods for address. While both HIPAA and the FCRA

have been somewhat limited in their effectiveness, they do provide scope for understanding methods by which privacy regulations have been established in the governmental framework. The next chapter will address current practices regarding the provision of individual privacy policies in the location environment within the context of consumer expectations as framed by these policies.

3.7 Technological Approaches to Privacy Protection

Technological methods of privacy protection may be implemented in part to ensure that the privacy protections guaranteed in policies are met. These technological methods of privacy protection may be either security enhancing (such as Secure Sockets Layer (SSL)) or specifically designed for privacy protection (such as pseudonyms and cryptographic keys, described below). In some cases, the privacy policies of LBS agencies refer to these technologies, especially SSL, in order to assure users that measures are being taken to protect their private information. While use of and research on privacy enhancing technologies in mobile applications have been growing in recent years, it is still often addressed at the back end of development instead of being included in the development process. Blumberg and Eckersley (2009) state that privacy needs to be built in to systems as part of the original design, while noting that the easiest solution would be to not collect data at all

Addressing the protection of personally identifying information in the mobile environment has been approached from a number of different technological angles. As noted above, the ability to gather data from mobile sources allows for placement of vehicles or persons in both space and time. In reaction to this, a number of technological approaches to

privacy protection have used either spatial or temporal "cloaking" in addition to other methods of making position and temporal data fuzzy. Another issue is that of the identification of the individual vehicle, which may be necessary to ensure the validity of the user and not a malicious attacker, but which may leave a large amount of personal data open to violation. Methods such as the separation of communication and authentication steps in data collection and pseudonyms have been proposed to address this problem. The following sections will outline these and other common suggestions for protecting the privacy of the traveler from a technological standpoint.

3.7.1 Protecting Identification Data via Pseudonyms and Cryptography

The issue of identifiers is perhaps one of the most critical in mobile privacy. If vehicles are validated in a way that links to their real-world identities, it may be possible to gain a large amount of data on specific travelers and their habits, and use that to mine other data sources. A number of researchers (including Kamat, *et al.* (2008), Tang, *et al.* (2008), and Dötzer (2005)) have proposed that utilization of pseudonyms validated by trusted third parties may be one effective method of protecting private data while not compromising security needs (such as protection from Sibyl attacks). One limitation to such an approach is that if the pseudonym is static, it may still be possible to accurately identify the unique vehicle by linking to physical surveillance or other travel data. To address this issue, some studies have proposed the use of multiple pseudonyms. For example, Dötzer (2005) recommends that pseudonyms be changed periodically to reduce the potential for identification.

A different type of pseudonym proposal is set forth by Kamat, *et al*. (2008). The authors suggest a pseudonym generation system that, "allows for user-controlled levels of privacy

(pseudonymity) and yet provides non-repudiation because only a Trusted Arbiter (TA) can reconstruct the true identity of a vehicle from its pseudonym." In the proposed configuration, the authors allow a TA (located at a base station) to provide new pseudonyms at the request of the user, thus giving the user the ability to personally decide the level of privacy he or she desires based upon current actions and network characteristics. While this approach may address some of the limitations experienced in a sparse network (as the user may choose to change pseudonyms when in a more densely populated link), it is still subject to a number of concerns, including unintentional privacy loss due to a lack of understanding or knowledge of the system. While this may, to some extent, reflect privacy preferences (i.e., those persons who are least concerned about locational privacy will change their pseudonyms less frequently than those who are highly concerned), it may also be a function of ignorance as to the potential privacy impacts of less frequent changes. Secondly, if a user is not adequately familiar with the system, he or she may make decisions relevant to pseudonym changes that will be identifiable and traceable to certain sensitive travel activities. This concern is particularly troubling given that many travelers tend to gravitate towards habitual routes and activity points.

A second approach that has been suggested by researchers focuses on group-based authentication and identification using cryptography. Under such a scheme, individual vehicles are not identified, but are rather able to send messages from within a group. Guo, *et al.* (2007) have proposed a group-based signature security framework that they argue provides both security and privacy benefits. In the system, group members are assigned a small number of secret keys/public group key pairs (stored in a tamper-resistant on-board module), and a group manager (or managers) is assigned an additional key (*gmsk*) in order that these keys may be

validated and traced to an individual group member if needed to provide accountability or investigate improper use. While this system addresses several limitations of the above proposals (such as identifiability in a sparse network), the authors do note a number of other limitations to the proposal, such as loss of efficiency and the possibility of identification of vehicles traveling outside of their region. The authors suggest maintaining a hierarchy of group relationships (for example, national, state, region and role) that may be changed based on travel realities, but note that this approach may also have limitations in terms of efficiency. Additionally, due to the responsibilities of the group manager, it will be necessary to ensure that that person is trustworthy. The authors propose that no individual entity be provided with *gmsk*, but rather that it be portioned among a set of individuals who would need to collude in order to open a message to identify the individual sender. Finally, key storage will need to be thoroughly addressed, as onboard systems may be vulnerable to attack or manipulation.

Raya and Hubaux (2007) propose still a different method of responding to potential privacy threats by suggesting a protocol based on digital signatures under a public key infrastructure (PKI). In this system, each vehicle is assigned a limited number of public/private key pairs and key/certificate sets by a Certification Authority (CA) such as the vehicle manufacturer or a governmental entity. Each message sent from the vehicle is appended with a key certificate and a digital signature verifying its authenticity. Privacy in this system is established via the frequent change of the key pair and a relatively short lifespan for certificates, thus making identification of an individual vehicle unlikely. The authors suggest that anonymous keys should be changed only after a certain number of messages have been sent, though they indicate that this may result in key changes every minute or so of driving

time. The greatest liability of this proposal has to do with the overhead required in terms of communication and computation. Guo, *et al.* (2007) also note that the inclusion of the message, the digital signature and the certificate will greatly increase the size of any communication, thus perhaps straining the efficiency of the system. The authors propose that Public Key Cryptosystem (PKCS) with a compact signature size be used to address these issues, but do not provide evidence that this will be adequate to ensure overhead efficiency. As above, a secondary concern is that of the security hardware. Particularly because of the large number of key pairs and certificate sets stored onboard the vehicle, it will be critical to address issues related to potential tampering or modification of the storage.

3.7.2 Protection of Spatial and Temporal Data

In addition to issues related to identification information, the ability to precisely identify the locations of individual travelers at specific times is of potential concern. As noted in Section 3.4.3, the potential ramifications of collecting such detailed data may include many unintended consequences, such as identifying political leanings, personal associations, and many other identifying concerns. In response to these concerns, a number of authors have proposed methods of blurring the spatio-temporal data gathered in the mobile network. For example, Ruan, *et al.* (2007) suggest the use of a secure privacy-preserving hierarchical location service (SPPHLS) in mobile ad hoc networks (MANETs). In this system, the area occupied by the network is divided into a hierarchy of regions, which are aggregated to form a region on the next higher level. For any node *A*, one cell in each level of the hierarchy is selected by a hash function as the responsible cell. As *A* moves through the area covered by the network, it updates its responsible cells with its current position. When another node *B* needs *A*'s position,

it uses the hash function to determine these cells which may be responsible for *A*. *B* then queries those cells in the order of the hierarchy, starting with the lowest level region. On the first level that contains both nodes *A* and *B* the query will arrive at a responsible cell of *A* where it can be answered.(Ruan, *et al.* 2007) In this manner, general information about the node's location can be obtained, but the specific location is hidden to the user's preference.

A second method of preserving the privacy of spatial and temporal information is the use of spatial and temporal cloaking, as suggested by Gidófalvi, *et al.* (2007) Location-Based Services (LBS) depend in large part on context information for success, and data mining is often used in this framework to determine patterns of behavior and location. However, there is potential for conflict, as data mining methods tend to best use precise information, while users likely have some desire for privacy protection. The framework proposed in this paper allows user location data to be anonymized, while still allowing for interesting pattern discovery via the use of spatial cloaking, which uses anonymization rectangles to aggregate and "hide" the specific location of the user. The system may be developed in such a way that the user may define his or her degree of privacy protection as based upon the size and pattern of the rectangles. (Gidófalvi, *et al.*, 2007) Hoh, *et al.* (2006) propose the use of virtual trip lines (VTLs) for privacy protection. A VTL is a geographic marker stored in a client (such as a GPS-enabled cell phone), which triggers a position and speed update whenever a probe vehicle passes. According to the authors, privacy is preserved by updating in space rather than time, as areas of high privacy concern (such as freeway on-ramps or red light districts) may be avoided.

3.7.3 Conclusions on Technological Approaches to Privacy Protection

The methods identified above review only a small portion of available technological approaches to privacy protection in the mobile environment. While the methods identified, including pseudonyms, cryptography, and spatio-temporal blurring, are not comprehensive, they do reveal a number of solutions that have been proposed for some of the most pressing problems in mobile privacy. It is necessary to observe, however, that the issues of trust and security are also important components of privacy, thus it will be necessary to ensure the managers of these systems are trusted by the users.

3.8 Privacy Policies in the Mobile Environment

In February 2010, at a joint hearing on privacy in location-based services (LBS) held by the Congressional Committee on Energy and Commerce's Subcommittees on Commerce, Trade, and Consumer Protection and Communications, Technology, and the Internet, Congressman Bobby Rush stated the following:

> Yesterday there was Facebook and in the not too distant future we will be encountering something more akin to a placebook. Location-based services and the applications that ride on these services utilize a number of different tracking technologies which can make it easy to track the whereabouts of an estimated 100 million individuals around the world. By the year 2013, it is estimated that the precise whereabouts of over 800 million individuals will be readily discernible at any given moment in time.

In the hearing, Congressman Stearns (R – FL) expanded upon privacy concerns when he stated that, "…wireless carriers are generally prohibited from using location-based information for commercial purposes without the express prior consent of the consumer. However, application providers are subject to no such requirement even though their applications are being

downloaded on the devices of wireless carriers. (2010)" The lack of consistency between policies regarding the treatment of private information is particularly germane to the discussion of LBS. While users may be unaware of the distinction, it is possible that many will assume that such protections are applicable both to the information shared with wireless providers and that shared with LBS application providers. As it stands, however, providers of location aware applications are not specifically subject to these regulations, but rather are subject to the recommendations and guidelines such as those propagated by the FTC, ITS America, and CTIA – The Wireless Association, though many do many do include a privacy policy of some sort in relation either to (a) use of their product, or (b) website access related to the product.

Such lack of specific oversight is particularly relevant to discussions regarding the use of data collected by these agencies in the public realm. For example, John Morris (2010) stated in the February hearings that, "A lack of clear rules about law enforcement access to location information held by service providers has left location technology without sound legal footing." Existing regulations regarding access to data lag in their relation to use of location data collected by service providers for purposes of law enforcement, a gap that is shared by other public agencies. In short, while the collected data may be of use to public agencies, no clear direction exists on the legality of access or the consistency of collection and sharing. Such issues will become of increasing importance as public agencies seek to expand their knowledge base within budgetary constraints, as in some cases it may be more fiscally possible to obtain or buy existing data from private companies rather than participating in collection activities. Technological methods of privacy protection, described below, may be of particular interest in

this realm, as such methods will likely be implemented at the level of the collecting agency and their efficacy may need to be evaluated before data are used by other interested agencies.

The use of privacy principles and policies provide the opportunity for companies and organizations to indicate how private information will be treated, and what the consumer may expect. While a number of US laws, such as the aforementioned Privacy Act of 1974 and Communications Act, are broadly applicable to some providers or carriers of ITS and LBS services, in some cases the guidelines or principles of organizations such as CTIA – The Wireless Association and ITS America, the Vehicle Infrastructure Integration (VII) initiative (now IntelliDrive) consortium, or the FTC may be more relevant. These, often voluntary, guidelines provide an indication both of the information these organizations feel it is necessary for their members to share with consumers, and how collected data should be treated. Because federal regulations do not necessarily apply to all providers of ITS and LBS services, it is appropriate at this time to evaluate these industry approaches to privacy to determine how well companies within these industries are meeting the recommended guidelines. As noted above, the FTC's "Fair Information Practice Principles" provide a fairly comprehensive approach to treatment of private information. Because the CTIA, ITSA, and VII guidelines all more or less reflect some or all of the FTC's principles, these will be used as a guideline for evaluation.

As noted above, Notice/Awareness comprises the underlying framework of the FTC's privacy guidelines, forming the basis for the remaining principles. Notification as defined by the FTC consists not only of informing the customer that data are being collected, but also who is collecting the data, to what uses it will be put, and with whom the data may be shared. These notifications are particularly essential in regards to data being collected in the mobile

environment, or in relation to location information, as users may not fully comprehend either the degree to which data are collected, or the purposes for which they may be used. Each of the overarching privacy policies under consideration here acknowledges that notice is both a critical component in the collection and dissemination of data, and a potentially changing and challenging step.

Also of note is that while the CTIA observes that notice is the hallmark of its recommended practices, they also focus on the need for clarity in privacy policies. In a 2004 study of online privacy policies, Jensen and Potts (2004) found that, "…only 6% of policies are readable by the most vulnerable 28.3% of the population, and…13% of policies were only readable by people with a post-graduate education…" Such findings indicate that it is not only provision of a privacy notification that is important, but also the ability of the user to comprehend the policy. CTIA (2010) acknowledges this necessity when it states that, "Any notice must be provided in plain language and be understandable." While both the ITS America and VII Privacy Policies Framework acknowledge the necessity of notifying consumers that their data will be collected, neither explicitly address the question of ensuring that such notice is understandable by the consumer.

The second principle included in the FTC's policy is that of choice/consent, which builds upon the principle of notification, as it assumes that consumers have been provided information regarding which data are being collected, but extends it to apply to treatment of data by third parties. Also of note here is the question of opt-in versus opt-out conditions, which are differentiated by the FTC (2007) dependent upon if affirmative steps are required by the user to allow or disallow collection of data. While some entities and organizations, such as

the Direct Marketing Association, favor an opt-out scenario, others, such as Godin and Don (1998), indicate that an opt-in scenario may be favored by consumers. Again, one issue at question is whether, in an opt-out scenario, consumers are aware (1) of the amount of data collected, (2) the third-party entities with which that data may be shared, (3) the privacy policies of those third parties, and (4) methods by which one may opt out of that sharing. Such knowledge requires not only notification of policies, but also comprehension of the policy and ability to determine methods of opting out.

In its Privacy Guidelines, CTIA explicitly addresses the issue of choice/consent, and includes language pertaining to opt-in and opt-out conditions. The Guidelines (2010) indicate both that use of opt-out mechanisms would, "ordinarily be insufficient to express user consent", and that adequate consent should be consistent with the notice principle. ITSA addresses choice and control in its "Individual Centered" principle, which states that ITS must respect the user's interest in privacy and use of information by allowing for disclosure to the individual and the provision of choice in data collection. Consent is further addressed in the eighth principle of "Commercial or Other Secondary Use". While ITSA indicates that informed consent should be obtained, the Principles do not provide guidance on how this should occur, thus placing the burden on the providers of ITS services.

The VII Framework (2007) addresses the issue in its second principle, "Information Purposes," which states that, "A personal information user should…inform a personal information subject about the purposes for which personal information will be collected, used or disclosed before collecting personal information from that subject so that the personal information subject can decide whether or not to agree to use of their personal information for

those purposes…" In this framework, choice and notification are, again, co-dependent upon one another, as the user's ability to make an informed choice about the provision and sharing of his or her personal information is dependent upon having adequate information about the purposes for which information is gathered. Control, however, is largely left to the discretion of the system administrators, who are responsible for enforcing adequate control over collected data. Here, control is a function of choice on the part of the consumer, as the framework primarily addresses the consumer's control over his or her collected data at the point of choosing whether or not to share data at the outset.

The third FTC principle, Access/Participation, is concerned with an individual's ability to access his or her data and to contest the data's completeness and accuracy. This principle acknowledges both that consumers may have interest in the data that are collected about them, and that this data may be incorrect. Additionally, the FTC guidelines focus on ease of access for consumers, stating that, "To be meaningful, access must encompass timely and inexpensive access to data, a simple means for contesting inaccurate or incomplete data, a mechanism by which the data collector can verify the information, and the means by which corrections and/or consumer objections can be added to the data file and sent to all data recipients (FTC, 2007)." Such guidance establishes that the onus of providing useful and understandable access to information rests on the backs of the providing entities.

The other three privacy policies evaluated here do an inconsistent job of addressing the concerns addressed under this principle. The CTIA, for example, does not directly address access and participation as described by the FTC. The guidelines do indicate that LBS users should be informed of privacy options and controls provided by the LBS providers, but they

neither explicitly address the need for users to have access to their data for the purposes of ensuring correctness, nor establish guidance for allowing corrections. The VII Privacy Principles, on the other hand, provide a "Participation Principle," which addresses the need for service providers to provide users with the ability to access collected data, correct that data, object to improper use, and remain anonymous; however, the principle also states that, "…each personal information subject should be expected to protect his or her own privacy (VII, 2007)." ITSA (2001) acknowledges that ITS systems, "…will be built in a manner 'visible' to individuals," which indicates that individuals should have access to collected information and knowledge as to its uses, but does not address the need to allow consumers the ability to review the data that have been collected and ensure its validity. In short, each of the ITS and LBS affiliated entities include some acknowledgement of the principles of access and participation, but none seem to fully address the needs embraced by these principles in the FTC documentation.

The fourth FTC principle, integrity/security, establishes that collected data should be "accurate and secure (FTC, 2007)." This principle focuses on the interaction between system administrators and data rather than on the interaction between users and providers, noting that collectors of data should take steps to ensure data integrity and security, including using only reputable sources, cross-referencing, providing consumer access, limiting access to data, and ensuring secure storage. This principle enjoys perhaps the most consistent acknowledgment across the industry policies evaluated here, as each includes language pertaining to ensuring the security of collected data, though, again, there is limited language pertaining to access. Assurances from each of the industry associations indicate the degree to which securing private information is regarded as critical to preserving privacy. It is also

possible, however, that the nature of the industries represented (i.e., heavily technologically savvy organizations) makes provision of security assurances a relatively simple element to fulfill.

The final element of the FTC's principles is "Enforcement/Redress". The FTC notes that this principle is particularly significant insofar as the remainder of the principles have little weight if there is no method for their enforcement. The FTC identifies the following three methods for enforcement:

- Self-regulation: Should include mechanisms for both ensuring compliance (enforcement) and to allow recourse by injured parties (redress).
- Private remedies: Such a scheme would create private rights of action for consumers if they experienced harm due to an entity's unfair information practices
- Government enforcement: Enforce of fair information practices by means of civil or criminal penalties.

CTIA's Guidelines, while they do not have the ability to impose government enforcement, do include language pertaining to the reporting of abuse and compliance with laws, an approach that is mirrored by the VII's Accountability Principle. ITSA's Principles address the question of enforcement in two ways, namely: 1) Compliance: ITS will comply with applicable laws governing the use of information and privacy, and 2) Oversight: Recommends that jurisdictions and companies operating or deploying ITS have an oversight mechanism to ensure compliance with privacy regulations (ITSA, 2001).

The uneven way in which the three industry associations evaluated here address privacy relative to the FTC's approach indicates that there is still a gap in how privacy is viewed in the mobile environment. The lack of clear and universally applicable regulations exacerbates this gap, as it places different restrictions on the treatment of privacy from service provider to

service carrier. Chapter 5 will provide a detailed overview of how these concerns emerge in a detailed review of ITS and LBS privacy policies.

3.9 Conclusion

This review and analysis of relevant law, policy and technology is intended to provide a general framework on which may hang a more targeted discussion of interest within the realm of privacy in ITS, as well as identifying some of the existing gaps in how privacy is currently treated in the locational environment. By establishing the legal basis for privacy in public, the need to address the issue is given due acknowledgement. Briggs and Walton's review of relevant issues within the larger framework sets forth the areas of primary concern and creates a greater awareness of the overall scope of the problem. The technological approaches suggested by Kamat, *et al.* (2008), Tang, *et al.*(2008), Guo, *et al.* (2007) and Dötzer (2005) provide examples of methods by which the ITS industry is addressing the protection of individual privacy, while the VII Privacy Policies Framework signals the initiation of more specific discussions at the confluence of government and industry. In general, the discussion of privacy within the context of ITS is in a nascent stage, with little being known of the actual willingness of people to accept ITS applications that may require great outlays of personal data.

The above analysis reveals the following specific implications and findings relevant to the research proposal:

1. Privacy may be viewed as a multidimensional construct, and locational privacy may be significantly from static privacy constructs. Spatiotemporal elements should be considered when determining locational privacy preferences.

2. Willingness to participate in ITS and LBS applications may be a function of general privacy preferences, though it will be necessary to determine tradeoffs between privacy and potential benefits related to expectations in the mobile environment.
3. A number of rules, regulations and policies exist relevant to privacy, and there are clear directives for privacy preservation. However, there appear to be no clear set of standards or understanding of how privacy functions in the mobile environment.
4. A number of privacy-preserving technologies and methodologies have been tested in the mobile environment, but these all have issues that must be addressed or studied more closely.

The work of this dissertation will focus on determining the willingness of travelers to trade data for congestion, safety or other benefits, as well as establishing better guidelines to ensure the correct collection, retention and use of collected data. The use of ITS surveillance techniques has the potential for great benefit, but it must be ensured that these benefits are not achieved at the expense of our ability to retain a sense of privacy while in public. For purposes of this dissertation, attempts will be made both to address the limitations identified as well as provide more of a linkage between the identified areas of interest. The key limitation of each of the studies reviewed above is that they are restricted to one area (law, technology, or policy) of interest within the realm of privacy in relation to ITS. This dissertation will attempt to provide a more holistic approach that addresses the areas of overlap between each.

CHAPTER 4: RESEARCH DESIGN AND METHODOLOGY

4.1 Introduction

As noted in Chapter 1, this dissertation is intended to analyze questions related to the topic of privacy in the context of ubiquitous mobility technology, in particular in relation to ITS and LBS applications. The main topics of interest relate to an analysis of the methods by which privacy may be protected in the mobile environment, in particular policy-oriented approaches, and the relationships between individuals' privacy preferences and their willingness to trade information related to their personal selves (such as name or address) and their mobility information (such as trip origins and destinations, time of travel, and route information). These questions will be evaluated in relationship to the research topics outlined in Chapter 1, by evaluating current privacy practices and consumer expectations and preferences related to privacy in the mobile environment.

This chapter is intended to present an overview of the data and methodologies which were used to study the research questions in the context of the current dissertation. As mentioned in Chapter 1, two primary sources of data were used, namely, an archive of existing privacy policies and a stated preference survey resulting from a primary data collection effort in order to ascertain preferences regarding willingness to trade data in the mobile environment. These data sets will first be described and contextualized. Next, the methods used to analyze the privacy policy archive and the survey will be described, in order to present the overall research design. Finally, expectations regarding the application of findings will be described, especially as they relate to steps that may be taken to preserve privacy in the future mobile environment.

4.2 Data Sources

Two primary data sources were compiled for purposes of the dissertation research. First, an archive of transportation related privacy policies was compiled as described below. Next, a stated preference survey was developed and administered to test the willingness to trade personal information for various transportation benefits that users perceive as receiving from mobility technologies, along with questions related to expected compensation, risk and overall expectations and practices regarding privacy in the mobile environment. Information related to survey development and dissemination will be described in detail below.

4.2.1 Policy Archive Development

A key strategy utilized by government agencies and private companies which offer mobility services to users is to require users to read and accept the terms and conditions in "privacy policies", which are legal documents that disclose a web site or a service's practices regarding the ways in which the site or service gathers, uses, discloses, and/or manages a customer's data. In order to evaluate current approaches to privacy protection in the mobile environment, an extensive archive of public and private privacy policies related to location and transportation services was developed. Though the dissertation focuses primarily on mobility technology, it is acknowledged that many transportation services are available online, thus policies of private and public agencies related to data collected online were also obtained. Policies collected, described in detail below, will be subject to content analysis, described further below.

Due to the rapid emergence of LBS applications and services, as well as increasing adoption of ITS technologies by public agencies, policies selected for analysis within this dissertation represent only a snapshot of possible data for analysis. To select privacy policies of

public agencies, the Federal Transit Administration's website was first used to identify public transit agencies. These agencies were then assessed to determine if (a) they provided privacy policies, and (b) if they currently use electronic transit (or fare) cards. Privacy policies of selected transit agencies were collected and sorted into general policies and policies related to use of electronic fare collection media. For electronic toll collection (ETC) policies, a web search was conducted to identify current implemented ETC systems in the United States. Of the ETC systems identified, seven had available privacy policies which were gathered and preprocessed according to the steps outlined in Section 4.3.1.4.

A directory of LBS companies available from Directions Magazine (2010) was next used to identify specific mobility information companies. The websites of all companies in that list were visited for three purposes: first, to identify the types of services and/or products offered by that company, which was used to exclude companies that did not meet our requirements; and second, to access privacy policies related to the services and/or products. A total of 48 companies were retained from that list for the rest of the analysis presented here. A majority of the companies are headquartered in the U.S, and all operate within the U.S. While the Directions Magazine listings may not be a fully accurate sampling frame of the universe of such companies operating within the U.S., it did provide a fairly comprehensive overview of companies of interest.

In total, 101 policies were collected and evaluated. The breakdown of company or agency type is:

- Private ITS or LBS Service Provider: 48
- Public Service Provider: 34
- Electronic Toll Collection (ETC): 7
- Electronic Transit Fare (ETF): 12

Public service providers include both transit providers and public Departments of Transportation (DOTs) that may be affiliated with transit providers or that host the privacy policy for transit providers. In the latter case, While the policies of ETC and ETF providers could also be construed to fall into the category of "public service providers," as they are guided by overarching policies applicable to the relevant Departments of Transportation or city or state governments, because of the addition of detail related, in particular, to financial considerations and additional personally identifying information they have been treated as separate categories for the purposes of this analysis. Additionally, overarching privacy policies for agencies that provide services such as trip-planning targeted at transit service users are generally accessible from a number of platforms (such as on paper, on a service website, or via mobile phones), while those geared for ETC and ETF use are generally included in terms of service or application agreements, which may be accessed only at the point of application. Because of these differences, ETC and ETF policies have been analyzed separately from overall public service policies.

The types of private ITS and LBS service provider agencies may, likewise, be categorized into a number of differing types of agencies. While differences in characteristics of policies of differing agency types will not be evaluated here (for more information, see Cottrill and Thakuriah, 2011), the types of companies whose policies were considered in the evaluation (a classification based on Heinonen and Pura (2006)) may be seen in Table III.

Table III: Types of Mobile Service Companies Considered in Content Analysis of Privacy Policies

Type of Mobile Services Companies	Brief Description	Examples of mobile services
Multiple-services technology companies	Multiple technology services, typically web-based	Typically search engines, email and other communication services; news; multiple other services; also web-based mapping, routing and location-based information
Location-Based Services companies	Services integrating a mobile device's position and other information to provide added value	mobile social network, friend/people finder, health and activity, green services, mobile commerce and location-based alerting, sports/off-road tracking, green services
Travel, Tourism and Entertainment companies	Travel and tourism agencies and entertainment/recreation	Local entertainment guides; location-based hotel, car-rental, restaurant, shopping guides; travel and ticketing
Traffic information and Navigation companies	Routing and navigation services	Navigation and routing solutions; real-time traffic information; real-time transportation sharing and multi-modal travel information
Map and Geospatial Infrastructure Companies	Developers of digital map databases	Proprietary and open-source editable map databases; mapping and mobile services
Location Communications Technologies Companies	Positioning system manufacturers	Location aware hardware; positioning systems
Geographic Analysis Services	Analysis for commercial locational decision-making	Geomarket analysis for retail; marketing solutions for geospatial customers
Security and Safe Location Management	Services using positioning for vehicle and personal security	Vehicle security and remote vehicle management; remote family location management

As seen here, there are a large number of private ITS and LBS agency types, serving both individuals and manufacturers of services targeted at individuals (such as provision of background maps for mobile navigation systems). While the privacy policies of these agencies may have differing impacts on the individual at the personal level, the ability of such service providers to collect both primary and secondary data (as third-parties to other services) makes the applicability of these policies to individuals in the locational environment of equal worth. Because many of the agencies that partner with other agencies to provide enhanced mobile services for consumers may have access to data collected by partnering agencies, the privacy policies of these agencies are relevant to consumer privacy concerns. Here, as with general

public agency policies described above, policies tend to be accessible online or per use, as opposed to being overtly stated primarily at time of initial application. Thus, the characteristics of policies of all agency types have been treated as broadly similar for purposes of the dissertation analysis.

4.2.2 Survey Development and Distribution

As seen in the research questions, one of the primary topics of interest to be addressed in this dissertation is that of the willingness of consumers to trade private data in exchange for assorted transportation benefits. As discussed in Section 4.4.1, there are numerous issues associated with stated preference surveys and contingent valuation (such as inconsistencies between a person's stated preference and how they will actually act (the "hypothetical bias"), and difficulties in assigning accurate values to goods that are generally not market priced); however, in the context of this dissertation, the need to evaluate traveler preferences regarding willingness to trade in reference to proposed or hypothesized applications of technology indicates that a stated preference survey including contingent valuation measures is the most reasonable method for ascertaining likely trade-off preferences. This section will describe both survey development and sampling.

4.2.2.1 Survey Design and Development

The survey design process consisted, first, of identifying the major topics of interest related to the willingness to trade privacy for perceived transportation benefits. It was determined that information was needed related to current use of general and transportation technologies, current consumer concerns related to privacy in the mobile environment, and expectations

related to benefits, in addition to general demographic information. In accordance with the research questions identified in Chapter 1, a questionnaire was developed to respond to these topics of interest. Table IV reflects how developed questions respond to items of interest (a copy of the final survey instrument may be found in Appendix 5).

Table IV: Relationships between research questions and survey items

Overall Research Question: Does a relationship exist between individuals' privacy preferences and their willingness to trade private information for use in ITS and LBS technologies and applications?		
Research Question	**Topic of Interest**	**Survey Item**
What expectations do travelers currently have regarding privacy protections afforded to them by transportation agencies and service providers?	Technologies used	See Questions 5.1, 5.2, 5.3 - Transportation and other technologies used
	Privacy expectations	See Questions 7.1 and 8.1 - Importance of information and risk perception
Do travelers currently demonstrate privacy concerns in the mobile environment?	Reading privacy policies	See Questions 6.1, 6.2, 6.3, and 12.1 - Reading and understanding of privacy policies
	Importance of protection of private information	See Question 10.1 - Privacy preferences
What are the component parts of locational privacy preferences, such as personal information (including name, address, or vehicle information) and travel information (such as origins, destinations, and other travel details)?	Privacy of personal information	See Question 11.1 - Compensation
	Privacy of travel data	See Question 11.1 - Compensation
How much influence, if any, does the desire to protect private information have on the willingness of persons to trade this information for various transportation benefits, including cost, efficiency, or safety benefits? Do any of the component parts of l	General privacy preferences	See Question 11.1 - Compensation
What types of expected benefits may most impact travelers' willingness to trade privacy components? Can these characteristics allow us to cluster potential users and determine ways to balance privacy preferences and application efficiency?	Willingness to trade data for efficiency benefits	See Question 11.2 - Compensation
	Willingness to trade data for economic benefits	See Questions 11.3 and 11.4 - Compensation
	Willingness to trade data for safety benefits	See Questions 11.5 and 11.6 - Compensation
	Willingness to trade data for third party benefits	See Question 11.7 - Compensation

Survey Item Development

Due to difficulties in identifying specific valuations of privacy on the part of consumers, Likert scales were used throughout the survey in order to query respondents on their relative degrees of importance and interest in privacy and benefit matters. While the use of Likert scales limits the ability to assign absolute values of importance or concern, it allows for relative measures of importance and concern to be determined. In general, five point scales were used, ranging from "Strongly Disagree" to "Strongly Agree" and "Not Important" to "Very Important." A seven point scale was used for purposes of determination of personality characteristics (Extraversion, Agreeableness, Conscientiousness, Emotional Stability, and Openness to Experiences) based on research by Gosling, *et al.* (2003). For one question, "How much, in general, would you have to be compensated to provide the following information to these [transportation] agencies?" respondents were asked to quantify the amount of compensation they would require to provide specific types of personal (name and address) or travel (origin, destination, travel time and route data) information to assorted transportation agencies. They were then asked to indicate whether this amount would increase, decrease, or stay the same under various scenarios evaluating the valuation of economics, safety, and efficiency. Though respondents were asked to provide specific valuation amounts, difficulties in ascertaining the value of various types of information has resulted in these valuations being treated as ordinal scales in keeping with evaluation of other questions.

Survey Pre-testing and Cognitive Interviewing

The survey was pre-tested by conducting cognitive interviews with five persons. In each case, participants were instructed to respond to survey questions online in the presence of the researcher. Participants were asked to inform the researcher of any difficulties in question phrasing or response while taking the survey, and were then asked a series of follow-up questions regarding their experience with the survey after completion. In response to these interviews, several changes were made to question phrasing (such as splitting the question, "How often do you notice or read privacy policies before using the following types of services" into two separate questions to test awareness and interest in the policy separately), response options (such as the addition of the option, "Would not sell" to questions regarding compensation required to be willing to sell personal information), as well as to question ordering. Participants were queried as to whether they felt that questions were "leading" or if they had any difficulties in determining what the questions were specifically referencing. If participants reported difficulties, adjustments were made to clarify those questions reported as unclear. The revised survey was submitted along with a Claim of Exemption to the University of Illinois, Chicago Institutional Review Board (IRB), and approved on November 9, 2010.

4.2.2.2 Survey Sampling

In order to reach a geographically diverse set of respondents, it was decided to conduct the survey online. The survey questionnaire was developed in the online survey program SurveyMonkey (http://www.surveymonkey.com/), and on November 15th, 2010, notifications

of the posting of a privacy survey were distributed in a number of social networks, including the following:

- Facebook
- TheChainlink.org (a Chicago bicycling social network)
- The email distribution list of the University of Illinois, Chicago's (UIC) Urban Transportation Center (UTC)
- The email distribution list of UIC's College of Urban Planning and Public Administration

Persons who received notification of the survey were additionally asked to forward the survey on to friends, colleagues, and others who might be willing to participate. In such a manner, a snowball survey was obtained. While the limitations of the survey methodology indicate that the findings are not statistically random, they do provide a basis for exploration of attitudes and expectations regarding privacy in the locational environment. It is not possible to determine the sampling rate for the survey due to a lack of information regarding the number of visitors to thechainlink.org, persons who viewed the Facebook announcement, persons who viewed the UIC email, or number of persons who received a survey announcement via email forwarding. Additionally, it is not possible to determine non-response rates, as the number of persons who received notification about the survey but did not participate can not be calculated. Persons who visited the survey but chose not to participate were asked to fill out an exit survey in order to provide general demographic information; however, no exit surveys were completed, thus no information was gathered on persons who chose not to participate in the survey.

Due to both the methods used to attract survey participants, as well as the nature of internet surveys, issues of selection bias and methodology must be addressed. Matsuo, *et al*. (2005) have identified the following issues in conducting online surveys:

- Sample representativeness: It is difficult to generate a sampling frame from which participants may be recruited, as there is no method for the researcher to access the entire population within the web environment.
- Response and non-response: As indicated above, it is difficult to identify response and non-response rate, as the researcher likely does not know the number of persons who potentially had access to the survey.
- Controlled testing conditions: The researcher has little or no control over the testing environment and the order in which participants respond to questions.

In the case of the current survey, as seen in Chapter 6, it is clear that the sample obtained is not representative of the population at large from a demographic standpoint, though it does reflect characteristics of first adopters of technology as discussed above. Another issue is that of nonprobabilistic sampling. Couper (2000) has identified the type of survey conducted here as an "unrestricted self-selected survey," as it relies upon open invitations on web portals and frequently visited web sites. Here, concerns may be raised regarding the representativeness of the sample in regard to the overall population. To address these concerns, the survey collected comprehensive socio-demographic data (including age, gender, income, education, and geographic location) in order that the sample population might be compared to the overall United States population. While, as noted, the surveyed population did not reflect general characteristics of the overall population, it was possible to determine initial findings regarding likely influences on privacy preferences in relation to experience with location and mobility technologies based on specific socio-demographic profiles.

4.3 Methodologies

In this section, the methods used to analyze the privacy policy archive and the survey data are discussed. The primary method used to analyze the privacy archive is content analysis

(described in Section 4.3.1.1). A series of statistical models were used to analyze the survey data (given in Section 4.4).

4.3.1 Analysis of Privacy Policies

The analysis of privacy policies was designed to respond to Research Question 1, identified in Chapter 1, "To what extent and by what methods should privacy in ITS and LBS be protected?". Privacy policies were chosen as the units of analysis here, as these are the primary methods by which agencies and organizations inform consumers as to how their personal data will be collected, stored, managed, shared and protected. Because there is no overarching privacy policy relevant to location data currently in place, the Federal Trade Commission's Fair Information and Privacy Principles, described in Chapter 3, were used to develop baseline categories of interest, as described below. This section will describe in greater detail the methodologies used to perform the content analysis of identified policies in this context.

4.3.1.1 Content Analysis

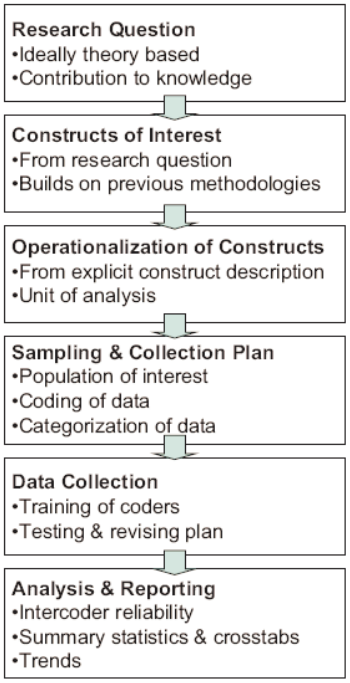
Content analysis is used here to evaluate, first, differences in the treatment of privacy notifications to consumers by public and private agencies and, second, to identify gaps in the information presented to consumers on the treatment and use of their travel data. In general, content analysis enables an objective description of the manifest or written content of textual material (Berelson, 1974), by determining the presence of themes, phrases or other related attributes of the material (Content Analysis, accessed June 24, 2009). Kassarjian (1977) notes that a measurement of the extent of emphasis or omission of any given analytic category distinguishes content analysis from ordinary critical reading. Content analysis begins with the

identification of the texts to be analyzed based upon a proposed research question followed by a sampling procedure to identify and select individual items of analyses. These items are then broken down into component units of analysis based upon the type of analysis desired. Kassarjian (1977) identifies the following possible units: words, themes, characters, items, and space-time measures. These units are then categorized and coded for purposes of analysis. For this analysis, words and phrases have been coded into categories of interest as identified by the FTC.

Subjectivity has been noted to be of concern in content analysis and Kassarjian (1977) stated that reproducibility of results by different analysts analyzing the same content is a key requirement of objectivity. For purposes of this dissertation, the issue of subjectivity has been addressed, in part, via the use of the automated content analysis software WordStat (described fully in Section 4.3.1.4). It should be noted, however, that a text will always involve multiple meanings and that there is always some degree of interpretation required on the part of the researcher (Graneheim and Lundman, 2003). Quantitative content analysis bypasses this difficulty to some degree by holding to the rigors of statistical analysis requirements, but it is difficult to argue that any content analysis will be completely free from the biases of the researcher and text coder.

While the interpretation may be somewhat subjective, a rational model underlies the process of textual coding. The approach described here follows that given by Holdford (2008), with steps given in the flowchart in Figure 3.

Figure 3: Steps in Content Analysis



The research question identified in Chapter 1 and restated above was determined based on current questions related to privacy in the locational environment raised in the literature described in Chapters 2 and 3. Constructs of interest were also developed from the literature review, and based on current questions raised in court cases and ongoing research. These constructs were then operationalized by reviewing the constructs and questions of interest in relation to the current treatment of privacy in regard to law and legal issues. These steps required identification of a broad underlying set of categories of interest in relation to privacy, for which the Federal Trade Commission's categories of interest related to privacy (Notice/Awareness, Choice/Consent, Access/Participation, Integrity/Security and Enforcement/Redress) were chosen. Here, other constructs of interest could have been identified (such as those identified in the VII Privacy Policies Framework, by ITS America, or by

CTIA); however, the analysis above noted that these policies were lacking in how well they addressed certain categories of interest. In acknowledgement of the more comprehensive set of constructs identified in the FTC policy, these categories were identified and established as categories of analysis for the identified policies.

The sampling and collection plan, described in detail in 4.2.1, first required identification of the populations of interest and a method for the collection of relevant policies. Coding categories were determined based on constructs and areas of interest identified in the FTC policy. Finally, policies were collected, coded and tested for consistency before performing the detailed statistical analysis. Such a process falls into a deductive approach as identified by Mayring (2000), which begins with the identification of the research question, from which one develops a theoretically based definition of analysis aspects, main categories and sub-categories. This step is followed by a theoretically based formulation of definitions, examples and coding rules, which are then collected into a coding agenda. Here, reliability checks are evident, as it is necessary to ensure that the categorical definitions developed accurately represent the meaning of the coded text. For qualitative content analysis, this is particularly necessity, as it may assist with minimization of subjective biases within the analysis. The use of manifest content as provided in the privacy policies of interest allowed for reliability checks to take place by comparing the consistency of coding between individual texts.

4.3.1.2 Cluster Analysis

In discussions of both content analysis and in survey evaluation, it is necessary to provide background on cluster analysis, as this method contributes to the underlying structure of

analysis. The main purpose of clustering in the context of the privacy policy analysis is to split a set of objects, here identified as privacy policies, according to some feature variables, in this case constructs of interest as identified in the FTC Fair Information an Privacy Principles, contained in the texts.

To compute the similarity of entities, it is first necessary to define the concept of similarity. Aldenderfer and Blashfield (1984) state that, "The terms 'case,' 'entity,' 'object,' 'pattern,' and 'OTU' (operational taxonomic unit) denote the 'thing' being classified; whereas 'variable,' 'attribute,' 'character,' and 'feature' denote those aspects of the 'things' used to assess their similarity." In short, the objects of interest in the analysis are assessed for similarity based upon the comparisons of attributes they contain. For the computational step in determining similarity, the following measures have been suggested by Sneath and Sokal (1973):

- Correlation coefficients,
- Distance measures,
- Association coefficients, and
- Probabilistic similarity measures.

For purposes of this analysis, Euclidean distance measures will generally be used to determine similarity. Euclidean distance may be represented in the following way:

$$d_{xy} = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2}$$

where d_{xy} is the distance between x and y, and x_i is the value of the k^{th} variable for the i^{th} observation. In the case of content analysis, distance is calculated based on "cases" as defined by the user (such as document, paragraph, or sentence), and then compared to other cases to

determine similarity. This type of method may be generally called an agglomerative hierarchical clustering method, as it begins with many individual elements which are successively combined based on distance measures until only one cluster remains, containing all elements of interest.

Cluster analysis methods use a variety of techniques to find groups of similar items within a dataset. As noted, an agglomerative hierarchical clustering method was used here, which produces a hierarchy of clusters from small clusters of very similar items to large clusters that include more dissimilar items. Agglomerative hierarchical clustering builds a hierarchy from identified individual elements by performing progressive merges of clusters based on distance (in this case, Euclidean distance). Once the initial distance between each element is calculated and the initial merge performed, distance calculations are updated between clusters using the average linkage clustering method as follows:

$$\frac{1}{|A||B|}\sum_{x\in A}\sum_{y\in B} d(x,y),$$

where *A* and *B* are the identified clusters of interest. The average linkage clustering step is iterated until all elements have been merged into one cluster.

Hierarchical clustering methods produce a graphical output known as a dendrogram or tree that shows this hierarchical clustering structure. For this analysis, agglomerative clustering took place by identifying subsets of variables (or categories within the overall classifications) and determining their rates of similarity to other categories. Clustering will be applied in a number of forms in the content analysis, including heatmaps and dendrograms.

4.3.1.3 Correspondence Analysis

Hoffman and Franke (1986) describe correspondence analysis as follows: “In correspondence analysis, numerical scores are assigned to the rows and columns of a data matrix so as to maximize their interrelationship. The scores are in corresponding units, allowing all the variables to be plotted in the same space for ease of interpretation. This representation then can be used to reveal the structure and patterns inherent in the data.” Graphical analysis allows one to observe interrelationships between identified terms and concepts in 2- or 3-dimensional space, with the objective of “[representing] the relationship among all entries in the table using a low-dimensional Euclidean space such that the locations of the row and column points are consistent with their associations in the table (Provalis Research, 2010).” Unlike pairwise comparisons, correspondence analysis provides a robust representation of the relationship of all variables of interest to one another, including the ability to analyze how such relationships are made.

Correspondence analysis, “… may be defined as a special case of principal components analysis (PCA) of the rows and columns of a table, especially applicable to a cross-tabulation. However CA and PCA are used under different circumstances. Principal components analysis is used for tables consisting of continuous measurement, whereas correspondence analysis is applied to contingency tables (i.e. cross-tabulations). Its primary goal is to transform a table of numerical information into a graphical display, in which each row and each column is depicted as a point. (Nagpaul, 1999)” According to Theus (1997),

The results of a correspondence analysis are often used to visualize categorical data. Given a $c \times r$ contingency table, one calculates the singular value decomposition

$$X = U \Lambda V$$

where U are the eigenvectors of XX' and V the eigenvectors of $X'X$. X is the matrix of the standardized residuals of Pearsons x^2 statistic.

$$x^2 = \sum_{i=1}^{r} \sum_{j=1}^{c} \frac{(o_{ij} - e_{ij})^2}{e_{ij}} \quad \text{and} \quad x_{ij} = \frac{o_{ij} - e_{ij}}{\sqrt{e_{ij}}}$$

The o_{ij} denote the observed values, whereas the e_{ij} denote the expected values under the assumption of mutual independence. This kind of decomposition is a categorical equivalent to the principal components for continuous data and hence is very popular.

