An Examination of How Latent Costs Shape Consumer Purchase Behavior

by

Michael Palazzolo

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Business Administration)
in the University of Michigan
2016

Doctoral Committee:

Professor Fred M. Feinberg, Chair
Professor Richard Gonzalez
Associate Professor Ryan Kellogg
Professor Peter Lenk
Assistant Professor A. Yeşim Orhun

Dedication

To my mother, Hilda Palazzolo.

Acknowledgements

Thanks first and foremost to my parents and brother, for always being a positive constant in my life. Thanks are also due to Olivier Rubel, who convinced me that starting my PhD at the age of 28 was not as risky as it seemed to me at the time, and worth doing in spite of my reservations; and to Ashwin Aravindakshan and Prasad Naik, who along with Olivier helped me prepare for this journey. I will be forever indebted to Fred Feinberg and Yeşim Orhun; to Fred for his sarcastic guidance and training (and in spite of the cat sounds he makes to distract me from my work), and to Yeşim for her non-sarcastic guidance and training. I also owe a debt of gratitude to my other commitee members Peter Lenk, Richard Gonzalez, and Ryan Kellogg, for providing guidance while being flexible enough to accomodate my ever-changing plans; to the marketing department at the Ross School of Business, which provided invaluable support and feedback (especially when it was needed most); and to the closest group of friends I have had since high school, who made my years in the PhD program the most fun I've ever had—but who will not be named here, because I'm afraid of forgetting someone and hurting their feelings (probably Adi).

Contents

List of Tables

List of Figures

List of Appendices

Abstract

This dissertation contains two essays that explore ways in which latent, short-term "costs" shape consumers' purchase behavior, and consequently influence long-term outcomes. The first essay, "Modeling Consideration Set Substitution," examines consumer search, which involves a trade-off of short-term effort for long-term utility. More specifically, it illustrates the importance of accounting for consideration set substitution when modeling demand in markets where consumers engage in search. Consideration set substitution refers to a consumer considering one alternative at the expense of considering another. For example, if an advertisement for the Ford Fusion induces a consumer to consider the Ford Fusion when he is on the market for a car, this may cause him not to consider another vehicle he otherwise would have, had he not seen the advertisement. The second essay, "Frugality is Hard to Afford," examines whether low income households are less able than higher income households to take advantage of intertemporal savings strategies commonly available in everyday purchase categories due to liquidity constraints. Two strategies are investigated: buying in bulk and accelerating purchase timing to take advantage of sales. Both involve trading off the cost of an increase in short-term expenditure for the reward of long-term savings. Because low income households are more likely to face liquidity constraints, they may be less able to utilize these two strategies. Together, these two essays contribute to our knowledge of how consumers' purchase decisions are influenced by "costs" other than price. Essay one aids our ability to measure consumers' responsiveness to marketing actions by more precisely modeling their decision-making process. Essay two contributes to our knowledge of low income consumers' limited ability to make intertemporal financial trade-offs—even for seemingly low-priced goods—and why they may be less responsive to some marketing actions as a result.

Chapter 1

Introduction

This dissertation contains two essays that explore ways in which latent, short-term "costs" shape consumers' purchase behavior, and consequently influence long-term outcomes.

The first essay, "Modeling Consideration Set Substitution," illustrates the importance of accounting for the limiting nature of (1) search costs and (2) decreasing marginal benefits to search, when modeling demand in product categories where consumers engage in limited search. Specifically, it uses stated consideration set and observed choice data from the automotive industry to explore an under-discussed and oft-overlooked consequence of these two latent "costs": consideration set substitution, which refers to a consumer considering one alternative at the expense of considering another. For example, if an advertisement for the Ford Fusion induces a consumer to consider the Ford Fusion when he is on the market for a car, this may cause him not to consider another vehicle he otherwise would have, had he not seen the advertisement (perhaps because he does not have time to test drive both). Consideration set models that can accommodate markets with many alternatives—the very markets for which modeling consumer search behavior is likely to be most crucial—often place strong restrictions on the degree to which consideration set substitution can occur (e.g., assuming it never or always occurs). Essay one develops a new consideration set formation model that flexibly accounts for the role of search costs and decreasing marginal benefits to search, in turn allowing the data to identify the degree to which consideration set substitution occurs. Further, essay one shows that models that do not account for consideration set substitution will systematically misestimate the impact of marketing actions and market events on consideration and choice probabilities—and, in turn, misestimate substitution patterns between alternatives in the market. The findings of this essay highlight the importance of accounting for consideration set substitution when gauging the effectiveness of marketing actions in industries where consumers engage in limited search.

The second essay, "Frugality is Hard to Afford," examines whether low income households are less able than higher income households to take advantage of long-term savings strategies commonly available in everyday purchase categories due to liquidity constraints. Each purchase a low income household makes forces it to sacrifice liquidity, limiting its financial flexibility in the (near) future. Consequently, low income households may be less willing to spend money on products they do not anticipate consuming soon. Using data on household purchases from the toilet paper category, this essay investigates households' tendency to use two strategies that offer long-term savings in exchange for an increase in short-term expenditure: buying in bulk and accelerating purchase timing to take advantage of sales. This essay first documents that compared to higher income households, low income households are less likely to utilize these strategies, even though they are more financially incentivized to do so. Moreover, a compounding effect is found. Because low income households buy in bulk less often, they carry lower inventories, and are less able to wait for a sale before purchasing again. Additionally, because low income households are less able to accelerate purchasing timing to buy on sale, they are less able to buy in bulk, as sales make large package sizes temporarily more affordable. Next, this essay provides causal evidence that liquidity constraints inhibit the use of these money-saving strategies by low income households, above and beyond the effects of income group differences in brand preferences, access to store types, access to transportation, storage space, or myopia. Finally, the potential financial losses that lower income households incur due to the role liquidity constraints play in shaping their purchase behavior is quantified. This essay's findings pertain to both the deal-proneness and welfare of low income households, and this essay consequently has both managerial and policy implications.

This dissertation contributes to our knowledge of how consumers purchase decisions are influenced by "costs" other than price. Essay one aids our ability to measure consumers' responsiveness to marketing actions by more precisely modeling their decision-making process. Essay two contributes to our knowledge of low income consumers' limited ability to make intertemporal financial trade-offs—even for seemingly low-priced goods—and why they may be less responsive to some marketing actions as a result.

Chapter 2

Modeling Consideration Set Substitution

2.1 Abstract

Consumers purchasing from a large set of alternatives often evaluate only a subset—a consideration set—in order to balance the expected benefits from search (e.g., finding a high-quality product) with costs (e.g., time). If the marginal expected benefit from search decreases in the number of considered alternatives, marketing actions that encourage consideration of one alternative may discourage consideration of another. This paper develops a model of consideration set formation that can account for this "consideration set substitution." Using stated consideration set and observed purchase data from the automotive industry, we measure the impact of marketing actions (vehicle redesigns) and market events (Toyota vehicle recalls, Tōhoku tsunami) on consideration and purchase. We benchmark our model against one that is commonly used in the literature but that does not account for consideration set substitution. We show the benchmark model misestimates the impact of studied marketing actions and events on market share by as much as 13%. Further, it underestimates the frequency with which a gained consideration is converted to a sale. Lastly, although the benchmark model often appears to "fit" the data well, its failure to account for the role search costs play in consideration set formation causes it to infer critical quantities incorrectly, such as the distribution of consideration set sizes and price elasticities, the latter of which are underestimated by nearly 10%.

2.2 Introduction

Consideration sets can be thought of as an intermediate outcome of a consumer's search process (Roberts and Lattin 1991, Mehta et al. 2003, Honka 2014). In categories with broad competitive landscapes, consumers may not expend the effort necessary to learn about all alternatives. They may instead construct a consideration set of alternatives to search over in an effort to balance expected benefits (e.g., finding a high quality product or low price) with costs (e.g., time needed to collect information or the mental cost of evaluating alternatives).

One consequence of consumers' limited willingness (or ability) to search is that marketing actions that increase consideration of one product may in turn decrease consideration of others. For example, if a consumer sees a commercial advertising the Ford Fusion, this consumer might (a) consider the Fusion in addition to any other vehicles s/he would have considered had the commercial not been seen (increasing consideration set size), or (b) consider the Fusion instead of another vehicle which s/he would have considered (keeping consideration set size constant). We refer to the latter of these possibilities as "consideration set substitution."

A wealth of literature has documented the importance of accounting for variations in which alternatives consumers consider when modeling demand. Failing to do so will lead to biased estimates of demand determinants such as brand valuations (Draganska and Klapper 2010) and price sensitivity (Mehta et al. 2003, Van Nierop et al. 2010). Moreover, optimal marketing strategies generated by models that do and do not account for consideration can differ substantially (Van Nierop et al. 2010). Consideration set substitution is a potentially important element of the consumer choice process, yet consideration set models in the literature often enact implicit assumptions about the frequency with which it occurs—at the extremes, either never (e.g., Goeree 2008, Terui et al. 2011) or always (e.g., Feinberg and Huber's 1996 'quota' model).

This paper develops a new two-stage, consideration and choice model that can flexibly measure consideration set substitution. Importantly for applications, the model admits a closed form solution for which the number of calculations does not increase exponentially in the number of available alternatives, so estimation is not impaired by the curse of dimensionality. This makes it an attractive alternative to models with simplifying assumptions that alleviate the curse of dimensionality but also restrict the degree to which consideration set substitution can (be modeled to) occur. The model is also a useful alternative to more structural models that do not admit a closed form solution.

A primary objective of this paper is to demonstrate the importance of flexibly and accurately accounting for consideration set substitution when modeling consumer demand. To this end, we model consideration and choice in the automotive industry. This industry provides a particularly appropriate empirical setting to examine consideration set substitution, as previous research has shown that automotive consumers engage in fairly limited search, typically considering only a small fraction of the (several hundred) available alternatives (Hauser et al. 2010). Our data consists of 634,539 responses to the New Vehicle Customer

Survey from 2009 through 2011. The NVCS is an industry standard regularly utilized by Ford and other vehicle manufacturers to gauge consumer preferences, measure the effectiveness of past marketing actions, and make decisions about future ones. The dataset contains respondents' stated consideration sets, observed purchases, demographics, and purchase history information.

We use our model to estimate how consumer demand for automobiles changed in response to marketing actions (vehicle redesigns) and market events (the 2009-2010 Toyota vehicle recalls and the 2011 Tōhoku earthquake and tsunami). We compare these estimates to those from a restricted version of the model that artificially constrains consideration set substitution. The parametric restrictions employed reduce the model to one with implicit assumptions about consideration set substitution that precisely mirror those of the model previously used in Van Nierop et al. (2010). The differences between the models' estimates highlight two consequences of not accounting for consideration set substitution. First, the restricted model underestimates the frequency with which a gained consideration leads to a sale, because it underestimates how often a gained consideration kicks a competing alternative out of a consumer's consideration set. Second, the restricted model misestimates the considerations gained or lost due to a marketing action or event. For example, we find that Toyota compact and mid-sized cars ("C" and "CD" vehicle classes) lost 4.9 considerations and 2.3 purchases per 100 consumers in our sample due to the recalls, while Japanese C and CD cars lost 9.4 considerations and 5.1 purchases. The restricted model overestimates these losses by as much as 14%.

Additionally, the restricted model's failure to account for search cost leads to an interesting result—though it accurately estimates the average consideration set size, it misestimates the distribution of set sizes. It consequently misestimates important quantities such as price elasticities, which are heavily dependent upon a consumer's set size. The restricted model underestimated the own-price elasticity of six redesigned vehicles by 9.9%, on average.

2.3 Related Literature

We discuss two streams of research to which the present paper hopes to contribute. The first concerns methodological approaches to modeling consumer search and consideration set formation; the second, the substantive literature on the role of search in the automotive industry.

2.3.1 Consumer search and consideration set formation models

Consumers in many categories do not consider all alternatives, instead choosing from a small subset—a consideration set (Hauser and Wernerfelt 1990, Roberts and Lattin 1991). How consumers construct this set has primarily been modeled in two ways. One stream of literature models consideration sets as the outcome of an optimal search process (Stigler 1961, Weitzman 1979). Empirical papers are typically structural and model the consumer's search process to be either simultaneous (e.g., Mehta et al. 2003, Seiler 2013, Honka 2014, Honka et al. 2014) or sequential (e.g., Hortaçsu and Syverson 2004, Kim et al. 2010, Kim et al. 2014), though some work has addressed which assumption is more appropriate for a given context (de los Santos et al. 2012, Honka and Chintagunta 2014). Consumers in structural search models are typically modeled as though they know a portion of their utility for a product and that the population distribution of the unknown portion (e.g., normal with known mean and variance, as in Kim et al. 2010) is known to each of them. In simultaneous structural search models, consumers are modeled as though they construct a consideration set to maximize the expected utility from their choice from the set, less search costs. In sequential structural search models, consumers are modeled as though they search alternatives one by one (revealing the utility of each) and stop when the expected value of continued search no longer exceeds a reservation utility.

A second stream of literature models the probability of a product being considered without specifying a search process through which this probability is generated. One common approach is to model the probability of observing each possible consideration set (as in Swait's seminal 1984 dissertation). The probability of an alternative being considered is then a summation of the probabilities associated with each possible set containing that alternative. This approach can quickly become infeasible as the number of available alternatives rises, though. A common alternate approach is to model consideration as an alternative-specific construct (e.g., Goeree 2008, Van Nierop et al. 2010, Terui et al. 2011). Consideration of each alternative is modeled with a latent variable and alternatives enter a consideration set if this variable exceeds some level (often set to zero for identification). We refer to these as 'level' models (as per Feinberg and Huber 1996). Models that do not specify a search process have typically not accounted for consideration set substitution, as they (implicitly) assume that consideration of one alternative does not affect consideration probabilities of others.

Both model types can accommodate markets with a large number of alternatives, but each sacrifices something to do so. The (non-structural) consideration set models abstract away from the cost-benefit trade-off associated with search, losing the ability to measure consideration set substitution. A common limitation of structural search models is that the probability of an observed consideration set being optimal typically does not have an analytic solution, and numerical integration is needed to calculate the likelihood function. An exception is the model used by Kim et al. (2010), but that model cannot be estimated using individual-level data.

Our model mostly follows in the tradition of the (non-structural) consideration set formation models, in that it necessitates neither an account of which attribute(s) a consumer is searching for information about nor an explicit model of consumer expectations. However, it does accommodate the possibility that the marginal expected benefit from search may be decreasing, and can thereby measure consideration set substitution.

2.3.2 Search and substitution in the automotive industry

Other research has explored the role that search plays in the automotive industry. For example, Ratchford et al. (2003) find that the availability of internet-based sources of information leads to less search (fewer alternatives are considered), Zettelmeyer et al. (2006) find that online information search helps consumers negotiate lower prices at dealerships, and Singh et al. (2014) find that four information sources (dealer visits, print advertising, dealer websites, and resale websites) serve as complements, rather than substitutes. Moraga-González et al. (2015) show that price elasticity estimates are lower when limited search is accounted for than when full information assumptions are applied. This is in line with past research in other categories finding that full information models produce biased price sensitivity estimates (e.g., Mehta et al. 2003, Goeree 2008, Koulayev 2009, Van Nierop et al. 2010, Draganska and Klapper 2011). Sudhir (2001) builds a structural model of competitive pricing behavior to better estimate how pricing decisions affect market share. Berry et al. (2004) use stated second choice data to improve measurement of substitution patterns in the automotive industry. Here, we intend to complement and enhance this line of work—specifically, to improve measurement of marketing mix effects (as well as the impact of market events not controlled by the manufacturer) by accounting for consideration set substitution.

2.4 Data

We avail of cross-sectional data consisting of 634,539 responses to the New Vehicle Customer Survey from 2009 through 2011. Each quarter, the survey is mailed by Maritz Research, Inc. to consumers who purchased a new (unused) vehicle in the United States during the previous quarter. The purchased vehicle is recorded by dealerships at the time of sale. Buyers of small-share vehicles are oversampled, and each observation is given a weight to ensure that the total proportion of weight assigned to a vehicle matches that vehicle's market share for a given quarter.

Survey respondents are asked to list any vehicles they considered buying other than the one purchased. The vehicle make, model, and class are collected for both the considered and purchased vehicles. 43.4% of respondents in our sample had a consideration set consisting of a single vehicle, with 33.7%, 16.1%, and 6.9% having consideration sets of two, three, and four.

The survey also asks respondents to list all other vehicles currently owned or leased and to identify whether one of those vehicles is being replaced by the purchased vehicle. We use these purchase history variables in our model. We also use a few sets of importance rating and product use questions: those used to identify how important fuel economy and brand loyalty are to a respondent and those that catalogue what a consumer uses their vehicle for (e.g., taking children to school or off-roading). The NVCS also collects demographic information. We use the respondent's age, sex, and household income, which were the demographics most highly correlated with vehicle class preference.

2.4.1 Sampling

Several screening criteria are used to reduce the dataset's size. We focus on compact and mid-sized cars (C and CD car vehicle classes)—by far the two most frequently purchased classes, accounting for 33.7% of new vehicle sales (and 82% of non-premium car sales) between 2009 and 2011. We therefore include only consumers who considered at least one car, as C/CD cars were unlikely to be co-considered with trucks, vans, or utilities (Table 2.1). We also remove erroneous or incomplete survey responses and exclude a few respondents who purchased vehicles with exceptionally small market share. We then sampled 9,000 respondents for estimation. A detailed description of our sampling approach can be found in Web Appendix 2.1.

2.4.2 Summary statistics: vehicle redesigns

Measuring the impact of vehicle redesigns serves as a natural means of demonstrating the importance of accounting for consideration set substitution, as redesigns are strategic marketing actions undertaken by firms. Vehicles are redesigned on a cycle, with smaller "refreshes" (e.g., changes to the lights or seat fabric) occurring annually or bi-annually and major redesigns occurring approximately every five years (an industry benchmark, according to Ford). A major redesign marks the beginning of a new "generation" for a nameplate (e.g., the Toyota Camry). Several vehicles transitioned to a new generation between the years of 2009 and 2011, with the Hyundai Elantra, VW Jetta, Subaru Legacy, Kia Optima, Subaru Outback, and Hyundai Sonata being a few for which the new generation saw particularly large growth in market share. The new generations of these vehicles garnered 95.1% more considerations and 86.1% more purchases per 100 consumers in our data sample (Table 2.2) than the previous generation during the three year period under study.[1]

The new generation of a nameplate often features superior styling as well as quality and technological improvements. Manufacturers make these improvements with the expectation that demand will increase, but the aim is not merely to increase market share—often manufacturers intend to charge a higher price commensurate with greater demand. This can be seen in the data: for the six redesigned vehicles of interest, the new generation saw an average inflation-adjusted increase in price paid of approximately 9.9% (Table 2.2) relative to the previous generation.

2.4.3 Summary statistics: market events

Since marketing actions like redesigns are intended to increase vehicle demand, we also measure the impact of two market events that decreased vehicle demand in order to provide a comprehensive overview of how consideration set substitution shapes substitution patterns.

Toyota recalled vehicles due to safety concerns on three separate occasions during 2009 and 2010, with the final recall conducted in the first quarter of 2010. In the second quarter of 2010, consideration of Toyota C and CD cars dropped sharply, from 28.4 considerations per 100 consumers (Q1 2009 through Q1 2010) to only 23.8 over the next year (a decline of 16.2%). Purchases also dropped by 12.0%, from 15.1 to 13.3 (Table 2.3).

In 2011, the Tōhoku earthquake and tsunami disrupted the ability of Japanese manufacturers to produce vehicles and deliver them to the United States. Though this disruption served as a supply shock, it also affected demand through consideration. From the second quarter of 2010 to the first quarter of 2011, Japanese C and CD cars received 68.4 considerations per 100 consumers. After the tsunami (Q2 through Q4 of 2011), these vehicles received only 61.9 considerations per 100 consumers, a decline of 9.5% (Table 2.3). Honda and Toyota took the brunt of the damage—considerations for their vehicles declined by 16.8%, while considerations for other Japanese manufacturers remained virtually unchanged. Though explaining why consumers considered Japanese vehicles less often post-tsunami is not the objective of this paper, there are reasonable explanations. For example, consumers may have decided not to risk expending effort to search for information about vehicles that they might not be able to purchase.

[1] The number of considerations and purchases a vehicle received per 100 consumers in our data sample are the primary measures of consideration and choice that we report. Both are weighted measures—a respondent with a weight of two will be counted twice. For ease of exposition, we will sometimes simply refer to "considerations" or "purchases," omitting the explicit reference to the scale and data sample.

Figures 2.1 and 2.2 make clear that the drop in considerations and purchases for Toyota C and CD cars post-recall and for Japanese C and CD cars post-tsunami are not merely the product of long-term trends. There are stark drops coinciding with the occurrence of each market event.

2.5 Model Development

In this section we develop a model that accounts for consideration set substitution. Each consumer i ($i = 1, 2, \ldots, I$) receives utility u_{ij} from purchasing alternative j ($j = 1, 2, \ldots, J$), but this utility is not fully known to the consumer. Consumer i will conduct a simultaneous search for information about all alternatives in his or her consideration set, S_i, revealing u_{ij} for $j \in S_i$. From that set, consumer i will then purchase the alternative offering the highest utility.

Our model is agnostic to which attribute(s) a consumer is searching for more information about. Consumers may be searching for information about prices (e.g., Kim et al. 2010, Honka 2014), product quality (e.g., Kim et al. 2010), specific product attributes (e.g., Koulayev 2014), among others. Instead of explicating consumer expectations over attributes, we model the probability that alternative j will be considered by consumer i in a manner that accounts for search costs and allows for the net marginal expected benefit from search to be decreasing.

2.5.1 A simple consideration set model

Consumer i's latent preference for information about alternative j is given by ω_{ij}. ω_{ij} (which we refer to as "consideration propensity") consists of a deterministic (w_{ij}) and stochastic component (ϵ_{ij}, observed by the consumer but not the researcher):[2]

$$\omega_{ij} = w_{ij} + \epsilon_{ij} \tag{2.1}$$

Recall that a consumer will learn his or her utility for an alternative j, u_{ij}, if j is included in S_i. A consumer's relative preference to learn his or her utility for any two alternatives a and b (that is, to reveal u_{ia} or u_{ib}) is represented by the levels of ω_{ia} and ω_{ib}. If $\omega_{ia} > \omega_{ib}$, then consumer i prefers to know u_{ia}, and alternative b will only be included in S_i if alternative

[2]Consumer i is searching to reveal u_{ij}, not ϵ_{ij} (which s/he already observes). w_{ij} does not represent the portion of u_{ij} observable to the consumer, nor does ϵ_{ij} represent the unobservable portion. w_{ij} is a collection of covariates that are predictive of consumer i's propensity to consider alternative j, and ϵ_{ij} represents the portion of this propensity that cannot be explained by the covariates available to the researcher.

a is included. More generally, the fact that consumer i has decided to search only for more information about the alternatives in S_i implies the following inequality:

$$\omega_{ij^+} > \omega_{ij^-} \ \forall \ j^+ \in S_i, \ j^- \notin S_i \tag{2.2}$$

Equation 2.2 does not provide insight as to why consumer i only considers the alternatives in S_i and no others. It only reflects one half of the cost-benefit trade-off underlying search and consideration set formation. However, if searching for information about alternative j (including it in S_i) incurs cost c_j, then the consumer's objective becomes clear: search for information about any alternative j for which $\omega_{ij} > c_j$.

In practice, alternative-specific search costs (c_j) will not be separately identifiable from consideration propensities (ω_{ij}). Instead, researchers commonly model consideration as an alternative-specific construct (e.g., Feinberg and Huber 1996, Van Nierop et al. 2010), where j enters S_i if his or her consideration propensity for that alternative exceeds some level (often set to zero for identification). The probability S_i is optimal is then given by:

$$Pr(S^* = S_i) = \prod_{j^+ \in S_i} Pr(\omega_{ij^+} > 0) \prod_{j^- \notin S_i} Pr(\omega_{ij^-} < 0) \tag{2.3}$$

The "level" model in equation 2.3 has a particularly attractive feature: estimation of the model circumvents the curse of dimensionality. However, a notable limitation is that it cannot account for consideration set substitution. The determination of whether alternative j is worth considering depends only on ω_{ij}. Thus, a change to the deterministic portion of consumer i's consideration propensity for any one alternative will not affect the consideration propensity of any other. If, for example, consumer i sees a commercial advertising alternative k, increasing the deterministic portion of consideration ω_{ik} and (correspondingly) the probability that alternative k is considered, the probability of other alternatives being considered will not change.

2.5.2 Consideration set substitution

If the cost of searching for information about alternative j is c_j, then consideration set substitution can only occur if the net expected benefit from search ($\Lambda_{ij} = \omega_{ij} - c_j$) is decreasing in the number of other alternatives considered. To see why, consider an example: under a hypothetical set of market conditions, consumer i will consider alternatives x and y but not z ($S_i = \{x, y\}$). However, this consumer will consider z if s/he sees a commercial advertising it. If Λ_{ix} is decreasing in the number of alternatives considered, then:

$$\{\omega_{ix} - c_x | S_i = \{x, y, z\}\} < \{\omega_{ix} - c_x | S_i = \{x, y\}\} \tag{2.4}$$

It may be the case, then, that $\omega_{ix} < c_x$ when alternative z is added to the consideration set, even though $\omega_{ix} > c_x$ without z in the set. Alternative x may therefore be removed from the set as a direct consequence of z being added to it.

If Λ_{ij} is decreasing in the number of alternatives considered, then either ω_{ij} is decreasing or c_j is increasing (or both). That the marginal expected benefit from search (ω_{ij}) might be decreasing is fairly intuitive, and this effect has been referenced or modeled in past work (e.g., Roberts and Lattin 1991, Kim et al. 2010, Honka 2014). The more alternatives a consumer considers, the more likely it is that at least one will have a particularly low price, be of very high quality, or otherwise be an especially good fit for the consumer. For that reason, the expected benefit from adding alternative j to a set of size n is likely to be lower than the expected benefit from adding it to a set of size $n-1$. It is perhaps less intuitive why search costs (c_j) might be increasing, but situations where this might be the case are hardly uncommon. For example, choice overload might cause a consumer's psychological costs to increase superlinearly in the length of search or number of alternatives considered.

We cannot separately identify whether the expected benefit from search is decreasing, costs are increasing, or both. We can, however, account for the net effect—that Λ_{ij} may be decreasing in consideration set size. Since alternative-specific search costs are not identified, we instead model the marginal cost of considering an n^{th} alternative:

$$c_{N_{S_i}} = c + \sum_{n=2}^{N_{S_i}} \tilde{c}_n \tag{2.5}$$

$$\tilde{c}_n = c + \psi_n \qquad \psi_n \geq 0 \; \forall \, n \qquad \psi_{n+1} \geq \psi_n 0 \; \forall \, n$$

In equation 2.5, N_{S_i} is the size of consideration set S_i and ψ_n represents the degree to which Λ_{ij} is decreasing in n (ψ_n is consequently non-decreasing in n). c and ψ_1 must be normalized for identification (so $\tilde{c}_n = \psi_n$), but the parameters ψ_n are identified for all $n > 1$. Consumer i will add alternatives to S_i in decreasing order of consideration propensity, ω_{ij}, until no remaining alternative has a consideration propensity exceeding the marginal cost of adding it to the set.

