Privacy Concerns, Economic Benefits, and Consumer Decisions: A Multi-Period Panel Study of Consumer Choices in the Automobile Insurance Industry

Miremad Soleymanian
Beedie School of Business, Simon Fraser University
Assistant Professor of Marketing
Miremad_soleymanian@sfu.ca

Charles B. Weinberg
Sauder School of Business, University of British Columbia
Professor Emeritus of Marketing
charles.weinberg@sauder.ubc.ca

Ting Zhu
Krannert School of Management, Purdue University
Associate Professor of Marketing
zhu640@purdue.edu

Abstract

To examine the relationship between privacy concerns and consumer choices, we develop a finite time-horizon dynamic structural model to study consumers' adoption and use of Usage-Based Insurance (UBI). UBI enables auto insurers to collect individual-level driving data, provide feedback on driving performance, and offer individually targeted prices. Using detailed information on insurance premiums, adoption and retention decisions, and driving behavior (as measured by sensor data), we estimate the costs of using UBI including the privacy cost. During our study, the company (along with most competitors) announced an enhancement to privacy protection (limiting usage of location-based data); subsequently, there was a widely reported external data breach at a major, unrelated retailer. Our main empirical results indicate that (1) both initial and ongoing costs play crucial roles in customers' adoption and dropout decisions; (2) the enhancement in privacy policy reduces the adoption cost and is associated with higher UBI adoption; and (3) despite being an external event, the major data breach is associated with a decrease in retention rates among customers currently being monitored. Sensitivity to these events varies by gender. These results are consistent with the view that consumers trade off privacy costs for economic benefits.

Keywords: privacy; data breach; usage-based insurance; dynamic structural model

1. Introduction

Detailed data on individual consumers have become one of the most valuable assets for companies in almost all industries. The prevalence of wireless connectivity, the increasing usage of real-time sensor data, and machine-to-machine communication are presenting companies with unprecedented transformational opportunities and challenges. Firms are investing in their ability to collect, store, manage, and analyze vast amounts of valuable, individual-level data to serve customers better, gain a competitive advantage, and improve profitability (Wedel and Kannan 2016). Such vast amounts of collected data have substantial economic value. Current and historical data on individuals' traits, attributes, attitudes, and (particularly) behavior are increasingly regarded as business assets that can be used to target services or offers, to set prices, to present relevant advertising, or to be traded with other parties.

Among consumers, reactions to these developments are more ambivalent. While some appreciate the ability of companies to provide customized services and prices, there are concerns that this personal information will be used to earn higher profits at the expense of the consumer and, more importantly, that what was previously considered private will now become public. More generally, at both an individual and a societal level, individuals, government officials, and policymakers are worried about the loss of privacy as ever more personal information is collected. Some fear that such granular private information will be used in ways that are harmful to the individual in particular and society more generally.[1]

In light of these concerns, protecting consumer privacy can become a competitive advantage for data-driven companies. The companies that generate value based on collecting and processing customers' private data could better convince their customers to adopt and use their innovative products if they manage and protect consumer privacy well. In this regard, it is essential to understand how individuals perceive the value of and risk to their privacy when they decide whether or not to adopt new technology primarily driven by big data.[2] Despite the considerable discussion of privacy issues in both academic literature and the popular press,[3] little is empirically

[1] https://www.networkworld.com/article/3267065/
[2] https://digitalguardian.com/blog/us-consumers-ignorance-data-breaches-bliss
[3] For example: Acquisti et al. (2015),
https://hbr.org/2020/01/do-you-care-about-privacy-as-much-as-your-customers-do
https://www.mckinsey.com/business-functions/risk/our-insights/the-consumer-data-opportunity-and-the-privacy-imperative

known in field settings about whether the perceived loss of privacy influences consumption decisions to adopt and use new products and services.

In this paper, we examine the effect of privacy policy enhancement on the adoption and usage of an innovative program in the auto insurance industry that relies on collecting private and sensor-based data of customers. To this end, we study and quantify in a dynamic setting the trade-offs that consumers make between their cost of being monitored (including privacy) and the savings in auto insurance premiums. Such programs, known as Usage-Based Insurance (UBI), are widely offered, as we discuss more fully below, by most major US auto insurance companies. The UBI program uses location-based services (LBS) to measure different elements of actual driving behavior at the individual level. Prior to the introduction of LBS, insurance firms could not observe consumer actions and personal information at this detailed level. We build a dynamic structural model by using individual-level data-recording information and consumers' decisions about whether or not to allow their private driving behavior to be monitored over time in return for a long-term discount on their auto insurance premium. The discount is based on how well they drive while being monitored and how long they are monitored.

To control for the so-called "privacy paradox" observed in experimental studies, in which people underestimate their willingness to release private information in exchange for personalized services (Norberg et al. 2007), our empirical study is based on actual choice behavior in a field setting. Notably, during the time period of our study (2012 to 2014), the insurer, along with many other major auto insurance companies, enhanced its privacy policy to limit the use and retention of the location data of UBI customers. We are thus able to look at the effect of an enhancement to a firm's privacy policy. This event—which relates to the security of personal data and a person's privacy—allowed us to employ a quasi-experimental design and examine whether this enhancement had significant effects on the willingness of auto insurance customers in our sample to adopt, be monitored by, and to remain in the UBI program. In addition, we considered a widely reported data breach at Target department stores (announced in December 2013 during the time period of our study) to help control for and explore the effect that external factors could potentially have on the privacy-related decisions of customers. In brief, we find that the privacy policy enhancement was associated with a decline in the adoption cost of consumers and an increase of about 10% in adoption of UBI than would have otherwise occurred.

The auto insurance market (with total estimated premiums of $311 billion in 2021, https://www.ibisworld.com/industry-statistics/market-size/automobile-insurance-united-states/) is the largest insurance market segment in the US, and it is extremely competitive, as insurers try to attract the more profitable low-risk drivers and retain these customers over time. Hundreds of auto insurance writers are essentially competing for the same premium base, which is relatively stable over time, at least in the US. Auto insurance is fast becoming a big-data industry, with telematics-based auto insurance poised to potentially change the business of insurance. Usage-based insurance, an innovation that more closely aligns driving behaviors with premium rates, is a significant example of such a change. With telematics, a driver's behavior is monitored directly. Traditional rating factors, including drivers' demographics and past driving history, serve as proxies for actual driving behaviors and the risk of accidents and injury. The idea of UBI is to measure factors (such as number of hard brakes and daily mileage) that directly influence risk on an individual level and to set each policyholder's premium based on his or her monitored driving performance. By using UBI rating factors instead of traditional rating proxies people who drive very safely can earn considerable discounts on the rates offered by traditional policies. In the case of the company in our study, this option offered a maximum permanent discount of 25%, amounting to several hundred dollars annually to those who qualify. While the monitoring period lasts for a maximum of 6 months, the permanent discount is available to the UBI customer as long as she continues buying her auto insurance from the company.

UBI offers a great opportunity for insurance companies to increase their profits and market share (Reimers and Shiller 2019). For consumers, UBI programs boost affordability for lower-risk drivers, possibly compensating for the cost of losing privacy. Our empirical study focuses directly on how people's perception of privacy risk affects their use of new technologies that offer direct economic and safety benefits.

Inevitably, telematics devices involve loss of privacy for consumers who allow the insurance company to monitor where, when, and how they drive. The loss of privacy and the inconvenience of installing and using the device may limit the widespread adoption of telematics devices. Yet UBI is an excellent setting for studying the economic significance of privacy for several reasons. First, UBI is an option (as we discuss in more detail in section 3) that the customer can choose to enroll in or not. In other words, the customer can obtain the same auto insurance policy with or without agreeing to be monitored; the only difference is that the monitored customer is able to earn a discount on the premium paid. Moreover, while the maximum monitoring period in our study is 26 weeks, the customer can drop out of the program and still be covered by the same insurance policy. This is unlike many innovations, such as Google Maps, where disclosing private information is not optional if the full benefit of the service is to be realized. Second, the consumer knows what information is being monitored, as compared to many apps in which it is unclear what behaviors are actually being tracked. Third, the consumer receives a direct economic benefit, so that the cost of adopting and being monitored can be compared to the monetary value of adopting and continuing to be monitored for each individual. Fourth, given that adoption rates are sufficiently high (30% among the customers in our study), it's apparent that not only technophiles are adopting the program. Finally, during the time period of our study, the consumers experienced the insurance company's enhancement of its privacy policy, and we observe how they respond to this change.

In summary, we develop a dynamic structural model that allows us to estimate the (possibly heterogeneous) costs of using UBI, including the privacy cost, with a unique dataset of individual daily driving records for a company's new UBI customers. The clear trade-off between the expected premium saving as the long-term benefit of using UBI and the cost (including privacy) of UBI allows us to quantify the cost parameters of different groups of customers. Using these estimates as a baseline, we focus on examining the effects of the insurance company's privacy policy enhancement on changing the cost parameters, UBI adoption, and usage considering our dynamic structural model.

More specifically, in this paper we aim to answer the following research questions:

1- How can a privacy policy enhancement change the individual cost of adopting and using a new technology that relies on sharing private data? How can we quantify those changes?
2- What is the impact of a privacy policy enhancement on UBI adoption and usage? To what extent are these effects heterogeneous across demographic groups?

We use an internal database from a major US automobile insurance company to identify the effects of premium savings and costs of using UBI on participation and usage of this program based on the customers' actual behavior.[4] We have detailed, individual-level information from more than 130,000 new customers who submitted a quote request to purchase an insurance policy from March 2012 to November 2014. We observe the customers' UBI adoption decisions in addition to demographic information and premium rates of each customer. For all customers who adopted the UBI policy, we have daily information on their driving performance, which determines the discount they get on their automobile insurance premium. We also observe any dropouts from the UBI policy if the customers decide to cancel before 6 months of usage expire. (This company monitors customers for a maximum of 6 months; customers can drop out at any time.)

By developing a dynamic structural model, we consider the forward-looking behavior of customers and estimate the cost parameters. We find that the customers incur significant initial and per-period costs of using UBI, and these costs are heterogeneous across age and gender. More importantly, by considering a quasi-experimental setup due to changes in the company's privacy policy, we study the effect of internal privacy policy enhancement on UBI usage and cost parameters. The results indicate that the company's new privacy policy—which emphasizes protecting customers' location data in UBI—significantly increases the UBI adoption rate by reducing the initial cost of adopting this policy. However, for customers already in the UBI program, there is no significant increase in their willingness to continue in the program. Overall, at least in the case we study here, this suggests that improvement in privacy protection can lead to more new product adoptions.

Our research is the first marketing study in a non-digital product market to quantify the changes in individual cost (including the cost of privacy concern) of adopting and using a data-based new technology because of the company's privacy policy enhancement. This enhancement allows us to test if the consumers change their UBI usage decision in the short term due to the improvement in privacy protection. Although we cannot separate the cost of privacy from other usage costs, we are able to estimate the changes in privacy costs due to this event and find that the privacy policy enhancement is associated with a significant decrease in the perceived cost of adopting the UBI program. Using a counterfactual analysis based on our structural model, we

[4] We use the same data source as Soleymanian et al. (2019).

demonstrate that the enhanced privacy protection leads to an approximately 10% increase in consumer adoption of the UBI program compared to the original level of privacy protection.

The rest of this paper is organized as follows. After reviewing the literature related to our research questions, we discuss the sensor data used in our analysis and some key patterns observed in the data. We then outline in section 4 our dynamic structural model to capture the benefits and costs of using UBI by considering the forward-looking behavior of customers. The estimation and empirical results based on the proposed model are presented in section 5. In section 6, we describe a counterfactual analysis to examine the impact of the privacy enhancement compared to not changing the privacy policy. Finally, we offer some concluding comments on managerial and public policy issues related to the customers' responses to sharing the private data and their sensitivity to privacy concerns.

2. Literature Review

To our knowledge, our study is the first empirical research that uses sensor-based data to investigate how a privacy policy enhancement could affect the customers' decision to adopt and continue using a new usage-based pricing system. We build and estimate a model that quantifies the costs of adoption and being monitored, including privacy issues. Our paper is related to different streams of research including measuring privacy, the effect of privacy on innovations and customer responses, usage-based insurance, and dynamic structural modeling.

The challenges of defining and measuring privacy. Privacy is difficult to define, and the concept is evolving over time. There are different definitions in the literature. For example, Westin and Ruebhausen (1967) described it as the control over and safeguard of personal information. Later, Schoeman (1992) defined privacy as an aspect of dignity, autonomy, and ultimately human freedom. While seemingly different, these definitions are related, because they pertain to the boundaries between self and others, between private and shared or public features of one's life (Altman 1975). Consumers constantly navigate those boundaries, and the decisions made in this regard determine tangible and intangible benefits and costs for individuals and for society. It is often said that information is power, so control over personal information can affect the balance of economic power among parties. Lin (2019) proposes a framework for understanding why and to what extent people value their privacy. That paper distinguishes between two motives for protecting privacy: the intrinsic motive, that is, a taste for privacy; and the instrumental motive, which reflects the expected economic loss from revealing one's type specific to the transactional environment.