Here, the data analyzed is that of clusters of concepts defined within the content categories described above. Rows represent the content categories, while columns are comprised of varying types of policies as described in the content analysis chapter. Here, correspondence analysis is used to identify and describe how well policies of interest respond to constructs and categories of privacy interest identified in the FTC policy, as well as examining if there are patterns in how these constructs are treated by different types of agency or organization.

4.3.1.4 Computer Content Analysis

Computer-assisted content analysis has been noted to mitigate several pitfalls evident in human coding methods such as reliability, inter-analyst reliability and accuracy (Alexa and Zuell, 1999). The analysis undertaken here utilized the WordStat Content Analysis module (developed by Provalis Research), which operates from the base platform of either SimStat (a statistical software) or QDA Miner (a text management and qualitative analysis program). Documents of interest may be imported into either program in a variety of formats, including alphanumeric,

plain text, and rich text memos. For purposes of this analysis, privacy policies accessed online were first converted into Word documents, and then pre-processed, including spell checking, and removal of brackets, braces and hyphenation. Once all documents were imported into QDA Miner, the “Content Analysis” tab was chosen, which opens WordStat.

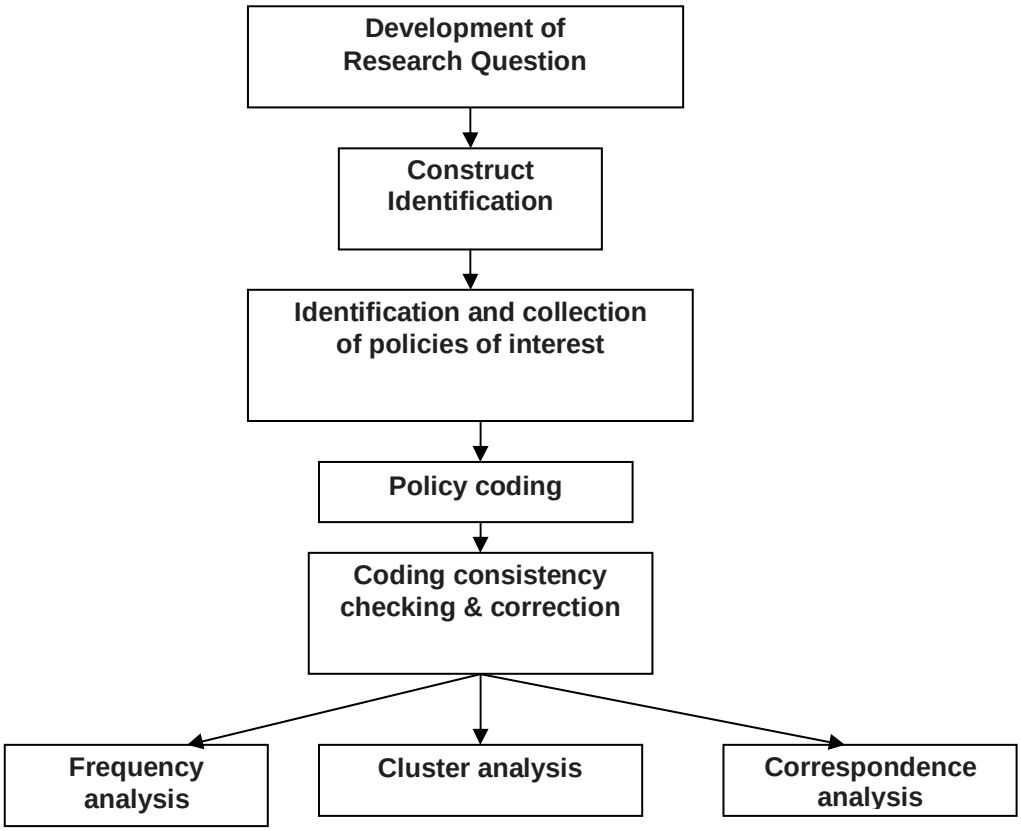
A number of processing features are available in WordStat, including the following:

- Lemmatization: Used to group different forms of a word (such as walk, walking, walked, etc.) in order to allow them to analyzed as a single term.
- Exclusion: An exclusion list is used to remove words that should not be included in the content analysis, such as pronouns, conjunctions, and other words or phrases with little discriminative value.
- Categorization: “Allows one to change specific words, word patterns or phrases to other words, keywords or content categories and/or to extract a list or specific words or codes.”

For this analysis, a series of preprocessing steps were conducted, including text “cleaning”, in which punctuation such as hyphens were removed and spelling was standardized; common terms (such as if, and, and the) were removed via use of an exclusion list; and all words were converted to uppercase. Such preprocessing allowed the text to be normalized for consistent coding and analysis.

The flowchart presented in Figure 4 describes the overall approach taken for the content analysis.

Figure 4: Content Analysis Process Flowchart



The first five steps have been described above. The frequency, cluster, and correspondence analysis steps were used for the detailed analysis of the collected and coded privacy policies in the context of overall privacy constructs of interest. The frequency analysis step was used to identify how often constructs of interest were responded to in privacy policies and resulted in an overall analysis of how comprehensively privacy policies address the variables of interest in the privacy context. Cluster analysis allowed for closer examination of similarity of policies, in part revealing issues of overall consistency of privacy expectations as presented to the consumer. Finally, correspondence analysis was used to provide information related to the co-occurrence of privacy topics in policies in order to allow for better understanding of related

privacy ideas within different types of agencies and organizations. Procedures and findings from each step are presented in more detail in Chapter 5.

4.4 Description of Survey Development

In this section, we focus on the analysis that was undertaken of the survey data, with the objective of identifying privacy preferences and the trade-offs that users of mobile information are willing to make for perceived utility received, and the perceived risks and expected compensation for giving up information.

4.4.1 General Survey Type

Approaches to measuring privacy preferences traditionally fall into one or more of the three forms of stated preference, revealed preference or willingness-to-trade. Perhaps the greatest limitation to stated preference and willingness-to-trade surveys is that they do not necessarily reflect how persons will act in a given situation, while revealed preference studies do not necessarily allow for such confounding factors as limited knowledge of technologies that may be used to prevent data sharing. Finally, as shown in 2.2.3, much of what is considered "private" is contextually based, which may limit the effectiveness of point-in-time studies.

A number of researchers, including Jensen, *et al.* (2005), Connelly, *et al.* (2007), and Sheedy and Kumaraguru (2007) have researched inconsistencies between stated and revealed privacy preferences, and in some cases have proposed methods by which stated preference surveys may be augmented by *in situ* scenarios. While moderately effective, such methods are

still subject to concern, as participant's actions may still be influenced by the testing context, by an inability to adequately recognize privacy implications of actions taken, or by inconsistent valuation of privacy variables. Given the presence of such concerns, as well as the hypothetical nature of many proposed ITS and LBS applications, the survey used for purposes of the dissertation is based on stated preferences.

A further consideration for the use of a stated preference survey is related to Slobogin's findings presented in section 3.4.2. Slobogin found that the public's expectations of privacy do not necessarily match with the Court's understanding, and he suggests that if legal decisions relevant to the Fourth Amendment are to be made on the basis of a "reasonable expectation" it may be necessary to reevaluate what is understood as "reasonable." If stated preference surveys related to privacy reflect a participant's overall expectations of how private data should be treated, separate from implications of individual actions, they are perhaps the best method by which to guide policy reflective of this overall criterion. Additionally, Wathieu and Friedman (2007) posit that an alternative reading of the "privacy paradox" of statement versus action may be that consumers do not feel capable of enacting their privacy preferences due to unanticipated or unknown consequences of sharing data in particular scenarios. Here, again, the consumer may not act in accordance with his or her privacy preference, but this should not necessarily suggest that their stated preferences are untrue or inconsistent.

While the noted limitations are of concern, survey information can be invaluable when determining privacy-utility tradeoffs. Krause and Horvitz's (2008) utility-theoretic approach to privacy and personalization provides a frame of reference for the expectations of the survey portion of the dissertation data collection. Here, a utility-theoretic approach was used to

ascertain a balance between the costs of sharing personal data online with personalization benefits gained by the sharing. Survey participants to use a Likert scale ranging from 1 (not very sensitive) to 5 (highly sensitive) to identify the sensitivity of attributes related to demographics, activities, or search topics related to their use of internet search engines. Additionally, participants were asked to provide preferences regarding the sharing of information relative to degrees of precision (region, country, state, city, zip code or address levels), levels of identifiability (whether they would be indistinguishable from at least *k* other persons), performance benefits, social groups (whether private data would be shared with family, friends, employers, or peers) and privacy-cost factors (such as paying more for increased privacy levels). The study here revealed methods by which varying privacy components could be evaluated, as well as identifying various constructs that may influence privacy preferences in a number of situations. The survey used for the dissertation has taken a similar approach by testing a number of constructs related to benefits, contexts, and economic factors in order to better evaluate privacy-utility constructs from a multidimensional context.

While many of the benefits evaluated in studies mentioned above focus primarily on economics and convenience, benefits of ITS technologies may also relate to safety and travel efficiencies. The utility-privacy trade-offs of such a system may be more difficult to gauge, as such benefits as safety improvements may require that the individual consider such things as value of life or physical well-being. Additionally, because ITS technologies will rely in part on fairly widespread deployment to attain maximum benefit, the user will need to evaluate not only personal utility (as in the cases above), but also societal utility. Keeping these factors in mind, an approach similar to that taken by the authors above would be appropriate to

determining a utility-privacy tradeoff curve for ITS technologies. The survey used for the dissertation (outlined in Chapter 6) addresses these needs by presenting survey participants with the opportunity to examine not only privacy in relation to individual benefits, but also system and network benefits. Previous research has indicated that a multi-step approach that evaluates (a) the individual's degree of sensitivity to the sharing of certain personal data, (b) the degree of anonymity that he or she prefers, and (c) the degree of benefit that he or she would expect in return for trade-offs of the first two elements may be the most effective way of accurately determining a user's privacy-utility tradeoff. Further, results from a study conducted by Danezis, *et al*. (2005) in conjunction with findings about concerns related to secondary usage indicate that asking participants questions regarding to what entities they would be willing to trade their data in return for what benefits will be a necessary part of the data collection.

In the context of locational privacy, the need to address the limitations of stated-preference approaches is particularly significant as much of the proposed technology and its implications are unfamiliar to the traveling public. Additionally, willingness-to-trade and privacy-utility evaluations are necessary insofar as many proposed and implemented ITS projects rely on the provision of personal information to create travel benefits in terms of time or cost. Statistical approaches to evaluating actual preferences in relation to general survey questions on stated preferences are described below.

4.4.2 Principal Component Analysis

While there are several variables of interest in the survey data, several are likely to be correlated. Principal Components Analysis (PCA) was used to reduce correlated variables into

sets of principal components in preparation for the Structural Equation Modeling step. PCA is used to explain variance-covariance structures in a variable set through a smaller number of uncorrelated linear combination of these variables. A principal component can be defined as a linear combination of optimally-weighted observed variables.

Let ***X*** be an *n* x *p* centered data matrix with zero empirical mean (where the mean values of each variable has been subtracted from each value along a column), where each of the *n* rows represents an observation, and each of the *p* columns is a variable. The Singular Value Decomposition (SVD) of the centered input matrix **X** is: ***X*** = ***UDV',*** where ***U*** is an *n* x *p* matrix of the eigenvectors of ***XX***', ***D*** is a diagonal matrix of dimension *p* x *p*, and ***V*** is an *n* x *n* matrix of eigenvectors of ***X'X***. The eigenvectors v_j (columns of ***V***) are called the principal components (or Karhunen-Loeve) directions of ***X***. The first principal component direction v_1 has the property that $z_1 = Xv_1$ has the largest sample variance amongst all normalized linear combinations of the columns of X. This sample variance is $\text{Var}(z_1) = \text{Var}(\boldsymbol{X}v_1) = d_1^2/n$ and $z_1 = \boldsymbol{X}v_1 = u_1 d_1$, where d_1 is the first diagonal element of ***D***. The derived variable z_1 is called the first principal component of ***X*** and u_1 is the normalized first principal component**.** Subsequent principal components z_j have maximum variance d_j^2/n, where d_j is the j^{th} diagonal element of ***D***, subject to being orthogonal to the earlier ones. Conversely the last principal component has minimum variance. In practice, the X matrix is first centered, and then the SVD or the eigenvectors are estimated, to determine the principal components.

By performing a principal component analysis, it is possible to calculate a score for each survey respondent on a given principal component. Whereas in reality, the number of components extracted in a principal component analysis is equal to the number of observed

variables being analyzed, in most analyses, only the first few components account for meaningful amounts of variance, so only these first few components are retained, interpreted, and used in subsequent analyses, such as the regressions used here to understand the nature of locational privacy. In particular, constructs of interest tested through questions related to willingness to trade private information, desire for compensation, context of interest (specifically, private or public services), and knowledge of privacy policies and practices will be subject to PCA in order to better evaluate related constructs to demographics, privacy expectations and risk perceptions.

4.4.3 Ordered Probit Modeling

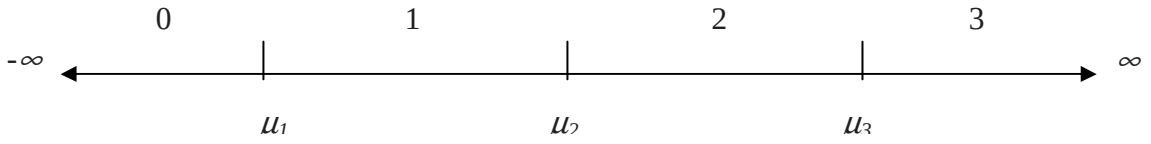
Respondents were queried about their preferences and values on a number of factors on Likert scales, with ordinal values. For example, one major factor of interest is a respondent's self-reported willingness to trade privacy for a number of benefits relating to their mobility conditions. To appropriately model the ordinal nature of such factors, ordered probit models were used. Following the example of the willingness-to-trade, the ordered probit model uses the following form:

$$y_i^* = \boldsymbol{\beta}' x_i + \boldsymbol{\varepsilon}_i$$

where y_i^* is the dependent variable (a continuous, composite score that gives the willingness-to-trade privacy of the i^{th} respondent), $\boldsymbol{\beta}$ is a vector of parameters to be estimated, $\boldsymbol{x}_i$ is a vector of independent variables and $\boldsymbol{\varepsilon}_i$ is the error term, which is assumed to be normally distributed with mean zero and unit variance, with cumulative distribution denoted by $\Phi(\cdot)$ and density

denoted by $\phi(\cdot)$. Given a decision regarding the use of mobile technologies, an individual falls in category *n* if $\mu_{n-1} < y_i^* < \mu_n$, where the μ's are cut-off thresholds to be estimated, along with $\boldsymbol{\beta}$. Figure 5 gives the relationship between latent, continuous underlying willingness-to-trade propensity and the observed willingness-to-trade category.

Figure 5: Relationship between latent, continuous underlying willingness-to-trade propensity and the observed willingness-to-trade category



The observed willingness-to-trade values, y_i, are related to the underlying latent variable y_i^*, as follows:

$$y_i = \begin{cases} 0 \text{ if } -\infty \le y_i^* \le \mu_1 (\text{willingness-to-trade is low}) \\ 1 \text{ if } \mu_1 < y_i^* \le \mu_2 (\text{willingness-to-trade is medium-low}) \\ 2 \text{ if } \mu_2 < y_i^* \le \mu_3 (\text{willingness-to-trade is medium-high}) \\ 3 \text{ if } \mu_3 < y_i^* \le \infty (\text{willingness-to-trade is high}) \end{cases}$$

The probability that respondent *i* falls into category *n* is given by:

$$\text{Prob}(y_i = n) = \Phi(\mu_n - \beta' x_i) - \Phi(\mu_{n-1} - \beta' x_i), \; n = 1, 2, 3$$

A positive coefficient means a higher underlying willingness-to-trade privacy in the use of location information as the value of the associated variable increases and a greater likelihood of reporting a higher category of self-assessed willingness-to-trade privacy risks in return for specifically queried transportation benefits. A negative value means a lower value of the latent variable y_i^* as the associated variable increases and a greater likelihood of reporting a lower

category of willingness to give up privacy for benefits. The estimated threshold parameters μ_1, μ_1 and μ_3, imply that a value of the latent variable less than $\hat{\mu}_1$ corresponds to the lowest willingness-to-trade, a value between $\hat{\mu}_1$ and $\hat{\mu}_2$ corresponds to low willingness-to-trade, a value between $\hat{\mu}_2$ and $\hat{\mu}_3$ corresponds to higher propensities of willingness-to-trade, whereas values above $\hat{\mu}_3$ corresponds to the highest levels of willingness-to-trade privacy.

Hausman tests (which test whether the potentially endogenous variable acting as a regressor and the disturbance are uncorrelated) were used to evaluate if willingness to trade privacy and the other constructs of interest are simultaneously determined, resulting in p values close to .1, indicating that the constructs are potentially exogenous. Since simultaneity was not strongly evident from the tests, single-equation ordered probit models were used to determine whether willingness to trade privacy by the respondents is related to utility derived from information, and other factors. However, due to the Hausman tests being significant at close to the 10% levels, multivariate Structural Equation Models (SEM) were used to further test relationships between collected data and constructs.

4.4.4 Structural Equation Modeling

It has been noted that the components of privacy in the mobile environment are not strictly organized in one to one correspondence, but rather are arranged in a system of interrelationships. One method that may be used to evaluate collected survey data of this form is structural equation modeling (SEM), which is a multivariate statistical approach of which factor analysis and path analysis represent special cases. SEM is commonly used as a

confirmatory technique used to ascertain model validity, though analysis may also include exploratory elements focusing on latent constructs composed of multiple measures. Used here, causal associations are hypothesized among variables and tested with a linear equation system.

According to Golob (2001), "An SEM with latent variables is composed of up to three sets of simultaneous equations, estimated concurrently: (1) a measurement model (or submodel) for the endogenous (dependent) variables, (2) a measurement (sub)model for the exogenous (independent) variables, and (3) a structural (sub)model, all of which are estimated simultaneously." For this analysis, the SEM has been conducted with observed variables measured via the stated preference survey. Lynch (2003) has described the basic formulation of an SEM as consisting of the following three equations and four matrices:

$$\begin{bmatrix} \eta_1 \\ \eta_2 \\ \vdots \\ \eta_j \end{bmatrix} = \begin{bmatrix} 0 & \beta_{12} & \cdots & \beta_{1j} \\ \beta_{21} & 0 & \cdots & \beta_{2j} \\ \vdots & \vdots & \ddots & \vdots \\ \beta_{j1} & \beta_{j2} & \cdots & 0 \end{bmatrix} \begin{bmatrix} \eta_1 \\ \eta_2 \\ \vdots \\ \eta_j \end{bmatrix} + \begin{bmatrix} \gamma_{11} & \gamma_{12} & \cdots & \gamma_{1k} \\ \gamma_{21} & \gamma_{22} & \cdots & \gamma_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ \gamma_{j1} & \gamma_{j1} & \cdots & \gamma_{jk} \end{bmatrix} \begin{bmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_k \end{bmatrix} + \begin{bmatrix} \zeta_1 \\ \zeta_2 \\ \vdots \\ \zeta_j \end{bmatrix}$$

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_p \end{bmatrix} = \begin{bmatrix} \lambda y_{11} & \lambda y_{12} & \cdots & \lambda y_{1j} \\ \lambda y_{21} & \lambda y_{21} & \cdots & \lambda y_{2j} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda y_{p1} & \lambda y_{p2} & \cdots & \lambda y_{pj} \end{bmatrix} \begin{bmatrix} \eta_1 \\ \eta_2 \\ \vdots \\ \eta_j \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_p \end{bmatrix}$$

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_q \end{bmatrix} = \begin{bmatrix} \lambda x_{11} & \lambda x_{12} & \cdots & \lambda x_{1k} \\ \lambda x_{21} & \lambda x_{21} & \cdots & \lambda x_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda x_{q1} & \lambda x_{q2} & \cdots & \lambda x_{qk} \end{bmatrix} \begin{bmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_k \end{bmatrix} + \begin{bmatrix} \delta_1 \\ \delta_2 \\ \vdots \\ \delta_q \end{bmatrix}$$

Φ=k-by-k covariance matrix of the ξ
Ψ=j-by-j covariance matrix of the ζ
θ_δ=q-by-q covariance matrix of the δ
θ_ε=pXp covariance matrix of the ε

Here, unobserved variables are represented by η (endogenous variables) and ξ (exogenous variables). According to Lynch (2003), "The coefficients that relate the η to each other are β, while the coefficients that relate the ξ to the η are γ." The first equation (the structural equation) relates the latent variables, providing an error term, ζ, for each η. The measurement equations (equations two and three) show how the observed *x* and *y* are related to the latent variables via the λ coefficients. Here, ε and δ represent the part of the observed variables that are unaccounted for by the influencing latent variables. The matrix referenced by Φ models covariances of exogenous latent variables, while the Ψ matrix models covariances of structural equation errors. Correlation error between errors in the measurement equations are allowed by the θ matrices (Lynch, 2003).

4.4.5 Conclusions

The PCA, SEM, and ordered probit models are used here to provide more detailed statistical analysis than will be possible through simple regression models. Because of the multi-dimensional nature of privacy as studied, PCA and SEM will be used to simplify the analysis. For example, a number of survey questions asked relate to knowledge of or attention paid to privacy policies of various services. Analyzing these items individually may reveal findings that would be better addressed in an overall construct of "knowledge" due to similarities of responses identified by survey participants; thus, they will be tested both individually and as a reflection of general attitudes. Additionally, a number of questions on the survey ask participants to provide responses to questions on a Likert scale, thus providing information that is ordered rather than cardinal (see Appendix 2 for additional details on question response

type). In these cases, the ordered probit model will provide a better estimation of variable influence.

4.5 Conclusion

Chapter 4 is intended to provide an overview of the research design and methodologies used for purposes of the dissertation. In each case, additional and more project-specific information will be presented in Chapters 5 and 6, which will present more detailed analyses of how these methods are applied to the collected data for purposes of responding to the developed research questions. It is hoped that this general outline will provide a context within which the methodologies proposed for the dissertation research may be adequately understood and their place in the data collection process generally accepted. Each approach will help address one or more of the primary research questions outlined above, specifically:

1. The content analysis of privacy policies from a range of agencies will help to determine current state-of-practice regarding privacy rights and responsibilities as enacted in the mobile environment from the point of view of both law/policy and technology.
2. The survey will be analyzed to examine evidence of privacy and privacy-utility tradeoffs from the user's point of view, thus allowing for a more targeted method of proposing techniques to alleviate user concerns.

CHAPTER 5: ANALYSIS OF PRIVACY POLICIES

5.1 Introduction to Content Analysis of Privacy Policies

While the overview of overarching privacy policies given above provides the general context for privacy preservation in the United States, the lack of consistent and comprehensive guidance provided for both public and private service providers leaves a gap in the ability of researchers to effectively determine how well individual organizations, or groups of organizations, respond to these recommendations. To begin the task of addressing this gap, we have taken the approach of evaluating the privacy policies of individual organizations via the use of content analysis. In accordance with the content analysis approach discussed in Chapter 4, the steps used to analyze and evaluate the collected privacy policies will be described, followed by a descriptive analysis of the policies to be examined. Subsequently, the results of the content analysis, including frequency analysis, cluster analysis, and correspondence analysis, will be presented, followed by a discussion of the findings. Finally, the relationship of the content analysis results to the overall theme of the thesis will be discussed.

5.1.1 Policy Analysis Results

5.1.1.1 Description of Privacy Policies

Privacy policies evaluated here fall into two general categories: first are those that deal generally with privacy issues over a range of services (such as website, services, and applications); second are those that specifically address the privacy issues inherent in use of a

mobile service or application. The distinction here is an important one, as data gathered from each of these sources will provide varying degrees of personal information related to a traveler's location, preferences, and patterns of use.

In collecting privacy policies for this analysis, efforts were made to concentrate on specific polices related to the use of applications and services offered by providers; however, in many cases available policies were generic and related to general data collection policies and website use. Efforts were made to contact service providers, many of which resulted in receipt of electronic copies of the standard privacy policies. In one case, for example, a public entity involved with electronic toll collection (ETC) was contacted for information related to privacy policies relevant to use of data gathered via ETC use. In this case, the researcher was instructed to file a Freedom of Information Act (FOIA) request for such information. In such cases, terms of use for services (such as transit cards and ETC transponders) were examined to determine if they referred to specific privacy policies related to their use. If these were not available or requests for access were not granted, privacy policies of agencies or organizations were evaluated for general practices, and to determine if they specifically referred to use of mobile or location-based devices.

In many cases, terms of use or service for publicly-managed ETC and electronic transit cards dealt only briefly with privacy, or referred one back to the general privacy policy for the agency, organization, or locale. To provide some analysis of these services, nineteen policies were obtained (seven from ETC service providers, and twelve from providers of electronic transit cards). Of the ETC service providers researched, none had a specific privacy policy related to the ETC service available. Each had terms and conditions with a section related to

privacy or confidentiality of information, but most referred to the providing agency or state Department of Transportation's general privacy policy for additional information. For providers of electronic transit passes, six provided privacy policies, while the remaining dealt with privacy along the lines of the ETC providers. In these policies, roughly 57% (10/19) dealt specifically with privacy related to the product or service offered, while only 16% (3/19) dealt with privacy related to location information gathered as part of use of the product or services. Overall, privacy policies of electronic transit cards were found to more consistently address both of these issues. These policies differ significantly from those of general public location service providers, and thus will be evaluated separately from the more generic policies grouped under the "Public" heading.

While all of the policies of private providers referred to use of their websites, only a few referred specifically to how the privacy policy will be implemented with respect to products, services or location information. Roughly 49% addressed the specific product or service, while 42% addressed location-specific information. This finding was concerning, as it reflects that many companies do not recognize and/or acknowledge the particular privacy concerns associated with location information. General privacy policies of public service providers are not linked to specific applications, and thus generally do not address location- or product-specific information. As noted above, those providers that do provide services that may be able to track location data (such as ETC or electronic transit cards) have had their policies evaluated separately.

Another concerning finding of privacy policies overall relates to reading levels. According to the Office of Educational Research and Improvement (2002), the average reading

level of U.S. adults is between the 8th and 9th grade. The Flesch-Kincaid Grade Level calculator, which uses average sentence length and average number of syllables per word along with weighting factors to determine a reading level, is a standard test used to ascertain the general readability of texts as measured by US grade level. The test was applied to each policy considered, resulting in the following averages:

- Overall average (all companies): 14.60 (range of 9.71 – 24.50)
- Overall average (public organizations): 14.33 (range of 11.43 – 19.02)
- Overall average (private companies): 14.27 (range of 9.71 – 24.50)
- Overall average (ETC providers): 13.87 (range of 10.09 – 17.69)
- Overall average (Electronic Transit Card providers): 15.34 (range of 12.09 – 23.06)

These figures indicate that the overall average reading grade of the privacy policies studied here is roughly that of a sophomore to a junior in college, well over that of the average US adult. Such a finding provides a baseline indication that the privacy policies used by mobile transportation service providers may not be understandable by the average user, and thus meet neither the CTIA's recommendation that policies be comprehendible, nor the FTC's general notice provision.

5.2 Categorization

Categorization allows for individual words or phrases to be grouped under a common heading, thus simplifying the process of analyzing text within headings of interest. For purposes of this content analysis, the following five categorizations were developed, based on the identified areas of concern in the Federal Trade Commission's (FTC) Fair Information Practice Principles:

- Access/Participation
- Choice/Consent
- Enforcement/Redress

- Integrity/Security
- Notification/Awareness

Words and phrases found within the identified privacy principles were then assigned to a category based upon their most common uses and meanings. To effectively categorize words, a number of policies were reviewed to determine how words and associated phrases were used within the context of the policies. Words and phrases within policies were next assigned a category and then these categorizations were compared across policies to ensure consistency. Once words and phrases contained in the document were categorized and subjected to the consistency check, a frequency analysis was run at the Category level to determine how often policies referred to the assigned topics. While some words tended to overlap in different categories, most were able to be effectively and accurately classified into overall categories. Table V outlines how frequently occurring words were categorized.

To ensure that words and phrases included in the categorization dictionary were not overemphasized due to recurrence within a unit of interest (for example, within a paragraph focusing on a certain category of interest), paragraphs and sentences were treated as thought units within policies and frequencies determined on the amount to which the category of interest was emphasized within paragraphs over the course of the document. For example, if a word were repeated within a sentence, it would count as only one occurrence, as it is likely that the second occurrence would be for purposes of emphasis or clarification. If, on the other hand, a large section of text were left uncoded, that section was reviewed to determine if a word, phrase or category had need of addition to the overall analysis. To ensure that these techniques were applied consistently, codings were reviewed after the initial coding had taken place, and recoded if inconsistencies were found between policy codings.

Two separate analyses were run, one on privacy policies of public companies and one of privacy policies of private companies. The next section reviews results of each of the two analyses.

Table V: Privacy Policy Word Categorization

Content Analysis Database Keyword Categorization					
Classification	**Category**	**Keywords**	**Classification**	**Category**	**Keywords**
Access/ Participation	Contest Change	Contest	**Notification/ Awareness**	About	Assurance
		Incorrect			Changes
	View	View			Children's Privacy
	Choice/Consent	Opt-in		Collected Data	Address
		Opt-Out			Birthdate
Enforcement/ Redress	Contact	Contact			Device
		Questions			Device ID
	Government Enforcement	Government			IP Address
		Legal Claims			Location
		Legal rights			Password
		Required by law			Personal Information
		Safe Harbor			Phone number
		TrustE			Name
	Private Remedies	Disclosure			Preferences
		Fraud			Username
		Illegal			Zip code
		Investigate		Cookie use	Cookie
		Prevent		Cookies and Clickthrough	Beacon
		Take action		Data Mining	Combine
		Transfer			Mine
		Violation		Data Use	Advertising
		Terms of use			Billing
	Self Regulation	Your responsibility			Contests
		You guard			Outreach
		Protect your			Publicity
Integrity/ Security	Managerial	Managerial			Support
		Administrative			User experience
		Procedural		Ownership	Ownership
	Technical	Electronic			Own
		Technical		Third Parties	Advertisers
		Physical			External links
		Protect			Government agencies
		Security			Partners
					Third parties
					Vendors
				User Action	Check in
					Download
					Make a call
					Report location
					Send a text message
					You send
					You tell us
					You click

5.2.1 Frequency Analysis

First, a general frequency analysis was conducted in order to describe the general landscape of the presence of identified privacy elements in the privacy policies of both public and private providers, as well as policies associated with ETC and electronic transit card services, as shown in Table VI. Clear differences and similarities in the topics addressed by policies in each category quickly emerge. Overall, it is evident that the “Collected Data” category is the most widely

addressed, with elements cited by roughly 97% and 98% of public and private companies, respectively. Such a finding is unsurprising for a number of reasons: 1) the number of words included in this category (13) is higher than that of any other category, and 2) The overarching category ("Notification/Awareness") is often considered the fundamental privacy principle, as it is notification of rights and responsibilities that encourages the development of privacy policies. This finding is not applicable, however, for ETC and transit card providers, with only 14.4% of these policies addressing this issue. This finding may be related to these policies addressing only general confidentiality practices, and relying on agency privacy policies for more substantive information.

Private companies were generally more likely to address more aspects of privacy as identified in the FTC guidance, with some exceptions such as Children's Privacy, language referring to notification of changes to the privacy policy, and actions related to third parties. Public agencies, on the other hand, fall short in their addressing of managerial means of privacy protection and the issue of how users may contest or change collected data. ETC and electronic transit card providers, in general, demonstrated overall inconsistency with policy framing when compared to both public and private policies. This may, again, be related to reliance on general privacy policies to address more specific areas of privacy (such as children's privacy or data collection and use policies), but may also reflect a difference in how privacy terms for these policies differ significantly from those of more comprehensive service providers.

Table VI: Category Presence Frequencies in Privacy Policies of Public and Private Transportation Service Providers and ETC and Electronic Transit Card Providers

Electronic Toll Collection and Electronic Transit Card Providers (19 Total)						
Category	**Sub-Category**	**FREQUENCY**	**% SHOWN**	**NO. CASES**	**% CASES**	**TF • IDF**
Access/ Participation	View	3	0.80%	3	0.60%	6.6
Choice/ Consent	Opt-In	9	2.50%	9	1.80%	15.7
	Opt-Out	1	0.30%	1	0.20%	2.7
Enforcement/ Redress	Contact	25	7.00%	25	5.10%	32.4
	Government Enforcement	5	1.40%	4	0.80%	10.5
	Private Remedies	48	13.40%	27	5.50%	60.6
	Self-regulation	2	0.60%	2	0.40%	4.8
Integrity/ Security	Integrity/Security	36	10.10%	28	5.70%	44.9
	Managerial	1	0.30%	1	0.20%	2.7
	Technical	40	11.20%	31	6.30%	48.1
Notification/ Awareness	Assurance	11	3.10%	11	2.20%	18.2
	Collected Data	116	32.50%	71	14.40%	97.7
	Cookie Use	31	8.70%	17	3.40%	45.4
	Data mining	12	3.40%	7	1.40%	22.2
	Data use	2	0.60%	2	0.40%	4.8
	Ownership	7	2.00%	7	1.40%	12.9
	Third Parties	1	0.30%	1	0.20%	2.7
	User Action	7	2.00%	7	1.40%	12.9

Frequency: Number of occurrences of the word or category names

% Shown: Percentage based on the total number of words displayed in the table

No. Cases: Number of cases where this keyword appears.

% Cases: Percentage of cases where this keyword appears.

*TF*IDF Term frequency weighted by inverse document frequency.*

Frequency analysis, on its own, is insufficient to evaluate the overarching comprehensiveness of privacy policies, as it does not allow for the context of identified elements to be adequately described. For example, while it is of use to determine how many organizations or agencies refer to either “Self-regulation” or “Contact” under Enforcement/Redress, it may be of more value to ascertain how many policies refer to both concepts, and to what extent the two are related, in order to determine the degree of input the consumer has on protection of his or her privacy. In order to evaluate such measures, clustering may prove a useful tool.

5.2.2 Cluster Analysis

Based on the frequency analysis, dendrograms were next created. In a dendrogram, the vertical axis is made up of the items and the horizontal axis represents the clusters formed at each step of the clustering procedure. Words or categories that tend to appear together are combined at an early stage while those that are independent from one another or those that don't appear together tend to be combined at the end of the agglomeration process (Provalis, 2010).

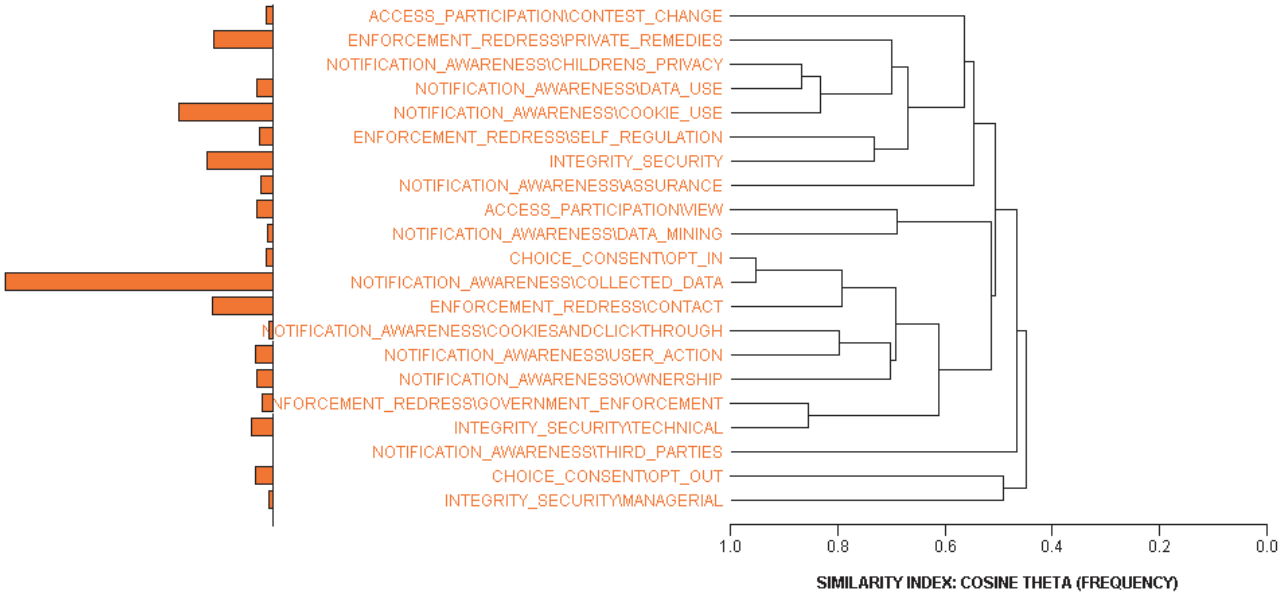
The following four options are given for indexing the dendrogram, allowing the user to select the similarity measure used for clustering and multidimensional scaling:

- Jaccard's coefficient - This coefficient is computed from a fourfold table as a/(a+b+c) where *a* represents cases where both items occur, and *b* and *c* represent cases where one item is found but not the other. In this coefficient equal weight is given to matches and non matches.
- Sorensen's coefficient - This coefficient (also known as the Dice coefficient) is similar to Jaccard's but matches are weighted double. Its computing formula is 2a/(2a+b+c) where *a* represents cases where both items occur, and *b* and *c* represent cases where one item is present but the other one is absent.
- Ochiai's coefficient - This index is the binary form of the cosine measure. Its computing formula is SQRT(a^2/((a+b)(a+c))) where *a* represents cases where both items occur, and *b* and *c* represent cases where one item is present but not the other one.
- Cosine theta - This coefficient measures the cosine of the angle between two vectors of values. It ranges from -1 to +1. (Provalis Research, 2010)

For purposes of this analysis, we have chosen to use cosine theta, as it takes into account not only the presence of a word or phrase in a case, but also how often the word or phrase occurs. This additional information will allow for better determinations of similarity to take place. Dendrograms using cosine theta for private and public privacy policies, as well as ETC and electronic transit card policies are shown in Figure 6.

Figure 6: Policy analysis dendrograms identifying clusters of related concepts

Private Policy Dendrogram



Public Policy Dendrogram

ETC and Electronic Transit Card Dendrogram

From the dendrograms above, we can determine weight, compactness, and distinctness, which are defined as follows:

- Weight - the rough percentage of all individuals that fall within each cluster
- Compactness - how similar to one another the elements of a cluster are
- Distinctness - how different one cluster is from its closest neighbor (ESRI, http://edndoc.esri.com/arcobjects/8.3/Samples/Analysis%20and%20Visualization/Cluster%20Analysis/CLUSTERANALYSIS.htm)

From the dendrograms above, one can see that public and private privacy policies and ETC/Electronic Transit Card policies have quite different structures regarding the presence of various privacy concepts. For example, for private policies the codes for “Opt-in” and “Collected data” are quite compact, indicating that they are fairly similar in presence. The weight of the primary cluster which contains these elements is also fairly high, with 38% (eight of 21) of concept elements falling within this cluster with a similarity index of 0.6. Such a finding indicates that the clustered elements (Opt-in, Collected data, Contact, Cookies and Click-

through, User action, Ownership, Government enforcement and Technical) are relatively similar in their presence within the privacy policies of the included private agencies. While at first glance the cluster may seem odd, it is clear that each of the included elements are, a) seen with relatively high frequency according to Table VI, and b) primarily related to Notification/Awareness. The clustering of these concepts indicates that they are closely related within the objectives of these privacy policies within the private sector.

A second fairly compact cluster consists of the following six elements: Private remedies, Children's privacy, Data use, Cookie use, Self-regulation and Integrity/security. While still primarily concerned with Notification/awareness, this cluster is more closely concerned with how consumer data are used and protected. The relatively higher inclusion of Enforcement/redress categories here indicates that those companies that include some information regarding the protection of consumer data are also likely to include information on the uses of those data. It is possible that the relative freedom of private companies as opposed to public companies to determine enforcement mechanisms may play a role in this clustering.

For public policies, the landscape is somewhat different. The closest cluster in this area is that of Assurance and Ownership, both included under Notification/awareness. The Assurance element is a somewhat general category, consisting primarily of statements related to general assurances of privacy for the consumer (such as, "Agency X values your privacy", etc.). Such assurances are fairly generic, thus their close association with the element of Ownership, which occurs in a relatively small number of cases, is somewhat surprising. In all, based on the clusters identified, there is less overall consistency of content in public privacy policies than in their private company counterparts; however, as with private companies, one

fairly large cluster does exist. The largest and most compact cluster is composed of the following elements: Government enforcement, Technical, Data use, Changes, Children's privacy, Third parties, Assurance, Ownership, and Cookie use. As above, this cluster has a similarity index of 0.6, indicating fairly close association. As with the private company policies, Notification/awareness is the primary category associated with this cluster; however, the types of information provided to the consumer is somewhat different. One possible rationale for this is that such elements as the protection of children's privacy are mandated under federal law, and are thus more likely to be included by government agencies. Also, more strict regulations regarding the sharing of data by public agencies may make more explicit information regarding this element more common. In general, however, it appears that like elements tend to be clustered with like, indicating that there are some limitations as to the comprehensiveness of policies in addressing each of the above-identified primary categories.

Clusters seen in ETC and electronic transit card policies are markedly different from those of both public and private service providers. Here, the closest cluster is that of "Managerial" and "Collected Data," with a similarity index of roughly 0.9, indicating that these two concepts are closely related in these policies. Such a close relationship reflects, in part, the reliance upon the concept of "data confidentiality" in these agreements, with many policies indicating that collected data will be kept confidential via limitations on the ability to share data by managerial means. The relatively close association of "Private Remedies," which generally indicates that the service provider will use internal methods to address confidentiality or privacy concerns, also indicates that such concerns are often viewed from the vantage point of agency responsibilities. The remaining clusters are far less compact, indicating a high level of

distinctness for each cluster. This is, in part, reflective of the findings from the frequency analysis, which indicated that there is little consistency in the degree of information provided to consumers in the framework of this analysis.

5.2.3 Correspondence Analysis

For purposes of this analysis, correspondence will first be looked at generally via the use of a heatmap plot and associated table. Next, 2- and 3-D correspondence plots will be generated to allow for graphical analysis based on the sub-categories identified in Table V above in the context of public, private, and ETC and electronic transit card policies. While individual keyword analysis can be performed, such analysis tends to create overly complex graphics, which are difficult to analyze. As this analysis is focused more generally on the presence of key concepts found within each type of policy, the decision has been made to focus on the overall clusters of interest found within each category. Additionally, case occurrences of sub-categories are analyzed in order to better identify how concepts correspond within policies. It is expected that there will be differences between the correspondence analyses of the policies analyzed, particularly in regard to enforcement and children's privacy, due, in part, to expectations reflected in policies relevant to public service providers. The following section will outline the findings of the correspondence analysis.

5.2.4 Heatmap

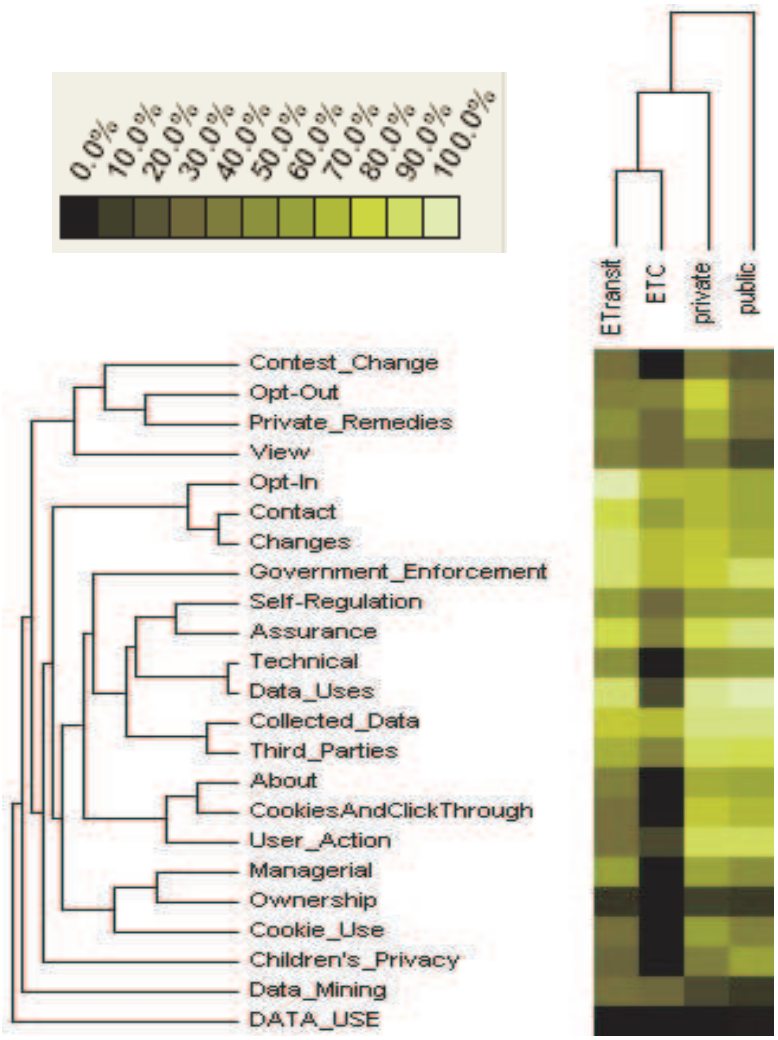
A heatmap shows relationships between keywords or keyword clusters and different categories of interest. We have examined relative frequencies of sub-categories of interest identified above in the context of public, private, and ETC/electronic transit card privacy policies. Table VII

shows the base data indicating the percentage of words included in the coding categories found in each group, while Figure 7 shows the heatmap generated from these percentages.