2.5.3 Likelihood function

With the net expected benefit from search decreasing in the number of alternatives considered, the probability that an observed consideration set S_i is optimal can no longer be cleanly expressed as a product of alternative-specific probabilities (as in equation 2.3). Instead, we identify a set of conditions under which S_i is optimal and calculate the probability that these conditions hold. S_i is optimal if consumer i cannot be made better off by perturbing the alternatives in the set. More specifically, consumer i cannot (1) remove an alternative from S_i, (2) add an alternative to S_i, or (3) swap an alternative in S_i with one excluded from S_i and be better off. These three conditions can be formally defined as follows:

No Removal:

$$\omega_{ij^+} > \tilde{c}_{N_{S_i}} \; \forall \; j^+ \in S_i \tag{2.6}$$

No Additions:

$$\omega_{ij^-} < \tilde{c}_{N_{S_i}+1} \; \forall \; j^- \notin S_i \tag{2.7}$$

No Swaps:

$$\omega_{ij^+} > \omega_{ij^-} \; \forall \; j^+ \in S_i, \; j^- \notin S_i \tag{2.8}$$

In other words: (2.6) consumer i's propensity to consider each alternative in S_i is greater than the marginal cost of including an $N_{S_i}^{th}$ alternative in the set, (2.7) consumer i's propensity to consider each alternative excluded from S_i must be lower than the marginal cost of adding it to S_i ($\tilde{c}_{N_{S_i}+1}$), and (2.8) consumer i's propensity to consider each alternative in S_i must exceed his or her propensity to consider each alternative excluded from S_i. The probability that these conditions hold for an observed set S_i is consumer i's likelihood function:

$$\pi(S_i) = Pr[\{\omega_{ij}\}_{j \in S_i} > \{\omega_{ij}\}_{j \notin S_i}, \; \{\omega_{ij}\}_{j \in S_i} > \tilde{c}_{N_{S_i}}, \; \{\omega_{ij}\}_{j \notin S_i} < \tilde{c}_{N_{S_i}+1}] \tag{2.9}$$

When ϵ_{ij} is distributed i.i.d. Gumbel, the model in equation 2.9 generalizes the exploded logit model in two specific ways—it incorporates partial rank orderings (we know only that alternatives in S_i are of higher rank than those excluded from S_i) and includes a means of estimating the truncation point of the consideration set (via the search cost variables, $\tilde{c}_n$).

Equation 2.9 has a closed form solution when ϵ_{ij} is distributed i.i.d. Gumbel (somewhat surprisingly, given the non-trivial additions it makes to the exploded logit model). The closed form solution is provided in equation 2.10. A proof of this solution can be found in Appendix 2.2. Appendix 2.3 also contains the results of simulation studies run to demonstrate that the parameters β_i and $\tilde{c}_n$ can be accurately retrieved in an estimation algorithm.

$$
\begin{aligned}
\pi(S_i) = \prod_{j^+ \in S_i} & [1 - exp(-exp(w_{ij^+})exp(-\tilde{c}_{N_{S_i}}))]exp(-a\, exp(-\tilde{c}_{N_{S_i}})) \\
& + exp(-a\, exp(-\tilde{c}_{N_{S_i}+1})) - exp(-a\, exp(-\tilde{c}_{N_{S_i}})) \\
+ \sum_{t=1}^{k} (-1)^t & \sum_{q=1}^{\frac{N_{S_i}!}{t!(t-N_{S_i})!}} \frac{a\,[exp(-(a+b_{qt})exp(-\tilde{c}_{N_{S_i}+1})) - exp(-(a+b_{qt})exp(-\tilde{c}_{N_{S_i}}))]}{a+b_{qt}}
\end{aligned} \tag{2.10}
$$

Here $a = \sum_{j^- \notin S_i} exp(w_{ij^-})$, $b_{qt} = \sum_{j \in G_{qt}} exp(w_{ij^+})$, and G_{qt} is the q^{th} subset of alternatives $j^+ \in S_i$ of size n (of which there are $\frac{N_{S_i}!}{t!(t-N_{S_i})!}$ in total).

Note that the solution, while complex, actually consists of only four unique parts: the exponentials of (1) the consideration propensities for all alternatives $j^+ \in S_i$, $exp(w_{ij^+})$, (2) the consideration propensities for all alternatives $j^- \notin S_i$, $exp(w_{ij^-})$, (3) the (negative of the) marginal search cost for the $N_{S_i}^{th}$ considered alternative $exp(-\tilde{c}_{N_{S_i}})$, and (4) the (negative of the) marginal search cost for the $(N_{S_i}+1)^{th}$ considered alternative, $exp(-\tilde{c}_{N_{S_i}+1})$.

Critically, the number of calculations in the likelihood statement does not increase exponentially in the number of available alternatives. Estimation is therefore not impeded the curse of dimensionality. Though the third line of equation 2.10 contains $2^{N_{S_i}} - 1$ calculations (and is thus exponentially increasing in the size of the consideration set, N_{S_i}), these calculations are merely $2^{N_{S_i}} - 1$ combinations of N_{S_i} terms; they are easily constructed in an estimation algorithm, and take little time to compute, even for relatively large values of N_{S_i}.[3]

2.5.4 Relationship to level model and exploded logit model

The model presented in equations 2.9 and 2.10 is directly linked to two commonly used models in the literature—the level and exploded logit models. If marginal search costs $\tilde{c}_n$ are estimated to be zero for all n, our model reduces to the level model in equation 2.3. Our model therefore generalizes the level model, and can be formally tested against it to see if the data support the hypothesis that consumers engage in consideration set substitution.

[3] This formulation works well in our empirical context, where set sizes range from one to four. There will be some limit to the size of consideration sets that can be modeled. That limit will depend on computing power and data set size.

Alternatively, if search costs are removed from equation 2.9 altogether (not merely set to zero), equation 2.9 becomes a simple rank order preference model:

$$Pr(\{\omega_{ij}\}_{j\in S_i} > \{\omega_{ij}\}_{j\notin S_i}) \tag{2.11}$$

Because the stochastic component of consideration propensity, ϵ_{ij}, is distributed i.i.d. Gumbel, the probability statement in equation 2.11 becomes a simple summation of exploded logit probabilities (Chapman and Staelin 1982). Specifically:

$$Pr(\{\omega_{ij}\}_{j\in S_i} > \{\omega_{ij}\}_{j\notin S_i}) = \sum_{r=1}^{N_{S_i}!} Pr(R_r) \tag{2.12}$$

In equation 2.12, R_r is the r^{th} rank ordering of all alternatives in S_i, and $Pr(R_r)$ is the exploded logit probability associated with this rank ordering. For example, if $S_i = \{a, b\}$ and A is the set of all alternatives, then there are two possible rank orders ($R_1 = (a, b)$, $R_2 = (b, a)$), and:

$$\sum_{r=1}^{N_{S_i}!} Pr(R_r) = \frac{exp(w_{ia})}{\sum_{j\in A} exp(w_{ij})} \frac{exp(w_{ib})}{\sum_{j\in\{A-a\}} exp(w_{ij})} + \frac{exp(w_{ib})}{\sum_{j\in A} exp(w_{ij})} \frac{exp(w_{ia})}{\sum_{j\in\{A-b\}} exp(w_{ij})} \tag{2.13}$$

2.6 Analyses

We estimate the impact of vehicle redesigns, the Toyota recalls, and the Tōhoku earthquake and tsunami using our model. These estimates are compared to those from a restricted version of the model that artificially constrains consideration set substitution (our "benchmark" model). Differences between the two models' estimates serve as a barometer for the importance of accounting for consideration set substitution.

2.6.1 Benchmark model

The level model is an attractive option for a benchmark, since it is commonly used, is properly nested in the proposed model, and does not account for consideration set substitution. Because our dataset consists only of consumers who considered at least one vehicle, we can enact parametric restrictions to both the full and benchmark model (in a manner similar to Van Nierop et al. 2010) to accommodate this feature. Specifically, the restriction $\tilde{c}_1 = -99$ imposes that consumers will always consider at least one alternative. For the benchmark model, we further restrict $\tilde{c}_n = 0 \,\forall\, n > 1$, which imposes that the first restriction is the only source of consideration set substitution. In other words, the restricted model implicitly assumes that consideration set substitution only occurs when a consumer with a consideration set of size one is induced to remove the lone considered alternative from his or her set.

2.6.2 Alternative space

We model 54 C and CD cars available from 2009-2011. For vehicles in other classes, we (mostly) model class-and-continent-specific "outside options." One exception is that we model Toyota and Honda vehicles from the B Car (super compact) and DE Car (full size) classes as individual alternatives to facilitate more accurate estimates of recall and tsunami effects (which primarily affected Toyota and Honda). For estimation speed and stability, we do not model a few C and CD cars, nor the European B Car and Truck outside options. These alternatives received exceptionally small market share. In total, we model 76 alternatives that consumers can consider and purchase. Details are provided in Appendix 2.4.

2.6.3 Empirical specification of consideration propensity

For both the full and restricted model, we take a Hierarchical Bayesian approach to modeling consumer i's propensity to consider alternative j from vehicle class l at time t:

$$\omega_{ijt} = \alpha_{ij} + \gamma D_{jt} + \delta R_{jt} + \lambda E_{jt} + \beta X_{ij} + \eta U_{jt} + \epsilon_{ijt} \tag{2.14}$$

$$\alpha_{ij} = \alpha_j + \xi_{il} \qquad \begin{pmatrix} \xi_{i,C} \\ \xi_{i,CD} \end{pmatrix} \sim MVN(0, \Sigma_\xi) \qquad \Sigma_\xi = \begin{bmatrix} \sigma_C^2 & \sigma_{C,CD} \\ \sigma_{C,CD} & \sigma_{CD}^2 \end{bmatrix}$$

It is not feasible to estimate 76 fully heterogeneous alternative-specific constants (α_{ij}) using only one observed consideration set per consumer (Andrews, Ainslie, and Currim, 2002), so we instead estimate heterogeneous preferences for our focal vehicle classes, C and CD Cars ($\xi_{i,C}$ and $\xi_{i,CD}$). A larger covariance matrix with more classes (e.g., super compact "B" cars) was not empirically identifiable. We therefore rely on the covariates in our model to explain heterogeneous preferences for other vehicle classes. σ_C^2 is fixed to one for identification.

Redesigns are modeled using the set of variables D_{jt}. We include a dummy variable for each redesigned vehicle. For any redesign variable, the variable is equal to one if alternative j is the associated redesigned vehicle and if consumer i purchased his or her vehicle during any quarter in which the redesigned generation of the vehicle was available.

R_{jt} is the Toyota recall variable. For any respondent that purchased during or after the second quarter of 2010 (after which all Toyota recalls had been announced), R_{jt} is equal to one if vehicle j was (a C- or CD-car) manufactured by Toyota. E_{jt} is a set of five separate manufacturer-and-class-specific variables that account for the effect of the earthquake and tsunami on (1) Toyota C/CD cars, (2) Toyota vehicles from other classes, (3) Honda C/CD cars, (4) Honda vehicles from other classes, and (5) C/CD cars from all other Japanese manufacturers (bundled together they have approximately the same market share as Toyota or Honda do individually). For any respondent that purchased during or after the second quarter of 2011, a tsunami variable is equal to one if vehicle j was made by the corresponding manufacturer(s) and comes from the corresponding vehicle classes.

X_{ij} consists of 69 consumer $\times$ vehicle class interaction terms designed to control for heterogeneous preferences. Consumer purchase history, demographic, importance rating, and product use variables are interacted with vehicle class dummy variables. U_{jt} is a "partial unavailability" variable, and can be thought of as a nuisance variable used to control for abnormally low consideration or purchase for a vehicle in the first quarter of its availability or at the tail end of its availability. Details about X_{ij} and U_{jt} can be found in Appendix 2.4.

Consistent with past research (Bronnenberg and Vanhonacker 1996), Ford Motor Company has stated that vehicle price tiers, rather than price, drive consideration decisions. For example, a low income household may consider several compact cars (knowing they are priced economically) but never think to consider premium SUVs. More granular price differences between vehicles are factored in by consumers at the choice stage. Since vehicles within a class are similarly priced, X_{ij} includes vehicle class $\times$ income interaction terms, while the specification for choice utility (see section 2.6.6) includes price $\times$ income interaction terms.

2.6.4 Empirical specification of "search costs"

Since we only observe consumers with consideration sets of sizes one through four, $\tilde{c}_1$ and $\tilde{c}_n$ for $n > 4$ are not identified. Recall that $\tilde{c}_1$ is set to -99. We set $\tilde{c}_5$ similarly to 99 (i.e. to values suitably close to -∞ and ∞ on the logit scale). We must fix one more parameter for identification, and so set $\tilde{c}_2$ equal to zero.[4] Lastly, we set $\tilde{c}_3 \geq \tilde{c}_2$ and $\tilde{c}_4 \geq \tilde{c}_3$, in line with our model's assumption that net marginal expected benefit from search is non-increasing in

[4]We could instead fix one alternative-specific constant from ω_{ijt}, but doing so leads to a great deal of autocorrelation in our Markov chain.

consideration set size:

$$\tilde{c}_3 = exp(\theta_3) \quad \tilde{c}_4 = \tilde{c}_3 + exp(\theta_4) \tag{2.15}$$

Heterogeneity (observable or unobservable) in search costs and consideration propensity are not separately identifiable. We include heterogeneity only in consideration propensities because doing so allows for greater flexibility. Variables intended to capture heterogeneous preferences for vehicle attributes like vehicle class can absorb the impact of heterogeneous search costs fairly easily (e.g., by estimating higher or lower consideration propensities across all vehicle types), but the reverse is not so.

2.6.5 Identifying marginal "search costs"

Because point estimates for marginal search costs are (a) at least in part identified by functional form assumptions placed upon the unobserved (to the researcher) component of consideration propensity, ϵ_{ijt}, and (b) the sole driver of consideration set substitution patterns in our model, exogenous variation in the attractiveness or number of available alternatives is critical. We have five primary shifters of alternative attractiveness or availability: the Toyota recalls, the tsunami, vehicle redesigns, vehicle discontinuations, and the launch of new vehicles.

If consideration set substitution does not occur in a market, then changes in the attractiveness or availability of one alternative should not affect the consideration frequencies of other alternatives, and $\tilde{c}_3$ and $\tilde{c}_4$ will be estimated to be zero. However, if consideration set substitution does occur, then $\tilde{c}_3$ and $\tilde{c}_4$ must necessarily be greater than zero. Consider the Toyota recalls: If consideration set substitution does not occur in this market, then when Toyota vehicles lost considerations after the recalls, consideration of other alternatives should not have changed (if there were no other changes in the market at the same time). $\tilde{c}_3$ and $\tilde{c}_4$ are identified (in part) by the degree to which non-Toyota alternatives saw their consideration frequency rise in the aftermath of the recalls (and in part by consumer response to other changes in the market).

2.6.6 Empirical specification of choice utility

Consumer i derives utility from the purchase of vehicle j, given by:

$$u_{ijt} = \alpha_j^c + \gamma^c D_{jt} + \delta^c R_{jt} + \lambda^c E_{jt} + \beta^c X_{ij}^c + \eta^c U_{jt} + \zeta^c P_{jt} + \varepsilon_{ijt} \tag{2.16}$$

$$\varepsilon_{ijt} \sim i.i.d.\ Gumbel$$

The choice utility specification includes all redesign, recall, and earthquake covariates that were included in consideration propensity, allowing for the identification of which affect choice, consideration, or both. There are three primary differences between the specification of consideration propensity and choice utility. First, alternative-specific constants are modeled as homogeneous in choice utility. Unobservable heterogeneity in vehicle class preference is not empirically identifiable in choice utility due to the small sizes of consideration sets. Second, X^c_{ij} (in choice utility) only includes a subset of the covariates from X_{ij} (from consideration propensity), because some of the interaction terms were not empirically identifiable at the choice stage. Third, a set of four price $\times$ income interactions (P_{jt}) are included in choice utility. Dummy variables for each of four household income groups ($0 - $24,999, $25,000 - $49,999, $50,000 - $79,999, and $80,000+) are interacted with a vehicle price variable. The price variable is defined as the (inflation-adjusted) average price paid by respondents during quarter t for vehicle j.

2.6.7 Bayesian estimation

We estimate the parameters of model using a Metropolis-Hastings algorithm with a normal random-walk proposal. The parameters of $\sum_\xi$ are sampled over the unidentified space using Gibbs steps, and posterior distributions for the identified parameters are obtained by dividing the draw for $\sum_\xi$ by the draw for σ^2_C (McCulloch and Rossi 1994). We draw 8,000 values for each parameter and dispose of the first 3,000 (burn in).[5] The remaining 5,000 are used to generate parameter estimates and credible intervals (Appendix 2.1). The consideration and choice levels of the model were estimated separately. Note also that the choice level of the model does not differ across the full and restricted models. Trace plots were checked for convergence. Our estimation algorithm can be found in Appendix 2.5.

2.6.8 Simulation studies

The objectives of these analyses are (1) to measure how vehicle redesigns, the Toyota recalls, and the tsunami affected consideration and choice, and (2) examine whether and how estimates from the proposed model (which accounts for consideration set substitution) differ from those from the restricted model. For each model, we run a baseline simulation (using

[5]We use parameter estimates from a logistic regression run on our data sample as starting values for our model's homogeneous parameters. This allows for fairly quick convergence.

the model's parameter estimates) and counterfactual simulations wherein an effect of interest (e.g., the recalls) is "removed" by setting the associated parameter(s) to zero. For the redesign counterfactuals we also reduce the price of the redesigned vehicles to pre-redesign levels. If consideration is higher (lower) for a counterfactual simulation than the baseline, the event associated with the counterfactual had a negative (positive) impact on consideration. The same is true for choice. Appendix 2.6 contains a detailed description of our simulation approach.

In table 2.3, we compare the simulated considerations and purchases (by vehicle type) from the baseline simulations to the summary statistics from our data sample. The simulated results strongly mirror the data sample's summary statistics, giving us confidence that our counterfactual analyses have an accurate baseline off of which to work from.

2.7 Results

2.7.1 Parameter Estimates

The full model's parameter estimates (Table 2.4) for marginal search costs are greater than zero ($\tilde{c}_3$ and $\tilde{c}_4$ are estimated to be .218 and .399, respectively), supporting the hypothesis that consideration set substitution does occur in this market. The 95% credible intervals for $\tilde{c}_3$ and $\tilde{c}_4$ are (0.189, 0.246) and ($\tilde{c}_3$ + 0.149, $\tilde{c}_3$ + 0.218), and the lowest draw (of 5,000) from the posterior distribution of $\tilde{c}_3$ was 0.175—suitably far from zero. If consideration set substitution did not occur in this market, one would expect a mass of points at zero in the posterior distribution for these parameters.[6]

For the consideration stage of the model, the redesign parameter estimates for the consideration level of the model were significant and positive for five of the six vehicles under study (all but the Subaru Legacy). The Toyota C/CD car recall parameter (-0.242) was also significant, as expected given the sharp drop in Toyota considerations post-recall. We find that the redesign parameters and the recall parameter were n.s. for the choice level of the model (Table 2.5).

The tsunami parameter estimates were also mostly significant at the consideration stage. We have evidence that the tsunami hurt consideration of Toyota and Honda, but not other Japanese manufacturers. One could have reasonably hypothesized that the tsunami would have primarily had supply-side consequences for these manufacturers, possibly manifesting

[6]In practice, we constrain the marginal search cost parameters below at zero by making these parameters exponential terms. Thus, we would not expect values of precisely zero for $\tilde{c}_3$ and $\tilde{c}_4$ if consideration set substitution did not occur, but values very close to zero (e.g., zero to the fourth or fifth decimal), and would expect that the posterior distribution for these two parameters would not converge in a region so far from zero.

as lower conditional choice probabilities. We do find this to be the case for Japanese manufacturers other than Honda and Toyota—consideration probabilities remained the same post-tsunami (the tsunami parameter was not significant in the consideration stage of the model), but conditional choice probabilities were lower post-tsunami (the tsunami parameter was significant and negative in the choice stage). Interestingly, however, we find the opposite for Toyota and Honda vehicles—consideration probabilities for these vehicles are lower post-tsunami, while conditional choice probabilities are unchanged. This result is striking. While we cannot draw a conclusion as to why consumers reacted this way, the result is consistent with at least one story: perhaps consumers anticipated limited supply, and adjusted their consideration sets accordingly.

We find that wealthier consumers are (unsurprisingly) more likely to consider expensive vehicle classes and that at the choice stage, higher income consumers are less price sensitive than consumers with lower levels of income (Appendix 2.7).

The parameter estimates for the various controls we've included in the model can be found in Appendix 2.7. Our estimates for the parameters intended to measure heterogeneity through past purchase history are mostly positive and significant for the consideration level of the model, indicating that (as would be expected) consumers who previously purchased a vehicle type are likely to consider it again in the future. For example, consumers who are buying a new vehicle to replace a previously purchased one are very likely to consider the same nameplate again—the parameter estimate for this variable is 1.171, a huge value on the logit scale. Some of these parameters are also significant for the choice stage of the model, but not as many. The parameter estimates for demographic and other interaction variables are fairly intuitive: consumers to whom fuel efficiency matters are more likely to consider smaller vehicles; men are more likely to consider premium cars and trucks than are women; consumers who rate themselves as brand loyal are more likely than others to again consider a previously purchased manufacturer; and consumers who use their vehicle for towing, hauling, and off-roading are more likely to consider larger vehicles such as vans, utilities, and trucks.

2.7.2 Full versus Restricted Model

Recall that we benchmark the performance of the proposed model against a restricted version of that model. The restriction employed—that the marginal search costs for all set sizes except a size of one are equal to zero—reduce the proposed model to a level model similar to the one used in Van Nierop, et al. (2010). The restricted model therefore can only account for consideration set substitution in one specific circumstance—when the lone alternative in a set of size one is removed and another is added in its place.

The restricted model's inability to fully account for consideration set substitution has two primary consequences. First, the restricted model underestimates the conversion rate of considerations gained by marketing actions or market events (the probability that a gained consideration will translate to a purchase). The restricted model cannot account for the fact that a consumer who adds an alternative to his or her consideration set (e.g., due to advertising) may in turn kick out a competing alternative from that set. The restricted model therefore overestimates the consideration set size of these consumers and underestimates the purchase probability for the added alternative.

Second, the restricted model misestimates the number of considerations gained (or lost) due to a marketing action (or any other effect of interest). This occurs because the restricted model cannot take into account that some of an alternative's period-to-period gains are substitutions coming from another alternative's losses (or vice-versa). For example, some of the considerations that Toyota lost after the recalls were stolen by redesigned vehicles, and would have been lost even if Toyota had not been harmed by the recall crisis.

Lastly, a third consequence of the restricted model stems from its failure to incorporate search costs: the parameters of consideration propensity must be fit to both (1) the distribution of consideration set sizes observed in the data and (2) the relative frequency with which each alternative is considered. The model may fit one or both poorly without the explanatory power of search costs (which truncate a consumer's consideration set). In our empirical context, the restricted model fits the distribution of consideration set sizes poorly, noticeably overestimating the proportion of sets of size one (54.9% in the restricted model's baseline simulation versus 45.8% in the data and the full model's baseline simulation) and of size four or greater (6.5% in the data and 6.8% for the full model versus 8.7% for the restricted model; Table 2.6). This issue is likely to be less problematic if fairly granular unobservable heterogeneity can be accounted for (e.g., fully heterogeneous alternative-specific constants for consideration propensity).

Because consideration set sizes are not well captured by the restricted model, predicting how the two models' estimates will differ is not straightforward. An additional complicating factor is that, because the restricted model lacks the explanatory power of consideration set substitution, other parameters will be used to "fill the gap." If consideration of Toyota vehicles is lower post-recall than pre-recall, the restricted model is forced to attribute increases in consideration to other vehicles post-recall to heterogeneous differences across quarters. Consequently, the parameters of Σ_ξ are the only ones for which the two models' 95% credible intervals do not overlap.

The remainder of section six is devoted to illustrating these consequences in the context of measuring the impact of vehicle redesigns, the Toyota recalls, and the Tōhoku tsunami.

2.7.3 Vehicle Redesigns

We find that the average redesigned vehicle gained 1.9 considerations and 1.0 purchases per 100 consumers due to the redesign. The estimated consideration gains were statistically significant for five of the six vehicles. The purchase gains were significant for all six (Table 2.7).

The restricted model underestimates the frequency with which considerations gained due to the redesign lead to purchase (the conversion rate). Table 2.8 illustrates this. Among all consumers with a consideration set of size 2 in the full model's no-redesign counterfactuals (averaged across all six vehicle counterfactuals), those who added the redesigned vehicle to their consideration set in the baseline simulation saw their consideration set rise to (on average) 2.70 (rather than 3.0, indicating that 30% of those additions supplanted a competing alternative). By contrast, in the restricted model's simulation, their consideration sets rose to 3.0 (because no consideration set substitution occurs). The full model has a correspondingly higher conversion rate, with 47.6% of considerations added to a set size of two translating to a purchase (versus 42.4% for the restricted model). Similar patterns are observed with larger set sizes. The one exception here is consumers who begin with a set size of one in the counterfactual simulations—both models allow for consideration set substitution for set sizes of one (because each consumer must consider at least one alternative).

Table 2.8 also highlights that the restricted model overestimates the proportion of consideration sets of size one (and greater than four). This can be seen in the last two columns: the proportion of consideration sets of size one is 44.4% in the full model counterfactual simulations and 51.0% in the restricted model's simulations. Because of this, the two models actually predict similar degrees of consideration set substitution overall, and similar gains in purchase due to the redesigns. The difference in predicted purchase gains by the two models is particularly small for the Outback, which was predicted to have gained 0.83 purchases by the full model and 0.84 by the restricted model (Table 2.7). In this particular example, two wrongs actually do seem to make a right. But we should not conclude that the restricted model is "good enough" for this empirical setting. The poor fit of consideration set sizes, while balancing the scales for our redesign effect estimates, leads to unambiguous mismeasurement of price sensitivity. We run a counterfactual (using both models) to estimate how sales would have differed had the six vehicle redesigns been priced \$1,000 lower. Table 2.9 provides the results. For any given consideration set size, the change (induced by the lower price point at the choice stage of the model) in conversion rate of a consideration is the same for both models (as expected—the choice level of both models is the same). However, the total (across all set sizes) conversion rate is higher for the full model, because the restricted model overestimates the proportion of singleton consideration sets. Consumers with a set of size one will, of course, not be swayed by any choice stage variables, including price. The restricted model consequently underestimates own-price elasticity for the redesigned vehicles by an average of 9.9%. This can be seen in Table 2.10, which (for each redesigned vehicle) lists its average post-redesign paid price, the percentage change that a $1,000 price decrease represents, the forecasted percentage change in market share due to this $1,000 price decrease (for both models), and the corresponding own-price elasticity estimates (for both models).