Acquisti et al. (2015) differentiate three themes to organize and draw connections among different streams of privacy research. The first theme is people's uncertainty about the nature of privacy trade-offs and their own preferences about them. People are often unaware of the information they are sharing. In our setting, the company specifically discloses the information to be monitored, so awareness is not likely to be a critical concern. The second theme is the powerful context-dependence of privacy preferences: The same person can in some situations be oblivious to privacy issues, and in other situations be strongly concerned about privacy. When people are uncertain about their preferences, they often search for cues in their environment to obtain guidance. In our study, the privacy policy enhancement and widely covered data breach at Target department stores provide cues to alert consumers to privacy risks and the need for protection. We test whether such events affect consumers' UBI usage and adoption behavior for car insurance. The third theme that Acquisti et al. (2015) identify is the malleability of privacy preferences. They suggest that with changes in the firms' and government's public policy, consumers' concern about privacy can be adjusted. In our study, we explore whether the effect of an improvement in the firm's privacy policy (not retaining LBS data) can affect UBI usage decisions.

While reports of attitudes and behavioral intentions are important, Norberg et al. (2007) highlight the difference between the intentions and behaviors of individuals in terms of sharing personal information; they call the relationship between individuals' intentions to disclose personal information and their actual personal disclosure behaviors the "privacy paradox." We attempt to quantify the costs of being monitored, including the privacy cost, using adoption and retention behavior of customers in the UBI context. Only a few studies examine actual behavior of customers in a field setting instead of customers' attitude and intention related to privacy. By running a field experiment, Tsai et al. (2011) found that consumers are sometimes willing to pay a price premium to purchase goods from merchants who offer more privacy-protective options. They designed an experiment in which a shopping search engine interface clearly and compactly displays privacy policy information. When such information is made available, consumers tend to purchase from online retailers who better protect their customers' privacy. Beresford et al. (2012) measured willingness to pay for privacy in a field experiment that studied the actual behavior of subjects buying a DVD from one of two competing online stores. One store consistently required more sensitive personal data than the other, but otherwise the stores were identical. In one treatment, DVDs were one Euro cheaper at the store requesting more personal information, and almost all buyers chose the cheaper store. Surprisingly, in the second treatment when prices were identical, participants bought from both shops equally often, which shows that people seem to show little sensitivity to privacy and sharing of personal data.

More closely related to our research, Kummer and Schulte (2019) studied the money-for-privacy trade-off in the market of smartphone applications. They found that cheaper apps use more privacy-sensitive permissions, and given price and functionality, demand is lower for apps with sensitive permissions. In general, their results portray the trade-off that both customers and firms consider when they decide about privacy aspects of the apps. The results in our paper extend the empirical literature beyond that of digital products.

In terms of the effect of privacy perception on changing customer behavior, Miller and Tucker (2017) explored how state laws on the privacy of genetic data affect the diffusion of personalized medicine, using data on genetic testing for cancer risks; empirical results show that approaches to genetic and health privacy that give users control over redisclosure encourage the spread of genetic testing, but that notification deters individuals from obtaining genetic tests. The authors found no effects of state genetic antidiscrimination laws on genetic testing rates. In our paper, we investigate the effect of changes in a company's privacy policy (restriction of access to the location data of customers) on customers' behavior in UBI adoption and usage.

In contrast to implementing a privacy protection policy, consumers' privacy can be violated due to harmful incidents, such as data breaches. Janakiraman et al. (2018) investigated the effect of a data-breach announcement on a multichannel retailer's customers' behavior. They found that although the data breach resulted in a significant decrease in customer spending in the physical store where the data breach occurred, customers of the firm migrated to the unbreached (internet and catalog) channels of the retailer. Martin et al.'s (2017) study of data security breaches affecting 414 public companies found that a data breach hurts the focal firm but may help rival firm. In contrast to these two papers about data breach, in our paper we examine the effect of a (widely publicized) data breach from an outside source on a company in another industry, along with identifying the effect of internal privacy policy enhancement on customers' behavior.

Privacy and innovation. While privacy has traditionally been an issue of interest to individuals and society (Bloom et al. 1994), the recent availability of low-cost technologies for data acquisition and analysis generates new concerns about personal information processing (Shapiro and Varian 1997). Laudon (1997) proposed the creation of information markets where individuals own their personal data and can transfer the rights to that data to others in exchange for some types of compensation. Following the widespread adoption of the internet and proliferation of databases containing consumer information, a number of studies documented the value to companies of detailed, individual-level behavioral data. In online advertising, by examining past surfing and click behavior, firms can learn about current needs as well as general preferences. Beales (2010) documented that in 2009 the price of behaviorally targeted advertising was 2.68 times the price of untargeted advertising. Lambrecht et al. (2011) further showed that the performance of behavioral targeting can be improved when combined with clickstream data that help to identify the consumers' degree of product search. In the health-care sector, Miller and Tucker (2011) noted that the use of patient data by hospitals helps to improve monitoring and the accuracy of patient medical histories.

The potential for a consumer's need to trade off innovation and privacy spans many industries. Bleier et al. (2020) examine how data-based innovation and marketing can trigger privacy concerns. The paper highlights several strategies firms can use to mitigate privacy concerns, and the authors observe that in some circumstances, privacy concerns may exert positive effects on data-driven marketing by stimulating privacy innovation and providing a source of competitive advantage.

Surveys of individuals repeatedly find that people are concerned about the sharing of their private information; see, for example, Westin (2005) in the health-care sector regarding digital medical records. Mao and Zhang (2014) examine the effect of privacy on location-based services (LBS) available on mobile phones and find in a survey-based study that higher privacy concern is negatively related to customers' adoption of LBS. In brief, survey-based studies show that consumers are concerned about the protection of their personal information, and this concern about privacy has a negative effect on adoption of new technologies and the consumer's relationship with companies that have access to private information. An interesting question is whether concern about privacy is increasing over time. Goldfarb and Tucker (2012) use respondents' willingness to disclose information about income in periodic surveys as a proxy for their changing concerns about privacy and find that refusals to reveal income information have risen over time. Additionally, people who are older, and females as compared to males, are consistently less likely to answer questions about their income. In our study, we also look at heterogeneity in response to perceived changes in privacy events by age and gender.

In summary, we study costs and concerns related to privacy in order to examine the effect of customers' perception about privacy cost on the success of a UBI program, and whether changing perceptions of potential privacy costs lead to differences in people's willingness to allow private behavior to be monitored.

Usage-based insurance. Current studies of UBI focus on two major benefits for the insurance companies and consumers. First, insurance companies can attract safe drivers to reduce costs of their service, while the customers self-selected into the UBI program pay a lower premium (Jin and Vasserman 2019). Second, Soleymanian et al. (2019) find that a UBI monitoring program and the economic incentives it provides can encourage UBI adopters to improve their driving behavior, which is heterogeneous across different groups of customers, and to thereby obtain a higher UBI discount. They find that the improvement in driving behavior is not limited to the period of monitoring; the UBI customers show long-term improvement even after the UBI monitoring period concludes. Reimers and Shiller (2019) investigate the value of telematics insurance for firms and customers. While innovating firms that introduced UBI experienced an initial increase in profits, the profits are eroded by entry of other insurers, which implies that this innovation does not raise novel antitrust concerns. Furthermore, they find a meaningful impact of UBI programs on reducing fatal car accidents. Shum and Xin (2020) study time-varying risk preferences among UBI drivers and show that they drive more conservatively following "near-miss" accidents. Structural estimation results of that paper indicate that such changes in behavior are consistent with an increase in risk aversion and a reduction in annual insurance premium cost. Choudhary et al. (2020) study the impact of the driver's decision to review immediate feedback on driving behavior in a UBI program. They find that, on average, users' driving performance after they review detailed feedback is nearly 14.9% worse than that of users who do not review their detailed feedback. Strong negative feedback (e.g., a sharp deterioration in performance) exerts a positive effect on short-term performance, but this only happens for very large drops in performance (3% of cases). Furthermore, these researchers demonstrate that drivers just below the insurance-incentive thresholds exert greater effort following immediate feedback.

UBI enables an insurance company to customize the product offering to each consumer based on his or her driving behavior by providing individualized price discounts. Our paper in the context of UBI goes beyond just the benefits of this program to identify the costs associated with the adoption and usage of this new policy in the trade-off with its benefits.

Dynamic structural model. To understand the consumers' UBI adoption and termination decisions, we develop a single-agent dynamic structural model in which agents are forward-looking and maximize expected intertemporal payoffs. An attractive feature of this model is that the structural parameters have a transparent interpretation within the theoretical model that frames the empirical investigation (Chintagunta and Nair 2011). Moreover, econometric models in this class are useful tools for the evaluation of alternative (counterfactual) policies.

More specifically related to adoption of new technologies and services, Yang and Ching (2013) develop a dynamic structural model to investigate consumers' adoption of and usage decisions for ATM cards when consumers are forward-looking and heterogeneous. These researchers use the nested fixed-point algorithm (Rust 1987) to estimate the structural parameters of the model. Considering the monetary benefits of adopting the innovation allows recovery of the monetary value of total adoption costs. Several other studies in the marketing literature use the discrete choice dynamic programming framework and individual-level data to assess consumers' technology-adoption decisions. Sriram et al. (2010) present a framework of durable goods purchasing behavior in related technology categories. Ryan and Tucker (2012) estimate the demand for a video-calling technology in the presence of both network effects and heterogeneity by considering a dynamic structural model. Unlike our framework, the framework of these latter two papers does not recover the monetary value of total adoption costs. As we discuss later in detail, our model can separately quantify the two cost parameters—initial cost of adoption and per-period cost of using UBI—because we observe the consumers' adoption and termination decisions along with the direct economic benefits customers expect to receive from a UBI policy.

In this paper we develop a finite-horizon dynamic structural model of consumers' decisions to adopt UBI and to allow themselves to be monitored for a period of up to 6 months at most.

3. Data

3.1. Data overview

We study customers' decision to adopt and keep the UBI policy based on data from a major US insurance company that offers the UBI program as an option alongside its traditional car insurance policy. The data cover all new customers that the company added in 15 states in a 32-month time period from March 2012 to November 2014. All new customers receive both a traditional premium quote based on a formula filed with each state's regulators and the offer of a discount if they enroll in the UBI program. Customers are free to leave the UBI program at any time and continue with the firm's traditional insurance, even though participation in the UBI program can lead to a lower premium. The UBI discount depends upon a score based on a number of factors related to actual driving behavior and how long (up to 6 months at most) the customer remains in the UBI monitoring program.[5] In other words, the customers choose to adopt and keep the UBI policy based on their belief about current driving behavior and expected future performance.

Like almost all the UBI policies in the US, this firm's UBI policy was introduced as an option that allows the customers to receive a personalized premium rate based on their actual driving behavior. The pricing strategy of the insurance company is to encourage new customers to sign up for a UBI program by offering an initial (temporary) discount (typically 5%) as soon as they enroll. The new UBI policyholder receives a telematics device to be plugged into the car, which enables the insurance company to monitor many aspects of the customer's driving behavior. The customer can monitor her own performance from real-time feedback: whenever she hard-brakes, the telematics device beeps to let her know, and she can monitor her performance on a daily basis via an app. If a customer withdraws from the UBI program before 75 days, she will no longer receive a discount. After 75 days of using the monitoring device, the customer receives an updated discount based on actual driving performance. From 75 days until 180 days, the customer can remove the telematics device and ask the company for a permanent UBI discount based on performance to date that will apply for her premium as long as she continues buying her auto insurance from the company. The monitoring period lasts for a maximum of 180 days, at which

[5] Although the UBI formula used by the company we study is confidential, The Co-operators (a major Canadian auto insurance company that offers UBI in the province of Ontario) discloses such information on its website. The Co-operators put the following weights on these four elements: sudden braking has the highest weight followed by distance travelled, late-night driving, and rapid acceleration (https://enroute.cooperators.ca/).

time the telematics device is removed, and the customer is offered a permanent UBI discount—up to 25% based on her daily driving scores after 6 months, but the average discount rate is 12% with a standard deviation of 5%. While some drivers (less than 1% in our sample) may be offered no discount, a surcharge is never imposed. Customers know the initial discount, the range of the discount, the average discount, and that no surcharge will be imposed because this information is provided on the company's website.

Our empirical research builds on a number of datasets containing information about individual drivers' auto insurance choices, their demographic characteristics, and premiums and risk scores defined by the insurance company. For the drivers choosing UBI, we observe sensor-based information on their actual daily driving behavior (UBI scores) and whether or not they drop out early (and when) from the monitoring program during the 6 months.

Table 1 reports some summary statistics about the customers in our sample. The first column shows a data summary for all customers; the second and third columns refer to non-UBI and UBI customers, respectively. The average UBI acceptance rate is about 30%. The program appears to be equally appealing to males and females, but the average age of the UBI policyholders (39.3) is much lower than that of the non-UBI customers (48.7), suggesting that the UBI program is more attractive for younger drivers. One possible explanation for this difference is that the insurance company assigns a relatively high risk to the young drivers due to the lack of sufficient driving history. Hence, this group pays a substantially higher initial premium. The UBI program offers a great opportunity for younger drivers to demonstrate their actual driving behaviors, and as a result they can receive a discount rate according to their performance. Therefore, the incentive for younger drivers seems to be higher to adopt the UBI program compared to older or experienced drivers. The higher average monthly premium of UBI drivers compared to non-UBI customers also shows that the program seems to be more attractive for the customers who are traditionally paying more, because their expected savings after using UBI can be greater.

Table 1: Summary statistics of UBI adoption

	Total	**Non-UBI**	**UBI**
Number of customers	135,540	95,013	40,527
UBI acceptance rate	0.30		
Fraction male	0.53	0.53	052
Average age	45.8	48.7	39.3
Average monthly premium ($)	109.1	107.6	112.4

In addition, we observe the daily driving score that all UBI customers enrolled in this policy receive at the end of each day as long as they are using the telematics device and don't drop out. This score represents daily driving performance by aggregating the measures of all factors that are considered to be important by the insurance company (mileage, hard braking, time of driving, etc.).