Table VII: Percentage of Coded Words per Category by Policy Type

	ETransit	ETC	private	public	Pearson's R	P value
Contest_Change	1.20%	0.00%	1.10%	1.00%	-0.219	0.015
View	1.30%	1.60%	1.70%	0.90%	-0.091	0.185
Opt-In	6.10%	11.30%	3.80%	2.60%	-0.124	0.110
Opt-Out	5.30%	7.90%	5.30%	3.20%	-0.327	0.000
Contact	2.80%	13.90%	2.80%	2.10%	-0.01	0.462
Government_Enforcement	8.60%	14.40%	5.40%	13.60%	0.012	0.455
Private_Remedies	4.00%	1.90%	5.20%	2.40%	-0.311	0.001
Self-Regulation	2.70%	7.50%	3.50%	3.50%	-0.191	0.029
Managerial	2.20%	0.00%	3.60%	2.60%	-0.235	0.010
Technical	5.40%	0.00%	2.80%	4.80%	-0.101	0.161
About	5.90%	0.00%	6.40%	5.40%	-0.262	0.004
Assurance	6.50%	3.40%	5.70%	4.90%	-0.258	0.005
Changes	4.00%	14.50%	4.00%	3.30%	-0.07	0.247
Children's_Privacy	1.40%	0.00%	3.20%	5.60%	-0.001	0.497
Collected_Data	14.20%	9.60%	6.40%	9.00%	-0.131	0.098
Cookie_Use	1.50%	0.00%	2.80%	3.30%	-0.166	0.050
CookiesAndClickThrough	2.50%	0.00%	3.70%	10.90%	-0.099	0.166
Data_Mining	2.10%	3.00%	0.50%	0.20%	0.033	0.374
Data_Uses	7.40%	0.70%	11.00%	7.40%	-0.431	0.000
Ownership	0.20%	0.00%	0.10%	0.60%	-0.03	0.384
Third_Parties	8.90%	6.10%	15.00%	7.10%	-0.433	0.000
User_Action	5.70%	4.10%	5.80%	5.50%	-0.246	0.007
DATA_USE	0.00%	0.00%	0.00%	0.10%	0.023	0.409

Figure 7: Privacy Policy Heatmap Generated from Categorical Word Occurrence



Some clear distinctions emerge as one evaluates the percentages shown here, although many values are in relative agreement. First, public organizations and providers of electronic transit cards are far more likely than private companies or ETC policies to refer to methods of government enforcement – often citing applicable laws and regulations that may provide consumers with additional information regarding data collection practices. Private companies,

on the other hand, are more likely to instruct consumers regarding “private remedies,” which generally includes investigations undertaken on the part of the company to evaluate potential privacy breaches. Such differences indicate competing methods of identifying and addressing privacy needs and concerns. For public agencies, ultimate authority rests in the structure of government hierarchy and laws to which they are subject. For private companies, which are subject to far fewer regulations and laws, much of the burden for addressing privacy concerns lies in the structure and methods of the company itself, though there is acknowledgement on the part of most companies that privacy may be intentionally breached if data requests are made through government warrant. Notifications regarding the need for “self-regulation,” which indicate measures that should be taken on the part of the consumer for self-protection (such as reading policies, setting browsers or other portals to reject “cookies,” or setting preferences in accordance with individual privacy preferences) are generally consistently referred to across both public and private organizations.

ETC policies, in general, are the least comprehensive of the policies evaluated here, as they tend to address the fewest number of categories of interest. Again, this reflects, in part, the fact that these policies are often contained within Terms and Conditions of use that may direct the consumer to view the general privacy policy of the providing organization. In these cases, the portion of the policy that deals directly with privacy may be fairly concise. This finding, however, raises the issue of comprehensiveness. Because ETC systems have the capability to collect large quantities of location data linked to financial and personally identifying information, they may not be adequately addressed in general privacy policies. While electronic transit cards may encounter some of the same issues, they are, generally,

more comprehensive and deal more directly with issues related to the collection of location data. As seen in the heatmap, these policies are most likely to address issues of government enforcement and data uses, issues of particular concern when addressing location data collected through a public agency.

Closely linked to differences seen in how privacy is treated within a legal and regulatory framework is the issue of children's privacy. Public agencies are far more likely to address issues of children's privacy in their privacy policies, in part due to federal regulations such as the Children's Online Privacy Protection Act (COPPA) of 1998. Such notice is generally relevant to online policies, as discussed above, which may account for this issue being treated most comprehensively within the privacy policies of public agencies and electronic transit card providers. ETC policies showed the lowest degree of addressing this aspect of privacy, which is likely related to age requirements for obtaining an ETC device based on driver licensing age requirements.

In addition to examining the percentages of coded words across provider categories, we also consider Pearson's R between the categories of keywords and the type of privacy policy evaluated. Findings from this analysis show that few variables have a significant correlation, with only "Data Uses" and "Third Parties" showing moderately negative correlations. The results, with p-values, are given in Table VIII below.

Table VIII: Significance Levels of Tested Privacy Policy Categories

Category	P value
Contest_Change	0.015**
View	0.185
Opt-In	0.110
Opt-Out	0.000*
Contact	0.462
Government_Enforcement	0.455
Private_Remedies	0.001*
Self-Regulation	0.029
Managerial	0.010*
Technical	0.161
About	0.004*
Assurance	0.005*
Changes	0.247
Children's_Privacy	0.497
Collected_Data	0.098
Cookie_Use	0.05**
CookiesAndClickThrough	0.166
Data_Mining	0.374
Data_Uses	0.000*
Ownership	0.384
Third_Parties	0.000*
User_Action	0.007*
DATA_USE	0.409

**Significant at 0.01 or above*

***Significant at 0.05 or above*

As seen here, there are a number of categories which vary significantly between policy types. Some of the strongest correlations are seen in areas where ETC policies differ significantly from the other policies (such as data uses, managerial, third parties, cookie use, and user action). In these cases, significance may be related to the differences in activity type and policy type. As noted, the ETC policies often present only basic information related to privacy, and direct the user to read the general privacy policy of the managing agency for additional information.

5.2.5 Graphical Representation of Correspondence Analysis

Graphical representations of correspondence allow for a more visual representation of the correspondences discussed above, and provide a clearer picture of how the privacy policies discussed here relate to one another and to the examined keyword categories. Two- and three-dimensional plots are created using cross-tabulations (or contingency tables) of rows and columns, where columns represent the category of policy studied, and rows represent the category of keyword (as seen in Table VII). The closeness of points in the chart represents similarity of row or column profiles. According to Lebart, *et al.* (1998), "According to usual notation, $f_{i.}$ designates the sum of the elements of row *i* and $f_{.j}$ is the sum of the elements of column j of this table. The profile of row *i* is the set of *p* values:

$$\left(\frac{f_{ij}}{f_i} \right), j = 1, \ldots, p$$

The profile of column *j* is the set of *n* values:

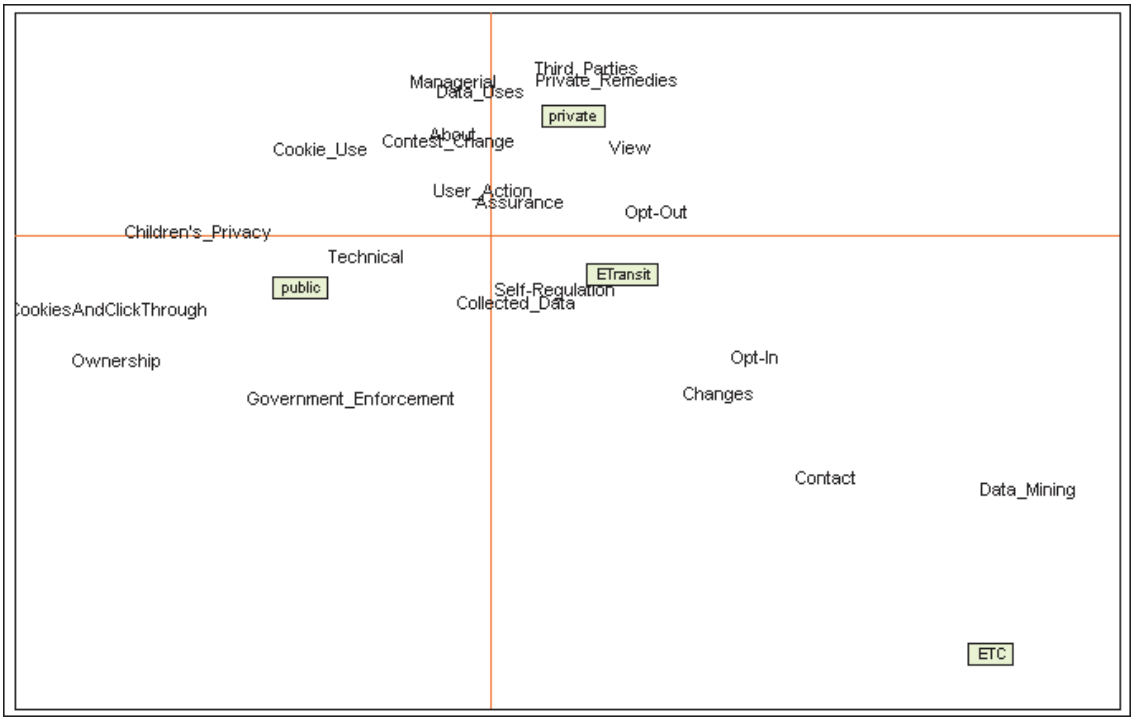
$$\left(\frac{f_{ij}}{f_{.j}} \right), i = 1, \ldots, n"$$

The origins of the axes in correspondence plots represent the marginals of the table of frequencies, with distance from the origin indicating singularity of items (in this case, privacy policy types). A visual representation of the contributing factors to this similarity or difference is found in the position of category keywords relative to the policies of interest. As with the policy categories, the location of a keyword category relative to the location of the origin or policy type indicates its' singularity relative to the overall distribution. Category and sub-category associations for public, private, ETC, and electronic transit card providers are presented in the

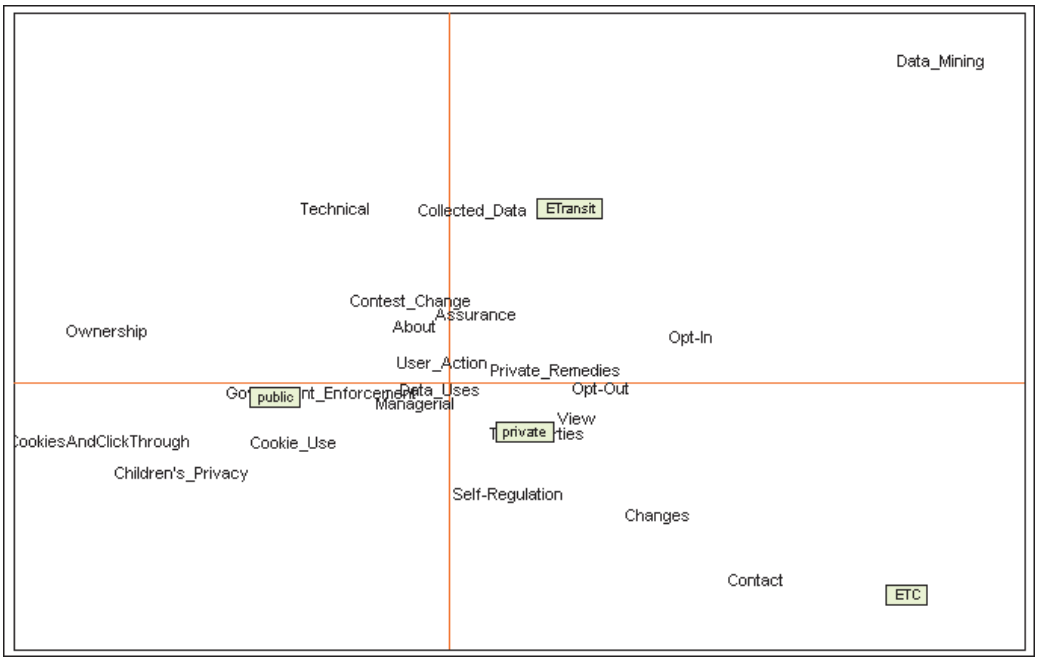
2D and 3D graphics below, with corresponding Eigenvalues shown in Table IX. The closeness and clustering of category keywords in relation to the policy types

Figure 8: Correspondence Plots of Policy Concepts with Policy Types

2-D Correspondence Plot: Axis 1 vs. Axis 2

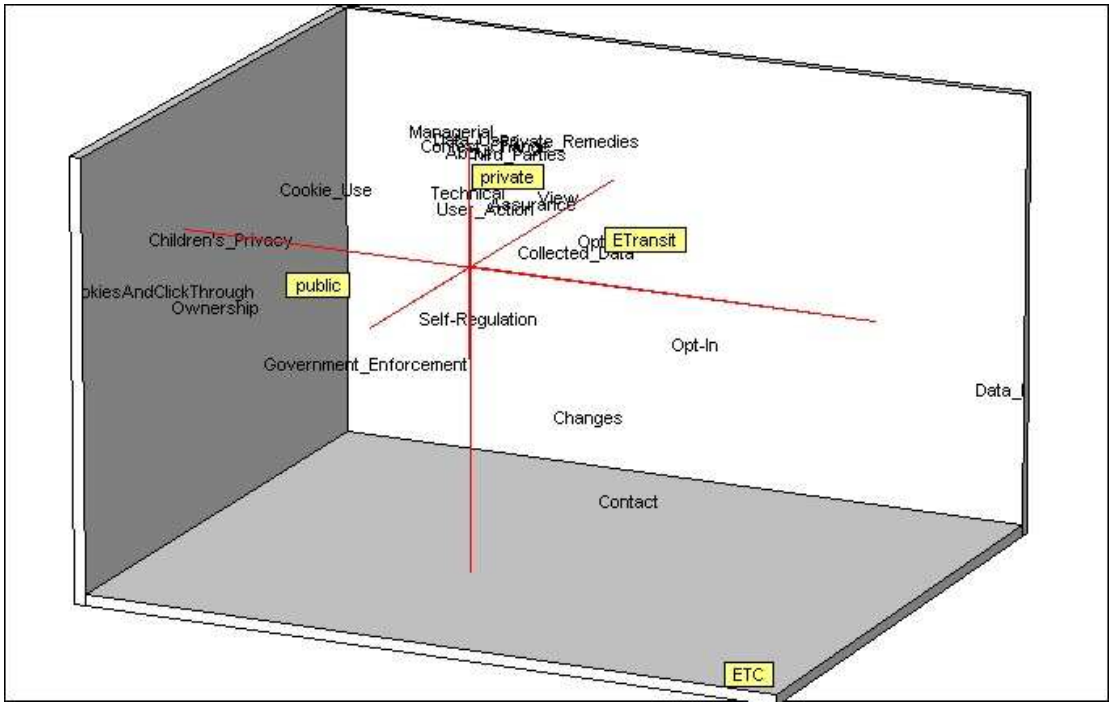


2-D Correspondence Plot: Axis 1 vs. Axis 3



2-D Correspondence Plot: Axis 2 vs. Axis 3

3D Correspondence Plot



It is evident from examination of these graphs that significant differences exist in the types of policies examined here. In particular, ETC policies vary significantly with respect to the origin of the remaining three policy types along all axes, indicating that they display the most dissimilarity and the least inclusion of categories of interest. Given the discussion above, this is not a surprising finding; however, the degree of difference seen in the four plots above indicates the degree of difference between each of the policy types in relation to the keyword categories studied.

In the plots above, distance from the origin indicates the relative singularity of items of interest. Table IX below provides an overview of relative distance from the origin for both categories and variables of interest.

Table IX: Relative Distance from the Axis of Origin of Privacy Policy Types and Content Categories

Variable Coordinates			
Item	**Axis1**	**Axis 2**	**Axis 3**
private	0.479	1.071	-0.546
public	-1.108	-0.528	-0.175
ETransit	0.763	-0.325	1.863
ETC	2.894	-3.409	-2.317
Category Coordinates			
Item	**Axis 1**	**Axis 2**	**Axis 3**
Contest_Change	-0.247	0.678	0.843
View	0.799	0.726	-0.431
Opt-In	1.533	-0.915	0.448
Opt-Out	0.954	0.087	-0.111
Contact	1.939	-1.906	-2.182
Government_Enforcement	-0.815	-1.411	-0.156
Private_Remedies	0.666	1.267	0.08
Self-Regulation	0.375	-0.574	-1.255
Managerial	-0.223	1.183	-0.28
Technical	-0.725	-0.233	1.836
About	-0.224	0.824	0.558
Assurance	0.158	0.348	0.683
Changes	1.317	-1.28	-1.49
Children's_Privacy	-1.692	0.008	-1.032
Collected_Data	0.145	-0.708	1.819
Cookie_Use	-0.992	0.578	-0.699
CookiesAndClickThrough	-2.225	-0.685	-0.682
Data_Mining	3.11	-2.006	3.436
Data_Uses	-0.058	1.183	-0.13
Ownership	-2.178	-1.145	0.508
Third_Parties	0.551	1.353	-0.6
User_Action	-0.053	0.216	0.169

From both this chart and the visual representation shown above, it is clear that the ETC category of privacy policies is the most singular of the policy types studied. In addition, contact, changes, opt-in and data mining are the most singular of categories. The charts and findings above indicate that while there is some consistency in topics addressed in overall privacy policies, there are some significant discrepancies in how well these topics are addressed by the various types of organization of interest. Such a finding, and the issues associated with these inconsistencies, indicates that there is scope for guiding policy that would bring more clarity and consistency to privacy policies. Particularly in the area of contact information and access to information, current practices leave the consumer with little information or ability to ensure

that collected data are accurate and being used correctly. The concerns that a consumer may have about this situation will be further described in the next section, which details findings related to the general survey conducted as part of the dissertation process.

5.3 Conclusions

A number of conclusions may be drawn from the preceding analysis, particularly in relation to consistency and comprehensiveness. One key finding is that it is often difficult to obtain application or service-specific information on privacy policies related to data collected in the mobile environment. The lack of policies dealing specifically with location information gathered as part of application use or electronic transportation services (such as position information, trip routes traveled, or origin and destination information) indicates that there is currently little attention being given to location based privacy. While consumers are generally assured that any personal information they provide (such as name, address, or financial information) will be protected by the collecting agency, non-personal data (such as IP address or patterns of use) are often considered anonymous, and thus consumers are informed that they may be shared with other agencies, or released in aggregated forms. While this type of protection may be sufficient for static data, it becomes more problematic if location data such as origins, destinations, or travel paths are not specifically defined as personal or anonymous. As shown in the literature review above, even “anonymous” travel data may be mined or analyzed in such a way that home and work addresses, among other locations, may be defined within a fairly accurate parameter. If data collected via the use of ITS or LBS technologies are treated as anonymous data, they may be subject to lesser degrees of privacy protection, thus

opening up the potential for misuse or loss of anonymity. The overall lack of policies specific to the treatment of these data is worrying, as it is likely that without specific guidelines directing appropriate uses, the minimal amount of protection will be afforded. Thus, a key finding of this study is that current policies are lacking in their treatment of location specific data.

A second key finding is related to the overall topic frequency analysis conducted. This analysis demonstrated that there is very little consistency across privacy policies in how well they address privacy concerns as outlined by the Federal Trade Commission (FTC). In particular, policies were lacking in providing information related to how consumers may view or correct data that have been collected; what data may be shared with third parties; what procedures consumers should follow if they feel that their data have been mishandled; and issues associated with data ownership and data mining. Again, these findings indicate that there is an overall lack of comprehensiveness associated with locational privacy policies, particularly in regard to consumer expectations of protection. While consumer expectations will be more thoroughly addressed in Chapter 6 via analysis of the privacy survey, it is reasonable to believe that consumers expect basic protections of personal data such as those guaranteed by HIPAA and the FCRA. While consumers may not demonstrate explicit awareness, privacy of personal data is a general expectation as shown by court findings related to the Fourth Amendment. If existing privacy policies do not demonstrate a comprehensive reflection of the expectations of the federal government and, in turn, consumers, it may be posited that a general framework for construction of privacy policies relevant to location information should be developed, in order that consumers may develop accurate expectations regarding treatment of their data.

A third key finding is related to the differences in the content of the types of policies evaluated here; namely, private, public, ETC, and Electronic Transit Card. A lack of consistency across the different types of policies indicates that agencies and companies tend to value different types of information provided to consumers. For example, public agencies consistently address the issue of children's privacy in their policies, as mandated by federal regulation. Without this requirement, however, private agencies are far less likely to address this issue. The cluster analysis conducted revealed significant differences in how well issues of interest are addressed across policy types, with those policies related to the use of Electronic Toll Collection systems showing the lowest degree of attention paid to overall privacy issues. Discrepancies across the range of policies analyzed indicate that consumers have very little consistent protection or information on which to base their expectations of privacy in the mobile network, thus it may be inferred that service agencies are not successfully meeting their responsibilities in regard to ensuring adequate protection of privacy. These findings will also be assessed in light of consumer expectations and preferences as revealed through the survey data analyzed in the next chapter.

CHAPTER 6: SURVEY ANALYSIS

6.1 Definitions

Concepts used in this dissertation such as benefits, risk, knowledge and willingness to trade have been defined in a number of ways in privacy literature. Due to the variety of definitions used as well as differing attributes assigned to each, questions posed in the survey were designed to test components of these concepts as identified in the literature review. For purposes of this survey analysis, the following operational definitions will be used:

- Benefits: Based on the survey instrument (shown in Appendix 5), benefits are generally defined here as advantageous transportation assistance provided to users, such as cost and time savings, or safety and security advantages, which provide utility to users.
- Risk: For purposes of this analysis, risk is defined by the user as perceived degree of danger associated with the sharing of location or mobility information with a variety of agencies or organizations for a variety of purposes. Here, risk may be understood as the potential that data will be misused by those agencies or organizations with whom data have been shared.
- Knowledge: Knowledge of privacy is here defined as a factor of the degree to which respondents report that they notice, read, and/or understand privacy policies for those transportation or location-based services or applications that they use.
- Willingness to Trade: Willingness to trade is defined by respondent answers to question 10.1 of the analyzed survey, which asks respondents about their degree of willingness to trade privacy for:
    - Transportation cost benefits
    - Transportation time savings
    - Transportation safety benefits, such as crash reduction
    - Transportation security benefits, such as terrorism reduction

These definitions are created via survey questions for purposes of the dissertation, and should be understood as pertaining to the specific items of interest evaluated here.

6.1 General Demographics

425 persons began the survey, with 382 (89.9%) completions. Table X provides a general overview of demographics of those who participated in the survey.

Table X: Overview of General Survey Demographics

Item of Interest		Response Count	% Response
Gender			
	Male	171	41.2%
	Female	240	57.8%
	Prefer not to answer	4	1.0%
Income Category			
	Less than $10,000	41	9.9%
	$10,000-$19,999	50	12.0%
	$20,000-$29,999	55	13.2%
	$30,000-$49,999	82	19.7%
	$50,000-$69,999	67	16.1%
	More than $70,000	102	24.5%
	Prefer not to answer	19	4.6%
Highest Level of Education			
	8th grade or less	0	0.0%
	Some high school	0	0.0%
	High school graduate or GED	3	0.7%
	Some college	25	6.0%
	Completed 2-year college degree	8	1.9%
	Completed 4-year college degree	146	34.9%
	Master's degree	195	46.7%
	Doctoral degree	30	7.2%
	Professional degree	11	2.6%
	Prefer not to answer	0	0.0%

As seen here, persons who participated in the survey are not reflective of the overall population. Survey participants were significantly more educated and in slightly different income brackets than the general population, as shown in Tables XI and XII. Given the distribution methods used for dissemination of the survey, such findings are not surprising, as the UIC mailing lists are centered, in particular, on current students or graduates of a master's

degree in urban planning or public administration, and the social networking contacts of the surveyor are reflective of the general demographics of her social contacts.

Table XI: Educational Attainment Levels of Survey Respondents and US Population

Educational Attainment	US Population*	Survey Sample
Less than high school graduate	15.7%	0.0%
High school graduate (includes equivalency)	29.7%	0.7%
Some college or associate's degree	29.6%	7.9%
Bachelor's degree or higher	25.1%	91.4%

** Source: 2005-2009 American Community Survey (ACS) 5-Year Estimates, S1501. Educational Attainment*

Table XII: Income Categories of Survey Respondents and US Population

Income Category	Survey Population	US Population
Less than $10,000	10%	10%
$10,000-$19,999	12%	14%
$20,000-$29,999	14%	16%
$30,000-$49,999	20%	26%
$50,000-$69,999	15%	15%
More than $70,000	24%	17%
Prefer not to answer	5%	

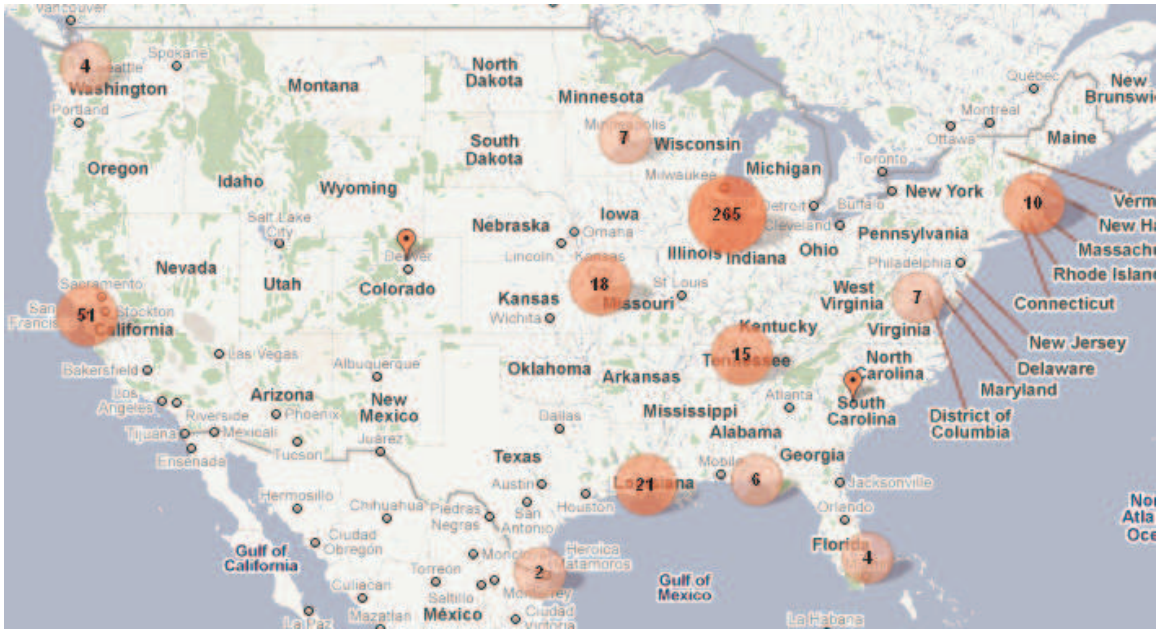
** Source: U.S. Census Bureau, Current Population Survey, 2010 Annual Social and Economic Supplement, PINC-10*

Here, we see that the survey population is both more educated and slightly wealthier than the general population, with a higher number of survey respondents reporting individual incomes of $70,000 or above than the general population.

Also of note is the geographic distribution of survey participants, shown in the map in Figure 9 below. As shown, most survey respondents (%) come from the Chicago region, with

smaller clusters seen in California, Louisiana, Missouri and Tennessee. In general, respondents tended to be overwhelmingly urban, which has implications for the survey as whole. For example, survey participants were asked to report their primary mode of transportation to work, shopping, social visits, and school. As shown in the table below, survey participants were significantly more likely to report “Public transit,” “Walking,” or “Biking” as their primary mode of transportation to work than the general US population. Such a finding is likely reflective of greater access to public transportation, biking and walking facilities within urban areas, as well as characteristics of social networking sites used to recruit survey participants (such as TheChainlink.org). Such findings are also seen in modes of transportation to other reported activities, as seen in Tables XIII and XIV, where public transit, walking and bicycling see a much higher proportion of travel than commonly seen in the US population.

Figure 9: Map of Survey Respondents



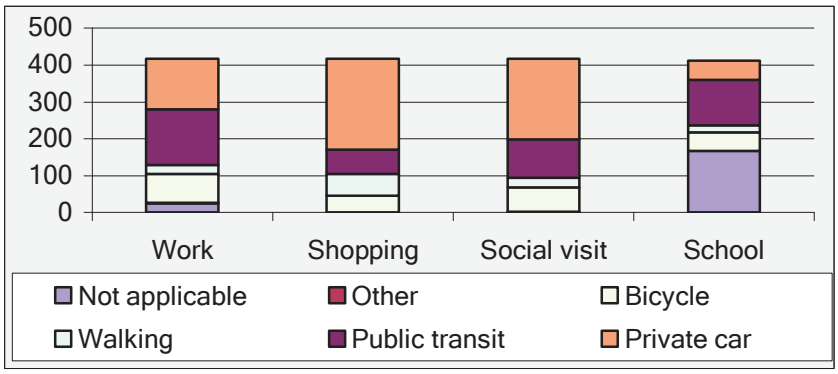
** Map Data ©2011 Europa Technologies, Geocentre Consulting, INEGI, Maplink, Tele Atlas; Google (http://www.batchgeo.com/map/025beea8f09c7f46fab763cc7b51e751)*

Table XIII: Mode of Transportation to Work

Mode of Transport to Work	US Population*	Survey Population
Private car/Car, Truck or Van	90.0%	34.9%
Public transit	5.2%	38.5%
Walking	3.0%	6.1%
Bicycle	0.5%	20.2%
Other (Taxicab, motorcycle, or other means)	1.3%	0.3%

** Source: 2005-2009 ACS 5-Year Estimates, S0801. Commuting Characteristics by Sex*

Table XIV: Mode of Transportation to Selected Activities

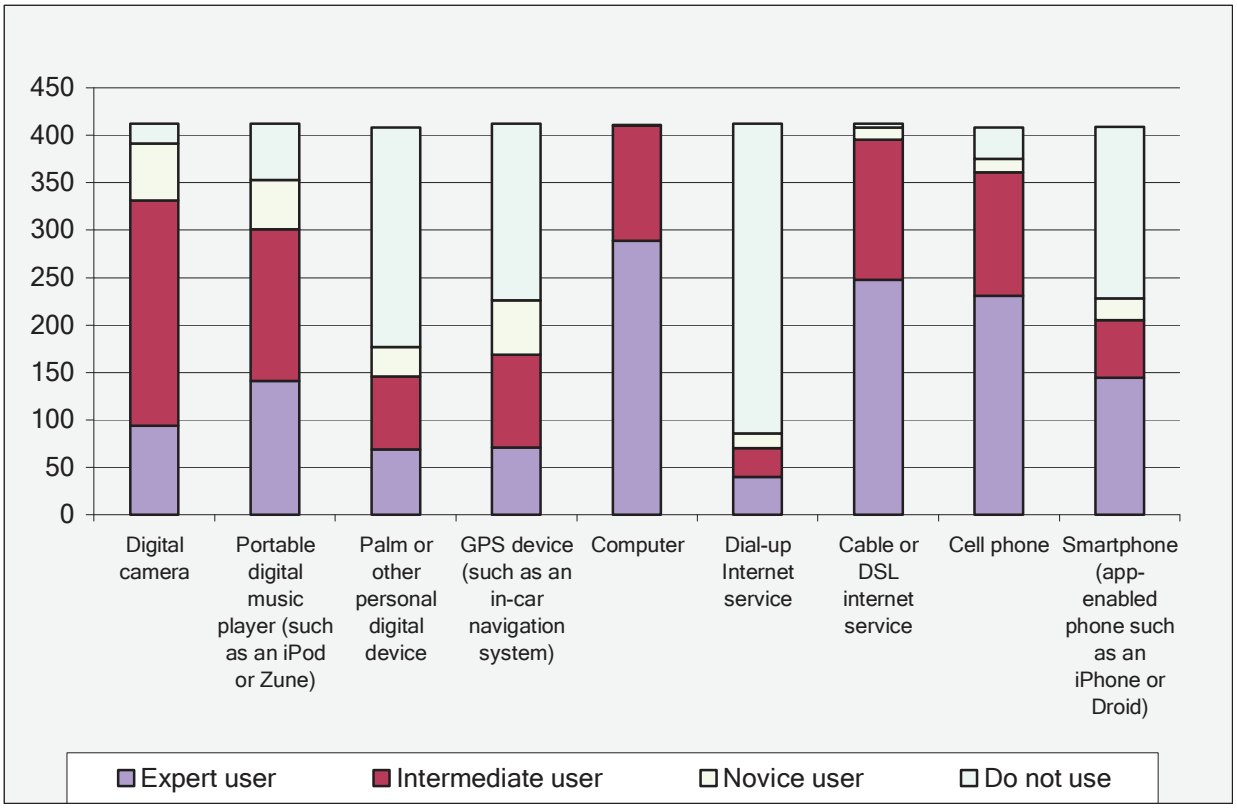


While the characteristics reported above do differ significantly from those of the general US population, it may be argued that they do reflect characteristics of those termed "first adopters," as described in Section 2.4. Bellman, *et al.* (1999) found, in a study of predictions of online buying behavior, that higher age, income, and education were somewhat predictive of propensity to make purchases online, a finding that reflects the influence of these factors on adoption of new technologies. In addition, Munnukka (2007) reports that, "Furthermore, studies by Ha and Stoel (2004), Perpermans et al. (1996), Leung (1998), Rogers (1995) and Goldsmith *et al.* (1995) collectively show that innovative consumers are in general better educated and younger than the general population, have higher incomes and occupational status, and are more often female than male." Such findings support the contention that, while the survey sample is not reflective of the general US population, it is somewhat more reflective of persons who will be more likely to be at the forefront of adoption of the types of technologies of concern to the overall thesis.

6.2 Use of Technology

A further category of interest in determining the characteristics of survey participants relates to use of technology. One question asked participants to self-report their level of expertise with various common technologies, such as computers and smartphones. Responses, seen in Table XV below, indicate that most survey participants were familiar with and considered themselves "expert" or "intermediate" users of many common personal technologies. Most respondents (nearly 92%) reported that they are "novice," "intermediate," or "expert" users of cell phones, while 55.7% (228) of respondents reported the same for smartphones (such as an iPhone or Droid). This number is significantly higher than the general US population of smartphone users (which, according to a Nielsen survey, stood at roughly 25% as of September 2010 (Dediu, 2010)). In general, survey responses indicate that participants consider themselves intermediate or expert users of the included technologies, with only dial-up internet service and Palm or other personal digital assistant (PDA) showing less than 50% adoption. For these two technologies, however, it is likely that survey respondents use other forms of technology, such as DSL or other high-speed Internet services, to perform the intended functions.

Table XV: Use of General Technology



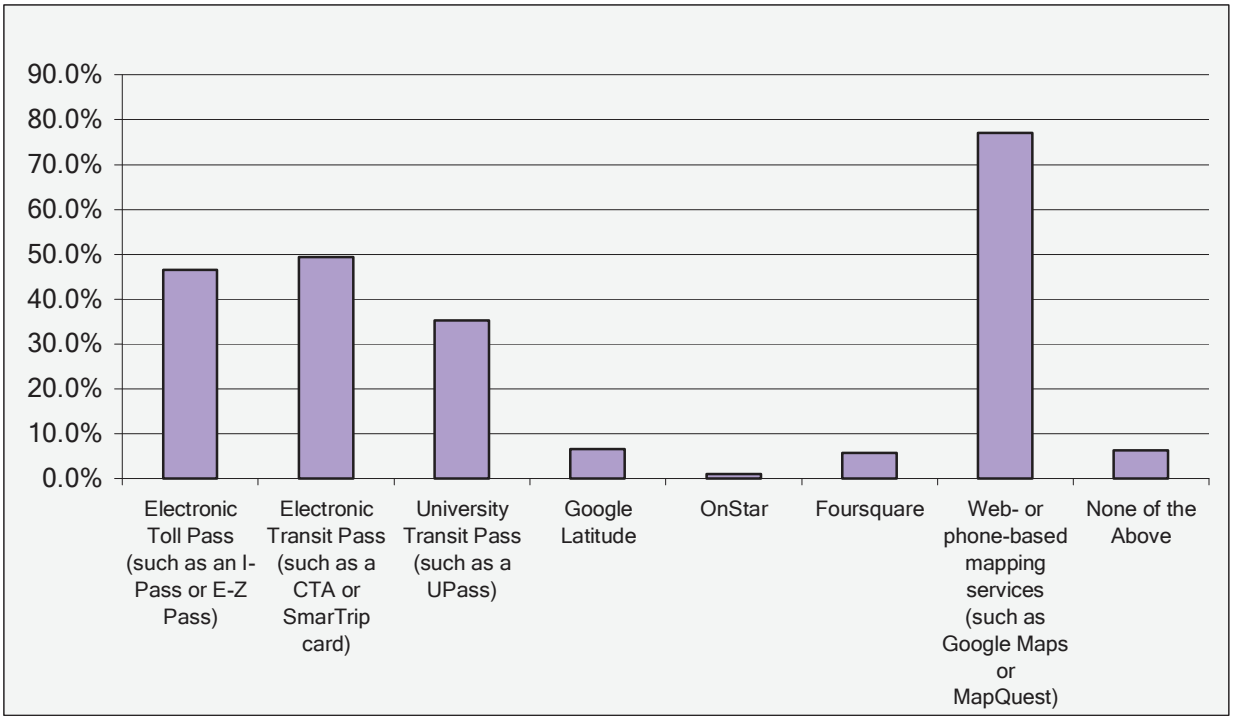
Respondents were also queried about use of a limited number of transportation or mobile technologies, including technologies aimed at payment (including electronic toll collection (ETC) passes, transit passes, and university passes), social networking (Foursquare and Google Latitude), safety (OnStar), and navigation (OnStar and web- or phone-based mapping services). Use of these technologies is significantly lower than that of the common, stationary technologies evaluated above. This may be due to a number of factors, including, but not limited to:

- Lack of necessity of some technologies depending upon travel behavior (for example, persons who do not utilize public transit would not need to use transit passes, while those who rarely or never drive would find ETC passes unnecessary)

- Cost of technology
- Novelty of technology
- Lack of easy access to technology (for example, persons who do not use smartphones may be less likely to use a service such as Foursquare or Google Latitude, as it would require application from a less easily-accessible technology)
- Lack of interest in technological services.

Only the use of web- or phone-based mapping services, such as Google Maps, showed usage by more than 50% of respondents, a finding that may be related to the relative ease of access from either a stationary (such as a personal computer) or mobile (such as a smartphone) technology.

Table XVI: Use of Transportation or Mobile Technologies



6.3 Attitudes and Actions Regarding Privacy

Respondents were next asked to report on how often they read or skim Terms of Use or Service when they use the above mentioned transportation or mobile technologies. Of those persons

who reported that they use these services, most reported that they “never” or “rarely” read the Terms of Use or Service. Of note is that persons reported a slightly higher rate of reading these Terms for those services provided by public or semi-public providers (Electronic Toll Passes, Electronic Transit Passes, and University passes) than those services offered by private providers. Each of the three public services evaluated showed a readership of over 50%, while none of the privately provided services showed this amount. Findings were similar in response to the question, “How often do you notice if there is a privacy policy before using the following types of services (For example, "Your privacy is important to Company X; maintaining your trust is important to us.")?” In this case, slightly fewer respondents reported noticing the presence of a privacy policy when using University Transit Pass, while more reported noting the presence when using a web- or phone-based mapping service. In general, a significant majority of respondents reported that they “never” or “rarely” read or skim Terms of Use or Service or notice the presence of a privacy policy.

Table XVII: How Often Respondents Read or Skim Terms of Use/Service

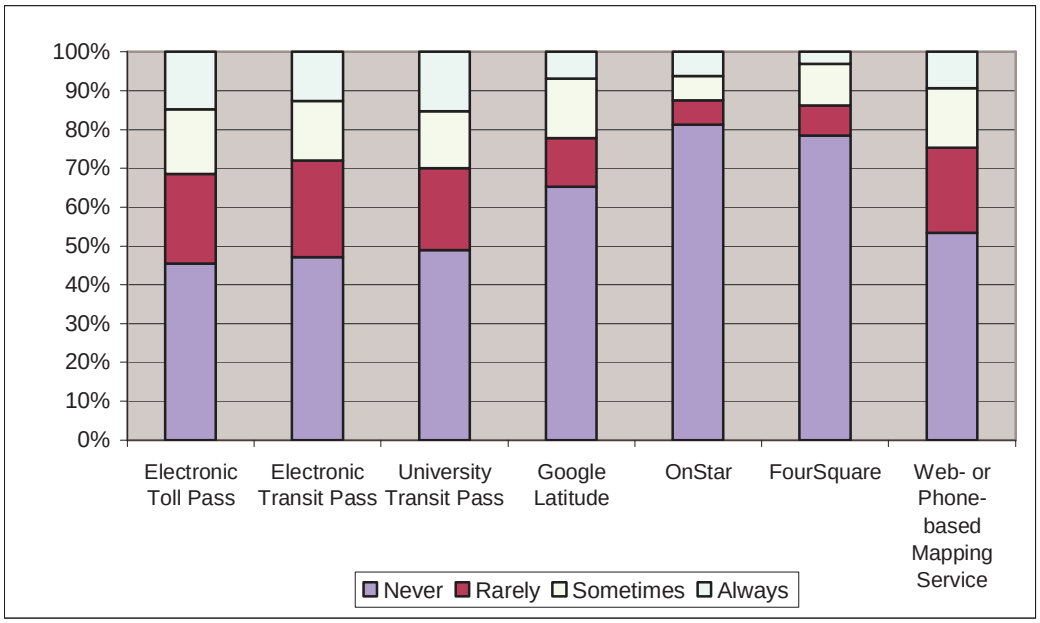
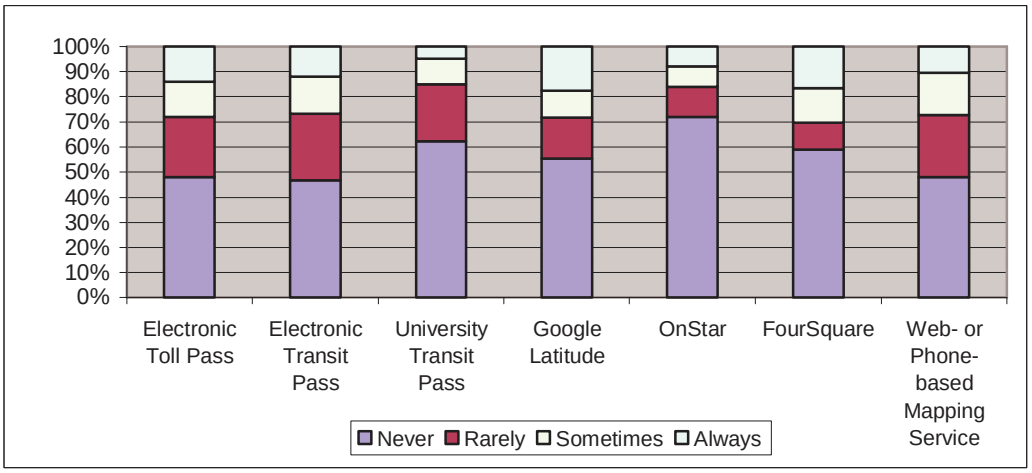


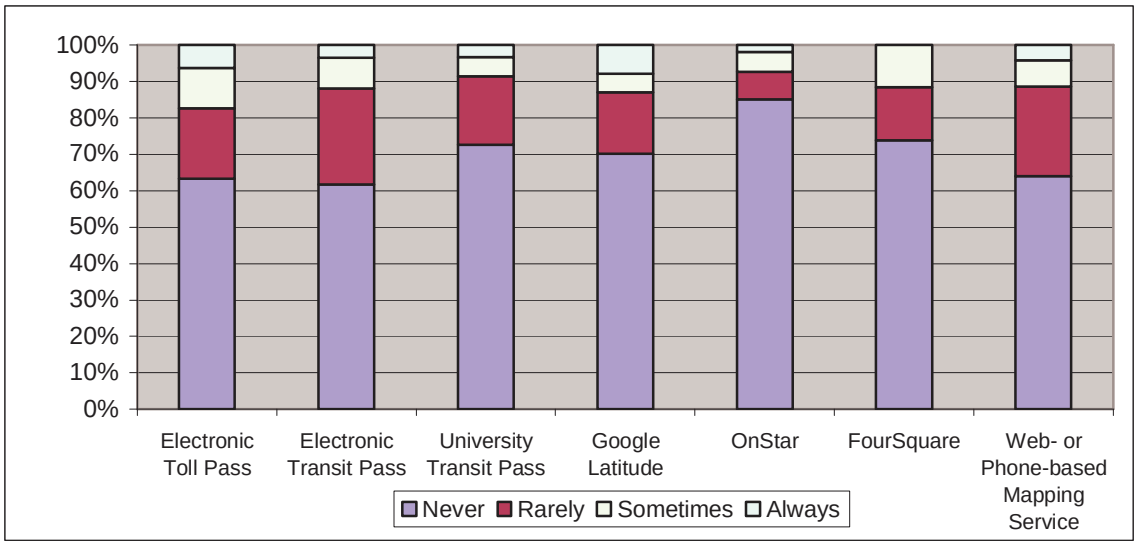
Table XVIII: How Often Respondents Notice Presence of a Privacy Policy



The findings reported above also hold when evaluating how often people read provided privacy policies or guidelines. As shown in Table XIX below, most people responded that they "Never" or "Rarely" read the privacy policies of the evaluated services. Again, levels of reading privacy policies tend to be slightly higher for publicly provided services than for those offered by

private providers, with the exception of University Transit Passes and web-based services such as Google maps.

Table XIX: How Often Respondents Read Privacy Policies Before Using Various Transportation Services



Participants were then asked to respond to a series of questions regarding their beliefs about the potential risk experienced in a number of scenarios. These scenarios were designed to test responses based on privacy risks related to data gathering, use and sharing by public and private agencies in terms of economics, travel efficiency, and law enforcement. The following table shows the responses obtained to these questions, with replies ranging from “Strongly Disagree” to “Strongly Agree”.