2.7.4 Market Events

We find that Toyota C and CD cars lost about 4.9 considerations and 2.3 purchases per 100 consumers as a consequence of the recalls (Table 2.11), declines of about 17.9% and 19.6% respectively. In total, 40.3% of Toyota's lost considerations were absorbed by other vehicles. By contrast, the rate of consideration set substitution for losses due to the Tōhoku earthquake and tsunami was only 36.3%. This is because the Tsunami affected many close substitutes—all Japanese vehicles. Of the considerations lost by Toyota due to the recalls, 12.2% were absorbed by other Japanese C or CD cars (28.1% by other vehicles). By contrast, only 5.4% of the losses due to the tsunami were absorbed by other Japanese C or CD cars (30.9% by other vehicles). With fewer attractive substitutes available post-tsunami, consumers were driven to reduce their consideration set size more frequently. We estimate that the tsunami lost Japanese C and CD cars 9.4 considerations and 5.1 purchases per 100 consumers (Table 2.11), declines of 13.2% and 13.6%.

The restricted model overestimated the considerations and purchases lost by Toyota due to the recall by 11.2% and 12.9%. It overestimated Japanese manufacturers' consideration losses due to the tsunami by 14.4%. Interestingly, even though lost considerations were noticeably overestimated in the restricted model's tsunami counterfactuals, lost purchases were only slightly overestimated (2.5%). This occurred because the restricted model underestimated the rate of consideration set substitution that occurred in response to the tsunami. More specifically, the restricted model underestimated the proportion of considerations lost by Japanese manufacturers that were replaced by considerations of North American, European, or Korean vehicles. For intuition, consider this example: the restricted model may have correctly estimated that some hypothetical consumer, X, did not consider the Toyota Corolla as a consequence of the tsunami, but failed to recognize that consumer X replaced the Corolla with the Ford Focus in his or her consideration set. If this hypothetical consumer also considered the Honda Accord, then the restricted model would predict a consideration set consisting only of the Accord (rather than a set of both the Accord and Focus). The restricted model would therefore overestimate the probability that consumer X would purchase a Japanese vehicle. This logic, applied in aggregate, explains how the restricted model might greatly overestimate the number of considerations lost by Japanese manufacturers while only slightly overestimating the number of purchases lost.

2.8 Limitations and Potential Extensions

One limitation of our model is that consideration set substitution is a function only of consideration set size, and not consideration set composition. By contrast, consideration set substitution in the structural search cost model used by Honka (2014) is driven by consideration set composition. That model can therefore account for some substitution patterns that our model cannot. For example, if a consumer has a consideration set $S_i = \{a, b, c\}$, and w_{ia} is increased due to some marketing action by the manufacturer of alternative a, Honka's model can account for the possibility that either b or c may no longer be worth considering (because the benefit from including either alternative is decreasing in w_{ia}).

Our empirical analysis is limited by at least two factors. First, NVCS survey respondents were given the option to list up to three vehicles they considered in addition to the one they purchased. Thus, the respondents who provided consideration sets of four vehicles may have considered more than four. We cannot know whether these consumers considered exactly four vehicles or more, and so use these stated consideration sets as given. We note also that there is a steep decline from the proportion of respondents who considered three vehicles (16.1%) to the proportion that considered (at least) four (6.9%), so it is not unreasonable to think that few (if any) respondents may have considered more than four vehicles.

Second, we excluded from our sample consumers who did not consider at least one car. Consequently, we cannot measure the degree to which market events caused consumers who would have considered at least one car to stop considering cars altogether (or the reverse). We decided that including a larger number of consumers who considered the focal vehicle classes (C and CD cars) in the sample was more critical to accurately assessing substitution patterns in these classes than including consumers who exclusively considered trucks and utilities. That cars were rarely co-considered with these other vehicle types made this decision easier (Table 2.1).

Lastly, one potentially valuable extension would be to model search over a specific attribute (e.g., price, as in Honka 2014). The proposed model is flexible in that it does not *require* this, but incorporating such behavior could nonetheless be useful (e.g., to disentangle the relative importance of price and quality search).

2.9 Conclusion

Ample work has illustrated that accounting for consideration is critical to properly modeling demand. In spite of this, there is a paucity of econometric models in the literature that can both account for the cost-benefit trade-off that underlies consideration set formation and be estimated for markets with a large number of alternatives. One especially under-discussed element of the consumer's cost-benefit trade-off is consideration set substitution.

This paper develops a new model of consideration and choice that accounts for the fact that an increase in consideration of one alternative may decrease consideration of others. Moreover, it documents two primary consequences of failing to account for consideration set substitution: (1) misestimation of the impact of marketing actions and market events, and (2) underestimation of the rate at which gained considerations are converted to sales. Additionally, this paper shows that models that do not account for search costs may poorly capture the distribution of consideration set sizes, and consequently misestimate important quantities such as price elasticities.

The model developed in this paper has several features (above and beyond its ability to account for consideration set substitution) that make it an attractive option for researchers and marketing managers. First, the model is directly related to two commonly used models—it serves as an extension of the exploded logit model and generalizes the level model. Moreover, it can be formally tested against the level model to determine whether the researcher's data provides evidence that consumers engage in consideration set substitution. Finally, and perhaps most critically, estimation of the model is not impeded by the curse of dimensionality. The model is therefore especially useful for modeling markets with many alternatives—the very markets for which consumers are most incentivized to engage in limited search.

2.10 Tables

Table 2.1: Co-consideration of vehicle classes

	C/CD	B/D	PrCar	SmUtil	Utils	Trucks	Vans
Sets containing C Car	75.4%	8.9%	6.0%	5.9%	2.4%	0.9%	0.6%
Sets containing CD Car	70.5%	8.1%	9.8%	5.7%	4.8%	0.8%	0.3%

This table provides the (weighted) proportion of vehicles from consideration sets containing at least one C (row 2) or CD (row 3) car belonging to a class. E.g., 0.6% of vehicles in sets containing a C Car were vans.

Table 2.2: Considerations, purchases, & price pre- and post-redesign

	Considerations / 100			Purchases / 100			Average Price Paid		
Vehicle	Pre-	Post-	Dif.	Pre-	Post-	Dif.	Pre-	Post-	Dif.
Elantra	2.3	4.7	93%	0.9	1.6	83%	$17.7	$18.9	6.5%
Jetta	3.2	4.3	37%	1.2	1.6	29%	$24.1	$23.4	-3.0%
Legacy	0.7	1.1	44%	0.3	0.4	41%	$24.9	$26.8	7.9%
Optima	0.6	1.8	137%	0.3	0.6	121%	$19.4	$23.6	21.7%
Outback	1.3	2.8	124%	0.4	0.9	120%	$27.2	$29.7	9.3%
Sonata	2.5	6.1	138%	0.9	2.0	123%	$20.5	$23.9	16.7%

Price Pre-RD and Post-Rd are the average of prices reported by buyers of a vehicle pre- and post-redesign during the three year period under study. Prices are adjusted for inflation, with Q1 2009 as the reference point.

Table 2.3: Considerations and purchases per 100 consumers, by period

Considerations per 100 Consumers									
	Data Sample			Full Model Sim			Restricted Sim		
Period	1	2	3	1	2	3	1	2	3
C & CD Cars	116	112	116	114	117	113	112	115	110
Toyota	28.4	23.8	20.8	28.4	23.8	20.9	29.0	24.1	21.1
Honda	25.6	22.7	18.4	23.9	25.1	18.4	24.5	25.8	18.6
Japanese	22.1	22.9	22.7	21.6	23.7	22.7	20.8	22.9	21.5
Other	39.8	42.2	54.3	40.3	44.1	50.9	38.1	42.5	48.8
Other Classes	69.3	66.5	67.6	68.8	67.6	68.1	68.0	67.0	67.5
Average Set Size	1.85	1.78	1.84	1.83	1.84	1.81	1.80	1.82	1.78

Purchases per 100 Consumers									
	Data Sample			Full Model Sim			Restricted Sim		
Period	1	2	3	1	2	3	1	2	3
C & CD Cars	60.6	61.9	61.6	61.0	62.4	60.7	60.9	62.3	60.3
Toyota	15.1	13.3	10.9	14.4	12.2	10.5	15.2	12.6	11.0
Honda	12.5	11.8	9.5	12.9	13.6	9.8	13.8	14.4	10.3
Japanese	12.0	13.4	12.7	12.3	13.5	12.4	11.8	12.9	11.8
Other	21.1	23.4	28.4	21.3	23.2	27.9	20.2	22.5	27.2
Other Classes	39.4	38.1	38.4	39.0	37.6	39.3	39.1	37.7	39.7

Period 1 is Q1 2009 - Q1 2010 (Pre-recall). Period 2 is Q2 2010 - Q1 2011 (Post-Recall, Pre-Tsunami). Period 3 is Q2 2011 - Q4 2011 (Post-Tsunami).

Table 2.4: Consideration set model parameter estimates

Parameters	Full Model			Restricted Model		
Search Costs	Est.	95% Cred. Int.		Est.	95% Cred. Int.	
θ_3	-1.522	-1.665	-1.402	-	-	-
θ_4	-1.713	-1.902	-1.524	-	-	-
$\tilde{c}_3$	0.218	0.189	0.246	0.000	-	-
$\tilde{c}_4$	0.399	$\tilde{c}_3$+.149	$\tilde{c}_3$+.218	0.000	-	-
Heterogeneity Cov Matrix	Est.	95% Cred. Int.		Est.	95% Cred. Int.	
Variance - Pref for C Cars	1.000	-	-	1.000	-	-
Covariance - C & CD Cars	-0.300	-0.422	-0.187	-0.508	-0.576	-0.435
Variance - Pref for CD Cars	1.252	1.100	1.419	1.199	1.037	1.358
Redesign	Est.	95% Cred. Int.		Est.	95% Cred. Int.	
Hyundai Elantra	0.745	0.511	1.003	0.767	0.556	0.992
VW Jetta	0.326	0.118	0.555	0.329	0.133	0.547
Subaru Legacy	0.305	-0.130	0.740	0.319	-0.194	0.831
Kia Optima	1.272	0.847	1.640	1.241	0.829	1.661
Subaru Outback	0.743	0.424	1.075	0.761	0.422	1.153
Hyundai Sonata	1.037	0.796	1.290	1.103	0.875	1.395
Recall and Tsunami	Est.	95% Cred. Int.		Est.	95% Cred. Int.	
Recall - Toyota C/CD	-0.242	-0.344	-0.132	-0.241	-0.355	-0.132
Tsu - Toyota C/CD	-0.150	-0.274	-0.029	-0.144	-0.273	-0.023
Tsu - Toyota B/DE	-0.769	-1.232	-0.345	-0.771	-1.192	-0.348
Tsu - Honda C/CD	-0.371	-0.497	-0.239	-0.379	-0.499	-0.264
Tsu - Honda B/DE	-0.423	-0.817	-0.062	-0.433	-0.807	-0.061
Tsu - Japanese C/CD	-0.098	-0.216	0.029	-0.087	-0.194	0.019

Table 2.5: Choice model parameter estimates

Parameters	Choice Model		
Redesign	Est.	95% Cred. Int.	
Hyundai Elantra	-0.232	-0.789	0.267
VW Jetta	-0.179	-0.652	0.284
Subaru Legacy	1.176	-0.009	2.332
Kia Optima	-0.814	-1.882	0.176
Subaru Outback	0.506	-0.377	1.599
Hyundai Sonata	-0.511	-1.184	0.011
Recall and Tsunami	Est.	95% Cred. Int.	
Recall - Toyota C/CD	0.020	-0.225	0.300
Tsu - Toyota C/CD	-0.244	-0.580	0.084
Tsu - Toyota B/DE	0.717	-0.869	2.494
Tsu - Honda C/CD	-0.196	-0.535	0.138
Tsu - Honda B/DE	0.220	-0.878	1.438
Tsu - Japanese C/CD	-0.366	-0.679	-0.074
Price	Est.	95% Cred. Int.	
Price - Income Group 1	-0.156	-0.222	-0.093
Price - Income Group 2	-0.099	-0.139	-0.061
Price - Income Group 3	-0.097	-0.138	-0.061
Price - Income Group 4	-0.074	-0.112	-0.039

Table 2.6: Distribution of consideration set sizes

Consideration Set Size	1	2	3	4+
Data Sample	45.8%	32.5%	15.2%	6.5%
Full Model Baseline Simulation	45.8%	32.1%	15.3%	6.8%
Restricted Model Baseline Simulation	54.9%	23.9%	12.5%	8.7%

Table 2.7: Simulated considerations & purchases per 100 consumers (redesigns)

Consideration Gain						
	Elantra	Jetta	Legacy	Optima	Outback	Sonata
Full Model Estimate	2.54	1.24	0.33	1.56	1.51	4.06
Restricted Model Est.	2.48	1.16	0.21	1.57	1.54	4.20
Difference (Pct)	-2.4%	-6.9%	-35.6%	0.8%	2.4%	3.4%
Full Model 95% CI	1.59	0.41	-0.16	0.90	0.87	3.07
	3.56	2.24	0.83	2.33	2.16	5.08

Purchase Gain						
	Elantra	Jetta	Legacy	Optima	Outback	Sonata
Full Model Estimate	1.45	0.65	0.34	0.81	0.83	2.00
Restricted Model Est.	1.39	0.61	0.25	0.84	0.84	2.17
Difference (Pct)	-3.8%	-6.3%	-27.9%	4.1%	1.9%	8.4%
Full Model 95% CI	0.76	0.05	0.04	0.36	0.41	1.26
	2.24	1.36	0.65	1.34	1.29	2.76

Consideration Set Substitution Rate						
	Elantra	Jetta	Legacy	Optima	Outback	Sonata
Full Model Estimate	43.8%	43.8%	42.6%	43.6%	43.2%	42.7%
Restricted Model Est.	26.9%	28.7%	23.1%	24.8%	24.5%	26.5%
Difference (Pct)	-38.6%	-34.5%	-45.9%	-43.1%	-43.2%	-38.1%
Full Model 95% CI	33.1%	27.6%	0.00%	28.5%	32.7%	35.6%
	54.6%	60.7%	75.7%	59.1%	53.7%	49.8%

For each vehicle, we list considerations and purchases gained due to the redesign and the rate of consideration set substitution estimated using both the full and restricted model. The restricted model's percentage over- or under-estimate is also provided ('Difference').

Table 2.8: Conversion rate of considerations gained by redesigns

(1) Set Size w/o RD	(2) Set Size w/ RD		(3) Convert Rt (Sale/Consid)		(4) Convert Rt Difference	(5) % of Sets by set size	
	Full	Rest.	Full	Rest.	Full - Rest	Full	Rest.
1	1.55	1.49	78.4%	80.5%	2.1%	44.4%	51.0%
2	2.70	3.00	47.6%	42.4%	-5.2%	30.4%	21.6%
3	3.65	4.00	35.5%	32.3%	-3.2%	16.9%	13.6%
4+	4.00	5.87	31.7%	22.8%	-8.9%	8.3%	13.7%
Wt. Avg:	2.46	2.76	57.9%	57.8%	-	-	-

For consumers who considered any redesigned vehicle in the baseline simulation, but did not consider that vehicle in the corresponding "no redesign" counterfactual, we report: (column 1) consideration set size in the "no redesign" counterfactual, (2) average set size in the baseline simulation for the full and restricted model, (3) conversion rate of the gained considerations (gained purchases divided by considerations), (4) the difference between the conversion rates for the full and restricted models, and (5) proportion of consumers that began with a set of the size indicated in column 1.

Table 2.9: Change in conversion rate of considerations $1K price reduction

Full Model			Restricted Model		
(1) Set Size	(2) $\triangle$Convert	(3) % of Sets	(4) Set Size	(5) $\triangle$Convert	(6) % of Sets
1	0.0%	20.2%	1	0.0%	25.8%
2	1.8%	33.0%	2	1.8%	24.8%
3	1.9%	27.3%	3	1.9%	21.5%
4	1.8%	19.5%	4	1.7%	13.6%
-	-	-	> 4	1.5%	14.2%

Weighted Averages:

Full Model			Restricted Model		
(1) Set Size	(2) $\triangle$Convert	(3) % of Sets	(4) Set Size	(5) $\triangle$Convert	(6) % of Sets
2.46	1.5%	-	2.80	1.3%	-

For consumers who considered a redesigned vehicle, we report: (columns 1, 4) consideration set size for full and restricted model, (columns 2, 5) the change in conversion rate for considerations in sets of that size due to the $1,000 price reduction, and (columns 3, 6) the proportion of consumers who had consideration sets of that size.

Table 2.10: Own-Price Elasticities for Redesigned Vehicles

	(1) Price		(2) Purch Gain		(3) % Change		(4) Elasticity		
	Price	% Chg	Full	Rest.	Full	Rest.	Full	Rest.	Diff.
Elantra	$18,895	-5.3%	0.076	0.067	2.4%	2.2%	-0.45	-0.42	-7.2%
Jetta	$23,402	-4.3%	0.070	0.059	2.3%	2.1%	-0.54	-0.49	-10.5%
Legacy	$26,819	-3.7%	0.020	0.016	2.7%	2.5%	-0.73	-0.66	-8.4%
Optima	$23,568	-4.2%	0.033	0.028	2.5%	2.3%	-0.60	-0.54	-9.0%
Outback	$29,718	-3.4%	0.044	0.037	2.8%	2.5%	-0.84	-0.73	-12.7%
Sonata	$23,937	-4.2%	0.099	0.088	2.6%	2.3%	-0.61	-0.54	-11.5%
AVERAGE		-4.2%	0.057	0.049	2.6%	2.3%	-0.63	-0.56	-9.9%

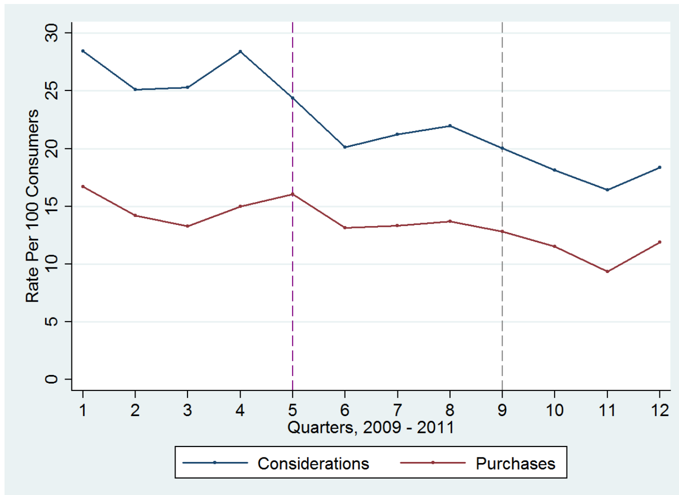
We report the (avg quarterly, post-RD) price for each vehicle and the % change a $1K price decrease represents (1); the estimated change in purchases resulting from this decrease, in absolute (2) and percentage (3) terms; and corresponding elasticity estimates (4), as well as the restricted model's % underestimation.

Table 2.11: Considerations and purchases lost due to Toyota recalls and Tōhoku tsunami

	Full Model			Restricted Model			Diff.
Recalls - Toyota C/CD	Est.	95% CI		Est.	95% CI		
Lost Considerations	-4.92	-7.21	-2.52	-5.47	-8.12	-3.25	11.2%
Lost Purchases	-2.25	-3.66	-0.88	-2.54	-3.98	-1.11	12.9%
CSET Substitution	40%	34%	47%	25%	20%	31%	-38.2%
Tsunami - Japan C/CD	Est.	95% CI		Est.	95% CI		
Lost Considerations	-9.43	-13.37	-5.52	-10.79	-15.06	-6.39	14.4%
Lost Purchases	-5.14	-7.31	-3.04	-5.27	-7.39	-3.10	2.5%
CSET Substitution	36%	30%	46%	23%	17%	31%	-37.3%

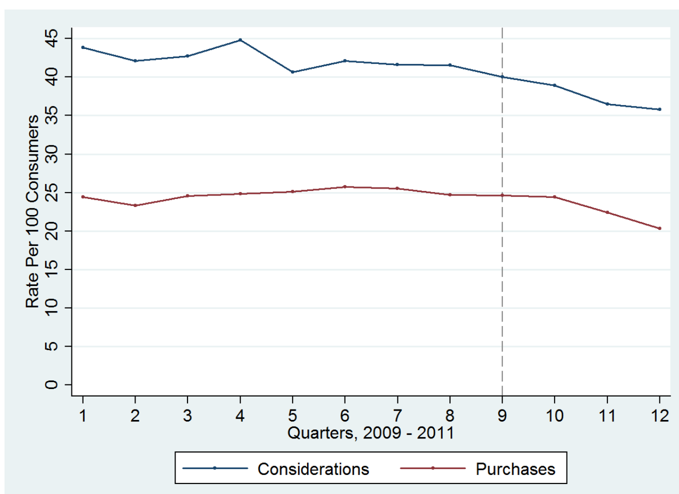
2.11 Figures

Figure 2.1: Toyota considerations and purchases per 100 consumers



Recall period begins after purple line, tsunami period after grey line.

Figure 2.2: Japanese considerations and purchases per 100 consumers*



* Excluding Toyota. Tsunami period begins after grey line.

Bibliography

[1] Andrews, Rick L., Andrew Ainslie, and Imran S. Currim. "An empirical comparison of logit choice models with discrete versus continuous representations of heterogeneity." Journal of Marketing Research 39.4 (2002): 479-487.

[2] Berry, Steven, James Levinsohn, and Ariel Pakes. "Differentiated Products Demand Systems from a Combination of Micro and Macro Data: The New Car Market." Journal of Political Economy 112.1 (2004): 68-105.

[3] Bronnenberg, Bart J., and Wilfried R. Vanhonacker. "Limited choice sets, local price response and implied measures of price competition." Journal of Marketing Research (1996): 163-173.

[4] Chapman, Randall G., and Richard Staelin. "Exploiting rank ordered choice set data within the stochastic utility model." Journal of Marketing Research (1982): 288-301.

[5] De los Santos, Babur, Ali Hortaçsu, and Matthijs R. Wildenbeest. "Testing models of consumer search using data on web browsing and purchasing behavior." The American Economic Review 102.6 (2012): 2955-2980.

[6] Draganska, Michaela, and Daniel Klapper. "Choice set heterogeneity and the role of advertising: An analysis with micro and macro data." Journal of Marketing Research 48.4 (2011): 653-669.

[7] Feinberg, Fred M., and Joel Huber. "A theory of cutoff formation under imperfect information." Management Science 42.1 (1996): 65-84.

[8] Goeree, Michelle Sovinsky. "Limited information and advertising in the US personal computer industry." Econometrica 76.5 (2008): 1017-1074.

[9] Hauser, John R., and Birger Wernerfelt. "An evaluation cost model of consideration sets." Journal of Consumer Research (1990): 393-408.

[10] Hauser, John R., et al. "Disjunctions of conjunctions, cognitive simplicity, and consideration sets." Journal of Marketing Research 47.3 (2010): 485-496.

[11] Honka, Elisabeth. "Quantifying search and switching costs in the US auto insurance industry." The Rand Journal of Economics 45.4 (2014): 847-884.

[12] Honka, Elisabeth, and Pradeep K. Chintagunta. "Simultaneous or sequential? Search strategies in the US auto insurance industry." Working Paper (2014).

[13] Honka, Elisabeth, Ali Hortaçsu, and Maria Ana Vitorino. "Advertising, Consumer Awareness and Choice: Evidence from the US Banking Industry." Working Paper (2014).

[14] Hortaçsu, Ali, and Chad Syverson. "Product Differentiation, Search Costs, and Competition in the Mutual Fund Industry: A Case Study of SandP 500 Index Funds." The Quarterly journal of economics 119.2 (2004): 403-456.

[15] Kim, Jun B., Paulo Albuquerque, and Bart J. Bronnenberg. "Online demand under limited consumer search." Marketing Science 29.6 (2010): 1001-1023.

[16] Kim, Jun B., Paulo Albuquerque, and Bart J. Bronnenberg. "The Probit Choice Model under Sequential Search with an Application to Online Retailing." Working Paper. (2014)

[17] Koulayev, Sergei. "Search for differentiated products: identification and estimation." The RAND Journal of Economics 45.3 (2014): 553-575.

[18] Luce, R. D., and Suppes, P. (1965). Preference, utility, and subjective probability. In R. D. Luce, R. R. Bush, and E. Galanter (Eds.) Handbook of Mathematical Psychology, Vol. III. New York: Wiley. pp.332-358.

[19] McCulloch, R., P. E. Rossi. 1994. "An exact likelihood analysis of the multinomial probit model." J. Econometrics 64(1–2) 207–240.

[20] Mehta, Nitin, Surendra Rajiv, and Kannan Srinivasan. "Price uncertainty and consumer search: A structural model of consideration set formation." Marketing Science 22.1 (2003): 58-84.

[21] Moraga-González, José Luis, Zsolt Sándor, and Matthijs R. Wildenbeest. "Consumer search and prices in the automobile market." (2015).

[22] Ratchford, Brian T., Myung-Soo Lee, and Debabrata Talukdar. "The impact of the Internet on information search for automobiles." Journal of Marketing Research 40.2 (2003): 193-209.

[23] Roberts, John H., and James M. Lattin. “Development and testing of a model of consideration set composition.” Journal of Marketing Research (1991): 429-440.

[24] Seiler, Stephan. “The impact of search costs on consumer behavior: A dynamic approach.” Quantitative Marketing and Economics 11.2 (2013): 155-203.

[25] Singh, Sonika, Brian T. Ratchford, and Ashutosh Prasad. "Offline and Online Search in Used Durables Markets." Journal of Retailing 90.3 (2014): 301-320.

[26] Stigler, George J. “The economics of information.” The Journal of Political Economy (1961): 213-225.

[27] Sudhir, Karunakaran. “Competitive pricing behavior in the auto market: A structural analysis.” Marketing Science 20.1 (2001): 42-60.

[28] Swait, J. D. (1984). Probabilistic choice set generation in transportation demand models (Doctoral dissertation, Massachusetts Institute of Technology).

[29] Terui, Nobuhiko, Masataka Ban, and Greg M. Allenby. “The effect of media advertising on brand consideration and choice.” Marketing Science 30.1 (2011): 74-91.

[30] Van Nierop, Erjen, et al. “Retrieving unobserved consideration sets from household panel data.” Journal of Marketing Research 47.1 (2010): 63-74.

[31] Weitzman, Martin L. “Optimal search for the best alternative.” Econometrica: Journal of the Econometric Society (1979): 641-654.

[32] Zettelmeyer, Florian, Fiona Scott Morton, and Jorge Silva-Risso. “How the Internet lowers prices: Evidence from matched survey and automobile transaction data.” Journal of marketing research 43.2 (2006): 168-181.