Figure 1: UBI policy timeline for adoption and dropout decisions

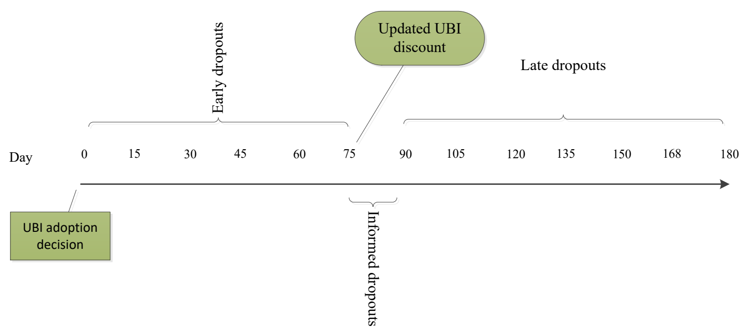


Figure 1 shows the timeline for the UBI policy. As we discussed above, the maximum monitoring time is 6 months (180 days), but the UBI customers can drop out before that. We label the UBI customers who drop out before 75 days of monitoring as "early dropouts." These customers receive the initial UBI discount for the period of using the telematics device, and after dropping out they will not receive the UBI permanent discount. The "informed dropout" UBI customers are those who drop out between days 75 and 90, just after being informed of their updated UBI discount based on their actual driving behavior in the first 75 days of monitoring. The third group of UBI customers are "late dropouts" who terminate the monitoring program later than the "informed dropouts" group after getting more feedback in UBI but don't keep the UBI policy for the whole 6 months. The last two groups, despite dropping out early (before 6 months), receive a permanent discount (adjusted by the time they remained in the UBI program) that applies to their automobile insurance premiums. Finally, the "loyal" UBI customers are those who keep the telematics device for the whole 6 months. They are monitored for 180 days and receive the permanent UBI discount based on their actual driving behavior during the 6 months.

Table 2 compares the four groups of UBI customers that we defined on a number of variables of interest.

Table 2: Summary statistics of “Loyal” and “Dropout” UBI customers

	Early dropouts (<75 days)	Informed dropouts (75-89 days)	Late dropouts (90-179 days)	Loyal
Fraction of enrollees	0.041	0.149	0.172	0.638
Average age at adoption	42.54	41.17	39.15	38.43
Fraction male	0.52	0.53	0.51	0.51
Average UBI score	62.47	65.09	66.33	66.89
Average updated UBI discount	0	0.074	0.081	0.092
Average permanent discount	0	0.075	0.113	0.164

The fraction of enrollees in Table 2 shows the proportion of UBI customers in each group. For example, about 4% of UBI customers drop out before obtaining the updated discount on day 75. About 15% of UBI customers drop out between days 75 and 90, just after getting the updated discount in this period and the opportunity to have a permanent discount; by contrast, around 64% of UBI customers stay in the program for the whole 6 months. The average age of loyal UBI drivers is significantly lower compared to dropouts, showing that the younger drivers tend to stay longer in the program. Average UBI score for loyal UBI customers is significantly higher than for dropouts, which shows that the actual driving performance may be associated with the customers' dropout decision. We also observe that the “Loyal” consumers receive higher updated UBI discounts after being enrolled in the UBI for 75 days. Those “Loyal” consumers receive a much higher permanent discount after the end of the program than other groups of consumers. In general, the summary statistics in Tables 1 and 2 suggest that the customers may systematically make adoption and dropout decisions.[6]

3.2. Privacy issues and consumer adoption and termination

As discussed above, the UBI program is based on the continuous monitoring of drivers for up to 6 months, so privacy concern may be a prominent factor for customers in choosing to adopt or to keep this optional policy. This concern may be affected by any significant changes in customer perception about privacy concerns during this time. In this paper we discuss two events that could affect the privacy perception of consumers in relation to using the UBI policy. As mentioned before, the main goal of this paper is to study the effect of the company's privacy policy enhancement as an internal shock on cost parameters, UBI adoption, and dropout rate. In

[6] For a reduced-form analysis of the adoption and dropout decisions and customers' characteristics using a logit model, please see the Online Appendix (Table A.1 and A.2).

addition, we consider one of the largest data breaches to occur in the US as an external incident to control for its effect on consumer behavior.

Enhancement in privacy policy. UBI works by collecting and analyzing real-time location-based information. Insurers' access to location-based information can be a privacy concern for customers at different levels. How the collected GPS data of customers are being stored and managed by the insurance company may affect the customers' decision to adopt and use the UBI policy. As we mentioned before, our datasets include all the customers who submitted a request for a quote from March 2012 to November 2014. In June 2013, the insurance company changed its data-monitoring terms for the UBI policy and made a public announcement stating that the customer location data would not be stored in the company's servers for any future usage, and the vehicles' location information by GPS is being used only at the real time for calculation of speed, hard brakes, and trip summary.

It's important to note that the change in the privacy policy that limits the usage and retention of location data gathered in the UBI program was common across the industry; almost all the other major insurance companies had been using the same privacy policy and made a similar change around that time. This change in privacy policy is noteworthy because there was a good deal of conversation and debate at the time about insurance companies' access to the location data of customers and how they might use it in the future.[7] This announcement made it clear that the company would not use the location data of its customers in any way beyond the immediate, real-time calculations of the UBI score. Our focal company simultaneously instructed its agents to let all current UBI customers and potential new customers know about this privacy policy enhancement. In this paper we study the effect of this change in privacy policy on UBI adoption and dropout rate. In this section we first show the data summary of UBI adoption and dropout rates before and after this privacy policy change.

The privacy policy change was announced in June 2013, so we first compare the adoption and dropout rates in the two months of April and May 2013 (before the event) with June and July

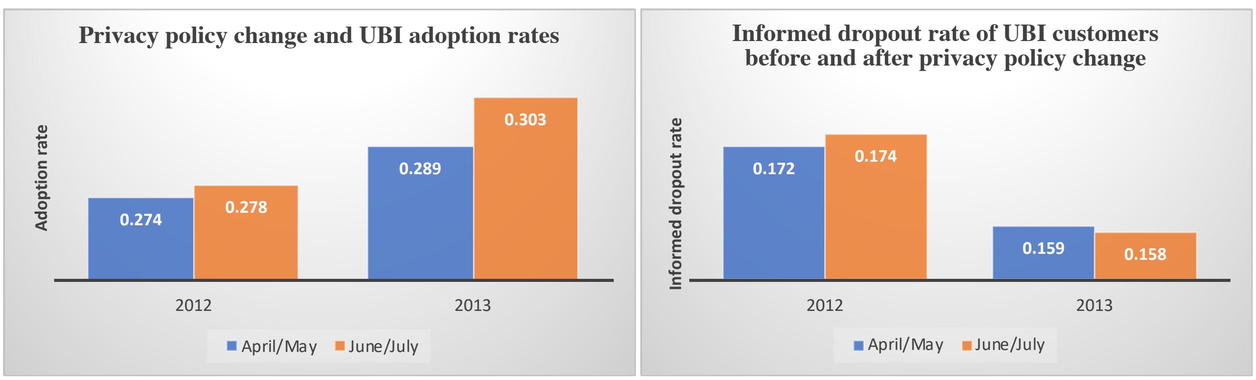
---

[7] "LexisNexis 2013 Insurance Telematics Study" http://solutions.lexisnexis.com/forms/IP13TelPIIP2013Research11757?source=RSpr&utm_campaign=telematics&utm_source=RS-pr

2013 (after the event). In addition, we plot the adoption and informed[8] dropout rates of customers in the same months of 2012 to control for the seasonal effects.

Figure 2 shows that there is a significant increase in the adoption rate of customers in June and July 2013 after the company announced the new privacy policy terms, compared to the two months before that. However, there is no significant increase in the adoption rate of customers in June and July 2012. For "informed dropout" rates, the plots show no significant changes. In the Online Appendix, we use a reduced-form logit model to test for the effect of the privacy policy change on UBI adoption and the "informed dropout" rate. The results suggest that after controlling for time trend, seasonal effects, and other demographics, the adoption rate after the privacy policy change is significantly higher than before (Table A.3 and A.6).

Figure 2: Comparison of UBI adoption and informed dropouts before and after the privacy policy change

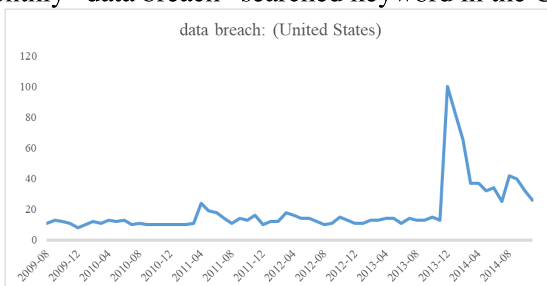


Data breach. In the US, a major data breach was announced on December 15, 2013, namely a breach of credit and debit card data at discount retailer Target that affected as many as 40 million shoppers who went to the stores in the three weeks after Thanksgiving in 2013. Target Corporation announced this event on December 15, and immediately afterward the news was widely reported. It's clear from the monthly data in Figure 3 that there is a spike in the "data breach" searched keyword in Google in December 2013. No other data breach of this magnitude was announced before or during the time period of our study. Although the data breach happened in an unrelated industry and there is no direct or immediate monetary loss to the auto insurance consumers, the event generated a heightened concern that extended beyond those directly affected.

[8] We consider only the "informed dropout" rate in our analysis, because in the period of two weeks after 75 days most of the observed dropouts occur.

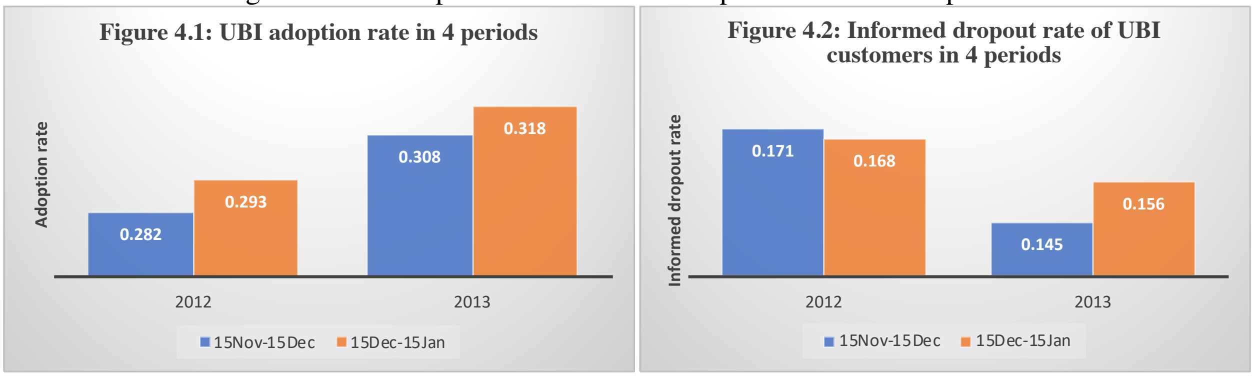
While the data-breach event pertains most directly to data cybersecurity, privacy and security are closely related. Privacy includes any rights that individuals have to control their personal information and how it is used. Security, on the other hand, refers to how their personal information is protected. Thus any threat to the security of personal information threatens a person's control over his or her personal data and hence privacy (Conger and Landry 2008).

Figure 3: Monthly "data breach" searched keyword in the US on Google



To provide a model-free analysis of the possible effects of the data breach, we compare the UBI adoption rates and the "informed dropout" rates of UBI customers in four periods. We consider the Periods 1 (November 15, 2012, up to December 15, 2012) and 2 (December 15, 2012, to January 15, 2013) as the benchmarks to account for possible seasonality effects. Our key focus is the change from Period 3 (November 15, 2013, up to December 15, 2013) before the data breach was made public to Period 4 (December 15, 2013, to January 15, 2014) after announcement of the data breach. To compare the UBI adoption rate of insurance policyholders, we compute the percentage of new customers who adopted the UBI policy in each period. Figure 4.1 shows the UBI adoption rate of customers who submitted an insurance quote request in each period. The difference in adoption rate of Period 4 (just after the data breach was reported publicly) compared to Period 3 is not significant, using a reduced-form logit model, as discussed in the Online Appendix (Table A.4).

Figure 4: UBI adoption and "informed dropout" rate across 4 periods



Next, we look at the "informed dropout" rate of UBI customers who adopted the UBI policy before the data breach and who can make their "informed dropout" decision after 75 days of using UBI in each of four periods shown in Figure 4.2. For example, the customer who adopted the UBI policy October 5, 2013, and kept the policy for 75 days should make her "informed dropout" decision December 20, 2013, so her decision is considered to be in Period 4. Figure 4.2 shows the "informed dropout" rate in each period. This rate in Period 4 (after the data breach) is higher than in Period 3 (before the data breach), while the dropout rate in Period 2 is lower than in Period 1, which we use to control for seasonal effects. So, the data-breach event is correlated with an increase in the "informed dropout" rate. The results of reduced-form logit models presented in the Online Appendix (Tables A.5 and A.7) show that the "informed dropout" rate is significantly higher in Period 4 after controlling for the demographics, drivers' performance, time trend, and seasonal effects.

In the next section, we model the customers' decisions in the UBI program to better understand the trade-off between the benefits of using UBI and the cost of it. Developing a dynamic structural model can help us identify the cost of using UBI for different groups of customers and the changes in the perceived privacy cost of using UBI after the privacy policy enhancement (internal shock) and data-breach (external shock) events.

