Essays on preference formation and home production

Xu, Yan

Publication date:
2017

Document Version
Publisher's PDF, also known as Version of record

Essays on Preference Formation and Home Production

PROEFSCHRIFT

ter verkrijging van de graad van doctor aan Tilburg University op gezag van de rector magnificus, prof.dr. E.H.L. Aarts, in het openbaar te verdedigen ten overstaan van een door het college voor promoties aangewezen commissie in de Ruth First zaal van de Universiteit op dinsdag 17 oktober 2017 om 16.00 uur door

YAN XU

geboren te Mianyang, China.

PROMOTOR:	prof.dr. B.J.J.A.M. Bronnenberg
COPROMOTOR:	dr. T.J. Klein
OVERIGE COMMISSIELEDEN:	prof.dr. J.H. Abbring
	prof.dr. A.T. Ching
	prof.dr. E. Gijsbrechts
	dr. G. Knox

Acknowledgments

This dissertation contains my work as a Ph.D. student at Department of Marketing in Tilburg University. It would not have been possible without the guidance and support from my mentors and friends at Tilburg. I would like to take this opportunity to express my gratitude.

First and foremost, my profound gratitude goes to my supervisors, Bart Bronnenberg and Tobias Klein, who have been outstanding mentors and examples to me. My development during the research master program and the Ph.D. program has benefited tremendously from their guidance and advice, and I have appreciated their support far more than words can convey. They not only provided advice to my research projects, but also educated me to think of problems in a scientific way and to learn how to become an independent researcher.

Since the very beginning of my Ph.D. study, Bart encouraged me to think broadly and follow my research interest. He has been a great mentor and motivator, who truly had my interests at heart. He is also a great example of academic integrity, and his advice on my research and career has been priceless. It has been an honor to be his Ph.D. student.

The first time I learned about structural empirical work was at Tobias's econometrics class during my research master study. Without knowing it, I made my first step towards this amazing research field. Under his guidance and support, I wrote my first Matlab code of a structural model and started to interact with seminar speakers on a regular basis. I learned countless valuable knowledge and his way of thinking about applied research from the magic white board in his office. For all these I am deeply thankful.

I want to thank the members of my doctoral committee Jaap Abbring, Andrew Ching, Els Gijsbrechts and George Knox. I greatly appreciate their expertise and the time and effort they put into their committee service. Jaap, thank you for all the detailed comments, questions and suggestions. Andrew, I very much appreciate your questions and the comments on future developments of the papers. Els, I am very grateful for your constructive comments and suggestions. George, thank you for your comments and all the nice discussions we had. Apart from my committee, I am also indebted to my reference Marnik Dekimpe, for your valuable comments and support throughout my job market.

During my Ph.D. study, I was blessed with a very supportive and talented group of colleagues and friends at both Structural Econometrics Group (SEG) and Marketing Department. I am grateful for the discussions and feedback from Roxana Fernández, Yufeng Huang, Nicola Pavanini, Yifan Yu, and many others. Apart from my advisors, committee members and references, my job market preparation has also greatly benefited from the feedback from Otilia Boldea, Pavel Cizek, Hannes Datta, Jochem de Bresser, Barbara Deleersnyder, Anne-Kathrin Klesse, Aurelie Lemmens, Martin Salm, Rik Pieters, Niels van de Ven. I am also grateful for my three awesome office mates Esther, Georgi, Max and my SEG siblings Bas and Liz. Thank you for all your support and the sweets we shared.

My Ph.D. time has been extremely valuable not only from an academic point of view, but it has also enriched my life with a number of amazing friends and great colleagues: Yifan Zhang, Hailong Bao, Yifan Yu, Yufeng Huang, Bowen Luo, Yi Zhang, Yi He, Elisabeth Beusch, Roxana Fernández, Emanuel Marcu, Alaa Abi Morshed, Renata Rabovic, Ittai Shacham, Chen Sun, Bas van Heiningen, Yuxin Yao, Jonne Guyt, Soulimane Yajjou, Esther Jaspers, Max Nohe, Kristopher Keller, Bernadette Van Ewijk, Ana Martinovici, Constant Pieters, Suzanne Bies, Nick Bombay, Pranav Desai, Georgi Duev, Astrid Stubbe, Samuel Levy and many others.

Finally, I want to thank my parents and my boyfriend. To my parents, who have fullheartedly support my decision to pursue my academic career. To my boyfriend Bo, whose faithful support during all stages of this Ph.D. has been so appreciated. We graduated from college together and traveled to The Netherlands to pursue our research master and Ph.D. study together. We shared all the ups and downs. Thank you for always being there for me and for all the courage you gave me.

Yan Xu

August 24, 2017

Contents

Chapter 1

Introduction

This Ph.D. dissertation is devoted to empirically quantifying the evolution of consumer brand preference under incomplete information and the effect of time on consumer purchase behavior. The first chapter studies consumer brand preference evolution in experience goods market and investigates brands' optimal temporary price promotion decisions. The second chapter examines the role of time in determining a household's use of the market by systematically studying the effect of time on household purchase behaviors. The last chapter focuses on information spillovers and consumers learning in typical repeat purchase experience goods market.

The first chapter studies consumer brand preference evolution and its implications for brands' optimal temporary price promotion strategies. We use a new consumer packaged goods category in the Netherlands as the empirical context. The long balanced panel data are well-suited for this purpose because we observe consumers making their first purchases in a typical repeat-purchase experience goods category. We look at the empirical patterns through the lens of a learning model in which consumers make purchase decisions under uncertainty about the values they attach to brands. Their initial prior beliefs regarding the consumption utility they will experience when purchasing products in the category, together with their sensitivity to marketing variables, determine their inclination to adopt. These beliefs are updated after each purchase. In our model, consumers do not only differ with respect to their prior beliefs, but also with respect to the value they attach to the products after learning has taken place, as well as their price sensitivity. We allow all of these to be related to the time at which they first buy a product from the category. From a firm's perspective, it is important that marketing variables–promotions and product line length–affect individual utility and, thereby, the inclination to buy a product, as well as the speed at which consumers learn about their preference for the product. No less of interest to a manager is to what extent consumer learning influences the outcome of firms' strategies, such as temporary price promotion. Estimating this structural learning model allows us to characterize learning effects and to perform counterfactual simulations. In our counterfactual simulations, we investigate the long-run effect of temporary price promotion, and provide suggestions on optimal temporary price promotion timing decisions.

The second chapter examines the role of time in shaping household purchase behaviors. Becker (1965) noted that consuming goods does not only cost money, but also time. This time is, for instance, spent on planning a shopping trip, going shopping, searching and evaluating products, preparing meals, and consuming them at home. Yet, most data sets do not contain information on the availability of time to take it into account when studying consumer behavior. We combine rich scanner data from a large consumer panel, detailed data on household characteristics, and product description data. Our long panel displays ample within-household variation in hours worked in the market and provides observations of household-level events such as individuals retiring or becoming unemployed. Controlling for the fixed effects, those observations provide exogenous shocks to a household's available time for home production and its opportunity cost of time. We find that increases in the availability of time lead to more shopping trips, more overall spending on groceries, larger quantities of grocery goods, and more purchased varieties. We also find that decreases in the household's opportunity cost of time leads households to buy more goods that require larger amount of time time for preparation and consumption. This provides direct empirical evidence supporting the idea that a non-negligible part of the cost of using the market is related to time. Therefore, it is important to account for the availability of time and the opportunity cost of time when analyzing household purchase behaviors. Based on our empirical findings, we discuss the desirability of retailers' and firms' policies in reducing the household's cost of using the market.

The third chapter focuses on modeling the information spillovers in consumer learning in typical Consumer Packaged Goods markets. It spells out a learning model that allows for information spillover and formally discusses the model properties. The correlated structure allows for information spillovers to take place across products through common product attributes and puts no restriction on the direction of information spillovers. Thereafter, I also formally characterize which normalizations are needed if one applies the model to standard consumer choice panel data. I then graphically present profile-likelihood-based confidence intervals for all the parameters of interest. Lastly, I show with a Monte Carlo study, that the model, under certain regularity conditions, can be easily applied to standard consumer choice panel data.

Chapter 2

The Relationship between Customer Value and the Timing of Adoption in a New Experience Goods Category

This chapter is based on joint work with Bart Bronnenberg and Tobias Klein.

2.1 Introduction

New products take time to diffuse because different consumers start purchasing them at different points in time. The decision to start buying a product depends on beliefs about the consumption utility that can be experienced after the purchase. Importantly, this decision can be influenced by marketing activities. For instance, a lower price can stimulate a marginal consumer's brand choice decision when this consumer is pessimistic about the brand. However, the same price promotion strategy may not be optimal when uninformed consumers tend to be overly optimistic about the brand. Firms' optimal marketing strategies in expanding experience goods market critically depend on how uninformed consumers perceive the brands and how they respond to different marketing variables.

Consumers differ from one another in their initial prior beliefs and their responsiveness to marketing variables, which leads to differences in adoption timing. Subsequently, the same marketing variables determine how often consumers buy the product and thereby how fast they learn about the value they attach to actually consuming it. Among the most important questions from the perspective of a firm offering a product are how inexperienced consumers perceive the products offered, e.g., whether beliefs are initially upward or downward-biased, and how fast learning towards true preferences takes place. No less of interest to a manager is the relationship between consumers' initial prior beliefs and their long-run tastes for the products, because those are closely linked to customer value. For instance, it could be that those consumers who adopt late do so because they have downward-biased beliefs, but consume the most after they have learned about their taste for the product. Also, if beliefs are downward-biased, then marketing activities can be seen as an investment of firms into their customers, which yields returns in the long run, because of the reinforcing effect that consumers positively update their beliefs the more they buy the product. Yet another important question for a firm is whether early or late adopters will have the highest willingness to pay for the product in the long run. Finally, managers of brands may be interested in whether the order in which they entered affects the perception and subsequent learning among inexperienced consumers.

In this paper, we study consumer behavior in a new repeat-purchase experience goods category with large category expansion in the extensive margin. The balanced panel data we use is well-suited for this purpose because they allow us to observe consumers purchase behavior from the moment at which they adopt the category. We characterize learning effects and separate them from individual heterogeneity by estimating a structural learning model in which initial prior beliefs regarding the post-adoption consumption utility determine the inclination to adopt. These beliefs are updated after each purchase.

In our model, consumers do not only differ with respect to their prior beliefs, but also with respect to the value they attach to the products after learning has taken place, as well as their price sensitivity and tastes for variety (responsiveness to product line length). On top of this, promotions and product line length affect individual utility and thereby the inclination to buy a product and thus the speed at which consumers learn about their preference for the product. The value consumers attach to the product, together with their sensitivity to marketing variables, ultimately determines customer value. Estimating this structural learning model allows us to characterize learning effects and to perform counterfactual simulations, in which we provide suggestions on optimal policy decisions.

We divide consumers into cohorts according to the time at which they first purchase a product from the category as the first step to examine the underlying factors that lead to observed heterogeneous adoption timing. Our results show that there are considerable differences both across adopter cohorts. All cohorts are optimistic towards the pioneer brand but pessimistic towards the follower brand.

In our counterfactual experiments, we then show that price promotions have different effects for the pioneer brand and the follower brand. Because of their dynamic effects, they may decrease profits of the pioneer brand but increase the profit of the follower brand. More generally, this shows that characterizing learning effects and estimating consumer preferences at the same time allows a firm to improve on its dynamic price and promotion strategy.

This paper relates to the literature on product diffusion, the literature on learning, and the literature on consumer brand choice. The literature on product diffusion seeks to describe and explain how markets respond to product innovation. Hauser et al. (2006) provide a recent survey. The central finding in this literature is that a plot of sales over time in the early years of the product life-cycle is generally S-shaped (Bass, 1969). Rogers et al. (1962) define five adopter categories: innovators, early adopters, early majority, late majority and laggards. Subsequently, adoption timing has been related to individual characteristics (see for instance Raju, 1980; Joachimsthaler and Lastovicka, 1984; Baumgartner and Steenkamp, 1996). We contribute to this literature by embedding a structural model about how beliefs evolve with experience into an innovation diffusion model.

Next, our paper is related to the literature on Bayesian learning. One strand of learning applications in empirical industrial organization and marketing, such as Ackerberg (2003); Coscelli and Shum (2004); Iyengar et al. (2007); Erdem et al. (2008); Marcoul and Weninger (2008); Ching (2010a); Chintagunta et al. (2009); Ching (2010b); Ching and Ishihara (2012); Sridhar et al. (2012); Szymanowski and Gijsbrechts (2012); Chan et al.

(2013); Ching and Lim (2016), assumes consumers learn about one common parameter, i.e., brand quality, drug quality or the true stock abundance of a fishing site. This type of modeling choice is sometimes because of the need for simplicity (e.g. Marcoul and Weninger (2008)), or because of the policy implications are about firms providing information to all the consumers/agents, for example through informative detailing or advertising (e.g. Ackerberg (2003); Chan et al. (2013)).

Another strand, such as Ackerberg (2003); Crawford and Shum (2005); Narayanan and Manchanda (2009); Chintagunta et al. (2012); Shin et al. (2012); Szymanowski and Gijsbrechts (2013), focuses on consumer learning about a individual specific value. Relatedly, there is a vast literature in labor economics on issues like job matching and turnover, and marriage (e.g., Jovanovic (1984); Moscarini (2005); Marinescu (2016)), which models an individual learning about his private value like match value with a job or match value with a marriage.[1] Our model specification relates to the first strand of literature – we assume consumers who adopt during a certain time window (i.e., adopter cohort) learn about a common parameter about each brand. This specification allows us to test whether there exist significant differences across adopter cohorts and whether firms can effectively target consumers based on observed adoption timing information. We build a bridge between the empirical Bayesian learning literature and the literature on product diffusion by formulating a model and empirically relating adoption timing to demand primitives such as the mean and the variance of the initial prior beliefs, price sensitivity and long-run preferences.

Finally, this paper also relates to the literature on consumers' brand choice. This literature goes back at least to Bain (1956), who raised the question why pioneer brands have a persistent advantage in the market. Shapiro (1982) subsequently related this to consumers having "better information" about the pioneering brand. Coscelli and Shum (2004) find that the slow diffusion pattern of new drugs can be attributed to higher uncertainty faced by the patients. Ching (2010a,b) show empirical evidence which suggests "patients/physicians generally have pessimistic initial priors about generic qualities". Bronnenberg et al. (2015) document that consumers are willing to pay more for national brands. By estimating our

[1] A special case is Ackerberg (2003), which allows both learning about a common parameter and learning about individual value under certain normalization assumptions with additional advertising data.

structural learning model, we provide an alternative explanation for consumer brand choice behavior in the CPG market: consumers who adopt early have a high long-run valuation for the brand, but learning actually leads them to downward-correct their initially upward biased beliefs about the utility they will experience when consuming the product. Nevertheless, it may reinforce their inclination to buy the product because it leads them to keep buying the brand rather than trying a competing brand's product.

The remainder of the paper is structured as follows. Section 2.2 describes the data. In Section 2.3, we present model-free evidence that motivates our structural model. Section 2.4 describes our structural learning model. Section 2.5 provides details on the empirical implementation and a discussion of identification. Section 2.6 presents the estimation results. Model predictions, counterfactual experiments and implications are collected in Section 2.7. Section 2.8 concludes.

2.2 Data

2.2.1 Product category

Our analysis focuses on a new product category in the Netherlands: boxed meals. A typical product in this category contains the dried ingredients for a main dinner course that the household needs to combine with fresh meats and produce.[2] The appeal is that it saves time to prepare a meal in that way while, at the same time, providing a good consumption experience. For instance, if a family wants to make a paella dish, they can source the recipe, rice, the spices, and other ingredients separately, or they can buy most of them bundled in the correct proportions pre-packaged as a boxed meal. Boxed meals exist in many varieties and different ethnic cuisines.

We chose the boxed meal category in the Dutch market for four reasons. First, boxed meals are a typical experience good, as consumers learn about their match values with the product from consumption experiences. Second, this category has witnessed a large expan-

[2] See http://www.knorr.nl/producten/categorie/303533/2-3-persoons-wereldgerechten (accessed June 2016) for an example of boxed meal product.

sion in the extensive margin with a large group of new consumers adopting the category during our observation window.[3] For new adopters, we can also track their purchases over up to eight years. Third, the boxed meal category is a repeat-purchase good allowing us to measure the evolution in purchases after adoption. Consumers' adoption decision and subsequent behavior is voluntary and not guided by the product being a necessity. Also, consumers' shopping trips are not likely determined by boxed meal purchases. Shopping trips can thus be viewed as exogenous to purchases in this category. Last but not least, the Dutch boxed meal market is dominated by the pioneer brand, Knorr. The other brands are younger, smaller national brands, and store brands. This relatively simple market structure facilitates the set up of consumer's brand choice problem—consumers choose between the pioneer brand and a follower brand. It provides us with the opportunity to provide evidence on pioneer brand advantage.

2.2.2 Scanner data

The data used in this study are from the Dutch 2001-2008 ConsumerScan purchase panel collected by GfK and provided by Aimark. Households in this panel scan the Universal Product Code of all consumer packaged goods products that they purchase on a given trip. GfK offers panelists weekly monetary incentives to join and remain active in reporting transactions.[4]

In addition to scanning items, households also record at which retailer the product was purchased and when the purchase took place. Thus, observations in our data contain a household identifier number, the trip date, a code for the retailer, and a UPC code. The variables that are collected at the transaction level are quantities and prices paid for those quantities. Therefore the data also contain information on when a household went shopping without buying any boxed meal in a certain retailer, as long as the household purchased at least one item on the trip.

We aggregated the data at the weekly level. Consumer in the sample do not appear to choose two brands in the same week very often. If they do, we use the brand with the higher

[3] Column 1 of Table 2.3 (discussed below) lists the penetration rates over the eight years.

[4] See https://www.consumerscan.eu/about/expectations/ (accessed June 2016).

spending associated to it.

In such a long panel, panel attrition may take place. Cross-sectionally, this panel contains 5000 to 7000 households per year. For the 8 years between 2001 and 2008, we observe a large balanced panel of 2244 households. 1737 out of 2244 households in this balanced panel have made purchases of boxed meals during our observation window. We use the full cross-section data to create price and brand characteristics measures, and use a balanced panel to construct consumer adopter cohorts and estimate the learning model.

The category was created before the start of our observation period. Therefore, we cannot assess whether purchases of households in the beginning of 2001 indicate adoption or not. This left truncation is a problem that is common in learning studies, and if one wants to estimate initial priors, it's necessary to account for this (Crawford and Shum, 2005). In our data, upon category adoption, a consumer purchases boxed meals every 17 weeks on average (the median is 7 weeks). We use 26 weeks as our "filter rule" to detect adoption. That is, if we do not observe any purchases for a consumer in the first half of 2001, then we say that he adopted the category as soon as we observe his first purchase.[5]

We observe 1599 consumers who adopt the category between 2001 and 2008.[6] Our identification strategy requires us to observe consumers long enough. Therefore, we restrict our analysis to consumers who adopted the category between July 2001 and the end of 2005, so that we can track each consumer for at least 4 years after he has adopted.[7] This left us with 825 households. For the same reason, we keep only consumers who buy at least three times in the first three years after adoption. Based on the above two selection rules, our final data set contains 550 households that we observe for 416 time periods (week), which means that we can draw on $228,800$ observations for our structural estimation.

We grouped the consumers into an early cohort, a middle cohort, and a late cohort based on their adoption timing. Our choice of three cohorts is a compromise between the empirical

[5]We did robustness checks by varying the filter rule between 20 and 36 weeks. The main patterns in our model-free description of the data (see Figure 2.1, Figure 2.2, and Figure 2.3) remain unchanged.

[6]138 households are dropped because of the 26-week "filter rule".

[7]Based on the summary statistics in Table 2.1 discussed below, 4 years is sufficient for an average consumer to make about 80 purchases and is sufficient for the least frequent buyer in our panel to make about 12 purchases. For a repeated purchase goods category, it is reasonable to take the time periods around 3 years upon adoption as long run steady state.

Table 2.1: Summary statistics for estimation sample

cohort	number of households	adoption time mean	total purchase events max	mean	min
early cohort (adopt in 2001)	203	38th week, 2001	184	29.0	3
middle cohort (adopt in 2002)	216	20th week, 2002	102	20.2	3
late cohort (adopt in 2003/2004)	131	37th week, 2003	177	16.0	3

Notes: This table shows numbers of households, the average adoption timing and information on the number of purchases for our estimation sample with 550 households. The information is presented by cohort.

goal of testing across adoption timing heterogeneity and the need for enough observations in each adoption window. Table 2.1 presents summary statistics at cohort level.[8] Figure 2.2 in Section 2.3 below shows a distribution of consumers adoption timing and the thresholds of adopter cohorts.

2.2.3 Summary statistics at the brand level

The market for boxed meals is very concentrated at the brand level with the pioneer brand, Knorr (manufactured by Unilever), accounting for roughly 75% of the market share in volume and revenue (on average across the eight years). The rest of the market is covered by several national brands and store brands, which are perceived as followers compared with the pioneer brand.[9]

The Knorr brand originally entered the Netherlands in 1957 as a brand that produces soups, bouillons, and sauces, and launched the first boxed meal product in 1987. However, only recently did the category develop into a major category.

Table 2.2 presents summary statistics of the boxed meal market at the brand level over the eight years. Retailers generally sell the boxed meal category and have been doing so from the start of our data in 2001. Most of the retailers provide both brands.

[8] In Appendix 2.9.1, we provide summary statics for the sample with 825 observations.

[9] Appendix 2.9.5 provides detailed information on the follower brand which is a composite with the other national brands and the store brands.

Table 2.2: Summary statistics of the Dutch boxed meal market at the brand level

	market share (units)	market share (euros)	availability	
year	pioneer brand	pioneer brand	pioneer brand	follower brand
2001	0.880	0.871	99.8%	94.8%
2002	0.812	0.800	99.8%	93.0%
2003	0.842	0.830	99.7%	95.8%
2004	0.804	0.798	97.6%	95.1%
2005	0.678	0.685	93.5%	97.8%
2006	0.631	0.637	95.2%	99.4%
2007	0.625	0.628	97.0%	99.6%
2008	0.603	0.612	99.7%	99.3%

Notes: The statistics in this table are based on cross sectional data for all 5000-7000 households per year. The pioneer brand's market share is calculated both in terms of units and euros (first two columns). The availability measure (last two columns) is calculated as the percentage of retailers that sell a specific brand versus the total number of retailers. There are 173 unique retailers in 2001. This number decreases to 153 in 2008.

2.2.4 Measuring price and product line length

In order to analyze consumer brand choice, we need to know the prices and other product characteristics faced by the consumer on a certain shopping trip. However, as GfK ConsumerScan data is at the household level, no store-level data set of price is available. Therefore, we infer prices from other purchases made in the same retailer chain, assuming that a retailer charges common prices across outlets of the same chain. If the consumer has visited multiple retailers in a certain week but purchased no boxed meals, we take the median of the prices of the brands he could have bought. Boxed meals are mainly available in two different sizes, as "2 to 3 person meals" and "4 to 5 person meals". The weight of each package may vary with cuisine type (e.g. per portion weight of staple may vary) or meal size, but one package needs to be consumed all at once. Therefore, we choose to use price per unit rather than price per weight.

We measure product line length of a given brand by the number of the unique brand UPCs in the assortment for a given retailer and year. We do so by year and not by week because all consumers we observe in our panel may not purchase all the available UPCs in a given week and the flavors offered for each brand are altered slowly over time. Thereby,

Table 2.3: Summary statistics at the household level

		price		Number of unique UPC's	
year	penetration	pioneer brand	follower brand	pioneer brand	follower brand
2001	0.36	1.96	2.50	9.2	1.9
2002	0.52	2.04	2.28	10.3	2.8
2003	0.59	2.01	2.31	10.3	2.4
2004	0.63	1.85	2.24	11.9	2.7
2005	0.68	1.71	1.94	11.3	3.4
2006	0.72	1.77	2.03	12.8	5.7
2007	0.75	1.82	1.98	13.6	7.0
2008	0.77	1.81	1.96	16.2	6.9

Notes: Penetration is calculated as the percentage of households who purchased boxed meals in a given year. Prices are weighted averages, as we divide total revenue by the total number of units sold. We use the balanced panel with 2244 households to calculate penetration and all available data to construct price and variety measures.

we can capture the observed trend in variety over the years.

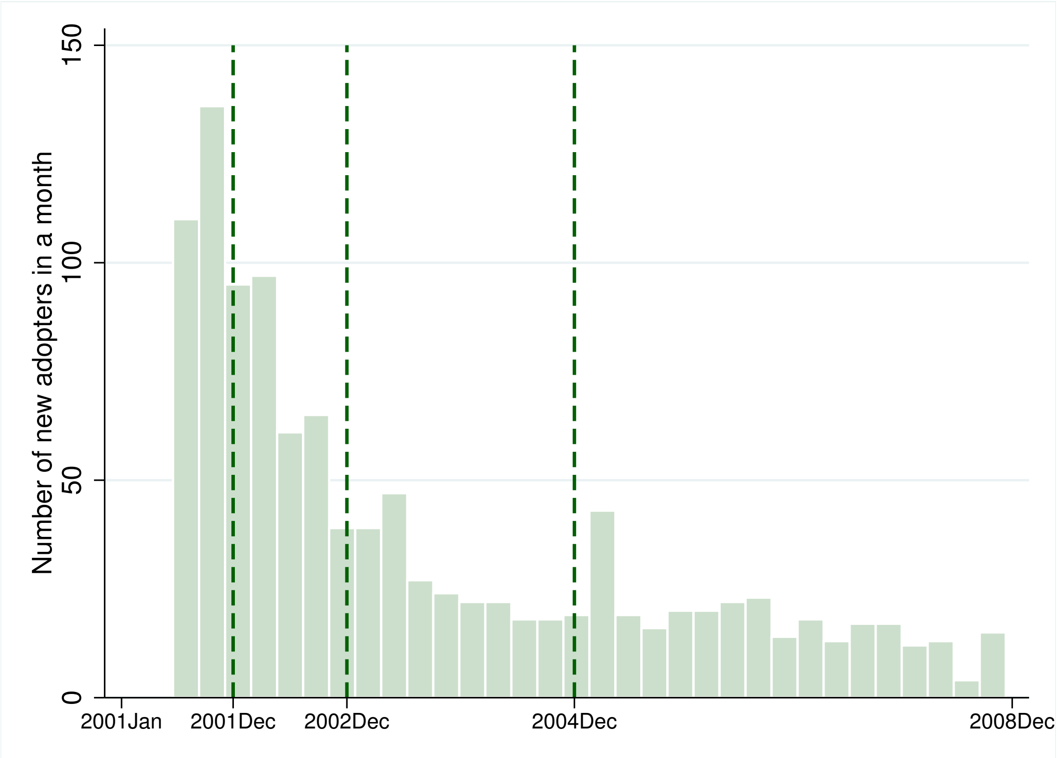
Table 2.3 reports summary statistics for the measures constructed using our household level data. Over the eight years, we see considerable growth in the extensive margin—36% of the consumers buy boxed meals in 2001 while 77% of the consumers buy them by 2008. The average transaction prices of both brands are decreasing over time.[10] Product line length as measured by the available variety increase over time and at each point in time, the pioneer brand offers more variety. The product line faced by an average consumer also increases over time.

2.3 Model-free evidence

In this section, we present the evolution of sales of the pioneer and the follower brand in the boxed meal category within and across these cohorts. We also use our individual-level choice data to report on model-free evidence for permanent taste heterogeneity and learning.

[10]The pioneer brand maintains regular price promotions, which leads to a lower average transaction price compared with the follower brand. The follower brand clusters the other national brands and the store brands, which rarely do price promotions or only mild price discounts. The decreasing pattern in the follower brand's price sequence is mainly driven by store brand entry.

Figure 2.1: Extensive margin expansion



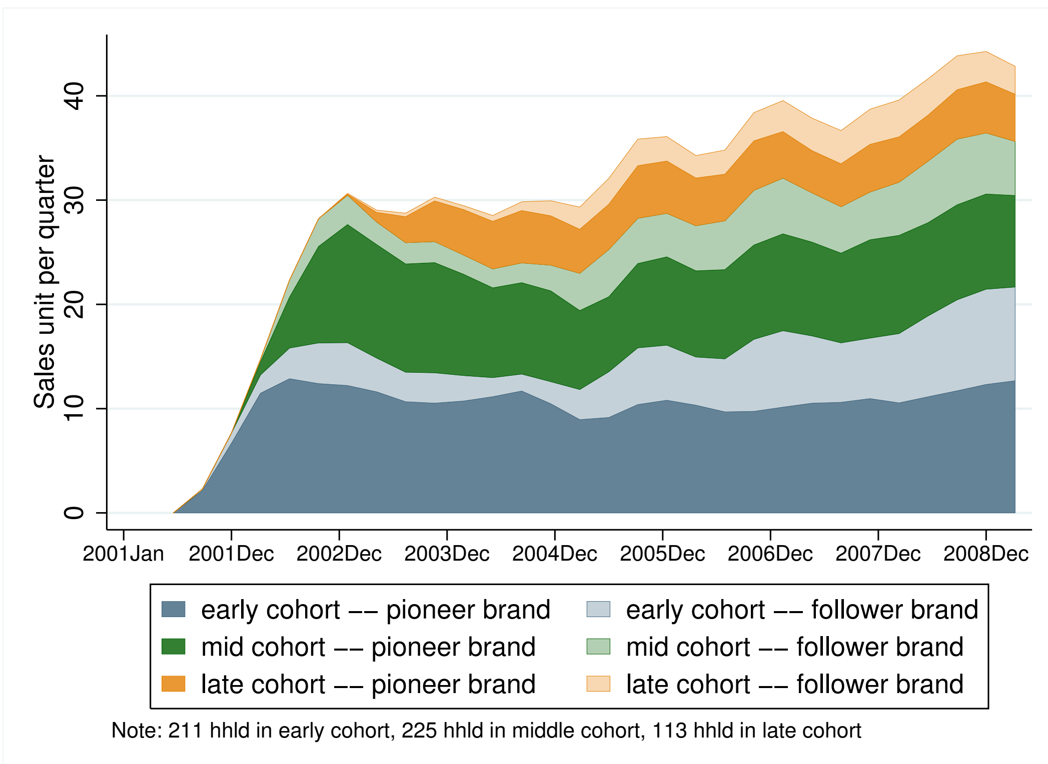
Notes: We grouped consumers who adopted the category between week 27 of 2001 and 2004 into three adopter cohorts and refer to them as early, middle and late cohort. Each cohort has a similar number of households.