Chapter 3

Frugality is Hard to Afford

3.1 Abstract

Abstract Households commonly utilize strategies that provide long-term savings for everyday purchases in exchange for an increase in their short-term expenditures. They buy larger packages of non-perishable goods to take advantage of bulk discounts, and accelerate their purchases to take advantage of temporary discounts. We document that low income households are less likely to utilize these strategies even though they have greater incentives to do so. Moreover, results suggest a compounding effect: the inability to buy in bulk inhibits the ability to time purchases to take advantage of sales, and the inability to accelerate purchase timing to buy on sale inhibits the ability to buy in bulk. We find that the financial losses low income households incur due to underutilization of these strategies can be as large as half of the savings they accrue by purchasing cheaper brands. We provide causal evidence that liquidity constraints inhibit the use of these money-saving strategies.

3.2 Introduction

Households have several strategies at their disposal to reduce per-unit spending on everyday products. One set of strategies offer immediate financial savings in exchange for a decrease in utility, such as buying cheaper brands, or shopping around to find better deals. In line with the intuition that individuals who are most incentivized to save money would be most likely to use these strategies, past research has shown that lower income households purchase cheaper brands (Kalyanam & Putler, 1997; Akbay & Jones, 2005; Griffith, Leibtag, Leicester, & Nevo, 2009) and search more for lower prices (Aguiar & Hurst, 2005) compared to their higher income counterparts.

A second set of money-saving strategies offers long-term savings in exchange for an increase in short-term expenditure. For example, households may purchase larger packages to take advantage of bulk discounts that offer a lower per-unit price. Similarly, households may accelerate their purchase timing in order to stock the pantry when they find an attractive deal. These intertemporal substitution strategies are commonly used for everyday purchases of non-perishable goods. However, it is unclear whether lower income households are able to utilize these money-saving strategies to their full potential. Even though lower income households are more incentivized to save, they may face constraints that inhibit them from increasing their short-term expenditure in order to obtain long-term savings.

This paper has three objectives. The first objective is to offer a thorough investigation of the degree to which lower income households differ from higher income households in their propensity to (a) buy larger packages that offer bulk discounts and (b) accelerate their purchases to take advantage of temporary discounts. Our focus is on the choice to utilize these money-saving strategies, conditional on their availability to the household. Therefore, we study product and timing choices while controlling for each household's access to store types and brands, as the prevalence of bulk and temporary discounts vary across both. The second objective is to examine whether liquidity constraints contribute to the discrepancies between lower and higher income households' propensities to buy in bulk or to accelerate purchase timing in response to sales, even after controlling for other inhibiting factors such as storage space, access to transportation, or myopic financial decision making. The third objective is to quantify the potential financial losses that lower income households incur due to the role liquidity constraints play in shaping their purchase behavior.

In service of these objectives, we analyze the toilet paper purchases of more than 100,000 households across seven years. We discuss the data source and important data patterns in Section 3.4. The toilet paper category is tailor-made for studying the intertemporal substitution strategies of consumers for the following reasons. First, bulk and temporary discounts are both common in the category. Second, toilet paper is storable, therefore purchase acceleration and bulk buying are feasible money-saving strategies. Third, toilet paper consumption is likely to be steady and not influenced by the household's inventory. Steady consumption is a significant benefit when studying decisions involving intertemporal substitution and inventory and consumption must be inferred from purchases (Ailawadi & Neslin, 1998; Silva-Risso, Bucklin, & Morrison, 1999; Erdem, Imai, & Keane, 2003; Hendel & Nevo, 2006a; Ailawadi, Lutzky, & Neslin, 2007). Fourth, a household's need for toilet cannot be satisfied by other product categories, therefore purchasing behavior in the category is unlikely to be influenced by substitution to or from other categories. Finally, the simplicity of attributes of products in this category allows us to pool our analyses across products using a standardized unit of toilet paper. Other categories, such as laundry detergent, soda, and diapers, feature greater differentiation between products based on a higher dimension of attributes, necessitating stronger assumptions regarding comparability across products.

Our first set of analyses examines the relative frequency with which lower and higher income households (a) buy in bulk and (b) accelerate purchase timing to take advantage of a sale. The results presented in Section 3.5 indicate that lower income households are less likely than higher income households to utilize both strategies, even after controlling for differences across income groups in households' access to store types and brands. Making matters worse for lower income households, these two purchase behaviors have compounding effects on one another. A high level of inventory provides a household with more time until they run out of inventory, allowing them to time their purchases to take advantage of a price discount. Lower income households maintain lower inventory levels because they do not buy in bulk. As a result, they are more likely to run low on toilet paper, and thus need to purchase even when the prices are not favorable. In sum, the inability to buy in bulk also prevents the lower income households from timing their purchases to take advantage of sales. In return, the inability of lower income households to time their purchases to buy on sale further inhibits their ability to purchase large package sizes, as large package sizes are more affordable when on sale.

To test whether liquidity constraints play a role in shaping these purchase behaviors, in Section 3.6, we compare each household's purchases made during the first week of the month—when its liquidity constraints are partially relaxed by the receipt of paychecks (and in some cases, food stamps)—to its purchases made during the rest of the month. We find that lower income households buy larger packages and accelerate their purchases in order to take advantage of a sale to a greater degree during the first week of the month than during the rest of the month. By contrast, the purchase behavior of higher income households differs to a lesser degree between these two periods. Importantly, the receipt of paychecks and food stamps does not affect the geographic location of a household, its ability to plan for future consumption, its storage space, its transportation options, or the set of brands available at nearby stores. Consequently, our empirical approach identifies the impact of liquidity constraints on purchase behavior, above and beyond any effect that these other factors may have. The results indicate that lower income households would utilize these money-saving strategies more often, if they could afford them.

In Section 3.7, we quantify a lower-bound of the potential for economic harm to low income households stemming from their relative inability to buy in bulk and to accelerate their purchase timing to take advantage of sales. We show that households with an annual income of \$20,000 or less pay at least 5.9% more per sheet of toilet paper due to their inability to utilize these strategies to the same extent as households with an annual income of \$100,000 or more. Interestingly, these low income households save approximately 8.8% per sheet of toilet paper by purchasing cheaper brands than high income households. Therefore, about two-thirds of the savings low income households accrue through brand choice is forfeited by their relative inability to utilize intertemporal money-saving strategies. We also show that during the first week of the month, low income households eliminate around 30% of their price-per-sheet deficit by buying in bulk and on sale more often.

Our results indicate that liquidity constraints hinder the responsiveness of lower income households to bulk and temporary discounts, even for seemingly low-priced, everyday purchases. Understanding households' intertemporal substitution patterns is critical to a retailer's promotion planning (Silva-Risso, Bucklin, & Morrison, 1999). Therefore, our findings point to a few retail recommendations. A retailer aiming to incentivize lower income families to consolidate their purchases in its stores, for example, could schedule temporary discounts more frequently during times of relatively higher liquidity. Additionally, they could offer liquidity assistance to encourage buying in bulk. Our findings also have policy implications. Policy makers should consider factors that inhibit lower income households from making financially beneficial choices, in addition to factors that limit the accessibility of stores (Kaufman, 1997; Chung & Myers, 1999; Talukdar, 2008) or factors that impede the development of financial literacy (Fernandes, Lynch, & Netemeyer, 2014). Public policy aimed at decreasing the costs of poverty has been especially focused on issues of store access. While helping improve lower income households' access to supermarkets could be beneficial by expanding their access to money-saving opportunities, providing liquidity relief would also help them take greater advantage of the money-saving opportunities already available to them.

3.3 Related Literature

This paper contributes to 1) literature that documents different financial burdens shouldered by the poor (the "poverty penalty") and 2) the marketing literature on drivers of deal-proneness. Within the poverty penalty literature, research on how income groups differ in their ability to buy in bulk has produced conflicting results. Some have found that the poorest households in developing (Rao, 2000; Attanasio & Frayne, 2006) and developed nations (Frank, Douglas, & Polli, 1967) pay a higher unit price for products as a consequence of buying in bulk less than their higher income counterparts, while more recent work using larger datasets from developed nations suggests the opposite (Griffith et al., 2009; Beatty, 2010). The results of these studies—and therefore the differences observed across them—may be influenced by differences in the types of stores or brands available to different income groups, as well as systematic differences between income groups in variables that are likely to affect package size choice, such as consumption or household type. Our empirical approach controls for such differences, allowing us to estimate how a low income household's package size choice would differ from that of a high income household with a similar demographic profile and consumption rate, purchasing at the same type of store, and purchasing the same brand. Given the wealth of literature showing that low income households have fewer store types and products available to them (Kaufman, 1997; Chung & Myers, 1999; Talukdar, 2008), we find such controls to be essential. We document that lower income households choose to buy smaller package sizes than higher income households, forgoing opportunities to receive bulk discounts.

The strategic timing of purchases in response to promotions is a central topic for marketers. However, to the best of our knowledge, previous researchers have not documented differences between income groups in their ability to accelerate purchase timing to take advantage of a sale. Neslin, Henderson, and Quelch (1985) were some of the first to provide a comprehensive study of consumers accelerating purchases in order to take advantage of sales, and the potential for an interaction between purchase acceleration and bulk buying, but they did not find significant differences between income groups among the sample of 2,293 consumers in their dataset. Even though larger datasets have since become available and allowed for larger scale studies, the literature has remained silent on whether lower income households are less able than higher income households to strategically time purchases to take advantage of sales. This paper shows that lower income households—those likely to be most price sensitive—are less likely to accelerate purchases in order to take advantage of sales than are other households, and also documents a feedback loop between the inability to accelerate purchases and the inability to buy in bulk.

The marketing literature studying differences in deal-proneness has often found that higher income households are more likely than lower income households to utilize deals such as coupons or temporary discounts (see Blattberg, Luesign, Peacock, & Sen, 1978; Lichtenstein, Burton, & Netemeyer, 1997; Bawa and Shoemaker, 1987). Our paper contributes to this literature by studying a process by which these differences may be generated. Our results suggest that the relative inability of lower income households to accelerate purchases is one potential reason low income households may be less responsive to temporary discounts.

An additional contribution of this paper is to go beyond documenting cross-sectional differences in these behaviors, to study a mechanism by which such differences may be caused. We focus on the role of liquidity constraints.[1] The fact that liquidity constraints may impact lower income household shopping behavior is intuitive, and has been hypothesized in earlier work (Kunreuther, 1973; Griffith et al., 2009; Beatty, 2010). Despite its appeal and importance, empirical evidence regarding the impact of liquidity constraints on the utilization of money-saving strategies is lacking. We provide causal evidence that liquidity constraints are at least partially responsible for inhibiting low income households from buying in bulk and accelerating purchases to take advantage of sales. We show that the effect of liquidity constraints on these behaviors transcends the effect of education, transportation, and storage constraints—factors that previous literature in marketing has shown explain some of the differences in deal-proneness between income groups. Clearly, these additional factors could also contribute to the behavioral differences across income groups. For example, lower income households tend to have less storage space. Bell and Hilber (2006) show that storage constraints inhibit the ability to buy in bulk. Lower income households also have fewer transportation options that allow them to access stores (Talukdar, 2008), limiting their options to buy in bulk or find sales. Additionally, low income households may have lower financial literacy or present-biased preferences (Delaney & Doyle, 2012), which may result in a lower propensity to engage in strategies that require intertemporal substitution in general. While these factors may be constraining low income households, our results show that liquidity constraints impede low income households above and beyond the impact of these other factors.

[1] We provide evidence regarding the role of liquidity constraints by comparing purchases made during the first week of the month, when households' liquidity constraints are partially relaxed by the receipt of paychecks (and in some cases, food stamps), to purchases made during the rest of the month. This identification strategy builds on the literature studying the sensitivity of total consumption to changes in liquidity: due to the receipt of financial resources such as paychecks (Stephens, 2006; Zhang, 2013), food stamps (Beatty & Tuttle, 2014), social security checks (Stephens, 2003), tax rebates (Johnson, Parker, & Souleles, 2004; Parker, Souleles, Johnson, & McClelland, 2011; Misra & Surico, 2014); or due to policy changes, such as those governing tax withholding (Shapiro & Slemrod, 1995) or credit markets (Leth-Petersen, 2010). Similar identification strategies have been used in recent work drawing causal connections between changes in the income or wealth of households and private label purchases (Dubé, Hitsch, & Rossi, 2015), and shopping behaviors (Nevo & Wong, 2014).

3.4 Data

We use the Nielsen consumer panel dataset provided by the Kilts Center for Marketing at Chicago Booth for our analyses. This dataset contains all purchases by a household during the period for which it was a member of the panel. For each purchase occasion, the data provides: the retail channel shopped at; the price, quantity, and package size of each UPC purchased; an indicator for whether each UPC purchased was on sale; and the purchase date.[2] We also observe a yearly survey of each household's demographics.

The data allows us to sort households into five annual household income groups that closely mirror income quintiles in the U.S.: (Income Group 1) Less than $20,000, (Income Group 2) $20,000 to $40,000, (Income Group 3) $40,000 to $60,000, (Income Group 4) $60,000 to $100,000, and (Income Group 5) greater than $100,000.[3] As Table 3.1 shows, the panel dataset provides fairly good coverage of each income group, although it slightly over-represents the middle income groups. For simplicity, we refer to income group 1 as "low income households" and income group 5 as "high income households" going forward. We use "higher income households" to designate all households except low income households.

In total, the panel dataset contains more than 110,000 households that purchased in the toilet paper category. These households made a total of 2.8 million purchases from 2006 to 2012 in this category, with the average household purchase pattern indicating a consumption rate of slightly more than two rolls of toilet paper per week. The most commonly purchased sizes are 1-, 4-, 6-, 9-, 12-, 24-, 30-, and 36- roll packages, which account for 92% of purchases. Products from the top five brands (Angel-Soft, Charmin, Kleenex, Quilted Northern and Scott) account for 74% of purchases, and private labels account for another 19%. The most commonly visited channels for toilet paper purchases were grocery stores (47.5%), discount stores (28.9%), warehouse stores (7.6%), drug stores (6.5%), and dollar stores (5.7%).

The toilet paper category is particularly suitable to studying differences in households' purchase strategies for non-perishable products, in part because bulk and temporary discounts are common. Table 3.2 shows just how common temporary discounts are. On average, 34% of purchases featured a temporary discount, though the frequency of temporary discounts varies greatly across sizes and brands. To illustrate the depth of bulk discounts in this category, Table 3.3 provides the average price-per-standardized roll (341.8 sheets) for six major products at three different sizes. The standardization of rolls helps us more accurately compare different products. The potential for savings by buying in bulk is substantial. For

[2] Package size must be constructed from multiple variables—see Appendix 3.1 for details.

[3] Actual quintiles, per http://en.wikipedia.org/wiki/Household_income_in_the_United_States, were $0-$25K, $25-$45K, $45-65K, $65-105K, and $105K + in 2011. The income buckets used by Nielsen, though fairly granular, do not allow for groupings that perfectly match these quintiles.

example, a standardized roll of Charmin Ultra-Soft costs 26.7% less if it comes from a 12-roll package than if it comes from a 4-roll package. Table 3.4 provides the sales-weighted average bulk discount provided by different package sizes across all products in a channel during a given quarter, compared to the 4-roll base size. Within a channel and quarter, the average 12-roll package offers a 19.8% discount per standardized roll compared to a 4-roll package of the same product, and the average 36-roll UPC offers a 44.3% discount.

These summary statistics indicate that both temporary and bulk discounts are highly prevalent in this product category. The raw data summarized in Table 3.5 reveals stark differences across income groups in the propensity to utilize these discounts. The households in the lowest income group (INC_1) buy on sale less often and purchase smaller sizes than higher income households. The highest income group (INC_5) purchases items on sale 38.9% of the time, while the lowest income group purchases items on sale only 28.3% of the time. The highest income group also buys UPCs that contain 4.8 more rolls and 1,558 more sheets, on average, than those purchased by the lowest income group. Importantly, however, low income households are not forgoing every opportunity to save. They are more likely to purchase the cheapest product available when buying a given package size.

These data patterns are consistent with the notion that low income households may have trouble utilizing intertemporal money-saving strategies, but could also be driven by a lack of access to the types of stores or brands that more readily offer bulk and temporary discounts (e.g. supermarkets and national brands). They could also be due to low income households purchasing cheaper brands (e.g., private labels), if those brands are less likely to be offered in larger sizes or on sale. Consistent with past research, the raw data shows that low income households are less likely to purchase at grocery and warehouse stores than higher income households and are more likely to purchase at drug and dollar stores (Table 3.6). UPCs sold at drug and dollar stores are considerably smaller than those sold at warehouse and grocery stores, indicating that the package sizes available at each differ. Additionally, the brands purchased at a greater relative frequency by low income households tend to have lower prices, but are purchased on sale less frequently (Table 3.7). Our analyses, therefore, control for the store types low income households visit and the brands they buy.

The observed differences between purchases made by low and high income groups could also be driven by factors that systematically differ across income groups, such as consumption rates and demographics. A household's consumption rate is particularly important to account for, as households that consume toilet paper faster are incentivized to carry larger inventories and may consequently be incentivized to buy in bulk or stockpile in response to sales to a greater degree. The data in Table 3.8 shows that low income households have different demographic profiles than higher income households. For example, they tend to have fewer occupants and lower levels of education. They also have lower average levels of daily consumption than higher income households (Table 3.9), though this is in large part driven by the greater proportion of single-person households at lower income levels. Our analyses control for differences across income groups in demographics and consumption, to ensure that inferences are drawn between households with similar levels of consumption and demographic profiles.

Einav, Leibtag, and Nevo (2010) note the importance of taking into account potential recording discrepancies in this dataset. Such discrepancies are apparent in the raw data (e.g., package sizes reported to contain 1,296 rolls of toilet paper). Another issue is that a few households do not seem to be reporting all of their purchases (e.g., for some households, several years elapse between reported toilet paper purchases). For these reasons, we make a (conservative) effort to clean the data, correcting or removing entries that suffer from severe discrepancies. We retain 97.4% of observations for our analyses, and our conclusions are not sensitive to the removed entries. Our cleaning approach may be useful to other researchers using this dataset, so a detailed description is provided in Appendix 3.2. The key results of our analyses using the full dataset can also be found in Appendix 3.2.

3.5 Differences in Purchase Decisions

In this section, we document differences in the propensity of households from different income groups to buy in bulk (Section 3.5.1) and to accelerate purchases in order to buy on sale (Section 3.5.2). We are investigating the purchase behavior of low income households within their typical shopping context, and therefore control for differences across income groups in households' access to different types of stores and brands, as well as brand preferences. In effect, we are comparing how likely a low income household is, relative to a higher income household, to engage in these intertemporal money-saving strategies when shopping at a given type of store (e.g. grocery store) and buying a given brand.

Additionally, in Section 3.5.3, we demonstrate a compounding relationship between these two intertemporal money-saving strategies. We show that an inability to utilize one impairs a household's ability to utilize the other.

3.5.1 Buying in Bulk

To determine whether low income households are less likely to utilize bulk buying as a money-saving strategy, we compare the package sizes purchased by different income groups while controlling for the types of products and stores they have access to. We regress the package size of product p purchased by household h during shopping trip t (S_{htp}) on income group indicator variables ($I[INC = i]$) and several sets of controls.

$$S_{htp} = \alpha + \sum_{i=2}^{5} \beta_i I[INC = i] + \eta_1 [Consumption]_h + \eta_2 [Demographics]_{ht} \\ + \eta_3 [Time]_t + \eta_4 [Product]_p + \eta_5 [Channel]_{ht} + \epsilon_{htp} \quad (3.1)$$

The error term ϵ_{htp} is assumed to independent across households but may be correlated within households. We therefore run statistical tests using cluster-robust standard errors, clustered by household (both here and for subsequent analyses).

The income group variable $I[INC = i]$ is equal to one if household h is a member of income group i during the year of purchase, where $i = 1$ corresponds to households with income less than \$20K, $i = 2$ to \$20K - \$40K, $i = 3$ to \$40K - \$60K, $i = 4$ to \$60K - \$100K, and $i = 5$ to greater than \$100K.

Product controls ($[Product]_p$) are fixed effects for the 97 products in the Nielsen dataset.[4] These control for the possibility that the products low income households have access to, or purchase most frequently (e.g., private labels, which are cheaper than national brands), may be more readily available in smaller sizes than those products purchased most frequently by higher income households. Similarly, channel controls ($[Channel]_t$) are fixed effects for each of the 66 channels in the Nielsen dataset and control for the possibility that low income households may have greater access to store types that carry smaller package sizes (e.g., dollar stores) than those that carry bulk sizes (e.g., warehouse stores).

Additionally, we control for consumption differences across income groups by including a third-order polynomial of a household's average daily consumption ($[Consumption]_h$) in our regression, which allows us to flexibly control for both linear and non-linear relationships between a household's consumption rate and the package sizes it purchases. Appendix 3.1 provides details for how this daily consumption rate is calculated. We also control for the annual demographic profiles of households ($[Demographics]_{ht}$), which also systematically vary across income groups, using dummy variables for its components: the categorical demographic variables previously described in Section 3.4 (e.g. household size, education). Lastly, we include time controls ($[Time]_t$) to control for seasonality and changes in the avail-

[4]We can identify 11 major products within our dataset, which account for 72.4% of purchases. From the remaining 27.6% of purchases, we construct an additional 86 conglomerate products, each of which consists of multiple, very small-share products for a given brand. The agglomeration is necessary when the data on attributes do not help distinguish between versions of products represented by different UPCs.

ability of different package sizes over time: "month" fixed effects for each four-week block between 2006 and 2012, and a weekend indicator variable.

Table 3.10 provides the parameter estimates from regression 3.1 (in the final column, "3.1D"). We find that low income households purchase UPCs with fewer sheets of toilet paper than higher income households, even after controlling for differences in consumption rates, demographics, time of purchase, brand purchased, and type of store visited. Low income households (income $<$ \$20K) purchase toilet paper UPCs containing 391.6 fewer sheets (1.14 fewer standardized rolls), on average, than those purchased by high income households (income $>$ \$100K) with similar demographics and consumption levels. Consequently, the average low income household makes 1.9 more purchase trips per year than the average high income household purchasing the same annual volume of toilet paper. Given the prevalence of bulk discounts in this category, low income households may be foregoing considerable savings by buying smaller package sizes.

Table 3.10 also provides the parameter estimates for versions of regression 3.1 that contain only a subset of our controls. This allows us to compare the degree to which various controls, and the assumptions that correspond to those controls, affect our estimates. Channel controls absorb differences between the average size purchased by low and high income households due to differences in the frequency with which each income group shops at various channels. As expected, the magnitude of the parameter estimates from the full regression (3.1D) is smaller than that from the regression that does not use channel controls (3.1C). This result suggests that accessibility of channels may play an important role in determining package size choice. Income difference estimates further increase when we remove product controls (3.1B), suggesting that the types of products low income households purchase are mostly purchased (and perhaps mostly available) in smaller package sizes. These results are consistent with the notion that issues of accessibility to channels that carry, and products that offer, larger package sizes may be a limiting factor for low income households. However, some of the attenuation of the income differences due to the inclusion of channel or product controls may be due to low income households *choosing* not to purchase at channels that carry larger UPCs, or not to purchase products that are available in larger package sizes. For example, a low income household may choose not to visit a warehouse store if it knows it cannot afford the 36-roll UPCs primarily sold there. To the extent that this is the case, the estimates in regression 3.1D can be thought of as conservative estimates of the impact of income on size choice.

3.5.2 Accelerating Purchase to Buy on Sale

Households that were not planning to purchase during trip t may be encouraged to do so if they observe a sale at the store they are visiting (Neslin, et al., 1985). If low income households have difficulty engaging in intertemporal substitution, they will naturally be less able to accelerate their purchase to take advantage of such opportunities. We test to see whether this is the case.

Hendel and Nevo (2006b) show that if a household is stockpiling for future consumption by accelerating its next purchase incidence to buy on sale, a researcher should observe two model-free patterns in the data: 1) that the interpurchase time preceding sale purchases should be shorter than the interpurchase time preceding non-sale purchases, which provides evidence that households are buying earlier than they otherwise would have; and 2) that the interpurchase time following sale purchases should be longer than the interpurchase time following non-sale purchases, which provides evidence that households are not merely buying earlier to consume more, but are storing for future consumption. By contrast, if household purchase timing is not influenced by the presence of a sale, then interpurchase times preceding sale and non-sale purchases should not differ.

If low income households are less likely to accelerate purchases than higher income households, then the difference between low income households' sale and non-sale interpurchase times should be less pronounced than the difference for higher income households. We test this hypothesis using a difference-in-differences regression, in line with Hendel and Nevo (2006b). For each product p purchased by household h during shopping trip t, we regress the time in days since the previous purchase, IPT_{htp}, on household fixed effects (α_h), a binary variable indicating whether the purchased UPC was on sale ($I[sale]_{htp}$), income group interactions with the sale variable, and our previously established controls (except consumption controls, which are redundant when household fixed effects are included).

$$\begin{aligned} IPT_{htp} = \alpha_h + \delta_0 I[sale]_{htp} + \sum_{i=2}^{5} \delta_i I[INC = i] I[sale]_{htp} \\ + \eta_2[Demographics]_{ht} + \eta_3[Time]_t + \eta_4[Product]_p + \eta_5[Channel]_{ht} + \epsilon_{htp} \end{aligned} \tag{3.2}$$

The parameter δ_0 measures the degree to which interpurchase times preceding sale purchases differ from those preceding non-sale purchases for households in the lowest income group. The parameters $\delta_i (i \geq 2)$ measure the degree to which households from higher income groups differ from the lowest income group on this measure.

We provide the results in Table 3.11. The estimate of δ_0 is negative, suggesting that the length of time between a sale purchase and a household's previous purchase is shorter than the length of time between a non-sale purchase and a household's previous purchase. In other words, low income households (income $<$ \$20,000) accelerate their purchase timing in response to sales. We find that δ_i is also negative for $i \geq 4$. This finding suggests that higher income households accelerate their purchase timing in response to sales even more than low income households do. More specifically, we find that low income households accelerate their purchase by 0.94 days on average when buying a UPC that is on sale—a 2.4% decrease in interpurchase time from that income group's average, non-sale interpurchase time (38.8 days; Table 3.5). High income households (income $>$ \$100,000) respond most aggressively, accelerating purchase by 1.62 days—a 72% increase over the number of days the lowest income group accelerated by—and a 3.3% decrease relative to the highest income group's average non-sale interpurchase time (48.4 days).

To check whether this difference in purchase timing is due, at least in part, to purchase acceleration, rather than an increase in consumption, we test whether the time until the next purchase occasion ($Duration_{htp}$) increases in response to purchasing on sale.

$$
\begin{aligned}
Duration_{htp} = \alpha_h + \delta_0 I[sale]_{htp} + \sum_{i=2}^{5} \delta_i I[INC = i] I[sale]_{htp} \\
+ \eta_3 [Time]_t + \eta_4 [Product]_p + \eta_5 [Channel]_{ht} + \epsilon_{htp} \quad (3.3)
\end{aligned}
$$

Consistent with stockpiling behavior, the results in Table 3.12 illustrate that households waited longer before purchasing again following sale purchases than they did following non-sale purchases ($\delta_0 > 0$). Interestingly, this difference is greatest for low income households ($\delta_i < 0$ for $i > 2$), even though higher income households display greater purchase acceleration. We will explore why this may be the case in Section **??**.