4. Model Setup

In order to answer the research questions presented in section 1, we first develop a baseline dynamic structural model to capture the consumers' trade-offs between the cost of being monitored when using UBI and their long-term saving on their insurance premium. In addition, to study the effect of privacy perception on customers' decisions, we consider a privacy policy enhancement and a major data-breach incident during the time period of our study. By including the effects of these two changes on cost parameters of our dynamic structural model, we extend our baseline model to capture possible changes in the customers' perceived cost of adopting UBI and of being monitored. The privacy policy enhancement and data breach both might change the perception of privacy from the customer's perspective, and we can test this possibility by modeling the customers' decision before and after these events to determine any significant effects on their behavior.

4.1.Baseline dynamic structural model.

In this part, we propose a dynamic approach to model the customer's decision process for UBI program adoption and termination. To receive a permanent price discount on her insurance premium, a consumer has to bear a short-term cost in using the UBI device; hence a consumer has to make a trade-off between a long-term benefit and a short-term cost. We believe that a dynamic setting can effectively capture the customers' forward-looking decision process in UBI program adoption and usage.

Although we have data on a daily basis, we aggregate our data into 15-day periods (semi-monthly) to limit the impact of random variation on a daily basis and for computational efficiency. We chose 15 days as one decision period, as 15 days is sufficiently long to achieve these goals and because 15 is an integer divisor of the 75 days when customers first receive a revised discount and can elect to withdraw from the program and still gain a permanent discount.

Since we observe the consumers' decision only within the insurance company, and virtually all customers (more than 97%, as discussed more fully below) remain with the company for at least one year, we focus on their choices between the UBI and the regular insurance without discount. If a consumer adopts UBI, we are interested in how long they stay in the program, with the maximum length of stay being 6 months. In our model, a consumer's decision process is defined as follows. At time $t = 0$, a consumer decides whether to adopt UBI or not. If a consumer does not choose UBI, then she will pay the full premium and has no further decisions to make afterwards. If a consumer adopts UBI, we assume that she makes decisions every 15 days after adopting the UBI policy. Specifically, at the end of each period, UBI customers observe their latest-period driving performance and decide whether to keep or drop out of the UBI policy. Figure

5 describes the timeline of the decision process, where we consider a finite time-horizon model to our problem since a customer will be monitored for 12 periods at most. We also note a few critical time points during the 180 days. In particular, at day 0, a consumer makes an adoption decision at the beginning of the decision process. After 75 days of monitoring (after adoption), right after a consumer makes the 5th keep or drop-out decision (d_5), she obtains an updated discount based on her 75 days' driving performance. As discussed in the data section, about 15% of UBI customers drop out at decision d_6. The final decision a consumer has to make is d_{12} before the maximum monitoring period is reached.

After 180 days of UBI usage, monitoring ends and the customers are required to return the device to the company, with no more monitoring after that. The customers will get permanent discounts for their insurance premium from the company. So, we have 13 decision points in this setting, including the UBI adoption decision in the initial period (labeled as d_0) and the keeping or dropping out of the UBI policy in the following periods (1 adoption and 12 dropout decision points). We formalize a consumer's decision as follows. For the initial decision point,

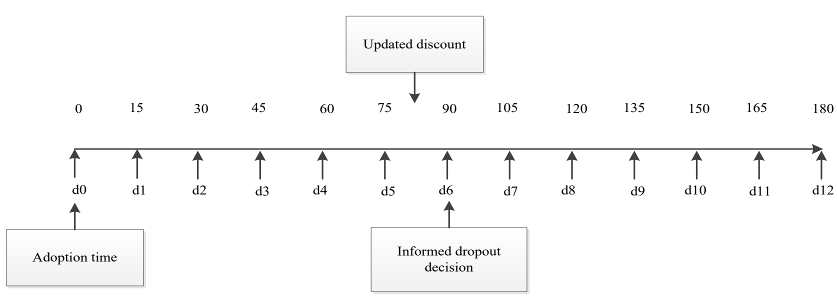
$$d_{i0} = \begin{cases} 1 & \textit{Adopt UBI} \\ 0 & \textit{Not adopt UBI} \end{cases}$$

If a consumer adopts UBI at t = 0, then her decision for the subsequent periods is

$$d_{it} = \begin{cases} 1 & \textit{Keep UBI} \\ 0 & \textit{Drop out at the end of period } t \end{cases} \qquad t = 1,2,3, \dots, T = 12$$

So d_{i0} represents the adoption decision of customer i, and $d_{i1}, \dots, d_{i12}$ represent the customer decisions to keep or terminate the UBI service. Once a customer drops out, she cannot return to being monitored.

Figure 5: Timeline of decision process in the model



To clarify the decision process in this context, we illustrate in Figure 6 the process of decision making for the customers at day 0 (d_{i0}) and day 15 (d_{i1}). At time t = 0, the customers decide whether they want to choose the UBI policy or the traditional auto insurance policy. As for the information available at this decision point, the customers know their demographic information, the initial premium to pay, and their past driving history, which helps them better predict their ability to get the benefit of a UBI policy, as well as the one-time cost of switching to the UBI policy and the cost of using UBI as a new and innovative technology to adopt.

After adoption of the UBI policy at t = 0, the customers plug in the telematics device and their driving behavior is monitored. So, in the first 15 days (day 1 to 15) the customers observe their actual driving performance (UBI score), and at the end of this time period they decide whether they want to keep the UBI policy or drop out and switch to the traditional policy from the company. Before making the decision at the end of the first period (d_1), the customers know their UBI score at period 1, the cost of using UBI, and the premium to pay.

The per-period utility functions at subsequent time periods will be different, and it depends on each specific period t.

$$U_{i0} = -\alpha.P_{i1}.\left(1 - discount_{i0}(d_{i0})\right) - (C_0 + C_1).I(d_{i0} = 1) + \varepsilon_{i1} \qquad (1)$$

$$U_{it} = -\alpha.P_{i1}.\left(1 - discount_{it}(d_{it}, S_{it})\right) - C_1.I(d_{it} = 1) + \varepsilon_{it} \quad t = 1,2,3,\ldots,11$$

where $\mathrm{P_{i1}}$ is the initial (first year) semi − monthly premium set for the customer i.

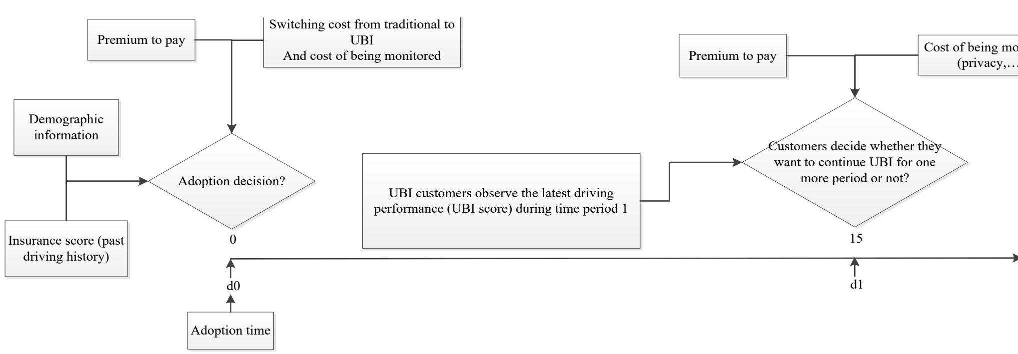
$\mathrm{discount_{it}}$: The UBI discount rate for customer i at period t.

C_0: The initial fixed cost for customers to adopt the UBI policy.

C_1: The semimonthly cost of UBI usage.

α: Price sensitivity coefficient.

Figure 6: Decision-making process at t = 0,1



The immediate utility at the first period (U_{i0}) depends on customer i's decision at the starting point (t = 0). If the customer adopts the UBI policy, there is a one-time adoption cost (C_0) to this new technology that does not depend on the length of UBI usage time and the semi-monthly cost (C_1) of being monitored by the telematics device. We should note that both cost parameters in the model may include the privacy cost. A market research company (Pinnacle)[9] studied the important factors that act as barriers to adopting and using UBI. For auto insurance customers who did not adopt UBI, the research company reported Privacy, Rate can go up, and Concerned about results[10] as major negative feelings among non-UBI customers that lead to not adopting UBI. Among UBI customers who used the device, Hard-brakes beeping, Not fair evaluation (UBI score), and Installation were reported as the major complaints. When a consumer decides whether to adopt UBI or not, she needs to consider whether she is willing to share personal, private information with the insurance company. After she has adopted UBI, she needs to consider whether she is willing to share additional private information about her behavior during the upcoming period. In general, the customer's immediate utility at each time period t (U_{it}) depends on the dropout decision at the beginning of that period (d_{it}) and the effective expected UBI discount that the UBI customer can receive in the current time period based on the driving behavior and the customer's decision. In the baseline model, although we consider the cost parameters (C_0, C_1) to be constant over time, we extend our dynamic structural model later by adding the time trends, the implementation of a privacy policy enhancement, exposure to the data-breach event, and observed heterogeneities in cost parameters.

Outside option. As we discussed before, in this paper we structurally model the customers' decisions in the first 6 months of UBI usage, including the customers' adoption decision. Since we observe only the subset of customers who purchased a policy from this specific insurer, we don't consider outside options in our model for the adoption stage. In other words, we assume that the customers decide to buy UBI or non-UBI policies given that they want to purchase a policy from this insurer. In addition, since the customers have a 1-year contract with the company, we assume there is no outside option within the first year. So, even after adoption of the UBI policy, there is no option in our model for the customers to switch to another insurer within the first 6

[9] Presented at Insurance Telematics USA 2015 Conference.
[10] That is, the respondent is concerned about how the company uses the results of monitoring later on.

months of UBI usage, and the customers simply decide between keeping the UBI policy or dropping out and switching to the company's traditional insurance policy. Our data show that 97.2% of initial customers who commit to an annual policy stay with the company for at least that first year.[11] However, at the time of annual renewal of the contract, the consumers are more likely to switch to other companies, which is considered in our model.

Time horizon and termination value. Since the consumers obtain a permanent discount after 75 days of UBI monitoring, we also need to consider their long-term savings after they are no longer being monitored. So, it's important to know the time periods (the number of years) that the customers would expect any potential long-term benefit from UBI. Ideally, we should structurally model a consumer's renewal decision to estimate the potential saving over the long term. However, since we don't observe the consumers' choices outside the company, it is impossible to build a full-fledged structural model by considering both annual renewal and UBI adoption and usage decisions. For our empirical application, we consider a 5-year time horizon for our model as the maximum time that the customers expect to stay with this insurer.[12] In addition, we model (in reduced form) and estimate the retention rate of customers for the first 5 years based on observed decisions of customers in the first two renewals.

In other words, in choosing whether to adopt and stay in the UBI program, a consumer considers a probability distribution for churning for up to 5 years. We calculate the expected present value of auto insurance cost for the remaining time after the first 6 months of UBI enrollment. We know that the insurance premium may vary over time for the policyholders beyond the benefit of the UBI program in the traditional form by having, for example, accident-free driving performance. So, we consider P_{i2}, P_{i3}, P_{i4}, and P_{i5} as the (semi-monthly) premium for customer i to pay at year 2, 3, 4, and 5, respectively. The permanent discount is the UBI discount

[11] The company offers a 1-year auto insurance policy and sets the rate for that year. The customer can opt out at any time, but there is usually a (sometimes implicit) charge for doing so. As our data indicate, very few customers do so. One major cause of switching is moving from one state to another, in which case a new auto insurance policy is needed. Our results are consistent with the annual state-to-state migration rate in the US, which is between 2% and 3%. https://www.census.gov/data/tables/time-series/demo/geographic-mobility/state-to-state-migration.html

[12] We also consider a 10-year horizon for our model and find similar estimation results for the impact of privacy policy enhancement and data breach on consumers' adoption and usage costs of UBI. The estimation results are reported in Online Appendix Table A13.

that each UBI customer can receive after using the UBI policy for at least 75 days, and this discount will be effective as long as the customer continues with this company.

We can calculate the residual utility or the present value of the total insurance cost of each customer (given the termination year) without considering any benefit and discount of the UBI policy. If we have $\mathrm{P}_{\mathrm{i1}}, \mathrm{P}_{\mathrm{i2}}, \mathrm{P}_{\mathrm{i3}}, \mathrm{P}_{\mathrm{i4}}$, and P_{i5} as the basic semi-monthly premiums that customer i should pay for the first 5 years without considering the UBI discount, then we have:

$$(residual\ value_i\ |termination_i = \mathrm{T}) = -\alpha * \left(\sum_{\varphi=0}^{11} \beta^{\varphi} * \mathrm{P}_{\mathrm{i1}} + \sum_{\tau=2}^{\mathrm{T}} \sum_{\varphi=(24*\tau)-36}^{(24*\tau)-13} \beta^{\varphi} * P_{i\tau}\right). \quad (2)$$

$$\beta\ is\ semi-monthly\ discount\ factor$$

$\sum_{\varphi=0}^{11} \beta^{\varphi} * \mathrm{P}_{\mathrm{i1}}$ is the present value of the premium a consumer has to pay in the remaining 6 months after the end of the UBI program; and $\sum_{\varphi=(24*\tau)-36}^{(24*\tau)-13} \beta^{\varphi} * P_{i\tau}$ is the present value of the premium customer i would pay in the entire year ($\tau = 2,3,4,5$).