2.3.1 Market expansion and the evolution of brand sales

Figure 2.1 describes the distribution of category adoption timing for each of the 825 households in our balanced panel. As defined above, we use three adoption segments, or cohorts, based on the timing of adopting the category being either in 2001 (early cohort), 2002 (middle cohort), or 2003-2004 (late cohort). Together these three cohorts make for 66.7% of the panelists who adopt the category between week 27 of 2001 and the end of 2008. This means that we can track each of these 550 consumers for at least 4 years.

Figure 2.2 shows the evolution of the brand sales based on our estimation sample of three cohorts. Plotting the sales per brand and cohort, we see three main patterns. First, in the short run, sales of the pioneer brand in any given cohort falls after adoption. Furthermore,

Figure 2.2: Growth in sales



Notes: This figure plots the total number of units sold in our sample, over time and by cohort, brand and week.

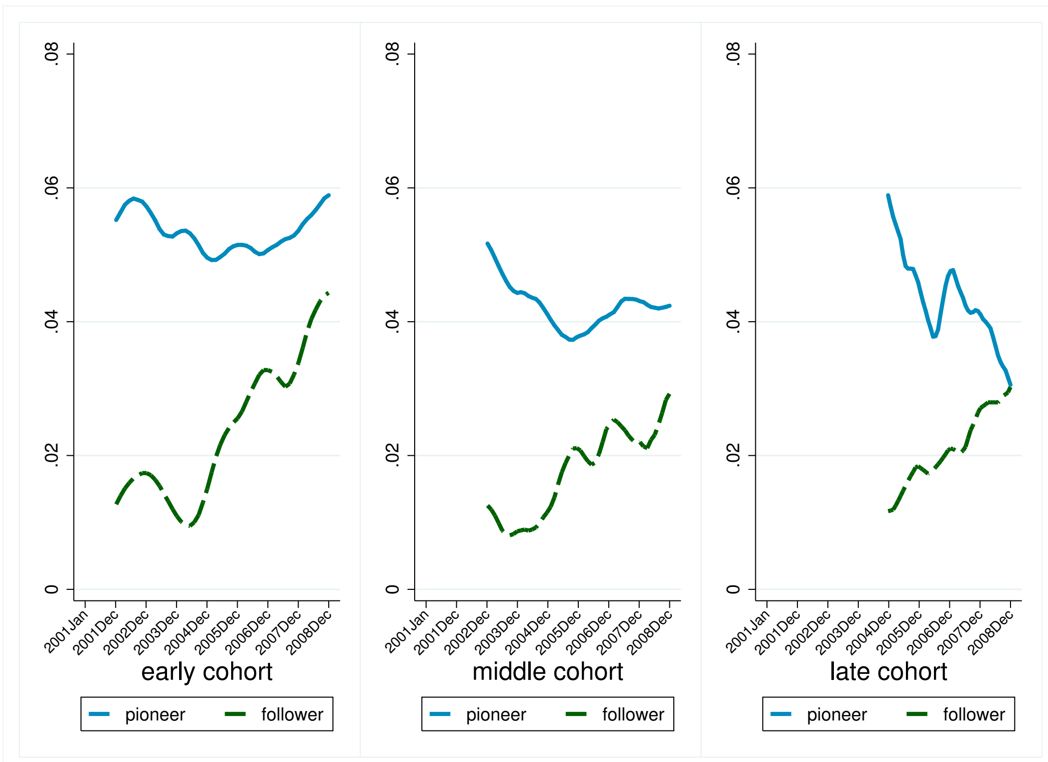
for the first two cohorts, even total sales in the category appear to fall after initial adoption and trial. Second, and in contrast, sales of the follower brand steadily increases over time. Third, sales of both brands are more stable in the long run than in the short run. In our model, we allow for these contrasting short run dynamics and subsequent stability to be the outcome of consumer learning about the true match value of these new brands.

Next, in Figure 2.3, we plot the unconditional purchase shares of both brands in any given week and for each cohort. We call attention to three features of these plots. First, the unconditional shares of the pioneer brand among all the adopter cohorts decrease over time, while the market shares of the follower brand increase over time. Second, the average market shares differ across cohorts, with the early cohort having higher purchase incidence than the middle and late cohort. Third, the rates with which the shares of the two brands changes also differ across cohorts, with the late cohorts changing more quickly than the early cohort.

These patterns are consistent with consumer behavior that displays both learning about the brands in a new category and permanent taste heterogeneity. Consumers might be optimistic about the pioneer brand and pessimistic about the follower brand at the initial stage. The observed market shares evolution could be explained if consumption experience makes consumers downward adjust their expectations about the pioneer brand and upward adjust their expectations about the follower brand. Different adopter cohorts may have different initial beliefs, so that their learning outcomes differ. This may lead to the observed heterogeneous market share evolution. Consumers may have different permanent brand match values, so that the long run market share distribution differs across cohorts.

However, consumers may also have different tastes for marketing variables, like price and product line length. Moreover, firms use time-varying price promotion strategies and expand their product lines at different rates. These factors constitute rival explanations for the pattern we see in Figure 2.3. To isolate effects that are due to consumer learning from those due to changing marketing variables, we specify a structural model with cohort-specific information priors and account for the concurrent cohort-specific responses to marketing investments for permanent taste heterogeneity.

Figure 2.3: Purchase shares by cohort and time



Notes: Figure 2.3 is a plot of a local polynomial smooth of brand choice indicators against calendar time, separately for each cohort. We do not use each consumer's first purchase incident, because the timing of a consumer's initial purchase is already used to define consumer cohorts. If a consumer made a trip to the supermarket but did not purchase any boxed meal, then we code this as him choosing the outside option. If a consumer had no supermarket visit in a given week, then we treated this as a missing observation.

2.4 The structural learning model

2.4.1 Brand choice decisions

The model introduced below is a Bayesian learning model of brand choice. Consumers are assumed to have heterogeneous valuations (i.e., match value) for each brand. They learn about those valuations over time, through consuming the products, and differ in their price sensitivity and taste for brand characteristics, such as product line length.

Consumers base purchase decisions on the current expected utility, i.e. their objective is to choose d_{ijt} to maximize the current period expected utility,

$$\mathbb{E}\left[\sum_{j\in\{0,\,1,\,2\}} u_{ijt} d_{ijt} \,|\, I_{ijt}\right]. \tag{2.1}$$

u_{ijt} is the consumer's consumption utility from consuming product j at time t ($j = 0$ denotes for the outside option); $d_{ijt} = 1$ indicates that alternative j is chosen by individual i at time t; and $d_{ijt} = 0$ indicates otherwise. We assume that $\sum_j d_{ijt} = 1$, such that consumers choose one option in each period. We also use the convention that the pioneer brand is denoted by the index $j = 1$ and the follower brand by the index $j = 2$.

The timing of a consumer's decision and information arrival is as follows. In the beginning of each period, when in the store, the consumer forms an expectation about the consumption utility for each brand, based on his prior beliefs (which is last period's posterior). The consumer next makes a purchase decision. If a consumer chooses to purchase from the category, consumption will result in a consumption experience signal for the purchased brand by the end of that period. The consumer then updates his beliefs about the brand purchased. Importantly, we allow beliefs of a novice consumer to be biased.

In the following subsections, we first present the specification of a consumer's consumption utility after purchase and his belief updating process. Then we describe the consumers' maximization problem in more detail.

2.4.2 Consumption utility specification

The utility for consumer i who consumes brand j at time t is given by the following expression:

$$u_{ijt} = \underbrace{q_{ijt} + \lambda_{ij}}_{\text{experienced match value}} + \alpha_i p_{ijt} + \omega_i x_{ijt} + \varepsilon_{ijt}, \tag{2.2}$$

where p_{ijt} is the price for brand j at time t and x_{ijt} is product line length. Further, α_i measures consumer i's price sensitivity, and ω_i measures consumer i's taste for product line length. Consumers can decide to buy neither brand and collect the utility of the outside good u_{i0t}. We normalize the constant in this utility to zero, and thus

$$u_{i0t} = \varepsilon_{i0t}. \tag{2.3}$$

ε_{i1t}, ε_{i2t} and ε_{i0t} are shocks known to consumers but unobserved to the analyst. These shocks are assumed to be drawn from a type 1 extreme value distribution, independently across consumers, brands, and time periods.

The permanent taste shock λ_{ij} is a normally distributed random coefficient that is normalized to have a mean of zero. It captures persistent unobserved differences in consumers' preferences for other brand characteristics that are observed to the consumers ex-ante, e.g., the overall package design of a brand's product line.

The match value q_{ijt} is the consumption experience that a consumer i receives when consuming brand j at period t. This consumption experience q_{ijt} is not observed by consumers when making a purchase. Instead, the consumer forms an expectation about q_{ijt} from the observed past consumption signals q_{ij1}, ..., q_{ijt-1}. Since the boxed meal category is a low stake environment, we assume that consumers are risk neutral.[11] Therefore, in Equation

[11] With choice panel data exclusively it is infeasible to separately identify the risk aversion parameter separately from the prior means ((Coscelli and Shum, 2004)). By assuming risk-neutrality one could identify the mean of initial beliefs with observed consumer purchase decisions. An alternative modeling approach is assuming consumers have rational expectation but are risk averse. We choose to assume risk neutral consumers mainly because the boxed meal category is a low stake environment. It is more logical to assume consumers are risk neutral towards a low stake product category but have biased perceptions about their true match values when they are inexperienced. This assumption is also consistent with the expected utility theory. The expected utility theory implies that people are approximately risk neutral when stakes are small (Arrow, 1971; Rabin, 2000). Bombardini and Trebbi (2012) shows, with designed experiment, that individuals are practically risk neutral at small stakes and risk averse at large

(2.2) q_{ijt} takes a linear form.

Thus, our model includes both a consumer's time invariant brand preferences λ_{ij} and brand tastes q_{ijt} that evolve from personal consumption experiences. We now provide a model of how experience and learning effects take place.

2.4.3 A Bayesian model of learning

We assume that consumers learn from past match value signals $q_{ijt'}$, $t' < t$, by combining new information into their best estimate of the true match value using Bayesian updating.

Let's assume that before a consumer i first purchases brand j, his initial belief of the match value is given by

$$I_{ij0} = \mathscr{N}\left(\mu_{ij0},\ \sigma_{ij0}^2\right) \tag{2.4}$$

in which μ_{ij0} denotes the prior mean of consumer i's initial belief of the match value for brand j, which may not equal to the true match value, μ_{ij}. A consumer i's initial beliefs may come from word-of-mouth information he has received prior to the initial purchase. This type of information could be "misleading" compared with the consumer's true match value, namely the means of the initial beliefs can be different from the true match values. The accuracy of those prior beliefs are measured by standard deviations, σ_{ij0}^2. The parameters of the initial distribution μ_{ij0} and σ_{ij0} are assumed to be known to consumers, but not to the researchers. In each subsequent period, the consumer will receive a signal q_{ijt} of the true match value μ_{ij}, if and only if he makes a purchase of brand j in period t. The consumption experience signal q_{ijt} is assumed to be unbiased, but noisy, and follows a normal distribution with mean μ_{ij} and variance σ_ν^2. Further, the signals are independent and distributed normally across periods and individuals, i.e.,

$$q_{ijt} = \mu_{ij} + \nu_{ijt}; \quad \nu_{ijt} \sim \mathscr{N}\left(0,\ \sigma_\nu^2\right). \tag{2.5}$$

The noisy consumption signals reflect the possibility that "consumers can randomly get

---

stakes. The average transaction price for one unit of boxed meal product is about €1.96, while the total household grocery expenditure that was scanned in a given year is €2843 on average.

lemons or windfalls" (Erdem and Keane, 1996).

After receiving a consumption signal q_{ijt}, consumer i updates his beliefs about j. Following the standard rules for Bayesian updating (e.g. DeGroot, 1970) for the conjugate pair of Normal distributions with a Normal prior, the following recursions for the expectation and the variance of the match value given consumption experiences from choices d_{ijt-1} are obtained:

$$\mu_{ijt} = \begin{cases} \left(\frac{1}{\sigma^2_{ijt-1}} + \frac{1}{\sigma^2_\nu}\right)^{-1} \left(\frac{1}{\sigma^2_{ijt-1}}\mu_{ijt-1} + \frac{1}{\sigma^2_\nu} q_{ijt-1}\right) & \text{if } d_{it-1} = j \\ \mu_{ijt-1} & \text{if } d_{it-1} \neq j \end{cases} \tag{2.6}$$

and

$$\sigma^2_{ijt} = \begin{cases} \left(\frac{1}{\sigma^2_{ijt-1}} + \frac{1}{\sigma^2_\nu}\right)^{-1} & \text{if } d_{it-1} = j \\ \sigma^2_{ijt-1} & \text{if } d_{it-1} \neq j \end{cases}. \tag{2.7}$$

From the above updating equations, we see that the uncertainty about the true match value σ^2_{ijt} diminishes from consumption as long as the signal variance σ^2_ν is finite. At the same time, given consumption, the expected match value μ_{ijt} is a weighted average of the previous expected match value μ_{ijt-1} and the most recent consumption signal q_{ijt-1}. The analyst does not observe the consumption signals q_{ijt}. Therefore one dimension of unobserved heterogeneity comes from the learning process itself, as a consumer's previous draws of q_{ijt} is his private information. Even when two consumers hold the same initial belief and have the same permanent tastes, their choice evolution paths may be different from different consumption experience signals they receive. In our model estimation, we account for this dimension of unobserved heterogeneity by integrating a large dimension integral.

2.4.4 Choice probabilities

The consumer makes a choice based on his expected utility given his prior information, before observing q_{ijt}. A consumer i's expected utility of brand j is

$$\begin{aligned} E(u_{ijt}|I_{ijt}) &= E(q_{ijt}|I_{ijt}) + \lambda_{ij} + \omega_i x_{ijt} + \alpha_i p_{ijt} + \varepsilon_{ijt} \\ &= \mu_{ijt} + \lambda_{ij} + \omega_i x_{ijt} + \alpha_i p_{ijt} + \varepsilon_{ijt} \end{aligned} \tag{2.8}$$

In the short run, a consumer's experiences influence his purchase decision of one brand through changing μ_{ijt}. In the long run, as the consumer accumulates experiences with brand j, the expected match value μ_{ijt} changes from μ_{ij0} for a novice consumer i to μ_{ij} for an experienced one.

Now, if the consumer is initially optimistic about a brand j, $\mu_{ij0} > \mu_{ij}$, then that consumer will initially buy more from the category and from that brand in the short run than in the long run. Furthermore, because the consumption signals are on average lower than the initial beliefs, purchasing and consuming the brand will lead to a purchase propensity of that brand that is lowered to meet the true match value. In the opposite case, $\mu_{ij0} < \mu_{ij}$, the consumer is initially too pessimistic about brand j and purchasing and consumption leads to upwardly adjusted expectations and ultimately a higher purchase propensity.

Given our assumptions for ε_{ijt} and ε_{i0t} as being drawn from the type 1 extreme value distribution, the probability that consumer i chooses brand j in period t takes a logit form:

$$\text{Prob}(d_{it} = j) = \frac{\exp(\mu_{ijt} + \lambda_{ij} + \omega_i x_{ijt} + \alpha_i p_{ijt})}{1 + \sum_{j=1,2} \exp(\mu_{ijt} + \lambda_{ij} + \omega_i x_{ijt} + \alpha_i p_{ijt})}. \tag{2.9}$$

2.5 Implementation

2.5.1 The empirical model

The goal of our empirical analysis is to characterize the evolution of brand preferences and test whether there exits significant differences between consumers who adopt at different points of time, controlling for differences in permanent taste; and to measure how the response to marketing activities differs across adoption timing. Guided by the empirical goal, we grouped the 550 households who adopted the category between 2001 and 2004 into three cohorts.[12]

Based on this, we specify our empirical model. For cohorts $c = 1,2,3$ we model the true match value (the intercept of the random utility) to be normally distributed with cohort-specific parameters, $\lambda_{ij} \sim \mathbb{N}\left(0, \sigma^2_{\lambda_{cj}}\right)$. Mean ($\mu_{ij0} = \mu_{cj0}$) and standard deviation ($\sigma_{ij0} = \sigma_{cj0}$) of the initial prior belief and long-run beliefs ($\mu_{ij} = \mu_{cj}$) are cohort-specific. The price coefficient is assumed to be normally distributed with cohort-specific mean α_c and variance $\sigma_{a_1} = \sigma_{a_2} = \sigma_{a_3}$ that is the same across cohorts. Also the taste for product line length (ω_c) is allowed to differ across cohorts.

2.5.2 Identification

We first briefly consider the variations in the data that identify the parameters we are interested in. To recap, the parameters we are interested in are: (1) μ_{cj0}, a cohort-specific mean of the consumer's initial belief of brand j; (2) σ^2_{cj0}, a cohort-specific variance of the consumer's initial belief of brand j; (3) μ_{cj}, cohort brand specific true match values; (4) the variance, $\sigma_{\lambda_{cj}}$, of the cohort-specific distribution of the consumers' unobserved brand taste, λ_{ij} ; (5) the mean, μ_{α_c}, and variance, σ_{α_c}, of cohort-specific distribution of price sensitivity, α_i; (6) cohort-specific coefficient of observed time trend—product line length, ω_c.

To identify the parameters that determine the well informed or experienced consumer's choice behaviors (μ_{cj}, $\sigma_{\lambda_{cj}}$, μ_{α_c}, σ_{α_c} and ω_c), the best data source is long term purchase

[12] We choose the cut-off point between cohorts so that all the cohorts have similar numbers of households.

data that cover later stages of the consumer learning cycle. This is because in the long run, the true match values, μ_{cj}, are revealed after accumulating sufficient experiences. Both the variance of the unobserved component of consumers' known taste for each brand, $\sigma_{\lambda_{cj}}$, and the true match value, μ_{cj}, are identified with consumers' long run purchase patterns. Price coefficient distribution parameters, α_c and σ_{α_c}, as well as consumer's taste for product line length, ω_c, are identified by the variation in observed price and product line length respectively.

Now we discuss the identification of the mean and variance of consumers' initial beliefs. The identification primarily comes from how a consumer's purchase behavior changes over time net of price changes, product line expansion over time, and the market level time trend. If there is no learning, a consumer's purchase patterns over time are fully explained by price, product line, and brand level common time trends. With learning, the choice patterns of one cohort depend on the initial beliefs the consumers in this cohort hold. Given consumers' true match values (identified from long run data), the purchase propensity of consumers in a specific cohort and the speed with which they adjust their beliefs to the true value identify the mean and variance of the cohort-specific consumer's initial beliefs about the brands. Intuitively, from the difference between the purchase propensity of the initial period and the long run periods, we can infer the mean of initial prior belief of a risk neutral consumer. Given the level of initial belief, we can infer the learning speed, namely the ratio of initial prior variance and the signal variance. Below we set the signal variance to a known constant and estimate the cohort-specific initial prior variance.

2.5.3 Estimation

The primary complication in estimation is consumer heterogeneity (λ_{ij}, α_i) and the realizations of the non-deterministic part (ν_{ijt}) in the consumption experience signals (q_{ijt}) after each purchase occasion. Those non-deterministic part in the consumption experience signals are observed by the consumers but not by the researcher. This implies that the likelihood function for a given sequence of purchase decisions for a given consumer involves a multivariate integral over the distribution of the non-deterministic part in the consumption signals and the unobserved heterogeneities. We use simulation techniques to evaluate these integrals and estimate our model using simulated maximum likelihood, where the likelihood contributions are at the individual level and given by the probability to observe the entire sequence of choices.

From the model discussion above, the probability a consumer chooses brand j depends on his preference, prior belief, prices, and product line length of both brands,

$$\text{Prob}\left(d_{it}=j|z_{it},\ I_{it},\ \kappa_i;\ \theta\right)=\text{Prob}\left(d_{it}=j|d_{it-1},\ z_{it},\ \nu_{i,t-1},\ \kappa_i;\ \theta\right), \tag{2.10}$$

where d_{it} is consumer i's observed choice in period t; $z_{it}=(z_{i1},\ ...,\ z_{it})$, where $z_{it}=\{p_{it},\ x_{it}\}$, the observed prices and product line lengths of two brands in period t; I_{it} is consumer i's prior belief at time t; $\theta\in\Theta$ denotes the parameters we want to estimate. The observed choice probability is equal to the model prediction given the set of parameters θ. Further, we define $\nu_{i,t-1}=(\nu_{i,1},\ \nu_{i,2},\ ...,\ \nu_{i,t-1})$, where $\nu_{i,t-1}=(\nu_{i1,t-1},\ ...,\ \nu_{iJ,t-1})$, and $d_{it-1}=(d_{i1},\ d_{i2},\ ...,\ d_{it-1})$. The vector κ_i contains the random effects. The elements of vector κ_i are drawn (identically across all consumers i) from the normal distribution with mean zero and variance $\sigma_{\lambda_{cj}}$ and σ_{α_c}, respectively.

The researchers cannot observe the unobserved heterogeneities κ_i and the realizations of the non-deterministic part ν_{it}. This implies that the likelihood function for a given sequence of consumption frequencies for a given consumer involves a multivariate integral over the distribution of the unobserved signals and the random coefficients:

$$\mathscr{L}_i\left(\theta|d_{iT_i},\ z_i\right)=\int\left(\Pi_{t=1}^{T_i}\text{Prob}\left(d_{it}|d_{it-1},\ z_{it},\ \nu_{i,t-1},\ \kappa_i;\ \theta\right)dF\left(\nu_{i,t-1},\ \kappa_i\right)\right), \tag{2.11}$$

where $z_i=(z_{i1},\ ...,\ z_{iT_i})$. We use simulation techniques to evaluate these integrals, and estimate our model using simulated maximum likelihood. We first draw S vectors of the unobservables, $(\nu_{i,t-1},\ \kappa_i)$. Secondly, for each set of draws $\left(\nu^s_{i,t-1},\ \kappa^s_i\right)$, we calculate the posterior mean and variance and the implied logit probabilities. Finally, we average the calculated logit likelihood function (one for each drawn sequence) over all the drawn sequences. In the results reported in this paper, we used $S = 40$:

$$\mathscr{L}_i\left(\theta | d_{iT_i},\ z_i\right) = \frac{1}{S}\sum_{s=1}^{S}\left[\left(\Pi_{t=0}^{T_i}\text{Prob}^s\left\{d_{it}|d_{it-1},\ z_{it}\right\}\right) | \left(\nu_{i,t-1}^s,\ \kappa_i^s;\ \theta\right)\right], \tag{2.12}$$

where s denotes the sth drawn vector of unobservables for consumer i, and Prob^s is the choice probability for consumer i and brand j during period t for the sth drawn.

2.6 Estimation results

Table 2.4 reports estimates of the parameters associated with learning. These parameters are the cohort (c)-brand (j) specific means of the initial beliefs about the match value (μ_{cj0}), the cohort-brand specific variance of the initial belief (σ_{cj0}), and the actual cohort-brand specific match value consumers learn about (μ_{cj}). Recall that we allow the price and variety coefficients to differ across cohorts. In order to interpret differences in intercepts across cohorts, we therefore demean price and variety measures. Recall also that we have normalized the value of the outside option to be zero.

Turning to the estimated values, we find that the mean initial beliefs for the pioneer brand are higher than the true match values for all cohorts (that is, $\mu_{c10} > \mu_{c1}$, for $c = 1, 3$ although not significantly so for $c = 2$). This means that consumers have higher preferences for the pioneer brand when they are novices than when they have accumulated experience from repeated consumption. In contrast, the initial beliefs about match values for the follower brand are lower than the true match values across all cohorts (that is, $\mu_{c20} < \mu_{c2}$, $c = 1, 2, 3$). Thus, for the follower brand, experienced consumers update positively.

While consumers in each cohort are initially optimistic about the pioneer brand in this category and pessimistic about the follower brand, the gap is reduced as they gain more experience. These effects are consistent with the idea that novice consumers perceive pioneer brands to be better than follower brands (Alpert and Kamins (1995)), even if the brands themselves are preferred similarly by experienced consumers (Golder and Tellis, 1993; Kerin et al., 1992). We believe that our empirical account of this process is novel, and leads to interesting implications of price and promotion induced consumption experiences

Table 2.4: Estimates (part 1 of 2): learning parameters

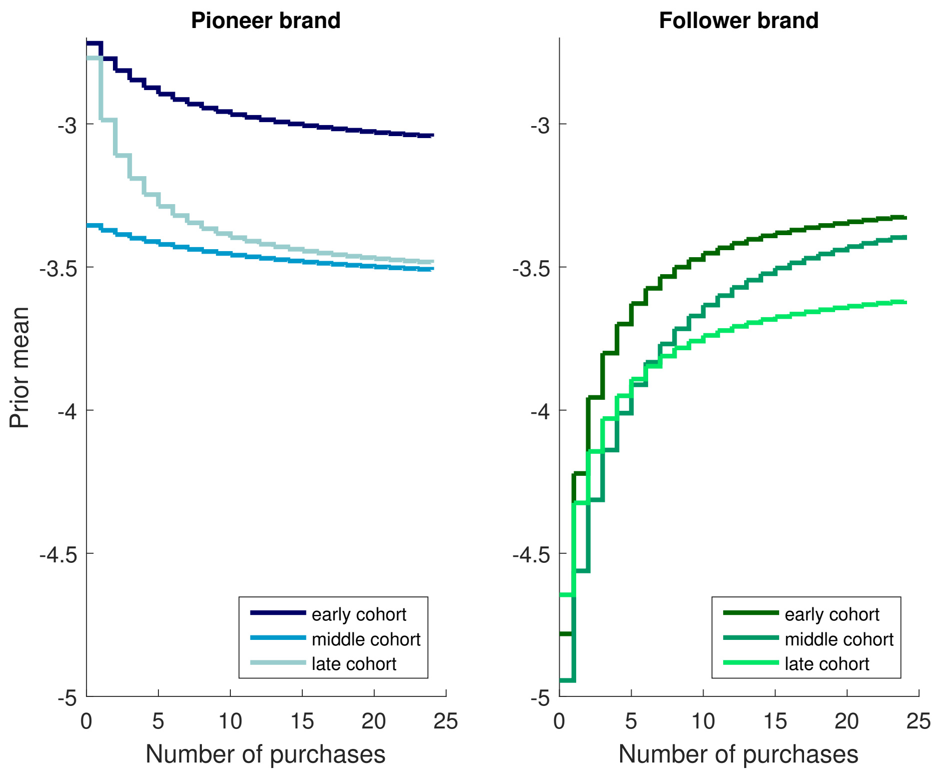
	par. est.	std. err.
initial brand-cohort belief mean:		
μ_{110}	-2.718	0.080
μ_{210}	-3.355	0.086
μ_{310}	-2.770	0.153
μ_{120}	-4.781	0.099
μ_{220}	-4.944	0.093
μ_{320}	-4.645	0.130
true brand-cohort match value:		
μ_{11}	-3.136	0.080
μ_{21}	-3.596	0.126
μ_{31}	-3.565	0.126
μ_{12}	-3.212	0.153
μ_{22}	-3.149	0.200
μ_{32}	-3.509	0.239
initial brand-cohort belief variance:		
σ_{110}	0.019	0.004
σ_{120}	0.037	0.006
σ_{210}	0.014	0.008
σ_{220}	0.026	0.004
σ_{310}	0.031	0.007
σ_{320}	0.031	0.010

Notes: We normalize the standard deviation of the signal, σ_v, to 0.5. The parameters regarding to the variances of the initial beliefs change with σ_v, while the rest of the parameters are not affected. In the existing empirical work, the range of the standard deviation of the signal is approximately from 0.1 to 1.4 (Crawford and Shum (2005); Szymanowski and Gijsbrechts (2012)).

across pioneer and follower brands (see below).

In addition, we find that the first and the last cohort initially value the category more than the second cohort, in the sense that they have the highest initial mean belief for both the pioneer and the follower brand. However, the true match values do not follow the same order. For the pioneer brand, the match value of the first cohort is higher than the ones for the other two cohorts, which are similar to one another; for the follower brand, consumers who adopt early are more likely to have higher brand match value. These patterns lend credence to the idea that adoption is partly based on preference for the category, with consumers who hold a high value adopting earlier and buying more.

Figure 2.4: Belief updating



Notes: Predicted means of the prior beliefs: using the model results. The y-axis is the predicted means of the consumer's prior beliefs. The x-axis is the number of purchases.

Table 2.5: Estimates (part 2 of 2): remaining parameters

	learning model estimates	
	par. est.	std. err.
price coefficient:		
α_1	-1.082	0.068
α_2	-0.790	0.074
α_3	-1.238	0.115
$\sigma_{\alpha_1} = \sigma_{\alpha_2} = \sigma_{\alpha_3}$	0.871	0.048
variety coefficient:		
μ_{ω_1}	0.029	0.002
μ_{ω_2}	0.020	0.003
μ_{ω_3}	0.028	0.004
heterogeneity in permenant taste:		
$\sigma_{\lambda_{11}}$	1.532	0.059
$\sigma_{\lambda_{12}}$	1.233	0.045
$\sigma_{\lambda_{21}}$	2.065	0.081
$\sigma_{\lambda_{22}}$	0.865	0.071
$\sigma_{\lambda_{31}}$	0.701	0.066
$\sigma_{\lambda_{32}}$	1.178	0.133
log likelihood	-45370.829	
number simulation draws	40.000	

The estimates of the variance of the initial beliefs, σ_{cj0}, suggest that for the first two cohorts, the initial uncertainty about the match value for the follower brand is higher than the one for the pioneer brand. This difference decreases with a consumer's adoption timing.

Thus, cohorts differ in the variance of their initial beliefs, and these different initial uncertainties imply heterogeneous learning rates. To illustrate this, in Figure 2.4, we plot the evolution of beliefs each cohort holds about a specific brand against the numbers of purchases. The level of the curves represents the mean of the beliefs about match value. The learning speed represents the level of the variance of the beliefs. This shows how consumers downward-adjust their beliefs about the pioneer brand and upward-adjust their beliefs about the follower brand. Learning ends after about 20 purchases. In the figure we see that, for example, the uncertainty level of the last cohort, as measured by the variance, halves after 3 to 4 purchases of a brand, while it takes 9 purchases for the first cohort until the variance of the belief is halved for the pioneer brand and $3-4$ purchases until the variance is halved for the follower brand.