Table XX: Reported Perceptions of Privacy Risk

Please indicate the degree to which you agree or disagree that the following actions will place your privacy at risk:					
Answer Options	**Strongly disagree**	**Disagree**	**Neutral**	**Agree**	**Strongly agree**
Having your location or travel data collected and stored by a private company (such as Google)	6.2%	17.6%	25.2%	34.9%	16.1%
Having your location or travel data collected and stored by a public agency (such as a transit agency)	7.4%	18.3%	30.0%	29.5%	14.9%
Sharing location or travel data with friends via an application such as Google Latitude or Foursquare	4.0%	12.2%	23.4%	40.4%	20.0%
Having your location or travel data shared for marketing purposes	8.2%	10.4%	11.7%	30.3%	39.5%
Having your location or travel data shared for legal purposes	6.5%	13.0%	20.1%	32.3%	28.1%
Having your location or travel data shared for purposes of transportation efficiency (such as providing real-time traffic data or alternate routes)	11.7%	32.8%	24.6%	19.1%	11.9%
Sharing identity and financial information for travel purposes (such as electronic toll collection)	8.5%	17.5%	23.8%	29.3%	21.0%
Having location or travel information gathered by a private company (such as Google, OnStar, or Orbitz) shared with law enforcement agencies after a warrant has been issued	6.2%	15.7%	21.4%	29.7%	26.9%
Having location or travel information gathered by a private company (such as Google, OnStar, or Orbitz) shared with law enforcement agencies with no warrant issued	8.8%	9.0%	10.8%	21.3%	50.1%
Having location or travel information gathered by a public agency (such as your state Department of Transportation) shared with law enforcement agencies after a warrant has been issued	7.7%	16.9%	22.9%	26.6%	25.9%
Having location or travel information gathered by a public agency (such as your state Department of Transportation) shared with law enforcement agencies with no warrant issued	8.2%	9.7%	14.2%	23.9%	43.9%

Findings here indicate that people believe that most of the scenarios stated have some possibility of putting one's privacy at risk. In all cases but "Having your location or travel data collected and stored by a public agency" and "Having your location and travel data shared for purposes of transportation efficiency," 50% or more respondents replied that they "Agree" or "Strongly Agree" that the scenario in question will place one's privacy at risk. The strongest responses regarding the potential for privacy risk involved the sharing of collected data with

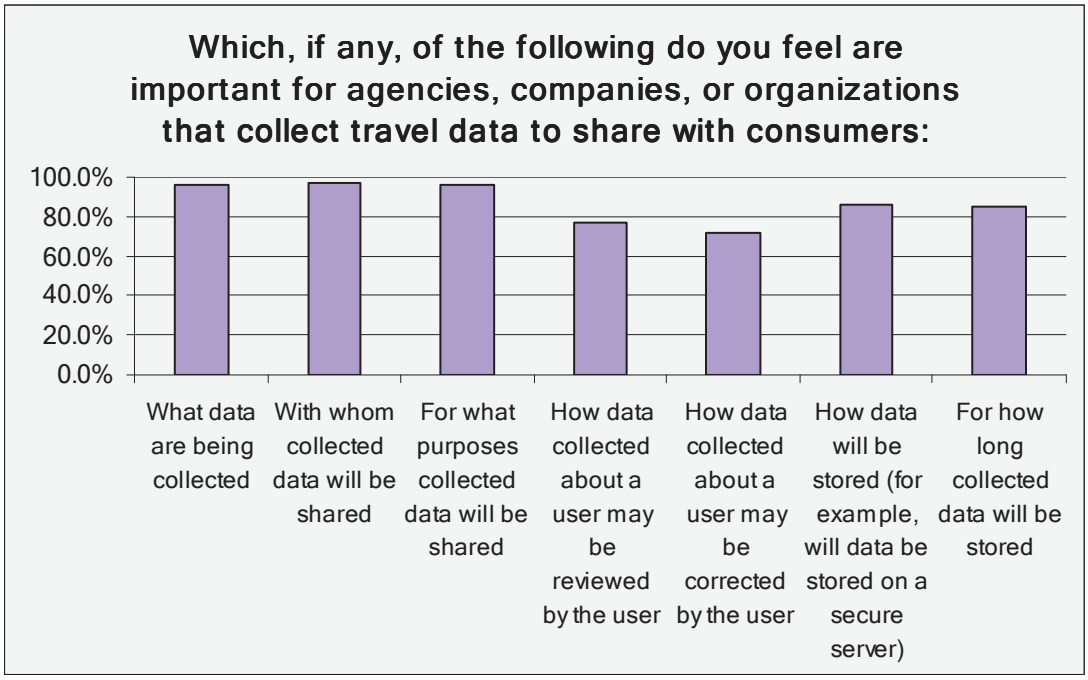
law enforcement agencies with no warrant issued. This finding was especially strong for private agencies, with over 50% of respondents replying that they “Strongly Agreed” that this would put privacy at risk. Responses were weakest for the privacy risk of having location or travel data shared for purposes of transportation efficiency, with nearly 45% of respondents indicating that they disagreed or strongly disagreed that such an action would place privacy at risk. Sharing data for marketing purposes also produced a strong response, with nearly 70% of respondents indicating that they agreed or strongly agreed that such actions would place their privacy at risk. The findings represented here indicate that people generally believe that sharing travel and transportation related data has the potential to place their privacy at risk, whether such data are shared and used by public or private agencies.

These findings are interesting when reviewed in the context of the privacy policy analysis conducted above. As noted in that analysis, privacy policies were remiss in providing information to consumers related to sharing of data with third parties and law enforcement, which would indicate, according to the findings shown in Table XX, that privacy policies are not addressing areas of risk identified by consumers. In particular, consumers demonstrate high perceptions of privacy risk relative to the sharing of data for marketing purposes and for law enforcement; however, many of the privacy policies analyzed do not address how these data will be shared, with what agencies, or under what circumstances. This lack of information places societal judgment of privacy expectations at risk, as it creates a gap in the knowledge base of consumers and does not provide adequate information for consumers to effectively evaluate the balance of benefit and risk. While it is true, as shown, that many consumers do not read privacy policies, in the event of perceived misuse or harm due to the release of data for

purposes regarded as “risky” by the consumer, having an established policy for proper use of data relative to perceived harm allows the consumer to feel that he has some control over how to indicate his displeasure and/or receive redress. In the absence of such control, the consumer may feel incapable of having the matter addressed adequately, thus heightening fears of privacy loss and leading to less willingness to participate in evolving mobile information systems.

Participants were next asked to indicate which, if any, of certain types of information they feel are important for agencies, companies, or organizations that collect travel data to share with consumers. Table XXI below shows the obtained responses.

Table XXI: Importance of Sharing of Travel Data



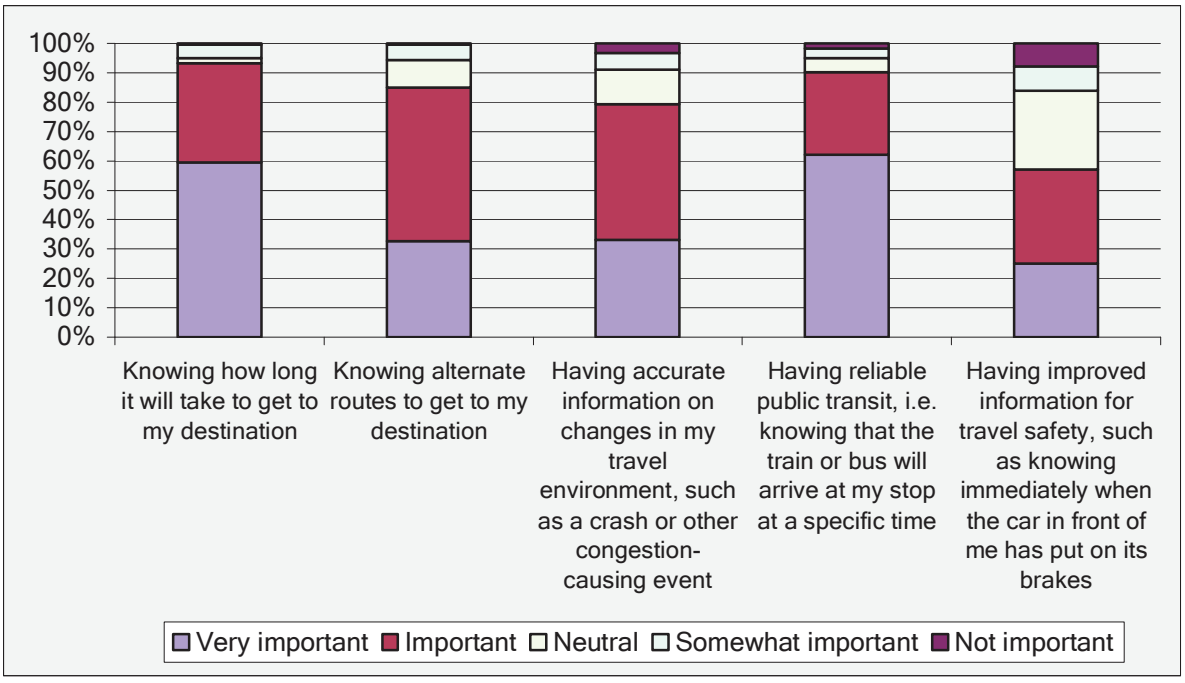
Findings indicate that consumers believe that a wide variety of data should be shared with consumers, with the strongest responses relating to the types of data being collected, with whom these data will be shared, and for what purposes. Weakest, though still strong, responses were shown for the provision of information regarding how collected data may be reviewed and corrected by the user. These findings are interesting in the context of the privacy policies reviewed above. While there is some overlap between consumer expectations and information provided in these policies (such as fairly strong coverage of the types of data to be collected), the policy analysis undertaken in Chapter 5 indicates that current privacy policies do not adequately address the types of information that consumers expect should be shared, in particular, information relating to the sharing and storage of data, as well as opportunities for review and correction. If we operate under the "reasonable expectations" framework, current privacy practices do not accurately reflect the public's expectations, and thus may be viewed as inadequate for the protection of personal information.

The strong results shown here are also interesting when reviewed in the context of the findings above regarding how often respondents review privacy policies, and how much they feel that certain agency or organization actions will place their privacy at risk. As shown above, the majority of respondents "never" or "rarely" read privacy policies or agreements prior to use of ITS or LBS services, yet they believe that certain actions by involved agencies or organizations may place their privacy at risk and that a broad spectrum of information regarding data collection, use and sharing should be provided to the consumer. Such findings may indicate a range of possibilities, including, (1) that consumers believe that the information they feel *should* be included in privacy policies *are* included in privacy policies; (2) that consumers are

concerned about privacy in the mobile environment but do not understand privacy policies as presented and so do not read them; or (3) that consumers do not believe that privacy policies adequately address relevant concerns, and thus simply agree to them. Because no question was asked regarding why privacy policies are not read if they are not, this question cannot be answered at this time.

In terms of expected benefits of transportation technologies, respondents were asked to rate how important they find a number of types of transportation information. Table XXII below indicates the responses and ratings.

Table XXII: Reported Importance of Transportation Information to Survey Respondents

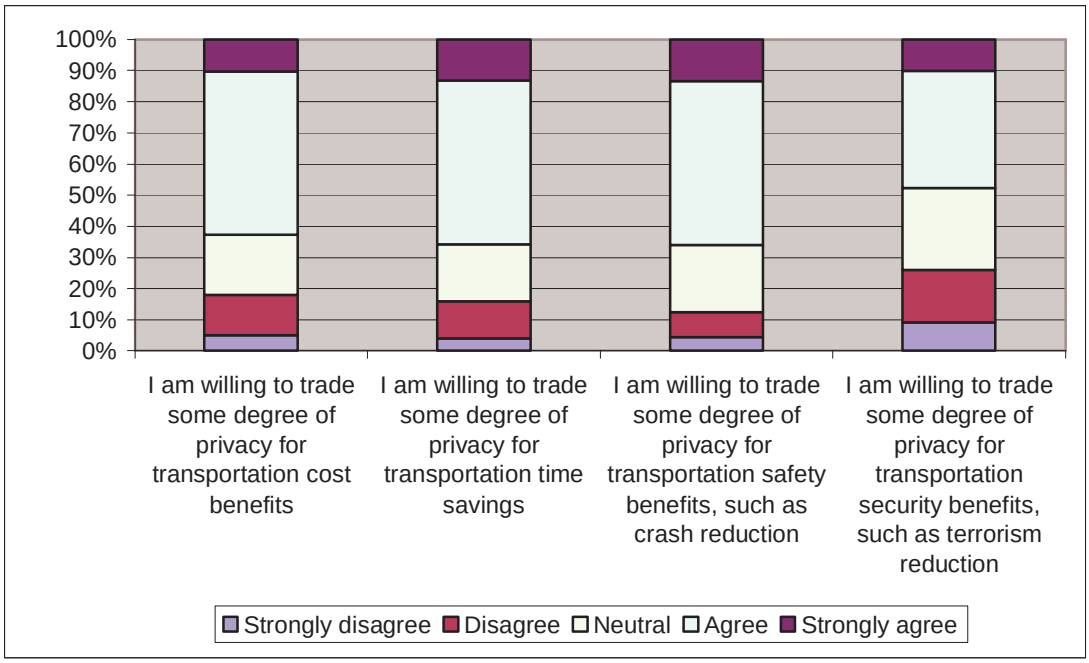


From the results obtained here, the most important information identified by respondents relate to reliability and assurance. While most information tested received at least 75% of

responses indicating that the information queried was “Very Important” or “Important”, information related to advanced safety features, such as braking information, which is a key factor of proposed vehicle-to-vehicle ITS applications, received only 57% response rates in these categories. This may relate to unfamiliarity with these technologies, or it may simply reflect that persons feel fairly well in control of their ability to react appropriately to conditions on the roadway.

A related question referring to willingness to trade travel information for transportation benefits was also asked, resulting in the findings shown in Table XXIII.

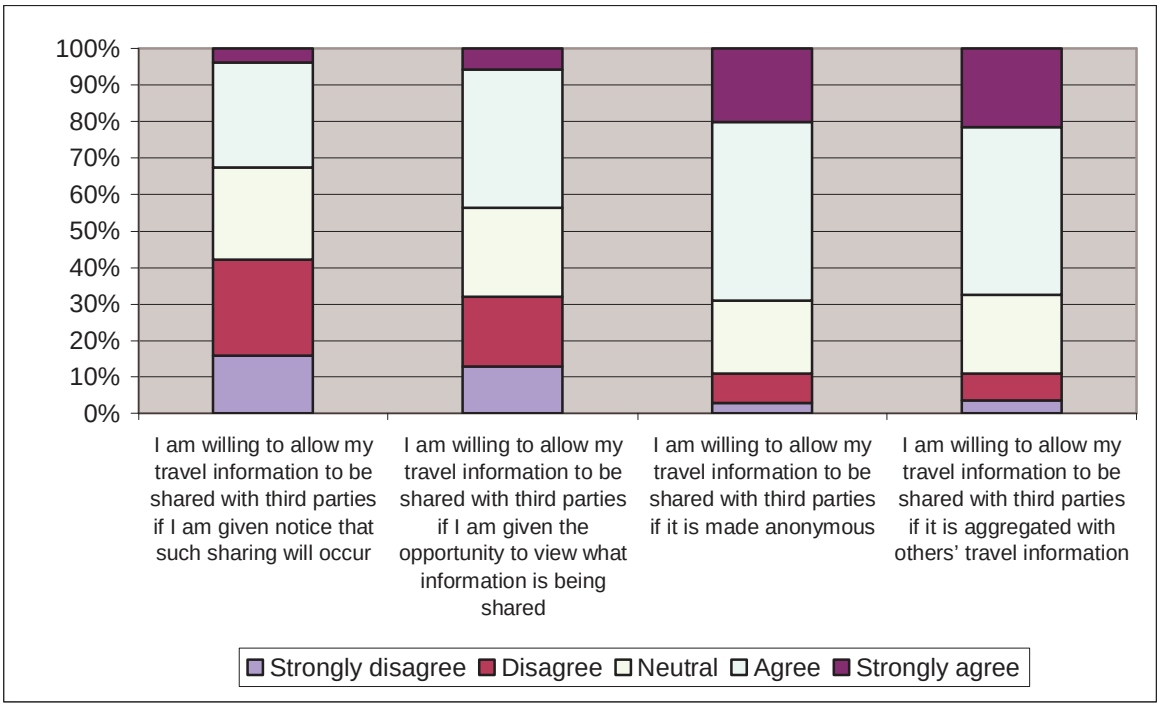
Table XXIII: Reported Willingness to Trade Privacy for Transportation Benefits



Findings across the first three categories – cost benefits, time savings, and safety benefits – were fairly consistent, while respondents appeared slightly less concerned about transportation

security benefits. Participants were also asked to indicate how willing they would be to share their travel information with third parties given certain caveats. This question provided the results shown in Table XXIV.

Table XXIV: Reported Willingness to Share Data with Third Parties Under Given Conditions



As seen here, travelers were most likely to agree that they would share their travel information with third parties if the data were made anonymous or were aggregated with other's travel information. Far fewer agreed that they would be willing to share their data if notice were given or if they knew what information was being shared. This finding indicates that anonymity may be the most important deciding factor for consumers when deciding whether or not to share travel information, regardless of whether they are notified or not.

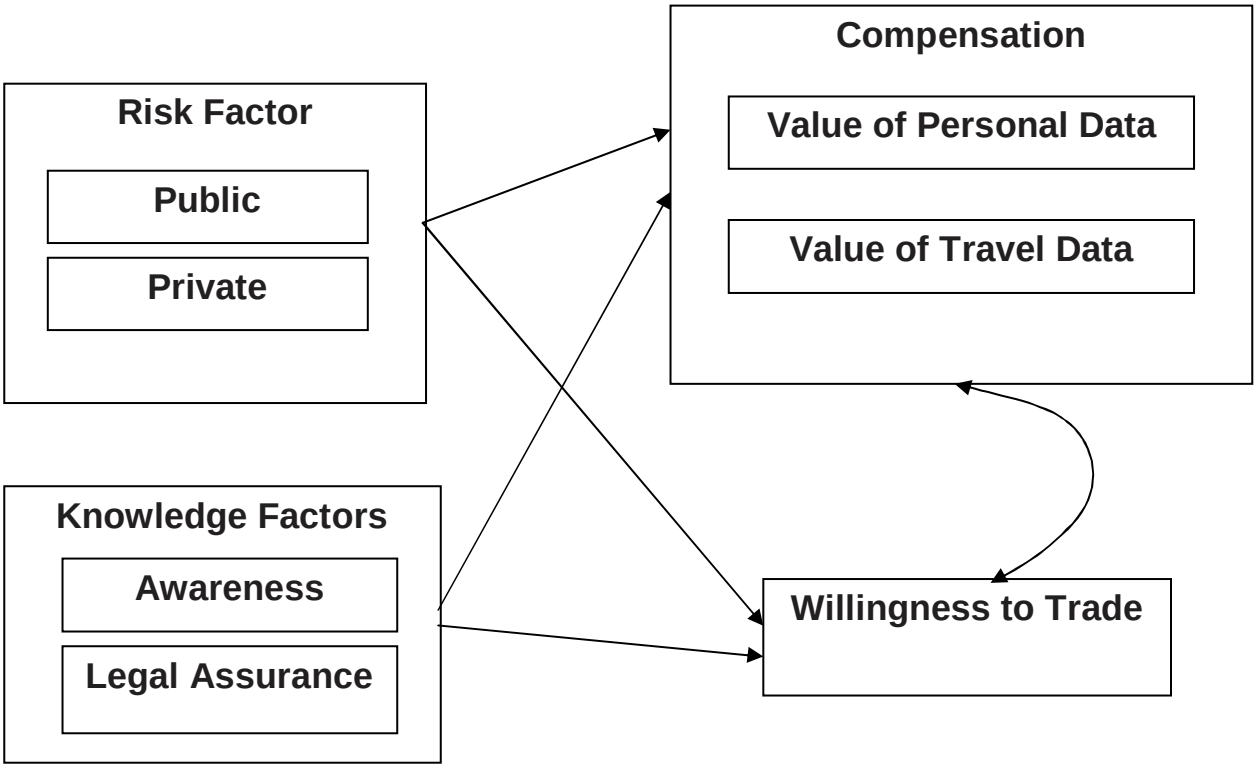
The brief analysis above indicates that privacy is of concern amongst consumers in the mobile environment. While adoption and use of certain general technologies, such as computers, cell phones and GPS-enabled devices, and certain mobile technologies, such as toll passes, transit passes, and web-based mapping tools, are shown to be quite high, question responses have also indicated a degree of concern related to privacy risk implications of use of transportation technologies. The following section will expand upon these findings, and look more closely at the trade-offs that consumers are willing to make between provision of personally identifying information related to transportation and potential mobility benefits.

6.4 Detailed Statistical Analysis

The above discussion provides a general overview of results obtained from the privacy survey. In addition to this analysis, however, we must next look at how respondents valued the trade-offs between sharing of personal data and concomitant transportation benefits. A number of methods were evaluated for their applicability to the data set and questions of interest. In part, the issue at hand is that “privacy” is not a static concept. As noted above, contextual and societal constructs will play a role in the desirability of the protection of private data dependent upon assessment of likely benefits, agencies or organizations involved in the management and use of data, and personal preferences relative to these factors. As such, simple regression analysis alone will not be adequate for evaluating survey results. As shown in Figure 10 below, the constructs of interest are varied and interrelated; thus, an approach should be used that will enable evaluation of these interrelationships. A full data dictionary indicating variables,

variable names and descriptions may be found in Appendix 1, and overall variable means and response type (ordinal, cardinal, dummy, or categorical) may be found in Appendix 2.

Figure 10: Conceptual Model of Location Privacy



A number of methods for conducting survey analysis were researched and evaluated. Based on the survey design and questions of interest, an approach using initial regression analysis followed by structural equation modeling (SEM, described below) in the context of privacy calculus were eventually selected for further analysis.

6.4.1 Data Pre-Processing

As noted in the literature review above, privacy preferences tend to vary from person to person, though various researchers (notably Westin, 2003 and Junglas and Spitzmüller, 2005) have noted that certain characteristics may indicate a greater or lesser preference for privacy. Westin, as noted above, has proposed three categories of consumer privacy sensitivity, namely, privacy fundamentalists, privacy unconcerned and privacy pragmatists. For purposes of this analysis, survey respondents have been clustered into three categories, following Westin's model and using the following as clustering characteristics:

- Privacy risk assessment
- Expected compensation for personal and travel data

Due to perceived similarities in ratings of privacy risk reported by respondents, a correlation matrix was developed for the perceived risk categories tested in the survey. Strong correlations were found for nearly all tested pairs, thus these factors were collapsed into one primary category. To perform this collapse, factor scores were added and then divided by 11 (the number of factors tested) to create an average score for privacy risk perception. No strong differences were seen in questions related to public or private risks, thus no differentiation was made based on these factors. In much the same way, correlations were checked for desired compensation for elements of personal data related to travel. Two factors (name and address) showed significant differences from other factors tested (vehicle information, starting point of a trip, ending point of a trip, time of day at which trips are made, and trip route and time of day), thus these were collapsed into two categories reflecting compensation for personal data and

compensation for transportation data. These two factors were then used to cluster respondents in the context of privacy risk assessment.

SAS statistical software was used to cluster the respondents. Two small (consisting of roughly 11% and 18%) and one large (consisting of the remaining 71%) clusters were developed using Proc ACECLUS and Proc FASTCLUS. According to The SAS Institute (1999), "The ACECLUS (Approximate Covariance Estimation for CLUStering) procedure obtains approximate estimates of the pooled within-cluster covariance matrix when the clusters are assumed to be multivariate normal with equal covariance matrices," while FASTCLUS, "finds disjoint clusters of observations using a k-means method applied to coordinate data." Table XXV provides variable descriptions, while Table XXVI shows median responses for each of the clustering elements for each cluster.

Table XXV: Descriptions of Population Clustering Variables for Privacy Preferences

Characteristic	Description	Score Description
Compensation for personal data	Average compensation required to share name and address information	1=$0.00 - $0.10
		2=$0.11 - $0.25
		3=$0.26 - $0.50
		4=$0.51 - $1.00
		5=$1.01 - $5.00
		6=>$5.00
		7=Would not sell
Compensation for travel data	Average compensation required to share travel data, including vehicle information, starting point of a trip, ending point of a trip, time of day at which trips are made, and trip route and time of day	Same as above
Average perceived risk	Composite variable composed of scores assigned to risk factors of various forms of sharing and use of data by public and private organizations	1=Strongly disagree
		2=Disagree
		3=Neutral
		4=Agree
		5=Strongly agree

Table XXVI: Median Response Rates for Privacy Clustering Characteristics

Variable by cluster	Statistical Results				
	N	Mean	Std Dev	Minimum	Maximum
Cluster 1					
Compensation for personal data	287	6.803	0.406	5.5	7
Compensation for travel data	287	3.860	2.199	0	7
Average perceived risk	287	3.514	0.855	0.545455	5
Cluster 2					
Compensation for personal data	45	4.000	0.584	2.5	5
Compensation for travel data	45	2.311	1.533	1	6.6
Average perceived risk	45	3.547	0.709	1.454546	5
Cluster 3					
Compensation for personal data	74	0.473	0.579	0	2
Compensation for travel data	74	0.473	0.580	0	2
Average perceived risk	74	2.424	1.721	0	5

The clustering technique, as reflected in Table XXVI, showed fairly distinct clusters. A regression analysis was conducted to determine the validity of the clustering technique, resulting in an adjusted r^2 of .9646, indicating strong correlation of clusters. Average perceived risk and compensation for personal data were shown to have the strongest correlations, thus these two factors most likely add most to the clustering. One item of note is that, in contrast to Westin's analyses, which generally reflect that most persons fall into the "privacy pragmatist" category, respondents to the current survey who reported high degrees of perceived risk and who required fairly high degrees of compensation for or unwillingness to sell personal information formed the largest cluster. This may indicate that respondents to the current survey have a higher perception of privacy concern than that of the general population. The following analysis will use these clusters to better examine how different segments of the population respond to further questions about privacy and travel preferences.

6.4.2 Description of Structural Equation Modeling (SEM)

The theory of "privacy calculus," described above, indicates that privacy is not a static concept, but rather a decision process subject to various inputs and contexts. A number of researchers (Culnan and Armstrong, 1999; Dinev and Hart, 2006) have examined the impacts of privacy calculus on online behavior, but have not specifically addressed the decision making process in a mobile environment. This research will use the privacy calculus model shown above to develop and test hypotheses related to the context of travel, potential benefits expected, privacy concerns, and overall willingness to share data. This model identifies interrelationships between constructs of interest, as described below.

Table XXVII: Constructs of Interest in Determining Privacy Preferences and Trade-offs

Construct Category	Elements	Acronym	Description
Context	Public Private	CO	Context in which data gathering or sharing takes place - collectors and users may be public agencies or private providers
Compensation	Economics Safety Efficiency	BEN	Benefits conferred by sharing of data
Risk	Perceived risk Value of personal data Privacy concerns	PI	Subjective issues that may contribute to a person's willingness or unwillingness to share data and participate in transportation programs
Knowledge	Awareness Legal assurance	TF	Issues that may impact a person's trust relative to the sharing of his or her data with transportation service providers
Willingness to Trade	Willingness to Share	WTS	Indicates a person's willingness to share data

Since many of the questions asked in the survey related to individual constructs shown above, Principal Component Analysis (PCA) was used to identify the primary components underlying the data related to these factors in order to reduce the number of variables of interest. SAS Proc Factor, which performs common factor and component analyses with rotations, was first used to conduct a principal axis method, followed by a varimax (orthogonal) rotation, which

imposes a restriction disallowing correlation of factors. Four factors were retained for rotation, and factor loadings are shown in Table XXVIII below. Though some factor loadings are somewhat low, each one is significantly different enough from each other loading that all variables have been retained.

Table XXVIII: Rotated Factor Loadings of Constructs of Interest

Variable	Description	Factor1	Factor2	Factor3	Factor 4
RiskPriv	Risk perceived by having location or travel data collected and stored by a private company	0.00	0.82	-0.10	0.06
RiskPub	Risk perceived by having location or travel data collected and stored by a public agency	0.05	0.73	-0.15	0.13
UseTTech	Use of transportation technology	-0.12	-0.05	0.09	0.06
TrCost	Willingness to trade some degree of privacy for transportation cost benefits	-0.07	-0.13	0.89	-0.28
TrSafety	Willingness to trade some degree of privacy for transportation time savings	-0.01	-0.15	0.55	-0.29
Tr3Anon	WIllingness to allow travel information to be shared with third parties if it is made anonymous	0.02	-0.09	0.20	-0.36
CmpPer	Compensation required to share personal data	0.02	0.02	-0.08	0.56
CmpTrvl	Compensation required to share travel data	0.03	0.08	-0.15	0.73
ReadPPApp	Reads privacy policy prior to adopting a new location service or application	0.62	0.03	-0.07	0.19
Understnd	Generally understands privacy policies	0.97	0.08	-0.03	0.07
Comfort	Degree of comfort reported with privacy protection offered by the providers of location services or applications	0.36	-0.18	0.13	-0.18
Overall variance explained by each factor:		**1.49**	**1.30**	**1.23**	**1.22**

Three variables relating to actions and attitudes towards privacy policies were found to load on Factor 1: (1) Reads privacy policies prior to adopting new location services or applications, (2) Perceived degree of general understanding of privacy policies, and (3) Perceived comfort regarding privacy protections. This factor will be referred to as the Knowledge Factor (KF). Two variables were found to load on Factor 2: (1) Risk perceived by having location data collected and stored by a private company, and (2) Risk perceived by having location data collected and stored by a public organization. This factor will be referred to

as the Risk Factor (RF). Factor 3 had four variables displaying load, namely: (1) Willingness to trade some degree of privacy for transportation cost benefits, (2) Willingness to trade some degree of privacy for transportation time savings, (3) Willingness to allow travel information to be shared with third parties if it is made anonymous, and (4) Use of transportation technologies. This factor will be referred to as Willingness to Trade (WT). Finally, Factor 4 had two loading items, namely: (1) Compensation required to share personal data, and (2) Compensation required to share travel data. This factor will be referred to as the Compensation Factor (CF).

Next, a regression analysis was conducted to determine how well the clusters identified earlier correlated with the identified factors. An adjusted R^2 of 0.4564 was obtained, indicating a reasonably good fit, given the constraints on the parameters tested. Next, factors were correlated with additional survey data within the identified clusters to test the following hypotheses:

- Hypothesis 1: Relationships will be seen between WTTP and self-reported personality characteristics.
- Hypothesis 2: Survey respondents will demonstrate a relationship between willingness to trade privacy (WTTP) and perceived utility of transportation benefits.
- Hypothesis 3: Correlations will be seen between WTTP and perceived degree of expected compensation.
- Hypothesis 4: Persons will display a relationship between WTTP and risk due to perceptions of perceived privacy loss.
- Hypothesis 5: Correlations will be seen between WTTP and extent of knowledge and concern related to privacy matters, as demonstrated by the reading and understanding of privacy policies.
- Hypothesis 6: Respondents will be more willing to reduce compensation for trading private information for purposes of safety or efficiency than for cost savings.
- Hypothesis 7: Respondents will require higher compensation if they are aware that their data will be used for commercial purposes.

The following section will review the findings associated with these hypotheses.

6.4.3 Hypothesis testing

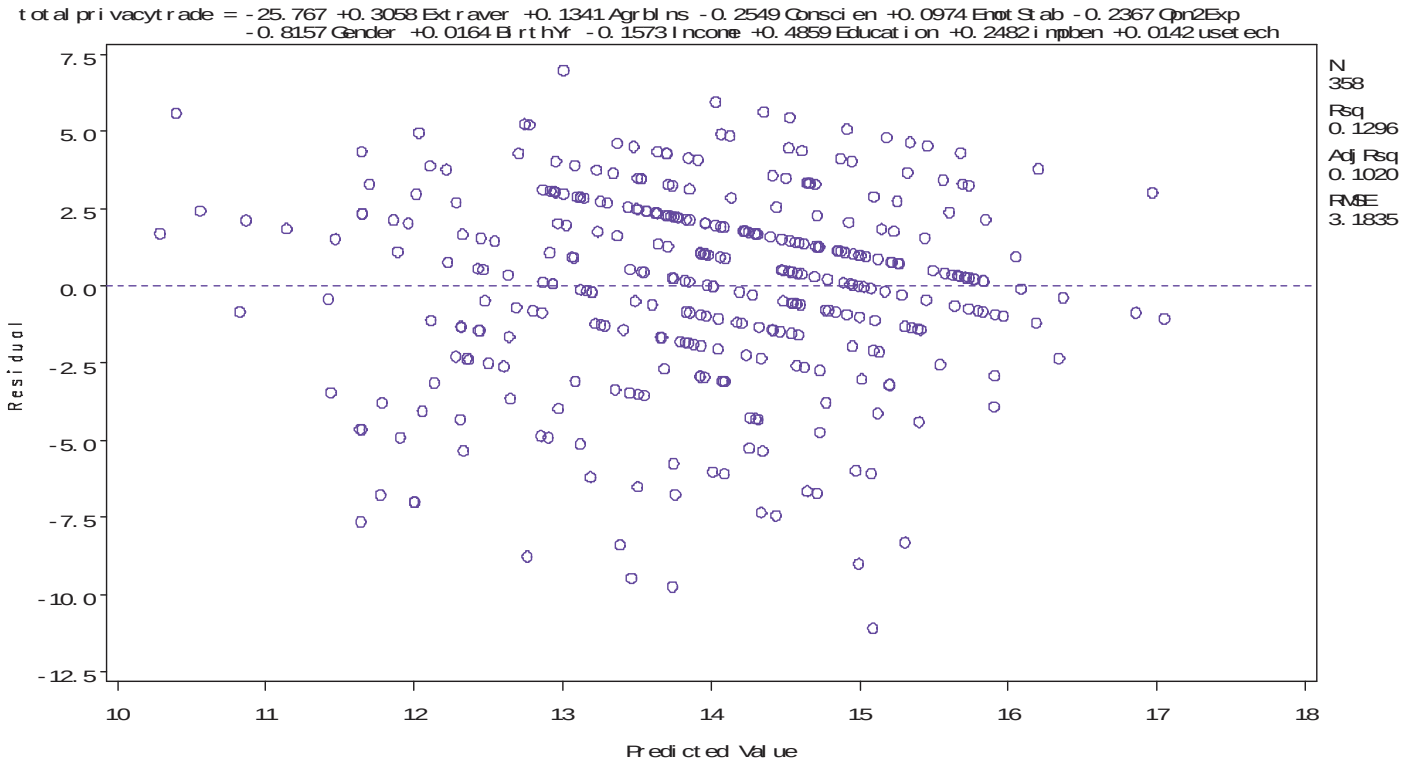
Hypothesis 1: Relationships will be seen between WTTP and self-reported personality characteristics.

Relationships between privacy preserving characteristics and personality characteristics have been analyzed as shown above in Section 2.4. Here, we are interested in evaluating how those traits may play out in the relationship between a person's characteristics regarding personality and lifestyle in regard to willingness to trade information. First, data were tested for heteroscedasticity, or non-constant standard deviations of a variable. A visual plot of the following regression was run:

$$y(totalprivacytrade) = \beta_1(extraversion) + \beta_2(agreeableness) + \beta_3(conscientiousness) + \beta_4(emotional\_stability) + \beta_5(openness\_to\_\exp erience) + \beta_6(gender) + \beta_7(birth\_year) + \beta_8(income) + \beta_9(education) + \beta_{10}(transportation\_\inf ormation) + \beta_{11}(use\_of\_techno\log y) + \varepsilon_i$$

yielding the plot of residual versus predicted values shown in Figure 11.

Figure 11: Plot of Residual versus Predicted Values for Willingness to Trade Privacy



Mild heteroscedasticity was observed, and confirmed via the White test, which resulted in a P-value of 0.0842. While mild, it was determined to use a weighted least squares (WLS) model in order to correct for this heteroscedasticity.

In WLS models, each term in the model includes an additional weight, ω_i, that determines how much each observation in the data set influences the final parameter estimate. Here, a WLS regression model was run to evaluate the personality characteristics of Extraversion (*extraver*), Agreeableness (*agrblns*), Conscientiousness (*conscien*), Emotional Stability (*emotstab*), and Openness to Experiences (*opn2exp*); demographic traits, including education (*education*), income (*income*), gender (*gender*) and age (*birthyr*); importance of travel information (calculated as a composite score of importance of receiving various types of travel

information (*impben*)); and comfort with technology (*usetech*) in relationship to willingness to trade information. The WLS regression resulted in the findings shown in Table XXIX.

Table XXIX: Findings of Weighted Least Squares Regression Analysis of Willingness to Trade Privacy

Variable	Parameter Estimate	Standard Error	t Value	Pr > |t|	95% Confidence Limits	
Intercept	-36.509	39.111	-0.930	0.351	-113.434	40.417
Extraver*	0.301	0.101	2.970	0.003	0.102	0.501
Agrblns	0.068	0.153	0.440	0.657	-0.233	0.369
Conscien**	-0.268	0.130	-2.070	0.039	-0.524	-0.013
EmotStab	0.001	0.130	0.010	0.996	-0.255	0.256
Opn2Exp	-0.240	0.161	-1.490	0.137	-0.557	0.077
Gender**	-0.728	0.360	-2.020	0.044	-1.436	-0.019
BirthYr	0.023	0.020	1.160	0.249	-0.016	0.061
Income**	-0.222	0.102	-2.180	0.030	-0.422	-0.021
Education**	0.419	0.168	2.500	0.013	0.089	0.749
impben*	0.244	0.056	4.380	<.0001	0.134	0.353
usetech	0.014	0.033	0.430	0.669	-0.050	0.078

** Significant at 0.01*
***Significant at 0.05*

Six characteristics modeled (Extraversion, Conscientiousness, Gender, Education, Income and Importance of Benefits – indicated in gray in the table above) were shown as having significant influences, with Extraversion, Importance of Benefits and Education having positive influences on likelihood of willingness to share data, and conscientiousness, income, and being male (gender=1) having negative influences. This is in keeping with prior research indicating that more extraverted persons tend to have a higher degree of trust value in other persons or agencies. The influence of education on willingness to trade has a number of potential explanations, including that persons with a higher level of education may have more access to or knowledge of technologies and their potential applications. It may also indicate that more

highly educated persons feel more comfortable with sharing data with technological organizations based on degree of familiarity with general policies and procedures relevant to risk management, as they have been generally well-served by such procedures. Junglas, *et al.* (2008) have hypothesized that conscientious individuals tend to be more concerned about privacy due to their concern for the actions of others, particularly if it will impact their own experience. This hypothesis is supported by the above analysis. Finally, the importance that respondents reported with respect to various travel information (such as time of trip and reliability of transit services) likely had a positive impact on willingness to trade, as individuals who value this information highly will assign higher value to the benefits than to the cost of sharing information. While the r^2 value obtained (0.1567) was fairly low, the overall indicators evaluated met the influences of the overall hypothesis. Here, it may be determined that there are missing variables that may have influence.

While this analysis has looked at the overall respondent population, further hypothesis testing will be conducted by segmentation in the clustering categories identified above.

Hypothesis 2: Survey respondents will demonstrate a relationship between willingness to trade private information (WTTP) and perceived utility of transportation benefits.

Privacy may be a difficult parameter to estimate, as it may be measured in a number of different ways. For purposes of this dissertation, privacy is measured in terms of "willingness to trade" via use of a question asking respondents to rate how willing they are to trade certain types of information for specific benefits, as seen in Table XXIII above. Such a measurement

may be seen in the form of a privacy-utility tradeoff, as described in Krause and Horvitz (2007): "A fundamental utility-privacy trade-off exists where the more information that is acquired, the higher the utility via, e.g., personalization, but, at the same time, the greater the privacy concerns." For purposes of this dissertation, general utility is measured in terms of cost benefits, transportation time savings, transportation safety benefits, and transportation security benefits.

The hypothesized relationship between WTTP and perceived utility is first explored in terms of perceived importance of transportation information. It is hypothesized that the higher a respondent rates perceived importance of receiving transportation information benefits (such as reliable travel times and incident occurrence), the more willing he will be to trade personal information. Such a hypothesis is in keeping with results found in Phelps, *et al.* (2000) and Krause and Horvitz (2007). Importance of travel information is estimated via a question asking participants to rate the importance of various types of transportation information.

To first explore the relationship between willingness to trade information and utility of transportation benefits, a series of regression analyses were run testing the relationships between the following:

- Willingness to trade privacy for benefits (*privtrade*): A composite variable composed of the summation of degree of willingness to privacy for cost benefits, time savings, safety benefits, and security benefits
- Utility (*utility*): Composite value composed of respondent ranking of importance of knowing travel time, alternate routes, changes in the travel environment, transit reliability, and immediate safety information
- Total risk (*totrisk*): Composite value composed of degree of risk that will be incurred by sharing location and travel data with a private agency, a public agency or with friends; having location and travel data shared for purposes of marketing, legal purposes and transportation efficiency; having data shared with law enforcement agencies by public or private agencies with or without a warrant; and for purposes of electronic toll collection

- Compensation (*compen*): Composite compensation desired to share name, address, vehicle information, trip origins and destinations, time of day of trip and trip route
- Use of Location Services:
    - Use of location based services (*LtechLBS*): Respondent reports use of one or more of the following: Google Latitude, OnStar, Foursquare, and Google Maps
    - Respondent uses no location based services (*LTechNon*)
    - Use of electronic payment methods (*LTechPay*): Respondent reports use of one or more of the following: Electronic toll pass, university transit pass, electronic transit pass
    - GPS expert user (*gpsexp*)
    - GPS intermediate or novice user (*gpsintno*)
    - Smartphone expert user (*spexpert*)
    - Smartphone intermediate or novice user (*spintnov*)
- Reported age of user (*age*)
- Mode of travel to shopping:
    - Car (*shcar*)
    - Bike or walk (*shnonm*)
    - Other (*shother*)
    - Transit (*shtran*)
- Mode of travel to work:
    - Car (*wkcar*)
    - Bike or walk (*wknonm*)
    - Other (*wkother*)
    - Transit (wktrans)

Hausman tests were run to test for biases or inconsistencies in the estimators. No evidence of simultaneity was found, thus it was not necessary to use an instrumental variables method or two-stage least squares. Instead, ordinary least squares (OLS) regressions were used for the initial analyses.

Three initial OLS models were run to test influences on the dependent variables of Utility (*utility*), Total Risk (*TotRisk*) and Compensation (*compen*). The resulting findings are shown in Tables XXX, XXXI, and XXXII below.

Table XXX: OLS Model of Utility

Dependent Variable: Utility					
Variable	**Parameter Estimate**	**Standard Error**	**t Value**	**Pr > |t|**	
Intercept	15.04	1.33	11.35	<.0001	
privtrade*	**0.22**	**0.05**	**4.64**	**<.0001**	
age**	**0.03**	**0.02**	**1.93**	**0.05**	
Gender	-0.24	0.31	-0.76	0.45	
Income	0.05	0.10	0.49	0.62	
Education	-0.12	0.16	-0.78	0.44	
LtechLBS	-0.09	0.44	-0.22	0.83	
LTechPay*	**2.08**	**0.47**	**4.38**	**<.0001**	
LTechNon	1.36	0.87	1.57	0.12	
wkcar	0.10	0.41	0.25	0.80	
wktrans	0.46	0.37	1.23	0.22	
spexpert	0.46	0.38	1.23	0.22	
spintnov	0.16	0.42	0.39	0.70	
gpsexp	0.06	0.46	0.13	0.89	
gpsintno	0.15	0.35	0.44	0.66	

Source	**DF**	**Sum of Squares**	**Mean Square**	**F Value**	**Pr > F**
Model	14.00	431.66	30.83	3.78	<.0001
Error	353.00	2881.51	8.16		
Corrected Total	367.00	3313.17			

Root MSE	2.86	**R-Square**	0.13		
Dependent Mean	20.65	**Adj R-Sq**	0.10		
Coeff Var	13.83				

** Significant at 0.01*
***Significant at 0.05*

Table XXXI: OLS Model of Total Risk

Dependent Variable: TotRisk					
Variable	**Parameter Estimate**	**Standard Error**	**t Value**	**Pr > |t|**	
Intercept	42.25	4.15	10.17	<.0001	
privtrade*	**-0.70**	**0.15**	**-4.79**	**<.0001**	
age**	**0.12**	**0.05**	**2.24**	**0.03**	
Gender	1.01	0.97	1.05	0.29	
Income	-0.10	0.31	-0.33	0.74	
Education	0.33	0.50	0.66	0.51	
LtechLBS	0.68	1.37	0.50	0.62	
LTechPay	-0.34	1.43	-0.24	0.81	
LTechNon**	**-5.04**	**2.61**	**-1.93**	**0.05**	
wkcar	0.42	1.25	0.34	0.74	
wktrans	-0.01	1.14	-0.01	0.99	
spexpert	0.50	1.15	0.43	0.67	
spintnov	0.25	1.27	0.20	0.84	
gpsexp	0.40	1.42	0.28	0.78	
gpsintno	-1.02	1.06	-0.97	0.33	

Source	**DF**	**Sum of Squares**	**Mean Square**	**F Value**	**Pr > F**
Model	14.00	3406.21	243.30	3.22	<.0001
Error	347.00	26200.00	75.50		
Corrected Total	361.00	29606.00			

Root MSE	8.69	**R-Square**	0.12
Dependent Mean	38.75	**Adj R-Sq**	0.08
Coeff Var	22.42		

** Significant at 0.01*
***Significant at 0.05*

Table XXXII: OLS Model of Compensation

Dependent Variable: Compen					
Variable	**Parameter Estimate**	**Standard Error**	**t Value**	**Pr > |t|**	
Intercept	58.85	5.89	9.98	<.0001	
privtrade*	**-1.40**	**0.20**	**-6.93**	**<.0001**	
age	0.01	0.07	0.12	0.91	
Gender	-2.06	1.36	-1.51	0.13	
Income	-0.26	0.44	-0.60	0.55	
Education**	**-1.28**	**0.70**	**-1.83**	**0.07**	
LtechLBS	-1.34	1.91	-0.70	0.48	
LTechPay	2.10	2.05	1.02	0.31	
LTechNon	-1.44	3.66	-0.39	0.69	
wkcar	-0.21	1.75	-0.12	0.91	
wktrans	0.24	1.61	0.15	0.88	
spexpert	-1.02	1.62	-0.63	0.53	
spintnov	-1.31	1.80	-0.73	0.47	
gpsexp	2.80	1.99	1.41	0.16	
gpsintno	-0.67	1.48	-0.45	0.65	
Source	**DF**	**Sum of Squares**	**Mean Square**	**F Value**	**Pr > F**
Model	14.00	9248.68	660.62	4.55	<.0001
Error	336.00	48790.00	145.21		
Corrected Total	350.00	58038.00			
Root MSE	12.05	**R-Square**	0.16		
Dependent Mean	29.44	**Adj R-Sq**	0.12		
Coeff Var	40.94				

** Significant at 0.01*
***Significant at 0.1*

For each model, willingness to trade privacy (*privtrade*) is seen as having a significant impact on the dependent variable. In the case of utility, the parameter is positive, indicating that the higher the willingness to trade, the greater the respondent views the utility of transportation benefits. The parameter is negative for both compensation and total risk, indicating that persons with a greater degree of willingness to trade privacy both require less compensation to give up personal data, and associate lower risk with giving up such data. These findings are consistent with the literature reviewed above, as well as with the hypothesis of this thesis.