As discussed above, we estimate the retention of insurance customers at the renewals and consider the estimated distribution in finding the expected residuals. As described more fully in the Online Appendix, we employ the shifted beta geometric model proposed by Fader and Hardie (2007) to estimate the retention behavior of customers. Estimating the retention probabilities helps us to estimate the unconditional expected residual values based on the definitions. In other words, $residual\ value_i$ shows the cumulative expected utility for customer i without considering the UBI discount and based on the customer's renewal expectation. We use this residual utility in calculating the valuation functions. For customers who receive a UBI discount, we adjust the residual values accordingly.

In the dynamic setting, the customer decides what to do based on the valuation function at each decision point. We can write the valuation function as a function of "state" and the customer's decision at each period. We start by writing the valuation function for the last decision point (T = 12).

$$\bar{V}_{i12}(d_{i12} = 1, S_{i12}) = E[residual\ value_i * (1 - E[permanent\ discount_i(d_{i12} = 1, S_{i12})])] \quad (3)$$

$$\bar{V}_{i12}(d_{i12} = 0, S_{i12}) = E[residual\ value_i * (1 - E[permanent\ discount_i(d_{i12} = 0, S_{i12})])]$$

where S_{i12} is the state level related to decision point 12 (last decision), and $permanent\ discount_i$ is the UBI permanent discount that customer i receives.

So, we can readily find the expected valuation functions at the last decision point for all the customers if we know the permanent discount function (we describe the permanent discount function below) and the state level for each customer. For all other decision points, the valuation function can be represented as follows:

$$V_{it}(d_{it}, S_{it}) = \begin{cases} \sum_{j=t}^{11} \beta^{j-t} * E[U_{ij} | d_{it} = 0, S_{it}] + \beta^{12-t} * residual\ value_i * (1 - E[permanent\ discount_i(d_{it} = 0, S_{it})]) & d_{it} = 0 \\ U_i + \beta . \bar{V}_{i(t+1)}(d_{i(t+1)}) & d_{it} = 1 \\ for\ t = 0,1, \dots ,11 & \end{cases}$$

and

$$\bar{V}_{i(t+1)}(d_{i(t+1)}) = E_{S_{i(t+1)}}\left[\max_{d_{i(t+1)}} V_{i(t+1)}(d_{i(t+1)}, S_{i(t+1)})\right]. \qquad (4)$$

Our dataset includes all the customers who had a one-time chance to enroll in a UBI policy by getting a 5% promotional discount at the time. In other words, we have a dynamic stopping problem in which there is no opportunity for the non-UBI (or UBI) customer to go back to UBI usage later if she decides to not adopt (drop out at any time period).

Discount function. We know from the company's UBI policy that UBI customers receive a 5% initial discount just for signing up, and this remains effective for the first 75 days of monitoring if they keep this policy. If they drop out before 75 days, there is no further UBI discount and the customers no longer get the initial benefit of the UBI policy. So, we have:

$$discount_{it} = \begin{cases} permanent\ discount_i = 0 & d_{it} = 0 \\ 0.05 & d_{it} = 1 \end{cases} \quad for\ t = 0,1,2, \dots ,5.$$

After 75 days and in time period 6, the UBI customers will receive the updated discount based on their driving performance. At the end of this period (d_{i6}), the customers can decide whether they want to keep the UBI policy or drop out. In both cases, they will receive the updated discount for the next period.

$$discount_{i6} = \begin{cases} permanent\ discount_i = f_{0,6}(UBIScore_{i6}) + \varepsilon_{i6} & d_{i6} = 0 \\ f_{1,6}(UBIScore_{i6}) + \varepsilon_{i6} & d_{i6} = 1 \end{cases}$$

After period 6, the customers can keep the updated UBI discount they received on day 75 if they want to continue using the UBI policy; however, if they decide to drop out, they will get an adjusted permanent UBI discount.

$$discount_{it} = \begin{cases} permanent\ discount_i = f_{0,t}(UBIScore_{it}) + \varepsilon_{it} & d_{it} = 0 \\ discount_{i6} & d_{it} = 1 \end{cases} \quad for\ t = 7,8, \dots ,11$$

$$permanent\ discount_i = \begin{cases} f_{0,12}(UBIScore_{i12}) + \varepsilon_{i12} & d_{i12} = 0 \\ f_{1,12}(UBIScore_{i12}) + \varepsilon_{i12} & d_{i12} = 1 \end{cases}$$

We don't know the exact formula used by the company for $f_{0,t}$ and $f_{1,t}$, but we can estimate these functions empirically based on the observed discount and customers' UBI score in our dataset. We assume the discount function defines a one-to-one mapping between UBI score and UBI discount. Considering the UBI score as the state variable in the state space, we use the expected value of discount in our model as the true value, which is not observed for some state levels in our sample.

State space and transition probabilities. After setting the valuation and discount functions, we now specify the state space. Figure 6 helps us to define the state variables. At the adoption time (d_{i0}), the UBI customers observe the demographic information, insurance score, and the premium. So, we define:

$$S_{i0} = (Premium_i, Demographic_i, Insurance\ score_i).$$

A consumer uses the observed information S_{i0} to predict her UBI score and the discount she can obtain based on her driving behavior.

After the initial period, the UBI customers observe additional information about their actual driving performance (UBI score) and then make the decision to drop out or not. So,

$$S_{it} = (Premium_i, Demographic_i, UBI\ score_{it}, t) \quad t = 1,2, \dots ,6.$$

Note that after UBI adoption, a consumer's insurance score is irrelevant after controlling for the traditional premium.

After being informed of the updated discount at period 6 and before the end of the 6 months, there is no new updated discount based on current UBI score if the customers decide to continue using UBI, and the UBI discount will be the last updated discount they receive after just 75 days ($f_{1,6}(UBI\ score_{i6})$). So, we should include the $UBI\ score_{i6}$ in addition to the most current UBI score in state space for t = 7,8,…, T = 12.

$$S_{it} = (Premium_i, Demographic_i, UBI\ score_{it}, UBI\ score_{i6}, t) \quad t = 7,8, \dots , T = 12$$

We assume a Markov process for the transition of UBI score.[13] We consider the UBI score of customer i at period t ($UBI\ score_{it}$) as a random variable that follows the Markov process. We assume this process has a fixed transition matrix M_S[14] and can be estimated empirically by using the UBI panel data.

$$M_S = pr\left(UBI\ score_t = S^j \middle| UBI\ score_{t-1} = S^i\right) \quad t = 2,3,\dots,12 \qquad (5)$$

We also need one more transition probability—the distribution of $S_{i1}|S_{i0}$, which is the mapping for a consumer's observed states (including the insurance score, premium, and demographics) to her UBI score:

$$M_0 = pr\left(UBI\ score_1 = S^j \middle| S_{i0}\right) \qquad (6)$$

The estimation procedure will be discussed in the next section.

This model setup enables us to capture the cost of adopting and using a UBI policy. We describe in the next section how we estimate the functions and parameters of the dynamic structural model.

4.2. **Extended dynamic structural model.**

In this subsection we introduce an extended version of the baseline model that can capture the time trends, the effects of observed heterogeneities, the privacy policy change, and the data-breach event on the cost parameters. To this end, we assume the cost parameters are time-variant, and we identify the changes in the two cost parameters because of the privacy policy change of the company and the data-breach event discussed earlier, after controlling for general time trends. In summary, we assume that the general time trends, privacy policy change in June 2013, and data-breach event in December 2013 are the three factors that make the cost parameters time-variant. In addition, we consider later the observed heterogeneity in the cost parameters across age groups and gender.

In the extended dynamic structural model, we add one more dimension to our state space in the model we proposed before. We now assume in addition to the state variables of UBI score, demographic information, and premium that the customers observe and consider the time ($month_i$) in which they make their adoption decisions. In our dataset we have 33 months of

---

[13] The UBI score is the only stochastic component in the state space that can change over time.

[14] We discretize the UBI score variable to form the discrete transition probabilities for the UBI score. We explain in detail the dimension of this matrix in the next section.

customers' data, and we assume that the adoption and dropout patterns can change over this time (month = 0,1,..., 32) due to customers' cost changes. The extended structural model specification that considers time-variant cost parameters is:

$$U_{i0} = -\alpha . P_{i1} . \left(1 - discount_{i0}(d_{i0})\right) - (C_{0i} + C_{1it}) . I(d_{i0} = 1) + \varepsilon_{i1} \tag{7}$$

$$U_{it} = -\alpha . P_{i1} . \left(1 - discount_{it}(d_{it}, S_{it})\right) - C_{1it} . I(d_{it} = 1) + \varepsilon_{it} \quad t = 1,2,3, \dots ,11$$

$$C_{0i} = C_0 + polynomial(month_i, 3) + \tau_0 * privacy\ policy_{i0} + \delta_0 * data\ breach_{i0}$$

$$month_i = 0,1,2, \dots , 32$$

$$C_{1it} = C_1 + polynomial(month_i, 3) + \tau_1 * privacy\ policy_{it} + \delta_1 * data\ breach_{it}$$

$$\boldsymbol{privacy\ policy_{it}} = \begin{cases} 1 & if\ customer\ i\ made\ decision\ in\ period\ t\ after\ privacy\ policy\ change \\ 0 & else \end{cases}$$

$$\boldsymbol{data\ breach_{it}} = \begin{cases} 1 & if\ customer\ i\ made\ period\ t\ decision\ in\ the\ period\ of\ two\ months\ after\ data\ breach \\ 0 & else \end{cases}$$

$$t = 0,1, \dots , 12$$

Note that we assume that the impact of the firm's policy change is long term and that the privacy policy enhancement affects all customers after its implementation. However, we assume that the data breach only has temporary effects on customers' insurance choices, as the media coverage of the Target data breach decreased over time. Since we assume the adoption and dropout may change over time as discussed above, we need to consider time in our state transition space as well. Therefore, we consider time trends, privacy policy change, and data-breach occurrence in the transition from S_{i0} to S_{i1}. We discuss this issue in more detail in the next section. It's also important to note our assumption that the customers cannot foresee the insurance company's policy change and the data-breach event. In other words, if the customers don't observe the privacy policy change or data breach until time t, they don't expect these events in the future when calculating the valuation function at time t. For example, if a customer makes the UBI adoption decision before the privacy policy/data-breach event, she doesn't expect these events to occur in the future.

In addition, we examine the observed heterogeneity of the effect of privacy policy and data breach on the cost parameters across different age and gender groups by extending our model. We allow for the interactions of customers' age and gender with policy change and data breach to explore how customers respond differently to these events. Specifically, we write our model as follows:

$$C_{0i} = C_0 + polynomial(month_i, 3) + \tau_0 * Privacy\ Policy_{i0} + \delta_0 * data\ breach_{i0} + \varphi_0 * Gender_i + \omega_0 * Agegroup_i + \tau_0' * Privacy\ Policy_{i0} * Gender_i + \delta_0' * data\ breach_{i0} * Gender_i + \tau_0'' * Privacy\ Policy_{i0} * Agegroup_i + \delta_0'' * data\ breach_{i0} * Agegroup_i \qquad (8)$$

$$month_i = 0,1,2, \ldots, 32$$

$$\boldsymbol{Gender_i} = \begin{cases} 1 & female \\ 0 & male \end{cases}$$

$$\boldsymbol{Agegroup_i} = \begin{cases} 1 & customer\ i\ is > 35 \\ 0 & else \end{cases}$$ [15]

$$C_{1it} = C_1 + polynomial(month_i, 3) + \tau_1 * Privacy\ Policy_{it} + \delta_1 * data\ breach_{it} + \varphi_1 * Gender_i + \omega_1 * Agegroup_i + \tau_1' * Privacy\ Policy_{it} * Gender_i + \delta_1' * data\ breach_{it} * Gender_i + \tau_1'' * Privacy\ Policy_{it} * Agegroup_i + \delta_1'' * data\ breach_{it} * Agegroup_i$$

In summary, we model consumers' UBI adoption and usage decision within the first 6 months after they made a quote request from the focal company.

5. Estimation and Empirical Results

In this section we first estimate the transition probabilities of the state variables, UBI discount functions, churn probabilities, and semi-monthly expected premiums at renewals (P_{i2}, P_{i3}, P_{i4}, and P_{i5}) based on our dataset. In the second part, we use the estimated parameters of the first step to estimate the parameters of the dynamic structural model.

Transition probabilities' estimation. As explained in the previous section, the UBI score ($UBI\ score_{it}$) is the only stochastic time-variant state variable that the customers have uncertainty about in their forward-looking behavior. We assume that they have a rational expectation concerning the transition of their UBI scores over time. Following the Markov property, we need to estimate the distribution of $(UBI\ score_{it})|(UBI\ score_{it-1})$ $t = 2,3, \ldots 12$ and of the initial

[15] We also consider three age groups including customers below 35 years old, customers between 35 and 55 years old, and customers above 55 years old. However, we don't find statistically significant differences in the cost estimates between the latter two groups. Hence, we only report the estimation results for the two distinctive groups: customers below 35 and customers equal to and older than 35.

UBI score conditional on a consumer's insurance score and other observed characteristics $(UBIScore_{i1})|S_{i0}$. The estimation of these distributions helps us define the expected driving performance (UBI score) at the next period given the current state level in order to solve the valuation functions in our dynamic setting.