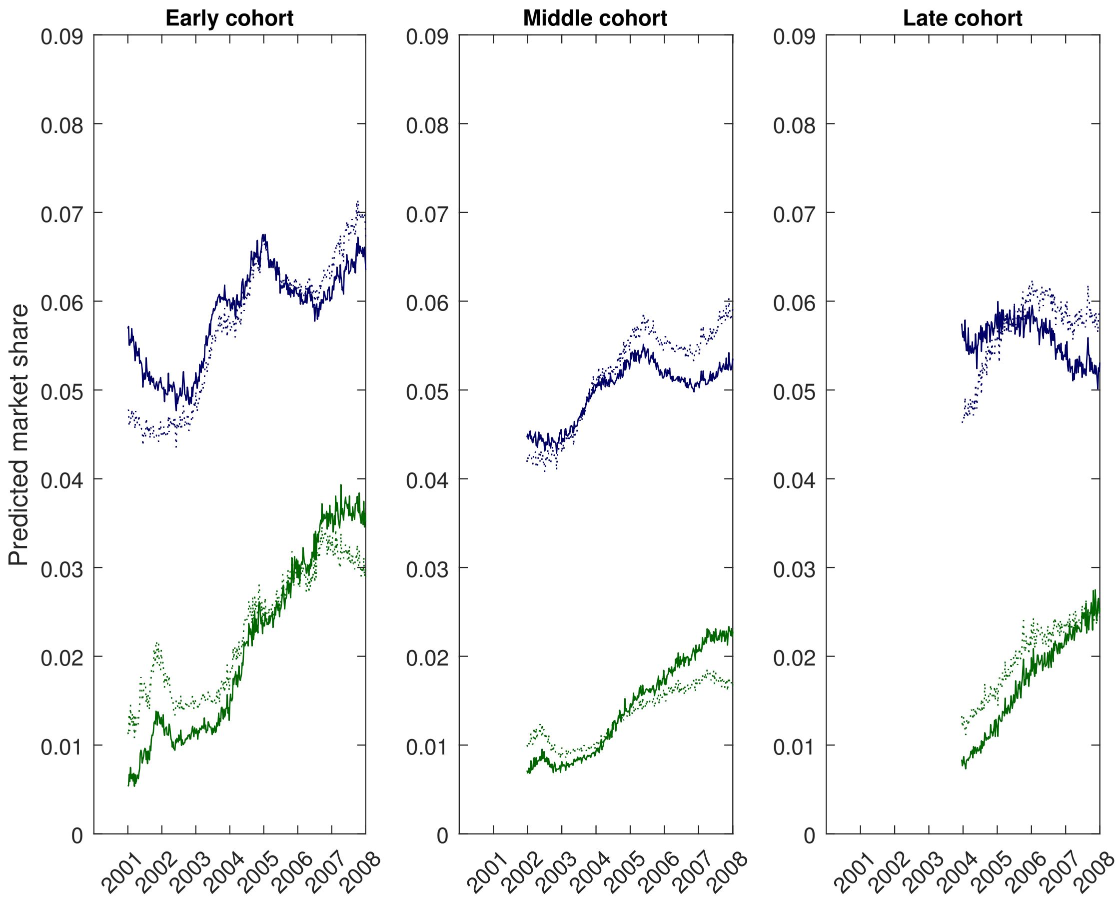
Turning to the static parameters of our dynamic learning model, recall that we allow for heterogeneity in match values within cohorts (captured by $\sigma_{\lambda_{cj}}$), heterogeneity in the price sensitivity across and within cohorts (captured by μ_{α_c} and σ_{α_c}), and heterogeneous tastes for variety across cohorts (captured by ω_c). The estimates of these parameters are reported in Table 2.5. They show that the second cohort is the least price sensitive, while the three cohorts are similar in their responsiveness to product line length. The static parameter estimates together with the estimates of the parameters regarding to consumer initial beliefs suggest that the observed marketing activities –price and variety– are not the only tools which sort consumers into the category. The observed adoption timing is the outcome of the consumer's responsiveness to marketing activities and the consumer's initial beliefs.

As a first way of illustrating the importance of learning to understanding consumer choice, we plot the prediction from our model against time and contrast it to the prediction of a static model. The specification of the latter resembles the former, with the difference that we (wrongly) impose that all learning has already taken place. Details are provided in Appendix 2.9.2. Figure 2.5 shows that the static model (in which consumers' brand match values are independent of their purchase experiences) will not be able to capture changes in market shares over time as well as the dynamic model does. The prediction from the static model (mistakenly) indicates that the sales of both brands increase over time, as the prices decrease over time and the product lines expend over time.

2.7 Counterfactual experiments

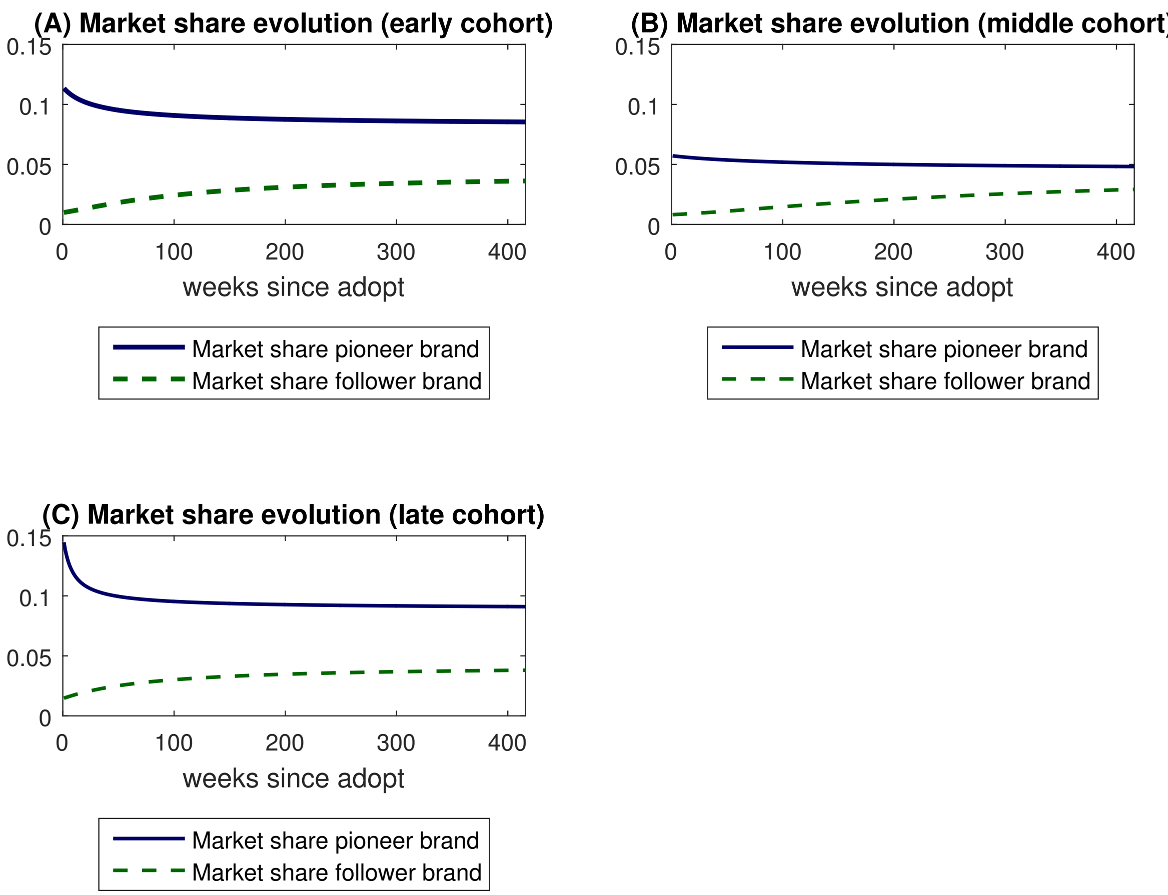
In this section, we study the implications of learning in more detail. In order to isolate them from time trends in prices and variety, we generate model predictions setting the price and the number of varieties of the two brands, respectively, to their sample averages. We then forward simulate consumer choice, keeping track of the number of times each brand has been purchased, and plot the resulting choice probability over time. Under the assumption that the quantity purchased given brand choice is the same for both brands and does not change over time, these are equal to market shares.

Figure 2.5: Predicted market share



Notes: Predicted market shares: using the model results and observed price and variety information. On the y-axis, predicted weekly market share of the two brands over time for the learning model and static model. Each subplot is one cohort. In each subplot, the upper group of lines are model predictions of the pioneer brand and the lower group of lines are the model predictions of the follower brand. The dotted lines are predictions using static model estimates, while the solid lines are model predictions with learning model estimates.

Figure 2.6: Predicted market shares holding supply side unchanged over time



Market share dynamics Figure 2.6 shows the resulting evolution of market shares over time, by cohort.[13] Any change over time is solely driven by learning. The market share of the pioneer brand declines over time, because consumers are initially too optimistic about the match value, while the market share of the follower brand increases, as consumers revise their beliefs.

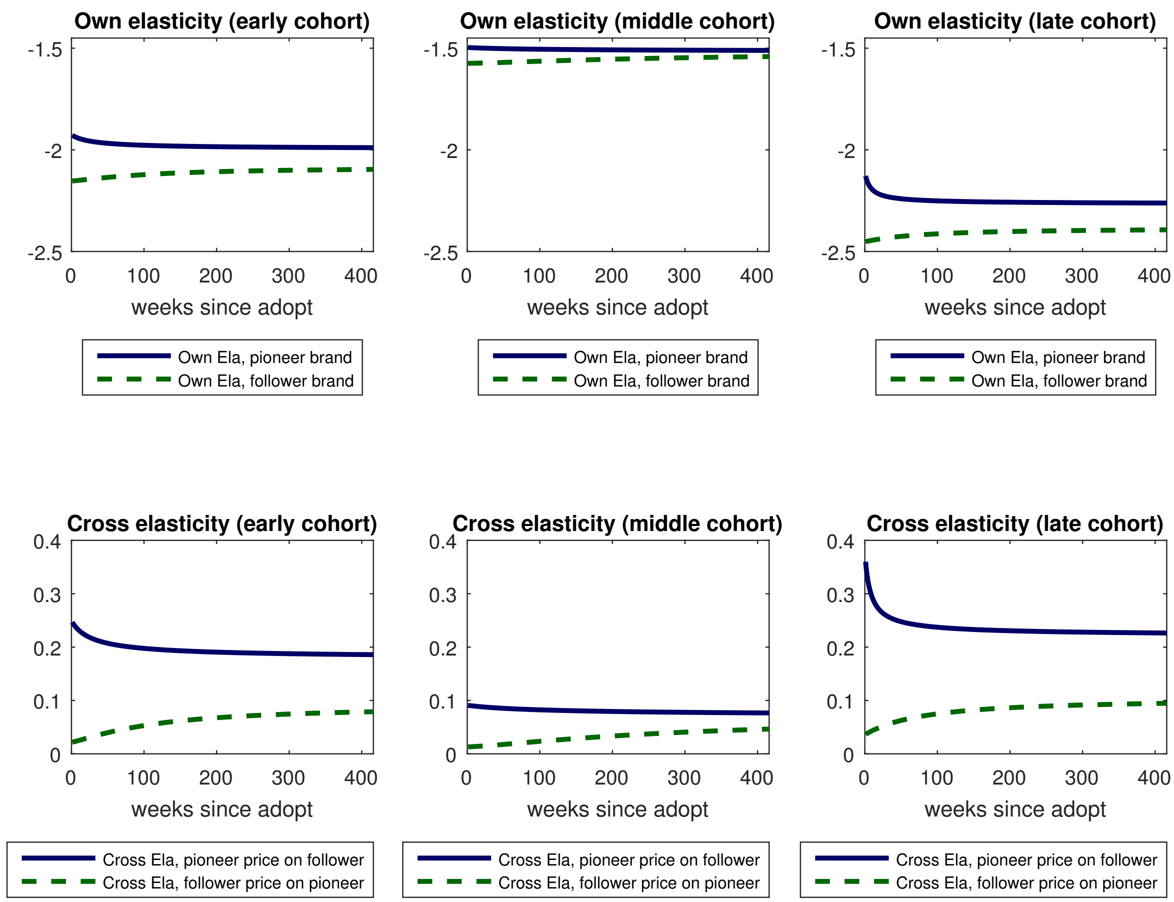
Using the last cohort (Panel (C) in Figure 2.6) as an example, consumer learning closes the market share gap between the two brands' by about 60%. This extends to other cohorts as well. That is, Figure 2.6 suggests that as the consumer gains more experience in the category, the pioneer and follower brands are less differentiated. Panel (B) presents the case where the market shares in the long run are no longer determined by different preferences for the brands but more by the marketing activities of two brands.

Elasticities and learning Next, we investigate the effects of price promotions and the implications of consumer learning for the price setting behavior by firms. To this end, we first compute price elasticities by cohort and plot them. In evaluating the elasticities, we set the level of prices and the number of varieties to a constant level for each brand. We next increase prices by a small amount and compute elasticities from the differences in predicted shares. We present the results in Figure 2.7.

The picture that evolves is that demand for the pioneer brand reacts less to own price than demand for the follower brand. At the same time, increases in the price of the pioneer brand have a larger percentage effect on demand for the follower brand, due to the fact that level of demand is lower in absolute terms. Over time, due to learning, demand for the pioneer brand becomes more price elastic, while the effect of the price of the pioneer brand on demand for the follower brand decreases in terms of magnitude. At the same time, demand for the follower brand becomes more responsive to price and the effect of the price on demand for the pioneer brand increases. Overall, we observe a move from an asymmetric setting to a more symmetric one.

[13] In Appendix 2.9.3, we present similar plots against purchase experience. The picture that emerges is slightly more nuanced, but the main conclusions remain the same.

Figure 2.7: Predicted elasticities holding supply side unchanged over time



Long run effect of temporary price cuts Investigating price effects in terms of elasticities does not allow us to paint the full picture, because price changes in a given week will have implications on future demand for both brands. The reason is that a price promotion will lead to changes in demand in that particular week, which will lead to learning, which will then in turn affect future demand. On top of that, the overall effects of price promotions also depend on the time at which they take place.

With this in mind, we next investigate the full dynamic effects of temporary price promotions as predicted by our learning model. Our previous discussion already suggests that the effects will be asymmetric. For both brands, the immediate effect of a price promotion is positive. However, for the pioneer brand, learning will lead to a downward adjustments of beliefs, while it will lead to upward adjustments for the follower brand.

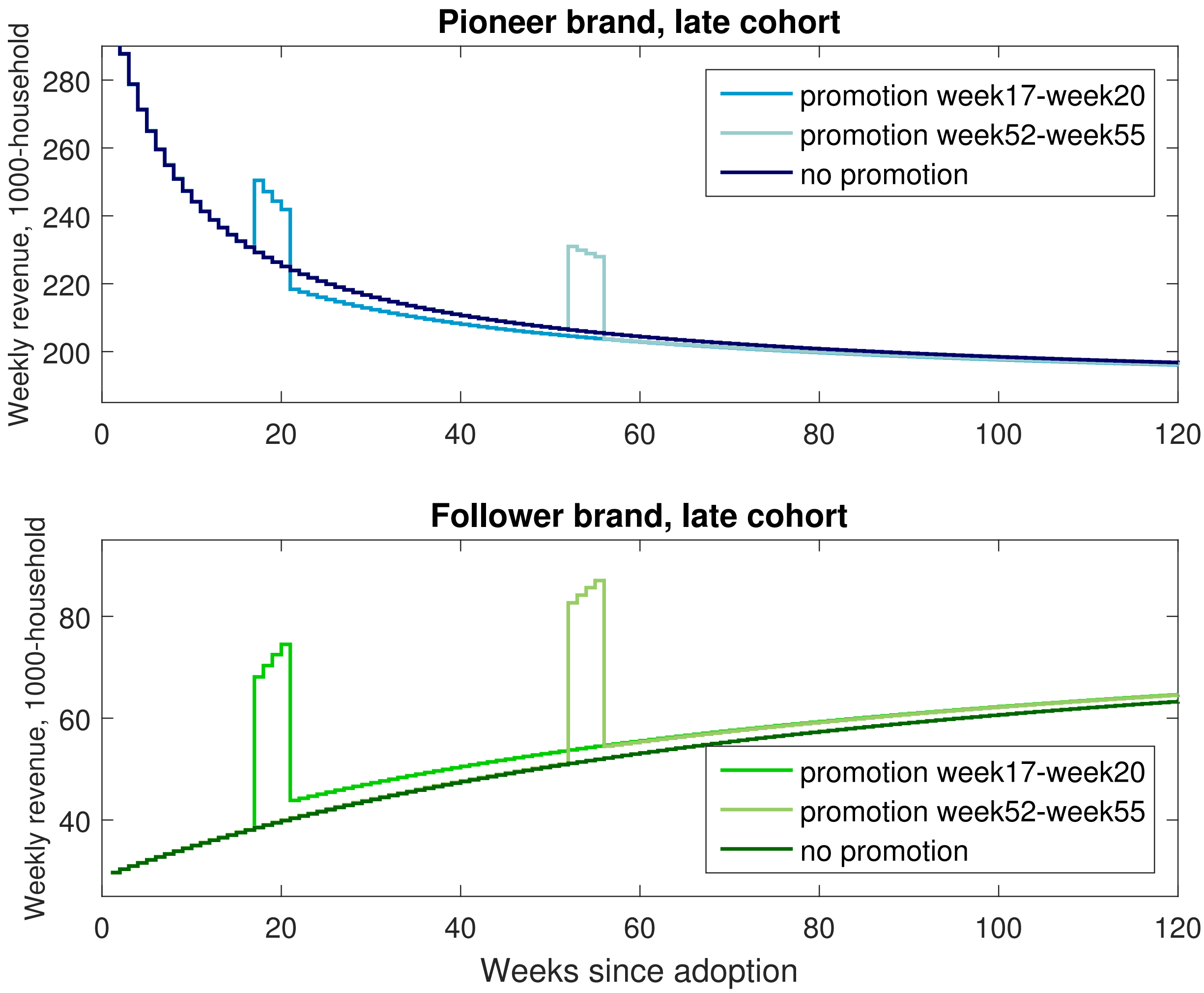
In our counterfactual experiments, we simulate weekly revenues for a hypothetical 1000 households per cohort (a market of 3000 households) under three scenarios. The first scenario establishes a baseline and calculates revenue at a regular constant price, holding variety constant at its sample average for each brand. Next, the second scenario is that promotion takes place during early periods of the consumer's life cycle. More practically, each brand (in turn) decreases price by 50% for 4 weeks, starting in week 17 after adoption, while the other brand's price remains unchanged. We call this condition the "week 17 promotion event". Finally, the third scenario simulates the price promotions to take place during later periods, in particular starting in week 52 after adoption ("week 52 promotion event"). These scenarios are carried through for each of our two brands and each of our three cohorts.

To see the long run effect of temporary price promotions, we calculate weekly revenue for the promoting brand from each cohort starting from the adoption week up to 250 weeks later.

Figure 2.8 shows the evolution of sales for the two experiments and the baseline, by brand and cohort. We see that the short run effect of a price promotion is always positive, but—due to learning—the dynamic effect is negative for the pioneer brand and positive for the follower brand.

These findings relate to the literature on long-run promotions effects, in particular the literature on promotion retraction effects (Dodson et al., 1978; Wathieu et al., 2004). In the

Figure 2.8: Effects of price promotions



Notes: The price promotion simulation outcomes are similar across cohorts, hence here we only plot the simulation outcome of the late cohort. See Appendix 2.9.4 for the promotion simulation plots for all three cohorts. The figures show the evolution of sales for a baseline scenario and two four-weekly 50% temporary price promotions. The promotions last for four weeks. Calculated using the estimated parameters and using 1000 households in the late cohort. (Regular) price and variety for each brand are set to their respective sample averages.

Table 2.6: Effects of price promotions

	Promotion window	Pioneer brand		Follower brand	
	(weeks since adoption)	SR Δrevenue [increase%]	LR Δrevenue	SR Δrevenue [increase%]	LR Δrevenue
Early cohort	week 17-week 20	82.19 [9.6%]	-97.09	88.94 [84.2]	192.09
	week 52-week 55	101.29 [12.9%]	-52.95	98.78 [64.7%]	156.80
Middle cohort	week 17-week 20	45.14 [9.8%]	-26.13	36.75 [49.9%]	73.16
	week 52-week 55	49.50 [11.2%]	-19.15	39.28 [42.6%]	63.50
Late cohort	week 17-week 20	75.43 [8.3%]	-157.13	128.60 [82.0%]	195.91
	week 52-week 55	93.90 [11.4%]	-66.15	133.39 [64.7%]	129.05

Notes: This table shows, by brand and cohort, the absolute and percentage (in square brackets) effects of a hypothetical 50% price promotion. The promotions last for four weeks. "SR Δrevenue" stands for "short run revenue change", which is the difference in revenue during the time of the promotion. "LR Δrevenue" is the effect in the following 1.5 years. Calculated using the estimated parameters and using 1000 households in each cohort. (Regular) price and variety for each brand are set to their respective sample averages.

literature on promotion retractions, some debate exists about whether such effects are positive or negative. In our learning framework, promotions stimulate consumption experiences. Whether these consumption effects are positive or negative is, in our model, fully dependent on the direction of the bias in initial match value. If a brand's match value corrects downward after consumption, then the after-effects of promotion induced consumption will be negative, relative to a regime where such promotions are absent. The opposite is true when the perceived match value is *ex ante* underestimated and the consumer updates her preferences for the brand positively. Obviously, rather than being promotion effects *per se*, these effects can alternatively be viewed as effects of promotion induced consumption and learning.

Table 2.6 summarizes the quantitative implications. We first turn to the short run effects in the third and fifth column. Consumers in the middle cohort react least to price promotions when looking at the absolute size of the effect. In percentage terms, the effects are similar across cohorts.

The absolute effects for the follower brand are comparable to the ones for the pioneer brand, but the percentage effects are much bigger. The reason for this is that individuals are pessimistic and therefore, in the absence of a promotion the probability to buy the follower brand is low, even at the later times. In general, the short run effects (in absolute terms) are slightly bigger when the price promotion takes place at a later point in time.

Turning to the long-run effects of price promotions, in column 4 and 6, we see that the long run effect is negative for the pioneer brand and positive for the follower brand. Moreover, the earlier the promotion takes place, the bigger the long run effect will be in terms of magnitude. Looking at short and long run effects in combination, we see that all the extra revenue the pioneer brand has gained from the early and late cohorts during promotion periods will be lost in post-promotion periods, and on top of that there will be an additional loss in revenues. This suggests that due to consumer learning, it is especially profitable for the follower brand to conduct price promotions, because they will reinforce learning, which in turn will lead to higher sales in the future.

2.8 Concluding remarks

In this paper, we investigate consumer brand choice behavior in the CPG market through the lens of a learning model in which we allow for rich heterogeneity across observed adoption timing. With the obtained parameter estimates from our structural Bayesian learning model, we provide suggestions on optimal temporary price promotion decisions.

In this study, we focus on the segment of consumers who have entered the market. We employ a long balanced panel, in which we observe a large number of households adopting, we estimate a structural brand choice model with Bayesian learning about utility. Our model allows novice consumers to have a biased perception about their post-experience match value and to be uncertain about their initial perception. We define consumer cohorts based on observed category adoption timing and incorporate cross-cohort heterogeneity in consumer's initial beliefs, true tastes, and responsiveness to marketing activities.

Estimates of our structural Bayesian learning model explicitly characterize the effects of learning in a new consumer's brand choice evolution. We first compare predicted market shares from our learning model with those from a static benchmark model where consumers' brand match values remain unchanged over time. This comparison shows that ignoring consumer learning leads to a biased view about how market shares evolve. To show how long the effect of consumer learning will last, we simulate the consumer's belief updating process. Given the average annual purchase incidence in this category, learning is non-negligible for 2 years after a consumer's category adoption.

Next, we show the effect of learning on different consumers' brand choice and how learning can shape the market structure among brands. We find that inexperienced consumers to have upward-biased beliefs about the pioneer brand and downward-biased beliefs about the follower brand.

We then take the readily observed consumer adoption decision as the segmentation scheme and estimate the demand primitives for each cohort. We find that consumer cohorts are different in their initial prior, permanent taste, and response to marketing activities. Early adopters are more certain about their initial perception for the pioneer brand than the follower brand. Consumers in the cohort of late adopters have the largest initial perception bias and are most uncertain about their initial beliefs. However, late adopters also have the fasted learning rate about the true match values. Earlier adopters have higher match values than later adopters. Overall, consumers continue to prefer the pioneer brand over the follower brand in the long run but consumption experience reduces the share gap substantially.

A consumer's initial perception bias gives the pioneer brand a clear short run advantage over the follower brand. To further show how demand side consumer learning can shape the market share evolution of the two brands, we simulate a counterfactual scenario, in which we hold the supply side changes constant. We also simulate the consumers' price elasticity matrix at each point of time after they adopt. We find that product experience makes consumers more price sensitive to the pioneer brand and less price sensitive to the follower brand. The cross price elasticities between two brands becomes more symmetric with a consumer's experience level, suggesting that consumption experiences in this category make the two brands more similar.

Brand managers in a new experience goods category should keep consumer learning in mind when planning price promotions. We found that the biased initial perceptions of the brands impact the efficacy of promotion policies. In particular, even though both brands experience a short run revenue gain during the promotion periods, the pioneer brand faces a long run revenue loss during the post-promotion periods while the follower brand has a large long run revenue gain after temporary price promotions.

At the same time, our promotion experiment also shows significant differences in promotion response across cohorts. This indicates brands may want to track the distribution of consumer adoption timing and incorporate this information in their marketing activity decisions. Interestingly, in the Dutch boxed meal category, the pioneer brand actively used price promotion in the early years in the category, whereas initially, the follower brand (which includes the private label brands) used a more even pricing strategy.

Our study has several limitations that set the stage for future research. First, a possible limitation of our model is that we treat the means of the distributions of the consumption experience signals as constant, as is usual in Bayesian learning models. In this paper, we assume the supply side innovations (caused by either existing brand's new product launch or new bran entry) can be largely captured by our product line measure, and therefore the means of the signals are viewed as constant. It would be interesting to further investigate whether the supply side innovations also affect consumer learning process. Second, we assume consumers only maximize their current period utility. In future research, an interesting model extension would accommodate forward-looking behavior, where consumers strategically try out brands to learn about their match values. Finally, in our current analysis, we took consumers' category adoption timing as given and estimate the cohort-specific primitives. In future work, we are interested in investigating why the cohort specific initial prior changes across cohort.

2.9 Appendix

2.9.1 Additional summary statistics

Table 2.7: Summary statistics of each cohort

cohort	number of households	adoption time mean	total purchase events max	mean	min
early cohort (adopt in 2001)	267	38th week, 2001	184	23.5	1
middle cohort (adopt in 2002)	330	22th week, 2002	114	15.1	1
late cohort (adopt in 2003/2004)	228	43th week, 2003	177	10.3	1

Notes: This table is the same as Table 2.1 in the main text, but without dropping households that did not have more than two purchases.

2.9.2 Static benchmark model

Figure (2.5) compares predictions from the learning model to the ones of a static consumer brand choice model. Here we provide more details on the latter.

To make the static model comparable with the learning model, we use a specification similar to the one in (2.13). Specifically, we use

$$u_{ijt} = \phi_{ij} + \alpha_i p_{ijt} + \omega_i x_{ijt} + \varepsilon_{ijt} \tag{2.13}$$

and

$$\begin{cases} \phi_{ij} \sim \mathcal{N}\left(\mu_{\phi_{cj}}, \sigma^2_{\phi_j}\right) \\ \alpha_i = \alpha_c + a_i, \qquad\qquad a_i \sim \mathcal{N}\left(0, \sigma^2_{a_c}\right). \\ \omega_i = \omega_c \end{cases} \tag{2.14}$$

That is, utility is assumed to consist of a time-invariant cohort-brand specific match value, and both the price and the variety coefficient are cohort-specific.

We estimate this static mixed logit model using the same data. Table 2.8 contains the resulting static model estimates.

Table 2.8: Static model estimates

	full	
	par. est.	std. err.
brand-cohort match value:		
$\mu_{\phi_{11}}$	-2.948	0.042
$\mu_{\phi_{12}}$	-3.187	0.050
$\mu_{\phi_{13}}$	-3.255	0.082
$\mu_{\phi_{21}}$	-3.963	0.064
$\mu_{\phi_{22}}$	-4.453	0.062
$\mu_{\phi_{23}}$	-4.187	0.080
heterogeneity in match value:		
$\sigma_{\phi_{11}}$	1.368	0.052
$\sigma_{\phi_{12}}$	1.352	0.047
$\sigma_{\phi_{21}}$	1.895	0.064
$\sigma_{\phi_{22}}$	1.320	0.073
$\sigma_{\phi_{31}}$	1.118	0.053
$\sigma_{\phi_{32}}$	1.582	0.117
price coefficient:		
$\mu_{\alpha 1}$	-1.021	0.060
μ_{α_2}	-0.826	0.070
μ_{α_3}	-1.264	0.108
$\sigma_{\alpha_1} = \sigma_{\alpha_2} = \sigma_{\alpha_3}$	0.824	0.052
variety coefficient:		
μ_{ω_1}	0.037	0.002
μ_{ω_2}	0.028	0.002
μ_{ω_3}	0.031	0.003
negLogLikelihood	45565.523	
SimulationDrawNum	40.000	

2.9.3 Effect of purchase experiences on market shares and price elasticities

In Section 2.7, we have presented plots of market shares and elasticities against time. We have obtained those by forward-simulating consumer choice. Alternatively, we can generate plots against purchase experience. The state variable here is a tuple and consists of the number of times each of the two brands has been bought up to that moment.

Figure 2.9 does so for the probability to buy either of the brands and Figures 2.10 through 2.12 contain the corresponding own and cross price elasticities. The picture that emerges is somewhat more nuanced, but the general pattern is already summarized in the figures presented in Section 2.7.