For utility, other variables of interest are age, which has a slightly positive impact, indicating that older persons may be more likely to place importance on having reliable and useful transportation information, and *LTechPay*, which indicates use of electronic transit fare cards or electronic toll passes. It is likely that those persons who use these methods of payment are interested in efficiency and expediency, thus their higher ranking of utility is unsurprising. For estimation of risk, age was again a factor, with an increase in age generally indicating a slight increase in estimation of risk of sharing private data. A surprising finding here is that those respondents who do not currently use location technologies report slightly less degree of risk estimation. This may be due to lack of knowledge of information that may be revealed through use of these services, or it may indicate a general unfamiliarity with such services. Educational levels had a slight positive impact on compensation requirements, indicating that more highly educated persons were likely to request less compensation to share data.

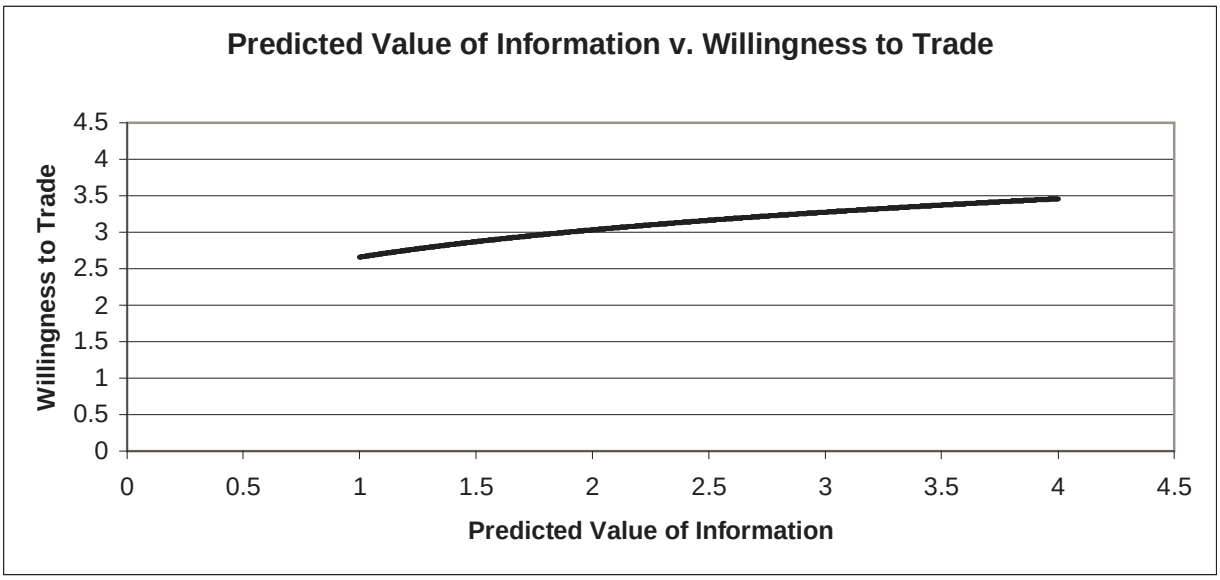
Neither travel modes nor user-reported degree of experience with GPS or Smartphone technologies were seen to significantly impact the model, indicating that these variables are not representative overall of person's privacy and utility preferences. Model fit and r^2 statistics shown in the model tables above indicate that, while influence is consistent among tested variables, the models overall do not account for a great deal of variance in the data. The following sections will use different methods and models to test the influence of variables in more detail.

The survey used a Likert scale to indicate degree of agreement, thus the resulting data take an ordinal, or ordered, form. An ordered probit model was thus next used to estimate relationships of interest. For this analysis, three ordered probit models were run, one for each

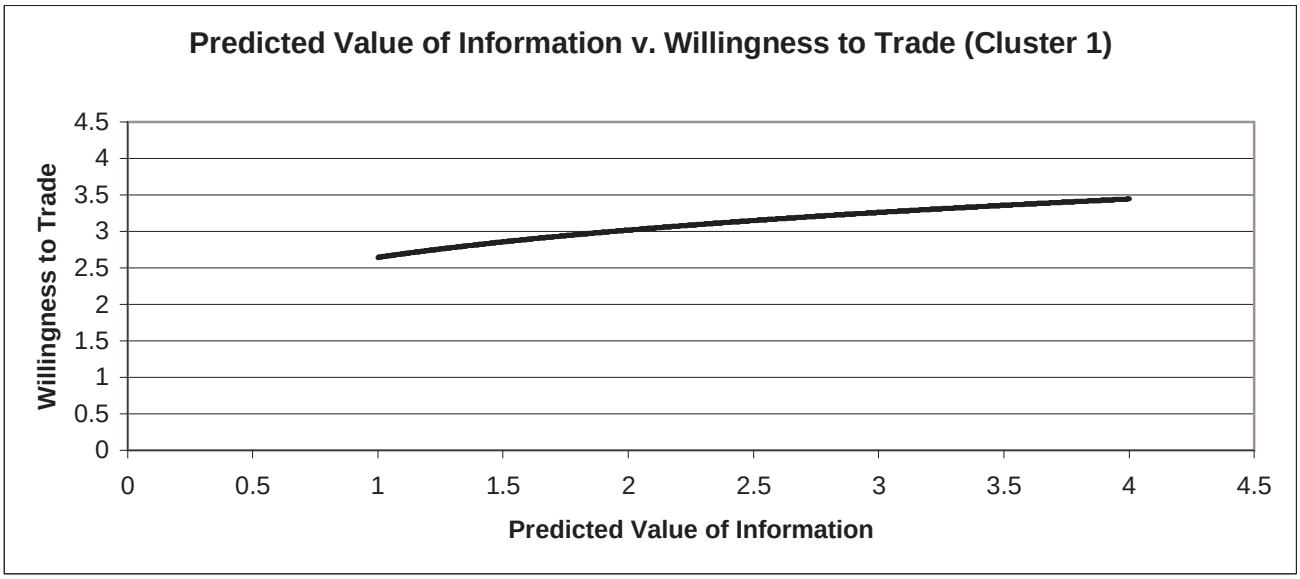
of the identified clusters, in addition to an overall model. While individual types of information were initially surveyed, a correspondence analysis was run on the responses that resulted in a standardized Cronbach's alpha of 0.651496, indicating strong correlation between the tested variables. Thus, the composite variable *totalinfovalue* was created to represent the general value that respondents assigned to the types of travel information surveyed. In a similar manner, the composite variable *totalprivacytrade* was created from the variables indicating respondent's willingness to trade private data for transportation benefits based on a standardized Cronbach's alpha of 0.837.

First, scatterplots were constructed for each of the three respondent clusters identified above. For the analysis, importance of travel information was plotted on the y-axis, and willingness to trade private information was plotted on the x-axis, with regression lines indicating the slope of responses. Estimated graphs are shown below in Figure 12. As shown here, the slopes of the estimated regression lines are somewhat similar, though cluster 2 is significantly steeper, indicating that for this group, willingness to trade private information in relation to perceived importance of transportation information is somewhat lower than for other clusters.

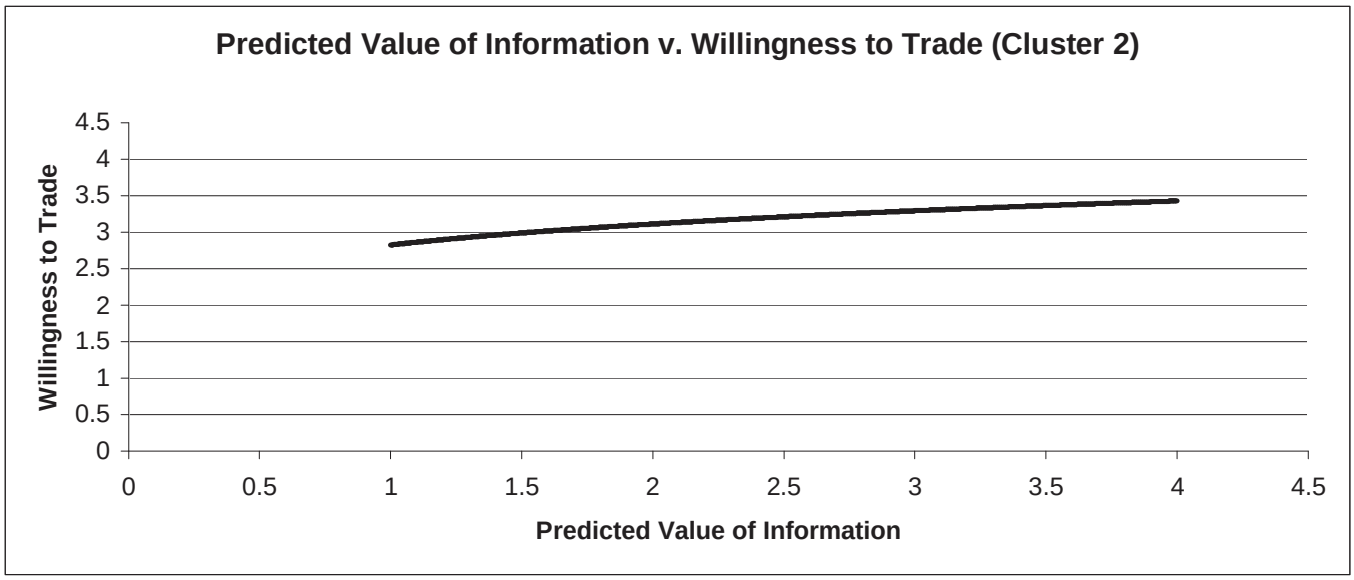
Figure 12: Privacy Cluster Analysis of *totalinfouse* by *totalprivacytrade*



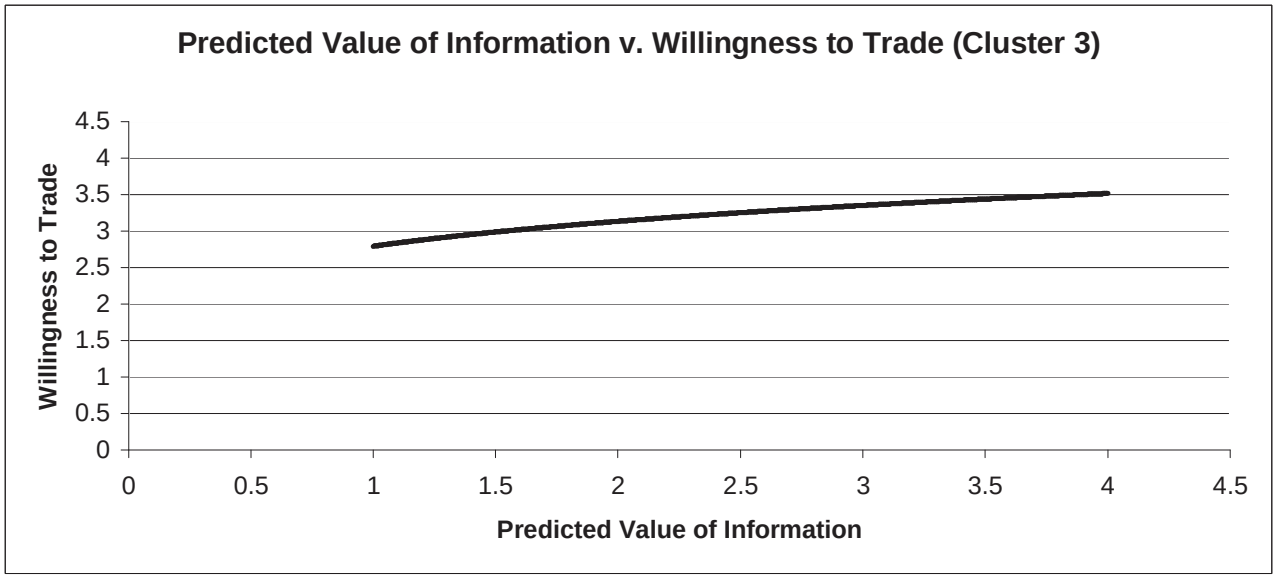
Cluster 1



Cluster 2



Cluster 3



Next, SAS proc QLIM (Qualitative and LImited dependent variable Model) was used to run the ordered probit models, described in 4.4.3 above. In order to limit the number of thresholds tested, raw scores for *totalinfovalue* were categorized into quadrants and then analyzed.

Here, ordered probit models were run first on the entire data set, and then according to the privacy preference clusters identified above, resulting in the findings shown in Table XXXIII.

Table XXXIII: Results of Ordered Probit Model (OPM) for Total Value of Information (*totalinfovalue*) Overall and for Clusters

	Overall		Cluster 1		Cluster 2		Cluster 3	
Parameter	Estimate	Approx Pr > |t|	Estimate	Approx Pr > |t|	Estimate	Approx Pr > |t|	Estimate	Approx Pr > |t|
totalprivacytrade	**0.09**	**<.0001***	**0.08**	**0.0002****	**0.22**	**0.038****	0.05	0.51
spexpert	0.16	0.29	0.19	0.29	0.68	0.32	0.02	0.96
spintnov	0.01	0.94	-0.24	0.21	**2.67**	**0.018****	0.11	0.85
gpsexp	-0.06	0.75	0.01	0.98	-0.17	0.87	-0.52	0.52
gpsintno	-0.04	0.79	0.08	0.61	**-2.06**	**0.005****	0.48	0.34
LtechLBS	-0.10	0.57	-0.09	0.67	-1.43	0.12	0.09	0.87
LTechPay	**0.90**	**<.0001***	**0.87**	**<.0001***	**5.13**	**0.01****	**1.22**	**0.04****
LTechNon	0.36	0.28	0.52	0.20	0.97	0.55	1.27	0.24
usetransit	-0.02	0.77	-0.05	0.49	-0.30	0.28	0.19	0.47
usewalk	0.01	0.95	0.05	0.70	**-1.11**	**0.098*****	-0.11	0.67
usebike	**-0.11**	**0.069*****	-0.10	0.15	**-0.70**	**0.072*****	-0.24	0.36
age	**0.01**	**0.064*****	**0.02**	**0.057*****	0.11	0.17	0.00	1.00
Education	-0.03	0.68	-0.04	0.55	0.33	0.35	-0.03	0.90
Gender	-0.09	0.45	0.01	0.95	-0.87	0.12	-0.25	0.55
impshr	0.01	0.85	-0.05	0.36	0.00	1.00	0.00	0.97
Income	0.01	0.79	-0.02	0.68	0.04	0.89	0.16	0.29
_Limit1	0.35	0.58	-0.16	0.84	**7.20**	**0.057*****	-0.08	0.97
_Limit2	0.92	0.14	0.40	0.60	**8.11**	**0.039****	0.94	0.62
_Limit3	**2.45**	**<0.0001***	**2.02**	**0.009****	**10.49**	**0.013****	2.55	0.18
Goodness of Fit Measures								
Log Likelihood	-372.348		-281.520		-22.050		-40.480	
AIC	782.695		601.040		82.099		118.960	
Schwarz Criterion	857.256		670.371		115.999		153.704	
Cragg-Uhler 1	0.139		0.138		0.627		0.286	
Adjusted Estrella	0.047		0.015		0.120		-0.552	
McKelvey-Zavoina Pseudo R2	0.168		0.170		0.870		0.351	

** Significant at <0.0001*

*** Significant at 0.05*

**** Significant at 0.1*

The McKelvey-Zavoina goodness of fit measures shown above indicated a somewhat weak fit for the overall model, though a fairly strong relationship was seen for Cluster 2. Additionally, threshold estimates (designated as _Limit1 (the deviation of τ_1 from τ_2), Limit2, which is the deviation of τ_2 from τ_3, and _Limit3, which is the deviation of τ_3 from τ_4) were insignificant with the exception of Limit3 for all models but that of Cluster 3, indicating that the greatest relationship between willingness to trade and perceived value of information occurs for those with the highest privacy preferences. Model fit for Clusters 2 and 3 was moderately high, while Cluster 1 indicated a weaker fit.

The significance of *totalprivacytrade* in relation to *totalinfovalue* is strong for all but Cluster 3, indicating that for persons with stronger privacy values (as represented by Clusters 1 and 2, as well as the overall model), there is a strong relationship between willingness to trade private information and the degree of importance assigned to various types of transportation information. Weaker relationships found for those with low privacy preferences indicate that low overall privacy preferences also indicates a low degree of relationship between willingness to trade information for transportation benefits, perhaps due to the relatively low value assigned to private information. Such a finding is consistent with overall expectations, as it indicates that those with higher privacy sensitivity are more concerned with receipt of potential benefits in order to be willing to trade their personal information. Additional variables of interest, highlighted in bold in the tables above, indicate that varying factors have influence on the perceived value of information for persons with different degrees of privacy sensitivity. For example, for all privacy preferences, use of electronic payment technology was positively associated with value determinations, supporting the finding under Hypothesis 1 that use of

such technologies is correlated with perceptions of transportation utility. For those with the highest privacy preferences (Cluster 1), age was correlated with perceptions of value, indicating that perceptions of value increased as age increased. For those with moderate privacy preferences (Cluster 2), persons who considered themselves intermediate or novice users of smartphones reported higher information values, while use of GPS was related to lower rankings of value. Persons who reported use of non-motorized travel modes were also less likely to value transportation information highly, perhaps as a result of feeling more in control of their experience of the transportation environment (for example, transit reliability and disruptions to motorized travel lanes will be of little value for persons walking or biking). For those with the lowest privacy preferences, seen in Cluster 3, only current use of payment technologies was significantly associated with perceptions of information value, indicating that these respondents have a fairly consistent valuation of information value and concurrent willingness to trade their private data.

Hypothesis 3: Positive correlations will be seen between WTTP and perceived degree of expected compensation.

It is hypothesized that persons who demonstrate a low willingness to trade private information will expect a relatively high degree of compensation for provision of this information. Such a hypothesis is consistent with Milne and Gordon (1993) and Sheehan and Hoy (2000), though here we are examining the issue in the context of transportation information. Here, we expect

that findings will be highly significant with respect to those respondents in Cluster 1, and less or not at all significant with respect to those respondents in Clusters 2 and 3.

To test this hypothesis, Factor 4, Compensation Factor (CF), was modeled in relation to *totalprivacytrade*, identified earlier. As with *totalinfovalue*, raw *totalprivacytrade* scores were clustered into quadrants for the evaluation, in order to provide appropriate thresholds for the OPM analysis. The table below provides an overview of model results and goodness of fit measures.

Table XXXIV: Results of Ordered Probit Model (OPM) for Total Willingness to Trade Privacy (*totalprivacytrade*) Overall and for Clusters

Overall Model of *totalprivacytrade* (Quadrants)								
	Overall		Cluster 1		Cluster 2		Cluster 3	
Parameter	Estimate	Approx Pr > |t|	Estimate	Approx Pr > |t|	Estimate	Approx Pr > |t|	Estimate	Approx Pr > |t|
Factor4 (Compensation)	-0.55	<.0001*	-0.58	<.0001*	-1.04	0.100***	1.44	0.36
spexpert	0.00	0.99	-0.02	0.89	0.02	0.97	-0.66	0.29
spintnov	-0.21	0.20	-0.31	0.10***	0.16	0.80	-2.21	0.071***
gpsexp	0.17	0.35	0.15	0.45	0.44	0.56	0.23	0.81
gpsintno	0.20	0.14	0.10	0.54	0.73	0.19	1.56	0.046**
LtechLBS	0.14	0.41	0.20	0.31	0.77	0.31	-1.63	0.073***
LTechPay	0.07	0.71	0.12	0.56	-0.17	0.89	-0.44	0.59
LTechNon	-0.13	0.69	-0.04	0.92	0.06	0.97	-1.79	0.21
usetransit	0.01	0.93	-0.01	0.83	-0.09	0.73	-0.40	0.25
usewalk	-0.23	0.023**	-0.31	0.012**	0.67	0.13	-1.15	0.04**
usebike	-0.05	0.39	-0.07	0.33	0.17	0.53	-0.75	0.14
age	0.00	0.59	-0.01	0.19	0.06	0.15	0.01	0.66
Education	0.09	0.14	0.07	0.31	0.03	0.93	0.94	0.034**
Gender	-0.21	0.087***	-0.30	0.03**	-0.22	0.66	0.12	0.86
impshr	-0.10	0.015**	-0.10	0.032**	-0.11	0.55	-0.09	0.55
Income	-0.04	0.32	-0.02	0.62	-0.36	0.098***	-0.65	0.034**
_Limit1	-1.83	0.001**	-2.19	0.001**	-0.86	0.69		
_Limit2	-0.94	0.096***	-1.28	0.059***	-0.26	0.90	-3.35	0.35
_Limit3	0.83	0.14	0.39	0.56	2.53	0.25	-0.20	0.96
Goodness of Fit Measures								
Log Likelihood	-378.57149		-311.44315		-29.49848		-19.50549	
AIC	795.14298		660.88631		96.99696		75.01098	
Schwarz Criterion	869.08422		730.21682		130.89656		102.48547	
Cragg-Uhler 1	0.1884		0.1671		0.3096		0.3878	
Adjusted Estrella	0.1008		0.0483		-0.542		-0.629	
McKelvey-Zavoina Pseudo R2	0.2264		0.1968		0.415		0.539	

** Significant at <0.0001*

*** Significant at 0.05*

**** Significant at 0.1*

Again, McKelvey-Zavoina goodness of fit measures showed a moderately good model fit, particularly for Clusters 2 and 3. Threshold estimates were significant only for the overall model and for Cluster 1, indicating that for those persons with less privacy sensitivity, willingness to trade privacy is less associated with increased compensation. Cluster three did not include any persons with low willingness to trade privacy (included in the first quadrant), thus no threshold is shown for _Limit1.

The lower overall model fit for Cluster 1 indicates that persons with higher privacy preferences may have additional socio-demographic or other personal characteristics that are not currently included in the mode. Compensation (*Factor4*), as expected, showed clear influence for all but persons in Cluster 3, indicating a greater relationship between the willingness of persons with higher privacy preferences and degree of compensation expected.

As with the models in Hypothesis 3, various factors indicated differing degrees of influence for each of the evaluated Clusters. For those persons in Cluster 1, a number of tested variables were seen to have significant impacts on willingness to trade, including gender, pedestrian travel, importance of information provided to consumers by ITS and LBS companies (*impshr*), and use of smartphones. Overall, each of these variables had a negative influence on willingness to trade information. For persons in Cluster 2, who reveal moderate privacy preferences, only compensation and income were shown to have significant influences, with persons reporting a higher income level being less willing to trade information. Persons with the lowest privacy preferences (Cluster 3) had a number of influencing factors, with persons reporting a higher educational level and those with some familiarity with GPS reporting higher willingness to trade, and those with higher income levels, who walk, and who currently use location based services being less willing to trade. For Cluster 3, these variables were more influential than compensation levels, indicating that compensation is less important than personal and socio-demographic characteristics.

The marginal effects of increases in Factor 4 and Cluster category on *totalprivacytrade* were calculated and are shown below in Table XXXV. The findings indicate that a one unit increase in Factor 4 (compensation), calculated from compensation required to share personal

and travel data, has varying degrees of influence on the likelihood of willingness to trade privacy. For lower willingness to trade (Quadrants 1 and 2), an increase in compensation will have a fairly significant positive influence on willingness to trade, indicating that economic benefits are of more importance to persons with higher privacy sensitivity. For those with higher willingness to trade, compensation is negatively associated with willingness to share, indicating, perhaps, that the higher the perceived value of information for persons with lower privacy sensitivity, the less they will be willing to share. The associated marginal effects of the Cluster categories reflects the direction, if not the scale, of these findings.

Table XXXV: Marginal Effects of Factor4 and Cluster on *totalprivacytrade*

Variable	totalprivacytrade Quadrant:	Mean
Factor4 (Compensation)	1	7.7%
Factor4 (Compensation)	2	8.8%
Factor4 (Compensation)	3	-4.0%
Factor4 (Compensation)	4	-12.5%
Cluster	1	0.9%
Cluster	2	1.0%
Cluster	3	-0.4%
Cluster	4	-1.4%

These findings indicate that there is a generally positive correlation between willingness to trade private information for degree of compensation expected for those persons with low to moderate degrees of privacy preferences. These effects are weaker for those with high degrees of privacy preferences, indicating that compensation will have little influence on those persons who have high privacy preferences. These findings indicate that there is a subset of the

population who will not trade their private information regardless of the amount of compensation provided.

Hypothesis 4: Persons will display a relationship between WTTP and risk due to perceptions of perceived privacy loss.

This hypothesis states that persons who have a lower degree of willingness to trade private information will also have higher degrees of perceived privacy loss relevant to certain situations. Respondents were asked to respond to a question inquiring their perceived risk of privacy loss in the situations described in Table XX above. As with earlier multi-part questions, a correlation analysis was run, resulting in a standardized Cronbach's alpha of 0.8834, indicating strong correlations between the tested variables. As a result, a composite variable of risk was developed titled *totalprivacyrisk*. As above, an ordered probit model was run to ascertain relationships for each of the three clusters indicated above. The results of the analysis are shown in Table XXXVI below.

Table XXXVI: Results of Ordered Probit Model (OPM) for Willingness to Trade Privacy (*totalprivacytrade*) vs. Perceived Risk Overall and for Clusters

Overall Model of *totalprivacytrade* (Quadrants)								
	Overall		**Cluster 1**		**Cluster 2**		**Cluster 3**	
Parameter	**Estimate**	**Approx Pr > |t|**	**Estimate**	**Approx Pr > |t|**	**Estimate**	**Approx Pr > |t|**	**Estimate**	**Approx Pr > |t|**
totalprivacyrisk	-0.026	0.0001*	-0.025	0.0011**	-0.072	0.038**	-0.029	0.163
spexpert	0.036	0.803	0.048	0.780	0.082	0.883	-0.004	0.994
spintnov	-0.130	0.418	-0.210	0.265	-0.158	0.798	-0.607	0.312
gpsexp	0.159	0.369	0.124	0.541	0.218	0.769	-0.308	0.687
gpsintno	0.283	0.037**	0.089	0.579	0.682	0.213	1.378	0.012**
LtechLBS	0.137	0.427	0.179	0.381	1.519	0.049**	-0.595	0.321
LTechPay	-0.104	0.569	-0.071	0.728	-0.534	0.677	-0.334	0.601
LTechNon	-0.377	0.250	-0.195	0.621	0.050	0.972	-2.403	0.041**
usetransit	0.054	0.350	0.031	0.635	0.105	0.677	-0.208	0.400
usewalk	-0.118	0.221	-0.279	0.024**	0.404	0.338	-0.013	0.961
usebike	-0.024	0.692	-0.008	0.905	0.199	0.460	-0.439	0.06***
age	0.005	0.447	-0.007	0.382	0.095	0.035**	0.022	0.336
Education	0.216	0.001**	0.165	0.026**	0.163	0.566	0.692	0.005**
Gender	-0.120	0.324	-0.235	0.09***	-0.337	0.515	-0.134	0.766
impshr	-0.091	0.028**	-0.087	0.079***	-0.137	0.456	-0.166	0.146
Income	-0.064	0.105	-0.025	0.588	-0.411	0.073***	-0.282	0.059***
_Limit1	-1.513	0.009**	-2.176	0.002**	-2.205	0.334	-0.579	0.747
_Limit2	-0.775	0.178	-1.308	0.066***	-1.605	0.480	-0.202	0.910
_Limit3	0.912	0.113	0.321	0.650	1.223	0.596	2.477	0.170
Goodness of Fit Measures								
Log Likelihood	-406.52446		-310.19295		-28.5998		-37.00255	
AIC	851.04892		658.3859		95.19961		112.0051	
Schwarz Criterion	925.40548		727.24223		129.09921		147.94969	
Cragg-Uhler 1	0.1198		0.127		0.3372		0.4572	
Adjusted Estrella	0.0247		-0.001		-0.494		-0.172	
McKelvey-Zavoina Pseudo R2	0.1436		0.1495		0.4734		0.5809	

** Significant at 0.0001*

*** Significant at 0.05*

**** Significant at 0.1*

As above, the overall fit of the models was moderate, with low McKelvey-Zavoina pseudo r^2 values for the overall model and for Cluster 1, but fairly high values for Clusters 2 and 3. Threshold estimates were significant only at Limit 1 for the overall model, and for threshold limits 1 and 2 for Cluster 1, indicating that risk perceptions are perhaps most differentiated and

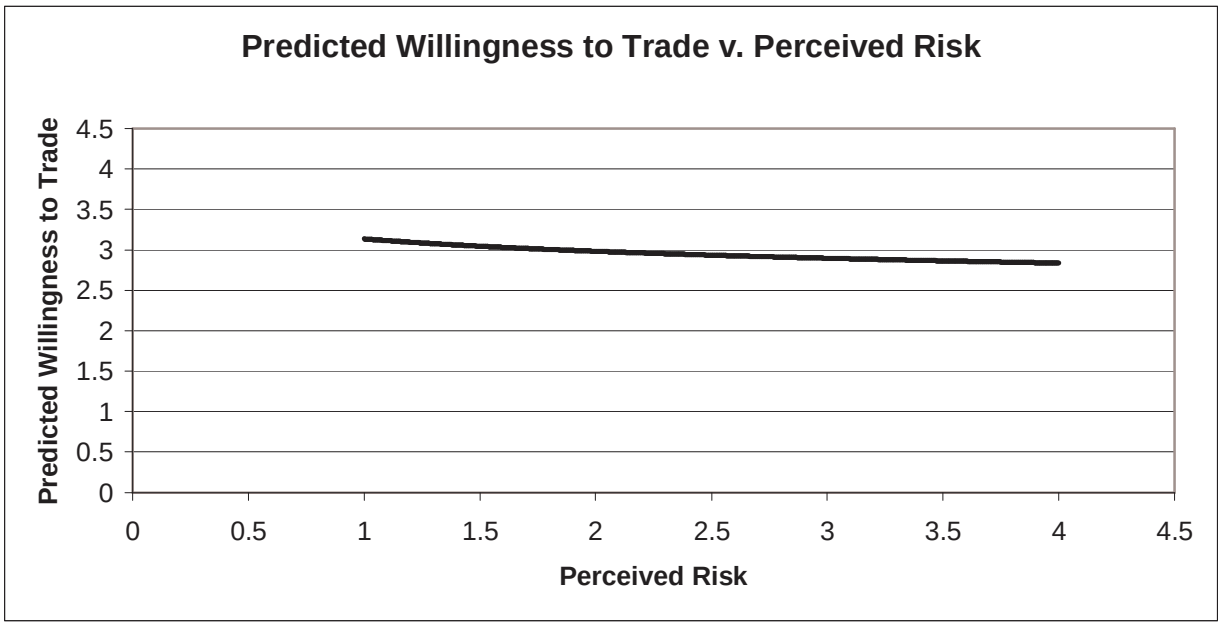
significant for those persons who have high privacy sensitivity. From these results, we see that perceptions of privacy risk have the most influence on willingness to trade privacy for persons in Cluster 1, who demonstrate the strongest privacy preferences; however, marginal effects, shown below, were also calculated, and indicate that there is little overall influence of perceptions of risk on willingness to trade. As above, the higher the risk estimation, the less likely the respondent was to indicate high willingness to trade privacy, a finding consistent with findings reported above. Also consistent were findings related to the influence of demographic variables (in particular education), current use of technology, and preferred transportation mode.

Table XXXVII: Marginal Effects of *totalprivacyrisk* on *totalprivacytrade*

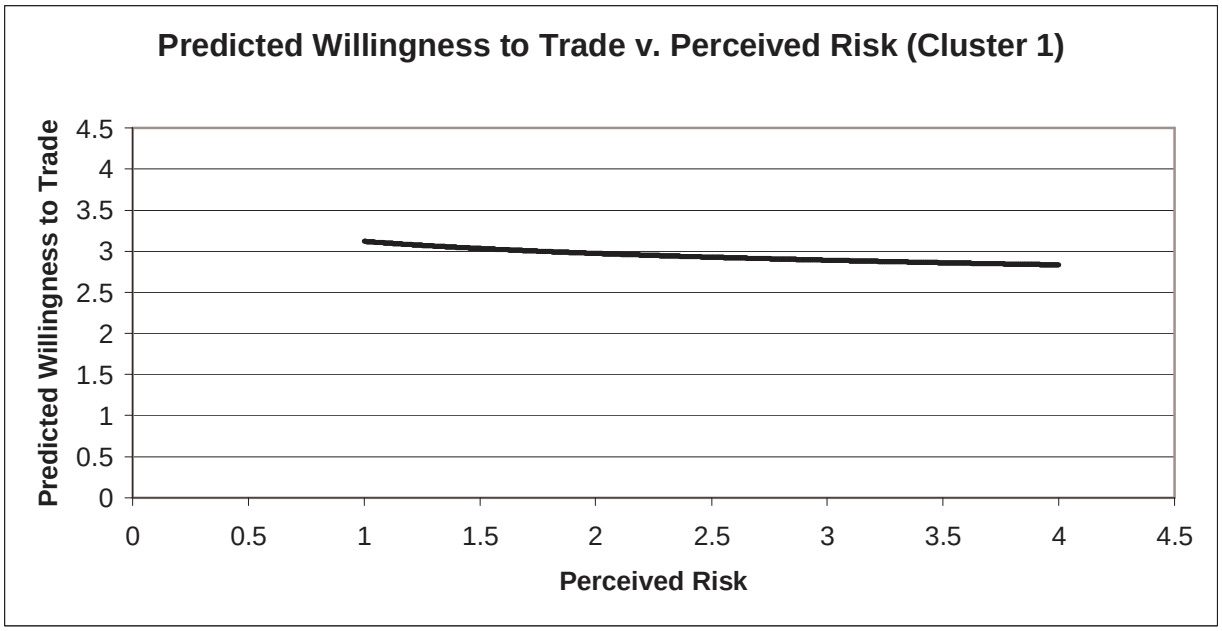
Variable	totalprivacytrade Quadrant:	Mean
totalprivacyrisk	1	0.49%
totalprivacyrisk	2	0.42%
totalprivacyrisk	3	-0.24%
totalprivacyrisk	4	-0.67%

Results were also plotted for each cluster, resulting in the graphs shown below in Figure 13.

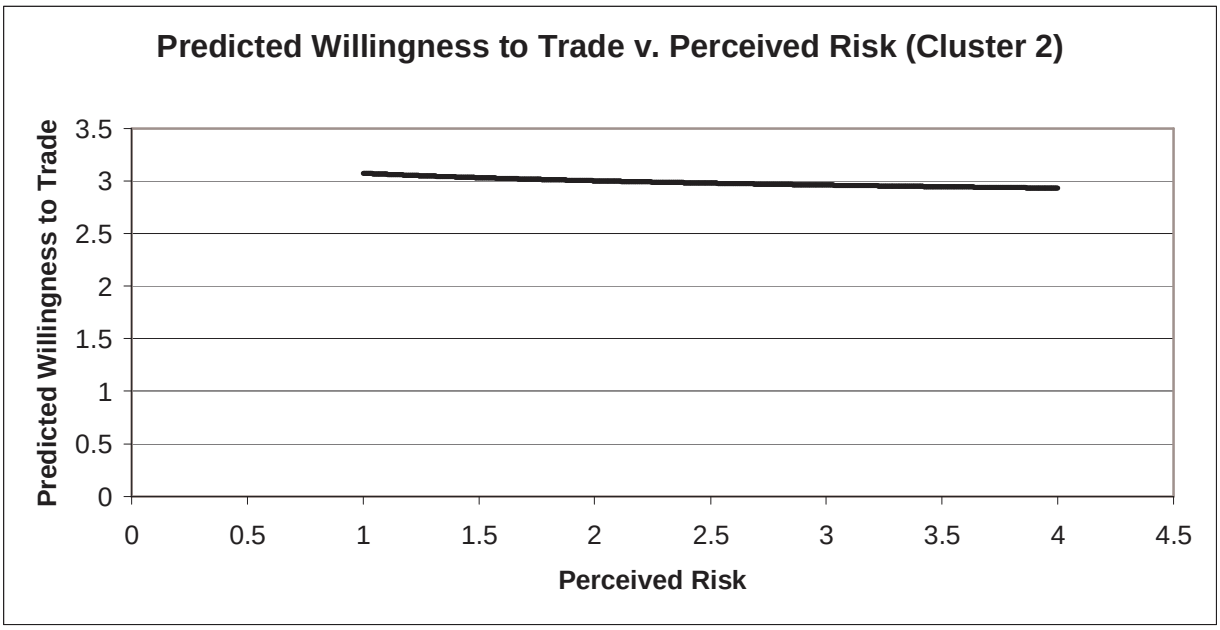
Figure 13: Plots of Predicted Willingness to Trade v. Perceived Risk



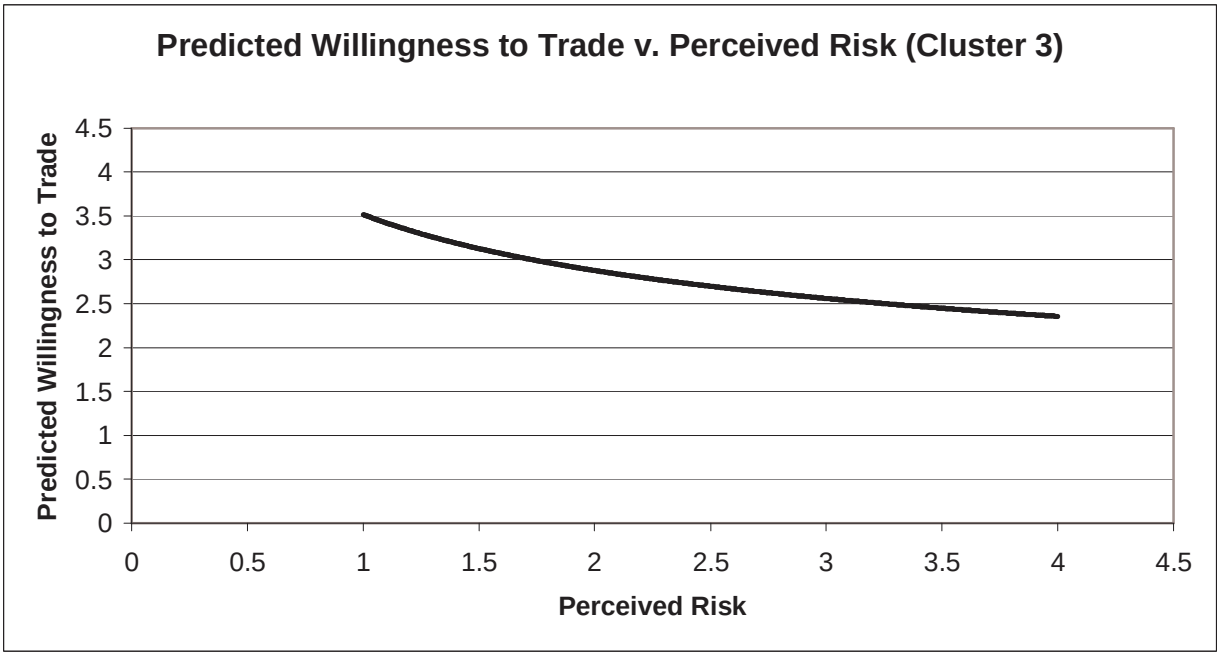
Cluster 1



Cluster 2



Cluster 3



The generally gentle downward slope of the line, combined with the calculated marginal effects, indicates that for each cluster there is a fairly steady relationship between perceptions of risk and willingness to trade. In part, this may be due to the fact that, as shown above in Table XX, perceptions of risk across the general spectrum of scenarios tested were fairly steady. Most respondents reported a fairly high perception of risk for each variable tested. One question arises, however, regarding the differences in perceptions of risk relative to public and private agencies. Thus, the differences in average means of marginal effects of risk perceptions were evaluated, resulting in the findings shown in Table XXXVIII. The absolute value of differences between the impact of perceptions of risk on willingness to trade between public and private agencies is small; however, the overall differences show some fairly clear patterns.

For the majority of respondents in all clusters, willingness to trade information increases as perceptions of risk increase. However, the higher a respondent's perception of risk, the likelihood of willingness to trade information decreases. Here, it may be interpreted that, up to a point, the possibility of benefits of trading information may offset risk perception for those with lower general propensity to trade information. However, as a person's perception of risk increases beyond a certain point, the likelihood of trading information will decrease, indicating that the benefits are not perceived as outweighing risk. Of note is that, especially in the case of those in Cluster 1, perceptions of risk between public and private agencies present an interesting pattern. Up to the point where the trend reverses, participants' willingness to trade information is slightly higher for increases in risk in the public sector than in the private sector. However, for those with higher perceptions of risk, the willingness to trade information becomes weaker relevant to risk in the public sector than in the private sector. This may

indicate that for those who have generally low privacy concerns, more risk may be associated with private corporations and agencies having access to personal information than for public agencies. However, persons with higher perceptions of risk, that risk is more closely associated with access to information by public agencies. Such a finding may reflect that persons who have higher concerns are generally less trusting of public organizations.

Table XXXVIII: Means of Marginal Effect of Risk Perceptions on Willingness to Trade Information

	Average Means of Marginal Effect of Perceptions of Risk on Willingness to Trade Information								
	Cluster 1			Cluster 2			Cluster 3		
Composite *totalprivacytrade* score	Private	Public	(Private - Public)/Private	Private	Public	(Private - Public)/Private	Private	Public	(Private - Public)/Private
4	0.17%	0.27%	-65.17%						
5	0.18%	0.29%	-63.90%						
6	0.08%	0.13%	-61.63%						
7	0.31%	0.51%	-61.17%	0.45%	0.39%	12.01%			
8	0.32%	0.52%	-62.18%				0.43%	0.54%	-26.52%
9	0.15%	0.25%	-62.27%						
10	0.33%	0.54%	-61.57%						
11	0.30%	0.48%	-61.35%				0.29%	0.40%	-40.80%
12	0.31%	0.48%	-57.02%	0.62%	0.56%	9.91%	0.45%	0.35%	21.38%
13	0.19%	0.32%	-69.49%	0.90%	0.79%	12.23%	0.68%	0.31%	54.83%
14	0.02%	0.05%	-126.08%	0.76%	0.66%	13.34%	0.77%	0.52%	32.35%
15	-0.16%	-0.25%	-52.05%	0.53%	0.47%	10.63%	0.12%	0.82%	-591.77%
16	-0.86%	-1.39%	-61.70%	-0.43%	-0.38%	13.02%	-0.74%	0.91%	222.78%
17	-0.37%	-0.64%	-72.00%	-0.36%	-0.32%	12.34%	-0.47%	0.12%	126.27%
18	-0.45%	-0.73%	-62.16%	-0.87%	-0.77%	11.49%	-0.19%	-1.20%	-525.72%
19	-0.25%	-0.40%	-62.19%	-0.23%	-0.21%	9.85%	-0.69%	-0.66%	4.33%
20	-0.26%	-0.42%	-60.59%	-1.35%	-1.20%	11.69%	-0.65%	-0.27%	59.03%

Hypothesis 5: Correlations will be seen between WTTP and extent of knowledge and concern related to privacy matters, as demonstrated by the reading and understanding of privacy policies.

It is hypothesized that those respondents most likely to be willing to trade private information will also be least likely to have knowledge related to the privacy protections offered by the services they used, as reflected in Factor 1 above (Knowledge Factor (KF)). Findings from an ordered probit model indicated a weak relationship between the two, as shown in Table XXXIX below, and interesting patterns in the marginal effects analysis, shown in Table XL.