First, we discretize the UBI score levels in our state space, which helps us to use the full solution methods for estimating the parameters of the dynamic structural model. The UBI score potentially can be between 0 and 100, but in our dataset the minimum average semi-monthly UBI score is 32.4 and the maximum is 99.1. So, we consider 68 integer levels for the UBI scores in the range of [32,99]. Each observed UBI score is approximate to the nearest integer between 32 and 99. Next, we estimate the distribution of $(UBI\ score_{it})|(UBI\ score_{it-1})$ $t = 2,3, \ldots 12$ for each age and gender group based on the observed UBI score for these UBI customers.[16] Considering the average semi-monthly UBI score improvement pattern (increasing exponential decay form), we try both exponential decay function and power function to capture this general pattern of our data. We found that the power function outperforms the exponential decay function. Specifically, we consider the model (9) to estimate the conditional distribution of transition $f(UBI\ score_{it}|UBI\ score_{it-1}, Gender, Agegroup\ )$,

by using the following log-log regression for each demographic group:

$$\begin{aligned}\log(UBI\ score_{it}) - \log(UBI\ score_{it-1})\\ = \mu_0 + \delta * (\log(t) - \log(t-1)) + \varepsilon_{it}\ , \varepsilon_{it} \sim N(0, {\sigma_1}^2). \qquad (9)\end{aligned}$$

This model can capture the increasing exponential decay pattern of the UBI score in our dataset. The estimated coefficients of model (9) for each age and gender group are reported in the Online Appendix (Table A.8). From model (9), we are able to predict the transition probability of UBI from t-1 to t for each age and gender group based on the estimated parameters of this model.

In addition, we need to estimate the initial distribution of $(UBI\ score_{i1})|S_{i0}$, which means the distribution of the first-period UBI score given the observed state variables before adopting the UBI policy. Model (10) assumes that all customers can predict their first-period UBI score based on a few observed state variables (insurance score, age, gender, and state of residence), but

[16] Since we observe only the loyal UBI customers' UBI scores for the entire 12 periods, here we assume that all customers' UBI scores follow the same transition process for identification purposes.

estimating model (10) based on just our observed first-period UBI scores may lead to a selection issue because we only observe the UBI scores (outcomes) for the customers who adopt the UBI policy. It is possible that UBI adopters are systematically different from non-adopters in their driving performance. This selection issue can result in biased estimation in the coefficients of our regression model.

$$UBI\ score_{i1} = \alpha_0 + \alpha_1 * insurance\ score_i + \alpha_2 * age_i + \alpha_3 * male_i + state_i + \varepsilon_i \quad \varepsilon_i \sim N(0, \sigma^2) \quad (10)$$

We use the Heckman approach to correct for possible selection bias in the estimated coefficients of model (10). We follow the two-step procedure, which is the most common method for estimating the Heckman model (Wooldridge 2010). In this approach, we first estimate the probit selection equation

$$P_i(S_{i0}) = probit\big(S_{i0} = (age_i, gender_i, premium_i, State_i, \text{Instrumental variables}_i)\big)$$

by MLE to obtain the estimates of UBI adoption probabilities. Then we estimate the coefficients of model (10) by OLS of $UBI\ Score_{i1}$ on covariates and $\widehat{\boldsymbol{P}}_{\boldsymbol{i}}$. In our specification we define two instrument variables (that do not appear in the second-stage equation) as follows:

$$\text{Instrumental variables}_i = (\#of agent_i\ , agent\ UBI\ performance_i)$$

$$\#of agent_i\text{: } the\ number\ of\ company's\ agents\ in\ customer\ i's\ city$$

$$agent\ UBI\ performance_i\text{: } the\ percentage\ of\ sold\ insurance\ policies\ with\ UBI\ for\ the\ agent\ of\ customer\ i$$

These two instruments are appropriate in our setup, because they are associated with the adoption decision of customer i, but they don't directly affect the UBI performance of customer i in the first time period. The details of this approach and the results can be found in the Online Appendix (Tables A.9 and A.10).

Discount functions estimation. As discussed earlier, the insurance company did not disclose its exact UBI discount functions to us. In this part, we estimate all the UBI discount functions defined in section 4, Model Setup. Considering the driving performance of customers in each period (UBI score), we observe the actual UBI discount rate that the customers receive based on their decisions at different time periods. So, we can specify a simple regression model to estimate the UBI discount functions. The updated discount at period 6 $\left(\hat{f}_{1,6}(UBI\ score_{i6})\ \right)$ and the discount for the loyal customers ($\hat{f}_{1,12}(UBI\ score_{i12})$) are estimated by a simple linear regression specification, as shown in equation (11).

$$discount_{i6} = \hat{f}_{1,6}(UBI\ score_{i6}) = \alpha_6 + \beta_6 * UBI\ score_{i6} + \varepsilon_{i6} \qquad (11)$$

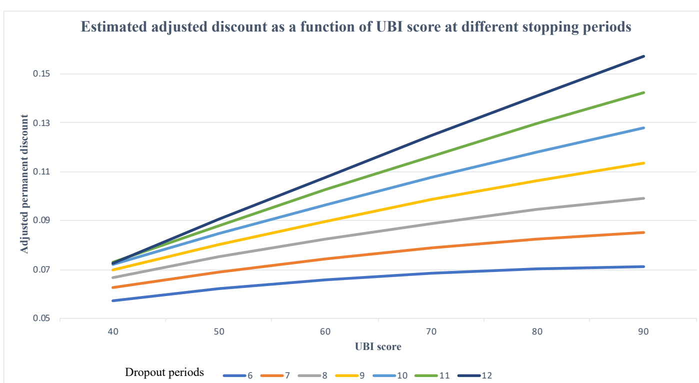
$$permanent\ discount_{i12} = \hat{f}_{1,12}(UBI\ score_{i12}) = \alpha_{12} + \beta_{12} * UBI\ score_{i12} + \varepsilon_{i12}$$

For the cases in which the UBI customers drop out before 6 months of UBI usage, the discount functions $f_{0,t}\ for\ t = 6,7, \ldots ,12$ can be estimated based on the discount received at the dropout time. We consider a functional form for the adjusted discount function as below.

$$f_{i0,t} = (\beta_0 + \alpha_0 * (t-6) + \rho_0 * (t-6)^2) + (\beta_1 + \alpha_1 * (t-6) + \rho_1 * (t-6)^2) * UBI\ score_{it} + (\beta_2 + \alpha_2 * (t-6) + \rho_2 * (t-6)^2) * UBI\ score_{it}^{\ 2} + \varepsilon_{it} \qquad (12)$$

Figure 7 shows the estimated plots for adjusted permanent discount at different stopping points based on the model (12) specification. The estimated plots show that the customers who drop out later receive a higher adjusted permanent discount than those who drop out earlier; and more interestingly, the effect of UBI score on the adjusted discount is greater when the UBI customers drop out at later periods. In other words, the customers can obtain a greater benefit (higher permanent discount) at any level of UBI score if they drop out later. The estimation results of model (12) can be found in the Online Appendix (Table A.11).

Figure 7: Estimated adjusted permanent discount functions



Churn probabilities and future premiums. In our dynamic setting, as we discussed earlier, all customers have expectations about the number of years they will stay with this insurer and their future premiums in the following years. These expectations help the customers to find their expected future valuations for all decision alternatives. First, we estimate the shifted beta geometric model proposed by Fader and Hardie (2007), as discussed in section 4. The estimated model, as shown in the Online Appendix, gives us the estimated churn probabilities for the customers at different times after the first year.

For future premiums, based on our dataset, we run a simple linear regression over all the company's customers to model future semi-monthly premiums ($\mathrm{P_{i2}}, \mathrm{P_{i3}}, \mathrm{P_{i4}}$, and $\mathrm{P_{i5}}$) considering the initial premium and the customers' age, gender, and state of residence.

$$\mathrm{P_{it}} = (\omega_{00} + \omega_{01} * (t-2)) + (\omega_{10} + \omega_{11} * (t-2)) * \mathrm{P_{i1}} + \alpha_{1t} * age_i + \alpha_{2t} * male_i + state_i + \varepsilon_i \qquad t = 2,3 \quad (13)$$

We use the estimated model (13) to find the expected value of annual premiums for years 2, 3, 4, and 5 for each customer. The estimation results of model (13) can be found in the Online Appendix (Table A.12).

Dynamic structural model parameters' estimation. In this section we discuss our estimation method for the single-agent dynamic structural model developed in section 4 and show the results of our baseline model based on the selected approach. In addition, we estimate the structural parameters of a set of extended models and discuss the results.

The low dimensionality of our state space, the UBI score discretization explained before, and the number of decision points (13) being reasonable, allow us to use the full solution method. With this method, we can solve the valuation functions for all possible state variables at each decision point by backward induction and maximize the likelihood function based on the expected valuation functions to estimate the parameters of the structural model. More specifically, in the backward induction, we start from the end and solve the valuation functions for all possible values of the state variables and decisions. For example, we start from T = 12 (the last decision point) and find the expected value function. We have $V_{i12}(d_{i12}, S_{i12})$ for both $d_{i12} = 0,1$ based on equation (3) in section 4. So, if we know the observed state variables at period 12 (S_{i12}), we can find the expected valuation function for each customer at the last decision point. Given the solved valuation functions at T = 12 and the equations (10) and (11) in section 4, we can find the $V_{i11}(d_{i11}, S_{i11})$ for both, keeping UBI and dropout decisions at this decision point. We repeat the same procedure to solve the valuation functions for all t = 0,1,…, 12.

If the error terms in model (1) follow the extreme value distribution, the probability of customer i's decision on the UBI policy will be as below.

$$P\big(d_{it} = 1 \big| S_{it}, \theta = (\alpha, C_0, C_1, \beta, \sigma^2)\big) = \frac{\exp\left(V_{it}(d_{it} = 1, S_{it})\right)}{\exp\left(V_{it}(d_{it} = 0, S_{it})\right) + \exp\left(V_{it}(d_{it} = 1, S_{it})\right)} \quad (14)$$

In this way, we assume that the customers are forward-looking in making their decision to keep the UBI policy or drop out.

Based on the observed state variables and the customers' decisions about staying in or dropping out of the UBI monitoring program, we can form the likelihood function and maximize it to estimate the parameters of the structural model $\theta = (\delta, C_0, C_1, \beta, \sigma^2)$. For identification purposes we need to fix β (semi-monthly present-value discount factor). Table 4 shows the estimated coefficients of the structural model in equation (6). We set $\beta = 0.995$, which will be equal to a 12% annual discount rate. To estimate the baseline model, we also assume the same state space transition distributions specified in equations (9) and (10) for all customers in each age and gender group.

Table 4: Estimated coefficients of structural model

5-year time horizon considering the retention decisions		
Parameters	Fixed	Estimate
$\boldsymbol{\beta}$	0.995	
$\boldsymbol{\alpha}$		0.48 (0.07)**
$\boldsymbol{C_0}$		72.59 (0.79)**
$\boldsymbol{C_1}$		8.72 (0.46)**
Number of customers	135,540	

(‘): *p*-value < 0.1, (*): *p*-value < 0.05, (**): *p*-value < 0.01, log-likelihood = -156,893

As shown in Table 4, all three estimated coefficients are significant at the 0.01 level. The difference between $\boldsymbol{C_0}$ and $\boldsymbol{C_1}$ indicates that the customers have a much higher initial cost for using UBI compared to the per-period cost of being monitored. There are many concerns from a customer's perspective that may be included in $\boldsymbol{C_0}$, such as the switching cost from traditional insurance to UBI, trust in the company's argument that there is no penalty, general concern about being monitored, and concerns related to privacy issues. On the other hand, $\boldsymbol{C_1}$ is the per-period cost of using UBI after adopting the policy. So, costs are more related to the customer's privacy concerns, ongoing inconvenience of using the telematics device, and the possibly annoying feedback system in which the UBI driver, for example, hears a beep every time she hard-brakes. We should note that our setting imposes $\boldsymbol{C_0}$ as a one-time cost while $\boldsymbol{C_1}$ is the cost of using UBI in each period. In terms of dollar value, we can interpret that the customers trade off the benefits of using UBI with the costs of $\frac{72.59}{0.48} = \$151.22$ initially and $\frac{8.72}{0.48} = \$18.16$ semi-monthly to choose the UBI policy and keep it. The high estimated initial cost ($\boldsymbol{C_0}$) of adopting UBI can help explain the relatively low adoption rate (30%), and $\boldsymbol{C_1}$ can explain the customers' dropout decision after adopting UBI.

Extended model results. In this part we estimate the parameters of the extended model introduced in section 4. In this way, we can capture the effect of privacy policy change and data breach after controlling for possible time trends in $\boldsymbol{C_0}$ and $\boldsymbol{C_1}$. We again use the full solution method to estimate the structural parameters. It's important to note that since we assume that time trends, privacy policy, and data breach can change the adoption decision of customers, we need to modify our initial state transition in the Heckman model, as explained in the Online Appendix.

If the error terms in equation (7) follow the extreme value distribution, the probability of customer i's decision about the UBI policy will be as below.

$$P\left(d_{it} = 1 \middle| S_{it}, \theta = (\alpha, C_0, \tau_0, \delta_0, C_1, \tau_1, \delta_1, \beta, \sigma^2)\right)$$
$$= \frac{\exp\left(V_{it}(d_{it} = 1, S_{it})\right)}{\exp\left(V_{it}(d_{it} = 0, S_{it})\right) + \exp\left(V_{it}(d_{it} = 1, S_{it})\right)} \qquad (15)$$

As before, we set $\beta = 0.995$. Table 5 shows the estimation results for the extended model. For simplicity, we omitted the polynomial coefficients of time trends in the estimation table.