Figure 2.9: Dependence on choice probability on purchase experience

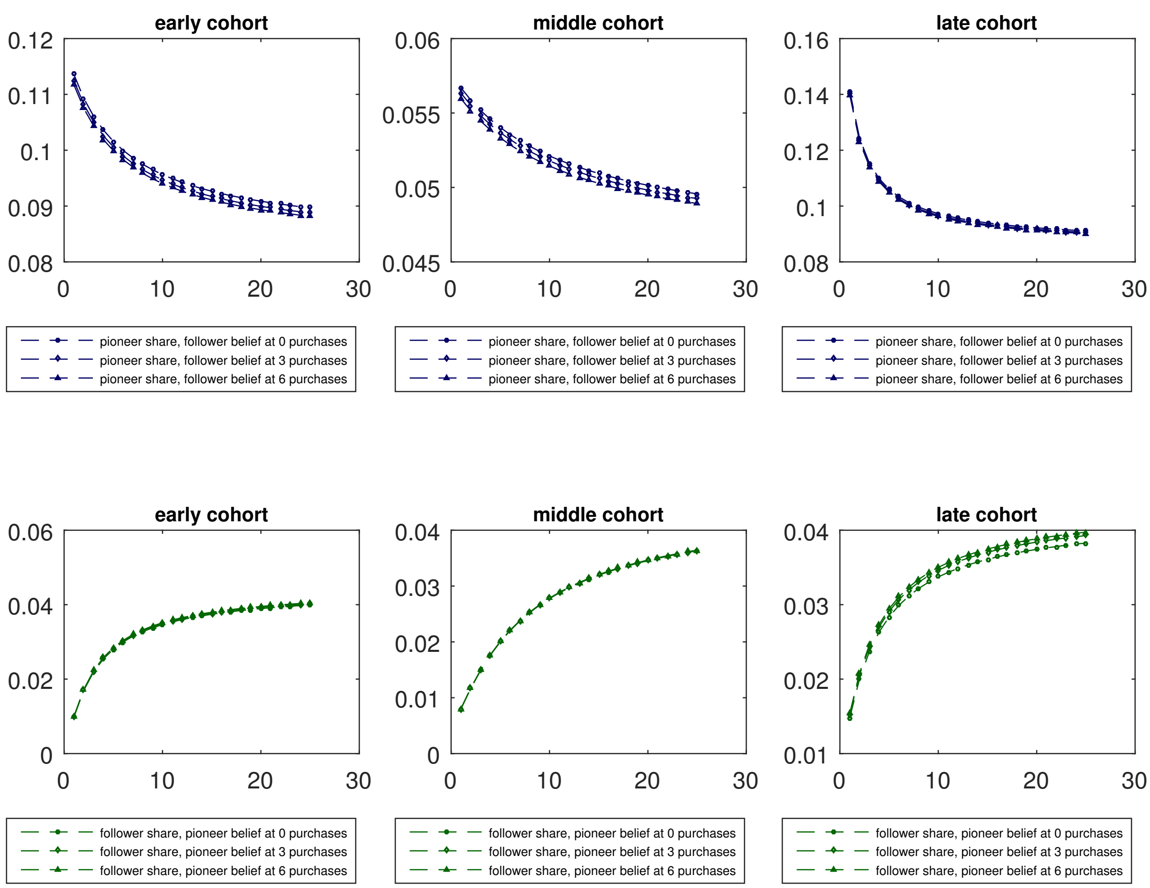


Figure 2.10: Dependence on elasticity on purchase experience—early cohort

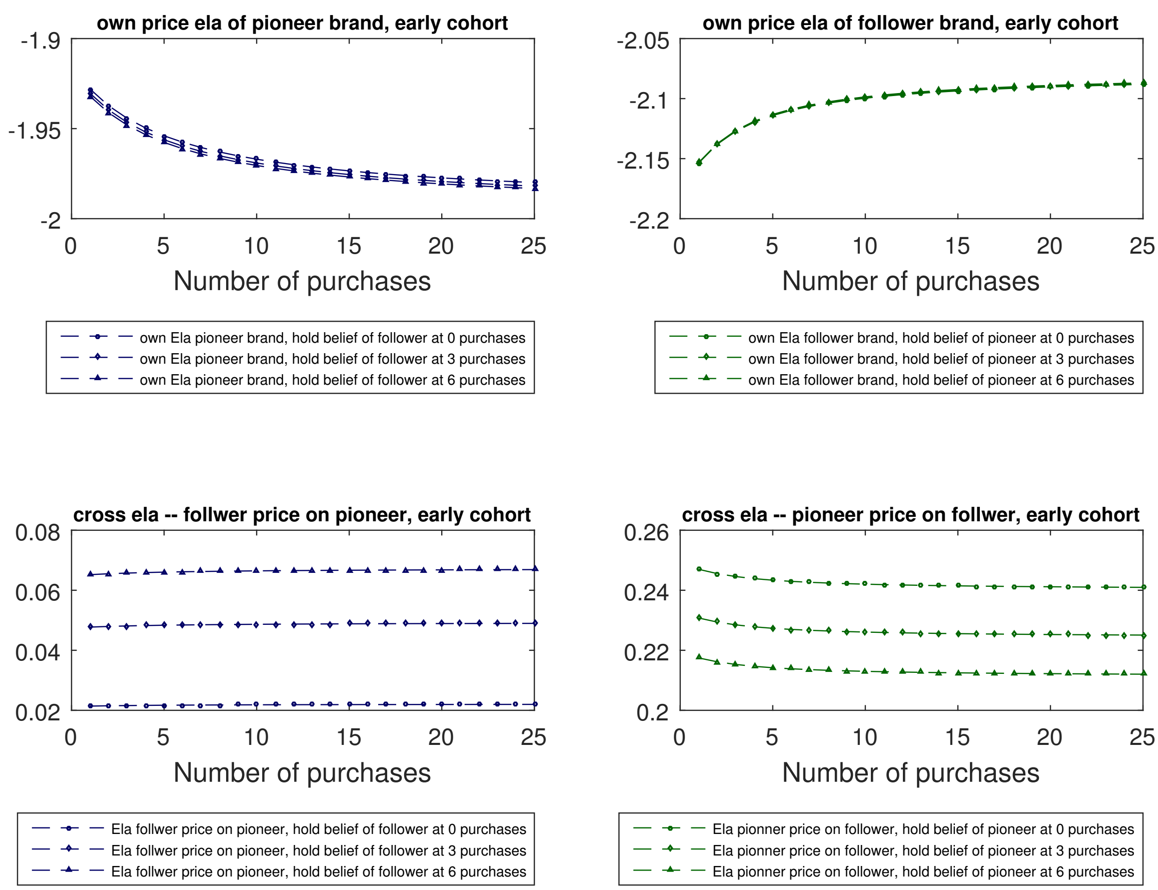


Figure 2.11: Dependence on elasticity on purchase experience—middle cohort

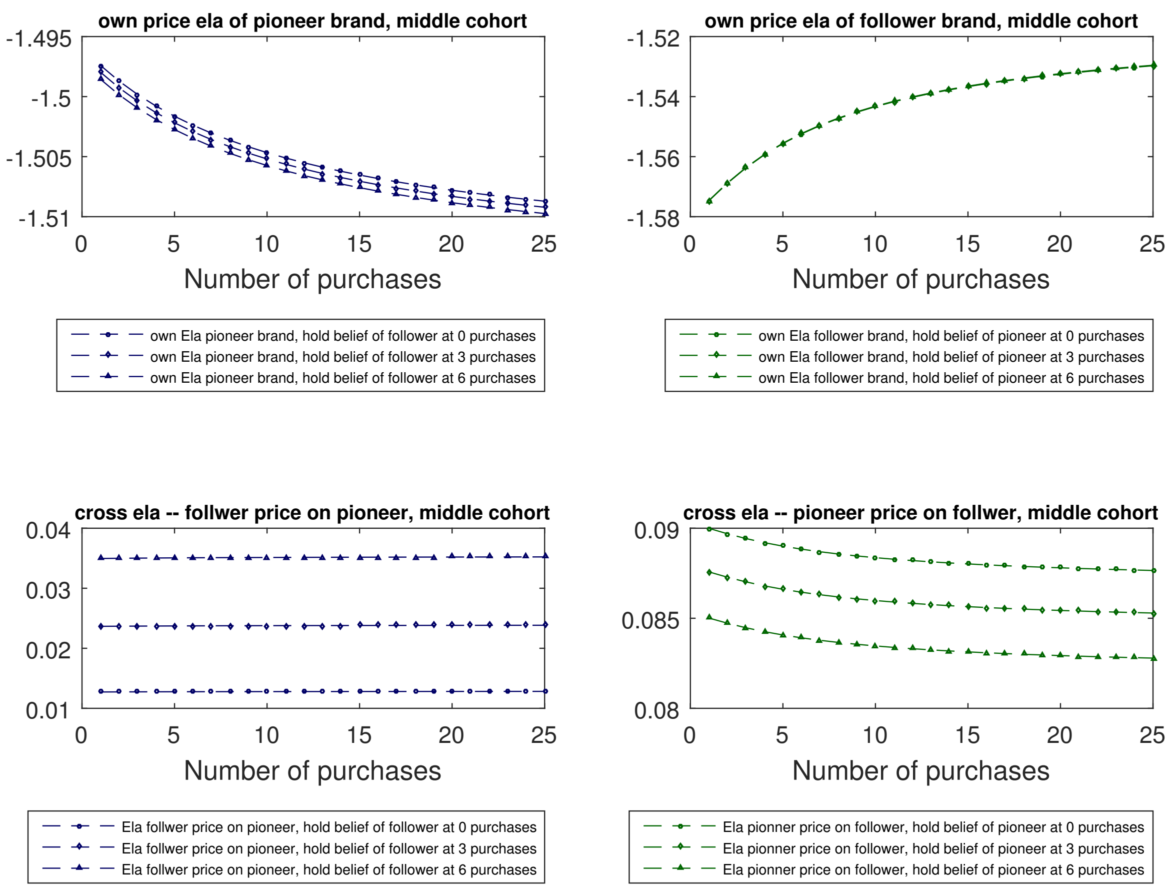
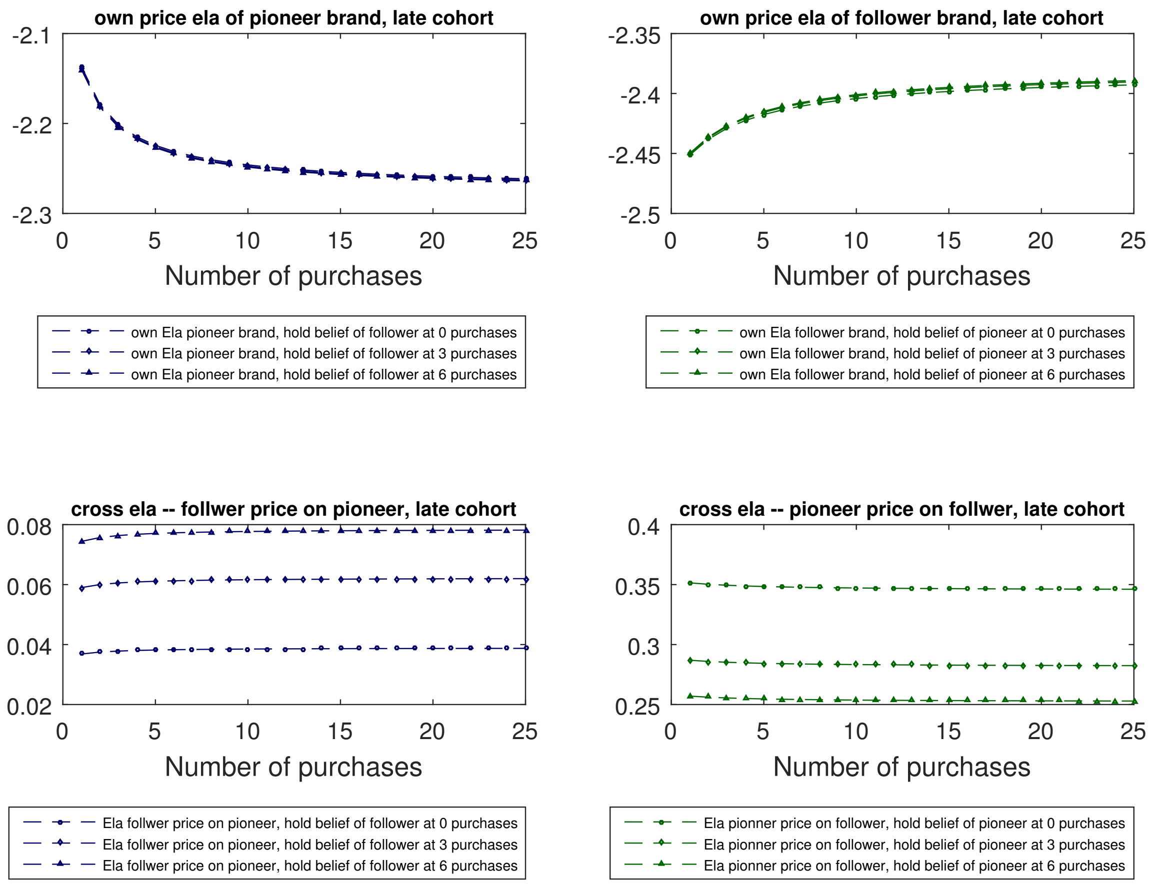
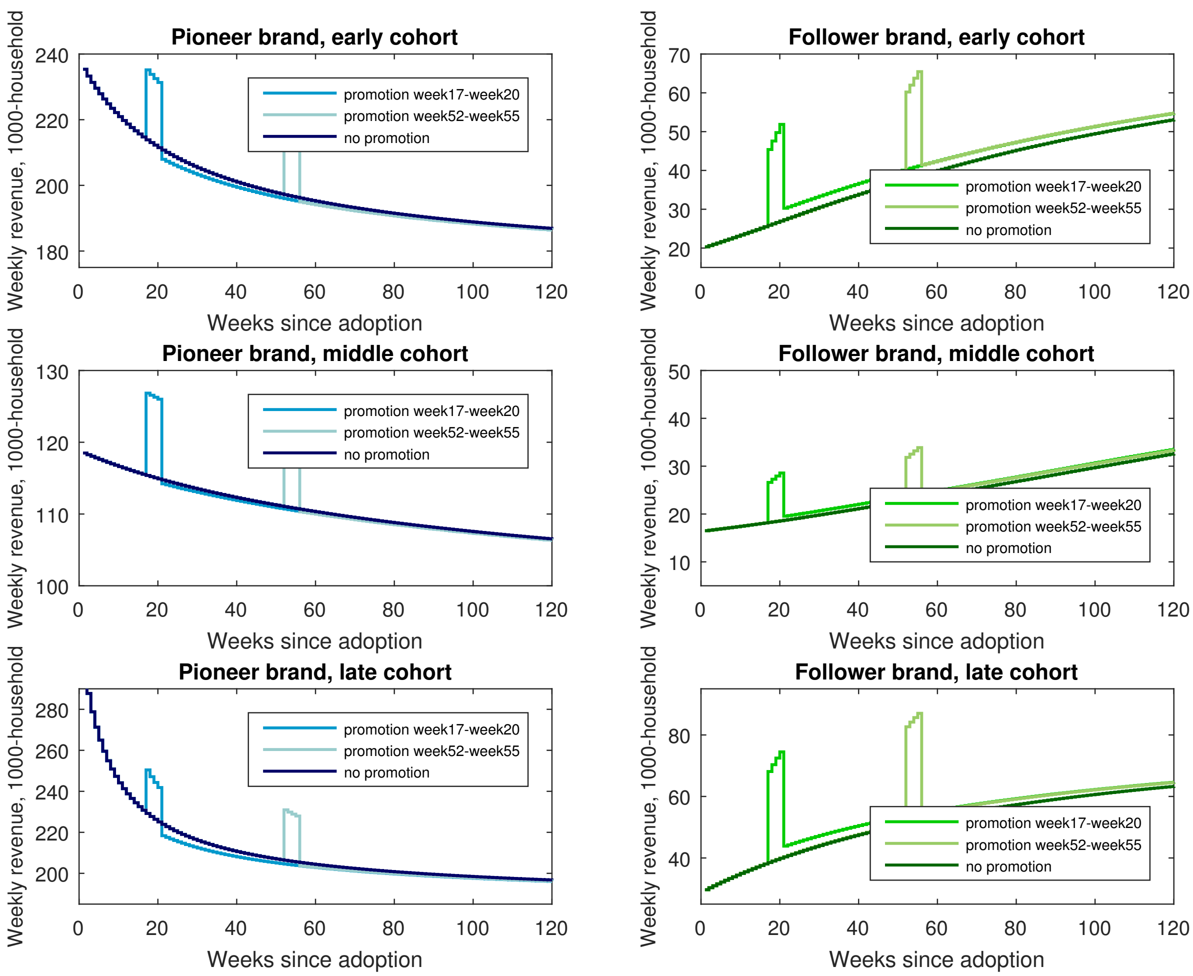


Figure 2.12: Dependence on elasticity on purchase experience—late cohort



2.9.4 Price promotion experiment

Figure 2.13: Effects of price promotions



Notes: The figures show the evolution of sales for a baseline scenario and two four-weekly 50% temporary price promotions. The promotions last for four weeks. Calculated using the estimated parameters and using 1000 households in each cohort. (Regular) price and variety for each brand are set to their respective sample averages.

2.9.5 Detailed description of the follower brand

The market for boxed meals is very concentrated at the brand level with the pioneer brand, Knorr, accounting for roughly 75% of the market share in volume and revenue (on average across the eight years). The rest of the market is covered by several other national brands (e.g., Honig and Conimex) and store brands. We define the follower brand as a composite with the other national brands and store brands. The other national brands enter the market later than the pioneer brand, but already exist at the beginning of 2001. However, the first store brand product was launched in 2002 by Albert Heijn, one of the biggest retailer chains in the Netherlands.

In this section, we provide more information on the follower brand. First, we use "other" brand to represent all the national brands that enter later than the pioneer, and use store brand to represent all the store brands produced by different retailers. Next, we present summary statistics for the pioneer brand, "other" brand and store brand, respectively. Finally, we plot the unconditional purchase shares of the three brands in any given week and for each cohort.

Table 2.9 reports summary statistics for the market shares and brand availability at the brand level over the eight years. Retailers generally sell the boxed meal category and have been providing pioneer brand and "other" brand from the start of our data in 2001. More and more retailers started to provide store brand since 2001. The store brand's market share level is similar to the "other" brand, and both of their market shares increase over the eight years.

Table 2.9: Summary statistics of the boxed meal market at the brand level (part 1 of 2)

	market share (units)		market share (euros)		availability		
year	pioneer brand	"other" brand	pioneer brand	"other" brand	pioneer brand	"other" brand	store brand
2001	0.898	0.102	0.889	0.111	99.8%	94.8%	0.00%
2002	0.814	0.110	0.800	0.117	99.8%	92.8%	25.4%
2003	0.853	0.085	0.844	0.091	99.7%	95.7%	27.6%
2004	0.805	0.098	0.801	0.101	97.6%	94.9%	62.3%
2005	0.662	0.082	0.675	0.087	93.5%	91.0%	74.8%
2006	0.615	0.184	0.627	0.186	95.2%	94.6%	76.0%
2007	0.625	0.182	0.633	0.183	97.0%	97.0%	90.3%
2008	0.600	0.174	0.615	0.172	99.7%	93.1%	90.3%

Notes: The statistics in this table are based on cross sectional data for all 5000-7000 households per year. The market shares of pioneer brand (Knorr) and "other" brand (non-Knorr national brand) are calculated both in terms of units and euros (first four columns), respectively. The availability measure (last three columns) is calculated as the percentage of retailers that sell a specific brand versus the total number of retailers.

We then present summary statistics on the number of unique "member" brands in pioneer brand, "other" brand and follower brand, and the corresponding concentration level in table 2.10. We find that more non-Knorr national brands and store brands gradually enter the market during 2001 and 2008. Both the store brand group and the "other" brand group become less concentrated over time, while at each point of time the "other" brand group is more concentrated.

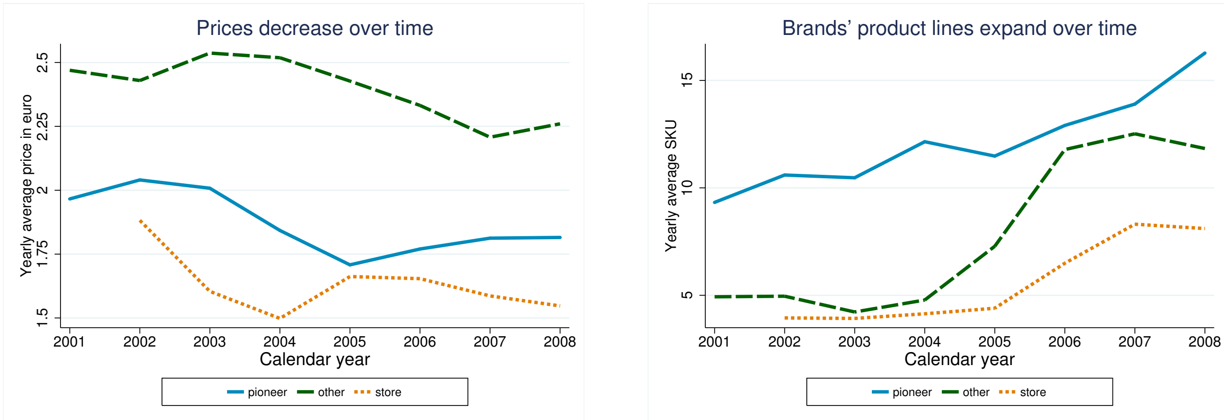
Table 2.10: Summary statistics of the boxed meal market at the brand level (part 2 of 2)

	Nr. brand			HHI		
year	pioneer brand	"other" brand	store brand	pioneer brand	"other" brand	store brand
2001	1	9	0	1	0.81	.
2002	1	8	1	1	0.83	1
2003	1	7	2	1	0.78	0.84
2004	1	9	6	1	0.68	0.50
2005	1	11	8	1	0.49	0.23
2006	1	12	8	1	0.37	0.17
2007	1	17	10	1	0.38	0.18
2008	1	18	9	1	0.40	0.15

Notes: Nr. brand is the number of unique brands that are clustered in the three brand alternatives – pioneer brand, "other" brand, and store brand. The last three columns present the concentration levels of the three brand alternatives defined in this paper. The concentration level is measure by Herfindahl-Hirschman Index (HHI). HHI is calculated by taking the purchase share of each brand within the three brand alternatives, respectively, squaring them, and summing the results.

Next, we present the two main supply-side trends –price variation and the variety expansion– over time in figure 2.14. The average transaction prices of all the three brands are decreasing over time, while the varieties are increasing over time. At each point of time, the pioneer brand offers more variety, but the "other" brand charges the highest price.

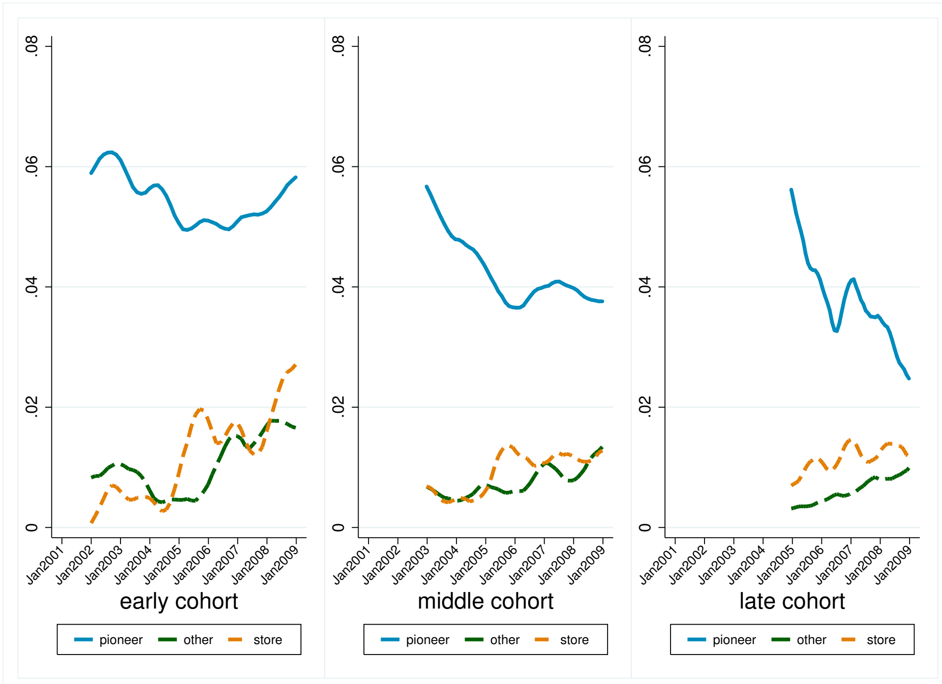
Figure 2.14: Price trend and variety trend by brand



Notes: The left subplot presents the price trend for each brand over time. The right subplot present the variety (measured by the number of unique SKUs) trend for each brand over time.

Finally, in figure 2.15, we present the unconditional purchase shares of three brands in any given week and for each cohort. We see that the purchase shares of "other" brand and store brand among all the adopter cohorts increase over time, and are at similar levels at each point of time.

Figure 2.15: Purchase shares of pioneer brand, "other" brand, and store brand by cohort and time



Notes: Figure 2.15 is a plot of a local polynomial smooth of brand choice indicators against calendar time, separately for each cohort. We do not use each consumer's first purchase incident, because the timing of a consumer's initial purchase is already used to define consumer cohorts. If a consumer made a trip to the supermarket but did not purchase any boxed meal, then we code this as him choosing the outside option. If a consumer had no supermarket visit in a given week, then we treated this as a missing observation.

2.9.6 Product line length

Our product line length measure is based on the transaction data. We use the number of the unique brand UPCs in the assortment for a given retailer and year as the product line length measure. Since this measure is based on the total number of UPCs that are observed to have positive sales, it will be an effective measure when we don't have large numbers of UPC exit during our observation window. In the table below, we verified that the number of UPC exit is almost negligible.

Table 2.11: Summary statistics of the Dutch boxed meal market at the brand level

	product exit	
year	pioneer brand	follower brand
2001	0	0
2002	0	1
2003	0	0
2004	0	0
2005	0	0
2006	0	3

Notes: The availability measure (the first two columns) is calculated as the percentage of retailers that sell a specific brand out of the total number of retailers weighted by the retailer market share. Product exist is defined as the number of SKUs that appear to have zero sales for at least 1.5 year.

Chapter 3

Time Use and Purchase Behavior

This chapter is based on joint work with Bart Bronnenberg and Tobias Klein.

3.1 Introduction

In his classic contribution, Becker (1965) argues that consumer purchase decisions should be analyzed within the broader context of time allocation and the use of market goods as inputs to home production. Home production is the conversion of market goods into consumption goods using time. It is of first-order economic importance: The home-produced output of consumption goods and consumption experiences is estimated to be 20-50 percent of the gross national product (Eisner, 1988; Benhabib et al., 1991).

Becker's seminal contribution has led to a wealth of insights that we summarize in Section 3.2. The common thread is that time costs are approximated as being proportional to the consumed quantity. We build on and relate to this literature in a conceptual framework that we spell out in Section 3.3. The main contribution of this paper is to present novel empirical evidence that a non-negligible part of the time is for instance spent on planning a shopping trip, going shopping, searching and evaluating products, preparing meals, and consuming them at home. A consumer's available time for home production activities largely influences his grocery expenditure decisions, demand for variety and product type preferences. We also present empirical evidence supporting the idea that part of the time cost is actually fixed to the amount of goods consumed. This gives rise to a set of costs of using the market, and therefore influences households' purchase behaviors.

A household's use of the market can be broken down into three different dimensions. First, the market distributes market goods to the households for them to undertake home production activities. However, the market is not free. For instance, a household faces transaction costs, e.g., in the form of transportation (Hotelling, 1929; Salop, 1979; Bronnenberg, 2017). Intuitively, one can view the transportation cost per shopping trip as the price of using the market distribution system. Second, households optimally choose the amount of time-intensive activities (requiring time-intensive market goods) and the amount of non-time-intensive activities (requiring goods-intensive market goods). The former means making less use of the market (shift to the home production sector from the market sector), and the latter means making more use of the market (Becker, 1965; Gronau, 1977). Finally, besides choosing the frequency of using the market and the time-intensity of the shopping basket, households can economize time on a third dimension: variety. Gronau and Hamermesh (2008) model time cost for household demand for variety as per variety-quantity cost. Hamermesh (2005) and Bronnenberg (2015) view it as per variety fixed cost. What connects these theoretical perspectives is the suggestion that the degree of a household's demand for variety is moderated by its time budget and cost of time.

Even though the home production theory provides a coherent framework to understand household purchase behavior and its subsequent influence on the market equilibrium, most of the *empirical* work on household purchase decisions has overlooked the impact of time on household purchase behaviors. This omission is partly caused by a dearth of data directly or indirectly measuring time availability for home production of consumption goods matched with purchasing decisions of the same households. In this paper, we have access to a unique balanced panel data set that combines scanner data of home-scanned grocery purchases with matched annual survey data covering reported weekly hours spent in the labor market and in household tasks. The matched annual survey data also covers household level events such as retirement and unemployment which lead to variations in a household's available time for home production and its opportunity cost of time (proxied by wage). As a result of observed work-hour change or events, our data contains ample within-household variation on available time for home production and opportunity cost of time. Our empirical setup further allows us to control for fixed effects and observed changes in household characteristics, so the variation in the time budget and opportunity cost of time are arguably exogenous to the grocery purchase decision, allowing us to give a causal interpretation to our estimates.

We first document household purchase behaviors across four dimensions of shopping: shopping activity, total grocery purchases, time-intensive goods versus non-time-intensive (convenience) goods preferences, and preferences for variety. Our measures of household purchase behaviors are computed from the GfK ConsumerScan Panel tracking the Dutch grocery market. The panel contains 4,358 households who record their purchases in each of the five years from

2009 to 2013, which forms a balanced panel. These households also participate in a yearly survey administered by GfK in the Netherlands. From the survey data, we retrieve information on household demographics, including employment status and parenthood status, as well as preferences for home production activities like cooking and shopping.

The main empirical findings in this paper are: Adjusting for income heterogeneity and time invariant household heterogeneity, there is a negative relation between a household's available time for home production and the number of shopping trips, the amount of grocery expenditure and the purchased quantity and the number of purchased varieties. We reach similar conclusions with household self-reported time allocation data and with household-level time budget shifting events (retirement and unemployment). Additionally, using the same shifting events as exogenous variations to the household's opportunity cost of time, we find a negative relation between a household's opportunity cost of time and the time-intensity level of its shopping basket.

This paper proceeds as follows. Section 3.2 contains an overview of the literature. Next, Section 3.3 discusses a simple theoretical framework using household production functions. In Section 3.4, we describe our data. Section 3.5 discusses the empirical approach and the results along with some robustness checks. Section 3.6 concludes.

3.2 Literature

Our work contributes to various strands of the literature. First, it relates to the literature on the economics of time use. In this literature, there has been a recent boom of articles using time diaries or macro level shocks to household budgets to address a variety of economic questions related to consumption and non-market work-time allocation.

A first set of contributions, which includes among others Aguiar and Hurst (2007b); Kimmel (2008); Ramey (2009); Ramey and Francis (2009); Aguiar et al. (2012); Lee et al. (2012); Aguiar et al. (2013); Kawaguchi et al. (2013); Duernecker and Herrendorf (2015), documents the trends in household time use over long periods of time and/or documents stylized patterns in multi-nation time use data. A consistent set of estimates of long-run trends in time use is demanded by the recent interest in macroeconomic and growth models that incorporate home production (see for example Benhabib et al., 1991; Greenwood et al., 2005; Francis and Ramey, 2009). The effects of marketing and retailing on household production and time use are likely important, but are absent from most analyses.