Table XXXIX: Ordered Probit Model of Knowledge Factor and Willingness to Trade Information

Parameter Estimates				
Parameter	**Estimate**	**Standard Error**	**Approx t Value**	**Pr > |t|**
Cluster 1				
Intercept	2.209232	0.197346	11.19	<.0001
Factor1	-0.06813	0.061879	-1.1	0.2709
Cluster 2				
Intercept	2.031106	0.424555	4.78	<.0001
Factor1	0.110597	0.159142	0.69	0.4871
Cluster 3				
Intercept	1.92594	0.431804	4.46	<.0001
Factor1	0.164931	0.166362	0.99	0.3215

Table XL: Marginal Effects of Knowledge Factor on Probability of Willingness to Trade Data

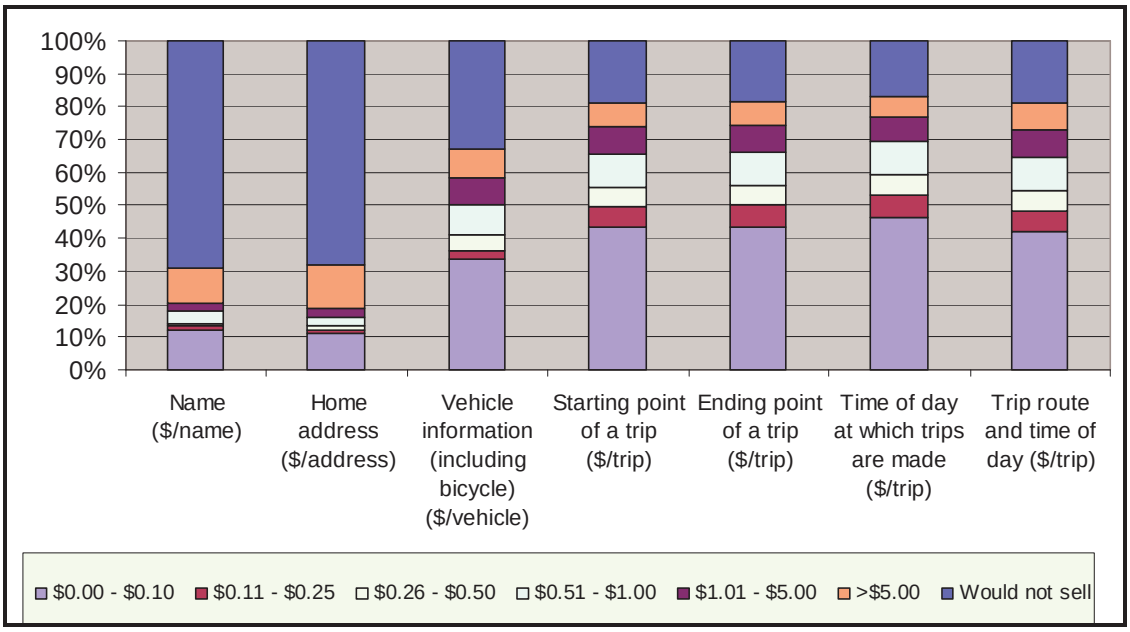
Willingness to Trade value	Cluster 1	Cluster 2	Cluster 3
4	0.24%		
5	0.20%		
6	0.09%		
7	0.36%	-0.56%	
8	0.36%		-1.09%
9	0.17%		
10	0.37%		
11	0.33%		
12	0.34%	-0.86%	-1.47%
13	0.22%	-1.22%	-1.61%
14	0.03%	-0.98%	-1.64%
15	-0.17%	-0.68%	-0.37%
16	-0.99%	0.55%	1.23%
17	-0.46%	0.48%	0.78%
18	-0.52%	1.19%	0.46%
19	-0.28%	0.31%	1.78%
20	-0.30%	1.78%	1.93%

The marginal effects analysis, shown above, reveals an interesting pattern in the data. For Cluster 1, as willingness to trade private information increases, likelihood of having a high knowledge factor decreases as reported by respondents; however, the pattern reverses for persons in Clusters 2 and 3, who have reported that as their willingness to trade information increases, their likelihood of having a high knowledge factor increases past a certain level. This may, as outlined in Hypothesis 3, reflect issues of trust. For persons who generally reflect higher privacy concerns (as seen in Cluster 1), as one's privacy concerns lessen, one's concern for knowledge lessens. This may reflect either a certain degree of trust in the system itself, or it may indicate an "ignorance is bliss" attitude. For those with lower privacy concerns (as shown in Clusters 2 and 3), lower privacy concerns reflect a lower knowledge factor; however, as privacy concerns increase, likelihood of having an increased Knowledge Factor increases. This may indicate a more pragmatic attitude, as consumers with lower privacy concerns may be more concerned with making informed decisions based on knowledge regarding privacy practices of the applications or services they are using.

Hypothesis 6: Respondents will be more willing to reduce compensation for trading private information for purposes of safety or efficiency than for cost savings.

Respondents were asked to indicate how much compensation they would require in order to trade various types of personal information as shown in Table XLI.

Table XLI: Compensation Required to Provide Data



Persons showed most reluctance to sell name and address information, with nearly 70% of respondents reporting that they would be unwilling to sell these data, and prices required to provide such data were significantly higher for these than for other data types. Despite these differences, a correlation procedure was run in SAS, resulting in a standardized Cronbach's alpha of 0.9051. Due to this finding, a composite variable was created in order to test how respondent requirements for compensation changed due to various scenarios.

Respondents were presented with the following six scenarios and asked to indicate if their compensation requirements would increase, decrease, or stay the same under the given circumstances:

- Scenario 1: Reduction of travel time by an average of 15% per trip
- Scenario 2: Reduction of gas tax for all persons by $0.01/gallon
- Scenario 3: Reduction of gas tax for all persons by $0.02/gallon
- Scenario 4: Decrease in vehicular fatalities by 100 persons per year
- Scenario 5: Decrease in vehicular fatalities by 1,000 persons per year

- Scenario 6: Collection agencies will sell information to third parties (such as Google, NAVTEQ, or Ford)

First, a crosstabulation was developed to determine general patterns of increases or decreases in expected compensation across scenarios. The generalized results are shown in Table XLII.

Table XLII: Crosstabulation of Compensation Changes Across Scenarios

		Increase		Decrease		No Change		
Scenario	Variable	# of respondents	%	# of respondents	%	# of respondents	%	Total
Scenario 1: Reduction of Travel time by average of 15%	Name	5	1%	27	7%	331	91%	363
	Address	6	2%	32	9%	325	90%	363
	Vehicle	14	4%	68	19%	277	77%	359
	Origin	16	4%	93	26%	252	70%	361
	Destination	17	5%	92	25%	253	70%	362
	Time of Day	15	4%	91	25%	254	71%	360
	Trip Route	16	4%	93	26%	248	69%	357
Scenario 2: Reduction of gas tax by $0.01 per gallon	Name	21	6%	16	4%	324	90%	361
	Address	21	6%	19	5%	321	89%	361
	Vehicle	29	8%	27	8%	301	84%	357
	Origin	28	8%	38	11%	294	82%	360
	Destination	28	8%	38	11%	294	82%	360
	Time of Day	29	8%	40	11%	289	81%	358
	Trip Route	29	8%	41	11%	288	80%	358
Scenario 3: Reduction of gas tax by $0.02 per gallon	Name	19	5%	19	5%	322	89%	360
	Address	19	5%	20	6%	321	89%	360
	Vehicle	25	7%	36	10%	293	83%	354
	Origin	29	8%	44	12%	285	80%	358
	Destination	29	8%	45	13%	284	79%	358
	Time of Day	31	9%	46	13%	279	78%	356
	Trip Route	32	9%	46	13%	278	78%	356
Scenario 4: Decrease in vehicular fatalities by 100 persons per year	Name	12	3%	74	21%	270	76%	356
	Address	12	3%	78	22%	266	75%	356
	Vehicle	18	5%	114	32%	220	63%	352
	Origin	20	6%	143	40%	192	54%	355
	Destination	21	6%	143	40%	190	54%	354
	Time of Day	21	6%	142	40%	190	54%	353
	Trip Route	22	6%	141	40%	186	53%	349
Scenario 5: Decrease in vehicular fatalities by 1,000 persons per year	Name	12	3%	104	29%	241	68%	357
	Address	12	3%	109	31%	236	66%	357
	Vehicle	19	5%	134	38%	202	57%	355
	Origin	20	6%	165	46%	172	48%	357
	Destination	20	6%	164	46%	173	48%	357
	Time of Day	20	6%	164	46%	172	48%	356
	Trip Route	21	6%	161	46%	170	48%	352
Scenario 6: Collecting agencies sell data to third parties	Name	125	35%	12	3%	221	62%	358
	Address	129	36%	13	4%	217	60%	359
	Vehicle	183	51%	19	5%	154	43%	356
	Origin	217	60%	19	5%	123	34%	359
	Destination	217	60%	20	6%	122	34%	359
	Time of Day	212	59%	20	6%	126	35%	358
	Trip Route	216	61%	18	5%	123	34%	357

Here, it is evident that for most scenarios respondents reported that their compensation requirements would not change despite various benefits evaluated. The clearest digression from this general pattern occurs for Scenario 6 (Collection agencies will sell information to third parties (such as Google, NAVTEQ, or Ford)), where a majority of respondents reported that their compensation requirements for vehicle, origin, destination, trip time of day, and trip route information would increase. Clear patterns for compensation decreases are also seen for the same variables in Scenarios 4 and 5 (decrease in vehicular fatalities), with a slightly higher incidence of compensation decreases seen for Scenario 5, which tested a higher degree of fatality savings.

As with the overall compensation factors, Cronbach's alpha was tested for correlation of responses to each of the above scenarios, and were found to be significant. Thus, responses were combined into composite variables for each scenario tested.

To test the hypothesis that respondents will, in general, require lower degrees of compensation for safety and efficiency benefits than for economic benefits or if they are aware that companies will receive monetary benefits for their information, a multivariate analysis of variance (MANOVA) test was run. A MANOVA tests for the difference in two or more vectors of means, allowing us to: (1) explore one statistical test on several correlated dependent variables (scenario compensation levels required), and (2) explore how the independent variable (baseline compensation) influences the patterns of the dependent variables. In this analysis, we have the vectors shown in Table XLIII included in the data matrix of interest:

Table XLIII: Vectors of Interest in the Data Matrix

Respondent	Respondent Cluster	Baseline Compensation	Scenario 1	Scenario 2	Scenario 3	Scenario 4	Scenario 5	Scenario 6
1	1	$0.00 - $0.10	Increase	Increase	Increase	Increase	Increase	Increase
2	2	$0.11 - $0.25	Decrease	Decrease	Decrease	Decrease	Decrease	Decrease
3	3	$0.26 - $0.50	No Change	No Change	No Change	No Change	No Change	No Change
.		$0.51 - $1.00						
.		$1.01 - $5.00						
.		>$5.00						
n		Would not sell						

Here, we are interested in two primary relationships: 1) The relationship between the respondent cluster and overall compensation requirements, and, 2) The relationship between changes in desired compensation and type of scenario tested (safety, efficiency, or economic). For the MANOVA analysis, three separate models were run (one for each privacy cluster) with the following structure:

$$V_t = V_{ef} + V_s + V_c + V_{tp} + V_{(ef*s*c*tp)} + V_e$$

where V refers to the values in the cluster matrix, and:

V_t =Total variability
V_{ef} =Variability due to efficiency savings
V_s =Variability due to safety benefits
V_c =Variability due to cost savings
V_{tp} =Variability due to third party access
$V_{(ef*s*c*tp)}$ =Interaction variability
V_e =Error variability

The results of the MANOVA analysis are shown in Table XLIV below.

Table XLIV: Results of MANOVA Analysis Testing Cluster in Relation to Compensation Changes

Dependent Variable	Pr > F	R-Square	Coeff Var
Compensation, scenario 1, personal information	<.0001	0.45	27.27
Compensation, scenario 1, travel information	<.0001	0.32	33.47
Compensation, scenario 2, personal information	<.0001	0.35	32.28
Compensation, scenario 2, travel information	<.0001	0.29	35.60
Compensation, scenario 3, personal information	<.0001	0.37	31.94
Compensation, scenario 3, travel information	<.0001	0.29	36.56
Compensation, scenario 4, personal information	<.0001	0.33	34.75
Compensation, scenario 4, travel information	<.0001	0.25	39.24
Compensation, scenario 5, personal information	<.0001	0.27	36.52
Compensation, scenario 5, travel information	<.0001	0.21	39.25
Compensation, scenario 6, personal information	<.0001	0.38	44.81
Compensation, scenario 6, travel information	<.0001	0.19	58.41

Here, it is seen that changes in compensation across all tested scenarios were significant with respect to the Cluster categorization of the respondent. Coefficients of variance were highest for scenario 6, which tested variations in compensation levels requested if collected data were to be sold to third-parties, though the r^2 obtained for this statistic was fairly low with respect to travel data. Overall, correlations between cluster type and compensation required were lower in general for travel data and higher for personal data (name and address) across all scenarios, indicating that privacy preferences are likely more tied to what consumers perceive as personally identifying information. This finding, however, is somewhat disturbing given that travel data, which consumers seem more likely to sell for lower costs, may be used to

determine personally identifying information, as described above. General goodness-of-fit measures, including Wilks' Lambda and the Hotelling-Lawley Trace, were significant at <0.001, indicating that we may reject the null hypothesis that cluster type (used here as a proxy for overall privacy preferences) will not impact compensation required for the sharing of data across scenarios. Less statistical variation is seen across how respondents reacted to scenario types here, though overall variations were seen in analysis of the raw data above.

Hypothesis 7: Willingness to Trade and Compensation Factors will be influenced by knowledge, risk, and privacy cluster characteristics

Figure 10 hypothesizes a multipath model linking context, benefits, knowledge factors, privacy issues and willingness to trade. As shown in Table XXVII, a number of factors have been developed from surveyed elements to reflect these constructs. Here, we test the hypothesized relationships via the use of simultaneous equation modeling, a form of structural equation modeling (SEM – described in Section 4.4.4). The model was estimated by means of the following relationships:

$$\text{Compensation Factor} = \beta_1 \text{ Risk Factor} + \beta_2 \text{ Willingness to Trade} + \beta_3 \text{ Knowledge Factor} + \xi$$

$$\text{Willingness to Trade} = \beta_5 \text{ Knowledge Factor} + \beta_6 \text{ Risk Factor} + \beta_7 \text{ Compensation Factor} + \xi$$

$$\text{Risk Factor} = \beta_8 \text{ Knowledge Factor} + \beta_9 \text{ Willingness to Trade} + \beta_{10} \text{ Compensation Factor} + \xi$$

$$\text{Knowledge Factor} = \beta_{11} \text{ Willingness to Trade} + \beta_{12} \text{ Risk Factor} + \beta_{13} \text{ Compensation Factor} + \xi$$

Here, β (the slope coefficient) is calculated based on the covariance matrix of endogenous variables, or variables in the causal model whose values are determined by the states of other variables in the system, while ξ corresponds to the error variable. Simultaneous equations are run due to the overlapping paths between variables, which are hypothesized to have similar influences on each of the endogenous variables. In this system, we hypothesize that the covariances of the endogenous variables will balance for each equation.

SAS Proc Calis was used to estimate the coefficients. In this model, a linear equations (LINESQ) model has been used, which uses the following form for each of the simultaneous models:

$$\eta = \beta\eta + \gamma\xi$$

Here, β and γ are coefficient matrices, and η and ξ are vectors of random variables. The components of η correspond to the endogenous variables of Willingness to Trade and Compensation, while the coefficient matrix of β describes the relationships among the endogenous variables of η. The default estimation method used in Proc Calis is maximum likelihood, which assumes a multivariate normal distribution of the manifest variables.

In this model, there are two endogenous variables, namely, Factor3 (willingness to trade) and Factor4 (compensation factor), and five exogenous variables, namely, the two manifest variables of Factor1 (knowledge) and Factor2 (risk), as well as two latent error terms of e_1 and e_a. The convergence criterion was satisfied in the model, indicating that a mathematical solution has been found.

Predicted and calculated covariances were obtained for the model, and residuals were calculated, resulting in the residual matrix shown in Table XLV.

Table XLV: Residual Matrix of SEM Model

Variables		V1	V2	V3	V4	V5	V6	V7	V8	V9	V10	V11
V1	riskpriv	0.000	0.000	-0.021	-0.026	-0.040	-0.039	-0.126	-0.061	0.023	-0.084	-0.208
V2	riskpub	0.000	0.000	0.013	0.000	0.010	-0.012	-0.126	0.000	-0.002	-0.035	-0.196
V3	usettech	-0.021	0.013	0.000	0.061	-0.073	0.020	0.721	0.353	-0.255	-0.089	-0.013
V4	trcost	-0.026	0.000	0.061	0.000	0.006	-0.020	-0.009	0.026	-0.136	-0.024	0.161
V5	trsafety	-0.040	0.010	-0.073	0.006	0.000	-0.006	-0.042	0.019	-0.104	0.028	0.138
V6	tr3anon	-0.039	-0.012	0.020	-0.020	-0.006	0.000	-0.084	-0.288	-0.114	-0.003	0.226
V7	cmpper	-0.126	-0.126	0.721	-0.009	-0.042	-0.084	0.000	0.000	0.427	0.018	-0.205
V8	cmptrvl	-0.061	0.000	0.353	0.026	0.019	-0.288	0.000	0.000	0.390	0.051	-0.318
V9	readppapp	0.023	-0.002	-0.255	-0.136	-0.104	-0.114	0.427	0.390	0.000	0.013	0.036
V10	understnd	-0.084	-0.035	-0.089	-0.024	0.028	-0.003	0.018	0.051	0.013	0.000	-0.009
V11	comfort	-0.208	-0.196	-0.013	0.161	0.138	0.226	-0.205	-0.318	0.036	-0.009	0.000

Average Absolute Residual: 0.083
Average Off-diagonal Absolute Residual: 0.1

The relatively low residuals obtained indicate a moderately good model fit, a finding corroborated by a Bentler & Bonnett's normed fit index (NFI) of 0.8758.

Correlations among exogenous variables are shown in Table XLVI below:

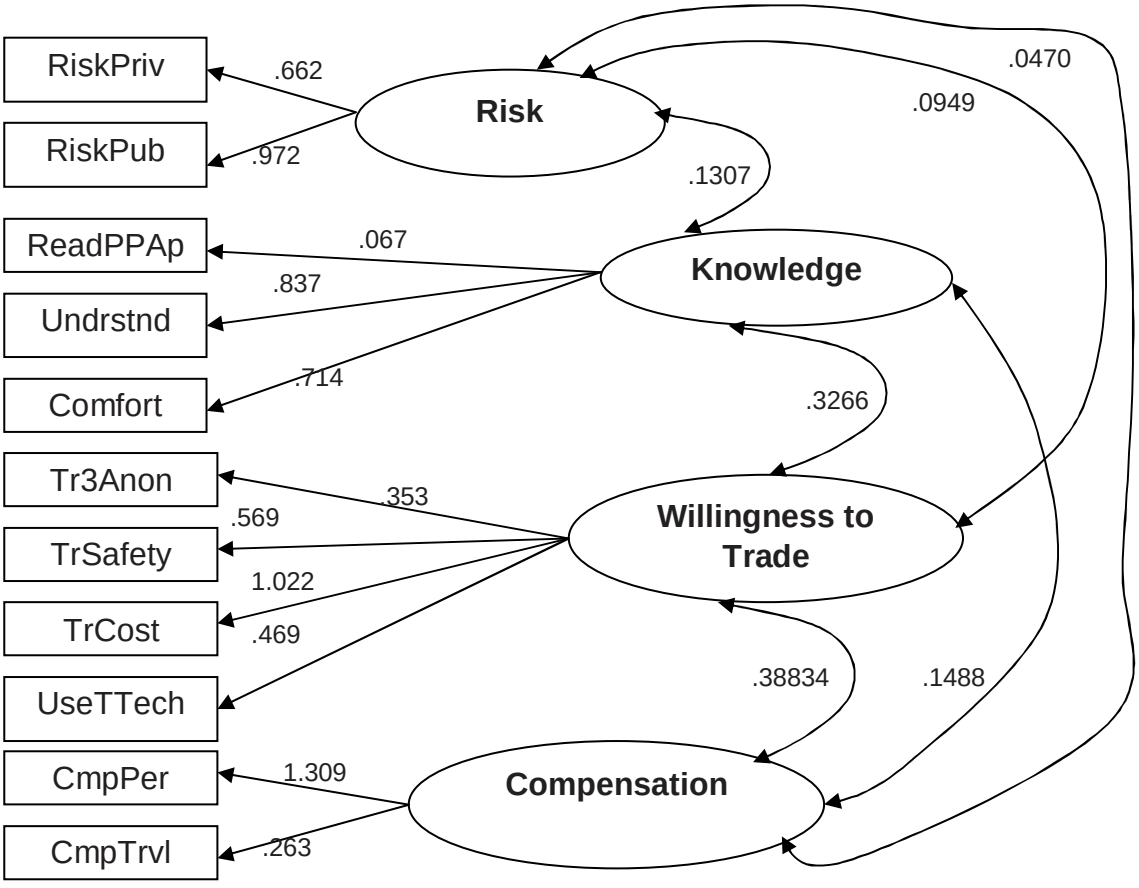
Table XLVI: Correlations Among Exogenous Variables

Var1	Var2	Estimate
Risk	Willingness to Trade	0.327
Risk	Compensation	0.149
Willingness to Trade	Compensation	0.388
Risk	Knowledge	0.131
Willingness to Trade	Knowledge	0.095
Compensation	Knowledge	0.047

These correlations indicate that the most significant relationships exist between perceived risk and willingness to trade, and between willingness to trade and compensation. Figure 14 presents the full calculated SEM model. Standardized factor loadings are shown along the arrows connecting the measured variables (shown in the rectangles) to the exogenous variables (shown in ovals), while calculated R^2 are shown along the paths connecting the exogenous

variables. Here, one may also see that the strongest connections exist between compensation and risk, and risk and willingness to trade privacy. As noted above, the fit of the overall model is moderate, with an Adjusted Goodness of Fit Index (AGFI) of 0.9074 and Root Mean Square Error (RMSEA) of 0.0779.

Figure 14: Obtained Full Structural Equation Model



Generally, findings from the models shown here indicate that the initial formulation of the model is sufficient, but could be improved. It would likely be helpful to include additional

information such as additional knowledge factors, impact of specific benefits, and personality characteristics.

6.5 Conclusions

The preceding analysis has revealed a number of findings relevant to consumer perceptions of privacy in the mobile environment. First is that, while consumers believe that sharing data in the mobile environment poses privacy risks, they do not generally take the steps necessary to address these risks, such as ensuring that a privacy policy exists or reading the applicable policy. Such a finding may indicate that privacy policies are not adequate methods of informing consumers about the risks their use of certain applications and services may pose to the privacy of their personal and travel information. This finding indicates a disconnect between consumer perceptions of privacy in the mobile environment, and the steps they take to protect this privacy.

A second finding is that privacy preferences are impacted in large part both by personal characteristics, such as gender, education, and degree of extraversion, and by the contextual factors surrounding the potential sharing of location data – such as whether data are to be shared with public or private agencies, with law enforcement, or within a social network. This finding indicates again that privacy policies may not be the most adequate default method of ensuring privacy protection, as they are static documents that may not effectively respond to privacy concerns in all situations.

Findings also indicate that willingness to trade private travel data is dependent upon a number of factors related to context, personal characteristics, expected benefits and degree of

trust in the collecting agency (as indicated by risk perceptions). Here, acknowledgement that willingness to trade is, again, not a static concept (though related to a person's overall demographics and characteristic) supports the need to expand consumer-driven or controlled methods used for privacy protection beyond traditional privacy policies. Such methods could include both technological means of privacy protection as well as policies that acknowledge contextual conflicts and promote consumer awareness and control dependent upon situational preferences. Evidence that willingness to trade may be influenced by degree and type of compensation and benefits also supports the need to incorporate appropriate incentives into LBS and ITS systems in order to encourage adoption and use. While further research is needed to address specific systemic needs, initial findings suggest that safety and efficiency benefits may have the greatest impacts on consumer perceptions of risk mitigation, particularly if combined with attention paid to types of mobility information and benefits identified as important by the consumer.

CHAPTER 7: FINDINGS AND RECOMMENDATIONS

7.1 Introduction

The analysis above reveals several underlying conclusions related to the issue of privacy in the mobile environment. The following four key findings have emerged:

- There is an overarching lack of consumer awareness related to privacy concerns in the mobile environment;
- There is a lack of consistency and comprehensiveness seen in how privacy is treated in the mobile environment;
- Perceptions of privacy in the mobile environment are heavily dependent on contextual factors; and
- Determining effective incentives for encouraging the sharing of private personal information in the mobile environment will require effectively addressing the issues above.

These major findings respond to the research questions identified above, and provide scope for further research. These key conclusions and their component findings will be further discussed below.

7.2 Consumer Awareness and Concern

The analysis above has shown that there is an overarching lack of consumer awareness of privacy issues in the mobile environment. According to the survey results discussed in Chapter 6, consumers rarely notice or read privacy policies associated with mobile technologies and services, despite believing that there are privacy risks associated with sharing data in the mobile environment. This finding may have a number of contributing factors, including the following:

- Readability: As shown in Chapter 6, privacy policies, in general, have a demonstrably higher reading level than the average U.S. consumer. Consumers may have been conditioned to expect that they will not understand the privacy policies for mobile services based on prior experience with such policies, and thus may avoid reading.
- Accessibility: Privacy policies may be included with terms of use or terms and conditions, available on websites, or linked to from a separate form. If consumers are not immediately presented with a privacy policy for a service or application, they may not take the time necessary to find and read the policy.
- Unconcern: Some consumers may assume that agencies or organizations will, by default, treat private data in accordance with consumer expectations. If consumers have a fairly high degree of trust in the agencies or organizations with which they are sharing data, they may not assume that risk is high for misuse in those instances.
- Lack of control: Users may feel that the reading of privacy policies will not provide them with any degree of control over the use of their personal data, and will thus balance their expected privacy risk with the benefits of using the service. As seen in the survey results, consumers value information related to travel efficiency and reliability, and this value may outweigh concerns related to lack of data control once information has been divulged. In this case, when consumers feel that they have little control over how their data may be used once traded they may be more willing to enter into agreements with little knowledge.

These four factors may explain some of the rationale behind the degree of consumer perception of risk associated with sharing data in the mobile environment contrasted with steps taken to educate or inform oneself via the reading of privacy policies. A further factor, however, may be that consumers are not able to make the link between their actions and the possible privacy implications. For example, one key finding from the survey analysis is that consumers overall are highly unwilling to release name and address information, but will sell or trade trip origin, destination, and route data. As shown in the literature review, such data may be easily used (depending on degree of cloaking or perturbation used) to aggregate and identify individual trip patterns. With this information, it is relatively simple to identify individual travelers and their route destinations if data are mined, or combined with publicly accessible

records such as White Pages. Thus, it is seen that travelers are unaware of the true "cost" of providing personal mobility information.

Particularly in conjunction with consumer concerns related to privacy risks in the mobile environment, such a finding is worrisome as it indicates a disconnect between consumer perception of privacy protection and actual privacy protections in the mobile environment. As consumers may assume that information related to their travel patterns may, and will, be effectively anonymized, they may have a higher tolerance threshold for sharing such data. As noted by the appellate courts, however (as reviewed in Chapter 3), access to such data provides a wealth of information on consumer habits and preferences. If consumers are unaware of such potential uses, they are incapable of effectively valuing these data, and may share more than they would otherwise be willing. Particularly in regard to the recent increase in the use of GPS enabled cell phones and smartphones, the amount of travel data available to public and private transportation and mobility agencies is growing exponentially, and with it the potential for misuse. If, for example, consumers begin to receive marketing information or political literature based on patterns of behavior or travels determined from location data sold to a third party, the public's concerns may rise and usage levels of ITS and LBS services may drop.

As shown in the analyses of privacy policies and the consumer survey, there are also multiple issues that may impact the usefulness of existing privacy policies in the context of consumer awareness and concern. A lack of standardization and readability in existing privacy policies limits the positive impacts that such policies may have on consumer expectations and trust, especially for those persons with the highest degree of privacy concerns. For those persons identified as having the highest privacy concerns (Cluster 1 in the analysis shown in

Chapter 6), a strong relationship was shown between willingness to trade private information and the total perceived risk. In short, if persons had a high perception of risk, the likelihood of their being willing to trade data was reduced.

In the context of trust issues outlined above, such a finding indicates that with a lack of awareness and guidance regarding specific privacy concerns relevant to use of collected data, persons with higher indicated privacy concerns may be less willing to adopt proposed and implemented transportation technologies due to a perceived risk of privacy invasion. Here, awareness of specific location-related privacy concerns may be subsumed by general feelings of distrust or privacy concern. As proposals are made for implementing ubiquitous mobile transportation technologies at the vehicular level, such a finding may have serious implications. If persons do not feel that they can trust the handling of their data by the manufacturers and/or mechanics of such vehicles, they may strive to "opt-out" of the system. This would, in turn, have implications for the efficacy of the system, as widespread adoption will be necessary to provide an adequate amount of real-time data. Such concerns are especially relevant in the context of public investments in mobile technologies such as networked ITS systems, as consumers with especially high degrees of privacy concerns may have an underlying distrust in public agencies. Here, lack of awareness of factors that might mitigate privacy concerns in the mobile environment (such as anonymizing techniques, cloaking, pseudonyms, and other such technological methods of privacy protection) may lead consumers to have a more heightened degree of privacy concern than is reasonable.

This lack of awareness, both of privacy concerns and protections, is a difficult issue to address. Currently, privacy policies are the primary method used to inform consumers about

the collection, use and management of their data in the mobile environment. As consumers often do not read these policies and the policies are, as shown in Chapter 5, often lacking in their coverage of issues identified as critical for privacy protection, the ability of consumers to accurately evaluate the privacy risks of sharing mobile data is compromised. Privacy protecting approaches reviewed above have focused primarily on technological and policy-oriented solutions, which essentially place the onus of privacy protection in the hands of policy-makers and technologists; however, public expectations of privacy are not necessarily reflected in the decisions of these agents.

The disconnect seen here indicates that more efforts are needed to move privacy protection "downstream," via such methods as education and grassroots efforts. While there are public-interest research groups and non-profits, such as the Electronic Privacy Information Center (epic.org) and the Privacy Rights Clearinghouse, that focus on the issue of privacy, few high-profile grassroots or consumer-driven location privacy advocacy groups exist. As consumers become more aware of privacy issues in the mobile environment, this may change; however, basic educational tools will be needed to ensure that correct information is disseminated and that the risks and benefits of sharing privacy in the mobile environment are accurately represented. Incorporating the risks of sharing data via the use of apps such as Foursquare, Facebook Places, and Waze into basic technological and computing classes would be one method of promoting greater consumer awareness, or incorporating such information into driver's education classes. As this issue gains more visibility, grassroots groups promoting greater awareness may naturally evolve, but as a complex issue, much education will be needed.

Another useful tool would be transparent and comprehensive policies related to the collection, storage, sharing, and use of data. Policies that would require implementing agencies to provide clear "opt-in" choices based on types of data to be collected and disseminated and their probable uses would encourage greater consumer awareness and understanding of the relative risks and benefits of sharing data in the mobile environment. An additional step that could be taken would be analogous to the FCRA requirement that consumers have free access yearly to credit reports, including information regarding entities that have requested a person's credit history. In this scenario, for example, location-based application providers would be required to disclose to wireless service providers those third parties with whom they have shared customer location data. This information could be aggregated and provided to customers on a yearly basis by request, in order to ensure that consumers have a clear understanding of the amount of data that has been disseminated. Taking such actions would place more onus on application developers and companies to provide clear data to consumers, and would enable consumers to make better decisions regarding sharing of data in the mobile environment. Combining these two factors would, in addition, provide consumers with a greater perception of control over their data (as they would be able to opt-out of those services which they feel are privacy invasive), thus potentially alleviating some degree of risk perception and encouraging adoption of ITS and LBS services.

Such concerns may also be allayed by ensuring that manufacturers and distributors of ubiquitous mobile transportation technologies are bound by comprehensive rules and regulations guiding the treatment and handling of data collected in the mobile environment. By providing a groundwork of trust, as demonstrated by HIPAA and the FCRA, consumers will have

a basis for determining their use of mobile technologies in accordance with reasonable expectations of privacy as codified by law. Such a step should be augmented by providing transparent and adequate information regarding degree of data protection, including appropriate information for steps to be taken if consumers are uncomfortable with protections given, or feel that their data have been mishandled. If consumers are provided with adequate and transparent methods of addressing privacy concerns, this may help increase their trust in the implemented systems, and their overall willingness to adopt mobile technologies. Additionally, if consumers are kept informed of methods being used to protect their private data, they will be more likely to trust in collecting entities.

7.3 Inconsistency and lack of comprehensiveness

As shown in Chapter 5, there are large discrepancies in how agencies, organizations and companies involved with the collection of travel data treat those data. A lack of consistent regulatory guidance forms the basis for much of this discrepancy, as voluntary guidance has not resulted in consistent attention being paid to consumer needs. Significant findings from the content analysis of privacy policies include recognition that considerable differences exist in how privacy is treated in public and private contexts, which may lead to difficulties in leveraging the use of one for benefits for both. If public agencies, for example, are to access and use data collected by private agencies, or vice-versa, significant problems may be encountered related to the potential for mining consumer information and revealing potentially sensitive information. With a lack of consistent guidance related to all aspects of data privacy in the mobile environment, including notification/awareness, choice/consent, access/participation,

integrity/security, and enforcement/redress, it is difficult for providers of ITS and LBS services to effectively plan and prepare for effective data protection.

The lack of consistency evident in the policy analysis carries with it implications for future implementation of ITS and LBS systems, as well as ubiquitous networked mobility systems. If policy-makers and the general public are not confident that collected data will be treated in a manner in keeping with privacy expectations of the general public, it is likely that funding and implementation of such systems will be stymied until such time as adequate protections are in place. By acknowledging the failures of the current approach, it may be possible to identify needed protections and begin the process of developing both technological protections that may be built into future systems, as well as developing policy guidance to ensure that these protections are adhered to. By reviewing the HIPAA and FCRA guidelines as models, it may be possible to develop consistent regulatory guidance to ensure adequate protection of privacy in the mobile environment. Addressing needs associated with differing operational models in the public and private environments will also be necessary if public-private partnerships are to be developed.

Such inconsistencies are also detrimental to the ability of consumers to determine accurate expectations of privacy in the mobile environment, or effective ways of mitigating these risks. As shown in the analysis above, there is a lack of correspondence between the content of reviewed privacy policies and consumer expectations of risk. Current legal policy regarding privacy protections in public are founded in part upon a reasonable expectation of privacy. Slobogin has shown that current practices related to the protection of privacy in the public environment do not necessarily reflect the public's expectations. The survey evaluation,

in conjunction with the analysis of privacy policies, indicates that this is also true in the mobile environment. Consumer identified risks associated with the sharing of data with third parties and law enforcement agencies indicate that they feel that these actions, particularly if undertaken without a warrant, place privacy at risk; however, many policies inform consumers that such sharing will occur if needed. Methods to encourage consumer awareness discussed above will be helpful in managing expectations of privacy in the mobile environment, as would more consistent regulation of mobile privacy policies generally.

Also as noted above, current court cases associated with mobile technologies have received a mixed reception. Technological advances have dramatically increased the availability of traveler data at a reduced cost, and legal precedent has not yet reached a point where we have adequate methods for addressing the issue in the proper context. While the February 2010 hearings raised a number of important issues associated with data collection in the commercial context, an equal number of important issues were not addressed. The first issue, as noted in the discussion of recent appellate court cases presented in Chapter 3, is the applicability of the Fourth Amendment to data collected by GPS units. The divergent findings but similar arguments presented by the two Courts indicate that there is no agreed-upon approach to determine how privacy regulations and rules are applied in the context of the current mobile environment.

The first particular concern is that of *degree.* As noted in Chapter 3, traditional expectations regarding the collection of travel data in the public environment are generally limited to physical following and detection of individual travelers only while they are on public roadways, or in the context of a warrant obtained through proper legal channels. The emergence of low-

cost GPS technologies and mobile applications on smartphones has brought the public environment into the private sphere, which creates conflict in relation to contextual integrity – spheres become muddled, and chains of data creation, awareness, ownership and sharing are left unclear. The lack of knowledge regarding ownership, in particular, is a difficult matter to address here, as most privacy policies evaluated either do not refer to who owns the data, or explicitly state that collected data are owned by the collecting entity. While most policies do indicate that collected data may be shared by the collecting entity for purposes of law enforcement, the degree of information collected may not be adequately presented to the consumer. In such cases, the expectations of the service user may be at odds with the practices of the collecting entity. Here, additional clarity with respect to collected data and potential uses would be of use to the consumer and the courts, as it would allow for more reasonable expectations to be developed. These expectations could be managed in the following ways:

1. Inform consumers of specific types of data that may be collected: While many privacy policies inform consumers that their name, email address, and various travel data may be collected, many others make only vague references to types of data that may be obtained via use of the service or application. Provision of more specific data regarding what data may be collected may provide consumers with the ability to develop more informed expectations regarding types of data that may be collected.
2. Inform consumers of potential for data uses: As shown in the analysis of privacy policies, consumers are currently not provided adequate information regarding how their data may be used by collecting agencies. Additional information regarding the potential for use by third parties, in legal contexts, and for transportation benefits (such as safety increases and efficiency improvement) may give consumers scope for making more informed decisions regarding the sharing of data. Publishing agreements between collecting entities and those with whom they share data would also be useful in this context, as this would provide consumers with better information with which to make decisions relevant to sharing of data and expectation of privacy risks.
3. Provide consumers with clear information regarding data ownership: Indicate to consumers what data will be generated via use of the service or application, and indicate specifically the agency that will be considered to be the “owner” of said data. Provide specific information regarding the extent of this ownership, including allowed uses and management in the case of account termination.

4. Provide companies with clear direction regarding federal expectations towards data collection and use: Here, it may be helpful to develop an overarching policy (such as HIPAA or the FCRA) that explicitly addresses expectations of collection, storage, management and use of data obtained through consumer use of mobile applications and services.

The methods proposed here are contingent upon the increase of consumer awareness of privacy risks and protections as more fully described in 7.2.

7.4 Contextual and Situational Factors

As described in Chapter 2, privacy has often been understood from the vantage point of context. Because of changes in spheres of influence over time and space, this understanding and findings above are particularly relevant in the locational context. Current privacy policies, described above, present a generally static method of addressing privacy, with little information provided about specific circumstances under which private data will be shared (excepting such situations as “when required by law” or “if the company is sold”) with third parties outside of the consumer’s direct control. While many LBS applications allow consumers to specify with whom certain types of information will be shared within their social network, aggregated and “anonymized” data may be shared for secondary purposes with third parties beyond the control of the user. In addition, several recent finding related to perceived privacy violations in the mobile environment (such as collection and storing of user’s location data by Apple and Google) indicate that existing policies do not adequately cover the privacy rights of individuals based upon their expectations within certain situations.

Varying degrees of privacy risk identified by consumers pursuant to the collection, storage, use and accessibility of location information by friends, government agencies,

marketers and private firms, and in the context of marketing, efficiency and law enforcement purposes indicate that, while these may all be identified as potentially risky, consumers have different degrees of comfort associated with each, or different evaluations of the potential benefits and costs. By relying primarily on static conceptions of privacy, these beliefs are devalued, thus leading to a situation where privacy is generally either over- or under-valued. In addition, it again reduces a consumer's perception of control over data, as she is not allowed to make decisions based upon her beliefs pursuant to the current context or situation, including actions, actors, and potential implications.

Addressing this issue may most beneficially be addressed via technological methods. If ITS and LBS technologies are enabled such that they allow users to set contextual preferences, users would feel a greater degree of control over their locational data within their preferred contextual constructs. In such a formulation, for example, an application could ask users a series of questions pertaining to data uses and sharing with reference to:

- Space:
    - Current location
    - Previous locations
    - Estimated future locations
- Time:
    - Time of current travel
    - Time of previous travel
    - Estimated future travel based on prior habits
- Agencies and Organizations:
    - Law enforcement
    - Emergency services
    - Marketing organizations
    - Transportation service providers
- Social spheres:
    - Family
    - Friends
    - Co-workers
    - Employers

- Employees

While this list is not complete, it provides a range of considerations and contextual cues that may be used by both service providers and consumers to better address contextual and situational realities of sharing data in the mobile environment. Development of consistent standards allowing consumers to make such determinations as a matter of rule would further assist with enhancing awareness and responding to the needs of consumers to develop reasonable privacy expectations.

This finding and associated recommendation also touches on the need to use both policy and technological methods to address privacy in the mobile environment. As noted in the literature review, a number of technological methods have been proposed to augment the possibility for privacy protection in the mobile environment. Also noted is that it is, at times, difficult for policy directives to keep pace with technological development. Here, it is recommended that policy directives developed to protect privacy in the mobile environment do not refer to specific technologies, but rather reflect the potential uses of those technologies (such as encouraging the ability to make contextual decisions). For example, it would be unwise to specify a particular type of methodology to assist consumers with constructing their contextual preferences, as these technologies will evolve over time, and may not be applicable to the next wave of ubiquitous mobile technologies. However, it will be critical for policies to address such issues as content, management, enforcement, and notification issues, as outlined in the FTC fair information policy, even if they do not make reference to specific technologies to achieve these ends.

Technological methods should be developed and implemented in such a way that policy directives are met, and the overall policy should be reviewed periodically to ensure that emerging technological innovations have not significantly changed the landscape in which privacy concerns are eminent. In short, policy should inform technology, and vice-versa. For consumers, implementing such a bi-directional system will ensure that basic standards of privacy are met in accordance with expectations developed in such areas as health care and financial matters, while also allowing scope for the development of more efficient and effective mobile technologies. In addition, by tying technological and policy methods of preserving privacy to one another, collaboration between technology developers and policy makers will be encouraged, thus ensuring that both sides are kept abreast of developments that will impact development.

7.5 Incentives

Current incentives for sharing private data in the mobile environment tend to focus on efficiency savings and social networking benefits. As seen in Chapter 6, however, survey respondents were most apt to reduce their compensation desired for access to private data in return for safety and efficiency benefits, and least likely to do so for economic benefits. Such a finding has several implications. First, it indicates that consumers are more concerned with safety improvements than economic benefits when it comes to sharing data in the mobile environment, though this finding may in part be attributable to the survey population (a limitation described in more detail in Chapter 8). If generalizable, such a finding is not surprising, as values placed on human lives tend to be higher than concerns one attributes to

paying $0.01 or $0.02 less per gallon for gas. However, this also indicates that travelers are more concerned with system-level benefits than with individual level benefits in transportation, as the questions regarding economic compensation would indicate guaranteed personal benefits, while safety benefits were presented on a system level. Here, we can infer that consumer willingness to trade will be most positively impacted if they are made aware of potential safety benefits for the planned technological systems.

Also of note is that efficiency benefits also showed an overall degree of decrease in compensation expected, indicating that consumers link mobility benefits to willingness to trade personal data. Here, it may be inferred that consumers also value time savings over economic benefits. It should be noted, however, that these results may in part reflect the makeup of the survey population. The large number of cyclists and transit users represented, along with the email notification targeting current students of a Planning program, may have impacted the degree of desirability for lowered gas prices. Gas prices tend to be fairly elastic for cyclists and users of public transit, as many do not own cars or can use alternative modes if gas prices are regarded as too high. In the case of planning students, recent “Livable Cities” and “Smart Growth” movements have highlighted the need to reduce vehicular traffic, thus persons studying these movements may be less inclined to view reductions in gas prices as a positive incentive. Efficiency savings, on the other hand, could be seen as beneficial for all users of the system (including those who travel primarily by bicycle or public transit), as well as having potential positive environmental impacts.

These findings indicate that, for applications and services that will require the consumer to share personal data, highlighting the safety and efficiency benefits will likely have the most

impact on potential adoption. If acknowledgement of these benefits is combined with clear information regarding the use and protection of consumer data (as recommended above), potential for adoption and use of services will likely be greater than if benefits are framed in terms of cost impacts. While the findings reported here do not provide a degree of increase or decrease, they do indicate that name and address information (viewed by most as more personal) are viewed as more valuable than travel data. This may be attributable to the fact that personal information is viewed as more revealing (as name and address information may be linked to multiple data sources and do not tend to change), or it may reflect that trip data are viewed as less "private" as they are generally created in the public environment. In either case, developing appropriate incentives to encourage travelers to use ITS and LBS services will require effectively valuing the benefits of services versus the cost of privacy risks encountered, and will likely be most effective if consumers are given some degree of control over the costs incurred relative to the benefits gained.

It should also be noted here that if consumers are aware that collecting agencies or third parties will benefit financially from their data, they will demand higher levels of compensation. Such a finding is consistent with previous studies reviewed in section 3.2.2, and touches on the conceptions of data production and ownership. If consumers are aware that the data they are producing via use of services or applications have a quantifiable value to other parties, they may in turn expect that they should be compensated for use of this data as they may consider themselves "owners" of this data. If policies regarding ownership are not explicitly stated (as they are often not in the policies reviewed above), consumers may feel that they have been lied to or cheated if they find out that others have benefited financially from

the use or sale of their data to third parties. Such a finding highlights the need to make explicit policies regarding ownership and use of data in the mobile environment.

7.6 Implications for Ongoing Research

The dissertation findings reported here have clear implications for ongoing and emerging research, particularly in the areas of Connected Vehicles and proposed Vehicle Miles Traveled (VMT) based taxes. While concerns related to privacy have been reviewed in the context of many mobile and location applications, the recent emergence of these two topics on the transportation and mobility landscape and the privacy implications that have been noted indicate that privacy will play a key role in both their development and acceptance. In both cases, proposed initiatives have developed from ongoing research and prototype testing, which have identified a number of privacy concerns, including, but not limited to:

- Balance between privacy and security
- Ensuring protection from malicious attacks
- Ensuring comfort of the traveling public with amount of travel data shared, and the entities with which those data are shared
- Ensuring an acceptable balance between data shared and benefits received (incentives) (CBO, 2011)

The second two concerns noted here directly apply to the topics covered in this dissertation, and, as such, it is hoped that the research conducted here may help to inform and guide development of policies and techniques applicable to the implementation of these projects.

The Connected Vehicle program, which grew out of the earlier Vehicle Infrastructure Integration (VII) and IntelliDrive initiatives, is intended to support research and development in the use of vehicle-to-vehicle (V2V) and vehicle-to-infrastructure (V2I) communications for safety purposes, real time data capture and management, dynamic mobility applications and

road weather management, among other topics. Privacy has been identified as a key topic of consideration in this development, and privacy risk analysis has been slated to take place in 2011-2013, with a final privacy analysis report due in 2013 (RITA, 2011). The goal of this research, according to RITA (2011) is to, "Analyze privacy requirements to determine (a) what level of privacy is acceptable and (b) what privacy can be addressed through technical means vs. policy. Conduct a privacy roundtable to engage experts and coordinate outreach to privacy advocacy groups." The findings described above, particularly in relation to consumer awareness, policy development, and the need to consider both technological and policy-based means of protecting privacy, both support the need to conduct such research and enhance the baseline from which such discussions may take place.