Low income households' relative inability to accelerate purchase timing to buy on sale implies that they should purchase on sale less often than high income households. Consistent with Hendel and Nevo (2006b), we test this conjecture by modeling the probability that product p purchased by household h at time t was made on sale using a linear probability model. We regress the sale variable $I[sale]_{htp}$ on income group variables and the full set of controls. We also introduce a new set of controls here—package size controls ($[Size]_p$). The frequency with which sales are offered may vary not only by channel and product, but also by package size. The data in Table 3.2 indicates that some package sizes may be offered on sale more frequently than others. Consequently, we include dummy variables for each package size in the data (in terms of rolls), as well as a third-order polynomial of the number of sheets each roll in the purchased product p contained.

$$I[sale]_{htp} = \alpha + \sum_{i=2}^{5} \beta_i I[INC = i] + \eta_1[Consumption]_h + \eta_2[Demographics]_{ht} + \eta_3[Time]_t + \eta_4[Product]_p + \eta_5[Channel]_{ht} + \eta_6[Size]_p + \epsilon_{htp} \quad (3.4)$$

The results presented in Table 3.13 suggest that low income households do not buy on sale as often as higher income households, even when they are purchasing the same product, of the same size, and in the same channel as higher income households. More specifically, a purchase by a low income household is 2.3% less likely to take advantage of a temporary discount than a purchase by a high income household. This 2.3% deficit is a sizable one, given that only 28.3% of low income household purchases are made on sale. Low income households appear to be foregoing the cost savings associated with temporary discounts with non-trivial frequency, even though they are generally more incentivized to save money. This result is consistent with the notion that an inability to strategically time their purchases inhibits the ability of low income households to buy on sale.

3.5.3 Compounding Relationship

A household's ability to buy in bulk has the potential to indirectly affect its ability to buy on sale, and vice-versa. For example, the purchase of a large UPC provides a bigger boost to inventory than the purchase of a small one. This provides a household with more time until they run out of inventory and must purchase again, increasing the likelihood that they will encounter a sale and be able to take advantage of it. Thus, buying in bulk may help a household strategically time its purchases to take advantage of sales. Similarly, sales on large UPCs make them more affordable and may put them within reach of households that would not be able to afford them at their regular price. Consequently, the ability to accelerate purchase timing to buy on sale helps facilitate buying bulk items. In this section, we test for evidence of this compounding relationship between buying in bulk and on sale.

We first examine whether buying in bulk helps households wait for sales. We test whether (a) low income households have lower inventories, on average, and (b) whether sale purchases are likely to be made at higher inventory levels, since purchases at low inventory levels are more likely to be induced by necessity than purchases made at higher inventory levels. If both tests are successful, this would provide support for the hypothesis that low income households are less able to wait for sales because they carry lower inventory levels.

Next, we examine whether buying on sale allows low income households to purchase larger package sizes. We test to see if low income households are disproportionately likely to increase their package size purchased when buying on sale. A successful test of this hypothesis would imply that low income households' relative inability to accelerate purchase timing to take advantage of sales also inhibits their ability to buy in bulk.

To investigate whether lower levels of inventory make it more difficult for low income households to wait for sales, we test whether (a) low income households have lower levels of inventory than higher income households, and (b) purchases made on sale are more likely to be made at higher inventory levels.

To test (a), we regress household h's average daily inventory on income dummy variables and consumption controls.[5] Because inventory is unobserved, we construct a household inventory variable in a manner consistent with past research (e.g. Neslin et al., 1985; see Appendix 3.1 for details).

$$AVGINV_h = \alpha + \sum_{i=2}^{5} \beta_i I[INC = i] + \eta_1 [Consumption]_h + \epsilon_h \tag{3.5}$$

To test (b), we use a linear probability model, regressing $I[sale]_{htp}$ on household h's inventory level (in sheets of toilet paper) at the time of purchase (INV_{htp}), household fixed effects (α_h), time controls, and controls for product, channel, and package size:

$$\begin{aligned} I[sale]_{htp} = \alpha_h + \rho INV_{hp} + \eta_2 [Demographics]_{ht} \\ + \eta_3 [Time]_t + \eta_4 [Product]_p + \eta_5 [Channel]_{ht} + \eta_6 [Size]_p + \epsilon_{htp} \end{aligned} \tag{3.6}$$

The results are presented in Table 3.14.[6] We find that low income households have 814 fewer sheets of toilet paper (2.4 fewer standardized rolls) in inventory than the highest income households, on average. Additionally, a household's probability of buying on sale is found to be increasing in their inventory at the time of purchase, consistent with the hypothesis that a purchase made at low inventory levels is more likely induced by the prospect of running

[5] For this regression, we exclude households that switched income groups during their time in the panel, because the dependent variable is household specific and the regression is thus cross-sectional. Using all households and assigning the first income group of the household for the entire time in the panel as an alternative specification produces results consistent with those presented here.

[6] As a robustness check, we also run variants of regression 3.6 on observable proxies for inventory, the results of which are consistent with the results presented here and can be found in Appendix 3.3.

out of inventory, and less likely to be strategically timed. Together, these results imply that low income households are less able to wait for a sale than higher income households.

We next use a difference-in-differences regression to test whether buying on sale helps low income households purchase bulk items. If, as hypothesized, sales on larger package sizes can make previously unaffordable sizes affordable to low income households, then (1) low income households' sale purchases should be larger than non-sale purchases, and (2) this difference should be greater than the corresponding difference between a higher income households' sale and non-sale purchases.

We regress the package size of product p purchased by household h during trip t (S_{htp}) on household fixed effects (α_h), the sale variable ($I[sale]_{htp}$), income group interactions with the sale variable, and our previously defined controls. The parameter δ_0 measures the degree to which households from the lowest income group increase the size of a product purchased when buying on sale (versus not). The parameters δ_i ($i \geq 2$) measure the degree to which households from other income groups differ from the lowest income group on this dimension.

$$\begin{aligned} S_{htp} = \alpha_h + \delta_0 I[sale]_{htp} + \sum_{i=2}^{5} \delta_i I[INC = i] I[sale]_{htp} \\ + \eta_2 [Demographics]_{ht} + \eta_3 [Time]_t + \eta_4 [Product]_p + \eta_5 [Channel]_{ht} + \epsilon_{htp} \end{aligned} \tag{3.7}$$

Table 3.15 summarizes the results. Low income households purchase UPCs containing 420 additional sheets (1.2 standardized rolls) of toilet paper when buying on sale than when not. This increase is 172.5 sheets greater than the increase seen by higher income households. All households tend to increase the amount of toilet paper that they buy in response to sale, but sales allow low income households to partially "catch up" to higher income households. This is consistent with the notion that low income households are cash constrained, as a sale on a large item might make it temporarily affordable to a cash constrained household that would otherwise be unable to afford it. Recall that, overall, low income households purchase UPCs that contain 391.6 fewer sheets than those purchased by higher income households, on average (Table 3.10). When buying on sale, low income households make up ground, erasing nearly half of this deficit by increasing their average package size purchased by 172.5 more sheets than the highest income households do.[7]

[7] In Appendix 3.4, we also investigate whether low income households are more or less likely than high income households to stockpile by purchasing multiple copies of a UPC (quantity), and whether low or high income households purchase greater total volume (package size × quantity). The results for both quantity and volume are consistent with the results presented here for package size—low income households do more stockpiling, conditional on being able to buy on sale in the first place.

3.6 Liquidity Constraints

Having established that low income households purchase smaller package sizes and accelerate purchase timing in response to sales to a lesser degree than higher income households, and that these two behaviors have a compounding nature, we next investigate whether liquidity constraints play a role in driving this differential. To do this, we utilize the partial relaxation of liquidity constraints provided by the arrival of paychecks and food stamps, in line with Stephens (2006) and Zhang (2013). Since we do not observe the actual timing of paycheck arrival for households in our data, we compare purchases during the first week of the month to purchases during other weeks. Both the semi-monthly and monthly paycheck schedules pay employees on the last day of the month. Therefore, a majority of the households get an influx of cash at the beginning of the month. A small fraction of these households also may be receiving food stamps during that time.

Low income households should benefit from this temporary boost in liquidity to a greater degree than higher income households, which may not benefit at all. We therefore hypothesize that during the first week of the month, compared to the rest of the month, (a) low income households will purchase larger package sizes, (b) accelerate purchase timing in response to sales by a greater number of days, and (c) buy on sale more frequently, as a consequence of (b). By contrast, the purchase behavior of higher income households may not differ between the first week of the month and the other weeks (if they are unconstrained), or their behavior may change to a lesser degree than low income households (if they are merely less constrained than low income households).

Our objective here is to determine whether the data support a causal link between liquidity and the ability of low income households to utilize intertemporal money-saving strategies. However, the magnitude of our findings should be interpreted as lower bounds, as (a) the receipt of paychecks and food stamps only partially relax a household's liquidity constraints, and (b) we do not observe the exact date on which paychecks or food stamps are received, so the measure of this partial relaxation is very noisy.

Importantly, a household's living arrangements, transportation options, and ability to plan for future consumption, as well as the location of nearby stores and the products sold there, are presumably time invariant.[8] Thus, the findings in this section cannot be the result of differences across income groups in access to different types of stores or products (e.g. large sizes), storage costs, transportation options, or myopic decision making. This

[8]Even if the availability of bulk or temporary discounts differs systematically by the week of the month, the difference-in-differences approach measures the difference between the change observed for low income groups and the change observed for other income groups while controlling for common trends that vary across time and store types.

rules out several alternate hypotheses as the sole drivers of the purchase behavior witnessed, though past research does suggest that these factors may be acting in concert with liquidity constraints.

3.6.1 Bulk Discounts

We first test whether purchases during the first week of the month tend to be larger than during other times of the month using a difference-in-differences regression. We regress the package size of product p purchased by household h during shopping trip t (S_{htp}) on (i) household fixed effects, (ii) a variable indicating whether the purchase was made during the first seven days of the month ($I[Week1]_{ht}$), (iii) interactions between the "first week" variable and income group variables, and (iv) the previously defined controls.

$$\begin{aligned} S_{htp} = {} & \alpha_h + \psi_0 I[Week1]_{ht} + \sum_{i=2}^{5} \psi_i I[INC = i] I[Week1]_{ht} \\ & + \eta_2 [Demographics]_{ht} + \eta_3 [Time]_t + \eta_4 [Product]_p + \eta_5 [Channel]_{ht} + \epsilon_{htp} \end{aligned} \tag{3.8}$$

As hypothesized, we find that low income households purchase larger package sizes during the first week of the month ($\psi_0 > 0$), and to a greater degree than higher income households ($\psi_i < 0$). Specifically, the lowest income households purchase UPCs containing 60.0 more sheets during the first week of the month than those purchased during the rest of the month (ψ_0; Table 3.16), and this gain is 42.5 sheets greater than the gain for high income households (ψ_5). Recall that the analyses in section 3.5.1 found that the lowest income households purchase UPCs containing 391.6 fewer sheets than the highest income households. A gain of 42.5 sheets relative to the highest income households represents a 10.8% reduction of this deficit.

Note that Table 3.16 indicates that, as with previous results, the inclusion of channel controls diminishes the estimated differences between low income households and higher income households. As mentioned previously, channel controls are intended to control for differences in access that income groups may have to different types of stores. However, these controls may lead us to underestimate the magnitude of our results if low income households are choosing not to purchase at stores that offer bulk discounts because they know they cannot afford the larger packages available there. Interestingly, during the first week of the month, low income households are more likely to purchase at discount stores than during the rest of the month.[9] Discount stores offer larger package sizes than most channels. This lends

[9] Additional details can be found in Appendix 3.5.

credence to the notion that the results presented in this paper are conservative in nature, as low income households are choosing to shop at channels with larger package sizes during times of higher liquidity.

3.6.2 Temporary Discounts

We next investigate whether the liquidity boost received during the first week of the month allows low income households to accelerate purchase in response to sale to a greater degree than during the rest of the month, and, consistent with this, purchase on sale more frequently.

The previously defined interpurchase time variable IPT_{htp} (time, in days, since the previous shopping trip) is regressed on several sets of variables. The first set includes (i) household fixed effects, (ii) a variable indicating whether the UPC purchased was on sale ($I[sale]_{htp}$), and (iii) income group interactions with the sale variable. These variables will capture the baseline behavior of households outside of the first week of the month. Also included are (iv) a variable indicating whether purchase p was made during the first seven days of the month ($I[Week1]_{htp}$), (v) an interaction between the sale and "first week" variables ($I[sale]_{htp}I[Week1]_{htp}$), and (vi) income group interactions the "first week" and sale × "first week" variables. The "first week" variable and the income interactions with it measure the degree to which interpurchase time changes for a household during the first week of the month for non-sale purchases. The sale × "first week" interaction, and the further interactions with income, measure whether the response of interpurchase time to the relaxation of liquidity that occurs in the first week of the month differs between sale and non-sale purchases. Lastly, we include (vii) the previously defined time, product, and channel controls.

$$
\begin{aligned}
IPT_{htp} = \alpha_h + \delta_0 I[sale]_{htp} &+ \sum_{i=2}^{5} \delta_i I[INC=i]I[sale]_{htp} \\
&+ \psi_0 I[Week1]_{htp} + \sum_{i=2}^{5} \psi_i I[INC=i]I[Week1]_{htp} \\
&+ \gamma_0 I[sale]_{htp}I[Week1]_{htp} + \sum_{i=2}^{5} \gamma_i I[INC=i]I[sale]_{htp}I[Week1]_{htp} \\
&+ \eta_3[Time]_t + \eta_4[Product]_p + \eta_5[Channel]_{ht} + \epsilon_{htp} \qquad (3.9)
\end{aligned}
$$

Our previous results suggest that low income households accelerate purchases in response to sales to a lesser degree than higher income households ($\delta_0 < 0$). This should be most true outside of the first week, when low income households have less liquidity ($\delta_i > 0$ for

$i \geq 2$). We hypothesize that low income households will accelerate purchase in response to sale to a greater degree during the first week of the month than at other times ($\gamma_0 < 0$), and that this difference will be greater than the increase in acceleration (if there is one) for higher income households ($\gamma_i > 0$ for $i \geq 2$). Lastly, there is no reason to hypothesize a change in the interpurchase time for non-sale purchases during the first week of the month (ψ_0 and ψ_i should be n/s). The results in Table 3.17 largely support these hypotheses. Low income households accelerate their purchases in response to sale by an additional 0.8 days during the first week of the month ($\gamma_0 = -0.8$). By contrast, the highest income households do not accelerate more during the first week of the month. Outside of the first week of the month, low income households accelerate in response to sale by 0.74 days (δ_0), while the highest income group accelerates by 1.64 days ($\delta_0 + \delta_5$). However, during the first week of the month, low income households erase this deficit entirely—accelerating by 1.54 days ($\delta_0 + \gamma_0$) to the highest income group's 1.48 days ($\delta_0 + \delta_5 + \gamma_0 + \gamma_5$).

Given these results, it stands to reason that low income households may consequently be able to take advantage of temporary discounts more often during the first week of the month. To test whether they do, we regress the sale variable $I[sale]_{htp}$ on household fixed effects, the first week variable ($I[Week1]_{hp}$), interactions between the first week variable and income group variables, and our controls.

$$I[sale]_{htp} = \alpha_h + \psi_0 I[Week1]_{htp} + \sum_{i=2}^{5} \psi_i I[INC = i] I[Week1]_{htp}$$
$$+ \eta_2[Demographics]_{ht} + \eta_3[Time]_t + \eta_4[Product]_p + \eta_5[Channel]_{ht} + \eta_6[Size]_p + \epsilon_{htp} \quad (3.10)$$

The results (Table 3.18) support the hypothesis that low income households will purchase on sale more frequently during the first week of the month ($\psi_0 > 0$) and to a greater degree than higher income households ($\psi_i < 0$ for $i \geq 4$). The lowest income households purchase on sale an additional 0.36% of the time during the first week of the month, and the gap between these households and the highest income households shrinks by 0.49%. This erases roughly 20% of their deficit relative to the highest income households, which were found to purchase on sale 2.3% more frequently than the lowest income households in section 3.5.2 (Table 3.13).

3.7 Potential Financial Impact

In this section we run back-of-the-envelope calculations to gauge the financial impact of low income households' relative inability to buy in bulk and on sale. We calculate three values: (1) how much more low income households pay by buying in bulk and on sale less than high income households (their "foregone intertemporal savings"), (2) how much less low income households pay by buying cheaper brands than high income households ("brand choice savings"), and (3) how much less low income households pay during the first week of the month compared to the rest of the month, due to buying in bulk and on sale more often during this period ("pay period savings").

The "foregone intertemporal savings" calculation compares the purchases of low income households to those of high income households. Because low income households are more incentivized to utilize intertemporal money saving strategies than higher income households, we would actually expect unconstrained low income households to utilize these money-saving strategies more frequently than high income households, rather than at the same frequency (all else equal). Consequently, the foregone intertemporal savings calculation provides a lower-bound value for how much money low income households are leaving on the table due to their relative inability to utilize these strategies. Comparing low income households' foregone intertemporal savings to their brand choice savings gives us a sense of the relative importance of buying in bulk and on sale as money-saving exercises in the toilet paper category.

Lastly, comparing low income households' foregone intertemporal savings to their pay period savings helps us gauge how much of low income households' foregone intertemporal savings can be attributed to liquidity constraints, rather than other factors such as storage constraints. The calculated pay period savings provides a lower-bound value for how much money low income households can save if they were unconstrained, as liquidity constraints are only partially relaxed during the first week of the month.

3.7.1 Calculation Methodology

If we wish to calculate how much low income households are leaving on the table by not buying in bulk and on sale as often as high income households, what we really wish to know is this: How much would low income households ($i = 1$) have saved had they purchased package sizes (in rolls) r and bought on sale at the frequency that high income households ($i = 5$) did, while holding the low income households' choice of product, channel visited, and time of purchase constant? To answer this question, we calculate 1) the frequency with which low income households purchase product j from channel c during quarter q ($\pi_{jcq|i=1}$), 2) the frequency with which low income households purchased a UPC containing r rolls on sale ($s = 1$) and not ($s = 0$), conditional on a product, channel, and quarter ($\pi_{rs|jcq,i=1}$), 3) the frequency with which high income households purchased a UPC containing r rolls on sale and not, conditional on a product, channel, and quarter ($\pi_{rs|qcj,i=5}$), 4) the average price paid by low income households for each product, in each channel, during each quarter, given sale status and package size ($P_{jcqrs|i=1}$), and 5) the average number of standardized rolls contained within UPCs for a given product, containing a given number of rolls, in each channel, during each quarter (R_{jcqr}). Using these five values, we can calculate (a) the average price low income households paid per standardized roll, and (b) the average price they would have paid per standardized roll had they purchased larger sizes and sale items at the same relative frequencies as higher income households ($\pi_{rs|qcj,i=5}$). The foregone intertemporal savings per sheet of toilet paper is given by the difference between (a) and (b):

$$Foregone\ intertemporal\ savings$$
$$= \frac{\sum_{q,c,j} \pi_{jcq|i=1} \sum_{r,s} \pi_{rs|jcq,i=1}(P_{jcqrs|i=1})}{\sum_{q,c,j} \pi_{jcq|i=1} \sum_{r,s} \pi_{rs|jcq,i=1}(R_{jcqr})} - \frac{\sum_{q,c,j} \pi_{jcq|i=1} \sum_{r,s} \pi_{rs|jcq,i=5}(P_{jcqrs|i=1})}{\sum_{q,c,j} \pi_{jcq|i=1} \sum_{r,s} \pi_{rs|jcq,i=5}(R_{jcqr})} \quad (3.11)$$

Recall that the lowest income group ($i = 1$) and the highest ($i = 5$) differ in their consumption levels, in large part because low income households are more likely to have only a single occupant. It is important to control for these differences when estimating foregone bulk discounts and saved brand premiums. Consequently, we do not lump all households together when calculating equation 3.11. Instead, we calculate it twice, using two different sets of households: once using those households that always contained a single member and never switched income groups and once using households that always contained more than one member and never switched income groups. Low and high income households within each of these sets have similar consumption rates (see Table 3.9 in Section 3.4). We then calculate a weighted average of these two results, where the weight applied to each calculation is the number of low income households in the set used for the calculation.

A similar approach is used to calculate brand choice savings and pay period savings, except that for the former we compare the brand choice frequencies of low and high income households instead of package size frequencies, and for the latter we use the data only on low income households, and compare their purchases during the first week of the month to those during the rest of the month.

3.7.2 Results

We find that the average low income household saves \$0.047 per standardized roll by purchasing cheaper products than high income households. On the other hand, the average low income household could save an additional \$0.030 per standardized roll by buying in bulk and on sale at the same frequency as higher income households. These forgone intertemporal savings lead them to pay 6% more per standardized roll, and account for two-thirds of the \$0.047 savings households accrue by purchasing cheaper products. However, we do find that increased liquidity helps low income households better take advantage of intertemporal savings. When buying during the first week of the month, the average low income household saves \$0.009 per standardized roll, making up roughly 30% of their foregone intertemporal savings.

3.8 Discussion and Conclusion

This paper provides (a) a thorough documentation of the ways in which income groups differ in their ability to utilize intertemporal money-saving strategies, (b) causal evidence that the presence of liquidity constraints impedes low income households from utilizing these strategies, and (c) an examination of the financial losses imposed on low income households by liquidity constraints' inhibition of these strategies.

We study money-saving purchase behaviors in the toilet paper category. Our focus on this category is motivated by several factors: 1) bulk and temporary discounts are both common in the category, 2) the product is not perishable, 3) consumption is relatively steady, 4) substitution to or from other categories is unlikely, and 5) product differentiation can be accounted for with a few product dimensions. For these reasons, bulk buying and purchase acceleration reflect a household's intertemporal substitution behavior in the toilet paper category more directly than in most others. Future research can devise methods to analyze purchase behavior in categories that feature a more complex set of product attributes.

We first document that low income households purchase smaller package sizes and are less able to accelerate their purchase timing in order to take advantage of a sale. Worse, low income households' inability to take full advantage of bulk and temporary discounts has a compounding nature. Because low income households buy smaller package sizes, they carry lower inventories and have more difficulty waiting for sales. Additionally, because low income households have difficulty accelerating their purchase timing to take advantage of sales, they are further prevented from buying in bulk, as larger package sizes are made more affordable when on sale. As a consequence, low income households leave money on the table. The average standardized roll of toilet paper in our dataset costs low income households about 50 cents, but we find that low income households pay an additional 3.0 cents per standardized roll than they would if they bought in bulk and on sale as often as high income households. By contrast, low income households save 4.7 cents per standardized roll by purchasing cheaper brands than high income households. Low income households' limited ability to utilize intertemporal savings strategies, then, is forcing them to forfeit approximately two-thirds of the savings they accrue by purchasing cheaper brands.

We show that the infrequency with which low income households utilize these strategies is at least partially explained by liquidity constraints. Liquidity constraints has often been cited alongside storage constraints as a factor that might inhibit the ability of low income households to buy everyday products in bulk (e.g., Kunreuther, 1973; Griffith et al., 2009; Beatty, 2010). While strong empirical support has been provided for the role storage constraints plays in shaping intertemporal purchase behavior for everyday items (Bell & Hilber, 2006), the literature has been mostly silent regarding the role of liquidity constraints. Our results show that during the first week of the month—when households receive a temporary boost to liquidity from pay checks and food stamps—low income households (a) purchase larger package sizes and (b) accelerate their purchases in response to sale to a greater degree than they do during the rest of the month. Moreover, these improvements are greater than those seen for high income households during the first week of the month, as would be expected, given that higher income household are less constrained. By buying in bulk and on sale more often during the first week of the month, low income households make up 29.5% of their 3.0 cent deficit per standardized roll.

Our results should be interpreted as a lower-bound for the financial impact of liquidity constraints on low income households, because our measures related to liquidity are noisy. This noise stems from three sources. First, the data does not provide a household's annual income—only an income group to which it belongs. Second, income is only one component of liquidity. More precise measures of income and wealth, as well as measures that explicitly take into account cost of living differences across cities and states, could potentially provide greater insight into the behaviors being studied. Third, we do not observe the exact timing of each household's receipt of paychecks and/or food stamps. While our measures related to liquidity allow us to provide evidence of a causal effect of liquidity constraints on the behaviors studied, the noisiness of the measures, coupled with the fact that paychecks and food stamps only partially relieve liquidity constraints, impairs our ability to identify precise magnitudes. Thus, we caution the reader that our results are likely to be biased towards a null effect. Even though they may be biased downwards, the reported results are sizeable and highlight how liquidity constraints influence shopping behavior in everyday product categories. Future research that uses a more precise measure of liquidity, or explicit budgets for shopping trips (as in Stilley, Inman, & Wakefield, 2010), can improve the precision of these results, and go further in modeling the relationship between liquidity relaxation and a household's ability to utilize intertemporal money-saving strategies.

Our work contributes to an important debate regarding the financial decisions low income households make. Carvalho, Meier, & Wang (2014) note that, "The debate about the reasons underlying [differences in financial decision making behavior across income groups] has a long and contentious history in the social sciences; the two opposing views are that either the poor rationally adapt and make optimal decisions for their economic environment or that a 'culture of poverty' shapes their preferences and makes them more prone to mistakes." In support of the latter view, researchers have suggested that the attentional demands of poverty reduce the cognitive capacity of the poor (Mani, Mullainathan, Shafir, & Zhao, 2013), and that low income households may be more myopic or present-biased (Delaney & Doyle, 2012; Griskevicius, Tybur, Delton, & Robertson, 2011). The fact that low income households behave more like higher income households when their liquidity constraints are relaxed provides support for the former view.

The results presented in this paper demonstrate that liquidity constraints shape purchase behavior even for seemingly low-priced product categories. While this is discouraging from a social welfare point of view, these results highlight valuable implications for retailers that serve low income households. Our results imply that these retailers should be scheduling temporary discounts more frequently during times of higher liquidity in order to increase the sales lift. The results also suggest that retailers could potentially increase the responsiveness of households to price promotions and bulk discounts by offering them liquidity assistance. Though some financing programs have been offered by retailers in the past, these programs are typically targeted at households looking to make large purchases (e.g., TVs), but not the everyday purchases that make up such a large share of low income households' expenses.

Our work also has important implications for policy makers. Policy makers, as well as many researchers studying the costs low income households face (e.g., Kaufman, 1997; Chung & Myers, 1999), have often focused on the lack of access to supermarkets as a driver of increased costs for everyday living. Our paper demonstrates that low income households may have trouble buying in bulk and on sale even within the store types they already have access to. This highlights the importance of enacting policies that take into consideration factors that influence a household's decision making process. While providing greater access to stores that offer bulk and temporary discounts could be helpful, policies designed to provide liquidity relief may also be necessary. We hope that our results contribute to the conversation on how our society can alleviate the additional financial burdens shouldered by low income households.