Another set of contributions (including Biddle and Hamermesh, 1990; Solberg and Wong, 1992; Cutler et al., 2003; Aguiar and Hurst, 2005, 2007a; Bertrand and Schanzenbach, 2009; Meyer and Sullivan, 2008; Stratton, 2012; Stancanelli and van Soest, 2012; Nevo and Wong, 2015) uses information on consumption and time use (sometimes as macro-level shocks to households' time and money budgets) to answer questions related to household well-being, i.e., consumption, sleep, leisure, and gender differences in time use. Most of these papers use survey data on household consumption expenditure. The only exception is Nevo and Wong (2015), who exploit scanner data on detailed household shopping budget choices and shopping trip decisions. Unlike other time use papers, which use direct information on household time allocation (often times the data is cross-sectional, e.g., Hamermesh et al., 2005) or household level budget-shifting events, Nevo and Wong (2015) use macro-level shock, i.e., The Great Recession, as the household individual budget-shifting event.

We complement existing work, especially the second set of contributions, by providing household level evidence with detailed scanner data together with direct information on household time allocation and household level time budget shifting events. To this end, we use various measures of household consumption decisions and two empirical strategies—direct information on work hours and non-market work hours of the whole population sample, and exogenous time budget shifting events, retirement and unemployment, which mainly take place among the elder sub-population sample. With our rich survey panel data and scanner panel data, we are able to further control for home production preferences heterogeneity, income heterogeneity, and time-invariant unobserved household heterogeneity.

Our findings with the elder sample, who have experienced the retirement event, complement the literature on the "retirement consumption puzzle" summarized in Hurst (2008). The general finding in the U.S. context is that food expenditures fall, because individuals dramatically increase the amount of time they spend on home production, substituting expensive products by time-intensive ones, and because they buy the same products at lower prices (Aguiar and Hurst, 2007a). We find significant increases in household expenditure, purchase units, and purchase variety in food purchases and in the amount of shopping trips through retirement after adjusting for rich observed heterogeneity in demographics, home production preferences and income.

Further, our work relates the literature on time use to the literature on firms' product selection. Among earlier theoretical work, Spence (1976) investigates the effects of fixed costs on the selection of products and product characteristics. Fixed costs force firms to choose from the large set of all conceivable products. At the same time, households also face fixed costs that limit them to purchasing a subset of affordable varieties (Bronnenberg, 2015). Our results suggest that one source of this fixed cost is the household's time—when households have more available time, they will be able to complete more home production activities (more likely to backup the fixed cost) and, therefore, demand more variety; on the contrary, when households have less time, they will refuse to undertake home production activities because of the fixed cost and opt for the outside option, e.g., restaurant. Our results suggest that ignoring the time component in household demand for grocery goods may lead to a mismatch in the desired number of varieties by households and the number of varieties provided by retailers.

Finally, our work is related to a broad marketing literature on the effect of time on consumer behavior. Among earlier work, Kelley (1958); Jacoby et al. (1976) stimulate conceptual and empirical attention regarding the relationships between time and consumer purchase behavior by summarizing the types of convenience that can be provided by the firms, and reviewing what has been found in the fields of economics, sociology, home economics, psychology, and marketing. The empirical literature investigates the importance of consumer's time and/or the opportunity cost of time in various marketing issues. For example, Gross and Sheth (1989); Duncan Herrington and Capella (1995); Nowlis (1995); Kenhove and De Wulf (2000); Vermeir and Van Kenhove (2005); Baltas et al. (2010); Pozzi (2012); Vroegrijk et al. (2013) present empirical evidence, which suggest time is an important factor in consumer product choice, response to price promotion, response to advertisement, store loyalty, channel choice and shopping trip decisions. This paper systematically document consumer's purchase behaviors and provide direct evidence on the importance of time on consumer purchase decision making with large household level panel data.

3.3 Conceptual framework

The point of departure for our conceptual framework is the "new consumption" theory of home production in economics pioneered by Becker (1965) and Gronau (1977). Below, we spell out a model that explicitly features alternative uses of money and time, which allows us to derive empirical predictions related to the effects of additional costs of using the market without explicitly modeling the exact mechanism through which they influence consumer behavior. The main assumption of our model is that we assume the work-time change is exogenous to a household's grocery purchase decisions and that retirement and unemployment lead to exogenous variations in a household's opportunity cost of time. The model predicts that an increase in the available time will have positive effects on the amount of goods bought, the number of shopping trips undertaken, and that households will substitute towards more time-intensive goods.

Our modeling choice is, for the sake of tractability, to not model the number of varieties bought explicitly. However, if there is fixed cost in consumer's demand for variety (Bronnenberg (2015)), then it follows very naturally from the derived results that the number of varieties bought will increase if the total amount of purchased market goods increases. An example of the fixed cost in household's demand for variety is the cost of exploration/learning. Most of the products in the grocery market are experiences goods, and therefore it takes a household energy and time to learn about a new variety. Since this cost is a per variety cost, a time-scarce household will restrict its purchases to the familiar items instead of try new items to avoid additional cost.

3.3.1 A model of time allocation and demand for goods and time

A household maximizes its utility subject to not only a money budget constraint but also a time constraint. The utility function takes as arguments quantities of consumable commodities—the outputs of home production activities using market goods and time as inputs. We consider two types of time use that closely relate to household's grocery purchases—shopping and meal preparation. A consumable commodity can be produced by either non-time-intensive meal preparation, e.g., warming up a frozen pizza, or time-intensive meal preparation, e.g., making a pizza from fresh ingredients.

To fix ideas, we denote the commodity produced by non-time-intensive meal preparation activity with Z_m, and the commodity produced by time-intensive meal preparation activity with Z_h. Households derive utility from quantities of Z_m and Z_h, along with the residual money left, denoted by X. X can be viewed as the outside goods—the resources that are not used as inputs in home production enter the outside goods. The household's utility is

$$U = U\left(Z_m,\, Z_h,\, X\right), \tag{3.1}$$

and we assume that it is increasing in its arguments and quasi-concave. Like Becker (1965), we make the assumption that goods and time are perfect complements in the production of each activity and they are used in fixed proportions to produce this activity. Z_m and Z_h are produced using market goods X_m and X_h, and time T_m and T_h, respectively, as inputs. We further assume a linear technology

$$\begin{cases} X_j = c_j Z_j \\ T_j = b_j Z_j \end{cases}, \ j \in \{m,\ h\},\ c_m > c_h > 0, \text{ and } b_h > b_m > 0, \tag{3.2}$$

where c_j and b_j ($j = m,\ h$) are constant parameters that represent the fixed goods input and time input intensity required by each commodity produced by household home production activity, respectively. The commodities produced by time-intensive activity require more time input but less goods input compared with the ones produced by non-time-intensive activity. Therefore, we let $b_h > b_m > 0$ and $c_m > c_h > 0$.

In order to access market goods X_m and X_h, the household has to undertake shopping trips. We assume one unit of shopping trips brings a household $I > 0$ units of market goods, and takes b_s amount of time. We use Z_s to denote the total number of shopping trips and assume each unit of shopping trips consumes b_s units time. Therefore, a household spends $T_s = b_s Z_s$ amount of time if it demands Z_s shopping trips. Given a household's total demand for market goods, its shopping trip need is

$$Z_s = \frac{X_m + X_h}{I}. \tag{3.3}$$

The conventional household's monetary budget constraint is

$$p_m X_m + p_h X_h + X \leqslant V, \tag{3.4}$$

where p_m and p_h are the prices of market goods and V is the monetary income. A household's available time budget for shopping and meal preparation home production activities is T, and its time constraint is

$$T_m + T_h + T_s \leqslant T. \tag{3.5}$$

Combining the conventional budget constraint, Equation (3.4), with the individual's time constraint, Equation (3.5), the total budget constraint becomes

$$\sum_{j \in \{m,\, h\}} \left(p_j X_j + \omega T_j \right) + \omega T_s + X \leqslant \omega T + V, \tag{3.6}$$

where ω is a household's wage, which is used as a proxy for the household's opportunity cost of time. The left hand side is the total expenditure on activities and the utility from residual income. So, "time is money", as the expenditure on each activity has two components—the money cost of the goods used for this activity and the opportunity cost of the time used for this activity. The right hand side represents the total income of the household.

3.3.2 Model solution and predictions

We now present the model solutions. We proceed by assuming the residual income (the outside goods) is positive. Substituting in Equation (3.3), we can write the household's optimization problem as

$$\begin{aligned} \max_{Z_m,\, Z_h,\, X} \mathscr{L} &= U\left(Z_m,\, Z_h,\, X\right) + \lambda \Big(\omega T + V - \omega \tfrac{X_m + X_h}{I} \\ &\qquad - \pi_h Z_h - \pi_m Z_m - X \Big) \end{aligned}.$$

Here

$$\begin{aligned} \pi_h &= p_h c_h + \omega b_h; \\ \pi_m &= p_m c_m + \omega b_m \end{aligned} \tag{3.7}$$

represent the combined time and market goods cost of producing each unit of consumable commodity.

The Kuhn-Tucker first-order conditions are

$$\begin{gathered} \frac{\partial \mathscr{L}}{\partial Z_h} = \frac{\partial U}{\partial Z_h} - \lambda \pi_h \leq 0, \quad Z_h \geq 0, \quad \text{and } Z_h \frac{\partial \mathscr{L}}{\partial Z_h} = 0; \\ \frac{\partial \mathscr{L}}{\partial Z_m} = \frac{\partial U}{\partial Z_m} - \lambda \pi_m \leq 0, \quad Z_m \geq 0, \quad \text{and } Z_m \frac{\partial \mathscr{L}}{\partial Z_m} = 0; \\ \frac{\partial \mathscr{L}}{\partial X} = \frac{\partial U}{\partial X} - \lambda \leq 0, \quad X > 0, \quad \text{and } X \frac{\partial \mathscr{L}}{\partial X} = 0; \\ \omega \frac{X_m + X_h}{I} + \pi_h Z_h + \pi_m Z_m + X = \omega T + V, \end{gathered} \tag{3.8}$$

where $Z_h \frac{\partial \mathscr{L}}{\partial Z_h} = 0$, $Z_m \frac{\partial \mathscr{L}}{\partial Z_m} = 0$ and $X \frac{\partial \mathscr{L}}{\partial X} = 0$, are known as "complementary slackness" conditions, indicating that the constraints are binding whenever the demands are nonzero. From (3.8), we see that positive demand for Z_h or Z_m is associated with the equality constraint from the Kuhn-Tucker first-order condition, whereas zero demand results in an inequality condition. Further, as we assume the residual income (the outside goods) X has positive demand, the Lagrange multiplier λ is positive.

By combining the first three first order conditions, we see that the optimum $\left(Z_h^*, Z_m^*, X^*\right)$ occurs at a point where the indifference curve between any two activities and the budget constraint have the same slope

$$\begin{aligned} \frac{\partial U}{Z_h} \Big/ \frac{\partial U}{\partial Z_m} &= \frac{\pi_h}{\pi_m}; \\ \frac{\partial U}{X} \Big/ \frac{\partial U}{\partial Z_m} &= \frac{1}{\pi_m}. \end{aligned} \tag{3.9}$$

The fourth first order condition tells us the optimum must lie on the budget constraint. Taking these conditions together, we see that the optimum lies at the tangent point between the indifference curve and budget constraint. This simple model directly leads to the following testable predictions.

Total demand for grocery goods and shopping trips An increase in available time T implies that the household's budget shifts outward. Therefore, the household can afford producing more commodities, Z_m and Z_h, and subsequently the household also demands more shopping trips $Z_s = \frac{c_m Z_m + c_h Z_h}{I}$.

Household preference for time-intensive goods versus non-time-intensive goods Recall that the optimum $\left(Z_h^*, Z_m^*\right)$ occurs at a point where $\frac{\partial U}{Z_h} / \frac{\partial U}{\partial Z_m} = \frac{\pi_h}{\pi_m}$. Let $R = \frac{\pi_h}{\pi_m}$ denotes the relative price. Increase in the relative price of a certain product means a household will substitute away from that product. The partial derivative of the relative price with respect to the wage is

$$\frac{\partial R}{\partial \omega} = \frac{p_m c_m b_h - p_h c_h b_m}{\left(p_m c_m + \omega b_m\right)^2}. \tag{3.10}$$

Wage will increase the relative price if $p_m c_m b_h - p_h c_h b_m > 0$. This leads to the following condition for a positive effect of wage on the relative price of Z_h

$$\frac{b_h}{b_m} > \frac{p_h}{p_m} \frac{c_h}{c_m}. \tag{3.11}$$

Intuitively, this condition implies if the commodity Z_h is sufficiently more time intensive to produce than commodity Z_m, an decrease in wage will lead the household to substitute from Z_m

to Z_h. In reality, a wage decrease can be caused by retirement or unemployment. It might also be correlated with a reduction in work hours (e.g., occupation change from a full time job to a part time job, retirement, and unemployment).

Demand for variety The results presented so far follow directly from our model. Towards thinking about household behavior in the presence of fixed costs of household demand for variety, we now develop a set of predictions that, for reasons of tractability, are not explicitly derived. If we take as a starting point that the total demand for goods increases and that households have a taste for variety, but buying additional variety comes at a fixed cost as in Bronnenberg (2015); Hamermesh (2005); Bronnenberg (2017), then our first conjecture is that increasing the amount of time a household has available will also lead to the household buying more varieties, because the household will consume more in total, which means that the fixed cost per consumed unit will fall.

3.4 Data

3.4.1 Product characteristics

Our data originate from three sources.[1] First, we collect product characteristics data at the barcode level from the GfK's SKU directory. Each product is identified by a unique Universal Product Code (UPC). For each UPC, the barcode file provides a detailed product description. In total, there are 189,619 UPCs of food products,[2] which are classified into 41 food product

[1] All three data sets were provided by AiMark.

[2] We further exclude baby food and special diet food, as the demand for products in those categories is likely to correlate with retirement, unemployment and parenting event.

categories by GfK. We supplement the product description data with a classification of the degree of convenience of each category. Appendix 3.7.1 contains the survey. The classification of categories that is based on this survey can be found in Appendix 3.7.2.

3.4.2 ConsumerScan purchase data

Second, we use data from GfK's ConsumerScan panel, covering grocery purchases for a national sample of Dutch consumers from the beginning of 2009 until the end of 2013. Each household in the purchase panel is given a handheld device to scan all the barcodes of products that were purchased. Households record the barcode, from which retailer the product is purchased, and during which part of the day[3] of a specific date the transaction takes place.

The panel covers Dutch residents from various areas across the country and household purchases from all the active supermarkets. From these data, we construct a balanced purchase panel of 4,358 households who are active across all years.[4] GfK provides weekly monetary incentives to households to keep them remain active.

Unless mentioned otherwise, we report on the descriptive statistics of the balanced panel. Table 3.1 shows that total expenditure that was scanned in a given year is €2,843 on average. Of this, €1,074 is purchased from time intensive categories (e.g., fresh cuts, vegetables, baking and dessert ingredients) and €730 is purchased from goods intensive categories (e.g., meal, soup and broth). A typical household purchases 632 different products in a given year and makes 126 different shopping trips. From this we conclude that our balanced panel of households scan regularly and make a shopping trip to the grocery store about every three days on average.

[3]Each day contains three day parts in GfK's definition.
[4]The unbalanced panel contains 12,483 households.

Table 3.1: Descriptive statistics of purchases

variable	mean	10th percentile	90th percentile	standard deviation	number of observations
total expenditure	2570	943	4515	1448	34259
time intensive expenditure	706	209	1318	458	34259
goods intensive expenditure	612	208	1130	380	34259
number of varieties	597	274	965	272	34259
number of trips	122	46	222	70	34259

Note: All expenditure measures are in Euros per year. The number of varieties refers to the number of unique items in a household's annual shopping basket. The number of trips refers to the trips per year, whereby a trip is defined as a unique combination of household id, date, and daypart. Reported descriptive statistics are taken over households in the balanced panel and years 2009 to 2013.

3.4.3 GfK annual survey of panelists

Third, we use data from an extensive annual survey of Dutch GfK panelists in the Consumer-Scan panel over the years 2009-2013. GfK distributes the consumer survey to all the panelists every year in December and asks each member of the household to respond. GfK collect the answers of the survey in January of the next year. The survey contains information on the household in the previous year.

The survey includes information on the demographics, such as the head of the household's age, gender, education, occupation, employment status, family composition, and household monthly income. Table 3.2 gives descriptive statistics for four demographic variables: age, income, household size, and education, conditional on survey respondents being in the balanced panel or in the full unbalanced panel. The eldest person in the household is 51 years of age on average. Net family yearly income from market work is €26,800 and a typical household consists of 2.86 persons. The balanced sample has a very similar demographic profile as the unbalanced sample but is slightly older, poorer, larger, and less educated.

Table 3.2: Descriptive statistics households

	variable	mean	10th percentile	90th percentile	standard deviation	number of households
unbalanced	age	51.09	32.00	71.50	14.92	13170.00
	income	26.80	13.44	40.80	10.73	10108.00
	size	2.86	1.00	5.00	1.48	13170.00
	education	9.40	4.00	14.00	3.14	13092.00
balanced	age	57.15	39.80	75.00	13.60	4512.00
	income	26.04	14.40	38.88	9.61	4496.00
	size	3.11	1.00	5.00	1.51	4512.00
	education	8.87	4.00	13.00	2.91	4512.00

Note: Reported descriptive statistics are taken over households and years 2009 to 2013. Age refers to the age of the household head. Household income is reported in 1000 Euro. Size is the number of reported members in the household. Education is the number of years of education after elementary school of the household head.

Table 3.3 gives survey responses, for the balanced panel, on various shifters of time availability and self-reported time allocations to home production and market work. A typical household in our sample has a high likelihood of containing at least one retiree (31%). This is slightly higher than the national average of 27%.[5] About 5% of the households report to have at least one unemployed individual. About 5% of the households report to have at least one baby (age younger than 4 years) in the house and about 15% report to have children (age 4 years and up) in the household. Combined with the reported aged statistics, this suggests that the panel over-samples older individuals.

In addition to these indirect measures of time availability, the GfK survey also contains direct measures of time allocations and budgets. In particular, the survey asks how many hours a week are spent in the labor market and how many are spent doing work in the home, for

[5]The Central Bureau for Statistics (CBS) in the Netherlands, estimates that for 2017 a total of 2.12 million out of 7.81 million households contains a household member of retirement age. See, https://www.cbs.nl/nl-nl/nieuws/2015/51/tot-2040-verdubbelt-het-aantal-alleenwonende-tachtigplussers.

cooking, cleaning, shopping, etc. We label these tasks all as home production, but realize that not all of it is related to grocery purchases or use thereof.

Table 3.3 reports on the total household sum of these direct measures of time budgets. The average household in our sample reports to spend 27 hours a week in home production. There is appreciable heterogeneity in home production with the first decile of household years being at 8 hours home production and the ninth decile at 48 hours. A typical household works 49 hours in the market. Across household variation is very large, with the 10th decile not working in the labor market at all, and the 90th docile spending more than a 100 hours.

Table 3.3: Descriptive statistics of time resource shifters and time allocations

variable	mean	10th percentile	90th percentile	standard deviation	number of observations
retirement	0.31	0.00	1.00	0.46	22560
unemployment	0.05	0.00	0.00	0.22	22560
baby	0.05	0.00	0.00	0.21	22560
child	0.15	0.00	1.00	0.35	22560
total hours home production	27.14	8.00	48.00	15.52	22560
total hours market work	48.62	0.00	100.50	38.25	22560

Note: Retirement, unemployment, baby, and child are dummy variables. Time budgets are the sum of reported time allocations across adult members of the household. Reported descriptive statistics are taken over households in the balanced panel and years 2009 to 2013.

Table 3.4 confirms the heterogeneity in self-reported time allocations but additionally shows the decomposition of time allocations into within and across household variation. The results suggest that 14% of the variation in time used for home production is within household. For market work, this is 5%. This means there is ample within-household variation in hours worked in the labor market or at home, which is important for our identification strategy.

Table 3.4: Variance decomposition of time budgets

	household component	time component	household ×time component
total hours home production	85.98%	0.03%	13.99%
total hours market work	94.89%	0.01%	5.10%

Note: Reported are the percentages of variance in the time budgets explained by household dummies and by time dummies. The remainder of variation is interpreted as within-household across-time variation.

The intersection of the purchase data and the survey data contains 4,358 households. Among these households, there are 1,832 households in which at least one person retired or became unemployed. For each household, we have information on the year a time-budget shifting event took place. Therefore, for our analysis we aggregate the purchase choice data to the household-year level.

To provide some final descriptive statistics, we regress, controlling for household and time fixed effects, reported time allocations on time-shifting events like unemployment and retirement. We expect an unemployment event and an retirement event to be associated with less household level work-time and more reported time available for home production. For completeness, we report these results for market work hours and home work hours and for both the balanced and unbalanced sample.

From the first row in Table 3.5, we see that an additional unemployed individual in the household is associated with 21.4 hours less work in the market and 1.4 hour more in home production per week (balanced sample). The effect of retirement on hours worked in the market, is less strong—15.0 hours less work in the market and 1.6 hour more in home production per week—, perhaps because the employed scale down their labor participation prior to retirement. The number of children does not affect hours in the labor market or hours spent in home production nearly as much,[6] but the effects are significant and in the expected direction. These results hold equally for the unbalanced sample. Thus, unemployment, retirement, and having children are all associated with less hours worked in the market and more hours worked in home production, holding household and time effects fixed.

Table 3.5: The relation between time allocations and household events

	unbalanced		balanced	
variable	weekly market work hours	weekly home work hours	weekly market work hours	weekly home work hours
unemployed	-21.42	1.41	-21.94	1.10
	(0.36)	(0.26)	(0.44)	(0.33)
retired	-14.91	1.19	-15.29	0.92
	(0.31)	(0.22)	(0.34)	(0.26)
baby	-0.15	0.66	-0.53	-0.10
	(0.34)	(0.24)	(0.47)	(0.35)
child	-0.85	0.89	-1.44	0.57
	(0.24)	(0.17)	(0.30)	(0.22)
baseline	50.90	25.55	55.88	26.42
	(0.18)	(0.13)	(0.21)	(0.16)
household fixed effects	X	X	X	X
time fixed effects	X	X	X	X
R^2	0.964	0.904	0.959	0.862
N	39657	39657	22550	22550

Note: Standard errors in parentheses. The variables unemployed, retired, baby, and child refer to their number in a household-year.

3.4.4 Dependent variables

We focus on grocery purchases and are interested in investigating to how and to what extend time affects a household's use of the market. To begin with, we construct a set of measures that systematically summarize household purchase behaviors. We relate the measures with the

[6] Colella and van Soest (2013), Hurd and Rohwedder (2005), Kimmel (2008), Pashigian and Bowen (1994) and Stancanelli and van Soest (2012) suggest that households face a reduction in their disposable time budget after the birth of a baby, and face time budget expansion after retirement or unemployment.

theory predictions and discuss the hypotheses. With respect to shopping activity, we expect to see that households shop more when they have more time for home production as time is a complement of the goods input in the home production process. An alternative explanation is that households are motivated by profiting from temporal variation in prices. Given more time to shop, the consumer can convert some of the additional time into money savings when prices vary over time. Accordingly, we measure shopping activity with two measures: the annual number of shopping trips, and the share of annual purchases bought on a price promotion.

With respect to the total demand for grocery goods, we expect that households buy more grocery goods when they have more time available for home production. Comparing with eating out, buying groceries and undertaking home productions are more time intensive. We therefore expect households with more time to expand their quantity demand for grocery products which may also be associated with an increase in the grocery expenditures. We measure total demand for grocery goods as the annual expenditure on grocery goods and the annual purchased units.

With respect to variety, we expect that households who are endowed with more time, not only find lower prices but also expand their basket in terms of variety. We measure variety as the number of unique SKUs in the annual shopping basket and as the number of unique product (sub) categories.

With respect to shifting to more time intensive activities, our theory prediction suggests that households who have lower opportunity cost of time buy more time-intensive products because they are more likely to undertake time-intensive activities. We measure the time-intensity of a household's shopping basket as the ratio of annual purchased amount of time-intensive goods versus the annual purchased amount of non-time-intensive goods. We measure the preference for convenient goods versus time-intensive goods preferences as the total quantity share of households' purchases of convenient categories and time-intensive categories. To measure time intensity of a category, we have used a preliminary survey of 22 graduate students at Tilburg University, who individually rated the time intensity of all the grocery categories used in this study. Averaging scores across respondents, we rank grocery categories on average reported time intensity and classify the top third as time intensive, the bottom third as goods intensive and the remainders as neutral. We measure the time intensity of a household's annual shopping basket as the ratio between the expenditures to time intensive and the expenditures to goods intensive categories. As an alternative measure we use units instead of expenditures.

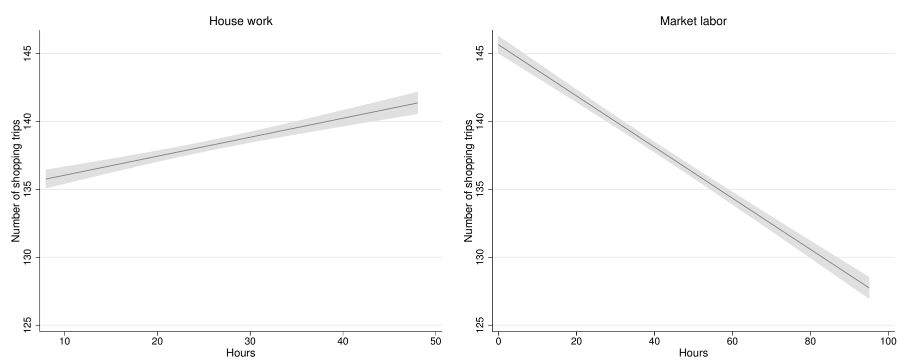
3.5 Empirical analysis

3.5.1 Preliminary evidence

We provide two graphs to show some preliminary evidence of a relation between time use and purchase behavior. More specifically, our aim is to visualize the relation between time availability and purchase behavior, net of household fixed effects. We, therefore, first estimate a regression of purchase behaviors as a function of household fixed effects, time fixed effects, and demographics. We then remove the estimated household fixed effects from the data. What remains are the raw data net of household fixed effects. Next, we plot the relation between the reported time spent in home production or market labor and purchase behavior net of these household fixed effects.

In Figure 3.1, we depict the relation between reported hours worked (either in home production or labor) and the number of shopping trips, controlling for household fixed effects. An additional hour in the labor market is associated with less shopping trips and an additional hour of household work is associated with more shopping trips. This is consistent with the theory that housework hours are complements of purchased market goods in home production activities. More home production hours are associated with higher demand for market goods, and, therefore, higher number of shopping trips. Figure 3.2 shows a similar result for the relation between reported hours worked and the amount of variety that consumers buy. An additional hour in the labor market reduces the amount of variety bought and an additional hour worked in home production increases it.

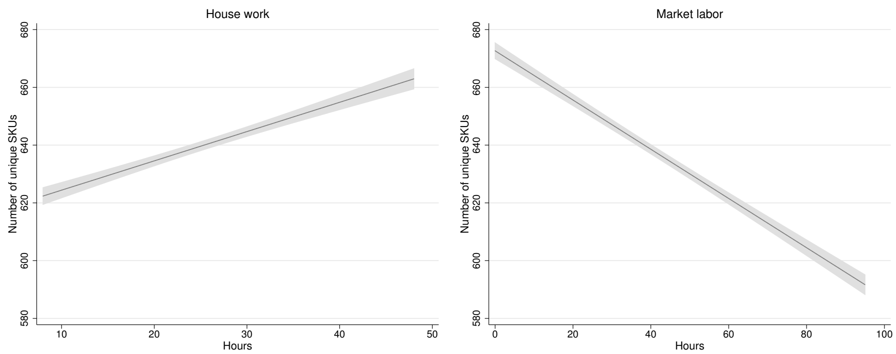
Figure 3.1: Shopping activity as a function of hours reported in house work and market labor versus



Notes: Relation between hours reported by consumers and the number of shopping trips households took. We control for household fixed effects.

Obviously, the relation depicted in the plots is not free from the spurious effects of the relation between hours worked and consumer demographics. For instance, the difference between the effects of hours worked at home versus in the labor market, and our possible explanation, motivates that in our empirical analysis we should account for income shocks. To separate the

Figure 3.2: Basket variety as a function of hours reported in house work and market labor versus



Notes: Relation between hours reported by consumers and the number of unique SKUs purchased annually by households. We control for household fixed effects.

spurious effects from the effects of interest, i.e., the relation between reported time availability and purchase behavior, we use the empirical approach explained next.

3.5.2 Empirical model

Our main empirical analysis exploits the two types variations –household self-reported time allocations and shifter events– to explain the purchasing outcomes of interest: (1) shopping frequency, (2) total grocery purchases, (3) the time intensity of the goods purchased, and (4) variety purchased. The empirical strategy is to use a regression framework, controlling for household fixed effects and making use of the balanced individual level data and a rich set of controls. Consider purchase outcomes y_{it}, e.g., total grocery expenditure, for household i in year t, we model these purchase outcomes as a function of a household fixed effects α_i, effects β of work-time or shifter event x_{it}, control effects γ, and time fixed effects δ_t,

$$y_{it} = \alpha_i + x_{it}\beta + z_{it}\gamma + \delta_t + \varepsilon_{it}, \tag{3.12}$$

with ε_{it} i.i.d. across i and t.

Our first empirical strategy applies to the whole sample of households for whom we observe the variations in the self-reported weekly work hours x_{it}. More work hours implies less available time for home production activities, vice versa. We investigate the relation between a household's available time for home production and its total grocery purchases, number of shopping trips and demand for variety, respectively. Our second empirical strategy applies to a sub-sample of households who have experienced retirement or unemployment during our observation window. A shifter event, i.e., retirement and unemployment, not only shifts a household's available time budget for home production but also leads to a decrease in the household's opportunity cost of time (proxied by wage), which allows us to further investigate the relation between the opportunity cost of time and the time intensity of a household's shopping basket. We then present a full set of results where x_{it} are dummies signifying unemployment or retirement.