Table 5: Estimated coefficients of extended structural model

5-year time horizon considering the retention decisions			
Parameters	Description	Fixed	Estimate
$\boldsymbol{\beta}$	Discount factor	0.995	
	****Time trend effect included****		
$\boldsymbol{\alpha}$	Price sensitivity		0.50 (0.06) **
$\boldsymbol{C_0}$	Cost of adoption at the beginning		92.32 (1.13) **
$\boldsymbol{C_1}$	Per-period cost of monitoring at the beginning		10.78 (0.65)**
$\boldsymbol{\tau_0}$	Effect of privacy policy on adoption cost		-4.39 (0.96)**
$\boldsymbol{\tau_1}$	Effect of privacy policy on cost of monitoring		0.14 (0.17)
$\boldsymbol{\delta_0}$	Effect of data breach on adoption cost		2.21 (1.32)'
$\boldsymbol{\delta_1}$	Effect of data breach on cost of monitoring		0.81 (0.26)**

No. of customers: 135,540, log-likelihood: -149,582
('): *p*-value < 0.1, (*): *p*-value < 0.05, (**): *p*-value < 0.01

Considering our specification of the extended model, the estimates of $\mathbf{C_0}$ and $\mathbf{C_1}$ in Table 5 show the UBI initial cost of adoption and per-period cost of monitoring, respectively, in the first month of our data (March 2012). The estimate of τ_0 shows that the improved privacy policy significantly reduces the UBI adoption cost, while the privacy policy improvement does not significantly reduce the monitoring cost. On the other hand, the data-breach event marginally (p-value < 0.1) increases the cost of adoption, but it significantly affects the monitoring cost of UBI usage, which means the existing UBI customers at the time of the data breach are more likely to drop out from the UBI policy in the two months after the data breach.[17] These estimates are consistent with our model-free results in the data section.

Extended model results with heterogeneity. As we discussed in section 4, considering the heterogeneities across age and gender groups helps us capture the differences in the cost parameters between males and females as well as the two age groups. Table 6 shows the estimation results of the structural parameters including the existing heterogeneity across age and gender groups. The results suggest that the UBI adoption cost for females is significantly higher than that for males. For the per-period cost of being monitored in UBI, females and older drivers both have a significantly higher cost of being monitored compared to males and younger drivers.

For the effect of the privacy policy enhancement on the adoption cost, the estimation results show that the males and younger drivers' adoption costs decrease more than that of females and older drivers after the new privacy policy is introduced in June 2013. In other words, the males and younger drivers seem to be more sensitive to the privacy policy enhancement compared to others.

By contrast, the data-breach event primarily affects the monitoring cost in the UBI program, and this effect is heterogeneous across different groups. The results in Table 6 show that females are more sensitive to the data breach, and their cost of being monitored increases significantly more compared to males during the two months immediately following the data-breach event. Older drivers' cost of being monitored also increases marginally more than that of younger ones due to the data breach.

Table 6: Estimated coefficients of extended structural model considering age and gender heterogeneities

5-year time horizon considering the retention decisions		
Description	Parameters	Estimate
Discount factor	$\boldsymbol{\beta}$	Fixed = 0.995
Price sensitivity	α	0.51**

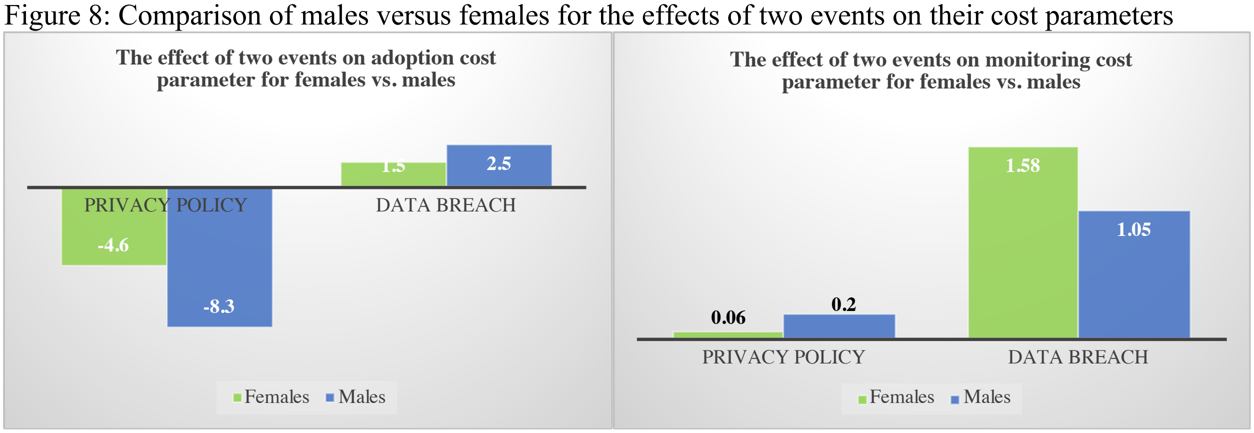
[17] As a robustness check, we change the time horizon to 10 years and estimate the model again. Table A.13 in the Online Appendix shows the results.

Base[18] cost of adoption at the beginning	$\boldsymbol{C_0}$	**87.32****
Cost of adoption (females – males)	$\boldsymbol{\varphi_0}$	**7.84***
Cost of adoption (older drivers – younger)	$\boldsymbol{\omega_0}$	2.29’
Base monitoring at the beginning	$\boldsymbol{C_1}$	8.38**
Monitoring cost (females – males)	$\boldsymbol{\varphi_1}$	0.94*
Monitoring cost (older – younger)	$\boldsymbol{\omega_1}$	1.83**
Base effect of privacy policy on adoption cost	τ_0	-9.72**
Effect of privacy policy on adoption cost (females – males)	τ_0'	3.81*
Effect of privacy policy on adoption cost (older – younger)	τ_0''	2.69*
Base effect of privacy policy on monitoring cost	τ_1	0.25
Effect of privacy policy on monitoring cost (females – males)	τ_1'	-0.14
Effect of privacy policy on monitoring cost (older – younger)	τ_1''	-0.09
Base effect of data breach on adoption cost	δ_0	2.10
Effect of data breach on adoption cost (females – males)	δ_0'	-1.04
Effect of data breach on adoption cost (older – younger)	δ_0''	0.97
Base effect of data breach on monitoring cost	δ_1	0.58’
Effect of data breach on monitoring cost (females – males)	δ_1'	0.53*
Effect of data breach on monitoring cost (older – younger)	δ_1''	0.95’
Number of customers	135,540	

(‘): *p*-value < 0.1, (*): *p*-value < 0.05, (**): *p*-value < 0.01, log-likelihood: -146,838

Figure 8 shows a clear comparison between males and females for the effects of the privacy policy enhancement and data breach on the cost parameters. Note that the numbers in this figure indicate the changes in the cost parameters due to the two events. As we can see, it’s clear that males are more sensitive to the privacy policy enhancement by the greater reduction on average (-8.3) of the adoption cost compared to females (-4.6). On the other hand, females show more sensitivity to the data-breach event according to the greater semi-monthly monitoring cost compared to males (1.58 versus 1.05).

Figure 8: Comparison of males versus females for the effects of two events on their cost parameters



[18] Base means the cases where $Gender_i = 0\ and\ Age\ group_i = 0$.

Combining the findings for age and gender across the two events of the privacy policy enhancement and the data breach, we find it interesting that females are more responsive to the increased perceived risk of losing privacy, whereas both males and younger customers are more responsive to the decreased risk associated with the gain in privacy protection.

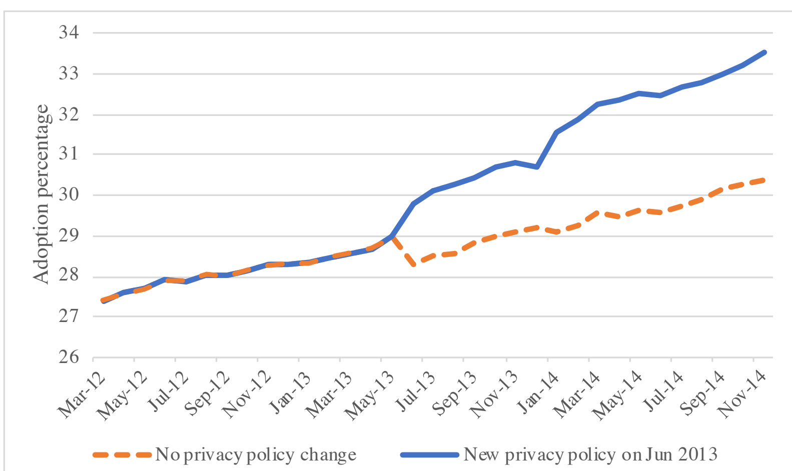
6. Counterfactual Analysis

A company can change its privacy controls, security, and usage of the data it collects from customers, but it cannot control the occurrence of outside events such as a data breach in another company. Therefore, in this section, we focus on a counterfactual analysis to evaluate the effect of changes in the privacy policy of the insurance company on UBI usage patterns, which is one of the main research questions of the paper. As we discussed in the data section, the insurance company enhanced its privacy policy in June 2013 to limit the access to and storage of customers' location data, so the customers may perceive the cost of adoption and usage of a UBI policy as being lower after the policy change. In our dynamic structural model, we estimated the value of the cost parameters before and after June 2013. The estimates of the structural parameters in Table 5 indicate that on average there is more than 5% reduction in the UBI adoption cost of customers after June 2013, but no significant change in the monitoring cost. In the counterfactual analysis, we consider the scenario that the insurance company doesn't change the privacy policy in June 2013 as compared to the current setting. The results provide valuable insights on how changing the data privacy policy of the insurance company could make a difference in the customers' adoption and usage of the UBI program.

More precisely, equation (7) in section 4 shows our specification for the cost parameters in the dynamic structural model. Based on the estimated value of the parameters, we can find the estimated cost parameters for each customer ($\hat{C}_{0i}, \hat{C}_{1it}$). The results in Table 5 show that the privacy policy change in June 2013 on average reduces the estimated adoption cost of customers ($\hat{C}_{0i}$) significantly by 4.39 and increases the cost of monitoring ($\hat{C}_{1it}$) by 0.14, which is not statistically significant. Considering these two estimated coefficients in equation (7) (τ_0, τ_1), we estimate the counterfactual cost parameters for each customer if the privacy policy isn't changed ($Privacy\ Policy_{it} = 0$ for all customers at all times). We simulate the adoption and dropout decisions of each customer in our dataset considering the state variables' information and estimated cost parameters in the two scenarios (current setup versus "not changing the privacy policy").

Figure 10 shows the adoption rates of customers in each month from March 2012 to November 2014 in the two scenarios discussed above. The blue (solid) line shows the monthly simulated adoption rates in the current setting, where the company improved the data privacy policy in June 2013. The orange (dashed) line, on the other hand, shows the counterfactual adoption rates of customers if the company didn't improve the data privacy policy. As we see in Figure 10, the adoption rates before June 2013 are exactly the same because the change in privacy policy occurred in June 2013. But after June 2013, as we expected, the orange dashed line is lower than the blue line, which shows that the adoption rate in the UBI program after June 2013 could be significantly lower if the company had not enhanced the data privacy policy to limit the storage and usage of customers' location data. At the end of our observation period, 18 months after the improvement in the privacy policy, our counterfactual analysis suggests that the adoption rate is 33.55% as compared to 30.39% if the privacy policy enhancement had not occurred. In other words, the customers' UBI adoption rate is estimated to have increased by 10% after the insurance company enhanced its privacy policy.

Figure 9: Comparing the adoption rates of customers in two scenarios by counterfactual analysis



In addition to the comparison of adoption rates reported in Figure 9, Table 7 provides a comparison of the two scenarios (after June 2013) based on other variables of interest in the counterfactual analysis of model (8), where we considered the heterogeneous cost parameters across age and gender groups. As indicated in Table 7, the average UBI adoption rate from June 2013 to November 2014 with the new privacy policy (31.9%) is significantly higher than the average UBI adoption rate without the privacy policy change (29.1%). The results also show that the proportions of young drivers and of males among the UBI adopters increase relative to older drivers and females after the privacy policy enhancement. It's interesting to note that the average initial insurance score and UBI score of UBI customers are both lower when the new privacy policy is in effect compared to the "no privacy policy change" scenario. That is, customers with lower driving scores and less to gain from UBI are more likely to enroll because the costs (in terms of privacy loss) are now lower. Consequently, and perhaps counterintuitively, the average permanent price discount is lower.

Table 7: Counterfactual analysis data summary

(June 2013 - November 2014)	No privacy change	New privacy policy	% change
UBI adoption rate	0.291	0.319	+9.6% **
Dropout rate among UBI adopters	0.329	0.334	+1.5% ‘
Young drivers UBI adoption rate	0.429	0.47	+9.5% **
Old drivers UBI adoption rate	0.238	0.253	+6.3% *
Males UBI adoption rate	0.287	0.316	+10.1% **
Females UBI adoption rate	0.296	0.322	+8.7% **
Average initial insurance score of adopters	61.25	59.47	-3% *
Average initial UBI score	64.57	63.10	-2.3% *
Average permanent UBI discount	12.9%	12.6%	-2.4% *

(‘): p-value < 0.1, (*): p-value < 0.05, (**): p-value < 0.01

We also performed a counterfactual analysis to see whether the policy change had an impact on the dropout rate among drivers who enrolled in the program. As indicated in Table 7, there is a marginally significant difference in the dropout rate, with an average dropout rate of 33.4% with the enhanced policy as compared to 32.9% if this improvement had not been made. This is likely because most of the effect of the privacy change occurs at the time of enrollment.

Overall, the results show that customers are sensitive to changes in the cost parameters. Companies can lower customers' privacy cost by setting privacy rules that limit the use and retention of personal information, resulting in higher adoption rates and more UBI-monitored customers at the end of the 6-month monitoring period, while also adding those with lower driving performance.