3.9 Tables

Table 3.1: Household-Years by Income Group

	N (HH-YR)	PCT	Census
Income Group 1 ($<$ \$20K)	38,196	10.6%	19.1%
Income Group 2 (\$20-40K)	91,545	25.3%	21.9%
Income Group 3 (\$40-60K)	83,775	23.1%	17.0%
Income Group 4 (\$60-100K)	99,646	27.5%	21.6%
Income Group 5 ($>$ \$100K)	48,950	13.5%	20.5%

Table 3.2: Average propensity of temporary discounts by brand and package size

	Angel-Soft	Charmin	Kleenex	Quilted N.	Scott	Pr. Label
1 Roll	n/a	40.2%	n/a	n/a	13.6%	32.2%
4 Rolls	23.0%	35.3%	52.3%	28.8%	28.9%	14.7%
6 Rolls	44.2%	36.0%	39.9%	31.5%	28.6%	15.5%
9 Rolls	38.5%	41.9%	58.5%	50.4%	14.1%	35.1%
12 Rolls	31.5%	42.3%	57.3%	45.7%	51.4%	28.1%
24 Rolls	17.9%	37.7%	45.7%	30.7%	43.1%	37.1%
30 Rolls	n/a	41.2%	64.4%	50.2%	28.7%	5.0%
36 Rolls	29.6%	26.3%	66.7%	4.4%	16.0%	2.6%

Table 3.3: Bulk discounts (per standardized roll) at grocery stores

	4-Roll	12-Roll	Discount	24-Roll	Discount
Angel Soft Classic White	$0.63	$0.55	12.4%	$0.52	16.8%
Charmin	$0.69	$0.55	20.7%	$0.53	24.0%
Charmin Ultra-Soft	$1.49	$1.09	26.7%	$1.03	31.0%
Charmin Ultra-Strong	$1.65	$1.15	30.3%	$1.09	33.9%
Kleenex Cottonelle Ultra	$1.43	$1.07	25.2%	$1.10	22.7%
Quilted Northern	$0.89	$0.68	24.1%	$0.60	32.7%
Scott 1000	$0.31	$0.25	18.1%	$0.12	61.4%

Table 3.4: Bulk discounts averaged across channels and products

Package Size	Δ Unit Price (%)
1 Roll	+29.4%
4 Rolls	—
6 Rolls	-8.4%
9 Rolls	-15.2%
12 Rolls	-19.8%
24 Rolls	-25.3%
30 Rolls	-43.3%
36 Rolls	-44.2%

Table 3.5: Purchase Behavior Across Income Groups

	INC_1	INC_2	INC_3	INC_4	INC_5
Pct of purchases cheapest* | Size (Rolls)	35.1%	31.8%	30.6%	28.7%	27.4%
Pct of purchases made on sale	28.3%	31.6%	33.9%	36.8%	38.9%
Average UPC size purchased (rolls)	9.8	10.9	11.7	13.1	14.6
Average UPC size purchased (sheets)	3,277	3,611	3,866	4,306	4,835
Average interpurchase time (days)	38.8	39.2	39.8	42.9	48.4

*Summary statistics of purchases that match to the retail scanner dataset.

Table 3.6: Purchase frequency by channel

	Avg Size*	INC_1	INC_2	INC_3	INC_4	INC_5
Discount	11.4	29.6%	31.1%	30.5%	28.4%	23.2%
Dollar	8.1	13.0%	8.2%	4.9%	3.0%	1.4%
Drug	9.6	7.5%	6.9%	6.2%	5.7%	5.5%
Grocery	9.4	43.2%	45.8%	47.9%	49.0%	50.8%
Warehouse	31.1	2.6%	4.5%	6.6%	10.2%	15.3%
Other	10.7	4.0%	3.6%	4.0%	3.7%	3.8%

* Average package size in standardized rolls

Table 3.7: Purchase frequency and price-per-standardize-roll by Brand

Brand	\$/Roll*	Sale %	Purchase Frequency				
			INC_1	INC_2	INC_3	INC_4	INC_5
Angel-Soft	\$0.55	27.8%	17.8%	17.7%	16.5%	14.9%	12.1%
Charmin	\$1.03	40.4%	18.2%	20.7%	22.3%	24.8%	28.2%
Kleenex Cottonelle	\$0.86	54.2%	9.4%	10.8%	12.1%	13.1%	13.7%
Quilted Northern	\$0.80	37.2%	9.8%	11.4%	12.6%	13.9%	14.8%
Scott	\$0.36	36.9%	11.3%	11.1%	10.7%	10.7%	10.9%
Store Brands	\$0.48	20.0%	24.1%	20.3%	18.5%	16.6%	15.8%
Other	\$0.50	20.6%	9.4%	8.0%	7.3%	6.0%	4.6%

* Average price-per-standardized roll for UPCs of given brand containing four rolls.

Table 3.8: Demographic Summary Statistics

	INC_1	INC_2	INC_3	INC_4	INC_5
Household Size					
One person	56.8%	34.8%	21.4%	12.1%	6.7%
Two people	27.7%	41.3%	45.7%	46.4%	46.3%
Three people	8.3%	11.6%	14.4%	17.2%	19.2%
Four people	4.4%	7.3%	11.2%	15.4%	18.4%
Five or more people	2.9%	5.0%	7.3%	8.9%	9.5%
Average household size	1.71	2.09	2.41	2.67	2.82
Female Household Head Education					
Did not graduate high school	7.6%	4.3%	2.3%	1.2%	0.6%
High school grad	37.5%	35.7%	28.4%	19.7%	10.8%
Some college	34.0%	34.3%	33.4%	30.8%	23.7%
College grad	17.6%	20.8%	27.4%	34.9%	40.4%
Post college grad	3.4%	4.8%	8.5%	13.4%	24.6%
Male Household Head Max Education					
Did not graduate high school	15.0%	10.3%	5.3%	2.6%	1.0%
High school grad	33.0%	36.0%	30.9%	21.4%	10.2%
Some college	29.6%	30.5%	31.6%	30.3%	21.3%
College grad	17.4%	18.8%	24.4%	32.4%	39.1%
Post college grad	5.0%	4.4%	7.8%	13.3%	28.4%
Marital Status					
Married	25.3%	49.3%	65.6%	77.4%	84.8%
Widowed	22.0%	13.6%	6.0%	3.3%	1.8%
Divorced	31.1%	20.7%	14.3%	8.8%	5.6%
Single	21.6%	16.4%	14.1%	10.5%	7.8%
Race					
White/Caucasian	85.7%	86.3%	84.9%	83.8%	81.5%
Black/African American	9.3%	8.7%	9.0%	8.9%	8.3%
Asian	0.7%	1.1%	1.9%	2.9%	5.8%
Other	4.3%	4.0%	4.2%	4.4%	4.4%
Type of Residence					
One Family House	58.8%	71.1%	79.8%	87.0%	91.0%
One Family House (Condo/Coop)	0.9%	1.1%	1.3%	1.2%	1.2%
Two Family	5.3%	4.3%	3.1%	2.1%	1.3%
Two Family House (Condo/Coop)	0.6%	0.7%	0.6%	0.5%	0.4%
Three+ Family House	18.6%	10.2%	6.9%	4.1%	2.6%
Three+ Family House (Condo/Coop)	4.0%	4.3%	4.0%	3.2%	3.0%
Mobile Home or Trailer	11.8%	8.4%	4.4%	1.8%	0.5%

Table 3.9: Household Average Daily Consumption (Sheets)*

	INC_1	INC_2	INC_3	INC_4	INC_5
Avg daily consumption, all HH	95.2	104.1	111.1	113.3	114.7
75th percentile	46.8	53.6	58.3	61.7	63.5
50th percentile	74.7	82.1	89.2	91.4	92.3
25th percentile	116.8	129.5	136.6	137.3	138.6
N (number of HH)	7,932	18,305	15,278	20,968	10,995
Avg daily consumption, single-person HH**	71.9	69.0	69.9	68.1	66.1
75th percentile	37.9	37.4	36.7	35.5	33.5
50th percentile	58.3	55.9	56.5	53.1	52.5
25th percentile	90.1	85.3	85.2	83.5	80.9
N (number of HH)	3,930	4,929	2,365	1,662	453
Avg daily consumption, multi-person HH***	121.9	119.0	119.8	117.8	117.3
75th percentile	62.5	64.0	65.4	65.3	65.4
50th percentile	95.4	95.4	96.4	95.0	94.3
25th percentile	145.3	147.3	145.7	141.3	140.7
N (number of HH)	3,388	12,438	12,515	18,878	10,358

* Data in this table excludes households that changed income groups during the panel

** For households that never had more than one person during their time in the panel

*** For households that always had at least two people during their time in the panel

Table 3.10: Purchased package size (number of sheets) differences across income groups

Regression	3.1A		3.1B		3.1C		3.1D	
	Est.	t-stat	Est.	t-stat	Est.	t-stat	Est.	t-stat
INC_2 (\$20-\$40K)	334.4	11.23	216.0	5.55	167.8	5.48	124.1	4.23
INC_3 (\$40-\$60K)	589.8	18.67	372.1	8.85	262.5	7.82	169.0	5.29
INC_4 (\$60-\$100K)	1029.4	31.31	696.2	15.44	461.2	12.59	270.5	7.94
INC_5 ($>$ \$100K)	1558.2	37.29	1104.2	20.70	720.3	16.36	391.6	9.88
$[Consumption]_h$			Yes		Yes		Yes	
$[Demographics]_{ht}$			Yes		Yes		Yes	
$[Time]_t$			Yes		Yes		Yes	
$[Product]_p$					Yes		Yes	
$[Channel]_{ht}$							Yes	
N (purchases)	2,729,285							

Table 3.11: Difference between interpurchase time (in days) preceding sale and non-sale purchases

Regression	3.2A		3.2B		3.2C		3.2D	
	Est.	t-stat	Est.	t-stat	Est.	t-stat	Est.	t-stat
Sale	-1.246	-5.96	-1.311	-6.28	-1.308	-6.29	-0.939	-4.53
$\times$ INC_2	-0.071	-0.30	0.053	0.22	0.134	0.57	0.207	0.89
$\times$ INC_3	-0.800	-3.28	-0.666	-2.74	-0.495	-2.05	-0.370	-1.54
$\times$ INC_4	-0.897	-3.66	-0.906	-3.71	-0.594	-2.45	-0.427	-1.77
$\times$ INC_5	-0.939	-3.14	-1.298	-4.37	-0.842	-2.86	-0.685	-2.34
$[Demographics]_{ht}$			Yes		Yes		Yes	
$[Time]_t$			Yes		Yes		Yes	
$[Product]_p$					Yes		Yes	
$[Channel]_{ht}$							Yes	
N (purchases)	2,556,138							

Table 3.12: Change in duration until next purchase, in response to sale

Regression	3.3A		3.3B		3.3C		3.3D	
	Est.	t-stat	Est.	t-stat	Est.	t-stat	Est.	t-stat
Sale	6.299	26.59	6.165	26.12	6.121	26.27	7.158	31.16
$\times$ INC_2	-0.954	-3.64	-0.895	-3.43	-0.643	-2.50	-0.294	-1.16
$\times$ INC_3	-2.255	-8.27	-2.170	-7.99	-1.634	-6.10	-0.989	-3.76
$\times$ INC_4	-2.963	-10.78	-2.909	-10.63	-1.978	-7.35	-1.132	-4.29
$\times$ INC_5	-3.138	-9.62	-3.148	-9.69	-1.794	-5.62	-0.911	-2.92
$[Demographics]_{ht}$			Yes		Yes		Yes	
$[Time]_t$			Yes		Yes		Yes	
$[Product]_p$					Yes		Yes	
$[Channel]_{ht}$							Yes	
N (purchases)	2,556,485							

Table 3.13: Proportion of purchases made on sale

Regression	3.4A		3.4B		3.4C		3.4D	
	Est.	t-stat	Est.	t-stat	Est.	t-stat	Est.	t-stat
Income Group 2	0.032	7.17	0.039	8.74	0.023	6.04	0.012	3.30
Income Group 3	0.056	11.72	0.063	12.99	0.038	8.79	0.020	4.96
Income Group 4	0.084	17.69	0.088	17.22	0.054	11.85	0.029	7.05
Income Group 5	0.106	18.95	0.097	15.90	0.057	10.39	0.023	4.58
$[Consumption]_h$			Yes		Yes		Yes	
$[Demographics]_{ht}$			Yes		Yes		Yes	
$[Time]_t$			Yes		Yes		Yes	
$[Product]_p$					Yes		Yes	
$[Size]_p$					Yes		Yes	
$[Channel]_{ht}$							Yes	
N (purchases)	2,729,285							

Table 3.14: Constructed inventory regressions

Average HH inventory (eq. 3.5)

	$AVGINV_h$	
	Est.	t-stat
Inc Grp 2	-3.38	-0.04
Inc Grp 3	130.78	1.51
Inc Grp 4	358.07	4.19
Inc Grp 5	814.35	8.31
$[Consumption]_h$	Yes	
N (households)	73,478	

Sale as function of inventory (eq. 3.6)

	$I[sale]_{hp}$	
	Est.	t-stat
Inventory	4.3E-07	11.37
$[Time]_t$	Yes	
$[Product]_p$	Yes	
$[Channel]_{ht}$	Yes	
$[Size]_{pt}$	Yes	
N (purchases)	2,729,285	

Table 3.15: Change in purchased package size in response to sale

	Package Size (Sheets)							
	Est.	t-stat	Est.	t-stat	Est.	t-stat	Est.	t-stat
Sale	63.6	2.84	78.1	3.49	203.1	9.80	420.3	21.68
$\times INC_2$	-119.7	-4.88	-126.4	-5.16	-82.1	-3.65	-29.2	-1.39
$\times INC_3$	-245.6	-9.24	-255.5	-9.63	-152.8	-6.34	-58.4	-2.62
$\times INC_4$	-401.8	-14.43	-420.1	-15.10	-237.5	-9.50	-112.4	-4.92
$\times INC_5$	-530.7	-15.27	-561.9	-16.17	-296.9	-9.99	-172.5	-6.46
$[Demographics]_{ht}$			Yes		Yes		Yes	
$[Time]_t$			Yes		Yes		Yes	
$[Product]_p$					Yes		Yes	
$[Channel]_{ht}$							Yes	
N (purchases)	2,729,285							

Table 3.16: Impact of relaxing liquidity constraints on package size (sheets)

	3.8A		3.8B		3.8C		3.8D	
	Est.	t-stat	Est.	t-stat	Est.	t-stat	Est.	t-stat
First Week (ψ_0)	53.1	4.39	59.8	4.94	59.9	5.39	60.0	5.65
$\times$ INC_2 (ψ_2)	-66.4	-4.52	-70.4	-4.81	-70.1	-5.23	-62.8	-4.93
$\times$ INC_3 (ψ_3)	-54.6	-3.62	-59.5	-3.95	-54.6	-3.98	-44.6	-3.44
$\times$ INC_4 (ψ_4)	-58.1	-3.86	-65.6	-4.37	-64.7	-4.76	-53.3	-4.16
$\times$ INC_5 (ψ_5)	-54.4	-2.87	-68.1	-3.59	-67.7	-4.01	-42.5	-2.71
$[Demographics]_{ht}$			Yes		Yes		Yes	
$[Time]_t$			Yes		Yes		Yes	
$[Product]_p$					Yes		Yes	
$[Channel]_{ht}$							Yes	
N (purchases)	2,729,285							

Table 3.17: Impact of relaxing liquidity constraints on purchase acceleration

	Purchase Acceleration							
	A		B		C		D	
	Est.	t-stat	Est.	t-stat	Est.	t-stat	Est.	t-stat
Sale (δ_0)	-1.038	-4.52	-1.121	-4.90	-1.103	-4.84	-0.737	-3.24
$\times$ INC_2 (δ_2)	-0.321	-1.24	-0.177	-0.68	-0.103	-0.40	-0.033	-0.13
$\times$ INC_3 (δ_3)	-0.975	-3.67	-0.807	-3.05	-0.644	-2.45	-0.523	-2.00
$\times$ INC_4 (δ_4)	-1.083	-4.05	-1.079	-4.05	-0.776	-2.93	-0.612	-2.32
$\times$ INC_5 (δ_5)	-1.125	-3.48	-1.491	-4.64	-1.047	-3.28	-0.899	-2.83
1st Week (ψ_0)	0.071	0.41	0.163	0.94	0.168	0.97	0.167	0.97
$\times$ INC_2 (ψ_2)	-0.102	-0.50	-0.066	-0.32	-0.070	-0.34	-0.064	-0.32
$\times$ INC_3 (ψ_3)	-0.129	-0.61	-0.025	-0.12	-0.029	-0.14	-0.025	-0.12
$\times$ INC_4 (ψ_4)	0.041	0.19	0.056	0.27	0.032	0.15	0.032	0.15
$\times$ INC_5 (ψ_5)	-0.050	-0.18	-0.245	-0.89	-0.277	-1.01	-0.260	-0.95
1st Week $\times$ Sale (γ_0)	-0.840	-2.41	-0.757	-2.18	-0.817	-2.36	-0.804	-2.33
$\times$ INC_2 (γ_2)	1.038	2.50	0.946	2.29	0.976	2.37	0.988	2.40
$\times$ INC_3 (γ_3)	0.744	1.79	0.549	1.33	0.579	1.40	0.592	1.43
$\times$ INC_4 (γ_4)	0.698	1.70	0.630	1.54	0.667	1.63	0.682	1.67
$\times$ INC_5 (γ_5)	0.763	1.51	0.869	1.73	0.923	1.84	0.958	1.91
$[Demographics]_{ht}$			Yes		Yes		Yes	
$[Time]_t$			Yes		Yes		Yes	
$[Product]_p$					Yes		Yes	
$[Channel]_{ht}$							Yes	
N (purchases)	2,556,138							

Table 3.18: Regression - Impact of relaxing liquidity constraints on proportion of purchases made on sale

	Buying on Sale							
	A		B		C		D	
	Est.	t-stat	Est.	t-stat	Est.	t-stat	Est.	t-stat
First Week	0.0031	2.04	0.0032	2.12	0.0023	1.55	0.0036	2.52
$\times\ INC_2$	-0.0033	-1.80	-0.0033	-1.80	-0.0026	-1.44	-0.0028	-1.63
$\times\ INC_3$	-0.0017	-0.92	-0.0018	-0.97	-0.0012	-0.67	-0.0018	-1.05
$\times\ INC_4$	-0.0031	-1.66	-0.0033	-1.77	-0.0026	-1.48	-0.0036	-2.07
$\times\ INC_5$	-0.0032	-1.47	-0.0038	-1.75	-0.0035	-1.63	-0.0049	-2.36
$[Demographics]_{ht}$			Yes		Yes		Yes	
$[Time]_t$			Yes		Yes		Yes	
$[Product]_p$					Yes		Yes	
$[Size]_p$					Yes		Yes	
$[Channel]_{ht}$							Yes	
N (purchases)	2,729,285							

Bibliography

[1] Aguiar, M., & Hurst, E. (2007). Life-cycle prices and production. American Economic Review, 97(5), 1533-1559.

[2] Ailawadi, K. L., Lutzky, C., Neslin, S.A. "Decomposition of the sales impact of promotion-induced stockpiling." Journal of Marketing Research 44.3 (2007): 450-467.

[3] Ailawadi, K. L., and Scott A. Neslin. "The effect of promotion on consumption: Buying more and consuming it faster." Journal of Marketing Research (1998): 390-398.

[4] Akbay, C., & Jones, E. (2005). Food consumption behavior of socioeconomic groups for private labels and national brands. Food Quality and Preference, 16(7), 621-631.

[5] Attanasio, O., & Frayne, C. (2006). Proceedings from EIGHTH BREAD Conference: Do the poor pay more?. Cornell University.

[6] Bawa, K., & Shoemaker, R. W. (1987). The coupon-prone consumer: some findings based on purchase behavior across product classes. The Journal of Marketing, 51, 99-110.

[7] Beatty, T. KM. (2010). Do the poor pay more for food? Evidence from the United Kingdom. American Journal of Agricultural Economics, 92(3), 608-621.

[8] Beatty, T. KM., & Tuttle, C. J. (2014). Expenditure response to increases in in- kind transfers: Evidence from the Supplemental Nutrition Assistance Program. American Journal of Agricultural Economics, 97 (2), 390-404.

[9] Bell, D. R., & Hilber, C. A. L. (2006). An empirical test of the Theory of Sales: Do household storage constraints affect consumer and store behavior?. Quantitative Marketing and Economics, 4(2), 87-117.

[10] Blattberg, R. C., Buesing, T., Peacock, P., & Sen, S. (1978). Identifying the deal prone segment. Journal of Marketing Research, 15(3), 369-377.

[11] Carvalho, L. S., Meier, S., & Wang, S. W. (2014). Poverty and economic decision making: Evidence from changes in financial resources at payday. Unpublished manuscript .

[12] Chung, C., & Myers Jr, S. (1999). Do the poor pay more for food? An analysis of grocery store availability and food price disparities. The Journal of Consumer Affairs, 33(2): 276-296.

[13] Delaney, L., & Doyle, O. (2012). Socioeconomic differences in early childhood time preferences. Journal of Economic Psychology, 33(1), 237-247.

[14] Dubé, J., Hitsch, G. J., & Rossi, P. E. (2015). Income and wealth effects on private-label demand: evidence from the great recession. (Chicago Booth Research Paper 15-18).

[15] Einav, L., Leibtag, E., & Nevo, A. (2010). Recording discrepancies in Nielsen Homescan data: Are they present and do they matter?. QME, 8(2), 207-239.

[16] Erdem, T., Imai, S., & Keane, M. (2003). Brand and quantity choice dynamics under price uncertainty. Quantitative Marketing and Economics, 1(1), 5-64.

[17] Fernandes, D., Lynch, J. G., and Netemeyer, R.G. "Financial literacy, financial education, and downstream financial behaviors." Management Science 60.8 (2014): 1861-1883.

[18] Frank, R., Douglas, S., & Polli, R. (1967). Household correlates of package-size proneness for grocery products. Journal of Marketing Research, 4(4), 381-384.

[19] Griffith, R., Leibtag, E., Leicester, A., & Nevo, A. (2009). Consumer shopping behavior: How much do consumers save?. Journal of Economic Perspectives, 23(2), 99-120.

[20] Griskevicius, V., Tybur, J., Delton, A., & Robertson, T. (2011). The influence of mortality and socioeconomic status on risk and delayed rewards: A life history theory approach. Journal of Personality and Social Psychology, 100(6), 1015-1026.

[21] Hendel, I., & Nevo, A. (2006). Measuring the implications of sales and consumer inventory behavior. Econometrica, 74(6), 1637-1673.

[22] Hendel, I., & Nevo, A. (2006). Sales and consumer inventory. The RAND Journal of Economics, 37(3), 543-561.

[23] Johnson, D. S., Parker, J. A., & Souleles, N. S. (2004). Household expenditure and the income tax rebates of 2001. (NBER Working Paper No. w10784). Retrieved from National Bureau of Economic Research: http://www.nber.org/papers/w10784.

[24] Kaufman, P. R., MacDonald, J. M., Lutz, S. M., & Smallwood, D. M. (1997). Do the poor pay more for food? Item selection and price differences affect low-income household food costs. (Economic Research Service No. 34065). Retrieved from United States Department of Agriculture: http://purl.umn.edu/34065.

[25] Kalyanam, K., & Putler, D. (1997). Incorporating demographic variables in brand choice models: An indivisible alternatives framework. Marketing Science, 16(2), 166-181.

[26] Kunreuther, H. (1973). Why the poor may pay more for food: theoretical and empirical evidence. Journal of Business, 46(3), 368-383.

[27] Leth-Petersen, S. (2010). Intertemporal consumption and credit constraints: Does total expenditure respond to an exogenous shock to credit?. The American Economic Review, 100(3), 1080-1103.

[28] Lichtenstein, D., Burton, S., & Netemeyer, R. (1997). An examination of deal proneness across sales promotion types: a consumer segmentation perspective. Journal of Retailing, 73(2), 283-297.

[29] Mani, A., Mullainathan, S., Shafir, E., & Zhao, J. (2013). Poverty impedes cognitive function. Science, 341(6149), 976-980.

[30] Misra, K., & Surico, P. (2014). Consumption, income changes, and heterogeneity: Evidence from two fiscal stimulus programs. American Economic Journal: Macroeconomics, 6(4), 84-106.

[31] Neslin, S. A., Henderson, C. & Quelch, J. (1985). Consumer promotions and the acceleration of product purchases. Marketing Science, 4(2), 147-165.

[32] Nevo, A., & Wong, A. (2014). The Elasticity of Substitution Between Time and Market Goods: Evidence from the Great Recession. (2014 Meeting Papers. No. 315). In 2014 Annual Meeting of the Society for Economic Dynamics. Toronto. Retrieved from https://ideas.repec.org/s/red/sed014.html

[33] Parker, J. A., Souleles, N. S., Johnson, D. S. & McClelland, R. (2011). Consumer spending and the economic stimulus payments of 2008. (NBER Working Paper No. w16684). Retrieved from National Bureau of Economic Research: http://www.nber.org/papers/w16684.

[34] Rao, V. (2000). Price heterogeneity and “Real” inequality: A case study of prices and poverty in rural south India. Review of Income and Wealth, 46(2), 201-211.

[35] Shapiro, M. D. & Slemrod, J. (1995). Consumer response to the timing of income: Evidence from a change in tax withholding. The American Economic Review, 85(1), 274-283.

[36] Silva-Risso, J. M., Bucklin, R. E. & Morrison, D.G. "A decision support system for planning manufacturers' sales promotion calendars." Marketing Science 18.3 (1999): 274-300.

[37] Stephens, M. (2003). "3rd of the Month": Do social security recipients smooth consumption between checks?. The American Economic Review, 93(1), 406-422.

[38] Stephens, M. (2006). Paycheque receipt and the timing of consumption. The Economic Journal, 116(513), 680-701.

[39] Stilley, K. M., Inman, J. J. & Wakefield, K.L. "Spending on the fly: Mental budgets, promotions, and spending behavior." Journal of Marketing 74.3 (2010): 34-47.

[40] Talukdar, D. (2008). Cost of being poor: retail price and consumer price search differences across inner-city and suburban neighborhoods. Journal of Consumer Research, 35(3), 457-471.

[41] Zhang, C. Y. (2013). Consumption Responses to Pay Frequency: Evidence from 'Extra' Paychecks. University of Pennsylvania.

Appendices

Appendix 2.1: Data cleaning and sampling

We exclude respondents who (a) did not consider at least one car, (b) did not answer the demographic, importance rating, or vehicle use questions included in our model, or (c) state that they considered more than one vehicle but do not list those vehicles. To ensure the proportion of consumers who considered multiple vehicles is not biased downward, we also randomly drop an appropriate number of respondents whose set size was one, so that the share of purchased vehicle classes among kept respondents is similar to that of dropped respondents (see table below).

We sample from the remaining respondents in a manner inversely proportion to their statistical weight (Maritz Research, Inc. provided us with an appropriately weighted sample, based on their proprietary weighting methodology). Each observation is assigned a random number from zero to one. That random number is divided by (WEIGHT/4) and the 750 observations with the lowest resulting numbers each quarter are kept for estimation. The sampled observations are re-weighted to ensure that the (weighted) proportion of respondents who purchased vehicle j in our data sample matches the (weighted) proportion among all respondents who considered at least one car.