Household fixed effects α_i in fact account for any unobserved taste differences that are constant through time but vary across households, including unobserved heterogeneity in shopping preferences, cooking preferences, etc. The variables z_{it} control for the effects of changes in demographic variables. In particular, we include the effects of household income, maximum age in the household, household size, and the maximum education level in the household, as those factors may vary over time and could have influence on a household's shopping outcomes. Finally, we control for time fixed effects to allow for general trends, such as the effect of the economic recession on purchasing of grocery items or supply side innovations in providing more convenience goods. Parameters are estimated by ordinary least squares (OLS). We think of the resulting estimates of β as causal, as the changes in x_{it} are viewed as exogenous to the household's purchase behaviors.

3.5.3 Main results

3.5.3.1 The relation between available time for home production and shopping activity

First focusing on the relation between time availability and shopping activity, Table 3.6 shows how the reported time spent at work affects the number of shopping trips made by a household. We view the number of shopping trips as a measure of the willingness to incur travel costs to access the market goods and view the change in a household's work hours as a shock to its available time for home production activities.

We find that the effect of the number of hours worked on the tendency to shop is strongly negative and significant. We also find, like Aguiar and Hurst (2005), that the age of a household (measured as the age of the oldest household member) has a strongly positive effect on the number of trips that are taken, adding an additional shopping trip with every year. Household size also positively impacts the number of trips to the store, but income and education do not.

We also follow up on Aguiar and Hurst (2005) and look at a possible motivation to take more shopping trips, namely to find more attractive prices by being able to monitor the basket items on deal more frequently. For this purpose, we report the effect of time on the fraction

Table 3.6: The relation between time allocations and shopping activity

variable	number of trips	share expenditure on deal
hours worked in the market	-0.150	-0.000
	(0.018)	(0.000)
baseline	85.983	0.124
	(17.850)	(0.041)
age	1.049	0.000
	(0.326)	(0.001)
income	-0.018	-0.000
	(0.034)	(0.000)
household size	2.834	-0.001
	(0.992)	(0.002)
education level	0.074	0.001
	(0.162)	(0.000)
household fixed effects	X	X
time fixed effects	X	X
R^2	0.942	0.895
N	19348	15769

Note: Standard errors in parentheses. Number of trips refer to the number per household year. Share of expenditure on deal is the part of recorded expenditure on deal (between 0 and 1).

of grocery products purchased on a price promotion. From the estimate of the constant in this model, consumers make about 12.4% of their grocery purchases when they are on a price promotion. However, we do not find the availability of time to have a strong impact on the share of grocery expenditures that is associated with price promotion. Further, none of the household demographic effects has a significant impact on this share of expenditure. Our findings suggest that the tendency to buy on promotion is a household trait that is insensitive to time allocations or changes in demographics in this panel with consumers in a developed European country.

Having presented evidence on the influence of a household time on its shopping outcomes with household self-reported work-hour changes, we then exploit household level time-budget shifting events to demonstrate the relation between a household's time and it shopping outcomes. We use a sub-sample of households who have experienced a retirement event or an unemployment event during our observation window. A retirement event or an unemployment event leaves a household more time for home production. Consistent with our theory predictions, we find that the number of retirees in a household and the number of unemployed members in a household both have strongly positive effects on the number of shopping trips made by a household. Thus, part of the time-budget windfall from spending less time working is used to make more shopping trips (see Table 3.7).

Table 3.7: The relation between time-budget shifting events and shopping activity

variable	number of trips	share expenditure on deal
number of retirees in household	2.693	0.002
	(0.867)	(0.002)
number of unemployed in household	3.274	-0.001
	(1.174)	(0.002)
baseline	87.690	0.134
	(17.899)	(0.036)
age	0.963	-0.000
	(0.328)	(0.001)
income	-0.033	-0.000
	(0.034)	(0.000)
household size	1.155	-0.000
	(0.980)	(0.002)
education level	0.075	0.000
	(0.163)	(0.000)
household fixed effects	X	X
time fixed effects	X	X
R^2	0.941	0.910
N	19348	15769

Note: Standard errors in parentheses. Number of trips refer to the number per household year. Share of expenditure on deal is the part of recorded expenditure on deal (between 0 and 1).

3.5.3.2 The relation between available time for home production and demand for grocery goods

Second, we hypothesize that time allocations affects the total demand for market goods—the demand for grocery goods increases with time availability. For this, we offer empirical supports on both the amount of money a household spends on the grocery goods and the total purchased quantities. Viewed as a more time intensive and less money intensive substitute for eating out, the effect of time on grocery expenditure speaks directly to its importance as an input to consumption, especially when controlling for income.

Table 3.8 shows that the effect of time spent working on grocery expenditure is negative. The effect is statistically and substantively strong even when controlling for (changes in) demographics and income. Further, expenditure on grocery goods is positively impacted by age, income, and household size, as one would expect.

Table 3.9 shows a similar pattern with a sub-sample of households who have experienced a retirement event or an unemployment event. Retirement makes consumers spend more on grocery shopping. The difference is estimated to be large. A 70-year-old retiree spends 287 (= $46.81 + 10 \times 24.02$) Euros more in a given year than a 60-year-old-worker, and buys 227 (= $44.45 + 10 \times 18.28$) more units.

The same conclusions hold for purchased quantities as opposed to expenditures. This can be related to the idea that many time costs of home production are fixed to quantity (Bronnenberg,

Table 3.8: The relation between time allocations and grocery goods purchasing

variable	euros	units
hours worked in the market	-1.730	-3.797
	(0.371)	(0.405)
baseline	1078.872	1531.761
	(368.912)	(402.834)
age	25.574	20.398
	(6.746)	(7.366)
income	4.659	2.607
	(0.697)	(0.761)
household size	148.146	181.886
	(20.499)	(22.384)
education level	-1.961	-0.621
	(3.354)	(3.662)
household fixed effects	X	X
time fixed effects	X	X
R^2	0.940	0.931
N	19348	19348

Note: Standard errors in parentheses.

2015), eroding a continuous relation between home production time allocations and quantity demand.

We conclude that the demand for grocery items increases with the time a household has for home production, which is just as the new consumer theory (e.g., Becker 1965; Gronau 1977) –the requirement of time as an input to home production– predicts.

3.5.3.3 The relation between available time for home production and the time intensity of shopping baskets

Third, we study whether, in addition to buying more, the time intensity of shopping baskets is also correlated with a household's opportunity cost of time. Our theory prediction suggests that a decrease in the household's opportunity cost of time leads to higher time intensity of the household's shopping basket. We first classify the categories as being time intensive, neutral,

Table 3.9: The relation between time-budget shifting events and grocery goods purchasing

variable	euros	units
number of retirees	46.805	44.445
in household	(17.889)	(19.578)
number of unemployed	-26.880	-41.282
in household	(24.220)	(26.506)
baseline	1120.998	1571.788
	(369.409)	(404.279)
age	24.017	18.283
	(6.759)	(7.397)
income	4.492	2.189
	(0.696)	(0.762)
household size	130.663	144.913
	(20.221)	(22.130)
education level	-1.963	-0.587
	(3.355)	(3.672)
household fixed effects	X	X
time fixed effects	X	X
R^2	0.940	0.931
N	19348	19348

Note: Standard errors in parentheses.

or goods intensive, and use the ratio of the time intense and goods intense portion of the annual shopping basket as a dependent measure (see Section 3.4.4 for definitions).

Table 3.10 gives the results. We find the quantity ratio to time intensive goods rises (non-significantly) in the reported hours spent on home production and drops in the reported hours spent working in the labor market. Consistent with our theory prediction, this effect replicates with and without controlling for non-labor income. These results clearly suggest that households that have low opportunity cost of time (proxied by wage), buy shopping baskets that are more time intensive. We conclude that time intensity of the grocery basket is shifted by the household's opportunity cost of time.

Among the demographic effects, aging leads a household to buy more from time-intensive

Table 3.10: The relation between time-budget shifting events and time intensity of basket

variable	time intensity of shopping basket (units)	time intensity of shopping basket (units)
number of retirees in household	7.014	7.425
	(3.750)	(3.858)
number of unemployed in household	-8.946	-5.832
	(4.909)	(5.224)
baseline	171.238	197.661
	(77.149)	(79.675)
age	5.452	5.158
	(1.422)	(1.458)
income	X	0.689
	X	(0.150)
household size	21.281	15.678
	(4.104)	(4.361)
education level	0.027	0.567
	(0.685)	(0.724)
household fixed effects	X	X
time fixed effects	X	X
R^2	0.908	0.923
N	22306	19348

Note: Standard errors in parentheses. Time intensity of the shopping basket is measured as the ratio of expenditures of time intensive over goods intensive categories. The definition of time intensive and goods intensive categories is based on a survey to consumers, as discussed in Section 4 of the paper.

categories relative to goods-intensive categories, perhaps as a consequence of a shift in taste towards fresh/unprocessed food. The same is true for income. Household size has a positive relation with the time-intensity of a household's shopping basket, which might because of the fact that the production process of time-intensive activities is more likely to have economies of scale.

3.5.3.4 The relation between available time for home production and purchased variety

Finally, we study the extent to which time allocations changes a household's demand for variety. A strong relation between time component and purchasing more variety indicates in our view that variety is costly in time to evaluate in the store, purchase, and use at home for a given household (see also Huang and Bronnenberg 2017).

Table 3.11 shows that the effect of the number of hours worked in the market has a large negative impact on the amount of variety bought by a household. Thus, the more hours a household spends in the labor market, the fewer varieties are bought in the grocery channel. We find that higher age, income, and larger household size increases the amount of variety strongly, both in a statistical and in a substantive sense. The reported effects of hours worked support the idea that (holding demographics fixed) the availability of time shifts the number of varieties that consumers buy.

The same pattern of effects completely replicates for the number of unique categories that consumers buy. The effect of hours worked in the house is positive but noisy and the effect of hours worked in the labor market is negative and significant.

Table 3.12 further confirms the findings above and shows the effect of retirement and unemployment on variety. Consistent with variety being costly to buy at the margin in terms of time, retirees buy more varieties than people who have not yet retired. To give an idea, a 70-year-old retiree buys 61 (= 7.580 + 10 $\times$ 5.304) varieties more in a given year than a 60-year-old worker.

Table 3.11: The relation between time allocations and variety

variable	number of unique SKUs	number of unique categories
hours worked in the market	-0.241	-0.046
	(0.069)	(0.010)
baseline	304.327	135.895
	(68.640)	(10.210)
age	5.540	0.706
	(1.255)	(0.187)
income	0.578	0.063
	(0.130)	(0.019)
household size	12.566	2.057
	(3.814)	(0.567)
education level	-0.046	0.013
	(0.624)	(0.093)
household fixed effects	X	X
time fixed effects	X	X
R^2	0.937	0.933
N	19348	19348

Note: Standard errors in parentheses.

The analysis replicates for the number of unique categories, although with a lower effect size.

We conclude that hours worked in the labor market significantly shifts the number of varieties that people buy, holding constant the number of hours spent on domestic tasks and changes in demographics. This suggests that market labor competes with some time input that is not reported as hours spent on domestic tasks, but which, nonetheless, shifts preference for variety in purchase behavior.

Taken together, with both the household self-reported work-hour information and the time-budget shifting events such as retirement and unemployment, we find that more available time for home production activities impact purchasing behaviors in the direction of making more shopping trips, increasing grocery expenditures, demanding more grocery goods and buying

Table 3.12: The relation between time-budget shifting events and variety

variable	number of unique SKUs	number of unique categories
number of retirees in household	7.580	0.160
	(3.328)	(0.056)
number of unemployed in household	-2.731	0.055
	(4.505)	(0.076)
baseline	310.965	34.013
	(68.717)	(1.162)
age	5.304	0.052
	(1.257)	(0.021)
income	0.556	0.005
	(0.130)	(0.002)
household size	10.054	0.141
	(3.762)	(0.064)
education level	-0.047	-0.002
	(0.624)	(0.011)
household fixed effects	X	X
time fixed effects	X	X
R^2	0.937	0.830
N	19348	19348

Note: Standard errors in parentheses.

more varieties. We also find that lower opportunity cost of time leads to higher preference for time-intensive goods. Hence, we conclude that time has profound effects on household purchase behavior.

3.6 Concluding remarks

This paper shows direct evidence that shifts in the amount of time a household has available for home production and the opportunity cost of time have important effects on household purchasing behavior. Using individual level survey panel data and scanner purchase panel data on the same sample of households from 2009 to 2013, we find that there is a negative co-movement between a household's available time and the amount of shopping trips, the grocery expenditures, purchased quantities and varieties. We further show that there is a negative relation between a household's opportunity cost of time and the time-intensity of its shopping basket.

Our first empirical strategy uses the whole sample of households for whom we observe both their grocery purchase decisions and reported work-time information. The self-reported work-time information provides ample within-household variance in the available time for home production. To confirm our findings with direct information on household time use, we use a sub-sample of households who reported to have experienced time budget shifting events, i.e., retirement and unemployment, and compare household purchase behaviors before and after the events. Retirement and unemployment not only provide exogenous shocks to the households' available time for home production, but also create exogenous variations to the household's opportunity cost of time. We further test whether there is a relation between a household's opportunity cost of time and the time-intensity of its shopping basket with this sub-sample of households.

Overall, the evidence presented in this paper suggests that more work is needed to understand whether and how structural demand estimates are biased if time availability and changes thereof over time are not incorporated in the model. In this paper, we show that changes in time availability have large effects on consumer behavior. This suggests that including time availability into a structural model would lead to more accurate counterfactual predictions that could be useful to firms optimizing their marketing strategies.

Moreover, our findings relate the time use literature, that has documented a trend to less available time for home production (see Aguiar and Hurst (2007b) and Gimenez-Nadal and Sevilla (2012) for multi-national evidences), to the optimal supply side activities. The trend moving to less available time for home production implies a greater demand for convenience and more sensitive to the cost of using the market. Therefore, one of the new goals for the firms is to make the market more accessible to the consumers by subsiding along the time dimension.

Specifically, our finding on the number of shopping trips supports the idea that one way for retailers to compete for consumers by investing in reducing consumer's cost in accessing the market goods (e.g., online channel, delivery service or extending store opening hours). Another way to subsidize the "time scarce" consumers is by providing higher ratio of convenience goods, which is supported by our findings regarding to the time-intensity of household's shopping basket. Lastly, subsiding consumers with "time" may increase the effectiveness of other retailer policies. For example, variety expansion may not be attractive to the consumers if the consumers' demand for variety is restricted by their time.

Our study has limitations that open interesting opportunities for future extension. First, in our current analysis, we assume the effect of time allocation on purchase outcomes is homogeneous across households. One interesting extension is to investigate whether households respond to time allocation change differently and which segment is the most "time-sensitive" segment. Second, we use household size as a proxy for the household's shopping needs. In future extensions, we will make use of household composition information to construct a more robust proxy for the household's shopping need.

3.7 Appendix

3.7.1 Original Survey

This survey asks you about your perceptions of how “time intensive” various product categories are in terms of converting them to something you can consume. First, we explain the meaning of the “time intensiveness” of a product group. Next, you will be asked to evaluate the time intensiveness of 30 product groups.

Households are engaged in various activities or consumption moments at home. A household “produces” consumption goods by combining goods bought in a store with personal time. Whereas some activities can be produced by the household using market purchased goods and time directly, others require households to produce an intermediate good first. In the survey below, you are asked to judge the degree to which a particular product group is “time intensive,” i.e., the degree to which a product groups requires extra time and effort to be turned into a consumption experience above and beyond the time required shopping for such products.

Below, you will see a list of 31 product groups. The task is to rate on a 9-point scale how time-intensive each group is based on your own intuition. To facilitate this task examples of subcategories of these groups are listed along with the group names. When judging the time-intensiveness of each product group, it may help you to first image how you typically use this product group. Then you complete the table by circling the number that best describes the time intensiveness of that product group.

Product group	Not at all time intensive								Very time intensive
feminine hygiene and diapers *(diapers, baby tissues, ladies related, etc.)*	1	2	3	4	5	6	7	8	9
alcoholic beverages *(wine port, beer-lager, spirits, etc.)*	1	2	3	4	5	6	7	8	9
smoking products *(tobacco products otherwise)*	1	2	3	4	5	6	7	8	9
cosmetics hair *(shampoo, conditioner, hair spray, etc.)*	1	2	3	4	5	6	7	8	9
cookies *(biscuit, pancake, cake freezer, etc.)*	1	2	3	4	5	6	7	8	9
cleaners *(cleaner powder, toilet cleaner, liquid soap, etc.)*	1	2	3	4	5	6	7	8	9
chicken and poultry *(chicken and poultry, fresh, and frozen, etc.)*	1	2	3	4	5	6	7	8	9
potatoes *(peeled, frozen, raw potatoes, etc.)*	1	2	3	4	5	6	7	8	9
animal nutrition *(cat food, dog food, etc.)*	1	2	3	4	5	6	7	8	9
edible oils and fats *(margarine, cooking oil, butte ,etc.)*	1	2	3	4	5	6	7	8	9
milk/replacements *(dry milk product ,milk replacer, etc.)*	1	2	3	4	5	6	7	8	9
cold cuts *(fresh, frozen, etc.)*	1	2	3	4	5	6	7	8	9
ice *(ice cream powder, ice cream, etc.)*	1	2	3	4	5	6	7	8	9
personal care products *(teeth maintenance ,makeup items, etc.)*	1	2	3	4	5	6	7	8	9
sandwich toppings *(jam, ,chocolate butter, sandwich litter, etc.)*	1	2	3	4	5	6	7	8	9
cosmetics skin *(bath-shower products, moisturizing lotion ,etc.)*	1	2	3	4	5	6	7	8	9
fish and seafood/shellfish *(preserve, fresh, frozen fish, snail, etc.)*	1	2	3	4	5	6	7	8	9
taste enhancers *(syrup ,vinegar, salt powder, etc.)*	1	2	3	4	5	6	7	8	9
Non-alcoholic drinks	1	2	3	4	5	6	7	8	9

(soft drinks ,iced coffee, soft drink powder, etc.)									
soups and broths *(wet, dry and frozen soup, dry broth, etc.)*	1	2	3	4	5	6	7	8	9
cosmetics fragrances *(aftershave, deodorant, perfume, etc.)*	1	2	3	4	5	6	7	8	9
other confectionery *(foam cake, chocolate kisses , etc.)*	1	2	3	4	5	6	7	8	9
diet/health food articles *(health food products, diet product, etc.)*	1	2	3	4	5	6	7	8	9
bread *(frozen bread, fresh bread, etc.)*	1	2	3	4	5	6	7	8	9
breakfast products *(muesli, corn flakes, oatmeal, etc.)*	1	2	3	4	5	6	7	8	9
cakes and pastries *(fresh /frozen pastry/pie/cake, etc.)*	1	2	3	4	5	6	7	8	9
vegetable *(froze/ preserves/mixed/fresh vegetable)*	1	2	3	4	5	6	7	8	9
bread sticks/crackers *(hash dishes, cream crackers, snack bottoms, etc.)*	1	2	3	4	5	6	7	8	9
milk products *(cream, yogurt, cheese , custard, etc.)*	1	2	3	4	5	6	7	8	9
baking and dessert ingredients *(ready cake bottom ,instant pudding, etc.)*	1	2	3	4	5	6	7	8	9
meals *(meal package, fresh/frozen pizza, etc.)*	1	2	3	4	5	6	7	8	9

Product group	Not at all time intensive								Very time intensive
dental hygiene *(toothpaste, mouthwash, etc.)*	1	2	3	4	5	6	7	8	9
flowers and plants *(flower bulbs ,indoor plants, garden plants, etc.)*	1	2	3	4	5	6	7	8	9
cleaning products *(furniture cleaners, soda powder, etc.)*	1	2	3	4	5	6	7	8	9
bread substitutes *(breadcrumbs, crackers, etc.)*	1	2	3	4	5	6	7	8	9
animal products *(animals food ,insect control, etc.)*	1	2	3	4	5	6	7	8	9
cheese *(cheese, various kinds of cheese, etc.)*	1	2	3	4	5	6	7	8	9
laundry detergents / fabric softeners *(detergent powder, softeners powder, etc.)*	1	2	3	4	5	6	7	8	9
fruit *(fresh fruit, fruit mix-salads, dried fruit, etc.)*	1	2	3	4	5	6	7	8	9
paper products *(toilet paper, kitchen rolls, tissues, etc.)*	1	2	3	4	5	6	7	8	9
promo-nonfood *(candles, coffee filters, lamps, batteries, etc.)*	1	2	3	4	5	6	7	8	9
meal components *(salad-dressings, dips, ketchups, frozen gravy, etc.)*	1	2	3	4	5	6	7	8	9
meal enhancers *(dry meal mixes sauces, wet hot sauces, etc.)*	1	2	3	4	5	6	7	8	9
meat *(frozen/fresh meat, sausage, etc.)*	1	2	3	4	5	6	7	8	9
game meat *(fresh, preserves, frozen)*	1	2	3	4	5	6	7	8	9
salads *(vegetable salads, snack salads, etc.)*	1	2	3	4	5	6	7	8	9
pastry *(raising pastry, frozen/fresh apple dumpling, etc.)*	1	2	3	4	5	6	7	8	9
pickles *(pickles)*	1	2	3	4	5	6	7	8	9

maintenance materials *(cleaning tools ,work cloths, etc.)*	1	2	3	4	5	6	7	8	9
sugar and sugar products *(syrup, sugar, sweeteners, etc.)*	1	2	3	4	5	6	7	8	9
household items *(domestic general items / broom / mop))*	1	2	3	4	5	6	7	8	9
milk and dairy drinks *(drinking yoghurt, chocolate milk, etc.)*	1	2	3	4	5	6	7	8	9
sugar confectionery *(jelly beans, peppermint, etc.)*	1	2	3	4	5	6	7	8	9
hot drinks *(coffee, cocoa liquid, tea, cocoa powder, etc.)*	1	2	3	4	5	6	7	8	9
detergents *(softeners powder, dishwasher powder, etc.)*	1	2	3	4	5	6	7	8	9
environment fresheners *(toilet freshener, air fresheners, deodorizers , etc.)*	1	2	3	4	5	6	7	8	9
rice and pasta *(pasta, lasagna ,rice ,etc.)*	1	2	3	4	5	6	7	8	9
herbs and spices *(fresh, dried, frozen herbs, etc.)*	1	2	3	4	5	6	7	8	9
egg and egg products *(egg products, eggs)*	1	2	3	4	5	6	7	8	9
medical products *(disinfectants cloth, vitamins, etc.)*	1	2	3	4	5	6	7	8	9
chocolate *(chocolate candy, chocolate ,etc.)*	1	2	3	4	5	6	7	8	9
savory snacks *(nuts, ,rice snacks ,potato chips, popcorn, etc.)*	1	2	3	4	5	6	7	8	9

3.7.2 Classification of categories and products

Table 3.13: Classification of product groups

category name	convenient in home production	convenient to purchase	total number of alternatives	category classification
alcoholic beverages	102	192	21587	neutral
baking and dessert ingredients	610	322	4918	time-intensive
bread	3490	11848	24050	neutral
bread sticks/crackers	0	0	898	convenient
bread substitutes	0	0	1167	convenient
breakfast products	0	0	1381	neutral
cakes and pastries	737	3594	7247	neutral
cheese	0	4567	10723	neutral
chicken and poultry	6802	1519	8626	time-intensive
chocolate	0	2105	11170	convenient
cold cuts	3764	7144	11233	time-intensive
cookies	2106	5260	16073	convenient
diet/health food articles	0	0	2837	neutral
edible oils and fats	28	1513	2776	convenient
egg and egg products	0	52	991	time-intensive
fish and seafood/shellfish	1333	1763	6430	time-intensive
fruit	0	3725	12059	neutral
game meat	330	359	429	time-intensive
herbs and spices	0	0	5576	time-intensive
hot drinks	2	639	6347	neutral
ice	0	0	4142	neutral
meal components	630	0	5399	neutral
meal enhancers	25	235	5106	time-intensive
meals	2358	4749	7720	time-intensive
meat	14081	4766	21065	time-intensive
milk and dairy drinks	975	1919	2466	convenient
milk products	1227	5024	6118	convenient
milk/replacements	49	324	1006	neutral
non alcoholic drinks	176	627	12746	convenient
pastry	30	238	1077	neutral
pickles	0	0	1315	convenient
potatoes	805	117	3843	time-intensive
rice and pasta	810	347	3983	time-intensive
salads	1	2476	4067	neutral
sandwich toppings	66	109	3901	neutral
savory snacks	1800	4166	9700	convenient
soups and broths	936	186	4192	time-intensive
sugar and sugar products	0	0	832	convenient
sugar confectionery	0	4490	11437	convenient
taste enhancers	0	0	1209	convenient
vegetables	1232	5902	19140	time-intensive

Note: The classification of a product in terms of convenience in home production and convenience in purchase are based on text mining the description of products in the SKU directory. The classicification of categories to be convenient or time intensive in home production is based on a consumer survey.

Chapter 4

Correlated Learning and Demand for New Products

4.1 Introduction

More than 3500 new Consumer Packaged Goods (CPG) products hit the shelf every year, of which half are me-too products, i.e., copycats of existing varieties, or brand extensions, i.e., existing brands introducing new varieties. Several questions arise immediately: How do new consumers respond to the two different types of new product launches? Can a consumer learn about his preference toward one product through his consumption experiences with other products? If yes, through which channel do these information spillovers take place? Does this type of information spillover have a larger impact on consumer demand for a certain brand producing different varieties than that for a certain variety produced by different brands, or vice versa? To answer these questions, one must first understand the evolution of consumer preferences and how the information spillovers take place in the consumer learning process.

There is a very large literature on the formation of preferences. For instance, the Bayesian learning model has been widely applied to investigate individual preference evolution. It extends the traditional discrete choice framework by allowing consumers to have incomplete information about the product attributes or about their own preferences (Roberts and Urban, 1988; Erdem and Keane, 1996; Keller and Lehmann, 2006). Over time, a consumer receives information signals from consumption experiences that enable him to learn about his true preferences. Different types of learning models have been applied to various marketing and economics problems (see Ching et al. 2013; Ching and Lim 2016 for literature reviews).

An important part of the learning literature concerns "information spillover across brands, or correlated learning, [which means] learning about a brand in one category by using the same brand in another category and/or learning about one attribute from another" (Ching et al., 2013). In particular, one strand of the correlated learning literature, e.g., Erdem (1998); Marcoul and Weninger (2008); Sridhar et al. (2012); Szymanowski and Gijsbrechts (2012); Che et al. (2015); Ching and Lim (2016), models the correlation in consumer beliefs about related options by specifying a more general prior belief structure. In these models, a consumer's initial beliefs follow a multivariate normal distribution. The intuition behind this type of model setup is that consumers tend to perceive products in the same class to have similarities (e.g., organic products, private labels, different versions of products from a given brand, or products that can heal the same disease) and, therefore, an (inexperienced) consumer's initial beliefs are correlated. As the evolution of consumer beliefs follow Bayes rule, all the subsequent beliefs are also correlated. The correlation in a consumer's beliefs leads to the possibility that a product may benefit from information spillovers from rival products. For example, if a consumer is risk-averse, his purchase experience of any of the related products will have a positive effect on his inclinations to purchase the products in the subsequent periods.

Another strand, such as Coscelli and Shum (2004); Sridhar et al. (2012); Thomas (2017), accommodates information spillover by including correlations across consumption/usage experience signals. A typical model setup in this strand of empirical work assumes that the signals from experiences are correlated instead of the initial beliefs. In other words, the signals an agent receives from experiences with different parties (e.g., patients, retailers, or courses) contain information on the value the agent has to learn about a certain choice alternative. For example, physicians learn about the match value between the drug and a certain disease from the signals experienced by different patients, or students learn about their abilities from exam results of different courses.

Adding to these two strands of literature, this paper specifies a correlated learning model that allows consumers to learn about a given product by consuming competing products that have the same product attribute (e.g, brand or variety). The model specified in this paper puts no restriction on the direction of the effect of information spillover. It further provides behavioral interpretation of the direction of the effect of information spillover, which is informative to the firms in order to make better product introduction decisions.

The model in this paper is intended to be applied to repeat-purchase experience goods markets (most of the products in the CPG market belong to this category, e.g., instant coffee and boxed meals). Such CPG goods form a good application domain because, without prior experience, a consumer's initial information about product attributes is typically incomplete and consumption experience helps resolve the associated uncertainty. The results from future application will be informative to the CPG firms' new product introduction timing decision and the store brands' optimal copycats action decision.

Similar to Chan et al. (2013), my model assumes a consumer has incomplete information about his true preferences of multiple product attributes—brand and variety (or type).[1] After

[1] Chan et al. (2013) assumes patients have incomplete information about the effectiveness and the side effects of a certain drug. Their paper focuses on how detailing and patient feedback help reduce risk-averse physician's uncertainty.

each consumption occasion, a consumer receives signals regarding to his brand preference and variety preference. Information spillovers occur because consumption experience of a given product not only changes the consumer's inclination to buy the product itself, but also changes his inclination to buy other products of the same brand or variety.

After formulating the model, I next discuss the identification of the parameters of interest. The identification discussion in this paper assumes that researchers have access to standard consumer choice panel data (e.g., Nielsen Homescan panel and IRI Consumer Panel), which contains information on each consumer's choice sequence from his first purchase onwards.[2]

I show the feasibility of applying the model to standard purchase panel data with artificial data experiments. The artificial data experiments model a CPG market for a set of low stake experience goods, where the consumers are immediate utility maximizers. The products are differentiated in three dimensions—brand, variety and price. Consumers observe the brand identities, the varieties, and the prices of all the products in each period. However, they have incomplete information on their brand preferences and variety preferences when they are inexperienced.

To begin with, I present evidence that each of the model's parameters can be estimated with simulated maximum likelihood method. I show, for each parameter of interest, that the calculated profile likelihood function value (Venzon and Moolgavkar, 1988; Cox and Snell, 1989) is maximized within a narrow range around the true value. Next, I perform a Monte Carlo study to demonstrate the feasibility of applying the model to standard purchase panel data.

The rest of this paper is organized as follows. Section 2 presents a model of learning

[2]In the CPG market context, scanner data is the most commonly available type of data to the researchers (see Einav et al. (2010) for a validation study of standard scanner data). In the drug market context, researchers can have information on the patient's feedback after each usage, which can enhance identification of learning model (e.g., Chan et al. (2013)). In the CPG market context, however, consumer feedback or rating information is very rare.

with information spillovers. Section 3 summarizes the estimation procedure. Section 4 provides formal discussions on identification and normalization. Section 5 presents the Monte Carlo study results. Having specified the parameter values in the Monte Carlo study, this section also revisits the identification of each parameter of interest using the profile likelihood method. Second 6 concludes with discussions and future research.