7. Discussion

Usage-based auto insurance (UBI) was introduced to help insurers improve their profits by better targeting pricing (premiums) to the actual driving behavior of their customers, to attract customers from other insurers that did not (yet) offer UBI, and to increase customer retention. In this paper we develop empirical models to better understand the adoption and retention of customers in the UBI program and the trade-off between savings in premium and cost of using UBI, including privacy considerations. The latter element was studied to determine the importance of privacy concerns overall and for different groups of customers. A particular concern of ours was to see whether an internal enhancement of the data privacy policy and a data breach outside the company could affect the customers' adoption and retention behavior. We used a unique dataset from a major US insurance company to address our research questions.

Our setting is an appropriate one in which to examine these effects, as the customer always has the option of obtaining the same product benefits—auto insurance—whether or not she is enrolled in the UBI program. This is unlike many other situations, such as Google Maps, where without yielding private information about your location and destination, you are unable to receive guidance. Thus, we have a non-digital, long-established environment to test the effects of privacy.

The model-free and reduced-form models' results in section 3 and in the Online Appendix suggest that the customers respond to positive and negative perceived changes in the privacy of their data. In addition, the estimated parameters of the baseline dynamic structural model developed in this paper indicate the crucial role of both initial and semi-monthly costs on the customers' adoption and dropout decisions.

Importantly, the quasi-experiment resulting from a major privacy policy enhancement by the insurance company that greatly limits the storage and usage of customers' location data helps us identify the effect of changing privacy perception on UBI usage. The results show that the initial cost of adopting UBI with the new data privacy policy is significantly lower compared to before, thus leading to a higher UBI adoption rate. However, the new privacy policy doesn't change the semi-monthly monitoring cost in UBI significantly and, hence, the dropout rates are statistically the same. In a further model allowing for heterogeneity, we find that males' adoption cost on average is more sensitive to enhancing the privacy policy (decreases to a greater extent) than that of females.

Our results indicate that in an actual field setting where consumers have a clear choice as to whether or not to share private information, consumers consider the economic benefits and the cost of sharing their private information. Moreover, males and females appear to differ in how they trade off these costs and benefits. Our study also shows that privacy concerns in the use of one company's products can be influenced both by internal policy changes and by external events occurring outside that company.

Implications. Big data and GPS tracking have provided the basis for a revolutionary new set of products and services that better target individual-level customer needs. The benefits of innovative products and services that rely on using vast amounts of private and sensor data manifest when customers adopt these services and share their private data with the companies, so the benefits must be offset against the cost of being monitored and sharing data. It's critical for firms offering these types of services to better understand the obstacles to adoption on the part of customers and their willingness to share their private data.

UBI insurance is a prominent example of a new technology providing personalized products and services based on a consumer's own usage experiences. However, consumers have to share their personal information in order to get the personalization benefits. The results of our study offer firms and policymakers better insight into the privacy issues associated with adoption of a new technology that relies on using customers' private data. These issues are not limited to a particular industry, as various types of companies are trying to design and implement business models involving new, improved, or better-targeted services based on using the customers' private data. For example, Amazon and Google are developing "smart home" devices and services using the collected sensor data of customers to optimize efficiency and product quality. In another recent example, John Hancock (a life insurance company) announced plans to transform the life insurance industry into a "wellness" model by using wearable devices to collect personal metrics involved in fitness and lifestyle of their customers.[19] Therefore, our results—which quantify and help explain the impact of privacy concerns and other costs of adoption and usage of new products—are relevant in a broader realm and not limited to the user-based auto insurance industry.

In addition to new-product adoption, our results can help firms better understand how to retain customers when they adopt new technologies that rely on their private data. While most

[19] https://money.cnn.com/2018/09/19/news/companies/john-hancock-life-insurance-vitality/index.html

studies of new technology focus on adoption costs, our study shows that there are significant costs to the consumer of remaining in the program and continuing to share private data. Encouraging customers to keep using the new products and share their private data for a longer period of time is crucial because it allows the firms to collect more information about actual behavior from each customer and thus continuously improve firm efforts to better target products and marketing programs. Considering both adoption and monitoring costs can help companies carry out CRM and monitoring strategies more efficiently, decrease the per-period cost of being monitored, and ultimately increase the firm's profit by reducing the dropout rate. Keeping costs down also encourages customers to stay longer in the program.

As IoT and collecting the private sensor-based data and location of customers are extensively being implemented in many industries, discussions have intensified among policymakers and companies about limiting the access to and usage of this sensitive information. Recently, to protect consumers, tech giants including Apple and Google instructed app developers to remove trackers; otherwise these developers will lose access to their operating systems.[20] One of the important topics is the response of customers if new restrictive rules apply for using and sharing the private data to better protect the customers' privacy. Our counterfactual results suggest that corporate programs designed to decrease privacy costs may significantly increase the adoption and retention of new technologies that rely on using private data. The results of the present work clearly indicate that companies should roll out their data-driven new products with privacy protection in mind. Numerous studies conclude that transparency about the use and protection of consumers' data reinforces trust and increases new-product adoption. However, our results also suggest that companies are potentially vulnerable to data breaches that are external to the firm.

There are a number of implications for public policymakers. First, the results indicate that beyond attitudinal and behavioral intentions surveys, people's actual consumption behavior is affected by privacy concerns. Regulations and laws that reduce the threat to privacy—by limiting the amount of information that companies can access, by improving the transparency in companies' data access and usage policy, by defining as illegal certain uses of private information, and by providing enforceable privacy protection that restricts the use and sharing of private information—not only help protect individuals but may also boost the adoption rate of new

[20] https://www.wsj.com/articles/apple-and-google-to-stop-x-mode-from-collecting-location-data-from-users-phones-11607549061

technologies. Recent years have seen an increased level of governmental regulation of data collection and use. A challenge for corporate management is to determine to what degree they should welcome such regulation, as it may lead to greater consumer trust without any one company having a competitive advantage. At the extreme, privacy protection may limit the effectiveness of new technology and thereby limit its adoption. Unlike companies, government regulators need to consider the effect not just of an individual sharing his or her own information, but the societal effects resulting from the widespread availability of information.

Limitations and future research. In this paper we have some limitations in our model that could be addressed by future research. First, we impose the fixed transition probabilities for UBI score in our model, as discussed earlier. In other words, we assume all the customers within a demographic group have the same belief and expectation to change (improve) their driving behavior. However, Soleymanian et al. (2019) discuss the mechanism of learning and improvement in driving behavior and find that the economic incentives and negative feedback can change the learning and improvement patterns in driving behavior. So, a future research project could be to extend our current baseline model by considering the effort of customers to change their driving behavior as an additional decision variable and model the learning of customers along with adoption and retention decisions in the dynamic structural setting. Such an extension would be quite challenging but could lead to a more realistic model and estimation of cost parameters.

In addition, in this paper we assume a fixed, maximum time horizon for the dynamic model and a reduced-form model of customer retention (churn). The renewal decision of customers to choose the insurance policy from their existing company or switch to another company can directly affect the expected benefit of customers from UBI, so incorporating the renewal decision of customers in the structural model could be another extension of this work to capture the endogeneity in renewal decisions of customers. However, such an extension would require data on the alternatives considered and choices made by consumers; such data were not available to us and would be difficult to obtain.

Our results can be important for the firm to set more efficient pricing and monitoring strategies to maximize the benefits of UBI for both customers and the company. Previous studies have indicated that adoption of UBI programs is associated with higher profits (Reimers and Shiller 2019), but not the benefits from specific pricing policies. Further corporate data would be required to investigate the profit implications of these policies. Such implications appear to be compelling, at least in the auto insurance industry: in July 2019, GEICO, a long-time holdout, became the last of the top 10 US auto insurers to offer UBI as an option to its customers.

Our analysis also reveals an interesting difference in responses to changes in perceived privacy costs by males and females. While males and females are equally likely to enroll in the UBI program, on average males have both a lower perceived initial cost and semi-monthly cost of being in the program. Perhaps more notably, males were significantly more sensitive than females to the privacy policy enhancement, but females were more sensitive than males to the data breach. Further research using additional data to investigate such differential responses would be warranted.

References

Acquisti, A., Brandimarte, L. and Loewenstein, G. 2015. Privacy and human behavior in the age of information. *Science* 347(6221): 509-514.

Altman, I. 1975. *The Environment and Social Behavior: Privacy, Personal Space, Territory, and Crowding*.

Beales, H. 2010. The value of behavioral targeting. Network Advertising Initiative.

Beresford, A.R., Kübler, D. and Preibusch, S. 2012. Unwillingness to pay for privacy: A field experiment. *Economics Letters,* 117(1): 25-27.

Bleier, A., Goldfarb, A. and Tucker, C. 2020. Consumer privacy and the future of data-based innovation and marketing. *International Journal of Research in Marketing*, 37(3): 466-480.

Bloom, P.N., Milne, G.R. and Adler, R. 1994. Avoiding misuse of new information technologies: Legal and societal considerations. *Journal of Marketing,* 58(1): 98-110.

Chintagunta, P.K. and Nair, H.S. 2011. Structural workshop paper—Discrete-choice models of consumer demand in marketing. *Marketing Science,* 30(6): 977-996.

Choudhary, V., Shunko, M. and Netessine, S. 2020. Does immediate feedback make you not try as hard? A study on automotive telematics. *Manufacturing & Service Operations Management*.

Conger, S. and Landry, B.J. 2008. The intersection of privacy and security.

Fader, P.S. and Hardie, B.G. 2007. How to project customer retention. *Journal of Interactive Marketing,* 21(1): 76-90.

Goldfarb, A. and Tucker, C. 2012. Shifts in privacy concerns. *American Economic Review,* 102(3): 349-353.

Janakiraman, R., Lim, J.H. and Rishika, R. 2018. The effect of a data breach announcement on customer behavior: Evidence from a multichannel retailer. *Journal of Marketing,* 82(2): 85-105.

Jin, Y. and Vasserman, S. 2019. Buying Data from Consumers.

Kummer, M. and Schulte, P. 2019. When private information settles the bill: Money and privacy in Google's market for smartphone applications. *Management Science*, 65(8): 3470-3494.

Lambrecht, A., Seim, K. and Tucker, C. 2011. Stuck in the adoption funnel: The effect of interruptions in the adoption process on usage. *Marketing Science,* 30(2): 355-367.

Laudon, K. 1997. Extensions to the theory of markets and privacy: Mechanics of pricing information.

Lin, T. 2019. Valuing Intrinsic and Instrumental Preferences for Privacy. Working Paper.

Mao, E. and Zhang, J. 2014. Gender differences in the effect of privacy on location-based services use on mobile phones. Paper presented at the 20th Americas Conference on Information Systems on Mini Track: Security/Privacy Issues in Mobile Computing. Savannah, GA.

Martin, K.D., Borah, A. and Palmatier, R.W. 2017. Data privacy: Effects on customer and firm performance. *Journal of Marketing*, 81(1): 36-58.

Miller, A.R. and Tucker, C.E. 2011. Can health care information technology save babies? *Journal of Political Economy,* 119(2): 289-324.

Miller, A.R. and Tucker, C. 2017. Privacy protection, personalized medicine, and genetic testing. *Management Science,* 64(10): 4648-4668.

Norberg, P.A., Horne, D.R. and Horne, D.A. 2007. The privacy paradox: Personal information disclosure intentions versus behaviors. *Journal of Consumer Affairs,* 41(1): 100-126.

Reimers, I. and Shiller, B.R. 2019. The impacts of telematics on competition and consumer behavior in insurance. *The Journal of Law and Economics,* 62(4): 613-632.

Rust, J. 1987. Optimal replacement of GMC bus engines: An empirical model of Harold Zurcher. *Econometrica,* Journal of the Econometric Society: 999-1033.

Ryan, S.P. and Tucker, C. 2012. Heterogeneity and the dynamics of technology adoption. *Quantitative Marketing and Economics,* 10(1): 63-109.

Schoeman, F.D. 1992. *Privacy and Social Freedom*. Cambridge University Press.

Shapiro, C. and Varian, H.R. 1997. US government information policy. University of California, Berkeley, 30.

Shum, M. and Xin, Y. 2020. Time-varying risk aversion? Evidence from near-miss accidents. *The Review of Economics and Statistics*, 1-45

Soleymanian, M., Weinberg, C.B., and Zhu, T. 2019. Sensor data and behavioral tracking: Does usage-based auto insurance benefit drivers? *Marketing Science,* 38(1), pp.21-43.

Sriram, S., Chintagunta, P.K. and Agarwal, M.K. 2010. Investigating consumer purchase behavior in related technology product categories. *Marketing Science,* 29(2): 291-314.

Tsai, J.Y., Egelman, S., Cranor, L. and Acquisti, A. 2011. The effect of online privacy information on purchasing behavior: An experimental study. *Information Systems Research,* 22(2): 254-268.

Wedel, M. and Kannan, P.K. 2016. Marketing analytics for data-rich environments. *Journal of Marketing*, 80(6): 97-121.

Westin, A.F. 2005. Public attitudes toward electronic health records. AHIP Cover, 46(4): 22-25.

Westin, A.F. and Ruebhausen, O.M. 1967. *Privacy and Freedom* (Vol. 1). New York: Atheneum.

Wooldridge, J.M. 2010. *Econometric Analysis of Cross Section and Panel Data*. MIT Press.

Yang, B. and Ching, A.T. 2013. Dynamics of consumer adoption of financial innovation: The case of ATM cards. *Management Science,* 60(4): 903-922.