Table 3.19: Proportion of purchases per class, by data group

	Dropped	Kept
C Cars	11.3%	12.7%
CD Cars	8.1%	9.0%
Non-Premium Cars	10.2%	9.5%
Premium Cars	17.2%	18.0%
Small Utilities	8.3%	9.1%
Med/Lg Utilities & Vans	24.0%	26.2%
Trucks	20.8%	15.6%

Appendix 2.2: Derivation of likelihood statement

There exists a set of alternatives $A = \{1, ..., K\}$. Consumer i has a consideration set S_i consisting of alternatives $1, \ldots, k$ (but not alternatives $k+1, \ldots, K$), where $\omega_n > \omega_{n+1}$ for all n. In other words, alternatives are sorted in decreasing order of consideration propensity. The consumer has consideration propensities ω_j for each alterantive j and also faces marginal search costs of considering the nth alternative in their set $\tilde{c}_n$. The consideration propensities are a combination of a deterministic (w_j) and random (ϵ_j) component: $\omega_j = w_j + \epsilon_j$. The probability that S_i is given by:

(1) $\Pr\left[\omega_{j\leq k} > \omega_{j>k}\ ,\ \omega_{j\leq k} > \tilde{c}_k\ ,\ \omega_{j>k} < \tilde{c}_{k+1}\right]$
Where $\omega_{j\leq k}$ refers to the consideration propensity of each alternative $j \leq k$
The probability statement from (1) can be shown to be equal to equation 2.10.

Since $\omega_{j\leq k} > \max(\tilde{c}_k, \omega_{j>k})$, and $\omega_{j>k} < \min(\tilde{c}_{k+1}, \omega_{j\leq k})$, then:

$\Pr\left[\omega_{j\leq k} > \omega_{j>k}\ ,\ \omega_{j\leq k} > \tilde{c}_k\ ,\ \omega_{j>k} < \tilde{c}_{k+1}\right]$
(2) $= \int_{m=-\infty}^{m=\infty} \Pr\left[\, \omega_{j\leq k} > \min(\tilde{c}_{k+1}, \max(\tilde{c}_k, m))\,\right] \Pr\left[\, \min(m, \tilde{c}_{k+1}) = \max_{j>k}(\omega_j)\,\right] dm$

Since $\Pr\left[\, \min(m, \tilde{c}_{k+1}) = \max_{j>k}(\omega_j)\,\right] = \frac{d}{dm} \Pr\left[\, \min(m, \tilde{c}_{k+1}) > \max_{j>k}(\omega_j)\,\right]$

(3) $= \int_{m=-\infty}^{m=\infty} \Pr\left[\, \omega_{j\leq k} > \min(\tilde{c}_{k+1}, \max(\tilde{c}_k, m))\,\right] \frac{d}{dm} \Pr\left[\, \min(m, \tilde{c}_{k+1}) > \max_{j>k}(\omega_j)\,\right] dm$

(4) $= \int_{m=-\infty}^{m=\tilde{c}_k} \Pr\left[\, \omega_{j\leq k} > \tilde{c}_k\,\right] \frac{d}{dm} \Pr\left[\, m > \max_{j>k}(\omega_j)\,\right] dm$
$+ \int_{m=\tilde{c}_k}^{m=\tilde{c}_{k+1}} \Pr\left[\, \omega_{j\leq k} > m\,\right] \frac{d}{dm} \Pr\left[\, m > \max_{j>k}(\omega_j)\,\right] dm$
$+ \int_{m=\tilde{c}_{k+1}}^{m=\infty} \Pr\left[\, \omega_{j\leq k} > \tilde{c}_{k+1}\,\right] \frac{d}{dm} \Pr\left[\, \tilde{c}_{k+1} > \max_{j>k}(\omega_j)\,\right] dm$

Since $\frac{d}{dm}$ is not a function of m from $(\tilde{c}_{k+1}, \infty)$, the third part of the integral in (4) is equal to zero.
Set $a = \sum_{j=k+1}^{j=K} \exp(w_j)$.
Since $\Pr\left[\, \omega_{j\leq k} > \tilde{c}_k\,\right] = [1 - \exp(-\exp(w_j)\exp(-\tilde{c}_k))]$,
And since $\Pr\left[\, m > \max_{j>k}(\omega_j)\,\right] = \exp(-\sum_{j=k+1}^{j=K} \exp(w_j)\exp(-m))$:

(5) $= \prod_{j=1}^{j=k} [1 - \exp(-\exp(w_j)\exp(-\tilde{c}_k))] \int_{m=-\infty}^{m=\tilde{c}_k} \frac{d}{dm} [\exp(-a\exp(-m))]\, dm$

$+\int_{m=\tilde{c}_k}^{m=\tilde{c}_{k+1}} \prod_{j=1}^{j=k} [1-\exp(-\exp(w_j)\exp(-m))] \frac{d}{dm} [\exp(-a\exp(-m))]\, dm$

Since $\exp(-\sum_{j=k+1}^{j=K} \exp(w_j)\exp(-m)) = 0$
$\int_{m=-\infty}^{m=\tilde{c}_k} \frac{d}{dm} \left[\exp(-\sum_{j=k+1}^{j=K} \exp(w_j)\exp(-m))\right] dm = \exp(-a\exp(-\tilde{c}_k)) - 0$, so:

$(6) = \prod_{j=1}^{j=k} [1-\exp(-\exp(w_j)\exp(-\tilde{c}_k))] \exp(-a\exp(-\tilde{c}_k))$

$+\int_{m=\tilde{c}_k}^{m=\tilde{c}_{k+1}} \prod_{j=1}^{j=k} [1-\exp(-\exp(w_j)\exp(-m))] \frac{d}{dm} [\exp(-a\exp(-m))]\, dm$

If we set $E_j = \exp(-\exp(w_j)\exp(-m))$
$T_t = \sum_{q=1}^{q=\frac{k!}{t!(k-t)!}} \prod_{j\in G_{qt}} E_j$,
Where G_{qt} is the q^{th} subset of alternatives $j \in S_i$, of size t, of which there exist $\frac{K!}{t!(K-t)!}$
Then:

$(7) = \prod_{j=1}^{j=k} [1-\exp(-\exp(w_j)\exp(-\tilde{c}_k))] \exp(-a\exp(-\tilde{c}_k))$

$+\int_{m=\tilde{c}_k}^{m=\tilde{c}_{k+1}} (1+\sum_{t=1}^{t=k} (-1)^t T_t) \frac{d}{dm} [\exp(-a\exp(-m))]\, dm$

$(8) = \prod_{j=1}^{j=k} [1-\exp(-\exp(w_j)\exp(-\tilde{c}_k))] \exp(-a\exp(-\tilde{c}_k))$

$+\int_{m=\tilde{c}_k}^{m=\tilde{c}_{k+1}} \frac{d}{dm} [\exp(-a\exp(-m))]\, dm$

$+\int_{m=\tilde{c}_k}^{m=\tilde{c}_{k+1}} \left[\sum_{t=1}^{t=k} (-1)^t T_t\right] \frac{d}{dm} [\exp(-a\exp(-m))]\, dm$

And, solving for the integral in the second part of (8):

$(9) = \prod_{j=1}^{j=k} [1-\exp(-\exp(w_j)\exp(-\tilde{c}_k))] \exp(-a\exp(-\tilde{c}_k))$

$+\exp(-a\exp(-\tilde{c}_{k+1})) - \exp(-a\exp(-\tilde{c}_k))$

$+\int_{m=\tilde{c}_k}^{m=\tilde{c}_{k+1}} \left[\sum_{t=1}^{t=k} (-1)^t T_t\right] \frac{d}{dm} [\exp(-a\exp(-m))]\, dm$

Set $b_{qt} = \sum_{j\in G_{qt}} \exp(w_j)$
Then $T_t = \sum_{q=1}^{q=\frac{k!}{t!(k-t)!}} \prod_{j\in G_{qt}} E_j = \sum_{q=1}^{q=\frac{k!}{t!(k-t)!}} \exp(-b_{qt}\exp(-m))$, and:

$(10) = \prod_{j=1}^{j=k} [1-\exp(-\exp(w_j)\exp(-\tilde{c}_k))] \exp(-a\exp(-\tilde{c}_k))$

$+\exp(-a\exp(-\tilde{c}_{k+1})) - \exp(-a\exp(-\tilde{c}_k))$

$+\int_{m=\tilde{c}_k}^{m=\tilde{c}_{k+1}} \left[\sum_{t=1}^{t=k} (-1)^t \sum_{q=1}^{q=\frac{k!}{t!(k-t)!}} \exp(-b_{qt}\exp(-m))\right] \frac{d}{dm} [\exp(-a\exp(-m))]\, dm$

Since $\frac{d}{dm} [\exp(-a\exp(-m))] = a\exp(-m) [\exp(-a\exp(-m))]$:

$(11) = \prod_{j=1}^{j=k} [1-\exp(-\exp(w_j)\exp(-\tilde{c}_k))] \exp(-a\exp(-\tilde{c}_k))$

$$+\exp(-a\exp(-\tilde{c}_{k+1})) - \exp(-a\exp(-\tilde{c}_k))$$

$$+\sum_{t=1}^{t=k}(-1)^t \sum_{q=1}^{q=\frac{k!}{t!(k-t)!}}(a)\int_{m=\tilde{c}_k}^{m=\tilde{c}_{k+1}} \exp(-m)\left[\exp(-(a+b_{qt})\exp(-m))\right] dm$$

And, solving for the integral:

$$(11) = \prod_{j=1}^{j=k}\left[1-\exp(-\exp(w_j)\exp(-\tilde{c}_k))\right]\exp(-a\exp(-\tilde{c}_k))$$

$$+\exp(-a\exp(-\tilde{c}_{k+1})) - \exp(-a\exp(-\tilde{c}_k))$$

$$+\sum_{t=1}^{t=k}\sum_{q=1}^{q=\frac{k!}{t!(k-t)!}}\left(\frac{(-1)^t a}{a+b_{qt}}\right)\left[\exp(-(a+b_{qt})\exp(-\tilde{c}_{k+1})) - \exp(-(a+b_{qt})\exp(-\tilde{c}_k))\right]$$

This gives us equation 2.10, though the notation in essay 2 is slightly different, as alternatives in our dataset are not number in decreasing order of consideration propensity.

Appendix 2.3: Validation of model via simulation

To demonstrate that our likelihood function can be used in an estimation routine to accurately retrieve the parameters governing the consideration set formation model developed in our paper, we include here the results of two simulation studies. First, we generate synthetic data for 30,000 consumers and 40 alternatives. Consideration propensity is a function of 40 (partially heterogeneous, as with our real data) alternative-specific constants and 10 randomly generated covariates. The marginal search costs for the first, second, and fifth alternative are (as with our real data) set to -99, 0, and 99. We simulate consideration sets and then estimate the consideration stage of our model to verify that our likelihood statement (equation 10 in the paper) can be used to accurately estimate the parameters of our model.

As a second test, we also use our data sample and the parameter estimates from our full model to simulate new consideration sets. We then re-estimate our model on these simulated consideration sets. In the interest of space, we only include the primary parameters of interest (since there are nearly 200).

Table 3.20: Parameter Estimates from Simulation 1 (synthetic data)

Param.	Actual	Est.	Param.	Actual	Est.	Param.	Actual	Est.
ASC 1	-6.000	-5.920	ASC 21	-5.000	-4.900	$\tilde{c}_3$	0.273	0.291
ASC 2	-5.950	-5.862	ASC 22	-4.950	-4.855	$\tilde{c}_4$	0.496	0.517
ASC 3	-5.900	-5.798	ASC 23	-4.900	-4.790			
ASC 4	-5.850	-5.711	ASC 24	-4.850	-4.753	$\sigma_{C,CD}$	-0.300	-0.296
ASC 5	-5.800	-5.757	ASC 25	-4.800	-4.722	σ^2_{CD}	1.250	1.231
ASC 6	-5.750	-5.618	ASC 26	-4.750	-4.604			
ASC 7	-5.700	-5.595	ASC 27	-4.700	-4.613	Covariate 1	0.664	0.654
ASC 8	-5.650	-5.536	ASC 28	-4.650	-4.549	Covariate 2	0.718	0.711
ASC 9	-5.600	-5.538	ASC 29	-4.600	-4.530	Covariate 3	0.354	0.360
ASC 10	-5.550	-5.443	ASC 30	-4.550	-4.462	Covariate 4	0.652	0.630
ASC 11	-5.500	-5.377	ASC 31	-4.500	-4.448	Covariate 5	0.211	0.209
ASC 12	-5.450	-5.396	ASC 32	-4.450	-4.364	Covariate 6	-0.688	-0.672
ASC 13	-5.400	-5.293	ASC 33	-4.400	-4.302	Covariate 7	-1.814	-1.793
ASC 14	-5.350	-5.236	ASC 34	-4.350	-4.267	Covariate 8	0.113	0.134
ASC 15	-5.300	-5.235	ASC 35	-4.300	-4.214	Covariate 9	-1.424	-1.142
ASC 16	-5.250	-5.179	ASC 36	-4.250	-4.156	Covariate 10	0.687	0.675
ASC 17	-5.200	-5.136	ASC 37	-4.200	-4.112			
ASC 18	-5.150	-5.039	ASC 38	-4.150	-4.084			
ASC 19	-5.100	-4.943	ASC 39	-4.100	-3.986			
ASC 20	-5.050	-4.971	ASC 40	-4.050	-3.995			

Table 3.21: Parameter Estimates from Simulation 1 (real data)

Parameters	Actual	Est.
Search Costs		
$\tilde{c}_3$	0.218	0.257
$\tilde{c}_4$	0.399	0.531
Heterogeneity Covariance Matrix		
$\sigma_{C,CD}$	-0.300	-0.347
σ^2_{CD}	1.252	1.210
Redesign Variables		
Hyundai Elantra	0.745	0.644
VW Jetta	0.326	0.340
Subaru Legacy	0.305	1.003
Kia Optima	1.272	1.340
Subaru Outback	0.743	0.545
Hyundai Sonata	1.037	1.117
Recall and Tsunami Variables		
Recall - Toyota C/CD Cars	-0.242	-0.337
Tsunami - Toyota C/CD Cars	-0.150	-0.057
Tsunami - Toyota B/DE Cars	-0.769	-0.481
Tsunami - Honda C/CD Cars	-0.371	-0.532
Tsunami - Honda B/DE Cars	-0.423	-0.429
Tsunami - Other Japanese C/CD	-0.098	-0.099

Appendix 2.4: Specifics of the empirical model

Alternatives modeled

We simplify the alternative space in the following ways: (1) We subsume all trims, model years, and fuel types of a nameplate into a single alternative. (2) We do not model Hyundai Veloster or Suzuki Reno, which accounted for 0.04% of purchases. (3) We condense the GLI, GTI, R32, and Rabbit (all variants of the VW Golf) into a single alternative. All four vehicles were replaced with a single VW Golf nameplate after the 2009 model year (which we treat as a separate alternative). (4) We condense all but three of the alternatives outside of the C and CD car classes into 19 vehicle class / continent of origin outside options. We have three continents of origin (Asia, Europe, North America) and seven vehicle classes (B Cars, DE Cars, Premium Cars, Premium Utilities, Small Utilities, Large Utilities / Vans, and Trucks), but do not model European B Cars or Trucks (0.02% and 0.05% of all purchases in our sample, respectively). We model the Honda Fit (B Car), Toyota Yaris (B Car), and Toyota Avalon (DE Car) independently from the outside options in order to facilitate more accurate estimates of substitution patterns in response to the recalls and tsunami, which primarily affected Toyota and Honda. In total we model 76 alternatives (54 C/CD cars and 22 other alternatives). Our decision to exclude or condense some alternatives was made for estimation speed and stability.

Observable heterogeneity (X_{ij})

The NVCS asks respondents if the vehicle they have just purchased was intended to replace another vehicle. If respondents answer "yes," they also list details about the vehicle being replaced (vehicle make, model, and class). Call the vehicle being replaced vehicle k. We include 17 interaction dummy variables equal to one if alternative j is (a) the same alternative as k, (b) made by the same manufacturer as k, (c) from the same vehicle class as k, or (d) from the same continent of origin as k. We also include variables equal to one if k is from vehicle class X and j is from another class Y in an effort to measure the likelihood of one class being replaced by another. X_{ij} also includes similar interactions for other vehicles owned but not being replaced.

A total of 38 demographic interaction terms are also included in X_{ij}. We measure how loyal various age groups are to a manufacturer, and include income group and gender $\times$ vehicle class interactions. We also include 30 importance and product use interactions. We measure whether self-reported brand loyalty influences a consumer's decision to consider or purchase from a manufacturer they have purchased before. We also measure the degree to which consumers who think the fuel used by a vehicle is important were more or less inclined to consider certain classes of vehicles. Another set of interaction terms were related to vehicle use. We interact consumers' self-reported tendency to use their vehicle to take their kids to school, tow, haul cargo, and go off-roading with vehicle class dummies.

Unavailability variable

$U_{jt}$is a nuisance variable that controls for abnormally low consideration/purchase during the first quarter or at the tail end of a vehicle's availability. This might occur if (for example) a vehicle was launched mid-quarter or if a limited supply of a discontinued vehicle remained at dealerships. We set $U_{jt} = 1$ under two conditions. First, if t was the first quarter that vehicle j was available and its market share was less than 50% of its market share in quarter $t + 1$. Second, if the market share for a vehicle in its final model year drops by more than 50% from t to $t + 1$, we set $U_{jt'} = 1 \forall t' > t$.

Appendix 2.5: Estimation algorithm

Recall that the specification of the consideration set formation level of the model is:

$$\omega_{ijt} = \alpha_{ij} + \gamma D_{jt} + \delta R_{jt} + \lambda E_{jt} + \beta X_{ij} + \eta U_{jt} + \epsilon_{ijt}$$

$$\tilde{c}_3 = exp(\theta_3) \qquad \tilde{c}_4 = \tilde{c}_3 + exp(\theta_4)$$

$$u_{ijt} = \alpha_j^c + \gamma^c D_{jt} + \delta^c R_{jt} + \lambda^c E_{jt} + \beta^c X_{ij}^c + \eta^c U_{jt} + \zeta^c P_{jt} + \varepsilon_{ijt}$$

$$\alpha_{ij} = \alpha_j + \xi_{il} \qquad \begin{pmatrix} \xi_{i,C} \\ \xi_{i,CD} \end{pmatrix} \sim MVN(0, \Sigma_\xi) \qquad \Sigma_\xi = \begin{bmatrix} \sigma_C^2 & \sigma_{C,CD} \\ \sigma_{C,CD} & \sigma_{CD}^2 \end{bmatrix}$$

Define $\Theta = \{\alpha, \beta, \gamma, \delta, \lambda, \eta, \theta\}$. We use diffuse multivariate normal priors (MVN(0,50)) for the homogeneous parameters in Θ. Additionally, we use a (conjugate) Inverse-Wishart prior $W^{-1}\left(\begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}, 2\right)$ for Σ_ξ. For each iteration of our estimation algorithm, we have three steps:

1. Draw Σ_ξ in a Gibbs step in unidentified space using the previous iteration's values for $\xi_{i,C}$ and $\xi_{i,CD}$, then post-process the draw by dividing Σ_ξ by σ_C^2.

2. Draw each heterogeneous parameter $\xi_{i,C}$ and $\xi_{i,CD}$ for each consumer in a Metropolis-Hastings step using the values for $\sigma_{C,CD}$ and σ_{CD}^2 from step one.

3. Draw each homogeneous parameter in Θ in a Metropolis-Hastings step using the most recently drawn values of all other parameters.

The restricted version of the model follows the same algorithm, except that search costs are set to zero rather than estimated. The choice stage of the model is estimated separately, using the same algorithm, but without search costs or homogeneous alternative specific constants (i.e., only step 3 of the algorithm is used for choice).

Appendix 2.6: Counterfactual simulation methodology

The data sample used for estimation is also used to run our simulations. We run a baseline simulation and counterfactual simulations using both the full and restricted model. The counterfactual simulations include "No Recall" and "No Tsunami" counterfactuals, where the effect of these events were removed by setting the associated parameters to zero. Also run were two counterfactuals for each of the six redesigned vehicles. One removes the effect of the redesign by setting the associated parameter to zero and reducing the price of the vehicle to pre-redesign levels. The second reduces price by $1,000 but does not set the recall parameter to zero.

We run each simulation 5,000 times, simulating a consideration set and calculating corresponding conditional choice probabilities for each respondent. For each iteration of the baseline simulation, consideration sets and choice probabilities are generated using one set of parameter estimates drawn from the MCMC chain from estimation. For our counterfactual simulations, we modify the parameter estimates to reflect the appropriate scenario. Consideration sets are simulated as follows. We construct each respondent's consideration propensity (ω_{ij}) by summing (1) random draws for heterogeneous preference for vehicle class, $\xi_{i,C}$ and $\xi_{i,CD}$, (2) the homogeneous portion of consideration propensity, and (3) the stochastic component, ϵ_{ijt}. Using ω_{ij} and our cost parameter estimates, $\tilde{c}_n$, we identify which alternatives a respondent considers. Lastly, using our parameter estimates for choice utility, we generate choice probabilities for each alternative in each consumer's consideration set. We weight each consideration set and all choice probabilities by the weight of the respondent for which it was generated. We calculate the number of considerations and purchases each vehicle received per 100 consumers. To identify the effect of a redesign, the Tsunami, or the Toyota Recalls, we subtract the results of the baseline simulation from the results of the counterfactual simulation.

Appendix 2.7: Secondary parameter estimates

Table 3.22: Secondary Parameter Estimates - Purchase History

Parameters	Consid (Full)		Consid (Rest)		Choice Model	
Purchase History	Est.		Est.		Est.	
Vehicle being purchased (j) shares attribute with replaced vehicle (k)						
Vehicle j & k same nameplate	1.171	*	1.193	*	-0.130	n/s
Vehicle j & k same US manuf.	0.905	*	0.912	*	0.038	n/s
Vehicle j & k same Euro manuf.	0.994	*	0.981	*	-0.015	n/s
Vehicle j & k same Japan manuf.	1.706	*	1.692	*	0.317	n/s
Vehicle j & k same Korean manuf.	1.574	*	1.575	*	0.281	n/s
Vehicle j & k both Honda	-0.926	*	-0.910	*	0.203	n/s
Vehicle j & k both Toyota	-0.992	*	-0.969	*	-0.190	n/s
Vehicle j & k same class - B Car	1.156	*	1.173	*	0.302	n/s
Vehicle j & k same class - C Car	-0.110	n/s	-0.097	n/s	0.185	n/s
Vehicle j & k same class - CD Car	0.527	*	0.489	*	-0.145	n/s
Vehicle j & k same class - DE Car	1.485	*	1.467	*	-0.132	n/s
Vehicle j & k same class - Prm Car	0.979	*	1.003	*	0.199	n/s
Vehicle j & k same class - Sm Util	0.728	*	0.726	*	0.234	n/s
Vehicle j & k same class - Util/Van	0.766	*	0.736	*	-0.045	n/s
Vehicle j & k same class - Truck	0.764	*	0.796	*	1.714	*
Vehicle j & k same cont. - Asia	0.247	*	0.254	*	0.028	n/s
Vehicle j & k same cont. - N.A.	1.007	*	1.008	*	0.063	n/s
Vehicle j & k same cont. - Euro	0.940	*	0.947	*	0.234	n/s
Vehicle being purchased (j) shares attribute with owned vehicle (k)						
Vehicle j & k same manuf.	0.796	*	0.800	*	0.245	*
Vehicle j & k same cont. - Asia	0.219	*	0.198	*	0.058	n/s
Vehicle j & k same cont. - N.A.	0.362	*	0.359	*	-0.206	n/s
Vehicle j & k same cont. - Euro	0.636	*	0.635	*	0.190	n/s
Vehicle k small car - j small car	-0.019	n/s	-0.035	n/s	-	-
Vehicle k small car - j large car	-0.695	*	-0.696	*	-	-
Vehicle k small car - j other	-0.261	*	-0.249	*	-	-
Vehicle k large car - j small car	-0.691	*	-0.694	*	-	-
Vehicle k large car - j large car	-0.441	*	-0.440	*	-	-
Vehicle k large car - j other	-0.198	*	-0.178	*	-	-
Vehicle k other - j small car	-0.147	*	-0.179	*	-	-
Vehicle k other - j large car	-0.224	*	-0.251	*	-	-
Vehicle k other - j other	0.449	*	0.449	*	-	-

Table 3.23: Secondary Parameter Estimates - Demographics

Parameters	Consid (Full)		Consid (Rest)		Choice Model	
Demographics	Est.		Est.		Est.	
Interaction: Income Group × Vehicle Class						
Inc Grp 2 × B Car	0.852	*	0.855	*	-	-
Inc Grp 2 × C Car	0.570	*	0.569	*	-	-
Inc Grp 2 × CD Car	-0.148	n/s	-0.118	n/s	-	-
Inc Grp 2 × DE Car	-0.173	n/s	-0.155	n/s	-	-
Inc Grp 2 × Prem Car	-1.479	*	-1.498	*	-	-
Inc Grp 2 × Small Util	-0.306	n/s	-0.289	n/s	-	-
Inc Grp 2 × Prem Util	-1.492	*	-1.528	*	-	-
Inc Grp 2 × Util / Van	-0.698	*	-0.719	*	-	-
Inc Grp 2 × Truck	1.097	*	1.059	*	-	-
Inc Grp 3 × B Car	0.683	*	0.690	*	-	-
Inc Grp 3 × C Car	0.431	*	0.427	*	-	-
Inc Grp 3 × CD Car	0.057	n/s	0.054	n/s	-	-
Inc Grp 3 × DE Car	0.094	n/s	0.103	n/s	-	-
Inc Grp 3 × Prem Car	-0.907	*	-0.918	*	-	-
Inc Grp 3 × Small Util	0.006	n/s	0.019	n/s	-	-
Inc Grp 3 × Prem Util	-1.540	*	-1.550	*	-	-
Inc Grp 3 × Util / Van	-0.264	*	-0.242	n/s	-	-
Inc Grp 3 × Truck	0.819	*	0.814	*	-	-
Inc Grp 4 × B Car	0.560	*	0.552	*	-	-
Inc Grp 4 × C Car	0.316	*	0.301	*	-	-
Inc Grp 4 × CD Car	0.144	*	0.135	*	-	-
Inc Grp 4 × DE Car	0.174	*	0.172	*	-	-
Inc Grp 4 × Prem Car	-0.553	*	-0.568	*	-	-
Inc Grp 4 × Small Util	0.023	n/s	0.019	n/s	-	-
Inc Grp 4 × Prem Util	-0.954	*	-0.945	*	-	-
Inc Grp 4 × Util / Van	0.119	n/s	0.120	n/s	-	-
Inc Grp 4 × Truck	0.642	*	0.614	*	-	-
Male × B Car	-0.252	*	-0.263	*	-	-
Male × C Car	-0.048	n/s	-0.052	n/s	-	-
Male × CD Car	-0.096	*	-0.093	*	-	-
Male × DE Car	0.190	*	0.207	*	-	-
Male × Prem Car	0.124	*	0.132	*	-	-
Male × Small Util	-0.450	*	-0.440	*	-	-
Male × Premium Util	-0.208	n/s	-0.202	n/s	-	-
Male × Util / Van	-0.310	*	-0.297	*	-	-
Male × Truck	1.258	*	1.215	*	-	-
Interaction: Age × Shared Manufactuer (j & k)						
$(50 > \text{Age} \geq 30)$ × same manuf.	0.099	n/s	0.133	n/s	0.318	n/s
$(\text{Age} \geq 50)$ × same manuf.	0.310	*	0.356	*	0.360	n/s

Table 3.24: Secondary Parameter Estimates - Product Use & Importance

Parameters	Consid (Full)		Consid (Rest)		Choice Model	
Interaction: Fuel Efficiency is important to consumer × Vehicle Class						
Important × B Car	0.527	*	0.513	*	-	-
Important × C Car	0.439	*	0.418	*	-	-
Important × CD Car	-0.010	n/s	-0.040	n/s	-	-
Important × DE Car	-0.893	*	-0.924	*	-	-
Important × Prem Car	-0.732	*	-0.745	*	-	-
Important × Sm Util	-0.001	n/s	-0.012	n/s	-	-
Important × Prem Util	-0.756	*	-0.780	*	-	-
Important × Util / Van	-0.738	*	-0.771	*	-	-
Important × Truck	-0.133	n/s	-0.184	n/s	-	-
Interaction: Stated Use for Vehicle× Vehicle Class						
Kids to School × Small Car	0.483	*	0.449	*	-	-
Kids to School × Large Car	-0.081	n/s	-0.118	*	-	-
Kids to School × Prem Util	-0.194	n/s	-0.224	n/s	-	-
Kids to School × Non-Pr Util	0.206	n/s	0.160	n/s	-	-
Kids to School × Truck	0.404	n/s	0.384	n/s	-	-
Towing × Small Car	-0.508	*	-0.525	*	-	-
Towing × Large Car	0.163	n/s	0.173	n/s	-	-
Towing × Prem Util	0.354	n/s	0.411	n/s	-	-
Towing × NonxPrem Util	0.529	*	0.512	*	-	-
Towing × Truck	0.981	*	0.957	*	-	-
Hauling × Small Car	0.405	*	0.363	*	-	-
Hauling × Large Car	-0.194	*	-0.225	*	-	-
Hauling × Prem Util	0.631	*	0.588	*	-	-
Hauling × NonxPrem Util	0.593	*	0.564	*	-	-
Hauling × Truck	0.400	n/s	0.310	n/s	-	-
Off-Roading × Small Car	-0.275	*	-0.304	*	-	-
Off-Roading × Large Car	0.041	n/s	0.008	n/s	-	-
Off-Roading × Prem Util	0.816	*	0.744	*	-	-
Off-Roading × Non-Prem Util	1.052	*	1.021	*	-	-
Off-Roading × Truck	1.222	*	1.248	*	-	-
Interaction: Stated Brand Loyal × Shared manufacturer (j & k)						
Loyal (4/5) × j & k same manuf.	0.447	*	0.465	*	0.535	*
Loyal (5/5) × j & k same manuf.	0.560	*	0.606	*	1.003	*

Appendix 3.1: Construction of variables

We describe in this section how we constructed three variables for our analyses: (1) Package size, (2) household consumption rate, and (3) household inventory.