Proposed VMT taxes, which would shift some of the transportation funding burden from fuel taxes to taxes based on use of the transportation network, will require monitoring of the number of miles traveled by users and, as such, open the possibility for user movements in the public sphere to be tracked and monitored for payment purposes. Some proposals have been made for minimizing the amount of data collected and used for VMT tax purposes, some in keeping with technological methods of privacy protection outlined in Section 3.7. Here, again, issues are raised in determining the balance between the willingness of persons to adopt VMT-based taxes, the need to explore new funding sources in the transportation realm, and potential privacy risks that could emerge. The CBO report identifies privacy concerns related to degree of detail of collected information, degree of access allowed by public and private agencies, and providing additional incentives to encourage adoption (such as more detailed travel information, lower costs for those who choose to participate in a VMT-based system, and

automated parking information). The report acknowledges, however, that one of the greatest determinants of adoption will be that of consumer perception of the safety of their private information should such a system be implemented. Recommendations made above for tactics that could be used to increase consumer awareness and comfort, including clear and transparent privacy policies, provision of options to allow consumers to determine with whom and for what purposes data are collected and used, and ensuring an adequate balance of incentives, may go far towards addressing this concern. It is hoped that the findings reported from the data analysis conducted above will assist with developing adequate responses and approaches towards encouraging both consumer adoption and adequate protection of private data.

7.7 Conclusions

Overall, the findings of the dissertation indicate that a great deal of work is needed to encourage consumer awareness of privacy issues in the mobile environment, and that a diversity of methods are needed to both inform consumers and ensure protection of data once collected. Implications for privacy loss in the mobile environment are concerning, particularly insofar as they relate to the overall security and use of ITS and LBS systems, and taking steps to bolster consumer trust and comfort with such systems will be necessary for the potential benefits of these systems to be fully realized. Balancing adequate incentives with accurate knowledge and transparent and consistent policies will encourage willingness to trade information in the mobile environment and allow for these systems to be implemented to the best of their potential.

CHAPTER 8: CONTRIBUTIONS, LIMITATIONS, AND DIRECTIONS FOR FUTURE RESEARCH

8.1 Summary of Findings

This dissertation has focused on location privacy issues in the realms of Intelligent Transportation Systems (ITS) and Location-Based Services (LBS). Through the literature review, review of privacy policies, and consumer survey, it has been found that, though the attention being paid to privacy in the mobile environment is growing, there is currently little understanding and even less consistency in how it is treated. A lack of consistent regulatory guidance has left the protection of privacy in the mobile environment largely in the hands of service providers, which may have negative implications for consumers. Inconsistent approaches to the protection of location privacy by both public and private providers indicates that there is currently little understanding of how location and mobility data should be managed in order to effectively ensure that consumers are comfortable with sharing this data.

As indicated in Chapter 7, the uncertainties resulting from this lack of consistency have implications for legal protections and for consumer perceptions of risk. Without consistent regulation ensuring protection of these data, the opportunities for misuse or uses not condoned by the consumer grow, which may in turn impact potential adoption of ITS and LBS technologies, as well as raising the threshold for willingness to trade information for transportation benefits. While safety and efficiency benefits may outweigh these concerns to some extent, balancing the benefits against perceptions of risk will require that consumers have adequate information to make decisions that accurately reflect their estimation of costs and benefits. It has been recommended that a mixture of technological and policy methods be

used for privacy protection in the mobile environment, and that consumers be better educated about the potential ramifications of sharing location data.

8.2 Contributions

This dissertation is predicated on the likelihood that mobile transportation services will continue to grow in ubiquity and scope. While the policies evaluated here refer to currently implemented services and applications, the findings are applicable to systems to be implemented in the future. By presenting information related to regulatory protection of privacy in the mobile environment along with an examination of consumer perceptions and trade-offs, the dissertation has endeavored to better link the issue of consumer expectations of privacy in the mobile environment to current practices related to mobile data in order to determine if there is an adequate relationship. According to the findings, implemented policies and procedures currently do not adequately protect consumer privacy to the degree expected by the U.S. population, indicating that legal obligations are not being met by transportation service and application providers. This disconnect is due, in large part, to a lack of adequate and comprehensive regulation guiding service and application providers as new technologies are developed and implemented.

Because the degree of information that can be gleaned from mobile location data is substantially greater than what can be gleaned from currently available data, this indicates a lack of adequate preparation on the part of government authorities. As noted in Chapter 7, perhaps the most effective way to address this gap is via the development of a comprehensive regulation, in the model of HIPAA or the FCRA, which would assure consumers of data

protection, and provide more adequate guidance to data collectors, users, and managers regarding appropriate uses of data and management of technology. In addition, education should be provided to consumers in order that they may better frame their expectations for privacy in the reality of the current context.

The dissertation has taken the novel approach of using content analysis to analyze existing privacy policies in the context of Federal Trade Commission regulations in order to ascertain how well service providers are currently addressing the multi-dimensional aspects of privacy. According to the findings, currently policies do not consistently and adequately address issues of notice and awareness, choice and consent, access and participation, integrity and security, or enforcement and redress. By subjecting policies to a thorough evaluation of how well they address key components of each of these categories, existing gaps in current protections may be easily seen. By, in turn, examining issues of consumer expectations and concern in relation to these gaps, we may then better ascertain what issues will prove most vital to address in the context of developing regulation. By combining analysis of public expectations and current practices as conveyed to the consumer, the dissertation has provided a more robust analysis of areas of concern regarding privacy practices in the mobile environment than has previously been conducted.

The framing of the content and survey analyses within the bounds of theory, law, and technology provides a contextual factor that addresses the socio-technological influences that must be addressed in order to guarantee: 1) that the risks associated with sharing data in the mobile environment are adequately understood by consumers, 2) that these risks are mitigated by adequate privacy protections, and 3) that the benefits of mobile technologies are not

negatively impacted by failure to adopt due to privacy concerns or by excessive limitations on data availability. By addressing these issues, the dissertation provides background and analysis that may be used in future discussions of privacy in the mobile environment, an issue that has recently been addressed by the U.S. Congress and Senate in subcommittee hearings. It is hoped that the findings of this dissertation may be used to move this discussion forward in order to address some of the practical issues associated with designing effective policy.

8.3 Limitations and Future Needs

While the dissertation has provided a fairly thorough overview of the history and current conditions regarding privacy within the United States, it is subject to some limitations. Perhaps the most obvious limitation to be seen is that of the survey population. Because a social networking and snowball approach was used, the surveyed population is not reflective of the overall makeup of the American population and the findings should be reviewed with some prudence. In the future, a more targeted survey should be conducted that reflects the overall composition of the United States population. In addition, the survey should be expanded to address questions pertaining to the importance of privacy issues in adoption of specific forms of technology as they are developed, in particular by using an approach that examines revealed preferences via *in situ* experiments, or analysis of current usage of mobile technologies in relation to privacy concerns. Such expansion of the current survey would provide more real-world examples of the sensitivity consumers have to privacy losses in the mobile environment.

A second limitation relates to the need to more effectively tie policy to technology. While a number of technological privacy protections have been reviewed, more work should be

done to adequately address how developing and planned ubiquitous mobile technologies may be designed to respond to the questions and issues raised above. In particular, this link between technology and policy should be more thoroughly examined if an overarching policy is to be put in place in order to ensure that the primary areas of concern are addressed in system design. While, as noted above, privacy is generally regarded as a secondary concern in the development of mobile technologies, a more thorough review of connections between policy and implementation would assist with ensuring that privacy is addressed on the front-end of development, thus providing more adequate protections.

A third limitation relates to the content analysis of privacy policies. Contractual language used in privacy policies is often designed for the purpose of ensuring the rights of the seller or service provider, and may include "loophole language" or purposefully vague or broad language that will limit the liability of the provider. Because the analysis used here was limited to those concepts identified in the FTC's Fair Information Practice Principles, such language has not been thoroughly addressed. Future research may be needed to adequately address this issue.

Finally, while legal issues of privacy of data in the mobile environment have been briefly reviewed, more work is needed to better ascertain consumer knowledge regarding the uses of data collected in the mobile environment for such matters as custody, divorce, or criminal cases. For these matters, it would be necessary to evaluate and analyze data agreements between companies, agencies, and organizations, policies which were not available for the current research. More information is also needed pertaining to the sharing of information between agencies and when and if this information is released to the consumer. Analysis of this information would provide better information upon which consumers could base their

expectations, and would allow for more transparency in the sharing and mining of data. Future research should more thoroughly address this issue, particularly in relation to agreements between public and private agencies.

8.4 Conclusion

This dissertation has endeavored to address overarching issues of technology, policy and personal preferences in regard to privacy in the mobile environment. The emergence of ubiquitous mobile technologies, including Intelligent Transportation Systems (ITS) and Location Based Services (LBS) brings with it the potential for great benefits, but also the potential for great misuse if privacy is not adequately addressed. Ensuring that the privacy rights of the American consumer are fully addressed in the design of these systems and the policies upon which they are based should be of highest priority, particularly as the boundaries for collecting, mining, and using data expand ever outward. The transportation network forms the backbone for much of public life, but the rights of the private citizens who travel in and on the network should be afforded great protection.

CITED LITERATURE

A. Barth, A. Datta, J. C. Mitchell, and H. Nissenbaum: Privacy and contextual integrity: Framework and applications. Proceedings of the 2006 IEEE Symposium on Security and Privacy, pages 184–198, Washington, DC, USA, 2006.

Aldenderfer, M.S. and Blashfield, R.K.: Cluster Analysis. Sage University Paper series on Quantitative Applications in the Social Sciences. Series No. 07-044. Newbury Park, CA. 1984.

Alexa, M. and C. Zuell: A review of Software for text Analysis. Mannheim, ZUMA.1999.

Ashworth, L. and C. Free: Marketing Dataveillance and Digital Privacy: Using Theories of Justice to Understand Consumers' Online Privacy Concerns. Journal of Business Ethics. Volume 67, Number 2, 107-123. 2006.

Bellman, S., G.L. Lohse, and E.J. Johnson: Predictors of Online Buying Behavior. Communications of the ACM. Volume 42 Issue 12. Dec. 1999

Berelson, B.: Content Analysis in Communication Research. New York: Free Press, 1974.

Berelson, B.: Content Analysis in Communication Research. The Free Press, Glencoe, Illinois, 1952.

Berendt, B., O. Gunther and S. Spiekermann: Privacy in E-Commerce: Stated Preferece vs. Actual Behavior. Communication of the ACM, Vol. 48, No. 3. 2005.

Bertini, R.L., C. Monsere and T. Yin: Benefits of Intelligent Transportation Systems Technologies in Urban Areas: A Literature Review. Portland State University, Center for Transportation Studies, Research Report, April 2005.

Blumberg, A.J. and P. Eckersley: On Locational Privacy and How to Avoid Losing it Forever. Electronic Frontier Foundation. Downloaded from http://www.eff.org/wp/locational-privacy. 2009.

Borgen, F.H. and D.C. Barnett: Applying cluster analysis in counseling psychology research. Journal of Counseling Psychology, Vol 34(4), 456-468, Oct 1987.

Bostrom, R. and J.S. Heinen: MIS Problems and Failures: A Socio-Technical Perspective. MIS Quarterly, Vol. 1, No. 3. 1977.

Briggs, V. and M. Walton. The Implications of Privacy Issues for Intelligent Transportation Systems (ITS) Data. Southwest Regional University Transportation Center. May 2000.

Caruso, J.B. and G. Salaway: The ECAR Study of Undergraduate Students and Technology. EDUCASE Center for Applied Research. 2007.

CIA World Factbook. Communications: United States. Downloaded from https://www.cia.gov/library/publications/the-world-factbook/geos/us.html. 2011.

Clarke, R.: Information Technology and Dataveillance. Communications of the ACM. 31,5. 498-512. May, 1988.

Congressional Budget Office (CBO): Alternative Approaches to Funding Highways. March 2011. Accessed online at: http://www.cbo.gov/ftpdocs/121xx/doc12101/03-23-HighwayFunding.pdf.

Connelly, K., A. Khalil and Y. Liu: Do What I Say?: Observed Versus Stated Privacy Preferences. INTERACT (1). Pp. 620-623.2007.

Couper, M. P.: Web surveys: A review of issues and approaches. Public Opinion Quarterly, 64, 464–494. 2000.

Court of Appeals of Wisconsin. Wisconsin v. Sveum. 2009 WI APP 81. Case No. 2008AP658-CR. May 7, 2009.

CTIA - The Wireless Association. CTIA's Best Practices and Guidelines for Location-Based Services. Available at http://files.ctia.org/pdf/CTIA_LBS_Best_Practices_Adopted_03_10.pdf. 2010.

Culnan, M.J. and P.K. Armstrong: Information privacy concerns, procedural fairness, and impersonal trust: An empirical investigation. Organization Science, 10(1), 104-115. 1999.

Culnan, M.J. and R.J. Bies: Consumer Privacy: Balancing Economic and Justice Considerations. Journal of Social Issues, 59: 323-342. 2003.

Danezis, D., S. Lewis and R. Anderson: How Much is Location Privacy Worth? Fourth Workshop on the Economics of Information Security (WEIS 2005). Harvard University. 2 - 3 June 2005.

Daykin, A.R. and P.G. Moffatt: Analyzing Ordered Responses: A Review of the Ordered Probit Model. Understanding Statistics, l(3), 157-66. 2002.

De Boni, M. and M. Prigmore: A Hegelian Basis for Information Privacy as an Economic Right. Contemporary Political Theory, Vol.3, No.2. 2004

Dediu, H.: Nielsen: Nearly 25 percent of US adults use smartphones. Asymco. Available online at: http://www.asymco.com/2010/09/14/nielsen-nearly-25-percent-of-us-adults-use-smartphones/. Sept. 14, 2010.

DeLong, J.B. and A.M. Froomkin: Speculative Microeconomics for Tomorrow's Economy. In Internet Publishing and Beyond: The Economics of Digital Information and Intellectual Property, Brian Kahin and Hal Varian, eds. Cambridge: M.I.T. Press. Pp. 6-44. 2000.

Dinev, T. and P. Hart: An Extended Privacy Calculus Model for E-Commerce Transactions. Information Systems Research, 17(1), 61-80. 2006.

Directions Magazine. Location Based Services. Available at http://www.directionsmag.com/companies/category/Location_based_Services_%28LBS%29/. 2010.

Dötzer, F.: Privacy Issues in Vehicular Ad Hoc Networks. Proceedings of the Workshop on Privacy Enhancing Technologies (PETs). 2005.

Edwards, J. R.: Multidimensional constructs in organizational behavior research: An integrative analytical framework. Organizational Research Methods, 4: 144-192. 2001.

Entner, R.: Quantifying the Mobile Data Tsunami and its Implications. NielsenWire. Accessible online at http://blog.nielsen.com/nielsenwire/online_mobile/quantifying-the-mobile-data-tsunami-and-its-implications/. June 30, 2010.

Epic.org. "The Privacy Act of 1974." Available online at http://epic.org/privacy/1974act/. Accessed February 2011.

Epic.org: The Fair Credit Reporting Act (FCRA) and the Privacy of Your Credit Report. Available online at http://epic.org/privacy/fcra/. Accessed February 2011.

Erickson, T.: Prologue: Socio-Technical Design. in Handbook of Research on Socio-Technical Design and Social Networking Systems, Brian Whitworth and Aldo de Moor. New York: Information Science Reference. 2009.

Fichman, R.G: The Diffusion and Assimilation of Information Technology Innovations. In Framing the Domains of IT Management Research: Glimpsing the Future Through the Past, R. W. Zmud, Ed. Cincinnati, OH: Pinnaflex Educational Resources. 2000.

Friendly, M.: Categorical data analysis with graphics (online). Available http://www.math.yorku.ca/SCS/Courses/grcat/grc5.html. 1995.

Fuchs, C.: The Internet as a self–organizing socio–technological system. Cybernetics & Human Knowing, volume 12, number 3, pp. 37–81. 2005.

Garrett, J.: How Rawls Could Support a Right to Privacy (online). Western Kentucky University, USA. http://www.wku.edu/~jan.garrett/ethics/rawlpriv.htm. 1995. Accessed January 2011.

Gauthier-Villars, D. and D. Ball: Mass Leak of Client Data Rattles Swiss Banking. The Wall Street Journal. Thursday, July 8, 2010.

Gidófalvi, G., X. Huang, and T. Bach Pedersen: Privacy-Preserving Data Mining on Moving Object Trajectories. Proceedings of the MDM, pp. 60-68. 2007.

Godin, S., and P. Don: Permission Marketing: Turning Strangers into Friends, and Friends into Customers. Simon and Schuster, New York, 1998.

Goldsmith, R.E., J.B. Freiden and J.K. Eastman: The generality/specificity issue in consumer innovativeness research. Technovation, Vol. 15 No. 10, pp. 601-12. 1995.

Golob, T.F.: Structural equation modeling for travel behavior research. Transportation Research Part B 37, 1–25. 2003.

Gosling, S. D., P.J. Rentfrow, and W.B. Swann: A Very Brief Measure of the Big Five Personality Domains. Journal of Research in Personality, 37, 504-528. 2003.

Graneheim, U.H. and B. Lundman: Qualitative content analysis in nursing research: Concepts, procedures and measures to achieve trustworthiness. Nurse Education Today, 24, Pp. 105–112. 2004.

Guo, J., J.P. Baugh and S. Wang: A Group Signature Based Secure and Privacy-Preserving Vehicular Communication Framework. Proceedings of the Mobile Networking for Vehicular Environments (MOVE) workshop in conjunction with IEEE INFOCOM. Anchorage, Alaska. May 2007.

Ha, Y. and L. Stoel: Internet apparel shopping behaviors: the influence of general innovativeness. International Journal of Retail & Distribution Management, Vol. 32 No. 8, pp. 377-85. 2004.

Haan, I.H., K.L. Hui, T.S. Lee, and I.P.L. Png: Online Information Privacy: Measuring the Cost-Benefit Trade-Off. Twenty-Third International Conference on Information Systems. 2002.

Halliday, J.: Google agrees to privacy reviews to settle Buzz complaint. guardian.co.uk, Wednesday 30 March 2011.

Harris Poll for TRUSTe: Smart Privacy for Smartphones: Understanding and delivering the protection consumers want. Aprl 2011.

Hatcher, L.: A Step-by-Step Approach to Using the SAS System for Factor Analysis and Structural Equation Modeling. SAS Institute, Cary, NC. 1994.

Hawkey, K. and K.M. Inkpen: Keeping Up Appearances: Understanding the Dimensions of Incidental Information Privacy. CHI 2006. April 22-27, 2006. Montréal, Québec, Canada.

Heffernan, W.C.: Fourth Amendment Privacy Interests. 92 Journal of Criminal Law and Criminology 1, 10, 12. 2001.

Helft, M.: Facebook Acknowledges Privacy Issue With Applications. The New York Times. October 18, 2010.

Herring, S.C.: Gender and Power in Online Communication. CSI Working Paper, No. WP- 01-05. University of Indiana, Bloomington. October, 2001.

Hoffman, A.: The Social Media Gender Gap. Business Week, May 19, 2008.

Hoffman, D.L., and G.R. Franke: Correspondence Analysis: Graphical Representation of Categorical Data in Marketing Research. Journal of Marketing Research, 23, 213–227. 1986.

Hoh, B., M. Gruteser, H. Xiong, and A. Alrabady: Enhancing Security and Privacy in Traffic-Monitoring Systems. Pervasive Computing. IEEE. Pp. 38-46. October – December 2006.

Holder, K.: Voting and Registration in the Election of November 2004. Current Population Reports. U.S. Census Bureau. March 2006.

Holdford, D.: Content analysis methods for conducting research in social and administrative pharmacy. Research in Social and Administrative Pharmacy. Vol. 4, Iss. 2. Pp. 173-81. 2008.

Hsieh H.-F. & S. Shannon: Three approaches to qualitative content analysis. Qualitative Health Research. 15, 1277–1288. 2005.

Hui, K.L. and I.P.L. Png: The Economics of Privacy. Economics and Information Systems, Handbooks in Information Systems, Vol. 1, Chapter 9, ed. Terrence Hendershott. Elsevier, 2006.

Intelligent Transportation Society of America. Fair Information and Privacy Principles. http://www.itsa.org/Fair_Privacy.html. Adopted January 11, 2001.

Intelligent Transportation Systems Joint Program Office (ITS JPO) of the U.S. Department of Transportation. ITS Benefits Database. Available online at: http://www.benefitcost.its.dot.gov/its/benecost.nsf/ByLink/BenefitsAbout. Last modified Oct. 16th, 2007.

Iqbal, M.U. and S. Lim: A Survey on User's Willingness-to-Pay for Privacy in Mobility Pricing Systems. International Journal of Liability and Scientific Inquiry, Vol. 1, No.3. 2008. Pp. 306 - 317.

Jensen, C. and C. Potts: Privacy Policies as Decision-Making Tools: An Evaluation of Online Privacy Notices. CHI 2004, April 24-29, 2004, Vienna, Austria. 2004.

Jensen, C., C. Potts and C. Jensen: Privacy Practices of Internet Users: Self-Reports Versus Observed Behavior. International Journal of Human-Computer Studies, 63. Pp. 203-227. 2005.

Jeyaraj, A., J.W. Rottman & M.C. Lacity: A review of the predictors, linkages, and biases and IT innovation adoption research. Journal of Information Technology, 21(1), 1–23. 2006.

Junglas, I. and C. Spitzmuller: A research model for studying privacy concerns pertaining to location-based services. HICSS '05: Proceedings of the 38th Hawaii International Conference on System Sciences, track 7 vol. 7. Washington: IEEE, p. 180.2. 2005.

Junglas, I.A. and R.T. Watson: Location Based Services. Communications of the ACM. 51-3: 65-69. 2008.

Junglas, I.A., N.A. Johnson, and C. Spitzmuller: Personality traits and concern for privacy: An empirical study in the context of location-based services. European Journal of Information Systems. 17, 387-402. 2008.

Kamat, P., A. Baliga and W. Trappe: Secure, Pseudonymous, and Auditable Communication in Vehicular Ad Hoc Networks. Security and Communication Networks, Vol. 1. Pp. 233-244. 2008.

Karvonen, H.: Different Aspects of Trust in Ubiquitous Intelligent Transportation Systems. Proceedings of ECCE 2010 Conference, Delft, The Netherlands. Pp. 311-314. August 2010.

Kassarjian, H.H.: Content Analysis in Consumer Research. The Journal of Consumer Research. Vol. 4, No. 1, pp. 8-18, June 1977.

Krause, A. and E. Horvitz: A Utility-Theoretic Approach to Privacy and Personalization. Proceedings of the 23rd Conference on Artificial Intelligence (AAAI), Track: AI and the Web. 2008.

Krippendorff, K.: Content Analysis. An Introduction to its Methodology. The Sage Commtext Series, Sage Publications Ltd., London. 1980.

Laufer, R.S. and M. Wolfe: Privacy as a Concept and a Social Issue: A Multidimensional Development Theory. Journal of Social Issues, Vol. 33, No. 3. Pp. 22-42. 1977.

Lebart, L., A. Salem and L. Berry: Exploring Textual Data. Dordrecht, The Netherlands: Kluwer Academic Publishers. 1998.

Leung, L.: Lifestyles and the use of new media technology in urban China. Telecommunications Policy, Vol. 22 No. 9, pp. 781-90. 1998.

Linowes, D.F., Chairman: Privacy Protection Study Commission. Personal Privacy in an Information Society. The Report of The Privacy Protection Study Commission. July, 1977

Liu, C., J.T. Marchewka, J. Lu and C.-S. Yu: Beyond concern: A privacy–trust- behavioral intention model of electronic commerce. Information & Management, 42, 127-142. 2004.

Lombard, M., J. Snyder-Duch, C.C. Bracken: Practical Resources for Assessing and Reporting Intercoder Reliability in Content Analysis Research Projects. (Online). Available at: http://www.temple.edu/sct/mmc/reliability/. Last Update June 1, 2010.

Lugano, G.: Mobile social software: Definition, scope and applications. eChallenges 2007 Conference, The Hague (Netherlands), pp. 1434–1441. 2007.

Lukibisi, F. B. and T. Lanyasunya: Using Principal Component Analysis to Analyze Mineral Composition Data. Proceedings of the 12th Kari Biennial Scientific Conference. Nairobi, Kenya. 8 – 12 November, 2010

Lwin, M.O. and J.D. Williams: A Model Integrating the Multidimensional Developmental Theory of Privacy and Theory of Planned Behavior to Examine Fabrication of Information Online. Marketing Letters, Vol. 14, No. 4. Pp. 257-272. 2003.

Lynch, S.M.: Course notes, Sociology 504, Princeton University. April 2003.

Magruder, J.: Law Experts Ponder Keeping Up with Technology. Arizona State University press release. Accessed online: http://www.nae.edu/Activities/Projects/CEES/26210/LawExpertsPonderKeepingUpwithTechnology.aspx. Dec. 8, 2008.

Margulis, S.T.: On the Status and Contribution of Westin's and Altman's Theories of Privacy. Journal of Social Issues, Vol. 59, No. 2. Pp. 411-429. 2003.

Markoff, J. and D. Barboza: Researchers Trace Data Theft to Intruders in China. The New York Times. Available at http://www.nytimes.com/2010/04/06/science/06cyber.html. 2010.

Matsuo, H., K.P. McIntyre, T. Tomazic and B. Katz: The Online Survey: Its Contributions and Potential Problems. Proceedings of the Joint Statistical Meetings of the American Statistical Association. 2005.

Mayring, P.: Qualitative Content Analysis. Qualitative Social Research. Vol. 1, No. 2. 2000.

McEneney, M.F. and K.F. Kaufmann: Fair Credit Reporting Act Developments. 59 Business Law 1215. 2003-2004.

Metzger, M. J.: Privacy, trust, and disclosure: Exploring barriers to electronic commerce. Journal of Computer-Mediated Communication, 9(4), article 1. 2004.

Milne, G.R. and M.E. Gordon: Direct Mail Privacy-Efficiency Tradeoffs within An Implied Social Contract Framework. Journal of Public Policy and Marketing, 12:2 (Fall), 206-215. 1993.

Morris, J.: Statement of John B. Morris, Jr.,General Counsel, and Director of CDT's Internet Standards, Technology & Policy Project Center for Democracy & Technology: The Privacy Implications of Commercial Location-Based Services. Available at http://www.cdt.org/files/pdfs/CDT-MorrisLocationTestimony.pdf. 2010.

Munnukka, J.: Characteristics of early adopters in mobile communications markets. Marketing Intelligence & Planning. Vol. 25, Iss. 7. Pp. 719-731. 2007.

Nagpaul, P.S.: Guide to Advanced Data Analysis Using IDAMS Software. UNESCO. 1999.

New York Court of Appeals. People v. Weaver. 2009 NY Int. 80. 2009 NY Slop Op 03762. May 12, 2009.

Nielsen. How Teens Use Media. Available at http://blog.nielsen.com/nielsenwire/reports/nielsen_howteensusemedia_june09.pdf. 2009.

Nissenbaum, H.: Privacy as contextual integrity. Washington Law Review, 79(1):119–158, 2004.

Nissenbaum, H.: Privacy in Context: Technology, Policy, and the Integrity of Social Life. Palo Alto, CA: Stanford University Press, 2010.

Nuendorf, K.A.: The Content Analysis Guidebook. Thousand Oaks, California: Sage Publications. 2002.

Ottens, M., M. Franssen and P. Kroes: Modelling infrastructures as socio-technical systems. International Journal of Critical Infrastructures. 2 (2–3), 133–145. 2006.

Palen, J.: The Need for Surveillance in Intelligent Transportation Systems. Intellimotion, Vol 6, No 1, University of California PATH, Berkeley, CA, pp 1-3, 10. 1997.

Parr, B. The Rise of Foursquare in Numbers (STATS). Mashable. Downloaded from http://mashable.com/2010/03/12/foursquare-stats/. 2010.

Pedersen, P.E.: Adoption of mobile internet services: an exploratory study of mobile commerce early adopters. Journal of Organizational Computing and Electronic Commerce, vol. 15, no. 3, pp. 203-221, 2005.

Perpermans, R., G. Verleye and S.V. Cappellen: Wallbanking, innovativeness and computer attitudes: 25 to 40 year-old ATM users on the spot. Journal of Economic Psychology, Vol. 17 No. 6, pp. 731-48. 1996.

Phelps, J., G. Novak, and E. Ferrell: Privacy Concerns and Consumers Willingness to Provide Personal Information. Journal of Public Policy and Marketing. 19, 27-41. 2000.

Pilon, M.: Data Theft Hits 3.3 Million Borrowers. The Wall Street Journal. Available at http://online.wsj.com/article/SB10001424052702304434404575150024174102954.html. 2010.

Popping, R.: Some views on agreement to be used in content analysis studies. Quality & Quantity. Volume 44, Number 6, 1067-1078. 2010.

Provalis Research: WordStat - Content analysis module for SIMSTAT and QDA miner. Montreal, QC. 2010.

RapLeaf: Rapleaf Study Reveals Gender and Age Data of Social Network Users. San Francisco, CA - July 29, 2008. Available from http://www.rapleaf.com/company_press_2008_07_29.html.

Rawls, J.: A Theory of Justice. Harvard University Press, Cambridge, MA. 1999.

Raya, M. and J.P. Hubaux: Securing Vehicular Ad Hoc Networks. Journal of Computer Security, 15. Pp. 39-68. 2007.

Research and Innovative Technologies Administration (RITA): V2V/V2I Safety Policy Roadmap, version 5.0. May 2011. Accessed online at: http://www.its.dot.gov/connected_vehicle/pdf/Safety_PolicyPublicRoadmap_v5.pdf.

Rogers, E. M.: New Product Adoption and Diffusion. Journal of Consumer Research (2), pp. 290-301. March 1976.

Rogers, E.M. and F.F. Shoemaker: Communication of Innovations: A Cross-Cultural Approach. New York: Free Press. 1971.

Rogers, E.M.: The 'Critical Mass' in the Diffusion of Interactive Technologies in Organizations. In The Information Systems Research Challenge: Survey Research Methods, Volume 3, K. L. Kraemer, J. I. Cash and J. F. Nunamaker (Ed.), Harvard Business School Research Colloquium, Boston, 1991, Rogers, E.M. Diffusion of Innovations, The Free Press, New York, 1995.

Ruan, X., B Yu, J. Xu and L. Yang: A Secure Privacy-Preserving Hierarchical Location Service for Mobile Ad Hoc Networks. Mobile Ad-Hoc and Sensor Networks. Volume 4864, Pp. 760-771. 2007.

Schneiderman, B. Foreword. In Handbook of Research on Socio-Technical Design and Social Networking Systems, Brian Whitworth and Aldo de Moor. New York: Information Science Reference. 2009.

Sheedy, C. and P. Kumaraguru: A Contextual Method for Evaluating Privacy Preferences. Policies and Research in Identity Management (IDMAN), Rotterdam, The Netherlands, October 11 - 12, 2007.

Sheehan, K.B. and M.G. Hoy: Dimensions of Privacy Concern Among Online Consumers. Journal of Public Policy & Marketing. Vol. 19(1). 62-73. Spring 2000.

Siegler, MG.: Google Latitude Has 3 Million Active Users, Check-Ins Likely On The Way. TechCrunch. Downloaded from http://techcrunch.com/2010/05/06/google-latitude-users-check-in/. 2010.

Singletary, M. W.: Mass communication research: Contemporary methods and applications. Boston: Addison-Wesley. 1993.

Slobogin, C.: Public Privacy: Camera Surveillance of Public Places and the Right to Anonymity. Mississippi Law Journal, 72: 213-299, 2002.

Sneath, P.H.A. and R.R. Sokal: Numerical Taxonomy. W.H. Freeman, San Francisco: 1973.

Steiniger, S., M. Neun and A. Edwardes: Foundations of Location Based Services. CartouCHel – Lecture Notes on LBS, V.1.0. 2006.

Stemler, S.: An overview of content analysis. Practical Assessment. Research & Evaluation, 7(17). Retrieved February 21, 2011 from http://PAREonline.net/getvn.asp?v=7&n=17. 2001.

Stevens, H. and L. Goasduff: Gartner Forecasts Worldwide Location-Based Services to Grow Nearly 170 Per Cent in 2008. Press Release, Gartner, Inc. Egham, UK, February 7, 2008.

Stewart, K.A. and A.H. Segars: An Empirical Examination of the Concern for Information Privacy Instrument. Information Systems Research, Vol. 13, No. 1. Pp. 36-49. March 2002.

Stigler, G.: An Introduction to Privacy in Economics and Politics. The Journal of Legal Studies, Vol. 9, No. 4, The Law and Economics of Privacy. Pp. 623-644. Dec. 1980.

Stoelting, R.: Structural Equation Modeling/Path Analysis. Online: http://www2.chass.ncsu.edu/garson/pa765/structur.htm. 2001.

Subcommittee on Commerce, Trade, and Consumer Protection joint with the Subcommittee on Communications, Technology, and the Internet; Committee on Energy and Commerce. The Collection and Use of Location Information for Commercial Purposes. Washington, D.C. Draft transcript downloaded from http://energycommerce.house.gov/Press_111/20100224/transcript.02.24.2010.cti.ctcp.pdf. 2010.

Sydell, L.: Facebook Flap Highlights Growing Privacy Concerns. Morning Edition, NPR. October 19, 2010.

Tang, L., X. Hong and P.G. Bradford: Privacy-preserving secure relative localization in vehicular networks. Security and Communication Networks, Vol. 1. Pp.195-204. 2008.

The Institutional Issues Subcommittee of the National VII Coalition. Vehicle Infrastructure Integration: Privacy Policies Framework Version 1.0.2. February 16, 2007.

Theus, M.: Visualization of categorical data. Advances in Statistical Software, volume 6, pp. 47–55. 1997.

U.S. Code. Title 47, Chapter 5, Subchapter II, Part 1,§ 222: Privacy of Customer Information. Available at http://www.law.cornell.edu/uscode/uscode47/usc_sec_47_00000222----000-.html.

U.S. Congress: Fair Credit Reporting Act, Title 15 § 1681 et seq. Congressional Record. Washington, D.C., 1970.

U.S. Congress: The Health Insurance Portability and Accountability Act of 1996 (HIPAA). P.L. No. 104-191, 110 Stat. 1938 (1996).

U.S. Department of Education. Adult Literacy in America (NALS). National Center for Education Statistics, U.S. Dept of Education, Office of Educational Research and Improvement (NCES 1993-275). 2002.

U.S. Department of Health and Human Services. "Summary of the HIPAA Privacy Rule." Revised May 2003.

U.S. Department of Justice, Office of Justice Programs. "Federal Statutes relevant in the Information Sharing Environment." Available online at: http://www.it.ojp.gov/default.aspx?area=privacy&page=1279. Last revised April 6, 2010.

U.S. Department of Justice, Overview of the Privacy Act of 1974, 2010 Edition: http://www.justice.gov/opcl/1974privacyact-overview.htm. 2010.

U.S. Federal Trade Commission Staff Report: Protecting Consumer Privacy in an Era of Rapid Change. Federal Trade Commission. December 2010.

U.S. Federal Trade Commission: Fair Information Practice Principles. http://www.ftc.gov/reports/privacy3/fairinfo.shtm. 25 June 2007.

U.S. General Accounting Office (1996). Content Analysis: A Methodology for Structuring and Analyzing Written Material. GAO/PEMD-10.3.1. Washington, D.C.

Ullman, J.B.: Structural equation modeling. In Using Multivariate Statistics, Third Edition, B.G. Tabachnick and L.S. Fidell, Eds. HarperCollins College Publishers. New York, NY. pp. 709-819. 1996.

Walzer, M.: Spheres of Justice. New York: Basic Books. 1983.

Wathieu, L. and A. Friedman: An Empirical Approach to Understanding Privacy Valuation. Harvard Business School Marketing Research Paper 07-075. May 18, 2007.

Weber, R. P.: Basic Content Analysis, 2nd ed. Newbury Park, CA. 1990.

Weber, R.: Computer-Aided Content Analysis: A Short Primer. Qualitative Sociology 7:126-47. 1984.

Westin, A. F.: Computers, health records, and citizen rights. National Bureau of Standards Monograph No. 157. U.S. Government Printing Office, Washington, D.C., (a) p. 308, (b) p. 21, (c) pp. 219-245, (d) p. 60. December 1976.

Westin, A.: Social and Political Dimensions of Privacy. Journal of Social Issues, Vol. 59, No. 2, Pp. 431-453. 2003.

Zahn, J. and V. Rajamani: The Economics of Privacy. International Journal of Security and its Applications, Vol. 2, No. 3. Pp. 101-108. July 2008.

Zimmer, M.: Surveillance, Privacy and the Ethics of Vehicle Safety Communication Technologies. Ethics and Information Technology, 7. Pp. 201-210. 2005.

Zyskowski, J.: Apple iPhone becomes lightning rod for public's privacy fears. Federal Computer Week. May 5th, 2011.

Appendix 1: Data Dictionary for Consumer Survey

Primary Question:	Multi-Part Question	Variable Name	Response Options
Please indicate your gender		Gender	1=Male 0=Female
Please indicate your year of birth		BirthYr	
Please indicate your income category		Income	1=Less than $10,000 2=$10,000-$19,999 3=$20,000-$29,999 4=$30,000-$49,999 5=$50,000-$69,999 6=More than $70,000 7=Prefer not to answer
Please indicate the highest level of education you have completed		Education	1=8th grade or less 2=Some high school 3=High school graduate or GED 4=Some college 5=Completed 2-year college degree 6=Completed 4-year college degree 7=Master's degree 8=Doctoral degree 9=Professional degree 10=Prefer not to answer
Please tell us what form of transportation you most frequently use to go to the following activities: (Converted to dummy variables, one response per activity per mode)	Work Shopping Social visit School	ModeWork ModeShop ModeSoc ModeEdu	1=Private car 2=Public transit 3=Walking 4=Bicycle 5=Other 6=Not applicable

Listed below are a number of personality traits that may or may not apply to you. Please indicate the extent to which you agree or disagree with each statement. You should rate the extent to which the pair of traits applies to you, even if one characteristic applies more strongly than the other.	Extraverted, enthusiastic	TraitEE	1=Disagree strongly
	Critical, quarrelsome	TraitCQ	2=Disagree moderately
	Dependable, self-disciplined	TraitDSD	3=Disagree a little
	Anxious, easily upset	TraitAEU	4=Neither agree nor disagree
	Open to new experiences, complex	TraitONEC	5=Agree a little
	Reserved, quiet	TraitRQ	6=Agree moderately
	Sympathetic, warm	TraitSW	7=Agree strongly
	Disorganized, careless	TraitDC	
	Calm, emotionally stable	TraitCES	
	Conventional, uncreative	TraitCU	
Variables created from personality traits - created from "Big 5" described in http://tinyurl.com/37kenmz	Extraversion	Extraver	
	Agreeableness	Agrblns	
	Conscientiousness	Conscien	
	Emotional Stability	EmotStab	
	Openness to Experiences	Opn2Exp	
Please indicate to what extent you use the following technologies:	Digital camera	UseCam	1=Do not use
	Portable digital music player (such as an iPod or Zune)	UseiPod	2=Novice user
	Palm or other personal digital device	UsePalm	3=Intermediate user
	GPS device (such as an in-car navigation system)	UseGPS	4=Expert user
	Computer	UseComp	
	Dial-up Internet service	UseDial	
	Cable or DSL internet service	UseDSL	
	Cell phone	UseCell	
	Smartphone (app-enabled phone such as an iPhone or Droid)	UseSmart	
Do you use any of the following transportation technologies or services?	Electronic Toll Pass (such as an I-Pass or E-Z Pass)	UseIPass	0=Do not use
	Electronic Transit Pass (such as a CTA or SmarTrip card)	UseETP	1=Use
	University Transit Pass (such as a UPass)	UseUPass	
	Google Latitude	UseGLat	
	OnStar	UseOnSt	

	Foursquare	UseFSq	
	Web- or phone-based mapping services (such as Google Maps or MapQuest)	UseGMap	
	Composite score based on how many Transportation Technologies used by respondent	UseTTech	
	None of the Above	UseNone	
If you use other transportation or location-based technologies, please tell us which ones below.	Open-Ended Response	UseOther	
How often do you read or skim the Terms of Use or Terms of Service before using the following types of services? (If you do not use the service, please mark "Do Not Use".) (For example, “The following are terms of a legal agreement between you and company X. By using this service, you acknowledge that you have read, understood, and agree to be bound by these terms and to comply with all applicable laws and regulations. If you do not agree to these terms, please do not use this service.”)	Electronic Toll Pass (such as an I-Pass or E-Z Pass)	ReadIPass	1=Never
	Electronic Transit Pass (such as a CTA or SmarTrip card)	ReadETP	2=Rarely
	University Transit Pass (such as a UPass)	ReadUPass	3=Sometimes
	Google Latitude	ReadGLat	4=Always
	OnStar	ReadOnSt	5=Do Not Use
	Foursquare	ReadFSq	
	Web- or phone-based mapping services (such as Google Maps or MapQuest)	ReadGMap	
How often do you notice if there is a privacy policy before using the following types of services: (For example, "Your privacy is important to Company X; maintaining your trust is important to us.")	Electronic Toll Pass (such as an I-Pass or E-Z Pass)	PPIPass	1=Never
	Electronic Transit Pass (such as a CTA or SmarTrip card)	PPETP	2=Rarely
	University Transit Pass (such as a UPass)	PPUPass	3=Sometimes
	Google Latitude	PPGLat	4=Always
	OnStar	PPOnSt	5=Do Not Use
	Foursquare	PPFSq	
	Web- or phone-based mapping services (such as Google Maps or MapQuest)	PPGMap	

How often do you read the privacy policy before using the following types of services:	Electronic Toll Pass (such as an I-Pass or E-Z Pass)	RdPPIPass	1=Never
	Electronic Transit Pass (such as a CTA or SmarTrip card)	RdPPETP	2=Rarely
	University Transit Pass (such as a UPass)	RdPPUPass	3=Sometimes
	Google Latitude	RdPPGLat	4=Always
	OnStar	RdPPOnSt	5=Do Not Use
	Foursquare	RdPPFSq	
	Web- or phone-based mapping services (such as Google Maps or MapQuest)	RdPPGMap	
Please indicate the degree to which you agree or disagree that the following actions will place your privacy at risk:	Having your location or travel data collected and stored by a private company (such as Google)	RiskPriv	1=Strongly disagree
	Having your location or travel data collected and stored by a public agency (such as a transit agency)	RiskPub	2=Disagree
	Sharing location or travel data with friends via an application such as Google Latitude or Foursquare	RiskShare	3=Neutral
	Having your location or travel data shared for marketing purposes	RiskMrkt	4=Agree
	Having your location or travel data shared for legal purposes	RiskLegal	5=Strongly agree
	Having your location or travel data shared for purposes of transportation efficiency (such as providing real-time traffic data or alternate routes)	RiskTrEf	
	Sharing identity and financial information for travel purposes (such as electronic toll collection)	RiskIDFin	
	Having location or travel information gathered by a private company (such as Google, OnStar, or Orbitz) shared with law enforcement agencies after a warrant has been issued	RskWrPr	

	Having location or travel information gathered by a private company (such as Google, OnStar, or Orbitz) shared with law enforcement agencies with no warrant issued	RskNWrPr	
	Having location or travel information gathered by a public agency (such as your state Department of Transportation) shared with law enforcement agencies after a warrant has been issued	RskWrPu	
	Having location or travel information gathered by a public agency (such as your state Department of Transportation) shared with law enforcement agencies with no warrant issued	RskNWrPu	
Which, if any, of the following do you feel are important for agencies, companies, or organizations that collect travel data to share with consumers:	What data are being collected	ShrData	0=Not important
	With whom collected data will be shared	ShrShare	1=Important
	For what purposes collected data will be shared	ShrPurShr	
	How data collected about a user may be reviewed by the user	ShrReview	
	How data collected about a user may be corrected by the user	ShrCrct	
	How data will be stored (for example, will data be stored on a secure server)	ShrStore	
	For how long collected data will be stored	ShrStTime	
How important, if at all, do you rate the following transportation information:	Knowing how long it will take to get to my destination	ImpTime	1=Not important
	Knowing alternate routes to get to my destination	ImpAlt	2=Somewhat important
	Having accurate information on changes in my travel environment, such as a crash or other congestion-causing event	ImpChange	3=Neutral

	Having reliable public transit, i.e. knowing that the train or bus will arrive at my stop at a specific time	ImpRel	4=Important
	Having improved information for travel safety, such as knowing immediately when the car in front of me has put on its brakes	ImpInfo	5=Very important
Please indicate to what degree you agree to each of following, if at all: (Note: For purposes of this question, "Travel Information" refers to such data as trip starting point, trip ending point, time of travel, route taken, and mode of transportation)	I am willing to trade some degree of privacy for transportation cost benefits	TrCost	1=Strongly disagree
	I am willing to trade some degree of privacy for transportation time savings	TrTime	2=Disagree
	I am willing to trade some degree of privacy for transportation safety benefits, such as crash reduction	TrSafety	3=Neutral
	I am willing to trade some degree of privacy for transportation security benefits, such as terrorism reduction	TrSecure	4=Agree
	I am willing to allow my travel information to be shared with third parties if I am given notice that such sharing will occur	Tr3Notice	5=Strongly agree
	I am willing to allow my travel information to be shared with third parties if I am given the opportunity to view what information is being shared	Tr3Info	
	I am willing to allow my travel information to be shared with third parties if it is made anonymous	Tr3Anon	
	I am willing to allow my travel information to be shared with third parties if it is aggregated with others' travel information	Tr3Agg	
How much, in general, would you have to be compensated to provide the following information to these agencies?	Name ($/name)	CmpName	1=$0.00 - $0.10
	Home address ($/address)	CmpAdd	2=$0.11 - $0.25
	Vehicle information (including bicycle) ($/vehicle)	CmpVeh	3=$0.26 - $0.50
	Starting point of a trip ($/trip)	CmpOrig	4=$0.51 - $1.00