4.2 A model of learning with information spillovers

4.2.1 Model setup

Consider a CPG category in which there are J unique brands and each brand can produce up to K different varieties. One can think of the flavored instant-coffee category as an example, where brands such as NESCAFÉ and Maxwell House produce different varieties, combining different roastings and flavors (e.g., vanilla, caramel, etc). Thus, in this category, a consumer can choose from $J \times K$ inside goods and one outside good — "not buying", denoted by good 0. For simplicity, I index products by jk to denote variety k produced by brand j throughout the paper.

A consumer does not have precise information about his brand preferences and variety preferences, and learns from experiences with product jk. He forms expectations about the ex-post consumption utility of each product based on his prior information at the beginning of each period t. The prior information set contains the consumer's beliefs about his preference for brand and variety. In making the product purchase decision, the consumer maximizes his expected utility for this single purchase occasion. A risk-neutral consumer i's consumption utility after consuming product jk (namely the ex-post utility) is specified as

$$u_{ijkt} = Q_{ijkt} + \lambda_{ijk} + \alpha_i p_{ijkt} + \varepsilon_{ijkt}, \tag{4.1}$$

where

- Q_{ijkt} is the consumer's noisy consumption experience (ex post) for product jk at time t, which is unknown to the consumers ex-ante.

- λ_{ijk} is the unobserved constant individual preference of product jk, which is known to the consumers but unknown to the researcher. I further assume $\lambda_{ijk} = \lambda_{ij} + \lambda_{ik}$, where λ_{ij} and λ_{ik} are modeled as normal random coefficients with mean 0 and variance σ_j^2 and σ_k^2, respectively. The specification of λ_{ijk} is to accommodate the possibility that a consumer's unobserved constant individual preference is conditional to the brand identity and the variety identity. [3]

- α_i is the price coefficient and is modeled as a random coefficient with mean μ_α and variance σ_α^2.

- p_{ijkt} is the transaction price of product jk paid by consumer i at time t.

- ε_{ijkt} is an identically and independently distributed error term with a Type I Extreme Value distribution.

The outside good has a utility of $u_{i0t} = 0 + \varepsilon_{i0t}$, which serves as a location normalization for the utilities. The Type I Extreme Value distribution gives the scale normalization for utility values.

4.2.2 Bayesian updating

At period $t = 0$, a number of new consumers make their first purchase in a CPG market. Those consumers, who adopt during the same period, are assumed to have the same initial beliefs about their brand preferences and variety preferences.

[3] For example, if λ_{ijk} represents consumers' preference for package size in the instant coffee market. The distribution of the consumer's preference of package size for the seasonal flavor can be different from that distribution for espresso.

A consumer i's initial belief about his brand preference is given by

$$\mu_{ij0}^{B} \sim \mathbb{N}\left(\bar{\mu}_{j0}^{B},\ \sigma_{Bj0}^{2}\right) \tag{4.2}$$

and his initial belief about his variety preference is given by

$$\mu_{ik0}^{C} \sim \mathbb{N}\left(\bar{\mu}_{k0}^{C},\ \sigma_{Ck0}^{2}\right), \tag{4.3}$$

where $\bar{\mu}_{j0}^{B}$ and $\bar{\mu}_{k0}^{C}$ are the means of prior beliefs about brand preference and variety preference, respectively. σ_{Bj0}^{2} and σ_{Ck0}^{2} are the variances of prior beliefs about brand preference and variety preference, respectively.

Consumer i's complete initial information set I_{i0} at $t=0$ can, therefore, be expressed by the following $(J+K)\times 1$ vector

$$I_{i0} = \begin{bmatrix} \mu_{i10}^{B} \\ \vdots \\ \mu_{iJ0}^{B} \\ \mu_{i10}^{C} \\ \vdots \\ \mu_{iK0}^{C} \end{bmatrix} \sim \mathbb{N}\left(\begin{bmatrix} \bar{\mu}_{10}^{B} \\ \vdots \\ \bar{\mu}_{J0}^{B} \\ \bar{\mu}_{10}^{C} \\ \vdots \\ \bar{\mu}_{K0}^{C} \end{bmatrix}, \begin{bmatrix} \sigma_{B10}^{2} & & & & & \\ & \ddots & & & & \\ & & \sigma_{BJ0}^{2} & & & \\ & & & \sigma_{C10}^{2} & & \\ & & & & \ddots & \\ & & & & & \sigma_{CK0}^{2} \end{bmatrix} \right). \tag{4.4}$$

Assuming independence of prior tastes for brands and varieties, consumer i's initial belief about his preference for product jk is

$$\mu_{ijk0} \sim \mathbb{N}\left(\bar{\mu}_{j0}^{B} + \bar{\mu}_{k0}^{C}, \sigma_{Bj0}^{2} + \sigma_{Ck0}^{2}\right). \tag{4.5}$$

The consumption signal Q_{ijkt} arrives after each consumption occasion. The specification for Q_{ijkt} needs to allow for the substitution between utility from brand and from variety. Therefore, I propose using the specification

$$Q_{ijkt} = q^B_{ijt} + q^C_{ikt}, \tag{4.6}$$

where q^B_{ijt} and q^C_{ikt} are the consumption signals regarding preference for brand j and preference for variety k, respectively. q^B_{ijt} and q^C_{ikt} take the form

$$q^B_{ijt} = \mu^B_j + \nu^B_{ij,t}, \text{ where } \nu^B_{ij,t} \sim \mathbb{N}\left(0, \sigma^2_{\nu B}\right) \tag{4.7}$$

and

$$q^C_{ikt} = \mu^C_k + \nu^C_{ik,t}, \text{ where } \nu^C_{ik,t} \sim \mathbb{N}\left(0, \sigma^2_{\nu C}\right), \tag{4.8}$$

where μ^B_j is the consumer's true preference for brand j, and μ^C_k is the consumer's true preference for variety k. $\nu^B_{ij,t}$ and $\nu^C_{ik,t}$ are normally distributed with zero mean, and are assumed to be i.i.d. over consumers, time, brands, and varieties. It is assumed that consumers know the distributions of $\nu^B_{ij,t}$ and $\nu^C_{ik,t}$. The consumption signal a consumer i receives after consuming product jk in period t can be expressed as the following $(J+K) \times 1$ vector

$$Q_{it} = [0 \cdots q^B_{ijt} \cdots 0 \quad 0 \cdots q^C_{ikt} \cdots 0]'_{(J+K)\times 1}. \tag{4.9}$$

Right after each consumption occasion t, a consumer receives a consumption signal Q_{it}.

Consumer i's choice in t is denoted by a $J \times K$ matrix D_{it}, where

$$D_{it} = \begin{bmatrix} d_{i11t} & \dots & d_{iJit} \\ \vdots & \ddots & \vdots \\ d_{i1Kt} & \dots & d_{iJKt} \end{bmatrix}. \tag{4.10}$$

d_{ijkt} is equal to 1 if product jk is chosen at purchase occasion t and 0 otherwise. The consumer will update his beliefs according to the Bayes rule (DeGroot, 1970) as follows: for $\forall k$,

$$\bar{\mu}^B_{ij,t+1} = \begin{cases} \left(\frac{1}{\sigma^2_{Bij,t}} + \frac{1}{\sigma^2_{B\nu}}\right)^{-1} \left(\frac{1}{\sigma^2_{Bij,t}}\bar{\mu}^B_{ij,t} + \frac{1}{\sigma^2_{B\nu}} q^B_{ij,t}\right) & \text{if } d_{ijkt} = 1 \\ \bar{\mu}^B_{ij,t} & \text{if } d_{ijkt} \neq 0 \end{cases}, \tag{4.11}$$

$$\sigma^2_{Bij,t+1} = \begin{cases} \left(\frac{1}{\sigma^2_{Bij,t}} + \frac{1}{\sigma^2_{B\nu}}\right)^{-1} & \text{if } d_{ijkt} = 1 \\ \sigma^2_{Bij,t} & \text{if } d_{ijkt} \neq 0 \end{cases}, \tag{4.12}$$

and, for $\forall j$

$$\bar{\mu}^C_{ik,t+1} = \begin{cases} \left(\frac{1}{\sigma^2_{Cik,t}} + \frac{1}{\sigma^2_{C\nu}}\right)^{-1} \left(\frac{1}{\sigma^2_{Bik,t}}\bar{\mu}^C_{ik,t} + \frac{1}{\sigma^2_{C\nu}} q^C_{ik,t}\right) & \text{if } d_{ijkt} = 1 \\ \bar{\mu}^C_{ik,t} & \text{if } d_{ijkt} \neq 0 \end{cases}, \tag{4.13}$$

$$\sigma^2_{Cik,t+1} = \begin{cases} \left(\frac{1}{\sigma^2_{Cik,t}} + \frac{1}{\sigma^2_{C\nu}}\right)^{-1} & \text{if } d_{ijkt} = 1 \\ \sigma^2_{Cik,t} & \text{if } d_{ijkt} \neq 0 \end{cases}. \tag{4.14}$$

After consumption occasion in t, consumer i's prior belief I_{it+1} at the beginning of period

t+1 can be defined as

$$I_{it+1}=\begin{bmatrix}\mu^{B}_{i1,t+1}\\ \vdots\\ \mu^{B}_{iJ,t+1}\\ \mu^{C}_{i1,t+1}\\ \vdots\\ \mu^{C}_{iK,t+1}\end{bmatrix}\sim\mathbb{N}\left(\begin{bmatrix}\bar{\mu}^{B}_{i1,t+1}\\ \vdots\\ \bar{\mu}^{B}_{iJ,t+1}\\ \bar{\mu}^{C}_{i1,t+1}\\ \vdots\\ \bar{\mu}^{C}_{iK,t+1}\end{bmatrix},\begin{bmatrix}\sigma^2_{Bi1,t+1} & & & & & \\ & \ddots & & & & \\ & & \sigma^2_{BiJ,t+1} & & & \\ & & & \sigma^2_{Ci1,t+1} & & \\ & & & & \ddots & \\ & & & & & \sigma^2_{CiK,t+1}\end{bmatrix}\right), \tag{4.15}$$

where $\mu^{B}_{ij,t+1}\sim\mathbb{N}\left(\bar{\mu}^{B}_{ij,t+1},\ \sigma^2_{Bj,t+1}\right)$ and $\mu^{C}_{ik,t+1}\sim\mathbb{N}\left(\bar{\mu}^{C}_{ik,t+1},\ \sigma^2_{Ck,t+1}\right)$. Given independence of the subutility for brands and varieties, consumer i's prior belief about product jk can be expressed as

$$\mu_{ijkt}\sim\mathbb{N}\left(\bar{\mu}^{B}_{ij,t}+\bar{\mu}^{C}_{ik,t},\ \sigma^2_{Bj,t}+\sigma^2_{Ck,t}\right). \tag{4.16}$$

Given the initial beliefs in 4.4, these prior beliefs μ_{ijkt}, for $t=0,\ 1,\ 2,\ ...$, can be defined recursively.

Consumers learn about their true preferences through consumption experiences. Let $N^{Bj}_{i,t-1}$ count of the number of purchases of all the products from brand j's by consumer i up to period $t-1$, and $N^{Ck}_{i,t-1}$ count the number of purchases of all the products that belong to variety k by consumer i up to period $t-1$. The transition rules for t $N^{Bj}_{i,t}$ and $N^{Ck}_{i,t}$ are simply:

$$N^{Bj}_{i,t}=\begin{cases}N^{Bj}_{i,t-1}+1 & \text{if } d_{ijkt}=1 \text{ for } \forall k\\ N^{Bj}_{i,t-1} & \text{if } d_{ijkt}\neq 0 \text{ for } \forall k\end{cases} \tag{4.17}$$

$$N^{Ck}_{i,t}=\begin{cases}N^{Ck}_{i,t-1}+1 & \text{if } d_{ijkt}=1 \text{ for } \forall j\\ N^{Ck}_{i,t-1} & \text{if } d_{ijkt}\neq 0 \text{ for } \forall j\end{cases} \tag{4.18}$$

In the long run, when $N_{i,t}^{Bj} \to \infty$ and $N_{i,t}^{Ck} \to \infty$, consumer i's ex-ante expectation about a product equals his true preference for product jk, which is

$$\begin{aligned}
\bar{\mu}_{ij,t}^{B} + \bar{\mu}_{ij,t}^{C} + \lambda_{ijk} =& \frac{\sigma_{\nu B}^{2}\bar{\mu}_{j0}^{B} + \sigma_{Bj0}^{2} N_{i,t-1}^{Bj}\mu_{j}^{B} + \sigma_{Bj0}^{2}\sigma_{\nu B}\sum_{t'=1}^{N_{i,t-1}^{Bj}} \nu_{ij,t'}^{B}}{\sigma_{\nu B}^{2} + \sigma_{Bj0}^{2} N_{i,t-1}^{Bj}} + \lambda_{ijk} \\
&+ \frac{\sigma_{\nu C}^{2}\bar{\mu}_{k0}^{C} + \sigma_{Ck0}^{2} N_{i,t-1}^{Ck}\mu_{k}^{C} + \sigma_{Ck0}^{2}\sigma_{\nu C}\sum_{t'=1}^{N_{i,t-1}^{Ck}} \nu_{ik,t'}^{C}}{\sigma_{\nu C}^{2} + \sigma_{Ck0}^{2} N_{i,t-1}^{Ck}} + \lambda_{ijk} \\
=& \mu_{j}^{B} + \mu_{k}^{C} + \lambda_{ijk}.
\end{aligned} \tag{4.19}$$

4.2.3 Consumer choice problem and the empirical model

In each period t, consumer i makes a choice based on his expected utility given his prior information I_{it}, before observing Q_{it}:

$$\begin{aligned}
E\left(u_{ijkt}|I_{it}\right) &= E\left(Q_{it}|I_{it}\right) + \lambda_{ijk} + \alpha_i p_{ijk} + \varepsilon_{ijkt} \\
&= E\left(q_{ijt}^{B}|I_{it}\right) + E\left(q_{ikt}^{C}|I_{it}\right) + \lambda_{ijk} + \alpha_i p_{ijk} + \varepsilon_{ijkt} \\
&= \bar{\mu}_{ij,t}^{B} + \bar{\mu}_{ik,t}^{C} + \lambda_{ijk} + \alpha_i p_{ijkt} + \varepsilon_{ijkt},
\end{aligned} \tag{4.20}$$

where $q_{ijt}^{B} \sim \mathbb{N}\left(\mu_{j}^{B}, \sigma_{B\nu}^{2}\right)$ and $q_{ikt}^{C} \sim \mathbb{N}\left(\mu_{k}^{C}, \sigma_{C\nu}^{2}\right)$. As $E(\cdot)$ takes expectation of q_{ijt}^{B} and q_{ikt}^{C} conditional on the distribution of I_{it}, the outcome of $E(\cdot)$ is a function of parameters that characterize I_{it}. The second equality follows from the specification of I_{it} and the assumption of risk-neutral consumers.

The utility level of the outside good is normalized to zero; i.e.,

$$u_{i0t} = \varepsilon_{i0t}. \tag{4.21}$$

The standard logit probability of consumer i purchasing product jk at time t is

$$Prob_{ijkt} = \frac{\exp\left(\bar{\mu}^B_{ij,t} + \bar{\mu}^C_{ik,t} + \lambda_{ijk} + \alpha_i p_{ijkt}\right)}{1 + \sum_{j'}\sum_{k'}\exp\left(\bar{\mu}^B_{ij',t} + \bar{\mu}^C_{ik',t} + \lambda_{ij'k'} + \alpha_i p_{ij'k't}\right)}, \tag{4.22}$$

and the probability of a consumer i getting the outside option is

$$Prob_{i0t} = \frac{1}{1 + \sum_{j'}\sum_{k'}\exp\left(\bar{\mu}^B_{ij',t} + \bar{\mu}^C_{ik',t} + \lambda_{ij'k'} + \alpha_i p_{ij'k't}\right)}. \tag{4.23}$$

In the section below, I will illustrate information spillovers with a simple example.

4.2.4 Model properties

4.2.4.1 An example

I now show how information spillovers affect the evolution of a risk-neutral consumer's beliefs and subsequently affect his choice sequences. It is easy to discuss this with a simplified case. Assume there are two brands in the market, brand A and brand B. Each brand produces two difference varieties—variety I and variety II. For simplicity, I label variety I produced by brand A as product AI. Similarly, the other three products are labeled as product AII, product BI, and product BII. To focus on the mechanism of information spillovers, in this example I assume consumers are homogeneous in their initial beliefs, their permanent preferences and their price sensitivities.

To make the example precise, I first consider the very stylized case of differences in consumers' initial beliefs about brands but no other utility differences. In this example, initial beliefs about preferences for brand A ($\bar{\mu}^B_{10}$) and brand B ($\bar{\mu}^B_{20}$) are chosen as 1.1 and 0.5, respectively. The true preferences for both brands are 0.8 ($\mu^B_1 = \mu^B_2$). Further, a consumer's initial beliefs about both varieties are set to be 0.4 ($\bar{\mu}^C_{10} = \bar{\mu}^C_{20} = 0.4$). Finally, his true preferences for both varieties are set to be 0.2 ($\mu^C_1 = \mu^C_2 = 0.2$). The difference between a consumer's initial preference and his true preference represents the consumer's perception bias caused by incomplete information. Note that the model imposes no restriction on the direction of the consumer's initial perception bias. I further set the variances of consumer's initial beliefs for variety I σ^2_{C10} and variety II σ^2_{C20} equal to 0.2^2 and set the variances of consumer's initial beliefs for brand A σ^2_{B10} and brand B σ^2_{B20} equal to 0.5^2. Higher variance of initial beliefs implies faster learning speed, all else being equal. All four products are sold at the same price level, which is set to be 1. The consumer's price sensitivity is set to be -1.

A new consumer enters the market with initial beliefs about his brand and variety preferences, which are governed by normal distributions $\mathbb{N}\left(\bar{\mu}^B_{j0},\ \sigma^2_{Bj0}\right)$ and $\mathbb{N}\left(\bar{\mu}^C_{k0},\ \sigma^2_{Ck0}\right)$, respectively. After each purchase, a consumer receives noisy signals regarding his brand and variety preferences. The noisy signals are drawn from normal distributions with means equal to the true brand and variety preferences and variances equal to 0.5^2.

4.2.4.2 Evolution of beliefs

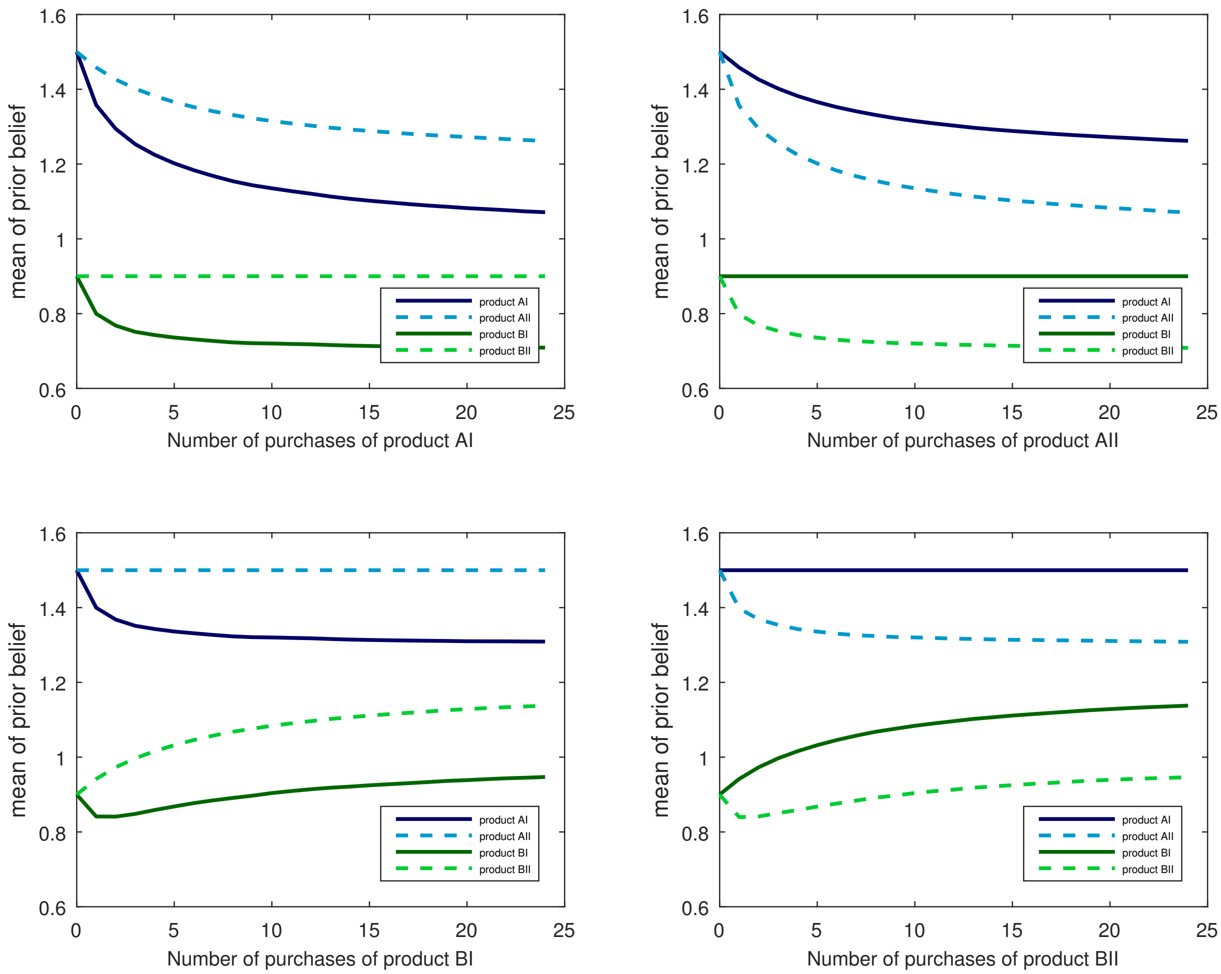
In the model proposed above, information spillovers take place through common product attributes, either brand or variety, among choice options. After consuming a certain product, a consumer uses the signals regarding his brand preference and variety preference to update his beliefs about those products that share the same brand attribute or variety attribute.

To directly present the effect of information spillover on the evolution of consumer beliefs, I first assume that consumers purchase one product exclusively and plot the evolution of his beliefs of all the available products. As Figure 4.1 shows, a consumer uses his purchase experiences of a given product to update the beliefs about other products that contain the same product attributes, while the beliefs of products with different product attributes remain unaffected. In Figure 4.1, the upper-left subplot shows the evolution of a consumer's beliefs when he purchases product AI exclusively. We can see the consumer's beliefs about product AI itself but also about product AII and product BI change with the number of purchases of product AI. The effect on the prior mean of AI is the standard learning effect—a consumer's inclination to purchase a certain product changes with his consumption experiences with that product. The effect on belief about product AII is caused by information spillover through the brand attribute, while the third one is caused by information spillover through the variety attribute. The upper-right subplot, the lower-left subplot, and the lower-right subplot are similar plots with respect to exclusive purchases of product AII, product BI, and product BII, respectively.

Furthermore, the evolution of a consumer's belief about a given product is a combination of two learning processes—learning about brand preference and learning about variety preference. Therefore, we may observe that a consumer's belief about a given product changes non-monotonously with his purchases of that product. Take product BI as an example: the mean of a consumer's initial belief about brand B ($\bar{\mu}_{20}^{B} = 0.5$) is lower than the true preference for brand B ($\mu_2^B = 0.8$), while the mean of consumer's initial belief about variety I ($\bar{\mu}_{10}^{C} = 0.4$) is higher than his true preference for variety I ($\mu_2^C = 0.2$). Then, what we observe is that this consumer's belief about product BI first decreases and then increases with his purchases of product BI (see lower-left subplot in Figure 4.1).

In sum, using this illustrative example, we see that information spillovers take place across products through common product attributes. In the absence of information spillovers, one can only observe how a consumer's belief about the chosen product changes over time. A consumer learns about both his brand preference and his variety preference. Therefore, the evolution of a consumer's belief about a given product may not change monotonously with his purchases of that product.

Figure 4.1: Prior mean evolution against purchases of one product



Notes: In each subplot, the solid lines denote type I variety and the dashed lines denote type II variety. Blue denotes the products produced by brand A and green denotes the products produced by brand B. The horizontal line is the number of purchases of a certain product. The vertical line is the mean of consumer's prior belief at the beginning of each purchase occasion.

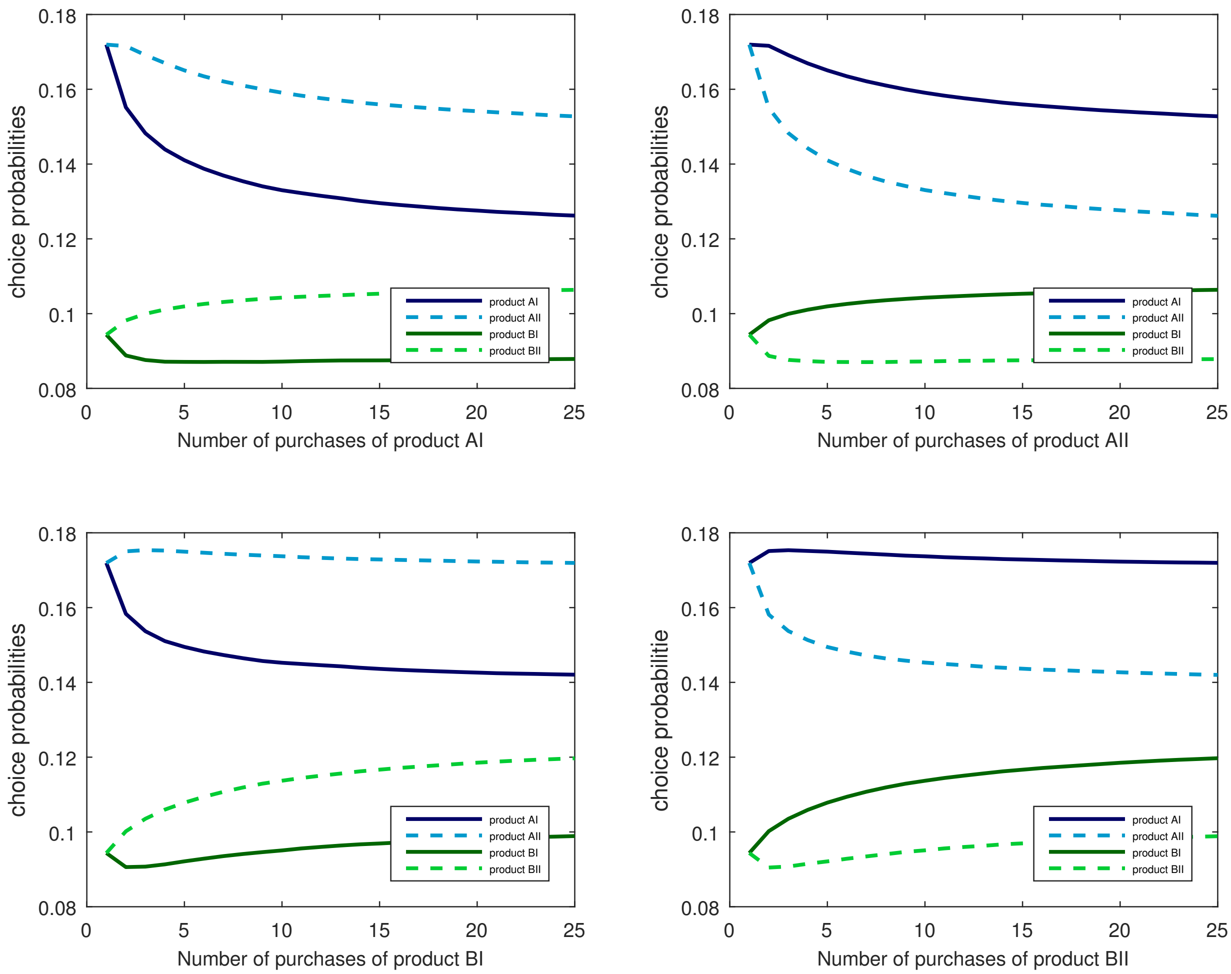
4.2.4.3 Evolution of choice probabilities

Information spillovers eventually affect the evolution of a consumer's purchase decisions. Figure 4.2 shows the evolution of a consumer's choice probability when he purchases one product exclusively.

From Figure 4.2, one can see two features. First, changes in this consumer's belief about any product will affect his inclination to purchase each product in the market. This is because the choice probabilities sum up to one in each period for a given consumer. Second, a non-chosen product may benefit more from consumer learning than the chosen product. For instance the window that belongs to product BI suggests that BII benefits more from purchases of BI than BI itself does. Recall that a consumer's initial belief about brand B is downward biased, while his initial belief about variety I is upward biased (implied by the parameter values). The evolution of a consumer's choice probability of product BI is a combination of learning about brand preference and learning about variety preference, which are in opposite directions. However, the evolution of consumer's choice probability about product BII is only affected by the upward adjustment in the consumer's belief about the brand preference.

We see that the evolution of the consumer's choice probabilities (the market share) largely depends on how the consumers perceive the product attributes initially and how fast learning takes place. If the consumers over-perceive the value of a new variety, the copycat strategy may be more effective when it is implemented at the early stage. On the contrary, if the consumers under-perceive the value of a new variety, the free-rider brands may benefit more to launch the copycat product when the consumers' uncertainty has been resolved.

Figure 4.2: Choice probability evolution against purchases of one product



Notes: In each subplot, the solid lines denote type I variety and the dashed lines denote type II variety. Blue denotes the products produced by brand A and green denotes the products produced by brand B. The horizontal line is the number of purchases of a certain product. The vertical line is the consumer's expected utility levels prior to each purchase occasion.