We construct the package size variable, S_{htp}, using three variables in the Nielsen dataset: "size1_amount," "multi," and "upc_descr." The variable "size1_amount" indicates the number of units in the UPC (for this category, the number of rolls), while "multi" indicates "the number of units in a multi-pack." An example of a multi pack would be a UPC containing 36 rolls, coming in six separately wrapped packages of six rolls. S_{htp} is equal to the product of "size1_amount", "multi", and the number of sheets per roll contained within the UPC, which is extracted from the variable "upc_descr." In the 36-roll UPC example given, "size1_amount" would be equal to six, as would "multi."

We calculate each household's average daily consumption rate (C_h) as the sum of all its purchases (in sheets) except its last, divided by the number of days between its first and last purchase. The implicit assumption here is that the household's consumption between the day of its first purchase and the day before its last purchase is equal to the amount purchased during that time; or, equivalently, that a household's inventory level at the time of its first purchase is equal to its inventory at the time of its last purchase. Because we do not observe how long a household waited before purchasing toilet paper again following its last purchase in the data, we do not use that purchase in the construction of average daily consumption.

This consumption rate is also used to construct a household's inventory, in line with Neslin et al. (1985). The inventory variable is constructed as follows. At the time of each purchase, t, a household's inventory (INV_{ht}) is calculated as the inventory at the time of the previous purchase ($INV_{h,t-1}$) plus the volume of toilet paper (size $\times$ quantity) previously purchased ($V_{h,t-1}$) minus the household's consumption (days since previous purchase times consumption rate: $IPT_{ht} \times C_h$):

$$INV_{ht} = INV_{h,t-1} + V_{h,t-1} - IPT_{ht}C_h \tag{3.12}$$

The household's average inventory variable ($AVGINV_h$) is constructed by taking the average of a household's inventory across all days between its first and last purchase in the panel. Note that the variable INV_{ht} is linked to purchase occasions t, and not days. However, daily inventory (INV_{hd}) for each day, d, is constructed in much the same way:

$$INV_{hd} = INV_{h,d-1} + V_{h,d-1} - C_h \tag{3.13}$$

Here $V_{h,d-1}$ refers to the volume purchased on day $d-1$, which more often than not will be equal to zero. The average inventory for a household whose first and last purchases were D_h days apart, then, is given by:

$$AVGINV_h = \frac{\sum_d INV_{hd}}{D_h} \tag{3.14}$$

A well-documented limitation of constructing inventory variables is the initial conditions problem (Hendel and Nevo, 2006a). A common solution is to use the first few observations for each household to initialize an inventory level. The downside to this approach is that a portion of our data cannot be used for estimation. Rather than sacrifice data, we set a household's starting inventory at zero. Note that we have two hypotheses related to inventory: (1) That higher income households have larger inventories, and (2) that the probability of purchasing on sale is increasing in inventory. If the first hypothesis is true, then higher income households should, on average, have larger starting inventories than low income households, and our assumption of zero starting inventory will lead us to underestimate the inventory of higher income households relative to that of low income households. Consequently, our assumption is a conservative one, and will not lead us to erroneously find support for our hypothesis. With respect to the second hypothesis, the impact of inventory on probability of purchasing on sale in a within-household regression is identified not by a household's absolute level of inventory, but by the differences in household inventory from purchase to purchase. Since initializing inventory merely adds a constant term (relative to the baseline assumption of zero) to a household's inventory level across a household's purchases, our assumption will not affect the test of hypothesis two.

Note also that our inventory variable is allowed to be negative. Our constructed inventory is actually a measure of the change in inventory relative to each household's "true" initial inventory level, which is unobserved and normalized to zero. Negative inventory, then, can be interpreted as a decrease in a household's inventory level since its first purchase.

Appendix 3.2: Data cleaning

We make an effort to conservatively correct two data issues: (1) recording discrepancies for package size, and (2) households that failed to report some of their purchases.

For 48 UPCs accounting for 8,616 purchases, the "size1_amount" and/or "multi" variables appear to be miscoded, generating unreasonably large values for S_{htp}. For example, there is one UPC for which the value of both "size1_amount" and "multi" are 36, suggesting that there are 1,296 rolls in the UPC. We attempt to correct these recording discrepancies manually, identifying the "correct" value for sheets and rolls by comparing the UPC to other UPCs from the same manufacturer and with other similar features. In the example just cited, the "correct" package size was 36. We determined that for this UPC, only one of "size1_amount" and "multi" should have been 36, with the other equal to one. The price of the UPC is similar to that of other 36-roll UPCs from the same manufacturer. The results presented in the paper are for these "corrected" values of package size. As a robustness check, we also run all package-size related analyses excluding all households with a problematic observation. The results of both approaches are not materially different, and the results of the latter approach are available from the authors upon request.

The second data issue is more complicated to address. The raw homescan data indicate that some households may have failed to scan at least one purchase during their time in the panel. The interpurchase time between two purchases for a household is often so large as to be clearly inaccurate, likely due to the household forgetting to scan a purchase (or multiple) at some point in between. For example, some interpurchase times are several years in length. This missingness affects three variables in our analyses:

1. Interpurchase time: Missing purchases bias the calculated interpurchase time (time since previous purchase) for the purchase following the missing purchase. E.g., if purchases A, B, and C are made six months apart, but B is missing, the interpurchase time preceding purchase C will appear in the data to be twelve months, rather than six.

2. Consumption rate: Missing purchases bias the consumption rate for a household downward.

3. Inventory: Because inventory is calculated as a function of purchases (some of which are missing) and consumption (which is biased downward), inventory is clearly biased. But since both factors contributing to the calculation of inventory are affected, it is unclear in

which direction inventory is biased.

To fix this issue, we identify "active periods" for a household—a set of consecutive purchases over the course of which we do not believe any purchase is "missing"—and use only these "active periods" for our analyses (how we determine missingness will be discussed shortly). To illustrate this approach, we provide a simple example. Consider a household that makes twelve purchases during two active periods ($a = 1, 2$), where active period $a = 1$ was from Jan 1, 2006 to Dec 31, 2006 and active period $a = 2$ was from Jan 1, 2010 to Dec 31, 2010. Rather than treating the five-year period spanning 2006 through 2010 as a household's time-in-panel, we treat the years of 2006 and 2010 as "active periods." Consequently, the four-year gap between the two active periods is treated as "missing," and (1) does not bias upward the interpurchase time of the first observation in the second active period, which we treat as "missing" (rather than taking the four-year interpurchase time at face value), and (2) does not bias our estimates of consumption rates or inventories, which are calculated using only the active periods.

To be more precise, our approach is as follows:

1: Identify "missingness"

We first identify observations that we believe follow a missing purchase (i.e., observations with unreasonably long interpurchase times). To do this, we identify the average duration until next purchase following purchase of each package size s: $E[Duration_s]$. For example, following a purchase of a UPC containing twelve rolls of toilet paper, households wait 46.5 days before purchasing again, on average. We also identify the standard deviation of these durations: σ_s.[10] For UPCs containing twelve rolls, the standard deviation of duration until next purchase is 61.3 days. If the time between two purchases, made during trips t and $t + 1$, exceeds the mean interpurchase time for the package size purchased during trip t by more than two standard deviations, then we conclude that the household failed to report a purchase between t and $t + 1$. Mathematically, if package size s was purchased during trip t, we conclude that there is a missing observation between t and $t + 1$ if:

$$IPT_{h(t+1)p} > E[Duration_s] + 2\sigma_s \tag{3.15}$$

If we conclude that there is a missing observation between t and $t + 1$, trip t marks the end of one active period and $t + 1$ marks the beginning of a new active period. If, for a given trip t , more than one UPC was purchased, and/or if the quantity of the UPC(s) purchased

[10]Note that $E[Duration_s]$ and σ_s are calculated using trips that feature only a single package size being purchased, in quantity 1.

was greater than one (e.g. three copies of a UPC), then the reference point used to identify whether a purchase is missing is a simple sum of the means ($E[Duration_s]$) and standard deviations (σ_s) of the purchase times corresponding to the purchased sizes. For example, if two packages of size 4 and one of size 12 were purchased during trip t, we conclude that there is a missing observation between t and $t+1$ if:

$$IPT_{h(t+1)p} > 2(E[Duration_{s=4}] + 2\sigma_{s=4}) + E[Duration_{s=12}] + 2\sigma_{s=12} \tag{3.16}$$

2: Calculate a household's daily consumption rate

For households that we believe have a missing purchase, we construct consumption as the average rate of consumption over "active periods." The daily consumption rate for a household with A active periods is calculated as:

$$\frac{\sum_{a=1}^{A} \sum_{p=1}^{P_a - 1} V_{hpa}}{\sum_{a=1}^{a=A} T_a} \tag{3.17}$$

Here V_{hpa} is the volume of toilet paper for purchase p during active period a, P_a is the total number of purchases made during period a, and T_a is the time between the first and last purchase during active period a. Recall that previously, consumption was constructed by summing all volume purchased by a household except for that of the last purchase, then dividing that sum by the number of days a household was in the panel. The logic behind this was that a household's consumption rate should reflect the relationship between volume purchased and the time over which the purchased volume was consumed—and since a household's "time in panel" ends with the last purchase, that last purchase should not be included, as it clearly would not be consumed during the "time in panel." We take a similar approach here, excluding the last purchase of each active period.

3: Calculate a household's inventory

Much like we did with consumption, we calculate inventory separately for each "active" period a household has. At the start of a household's first active period, we initialize inventory at zero (as we do for households without any missing purchases). For the second active period (and beyond), we use the inventory level at the end of the previous active period as the initial level of inventory for the current active period. The implicit assumption corresponding to our calculation of consumption and inventory is that the inventory level at the start and end of a household's time in the panel is the same—mirroring the implicit assumption for households without missing purchases. An additional implicit assumption specific to households with more than one active period is that a household's consumption and purchase volume are equal in the time between active periods.

Resulting summary statistics & robustness checks

We noted previously that interpurchase times were sometimes unreasonably long. Below is a chart detailing interpurchase times following purchases of sizes $s = 4, 12, 24$ for both the raw and cleaned data. It provides the interquartile range, minimum, and maximum interpurchase times, as well as the number of interpurchase times set to "missing" post-cleaning.

Table 3.25: Duration until next purchase summary statistics

	4-Roll		12-Roll		24-Roll	
	Raw	Clean	Raw	Clean	Raw	Clean
25th pct	7	7	17	17	23	22
50th pct	15	15	30	29	41	40
75th pct	30	28	55	51	71	68
Minimum	1	1	1	1	1	1
Maximum	2,269	137	2,379	169	2,039	216
N	421,314	411,075	817,833	793,869	201,876	195,536
% "Missing"	-	2.4%	-	2.9%	-	3.1%

Cleaning the data relabels 2-3% of interpurchase times to "missing" (last row in Table 3.25), but conservatively eliminates absurd data patterns (e.g. not purchasing toilet paper for 2,269 days after buying a 4-pack). Our approach is fairly conservative. Note that for 4-roll UPCs, the maximum IPT in the cleaned data is still 137 days.

In addition to cleaning the data, we also drop households that appear to be outliers. We identify outliers along three dimensions: A household's daily per-person consumption rate, average inventory, and maximum purchase volume per trip. Summary statistics for each are provided in Table 3.26, in terms of standardized rolls.

We drop households if:

(1) Their consumption rate is greater than one standardized roll per person, per day (>99th percentile); or less than 0.01 standardized rolls per person, per day (7.75 standardized rolls per year; 1st percentile).

(2) They purchase more than 175 rolls of toilet paper in a single day at any point during their time in the panel (99th percentile).

(3) They have an average (constructed) inventory >175 or <-175 rolls. In other words, their inventory was 175 rolls higher or lower, on average, than its starting point, which we

Table 3.26: Consumption, inventory, and purchase volume summary statistics

	Consumption		Average Inventory		Max Volume	
	Raw	Clean	Raw	Clean	Raw	Clean*
25th pct	0.06	0.07	2.33	2.58	16.94	-
50th pct	0.10	0.11	9.47	8.72	26.25	-
75th pct	0.16	0.17	21.87	19.24	44.63	-
Minimum	0.00	0.00	-1352.5	-1249.9	1.24	
Maximum	23.34	36.46	1044.0	999.4	2625.44	
HH	110,853	110,853	110,853	110,853	110,853	110,853

* Note that HH max purchase volume was not affected by our data cleaning

normalize to zero. (>99th pct)

(4) They have more than three missing observations.

(5) Consumption cannot be calculated, because the few observations we have for the household are spaced around missing observations.

As a result of these criteria, 4.4% of households and 2.6% of observations are dropped (Table 3.27). Note that our results are not sensitive to dropping these households–all hypotheses are supported even if the analyses are run on all data. Table 3.28 has the results for the set of analyses from section five, using all controls, without outliers dropped. In the interest of space, not all analyses are replicated here, but they are available from the authors upon request.

Table 3.27: Dropped observations

Criteria	HH dropped		Obs dropped	
	HH	% HH	Obs	% Obs
Consumption cannot be calculated:	2,015	1.8%	5,018	0.2%
Consumption > 1 roll per person/day:	2,478	2.2%	16,817	0.6%
Consumption < 7.75 rolls per person/year:	1,118	1.0%	5,500	0.2%
HH with >3 "missing" observations:	3,069	2.7%	51,969	1.9%
HH with > 175 rolls in a single trip:	1,343	1.2%	39,901	1.4%
HH w/ avg inventory > +/-175 rolls:	2,372	2.1%	29,585	1.1%
Total observations:	112,868	-	2,802,879	-
Total "outliers" dropped:	5,006	4.4%	73,564	2.6%

Table 3.28: "First week" analyses, no dropped observations

	Package Size		Purch. Accel.		Sale	
	Est.	t-stat	Est.	t-stat	Est.	t-stat
1st Week (ψ_0)	55.3	5.66	0.187	1.23	0.0035	2.71
$\times$ INC_2 (ψ_2)	-60.4	-5.23	-0.102	-0.56	-0.0028	-1.80
$\times$ INC_3 (ψ_3)	-44.4	-3.72	0.004	0.02	-0.0018	-1.15
$\times$ INC_4 (ψ_4)	-48.7	-4.07	0.039	0.21	-0.0037	-2.33
$\times$ INC_5 (ψ_5)	-38.2	-2.59	-0.272	-1.09	-0.0047	-2.47
Sale (δ_0)	-	-	-0.745	-3.76	-	-
$\times$ INC_2 (δ_2)	-	-	-0.017	-0.08	-	-
$\times$ INC_3 (δ_3)	-	-	-0.496	-2.18	-	-
$\times$ INC_4 (δ_4)	-	-	-0.587	-2.56	-	-
$\times$ INC_5 (δ_5)	-	-	-0.926	-3.35	-	-
1st Week $\times$ Sale (γ_0)	-	-	-0.912	-2.84	-	-
$\times$ INC_2 (γ_2)	-	-	1.084	2.84	-	-
$\times$ INC_3 (γ_3)	-	-	0.644	1.68	-	-
$\times$ INC_4 (γ_4)	-	-	0.746	1.96	-	-
$\times$ INC_5 (γ_5)	-	-	1.106	2.37	-	-

Appendix 3.3: Buying on sale as a function of inventory shifters

Buying in bulk allows households to carry higher levels of inventory. This, in turn, gives them more time until their current stock of inventory runs out and they must purchase again—and more time to wait for a sale to present itself. Following this logic, we tested to see whether a household's probability of buying on sale was increasing in its inventory (section **??**). However, because inventory is unobserved, we had to impute it. As a robustness check, we now test whether a household's probability of buying on sale is increasing in two observable boosters of inventory.

Greater purchase volumes provide larger boosts to inventory, and should allow a household more time until they must purchase again. A household with any current level of inventory INV_{ht} will run out of that inventory earlier if they purchase a 4-roll UPC during the current shopping trip than they will if they purchase a 24-roll UPC. Consequently, we can indirectly test whether a household's probability of buying on sale is increasing in its inventory by testing whether a household's probability of sale is increasing in its volume purchased during the last purchase occasion ($V_{h(t-1)}$).

Similarly, since sale purchases are typically of greater volume than non-sale purchases, we can also test whether a household's probability of buying on sale is higher if its last purchase was also made on sale ($I[sale]_{h(t-1)} = 1$).

We again employ the linear probability model, regressing the binary sale variable $I[sale]_{htp}$ on $V_{h(t-1)}$ or $I[sale]_{h(t-1)} = 1$, as well as our standard controls:

$$\begin{aligned} I[sale]_{htp} = \alpha_h + \rho V_{h(t-1)} + \sum_{i=2}^{5} \beta_i I[INC = i] V_{h(t-1)} \\ + \eta_3 [Time]_t + \eta_4 [Product]_p + \eta_5 [Channel]_p + \eta_6 [Size]_p + \epsilon_{htp} \end{aligned} \quad (3.18)$$

$$I[sale]_{htp} = \alpha_h + \rho I[sale]_{h(t-1)} + \sum_{i=2}^{5} \beta_i I[INC = i] I[sale]_{h(t-1)}$$
$$+ \eta_3[Time]_t + \eta_4[Product]_p + \eta_5[Channel]_{ht} + \eta_6[Size]_p + \epsilon_{htp} \quad (3.19)$$

We find that a household's propensity to buy on sale is indeed increasing in the volume of the last purchase and higher if the previously purchased UPC was on sale ($\rho > 0$, Table 3.29). Additionally, the increase in propensity to buy on sale is highest for low income households. We find $\beta_i < 0$ for several of the higher income groups. This finding is consistent with the hypothesis that low income households carry lower inventory levels. An increase in inventory of any amount will provide a more meaningful increase in a household's ability to wait for sale the lower its inventory was prior to purchase. Since low income households carry smaller inventories than higher income households, they should benefit the most from any given increase in inventory.

To illustrate this, we'll use a simple example. Compare two households with inventories of 1 and 30 rolls of toilet paper, each of which chooses to purchase a 24-roll UPC instead of a 4-roll UPC. The first household (inventory of 1) is much more likely to observe, and be able to take advantage of, a sale before their next purchase with an inventory of 25 rolls (1 + 24) than with an inventory of 5 rolls (1 + 4). By contrast, the second household (inventory of 30) may see only a marginal increase in their probability of buying on sale during their next purchase occasion as a consequence of increasing their inventory to 54 (30 + 24) rather than 34 (30 + 4), as both an inventory of 54 and 34 are likely to last long enough for the household to observe and take advantage of a sale before having to purchase again.

Table 3.29: Purchasing on sale as a function of past purchase

IV:	Sheets (t-1)		Sheets (t-1)		Sale (t-1)		Sale (t-1)	
	Est.	t-stat	Est.	t-stat	Est.	t-stat	Est.	t-stat
Prev. Purch.	1.8E-06	7.96	1.5E-06	7.09	0.095	30.72	0.088	29.74
$\times INC_2$	-3.7E-07	-1.52	-3.0E-07	-1.26	-0.009	-2.58	-0.008	-2.50
$\times INC_3$	-4.2E-07	-1.63	-2.7E-07	-1.09	-0.014	-3.86	-0.013	-3.70
$\times INC_4$	-7.6E-07	-3.01	-5.9E-07	-2.39	-0.022	-6.22	-0.021	-6.01
$\times INC_5$	-8.5E-07	-3.01	-6.2E-07	-2.28	-0.030	-7.59	-0.028	-7.27
$[Demographics]_{ht}$	Yes		Yes		Yes		Yes	
$[Time]_t$	Yes		Yes		Yes		Yes	
$[Product]_p$	Yes		Yes		Yes		Yes	
$[Size]_p$	Yes		Yes		Yes		Yes	
$[Channel]_{ht}$			Yes				Yes	
N (purchases)	2,618,573		2,618,573		2,618,573		2,618,573	

Appendix 3.4: Changes in quantity and volume in response to sale

Regression 3.7 was also run using quantity (Q_{htp}) and purchase volume (V_{htp}; package size $\times$ quantity) as DVs. It's important to note that UPCs with fewer rolls are easier to stockpile by purchasing higher quantities (e.g., two four-roll UPCs are easier, in terms of space and cost, to stockpile than two 24-roll UPCs). The data corroborate this—when a household purchases a single-roll UPC that is not on sale, for example, the average quantity purchased is 2.391, while the average quantity of twelve-roll UPCs purchased is 1.101. Because low income households are more likely to purchase smaller sizes, we must account for UPC size in our quantity regressions. We therefore include our size controls ($[Size]_{pt}$) in the quantity regression. The results for the volume and quantity regressions in table 3.30 are consistent with those from the size regression that were previously reported (and replicated here for ease of comparison).

Table 3.30: Regression - Change in purchase quantity and volume in response to sale

	Size (Sheets)		Quantity		Volume (Sheets)	
	Estimate	t-stat	Estimate	t-stat	Estimate	t-stat
Sale	420.3	21.68	0.18	29.72	778.7	34.14
$\times$ INC_2	-29.2	-1.39	-0.01	-1.54	-12.2	-0.49
$\times$ INC_3	-58.4	-2.62	-0.02	-2.99	-44.5	-1.70
$\times$ INC_4	-112.4	-4.92	-0.03	-5.02	-101.8	-3.82
$\times$ INC_5	-172.5	-6.46	-0.04	-5.37	-150.2	-4.84
$[Demographics]_{ht}$	Yes		Yes		Yes	
$[Time]_t$	Yes		Yes		Yes	
$[Product]_p$	Yes		Yes		Yes	
$[Channel]_{ht}$	Yes		Yes		Yes	
$[Size]_{pt}$			Yes			
N (purchases)	2,729,285		2,729,285		2,729,285	

Appendix 3.5: Channel switching at beginning of month

Discount stores are one of the primary channels through which households can purchase larger sizes. Households from each income group purchases larger sizes at discount stores than at any other channel except warehouse stores (which low income households rarely shop at). 30.4% of low income households' purchases are made at discount stores during the first week of the month, compared to 28.8% during other times of the month. This indicates that low income households may be choosing to purchase at discount stores more often when they have the liquidity necessary to purchase the larger package sizes that discount stores offer. However, the summary statistics are merely a cross-sectional snapshot of purchases during the first week of the month and during other weeks—they may not reflect within-household changes. They might instead (for example) indicate that households that purchase more frequently at discount stores are also more likely to purchase during the first week of the month. To check whether households are indeed more likely to purchase at discount stores during the first week of the month, we regress a binary variable $Discount_{htp}$, equal to one if product p purchased by household h during shopping trip t was made at a discount store, on household fixed effects, the "first week" variable and its interactions with income group variables, and time controls.

$$
\begin{aligned}
Discount_{htp} = \alpha_h + \lambda I[Week1]_{ht} + \sum_{i=2}^{5} \lambda_i I[INC = i] I[Week1]_{ht} \\
+ \eta_2 [Demos]_{ht} + \eta_3 [Time]_t + \epsilon_{htp} \quad (3.20)
\end{aligned}
$$

Recall from section 3.5.1 that the channel controls included in our previous analyses implicitly assume that low income households would not change channels in response to an increase in income or liquidity. The results of this regression (Table 3.31) indicate that this is not the case—low income households are in fact more likely to purchase at discount stores during the first week of the month, when they have greater liquidity. This, in turn, indicates that the inclusion of channel controls in our analyses may lead us to underestimate the magnitude of effects under study, and that the results presented in this paper are likely to be conservative in nature.

Table 3.31: Frequency of purchasing at discount stores

	Est.	t-stat
First Week	0.0047	3.45
$\times$ INC_2	-0.0020	-1.28
$\times$ INC_3	-0.0014	-0.83
$\times$ INC_4	-0.0036	-2.25
$\times$ INC_5	-0.0033	-1.70
$[Demographics]_{ht}$	Yes	
$[Time]_t$	Yes	
N (purchases)	2,729,315	