	Ending point of a trip ($/trip)	CmpDest	5=$1.01 - $5.00
	Time of day at which trips are made ($/trip)	CmpTime	6=>$5.00
	Trip route and time of day ($/trip)	CmpRtTm	7=Would not sell
	Name ($/name)	Cmp1Name	1=Increase
	Home address ($/address)	Cmp1Add	2=Decrease
	Vehicle information (including bicycle) ($/vehicle)	Cmp1Veh	3=No Change
	Starting point of a trip ($/trip)	Cmp1Orig	
	Ending point of a trip ($/trip)	Cmp1Dest	
Would this amount change were it to reduce your travel time by an average of 15% per trip?	Time of day at which trips are made ($/trip)	Cmp1Time	
	Trip route and time of day ($/trip)	Cmp1RtTm	
	Name ($/name)	Cmp2Name	1=Increase
	Home address ($/address)	Cmp2Add	2=Decrease
	Vehicle information (including bicycle) ($/vehicle)	Cmp2Veh	3=No Change
	Starting point of a trip ($/trip)	Cmp2Orig	
	Ending point of a trip ($/trip)	Cmp2Dest	
Would this amount change were it to reduce the gas tax for all persons by $0.01/gallon?	Time of day at which trips are made ($/trip)	Cmp2Time	
	Trip route and time of day ($/trip)	Cmp2RtTm	
	Name ($/name)	Cmp3Name	1=Increase
	Home address ($/address)	Cmp3Add	2=Decrease
	Vehicle information (including bicycle) ($/vehicle)	Cmp3Veh	3=No Change
	Starting point of a trip ($/trip)	Cmp3Orig	
	Ending point of a trip ($/trip)	Cmp3Dest	
Would this amount change were it to reduce the gas tax for all persons by $0.02/gallon?	Time of day at which trips are made ($/trip)	Cmp3Time	
	Trip route and time of day ($/trip)	Cmp3RtTm	
	Name ($/name)	Cmp4Name	1=Increase
Would this amount change were it to decrease vehicular fatalities by 100 persons per year?	Home address ($/address)	Cmp4Add	2=Decrease
	Vehicle information (including bicycle) ($/vehicle)	Cmp4Veh	3=No Change

	Starting point of a trip ($/trip)	Cmp4Orig	
	Ending point of a trip ($/trip)	Cmp4Dest	
	Time of day at which trips are made ($/trip)	Cmp4Time	
	Trip route and time of day ($/trip)	Cmp4RtTm	
Would this amount change were it to decrease vehicular fatalities by 1,000 persons per year?	Name ($/name)	Cmp5Name	1=Increase
	Home address ($/address)	Cmp5Add	2=Decrease
	Vehicle information (including bicycle) ($/vehicle)	Cmp5Veh	3=No Change
	Starting point of a trip ($/trip)	Cmp5Orig	
	Ending point of a trip ($/trip)	Cmp5Dest	
	Time of day at which trips are made ($/trip)	Cmp5Time	
	Trip route and time of day ($/trip)	Cmp5RtTm	
Would this amount change were these agencies to then sell your information to third parties (such as Google, NAVTEQ, or Ford)?	Name ($/name)	Cmp6Name	1=Increase
	Home address ($/address)	Cmp6Add	2=Decrease
	Vehicle information (including bicycle) ($/vehicle)	Cmp6Veh	3=No Change
	Starting point of a trip ($/trip)	Cmp6Orig	
	Ending point of a trip ($/trip)	Cmp6Dest	
	Time of day at which trips are made ($/trip)	Cmp6Time	
	Trip route and time of day ($/trip)	Cmp6RtTm	
Please indicate the degree to which you agree or disagree with the following statements:	I read privacy policies before I sign up for a new service or application ("app")	ReadPPApp	1=Never
	I understand the privacy policies of most services or applications	Understnd	2=Rarely
	I am comfortable with the levels of privacy protection offered by the providers of most services or applications	Comfort	3=Neutral 4=Sometimes 5=Always

Appendix 2: Survey Means

Variable	Label	Mean	Variable	Label	Mean
Gender	Dummy	0.4433	RiskIDFin	Ordinal	3.3675
BirthYr	Cardinal	1975.2500	RskWrPr	Ordinal	3.5576
Income	Ordinal	4.0806	RskNWrPr	Ordinal	3.9553
Education	Ordinal	6.4962	RskWrPu	Ordinal	3.4674
ModeWork	Ordinal	2.3568	RskNWrPu	Ordinal	3.8586
MdWrkCr	Dummy	0.3128	ComRisk	Ordinal	3.3189
MdWrkTr	Dummy	0.3670	ShrData	Ordinal	0.9015
MdWrkWlk	Dummy	0.0542	ShrShare	Ordinal	0.9138
MdWrkBk	Dummy	0.1872	ShrPurShr	Ordinal	0.9039
MdWrkOt	Dummy	0.0025	ShrReview	Ordinal	0.7241
MdWrkNA	Dummy	0.0567	ShrCrct	Ordinal	0.6700
ModeShop	Ordinal	1.7945	ShrStore	Ordinal	0.8153
MdShpCr	Dummy	0.5714	ShrStTime	Ordinal	0.8005
MdShpTr	Dummy	0.1576	ImpTime	Ordinal	4.4921
MdShpWlk	Dummy	0.1404	ImpAlt	Ordinal	4.1099
MdShpBk	Dummy	0.1108	ImpChange	Ordinal	3.9895
MdShpOt	Dummy	0.0025	ImpRel	Ordinal	4.4632
MdShpNA	Dummy	0.0000	ImpInfo	Ordinal	3.6026
ModeSoc	Ordinal	1.9020	TrCost	Ordinal	3.5079
MdSocCr	Dummy	0.5025	TrTime	Ordinal	3.5989
MdSocTr	Dummy	0.2488	TrSafety	Ordinal	3.6296
MdSocWlk	Dummy	0.0640	TrSecure	Ordinal	3.2354
MdSocBk	Dummy	0.1576	Tr3Notice	Ordinal	2.7778
MdSocOt	Dummy	0.0025	Tr3Info	Ordinal	3.0449
MdSocNA	Dummy	0.0049	Tr3Anon	Ordinal	3.7599
ModeEdu	Ordinal	3.7812	Tr3Agg	Ordinal	3.7434
MdEduCr	Dummy	0.1182	CmpName	Ordinal	5.9066
MdEduTr	Dummy	0.3005	CmpAdd	Ordinal	6.0027
MdEduWlk	Dummy	0.0394	CmpPer	Composite	5.3387
MdEduBk	Dummy	0.1182	CmpVeh	Ordinal	4.1444
MdEduOt	Dummy	0.0000	CmpOrig	Ordinal	3.3251
MdEduNA	Dummy	0.3916	CmpDest	Ordinal	3.3003
TraitEE	Ordinal	4.6599	CmpTime	Ordinal	3.1053
TraitCQ	Ordinal	3.5279	CmpRtTm	Ordinal	3.3684
TraitDSD	Ordinal	5.7724	CmpTrvl	Composite	3.0709
TraitAEU	Ordinal	3.4504	Cmp1Name	Ordinal	2.8981
TraitONEC	Ordinal	5.8444	Cmp1Add	Ordinal	2.8788
TraitRQ	Ordinal	4.2519	Cmp1Per	Composite	2.5825
TraitSW	Ordinal	5.4784	Cmp1Veh	Ordinal	2.7320
TraitDC	Ordinal	2.8542	Cmp1Orig	Ordinal	2.6537
TraitCES	Ordinal	5.3069	Cmp1Dest	Ordinal	2.6519
TraitCU	Ordinal	2.8087	Cmp1Time	Ordinal	2.6630

Variable	Label	Mean	Variable	Label	Mean
Extraver	Cardinal	4.0751	Cmp1RtTm	Ordinal	2.6518
Agrblns	Cardinal	4.8214	Cmp1Trav	Composite	2.3759
Conscien	Cardinal	5.2574	Cmp2Name	Ordinal	2.8393
EmotStab	Cardinal	4.7574	Cmp2Add	Ordinal	2.8310
Opn2Exp	Cardinal	5.3276	Cmp2Per	Composite	2.5209
UseCam	Ordinal	2.9822	Cmp2Veh	Ordinal	2.7639
UseiPod	Ordinal	2.9364	Cmp2Orig	Ordinal	2.7389
UsePalm	Ordinal	1.9666	Cmp2Dest	Ordinal	2.7389
UseGPS	Ordinal	2.1323	Cmp2Time	Ordinal	2.7278
UseComp	Ordinal	3.6990	Cmp2RtTm	Ordinal	2.7250
UseDial	Ordinal	1.4758	Cmp2Trav	Composite	2.4286
UseDSL	Ordinal	3.5471	Cmp3Name	Ordinal	2.8417
UseCell	Ordinal	3.4010	Cmp3Add	Ordinal	2.8389
UseSmart	Dummy	2.4128	Cmp3Per	Composite	2.5185
UseIPass	Dummy	0.4532	Cmp3Veh	Ordinal	2.7591
UseETP	Dummy	0.4828	Cmp3Orig	Ordinal	2.7151
UseUPass	Dummy	0.3424	Cmp3Dest	Ordinal	2.7123
UseGLat	Dummy	0.0616	Cmp3Time	Ordinal	2.6966
UseOnSt	Dummy	0.0099	Cmp3RtTm	Ordinal	2.6927
UseFSq	Dummy	0.0517	Cmp3Trav	Composite	2.3901
UseGMap	Dummy	0.7414	Cmp4Name	Ordinal	2.7247
UseTTech	Dummy	2.1429	Cmp4Add	Ordinal	2.7135
UseNone	Dummy	0.0616	Cmp4Per	Composite	2.3842
ReadIPass	Ordinal	3.3601	Cmp4Veh	Ordinal	2.5746
ReadETP	Ordinal	3.0951	Cmp4Orig	Ordinal	2.4845
ReadUPass	Ordinal	3.5979	Cmp4Dest	Ordinal	2.4774
ReadGLat	Ordinal	4.4016	Cmp4Time	Ordinal	2.4802
ReadOnSt	Ordinal	4.5726	Cmp4RtTm	Ordinal	2.4729
ReadFSq	Ordinal	4.4213	Cmp4Trav	Composite	2.1768
ReadGMap	Ordinal	2.1344	Cmp5Name	Ordinal	2.6397
PPIPass	Ordinal	3.3127	Cmp5Add	Ordinal	2.6257
PPETP	Ordinal	3.0954	Cmp5Per	Composite	2.3214
PPUPass	Ordinal	3.4234	Cmp5Veh	Ordinal	2.5153
PPGLat	Ordinal	4.4286	Cmp5Orig	Ordinal	2.4246
PPOnSt	Ordinal	4.5643	Cmp5Dest	Ordinal	2.4274
PPFSq	Ordinal	4.4974	Cmp5Time	Ordinal	2.4274
PPGMap	Ordinal	2.2391	Cmp5RtTm	Ordinal	2.4254
RdPPIPass	Ordinal	3.1016	Cmp5Trav	Composite	2.1527
RdPPETP	Ordinal	2.8290	Cmp6Name	Ordinal	2.2646
RdPPUPass	Ordinal	3.3115	Cmp6Add	Ordinal	2.2417
RdPPGLat	Ordinal	4.3368	Cmp6Per	Composite	1.9951
RdPPOnSt	Ordinal	4.4947	Cmp6Veh	Ordinal	1.9083
RdPPFSq	Ordinal	4.3937	Cmp6Orig	Ordinal	1.7361
RdPPGMap	Ordinal	1.9457	Cmp6Dest	Ordinal	1.7333
RiskPriv	Ordinal	3.3714	Cmp6Time	Ordinal	1.7556
RiskPub	Ordinal	3.2571	Cmp6RtTm	Ordinal	1.7361

Variable	Label	Mean		Variable	Label	Mean
RiskShare	Ordinal	3.6031		Cmp6Trav	Composite	1.5729
RiskMrkt	Ordinal	3.8177		ReadPPApp	Ordinal	2.8174
RiskLegal	Ordinal	3.6132		Understnd	Ordinal	2.9510
RiskTrEf	Ordinal	2.8750		Comfort	Ordinal	3.0847

Appendix 3: Privacy Risk Case Studies

Introduction

In April of 2011, outcry arose over the release of information that Apple iPhones and Google Android smartphones routinely transmit location information to Google and Apple as part of a strategy of creating databases able to pinpoint user locations. Release of this information heightened awareness of and concerns related to privacy risk in the mobile environment, as conflicting information was presented as to the actual rationale of collecting this information, as well as its' intended use. Concerns regarding actual and potential location privacy violations have recently been growing in number and scope. The proliferation of GPS enabled devices, such as smartphones, and increasing use of radio frequency identification (RFID) systems have heightened concerns regarding risks experienced when making use of such systems. This appendix will review a representative sample of cases that have taken place in recent times.

GPS Stalking

A number of cases have been brought before the courts involving stalking via the use of GPS devices or location-enabled smartphones. One such case involved Albert Belle, a former major-league baseball player, who stalked an ex-girlfriend via the use of a GPS device attached to her vehicle (IEEE, 2006: http://spectrum.ieee.org/aerospace/satellites/stalked-by-satellite). Belle was found guilty and sentenced to jail time. Such cases have been growing in number, with the Wall Street Journal (2010) reporting on a number of cases where persons have been tracked via smartphones, often ending in violence (http://online.wsj.com/article/SB10001424052748703467304575383522318244234.html).

Law Enforcement

In April of 2011, the Los Angeles Times reported that GPS navigation device maker TomTom had apologized to consumers for selling aggregate driver data to the Dutch government for use in setting speed traps (http://latimesblogs.latimes.com/technology/2011/04/tom-tom-gps-tracking.html). According to the Times, "Algemeen Dagblad, a newspaper in the Netherlands, reported that Dutch police had obtained traffic information from the government and were setting up speed traps based on the information. On the same day, TomTom slashed its 2011 sales forecast after a weak first-quarter earnings report and announced plans to bolster slipping demand for GPS devices by focusing on services such as selling traffic data." Other law enforcement uses of data that have been construed as privacy violations by consumers include the subpoenaing of electronic toll collection data by law enforcement agencies for use in civil and criminal cases. According to the Volpe Center, "Even though toll operators do not share traveler information unless subpoenaed, privacy concerns are still present and some customers do not want their movements tracked, yet still want the convenience of ETC." (http://www.i95coalition.org/i95/Portals/0/Public_Files/pm/reports/EPS%20Best%20Practices%20and%20Convergence%20White%20Paper%20Final.pdf)

Taxing

With emerging concerns over the stability of current gas-tax based funding sources for the United States Highway Trust Fund, some transportation experts have begun to explore the potential for a use-based tax, generally recommended to be based on vehicle miles traveled. To institute such a tax will require some sort of on-board vehicle tracking device, and privacy

concerns have emerged from these suggestions. The Privacy Rights Clearinghouse, a non-profit privacy activist group, has noted the following concerns affiliated with such use-based taxing:

- Lack of consumer choice in accepting on-board tracking;
- Lack of notice provided to consumers regarding types of data that might be recorded, how and for how long it will be stored, and potential uses by an insurer or other party.
- Lack of restrictions on the use of data collected through onboard technology for purposes other than verifying miles. (https://www.privacyrights.org/ar/PayAsYouDriveAutoInsuranceAugust09.htm)

Such concerns, though contingent upon the implementation of such programs, indicate the scope of privacy violations that may arise from the use of location tracking.

Conclusion

Privacy threats in the mobile environment are both present and emerging. The few cases reviewed above present a useful, though concise, overview of potential violations, including those linked to:

- Lack of specific information related to data management provided to consumers;
- Tertiary uses of collected data;
- Intentional malfeasance or misuse of collected data;
- Personal safety; and
- Lack of control.

These concerns indicate the real potential for risk in terms of locational privacy.

Appendix 4: IRB Approval Form

UNIVERSITY OF ILLINOIS
AT CHICAGO

Office for the Protection of Research Subjects (OPRS)
Office of the Vice Chancellor for Research (MC 672)
203 Administrative Office Building
1737 West Polk Street
Chicago, Illinois 60612-7227

Exemption Granted

November 9, 2010

Caitlin Cottrill, Ph.D.
Urban Planning and Public Affairs
412 S. Peoria St., Suite 340
CUPPA HALL, M/C 357
Chicago, IL 60607
Phone: (816) 506-9737 / Fax: (312) 413-0006

RE: Research Protocol # 2010-0920
"An Analysis of Privacy in Intelligent Transportation Systems (ITS) and Location Based Services (LBS): Policy, Technology and Personal Preferences"

Dear Dr. Cottrill:

Your Claim of Exemption was reviewed on November 9, 2010 and it was determined that your research protocol meets the criteria for exemption as defined in the U. S. Department of Health and Human Services Regulations for the Protection of Human Subjects [(45 CFR 46.101(b)]. You may now begin your research.

Please note the following regarding your research:

Exemption Period:	**November 9, 2010 – November 8 2013**
Sponsor:	U.S. Department of Transportation – Federal Highway Administration
PAF#:	2008-05388
Grant/Contract No:	E5425
Grant/Contract Title:	Eisenhower Graduate Fellow
Performance Site(s):	UIC
Subject Population:	Adult subjects only
Number of Subjects:	300

The specific exemption category under 45 CFR 46.101(b) is:
(2) Research involving the use of educational tests (cognitive, diagnostic, aptitude, achievement), survey procedures, interview procedures or observation of public behavior, unless: (i) information obtained is recorded in such a manner that human subjects can be identified, directly or through identifiers linked to the subjects; and (ii) any disclosure of the human subjects' responses outside the research could reasonably place the subjects at risk of criminal or civil liability or be damaging to the subjects' financial standing, employability, or reputation.

You are reminded that investigators whose research involving human subjects is determined to be exempt from the federal regulations for the protection of human subjects still have responsibilities for the ethical conduct of the research under state law and UIC policy. Please be aware of the following UIC policies and responsibilities for investigators:

1. Amendments You are responsible for reporting any amendments to your research protocol that may affect the determination of the exemption and may result in your research no longer being eligible for the exemption that has been granted.

2. Record Keeping You are responsible for maintaining a copy all research related records in a secure location in the event future verification is necessary, at a minimum these documents include: the research protocol, the claim of exemption application, all questionnaires, survey instruments, interview questions and/or data collection instruments associated with this research protocol, recruiting or advertising materials, any consent forms or information sheets given to subjects, or any other pertinent documents.

3. Final Report When you have completed work on your research protocol, you should submit a final report to the Office for Protection of Research Subjects (OPRS).

4. Information for Human Subjects UIC Policy requires investigators to provide information about the research protocol to subjects and to obtain their permission prior to their participating in the research. The information about the research protocol should be presented to subjects in writing or orally from a written script. When appropriate, the following information must be provided to all research subjects participating in exempt studies:
    a. The researchers affiliation; UIC, JBVMAC or other institutions,
    b. The purpose of the research,
    c. The extent of the subject’s involvement and an explanation of the procedures to be followed,
    d. Whether the information being collected will be used for any purposes other than the proposed research,
    e. A description of the procedures to protect the privacy of subjects and the confidentiality of the research information and data,
    f. Description of any reasonable foreseeable risks,
    g. Description of anticipated benefit,
    h. A statement that participation is voluntary and subjects can refuse to participate or can stop at any time,

i. A statement that the researcher is available to answer any questions that the subject may have and which includes the name and phone number of the investigator(s).
j. A statement that the UIC IRB/OPRS or JBVMAC Patient Advocate Office is available if there are questions about subject's rights, which includes the appropriate phone numbers.

Please be sure to:

→Use your research protocol number (listed above) on any documents or correspondence with the IRB concerning your research protocol.

We wish you the best as you conduct your research. If you have any questions or need further help, please contact me at (312) 355-2908 or the OPRS office at (312) 996-1711. Please send any correspondence about this protocol to OPRS at 203 AOB, M/C 672.

Sincerely,

Charles W. Hoehne, CIP
Assistant Director, IRB # 2
Office for the Protection of Research Subjects

Enclosure(s): (1) Optional Form 310 - Protection of Human Subjects, Assurance Identification/Certification/Declaration

cc: Kazuya Kowamura, Urban Planning and Policy, M/C 350
Piyushimita (Vonu) Thakuriah, Urban Planning and Policy, M/C 348

Appendix 5: Vita

Education

Present — **University of Illinois-Chicago** Chicago, IL
Ph.D. candidate, Department of Urban Planning and Policy and NSF-IGERT program on Computational Transportation Science.
Dissertation title: An Analysis of Privacy in Intelligent Transportation Systems (ITS) and Location-Based Services (LBS): Policy, Technology and Personal Preferences
Expected graduation date: August, 2011
Advisor: Piyushimita Thakuriah, Department of Urban Planning and Policy

May 2003 — **University of Tennessee** Knoxville, TN, Department of Urban and Regional Planning
Master of Science, Planning
Thesis: Guiding Development: An Analysis of Zoning and Four Alternatives

May 1999 — **Vanderbilt University** Nashville, TN
Bachelor of Arts, Philosophy

Honors and Awards

Dwight D. Eisenhower Transportation Graduate Fellowship (U.S. Department of Transportation), August 2008 – August 2011
National Science Foundation IGERT Fellowship on Computational Transportation Science, August 2006 – August 2008.
10th Privacy Enhancing Technologies Symposium Travel Award, Berlin, Germany, July 2010
WTS-Chicago, Helene M. Overly Memorial Graduate Scholarship, 2010

Refereed Publications

Published

Cottrill, C. and P. Thakuriah. Protecting Location Privacy: A Policy Evaluation. Forthcoming in the *Transportation Research Record*.

Thakuriah, P., S. Sööt, C. Cottrill, N. Tilahun, E.T. Blaise, and W. Vassilakis. Integrated and Continuing Transportation Services for Seniors: Case Studies of New Freedom Program. Forthcoming in the *Transportation Research Record*.

Cottrill, C. and P. Thakuriah(2010). Evaluating Pedestrian Crashes in Areas with High Low-Income or Minority Populations. *Accident Analysis and Prevention*. Vol. 42, Iss.1879-2057. Nov. pp. 1718-1728.

Cottrill, C. (2009). Overview of Approaches to Privacy Preservation in Intelligent Transportation Systems and Vehicle Infrastructure Integration Initiative. *Transportation Research Record*: Journal of the Transportation Research Board. Vol. 2129. pp. 9-15.

Cottrill, C. and P. Thakuriah. Privacy and Gender: Reviewing women's attitudes towards privacy in the context of Intelligent Transportation Systems (ITS) and Location Based Services (LBS) (forthcoming). *Women's Issues in Transportation, Vol. II*, National Research Council.

Tonn, B. and Cottrill, C. (2004). An Environmental Plan for the Middle Nolichucky River Area, *Environmental Practice*, Vol. 6, No. 1. pp. 50-62.

Currently under review

Thakuriah, P., C. Cottrill, N. Thomas, and S. Vaughn. A Sketch Planning Methodology for Determining Interventions for Bicycle and Pedestrian Crashes: An Ecological Approach. Submitted for consideration for publication in the *Journal of the Transportation Research Forum*, Aug. 2010.

Thakuriah, P., C. Cottrill, and T. Gustafson. Planning for Intelligent Transportation Systems Deployment: A Framework Utilizing Benefits, Risks, Opportunities and Costs. Submitted for consideration for publication in *Transportation Planning and Technology*, Dec. 2010.

Refereed Conference Proceedings

Thakuriah, P., C. Cottrill, N. Thomas and S. Vaughn. (2010) Sketch Planning Methodology for Determining Interventions in Bicycle and Pedestrian Crashes: An Ecological Approach. *Proceedings of 89th Transportation Research Board Annual Conference*. 2010.

Cottrill, C. and P. Thakuriah. (2009). Privacy and Gender: Reviewing Women's Attitudes Toward Privacy in the Context of Intelligent Transportation Systems and Location-Based Services. *Proceedings of the Fourth International Conference on Women's Issues in Transportation (forthcoming).*

Thakuriah, P., C. Cottrill and T. Gustafson (2008). A Networked Approach to Planning for Deployment of Intelligent Transportation Systems. *Proceedings of Intelligent Transportation Systems World Congress*.

Cottrill, C. and P. Thakuriah (2008). GPS Adoption by Households: Evidence from a Household Travel Survey. Proceedings *of MobiQuitous 2008, First International Workshop on Computational Transportation Science IWCTS*, held in conjunction with: The Fifth International Conference on Mobile and Ubiquitous Systems: Computing, Networking and Services.

Cottrill, C. and P. Thakuriah. Evaluating Pedestrian Risk in Environmental Justice Areas. *Proceedings of 87th Transportation Research Board Annual Conference*. 2008.

Presentations

Cottrill, C. and P. Thakuriah. (Jan. 2011) Protecting Location Privacy: A Policy Evaluation. *Transportation Research Board (TRB) Annual Meeting* Washington, D.C.

Thakuriah, P., S. Sööt, C. Cottrill, N. Tilahun, E. Blaise, W. Vassilakis. (Jan. 2011.) Integrated and Continuing Transportation Services for Seniors: Case Studies of The New Freedom Program. *TRB Annual Meeting* . Washington, D.C.

Cottrill, C. and P. Thakuriah. (Oct. 2010) Evaluating Pedestrian Crashes in Areas with High Low-Income or Minority Populations. *ACSP Annual Meeting*,. Minneapolis, MN.

Cottrill, C. (Aug. 2010) Examining Privacy and Surveillance in Urban Areas: A Transportation Context. *Hot Topics in Privacy Enhancing Technologies (HotPETs).* Berlin, Germany.

Thakuriah, P., C. Cottrill, N. Thomas, and S. Vaughn (Jan. 2010). Sketch Planning Methodology for Determining Interventions for Bicycle and Pedestrian Crashes: An Ecological Approach (poster presentation). Presented in *87^{9h} TRB Annual Conference*, Washington, D.C.

Cottrill, C. and P. Thakuriah. (2009) Privacy and Gender: Reviewing Women's Attitudes Toward Privacy in the Context of Intelligent Transportation Systems and Location-Based Services. *Fourth International Conference on Women's Issues in Transportation*. Irvine, CA.

Cottrill, C. (Jan. 2009) Overview of Approaches to Privacy Preservation in Intelligent Transportation Systems and Vehicle Infrastructure Integration Initiative (poster presentation). *TRB Annual Meeting*. Washington, DC.

Cottrill, C. (2009) Examining Surveillance in Urban Areas. *Urban Affairs Association Annual Meeting* . Chicago, IL.

Cottrill, C. (2009) Protecting Privacy in Public: Surveillance Issues in Intelligent Transportation Systems. *Surveillance Societies: What Price Security?* Macaulay Honors College, New York, NY.

Thakuriah, P. and C. Cottrill. (2008) A Networked Approach to Planning For Deployment of Intelligent Transportation Systems. *ITS World Congress* . New York, NY.

Thakuriah, P. and C. Cottrill. (2008) GPS Use by Households: Early Indicators Of Privacy Preferences (poster presentation). *ITS World Congress*. New York, NY.

Cottrill, C. and P. Thakuriah. (2008) GPS Use by Households: Early Indicators of Privacy Preferences Regarding Ubiquitous Mobility Information Access." *First International Workshop on Computational Transportation Science*. Dublin, Ireland.

Cottrill, C. and P. Thakuriah. (2008) Evaluating Pedestrian Risk in Environmental Justice Areas (poster presentation). *TRB Annual Meeting* . Washington, DC.

Thakuriah, P., C. Cottrill and T. Gustafson. (2007) A Networked Approach to Planning for Deployment of Intelligent Transportation Systems. *Association of Collegiate Schools of Planning Annual Conference*. Milwaukee, WI.

Cottrill, C. and P. Thakuriah. (2007) Evaluating Pedestrian Risk in Environmental Justice Areas. *Transport Chicago*. Chicago, IL

Cottrill, C. (2005) Smart Moves – Utilizing Census Data for Transit Planning (poster presentation). *TRB: Census Data for Transportation Planning: Preparing for the Future*. Irvine, CA

Experience

Aug. 2008 – Present **Dwight David Eisenhower Transportation Fellow,** UIC Urban Transportation Center, Chicago, IL

- Analysis of gender and stakeholder issues regarding privacy and surveillance concerns in Intelligent Transportation Systems (ITS) and Location Based Services (LBS)
- Analysis of privacy policies of public and private ITS and LBS service providers
- Comprehensive survey of ITS and LBS-related privacy beliefs and preferences

Aug. 2006 – Aug. 2008 **NSF-IGERT Fellowship on Computational Transportation Science** (Department of Urban Planning and Policy, Computer Science and Urban Transportation Center)

- Initial analysis of location privacy landscape
- Review of scientific and technological approaches to privacy preservation in the mobile environment

Aug. 2006-current **Urban Transportation Center** Chicago, IL

- Evaluation of Federal Transit Administration and US Department of Labor-sponsored Job Access and Reverse Commute and New Freedom program evaluation and associated CHSTP (full form etc)
- Evaluation of transportation decision-making using multi-criteria analysis
- Analysis of pedestrian-vehicle crashes in Environmental Justice areas
- Longitudinal analysis of bicycle crashes (work to start in Jan., 2011)

May 2008 National ICT Australia (**NICTA)** Sydney, Australia
Intern

- Researched privacy implications of planned ITS projects in Sydney, Australia

June-July 2008 **ITS America** Washington, DC
Intern

- Assisted with compilation of fact sheets related to ITS projects, including costs, safety implications, and involved parties.

Aug. 2003-Aug. 2006 **Mid-America Regional Council** Kansas City, MO
Transportation Planner II

- GIS mapping, community meetings, and financial analysis for regional transit plan
- Assisted with development of Long-Range Transportation Plan
- Staff contact for Transportation Enhancements (TE), Congestion Mitigation and Air Quality (CMAQ) and Transit committees
- DBE Liaison Officer, Title VI Coordinator

Aug. 2002-Aug. 2003 **Water Resources Research Center** Knoxville, TN
Graduate Research Assistant

- Assisted with comparative analysis of three statewide volunteer water monitoring programs
- Assisted with report for Tennessee: “A Comparative Analysis of Water Quality Monitoring Programs in the Southeast: Lessons for Tennessee.”

May-Aug. 2002 **Southern Appalachian Man and the Biosphere Program** Knoxville, TN
Graduate Research Assistant

- Assisted with project examining the potential for citizen volunteer environmental monitoring to be conducted along Appalachian Trail
- Contacted potential stakeholders, assisted with defining monitoring parameters, and conducted extensive research

Aug. 2001-May 2002 **Department of Urban and Regional Planning** Knoxville, TN
Graduate Assistant

- Assisted with examining economic development in Tennessee counties with high unemployment, low per capita income, and high poverty rates
- Results of study presented to state of Tennessee

Professional Affiliations

American Institute of Certified Planners
American Planning Association

Professional Activities

Women's Transportation Seminar, Programs Committee Co-chair (2010 – 2011)

- Responsible for organization and advertisement of monthly programs attracting 30-80 attendees.
- Assist with membership and other special events

College of Urban Planning Ph.D. Students (CUPPS), Chair (2008-2009)

- Led the formation of CUPPS, including development of bylaws.
- Assisted with planning of student meetings, pedagogy workshops, and social events.

Appendix 6: Dissertation Survey

1. Notification and Consent

***1. University of Illinois at Chicago**
Research Information and Consent for Participation in Social Behavioral Research

An Analysis of Privacy in Intelligent Transportation Systems (ITS) and Location Based Services (LBS): Policy, Technology and Personal Preferences

Lead Researcher:
Caitlin Cottrill, Ph.D. Candidate
Urban Planning and Policy, University of Illinois, Chicago
412 S. Peoria St., Suite 340, CUPPA Hall, Chicago, IL 60607

You are being asked to participate in a research study evaluating people's attitudes towards privacy and how they may or may not impact potential use of transportation- and location-related technologies and applications. Completion of the survey should take 15 - 30 minutes.

Four prizes will be given to randomly selected survey participants: one (1) $100, one (1) $50 and two (2) $25 Visa gift cards. Should you elect to participate in the prize drawing, we will ask you for your email address; however, email addresses will not be linked to survey responses. Once all prizes have been distributed, all email addresses will be erased from our records. No person other than the survey administrators will have access to these addresses.

Your IP address will not be collected as part of the data collection process. No personally identifying information, such as name or telephone number, will be collected as part of the survey process, unless you elect to participate in the prize drawing.

When the results of the research are published or discussed in conferences, no information will be included that would reveal your identity.

If you choose to participate in the survey and wish to receive further information about the study, you are welcome to contact the Lead Researcher at the above address.

Your participation in this research is voluntary. Your decision whether or not to participate will not affect your current or future dealings with the University of Illinois at Chicago. If you decide to participate, you are free to withdraw at any time without affecting that

relationship.

ELECTRONIC CONSENT: Please select your choice below.

Clicking on the "agree" button below indicates that:

• you have read the above information
• you voluntarily agree to participate
• you are at least 18 years of age and a current resident of the United States

If you do not wish to participate in the research study, please decline participation by clicking on the "disagree" button.

- ○ Agree
- ○ Disagree

2. Exit Survey

Thank you for taking the time to look at our survey. Though you have chosen not to participate, we would like to ask you to respond to some voluntary questions about your reason for refusal. These questions are voluntary and confidential.

1. Please indicate all of the reasons which led you to refuse to participate in the survey:

- [ ] I did not meet the age requirement
- [ ] I did not meet the residency requirement
- [ ] I have privacy concerns
- [ ] I do not currently have time to participate
- [ ] I am not interested in the survey
- [ ] Other

2. Please indicate your gender

- ○ Male
- ○ Female
- ○ Prefer not to answer

3. Please indicate your age category

4. Please indicate the highest level of education you have completed

3. Demographics

1. Please indicate your gender

- Male
- Female
- Prefer not to answer

2. Please indicate your year of birth

3. Please indicate your income category

4. Please indicate the highest level of education you have completed

5. Please tell us your zip code

6. Please tell us what form of transportation you most frequently use to go to the following activities:

	Private car	Public transit	Walking	Bicycle	Other	Not applicable
Work						
Shopping						
Social visit						
School						

4. Personality Traits

1. Listed below are a number of personality traits that may or may not apply to you. Please indicate the extent to which you agree or disagree with each statement. You should rate the extent to which the pair of traits applies to you, even if one characteristic applies more strongly than the other.

	Disagree strongly	Disagree moderately	Disagree a little	Neither agree nor disagree	Agree a little	Agree moderately	Agree strongly
Extraverted, enthusiastic	○	○	○	○	○	○	○
Critical, quarrelsome	○	○	○	○	○	○	○
Dependable, self-disciplined	○	○	○	○	○	○	○
Anxious, easily upset	○	○	○	○	○	○	○
Open to new experiences, complex	○	○	○	○	○	○	○
Reserved, quiet	○	○	○	○	○	○	○
Sympathetic, warm	○	○	○	○	○	○	○
Disorganized, careless	○	○	○	○	○	○	○
Calm, emotionally stable	○	○	○	○	○	○	○
Conventional, uncreative	○	○	○	○	○	○	○

5. Use of Technology

1. Please indicate to what extent you use the following technologies:

	Do not use	Novice user	Intermediate user	Expert user
Digital camera	○	○	○	○
Portable digital music player (such as an iPod or Zune)	○	○	○	○
Palm or other personal digital device	○	○	○	○
GPS device (such as an in-car navigation system)	○	○	○	○
Computer	○	○	○	○
Dial-up Internet service	○	○	○	○
Cable or DSL internet service	○	○	○	○
Cell phone	○	○	○	○
Smartphone (app-enabled phone such as an iPhone or Droid)	○	○	○	○

2. Do you use any of the following transportation technologies or services?

- ☐ Electronic Toll Pass (such as an I-Pass or E-Z Pass)
- ☐ Electronic Transit Pass (such as a CTA or SmarTrip card)
- ☐ University Transit Pass (such as a UPass)
- ☐ Google Latitude
- ☐ OnStar
- ☐ Foursquare
- ☐ Web- or phone-based mapping services (such as Google Maps or MapQuest)
- ☐ None of the Above

3. If you use other transportation or location-based technologies, please tell us which ones below.

6. Terms of Use & Privacy Policies

1. How often do you read or skim the Terms of Use or Terms of Service before using the following types of services? (If you do not use the service, please mark "Do Not Use".)

(For example, “The following are terms of a legal agreement between you and company X. By using this service, you acknowledge that you have read, understood, and agree to be bound by these terms and to comply with all applicable laws and regulations. If you do not agree to these terms, please do not use this service.”)

	Never	Rarely	Sometimes	Always	Do Not Use
Electronic Toll Pass (such as an I-Pass or E-Z Pass)	○	○	○	○	○
Electronic Transit Pass (such as a CTA or SmarTrip card)	○	○	○	○	○
University Transit Pass (such as a UPass)	○	○	○	○	○
Google Latitude	○	○	○	○	○
OnStar	○	○	○	○	○
Foursquare	○	○	○	○	○
Web- or phone-based mapping services (such as Google Maps or MapQuest)	○	○	○	○	○

2. How often do you notice if there is a privacy policy before using the following types of services:

(For example, "Your privacy is important to Company X; maintaining your trust is important to us.")

	Never	Rarely	Sometimes	Always	Do Not Use
Electronic Toll Pass (such as an I-Pass or E-Z Pass)	○	○	○	○	○
Electronic Transit Pass (such as a CTA or SmarTrip card)	○	○	○	○	○
University Transit Pass (such as a UPass)	○	○	○	○	○
Google Latitude	○	○	○	○	○
OnStar	○	○	○	○	○
Foursquare	○	○	○	○	○
Web- or phone-based mapping services (such as Google Maps or MapQuest)	○	○	○	○	○

3. How often do you read the privacy policy before using the following types of services:

	Never	Rarely	Sometimes	Always	Do Not Use
Electronic Toll Pass (such as an I-Pass or E-Z Pass)	○	○	○	○	○
Electronic Transit Pass (such as a CTA or SmarTrip card)	○	○	○	○	○
University Transit Pass (such as a UPass)	○	○	○	○	○
Google Latitude	○	○	○	○	○
OnStar	○	○	○	○	○
Foursquare	○	○	○	○	○
Web- or phone-based mapping services (such as Google Maps or MapQuest)	○	○	○	○	○

7. Attitudes Towards Privacy Risk

1. Please indicate the degree to which you agree or disagree that the following actions will place your privacy at risk:

	Strongly disagree	Disagree	Neutral	Agree	Strongly agree
Having your location or travel data collected and stored by a private company (such as Google)	○	○	○	○	○
Having your location or travel data collected and stored by a public agency (such as a transit agency)	○	○	○	○	○
Sharing location or travel data with friends via an application such as Google Latitude or Foursquare	○	○	○	○	○
Having your location or travel data shared for marketing purposes	○	○	○	○	○
Having your location or travel data shared for legal purposes	○	○	○	○	○
Having your location or travel data shared for purposes of transportation efficiency (such as providing real-time traffic data or alternate routes)	○	○	○	○	○
Sharing identity and financial information for travel purposes (such as electronic toll collection)	○	○	○	○	○
Having location or travel information gathered by a private company (such as Google, OnStar, or Orbitz) shared with law enforcement agencies after a warrant has been issued	○	○	○	○	○
Having location or travel information gathered by a private company (such as Google, OnStar, or Orbitz) shared with law enforcement agencies with no warrant issued	○	○	○	○	○
Having location or travel information gathered by a public agency (such as your state Department of Transportation) shared with law enforcement agencies after a warrant has been issued	○	○	○	○	○
Having location or travel information gathered by a public agency (such as your state Department of Transportation) shared with law enforcement agencies with no warrant issued	○	○	○	○	○

8. Attitudes Towards Privacy

1. Which, if any, of the following do you feel are important for agencies, companies, or organizations that collect travel data to share with consumers:

- [ ] What data are being collected
- [ ] With whom collected data will be shared
- [ ] For what purposes collected data will be shared
- [ ] How data collected about a user may be reviewed by the user
- [ ] How data collected about a user may be corrected by the user
- [ ] How data will be stored (for example, will data be stored on a secure server)
- [ ] For how long collected data will be stored

9. Transportation Preferences

1. How important, if at all, do you rate the following transportation information:

	Not important	Somewhat important	Neutral	Important	Very important
Knowing how long it will take to get to my destination	○	○	○	○	○
Knowing alternate routes to get to my destination	○	○	○	○	○
Having accurate information on changes in my travel environment, such as a crash or other congestion-causing event	○	○	○	○	○
Having reliable public transit, i.e. knowing that the train or bus will arrive at my stop at a specific time	○	○	○	○	○
Having improved information for travel safety, such as knowing immediately when the car in front of me has put on its brakes	○	○	○	○	○

10. Privacy Preferences

1. Please indicate to what degree you agree to each of following, if at all:

(Note: For purposes of this question, "Travel Information" refers to such data as trip starting point, trip ending point, time of travel, route taken, and mode of transportation)

	Strongly disagree	Disagree	Neutral	Agree	Strongly agree
I am willing to trade some degree of privacy for transportation cost benefits	○	○	○	○	○
I am willing to trade some degree of privacy for transportation time savings	○	○	○	○	○
I am willing to trade some degree of privacy for transportation safety benefits, such as crash reduction	○	○	○	○	○
I am willing to trade some degree of privacy for transportation security benefits, such as terrorism reduction	○	○	○	○	○
I am willing to allow my travel information to be shared with third parties if I am given notice that such sharing will occur	○	○	○	○	○
I am willing to allow my travel information to be shared with third parties if I am given the opportunity to view what information is being shared	○	○	○	○	○
I am willing to allow my travel information to be shared with third parties if it is made anonymous	○	○	○	○	○
I am willing to allow my travel information to be shared with third parties if it is aggregated with others' travel information	○	○	○	○	○

11. Compensation

Government agencies and other interested parties (such as planning agencies and public transit providers) can use information, such as travel time, transportation mode (vehicle, transit, bicycle, etc.), and the starting and ending points of a trip, to make travel more efficient (for example, by providing route suggestions).

The next seven questions will ask you to estimate how much money you would need to receive to provide certain information to these agencies based on the benefits you or the general American public would receive.

1. How much, in general, would you have to be compensated to provide the following information to these agencies?

	$0.00 - $0.10	$0.11 - $0.25	$0.26 - $0.50	$0.51 - $1.00	$1.01 - $5.00	>$5.00	Would not sell
Name ($/name)	○	○	○	○	○	○	○
Home address ($/address)	○	○	○	○	○	○	○
Vehicle information (including bicycle) ($/vehicle)	○	○	○	○	○	○	○
Starting point of a trip ($/trip)	○	○	○	○	○	○	○
Ending point of a trip ($/trip)	○	○	○	○	○	○	○
Time of day at which trips are made ($/trip)	○	○	○	○	○	○	○
Trip route and time of day ($/trip)	○	○	○	○	○	○	○

2. Would this amount change were it to reduce your travel time by an average of 15% per trip?

	Increase	Decrease	No Change
Name ($/name)	○	○	○
Home address ($/address)	○	○	○
Vehicle information (including bicycle) ($/vehicle)	○	○	○
Starting point of a trip ($/trip)	○	○	○
Ending point of a trip ($/trip)	○	○	○
Time of day at which trips are made ($/trip)	○	○	○
Trip route and time of day ($/trip)	○	○	○

3. Would this amount change were it to reduce the gas tax for all persons by $0.01/gallon?

	Increase	Decrease	No Change
Name ($/name)	○	○	○
Home address ($/address)	○	○	○
Vehicle information (including bicycle) ($/vehicle)	○	○	○
Starting point of a trip ($/trip)	○	○	○
Ending point of a trip ($/trip)	○	○	○
Time of day at which trips are made ($/trip)	○	○	○
Trip route and time of day ($/trip)	○	○	○

4. Would this amount change were it to reduce the gas tax for all persons by $0.02/gallon?

	Increase	Decrease	No Change
Name ($/name)	○	○	○
Home address ($/address)	○	○	○
Vehicle information (including bicycle) ($/vehicle)	○	○	○
Starting point of a trip ($/trip)	○	○	○
Ending point of a trip ($/trip)	○	○	○
Time of day at which trips are made ($/trip)	○	○	○
Trip route and time of day ($/trip)	○	○	○

5. Would this amount change were it to decrease vehicular fatalities by 100 persons per year?

	Increase	Decrease	No Change
Name ($/name)	○	○	○
Home address ($/address)	○	○	○
Vehicle information (including bicycle) ($/vehicle)	○	○	○
Starting point of a trip ($/trip)	○	○	○
Ending point of a trip ($/trip)	○	○	○
Time of day at which trips are made ($/trip)	○	○	○
Trip route and time of day ($/trip)	○	○	○

6. Would this amount change were it to decrease vehicular fatalities by 1,000 persons per year?

	Increase	Decrease	No Change
Name ($/name)	○	○	○
Home address ($/address)	○	○	○
Vehicle information (including bicycle) ($/vehicle)	○	○	○
Starting point of a trip ($/trip)	○	○	○
Ending point of a trip ($/trip)	○	○	○
Time of day at which trips are made ($/trip)	○	○	○
Trip route and time of day ($/trip)	○	○	○

7. Would this amount change were these agencies to then sell your information to third parties (such as Google, NAVTEQ, or Ford)?

	Increase	Decrease	No Change
Name ($/name)	○	○	○
Home address ($/address)	○	○	○
Vehicle information (including bicycle) ($/vehicle)	○	○	○
Starting point of a trip ($/trip)	○	○	○
Ending point of a trip ($/trip)	○	○	○
Time of day at which trips are made ($/trip)	○	○	○
Trip route and time of day ($/trip)	○	○	○

12. Privacy Practices

1. Please indicate the degree to which you agree or disagree with the following statements:

	Never	Rarely	Neutral	Sometimes	Always
I read privacy policies before I sign up for a new service or application ("app")	○	○	○	○	○
I understand the privacy policies of most services or applications	○	○	○	○	○
I am comfortable with the levels of privacy protection offered by the providers of most services or applications	○	○	○	○	○

13. Prize Participation

1. If you wish to enter the prize drawing for a $100 (1), $50 (1), or $25 (2) Visa gift card, please enter your email address below, otherwise leave blank. Your email address will be deleted after all prizes are distributed and will not be linked to your survey answers.

14. Thank You!

Thank you for participating in our survey! Should you have any questions or wish to learn more about the study, please feel free to contact Caitlin Cottrill at ccottr2@uic.edu.