4.3 Estimation

Consider a consumer choice panel data set that consists of observations of consumers' complete choice sequences in a given new experience goods category, $D_{it} = (D_{i1},\ D_{i2},\ ...,\ D_{it})$. The data set also contains information on observable product characteristics, e.g., price. Let $\theta \in \Theta$ denote the parameters of interest. From the previous discussion, the probability of a consumer purchasing product jk depends on his prior belief at the beginning of t and the observable product characteristics, i.e.,

$$\text{Prob}\left(d_{ijkt} = 1 |\ p_{it},\ I_{it},\ \kappa_i;\ \theta\right) = \text{Prob}\left(d_{ijkt} = 1 |\ D_{it-1},\ p_{it},\ \nu^B_{ik,t-1},\ \nu^C_{ij,t-1},\ \kappa_i;\ \theta\right), \tag{4.24}$$

$$\text{Prob}\left(d_{it} = j | z_{it},\ I_{it};\ \theta\right) = \text{Prob}\left(d_{it} = j | d_{it-1},\ z_{it},\ \nu_{i,t-1},\ \kappa_i;\ \theta\right),$$

where d_{ijkt} is the indicator that consumer i's chooses product jk in period t; $d_{it-1} = (d_{i1},\ d_{i2},\ ...,\ d_{it-1})$ is the observed choice sequence; $p_{it} = (p_{i1},\ ...,\ p_{it})$, where p_{it} is the observed prices of all the in period t; I_{it} is consumer i prior belief at time t; $\theta \in \Theta$ denotes the parameters we want to estimate. The observed choice probability is equal to the model prediction given the set of parameters θ. Further, we define $\nu^B_{i,t-1} = \left(\nu^B_{i,1},\ \nu^B_{i,2},\ ...,\ \nu^B_{i,t-1}\right)$, where $\nu^B_{i,t-1} = \left(\nu^B_{i1,t-1},\ ...,\ \nu^B_{iJ,t-1}\right)$, and $\nu^C_{i,t-1} = \left(\nu^C_{i,1},\ \nu^C_{i,2},\ ...,\ \nu^C_{i,t-1}\right)$, where $\nu^C_{i,t-1} = \left(\nu^C_{i1,t-1},\ ...,\ \nu^C_{iK,t-1}\right)$. The vector κ_i denotes unobserved heterogeneities. The elements of vector κ_i are drawn (identically across all consumers i) from the normal distribution with mean zero and variance σ_j and σ_k, respectively.

The researchers cannot observe the unobserved heterogeneities κ_i and the realizations of the non-deterministic part ν^B_{it} and ν^C_{it} in all the consumption experience signals consumer i has received. This implies that the likelihood function for a given sequence of consumption frequencies for a given consumer involves a multivariate integral over the distribution of the unobserved signals and the random coefficients. Hence, the likelihood contribution of consumer i is

$$\begin{aligned}&\mathscr{L}_i\left(\theta \mid d_{iT_i},\ p_i\right)\\&=\int\left(\Pi_{t=1}^{T_i}\mathrm{Prob}_i\left(d_{it}\mid d_{it-1},\ p_{it},\ \nu_{i,t-1}^{B},\ \nu_{i,t-1}^{C},\ \kappa_i;\ \theta\right)\right)dF\left(\nu_{i,t-1}^{B},\ \nu_{i,t-1}^{C},\ \kappa_i\right),\end{aligned}\tag{4.25}$$

where $p_i=\left(p_{i1},\ ...,\ p_{iT_i}\right)$. Because of the high dimensional integral, I use simulated maximum likelihood estimation to estimate θ. In the Monte Carlo study reported in this paper, the number of draws S is set to be 100. The likelihood contribution of consumer i is

$$\begin{aligned}&\mathscr{L}_i\left(\theta|d_{iT_i},,p_{iT_i}\right)\\&=\frac{1}{S}\sum_{s=1}^{S}\left[\left(\Pi_{t=0}^{T_i}\mathrm{Prob}_i^s\left\{d_{it}|,d_{it-1},p_{it}\right\}\right)|\left(\nu_{i,t-1}^{Bs},\ \nu_{i,t-1}^{Cs},\ \kappa_i^s;\ \theta\right)\right].\end{aligned}\tag{4.26}$$

where Prob_i^s is the choice probability of consumer i making the observed purchase decision during period t given the sth set of draws. The log-likelihood for all the observations of N consumers is

$$\log\mathscr{L}=\sum_{i=1}^{N}\log\mathscr{L}_i.\tag{4.27}$$

4.4 Identification and normalizations

4.4.1 General discussion

It is useful to first recall the definitions of identification and normalization. Koopmans (1949) describes the idea of identification as "inference from that distribution to the parameters of the structural equations describing economic behavior". In other words, the assumptions made about the data generating process that allows one to draw causal inference is identification assumption. The issue of normalization arises when two or more parameter values are observationally equivalent. Before discussing normalization assumptions used in this paper, we first recall the key issues associated with normalization regarding to the intercept of the utility function in the discrete choice model through the following example.

Suppose consumer i faces four products and one outside option, the utilities he can derive from choosing each product are:

$$\begin{aligned}
u_{i1} &= a_1 + b_1 + \varepsilon_{i1} &&= \delta_1 + \varepsilon_{i1} \\
u_{i2} &= a_2 + b_1 + \varepsilon_{i2} &&= \delta_2 + \varepsilon_{i2} \\
u_{i3} &= a_1 + b_2 + \varepsilon_{i3} &&= \delta_3 + \varepsilon_{i3} \\
u_{i4} &= a_2 + b_2 + \varepsilon_{i4} &&= \delta_4 + \varepsilon_{i4} \\
u_{i0} &= 0 + \varepsilon_{i0} &&= \delta_0 + \varepsilon_{i0}
\end{aligned}$$

where ε_j, for $j = 0,\ 1,\ 2$, are i.i.d. errors following Type I extreme distribution. The choice probabilities are

$$\mathrm{prob}\,(d_{ij} = j) = \frac{\exp(\delta_j)}{\sum_{j'} \exp\left(\delta_{j'}\right)}.$$

The observed choice of individual i is d_{ij}, where $d_{ij} = 1$ if j is selected and $d_{ij} = 0$ if j is not selected. The observed explanatory variable x_i is 1 for all the individuals. Let $\theta = (a_1, a_2, b_1, b_2)'$, the associated likelihood is

$$L(\theta) = \Pi_{i=1}^{N} \Pi_j \mathrm{prob}\,(y_i = j | x_i)^{d_{ij}},$$

and it henceforth gives the log-likelihood

$$\log L(\theta) = \sum_i \sum_j d_{ij} \log \Pi_{i=1}^N \Pi_j \mathrm{prob}\left(y_i = j | x_i\right) = \sum_i \left(\sum_j d_{ij} \exp\left(\delta_j\right) - \log \left(\sum_{j'=1}^J \exp\left(\delta_{j'}\right) \right) \right).$$

Note that any locus of pairs $(a_1 - \Delta, a_2 - \Delta, b_1 + \Delta, b_2 + \Delta)$ yields the same likelihood value contribution for individual i. Therefore, we need normalizations to restrict the parameter space. If we normalize one of the four parameters, the other three can be identified.

4.4.2 Arguments for the model in this paper

To start with, I make the identification and normalization arguments in the case of homogeneous consumers, i.e., consumers have the same true preferences and the same price sensitivity level.[4] Recall that a consumer maximizes his expected utilities at the beginning of each period (see (4.20)). When a consumer has had enough experiences in the long run, the uncertainty has been resolved and his expectations about his true preferences for brand and variety, respectively, equal to his true preferences (in other words, the return of an additional consumption experience is 0 in the long run). Therefore, the consumer's expected utility in the long run can be expressed as $\mu_j^B + \mu_k^C + \alpha p_{ijkt} + \varepsilon_{ijkt}$. We see that in the long run the consumer's choice problem has converged to a static discrete choice problem. Thus, the identification and normalization assumptions are also similar to a static consumer decision problem.

First, variation in observed transaction price p_{ijkt} identifies the price coefficient α. The rest of the parameters, μ_1^B, ..., μ_J^B, μ_1^C, ..., μ_K^C, cannot be identified all together. The intuition is similar to the illustrative example in Section 4.4.1: Any locus of pairs $(\mu_1^B - \Delta,\ ...,\ \mu_J^B - \Delta,\ \mu_1^C + \Delta,\ ...,\ \mu_K^C + \Delta)$ yields the same likelihood value. Therefore, one has to impose a

[4] Note that these consumers will still become heterogeneous during the learning process and appear to have different choice sequences, as they may receive different consumption experience signals over time.

normalization assumption — fix one of the parameters, say μ_1^C, to be a known constant. Then the rest of the parameters can be identified with long run choice data. From now on, we treat μ_1^B, ..., μ_J^B, μ_1^C, ..., μ_K^C, and α as known parameters and discuss the identification of the mean and variance of consumers' initial beliefs and the variances of signals.

The identification primarily comes from how consumer's purchase behavior changes over time before reaching the long run state. Recall that, in the short run, consumer i's brand preference and variety preference can be written as

$$\begin{aligned}\bar{\mu}_{ij,t}^B + \bar{\mu}_{ij,t}^C =& \frac{\sigma_{\nu B}^2 \bar{\mu}_{j0}^B + \sigma_{Bj0}^2 N_{i,t-1}^{Bj} \mu_j^B + \sigma_{Bj0}^2 \sigma_{\nu B} \sum_{t'=1}^{N_{i,t-1}^{Bj}} \nu_{ij,t'}^B}{\sigma_{\nu B}^2 + \sigma_{Bj0}^2 N_{i,t-1}^{Bj}} \\ &+ \frac{\sigma_{\nu C}^2 \bar{\mu}_{k0}^C + \sigma_{Ck0}^2 N_{i,t-1}^{Ck} \mu_k^C + \sigma_{Ck0}^2 \sigma_{\nu C} \sum_{t'=1}^{N_{i,t-1}^{Ck}} \nu_{ik,t'}^C}{\sigma_{\nu C}^2 + \sigma_{Ck0}^2 N_{i,t-1}^{Ck}}, \end{aligned} \tag{4.28}$$

where $N_{i,t-1}^{Bj}$ counts of the number of purchases of all the products from brand j's by consumer i up to period $t-1$, and $N_{i,t-1}^{Ck}$ counts the number of purchases of all the products belong to variety k by consumer i up to period $t-1$.

Variations in $N_{i,t-1}^{Bj}$ and $N_{i,t-1}^{Ck}$ across time and across individuals can identify ($\bar{\mu}_1^B$, ..., $\bar{\mu}_J^B$, $\bar{\mu}_1^C$, ... , $\bar{\mu}_K^C$) and $\left(\frac{\sigma_{B10}^2}{\sigma_{\nu B}^2}, \ldots, \frac{\sigma_{BJ0}^2}{\sigma_{\nu B}^2}, \frac{\sigma_{C10}^2}{\sigma_{\nu C}^2}, \ldots, \frac{\sigma_{CK0}^2}{\sigma_{\nu C}^2}\right)$ under a normalization assumption — fix one of the parameters, say $\bar{\mu}_{10}^C$, to be a known constant. The intuition is the same as for the previous discussion on the identification of $\left(\mu_1^B, \mu_2^B, \mu_1^C, \mu_2^C, \mu_3^C\right)$. Lastly, the identifications of two signal variances, $\sigma_{C\nu}$ and $\sigma_{B\nu}$, rely on the unexpected switches across brands and across variety, receptively. It is challenging to estimate two signal variances simultaneously when a consumer receives two signals in each period. In the Monte Carlo study, I normalize both signal variances to 0.5^2.[5] Normalizing both signal variances to a certain value is mainly to reduce computation time. The model itself does not require this

[5] In the existing empirical work, the range of the standard deviation of the signal is approximately from 0.1 to 1.4 (Szymanowski and Gijsbrechts (2012); Crawford and Shum (2005)).

assumption. By assuming both signal variances equal to 0.5^2 in the Monte Carlo study, I implicitly add one additional assumption—the signal regarding to brand preference and the signal regarding to variety preference have a similar precision level. In Section 4.5.2, I provide evidence on the feasibility of identifying signal variances. The variances can be identified together with other model parameters if desired.

4.5 Monte Carlo Study

4.5.1 Set-up

To demonstrate that the model can perform well with standard choice data, I construct a hypothetical Consumer Packaged Goods (CPG) market with $J = 2$ brands and $K = 3$ varieties. In total, there are $J \times K = 6$ products available to the consumers. In this market, there exists information on $N = 300$ new consumers' repeat purchase decisions in this market for $T = 100$ periods. Consumers make their initial purchases at the first period $t = 1$ and choose among the $J \times K$ products and a "not buy" outside option at each of the subsequent periods, $t = 2,\ 3,\ ...,\ T$.

The vector of observed product characteristics is $\left(dummy_j^B,\ dummy_k^C,\ p_{ijkt}\right)$, where $dummy_j^B$ and $dummy_k^C$ are dummy variables for brand and variety, respectively.[6] The transaction price p_{ijkt} of product jk in period t is known to consumer i before purchase and is drawn from a normal distribution, $\mathbb{N}(2,\ 0.3)$.

A consumer's beliefs about his brand preferences are modeled as normal distributions with means $\left(\bar{\mu}_{10}^B,\ \bar{\mu}_{20}^B\right)$ equal to $(3.3,\ 2.7)$ and standard deviations $(\sigma_{B10},\ \sigma_{B20})$ to $(0.3,\ 0.3)$.

[6]This paper is not aimed at separating learning about brand/variety as a vertical attribute from learning about consumer's brand/variety preference. I assume the brand identities and the variety identities are observed by consumers. But consumers do not have perfect information about their preferences, and have to learn through consumption experiences.

Similarly, a consumer beliefs about his variety preferences are modeled as normal distributions with means $\left(\bar{\mu}_{10}^{C},\, \bar{\mu}_{20}^{C},\, \bar{\mu}_{30}^{C}\right)$=(3.8, 4.3, 4.1) and standard deviations (σ_{C10}, σ_{C20}, σ_{C30} = (0.36, 0.36, 0.36). A consumer's true preferences for brand A, brand B, variety I, variety II, and variety III are given by $\left(\mu_1^B,\, \mu_2^B,\, \mu_1^C,\, \mu_2^C,\, \mu_3^C\right) = (3,\, 3,\, 3.8,\, 4,\, 3.8)$. The consumers' private information about their idiosyncratic preferences for brand and variety are modeled as normally distributed random coefficients with mean zero and standard deviations $\left(\sigma_{\lambda 1}^B,\, \sigma_{\lambda 2}^B\right) = (0.2,\, 0.2)$ and $\left(\sigma_{\lambda 1}^C,\, \sigma_{\lambda 2}^C,\, \sigma_{\lambda 3}^C\right)$=(0.15, 0.15, 0.15), respectively. The consumer's price sensitivity α_i is modeled as a normal random coefficient with mean -1 and standard deviation $e^{-1.5}$. The standard deviations of signals regarding to the consumer's brand and variety preference $(\sigma_{vB},\, \sigma_{vC})$ are set to be $(0.5,\, 0.5)$.

The parameters are set to mimic some realistic aspects of the CPG market. First, based on the chosen parameter values, the average market shares for brand A and brand B are about 0.3 and 0.6, respectively. In the U.S. CPG market, the market share of the leading brand is about 0.26 on average (Bronnenberg et al., 2007). Here, brand A represents the leading brand in a specific CPG market, while brand B clusters all the other brands. Second, it is realistic to assume the consumer's price sensitivity of CPG products to be around -1 (Heerde et al. (2013)). Lastly, on average, it takes about five purchases for a consumer to reduce his initial uncertainty (measured by the variance of initial beliefs) by half, which is a realistic amount for an average consumer in the CPG market.

4.5.2 Back to identification

In this section, I return to identification and discuss simulation results pertaining to the identification of the parameters of the learning model specified in this paper. Specifically, the discussion in this section is based on the same data set used in the Monte Carlo study.[7]

[7] To make the likelihood values comparable, the number of simulation draws is set to be 100 in both the simulations here and the Monte Carlo study.

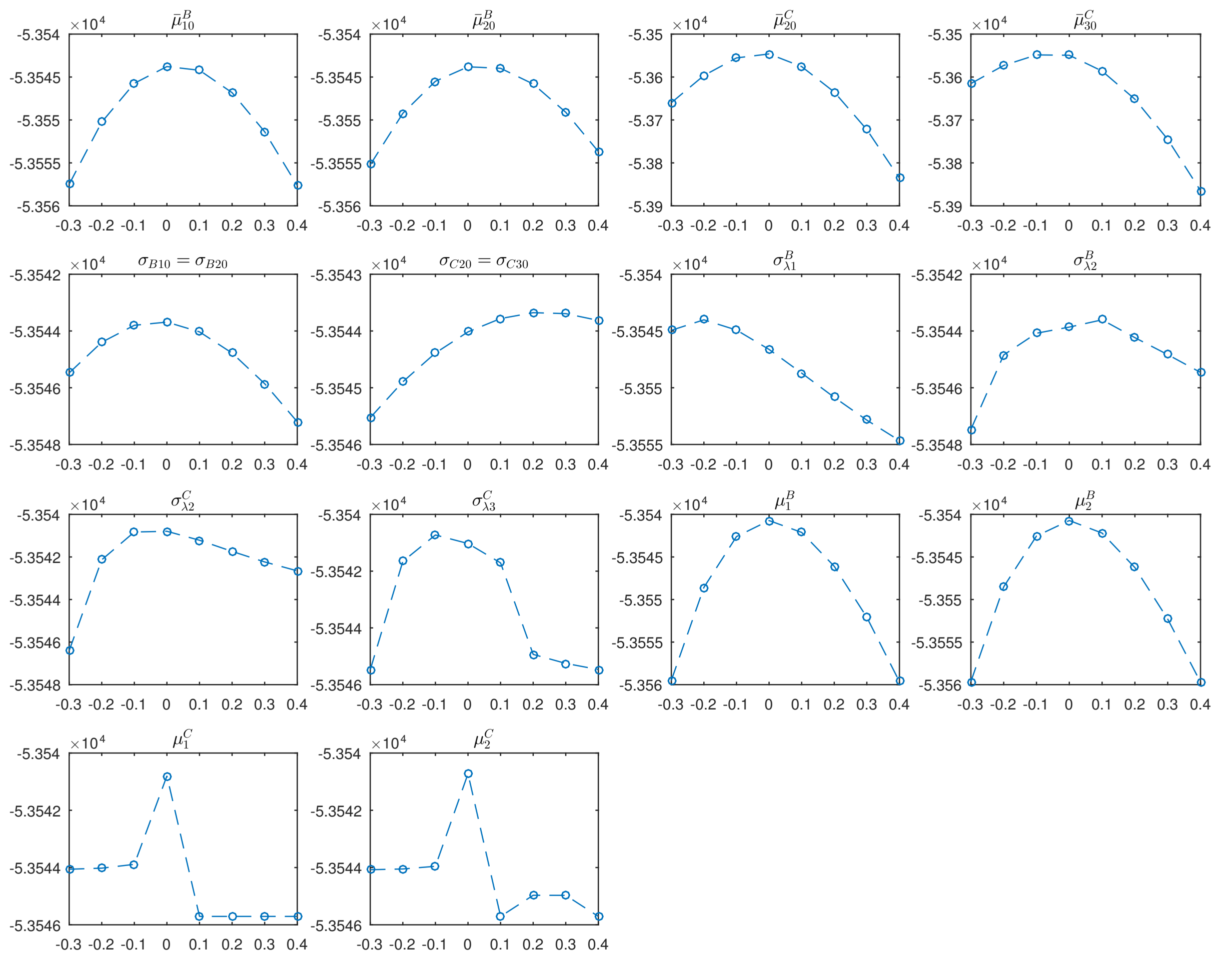
To this end, I plot the profile likelihood function (Venzon and Moolgavkar, 1988; Cox and Snell, 1989) for each parameter of interest in Figure 4.3. The profile likelihood method reduces the likelihood function to a function of a single parameter by treating the others as nuisance parameters and maximizing over them. Suppose that we partition the parameters of interest θ into $\theta_1 = (\theta_1,\ \theta_2)$, where θ_1 is a single parameter and θ_2 denotes all the other parameters. A profile likelihood function is $L(\theta) = L\left(\theta_1,\ \hat{\theta}_2\right)$, in which $\hat{\theta}_2$ is the value of θ_2 maximizing L for the assumed θ_1. I generate a sequence of values for θ_1 around the true value and maximize over θ_2 for each generated value. If a parameter can be well estimated, the profiled likelihood should be maximized within a narrow range around the true value.

In Figure 4.3, each subfigure plots the profile likelihood function for a certain parameter. The true data-generating value of this parameter is given by point 0 on the horizontal axis. I vary the value of θ_1 around its true data-generating value from $0.7\times$true value to $1.4\times$true value. For each assigned value of this parameter, I calculate the likelihood function by maximizing over the rest of the parameters, i.e., θ_2. The title of each subfigure in Figure 4.3 denotes which parameter is treated as θ_1 in that subplot.

Next, I describe the figure in more detail. In Figure 4.3, a parameter is not well identified if its associated profile likelihood function is flat (the maximum of the profile likelihood function cannot be found). A flat profile likelihood function indicates that a wide range of values of a given parameter can generate the same likelihood value. We see that the means and variances of consumer's initial beliefs $\left(\bar{\mu}_{10}^{B}, \bar{\mu}_{20}^{B}, \bar{\mu}_{20}^{C}, \bar{\mu}_{30}^{C}, \sigma_{B10}, \sigma_{B20}\right)$, the consumer's true preferences for brand and variety $\left(\mu_1^B, \mu_2^B, \mu_2^C, \mu_3^C\right)$, the variances of unobserved individual preferences for brand and variety $\left(\sigma_{\lambda 1}^{B}, \sigma_{\lambda 2}^{B}, \sigma_{\lambda 1}^{C}, \sigma_{\lambda 2}^{C}, \sigma_{\lambda 3}^{C}\right)$, and the mean and variance of the price coefficient $(\mu_\alpha, \sigma_\alpha)$ can be identified with standard purchase panel data. Among these parameters, the variances of consumer's initial beliefs $(\sigma_{C10}, \sigma_{C20})$ and the variances of unobserved individual preferences for brand and variety

$\left(\sigma^B_{\lambda 1}, \sigma^B_{\lambda 2}, \sigma^C_{\lambda 1}, \sigma^C_{\lambda 2}, \sigma^C_{\lambda 3}\right)$ are less well-identified.

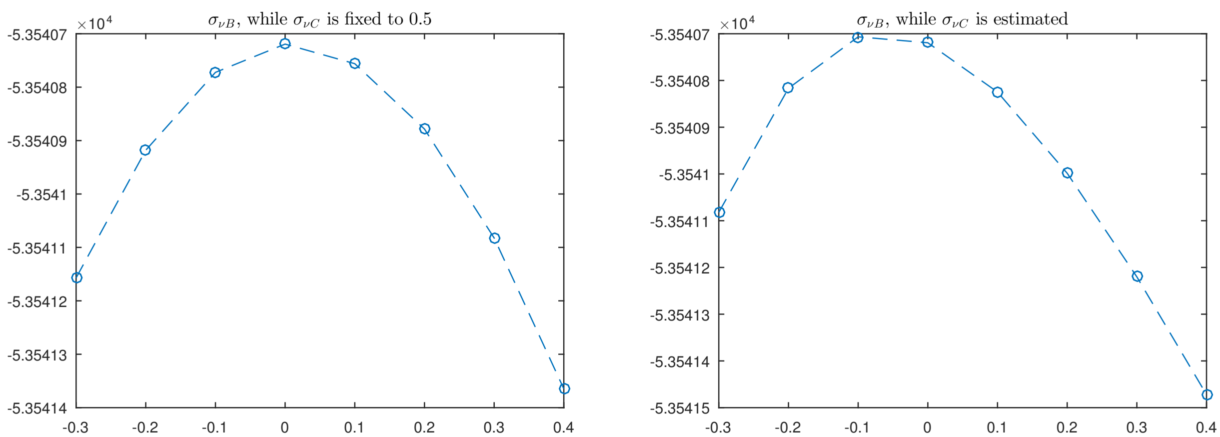
Figure 4.3: Profile likelihood function



Notes: On the x-axis, we consider proportional deviations of a given parameter away from the true value. Take the first subplot — profile likelihood function for $\bar{\mu}^B_{10}$—as an example. The x-axis contains a sequence proportional deviations of $\bar{\mu}^B_{10}$, i.e., $\bar{\mu}^B_{10} \times (1+\lambda)$, $\lambda = -0.3,\ -0.2,\ -0.1,\ 0,\ 0.1,\ 0.2,\ 0.3,\ 0.4$. $\lambda = 0$ denotes the true data-generating value for $\bar{\mu}^B_{10}$.

Lastly, I discuss the identification of two signal variances—σ^2_{Bv} and σ^2_{Cv}. First, I treat the standard deviation of signal regarding brand preference σ_{Bv} as a parameter to be estimated while normalizing σ_{Cv} to 0.5. The corresponding profile likelihood function is presented in the first subfigure in Figure 4.4. We see that σ_{Bv} can be identified under the normalization assumption imposed on $\sigma_{C\nu}$. Next, I treat $\sigma_{B\nu}$ and $\sigma_{C\nu}$ as parameters to be estimated and present the corresponding profile likelihood function for $\sigma_{B\nu}$ in the second subfigure. We see that it is feasible to estimate two signal variances together. However, the y-axis scale in Figure 4.4 is relatively narrow. It implies that when one reduces the number of simulation draws or include more brands/varieties (the parameter space increases), it may not be feasible to separately estimate two signal variances.[8]

Figure 4.4: Profile likelihood function of signal variances



Notes: On the x-axis, we consider proportional deviations of a given parameter away from the true value of $\sigma_{\nu B}$. The first subplot presents the profile likelihood function for the standard deviation (SD) of the brand signal, while the SD of the variety signal is normalized to 0.5. The second subplot presents the profile likelihood function for the SD of the brand signal, while the SD of the variety signal is estimated together with the other parameters. In each plot, the x-axis contains a sequence proportional deviations of $\sigma_{\nu B}$, i.e., $\sigma_{\nu B} \times (1+\lambda)$, $\lambda = -0.3,\ -0.2,\ -0.1,\ 0,\ 0.1,\ 0.2,\ 0.3,\ 0.4$. $\lambda = 0$ denotes the true data-generating value for $\sigma_{\nu B}$.

4.5.3 Monte Carlo results

I estimate the model with Maximum Simulated Likelihood method. Table 4.1 reports the Monte Carlo results with 124 repetitions.[9] The true values of parameters of interest are

[8] The findings presented here are based on a sample of 300 households who make repeat purchase decisions for 100 periods. The number of simulation draws is 100.

[9] I discarded results where the solver did not converge: 9 out of 133 repetitions were excluded.

denoted by θ_0. The starting value equals to $\theta_0\,(1+\omega)$, where ω is a sequence of normally distributed pseudo random numbers drawn from $\mathbb{N}\,(0,\ 0.5)$.

The first column presents the mean of estimates from the Monte Carlo study. The second column presents the true parameter values. The third column and fourth column report the standard deviation of Monte Carlo results and the mean of standard errors of the estimates from the Monte Carlo study, respectively. Comparing the first two columns, we find that most of the parameters have moderate small sample bias. Note that the validity of statistical inferences based on the reported standard errors is built on a relatively small sample size—300 households and each household is allowed to choose among six products in each purchase occasion. The sample size may not be large given the number of products and the number of random coefficients included in the current model. One may achieve better results with a larger sample size, but the computational burden of estimation also increases with the sample size.

Table 4.1: Results for the Monte Carlo Experiment

	par. est.	true values	std. err. MC	std. err. est
true brand match value:				
μ_{10}^{B}	3.005	3.000	0.004	0.167
μ_{20}^{B}	3.026	3.000	0.009	0.167
true variety match value:				
μ_{20}^{C}	3.999	4.000	0.042	0.027
μ_{30}^{C}	3.814	3.800	0.099	0.027
heterogeneity in permenant taste of brand:				
$\sigma_{\lambda_{1B}}$	0.216	0.200	0.022	0.029
$\sigma_{\lambda_{2B}}$	0.190	0.200	0.050	0.027
heterogeneity in permenant taste of variety:				
$\sigma_{\lambda_{2C}}$	0.102	0.150	0.022	0.033
$\sigma_{\lambda_{3C}}$	0.135	0.150	0.019	0.029
initial brand belief mean:				
μ_{10}^{B}	3.397	3.300	0.009	0.217
μ_{20}^{B}	2.800	2.700	0.016	0.203
initial brand belief variance:				
σ_{Bj0}	0.311	0.300	0.005	0.054
initial variety belief mean:				
μ_{20}^{C}	4.129	4.300	0.045	0.059
μ_{30}^{C}	3.884	4.100	0.023	0.059
initial variety belief variance:				
σ_{Ck0}	0.222	0.360	0.116	0.085
price coefficient mean:				
μ_{α}	-0.985	-1.000	0.002	0.024
price coefficient std:				
σ_{α}	0.196	0.223	0.010	0.033
negLogLikelihood	53545.156	53554.510		

4.6 Conclusion

In this paper, I specify a structural learning model that allows information spillovers across two product attributes—brand and variety. The model concerns typical repeat-purchase experience goods in the CPG market. A significant feature of the CPG market is that it has a large number of existing and entering products. Understanding how the evolution of consumer preferences takes place, and whether there is information spillover within a brand's product line or across brands, has important implications for the firms. The model in this paper is based on a Bayesian learning process, specified at the individual level. In this model, a consumer observes a product's brand identity and variety identity, but has to learn about his true preferences for brand and variety through consumption experiences. It explicitly models how the information spillover take place across products and to what extent the information spillovers affect a consumers choice sequences. The hypothesis behind the model is that the information spillovers across products through the common product attributes. To show the feasibility of the model, I formally discuss the identification issue and characterize the normalization assumptions that are needed to apply the model specified in this paper to standard consumer purchase panel data. I find that, under certain regular conditions, the parameters in the correlated learning model proposed by this paper can be identified. I further show with a Monte Carlo study that the model can be directly applied to standard consumer choice panel data.

In future research, I will apply the model to the CPG market and perform counterfactual simulations on the brand extensions (i.e., existing brands introducing new varieties) and me-too product launches (i.e., copycats of existing varieties). The results will be informative to the CPG firms' new product introduction timing decisions and the store brands' optimal copycat action decisions. For example, producing a me-too product may not be profitable if consumers adjust their perceptions about their preference of a new variety downwards, while copying existing products may be more beneficial if the brand is "under-estimated" by the consumers.

Lastly, in the current model, I assume all the consumers, who adopt in the same period, have the same initial beliefs and the same true preferences. Future extensions can relax this restriction and investigate the magnitude of information spillovers when the consumers appear to be very diverse.

Bibliography

Ackerberg, D. A. (2003). Advertising, learning, and consumer choice in experience good markets: an empirical examination. *International Economic Review 44*(3), 1007–1040.

Aguiar, M. and E. Hurst (2005). Consumption versus expenditure. *Journal of Political Economy 113*(5), 919–948.

Aguiar, M. and E. Hurst (2007a). Life-cycle prices and production. *The American Economic Review 97*(5), 1533–1559.

Aguiar, M. and E. Hurst (2007b). Measuring trends in leisure: The allocation of time over five decades. *The Quarterly Journal of Economics 122*(3), 969–1006.

Aguiar, M., E. Hurst, and L. Karabarbounis (2012). Recent developments in the economics of time use. *Annual Review of Economics 4*(1), 373–397.

Aguiar, M., E. Hurst, and L. Karabarbounis (2013, Aug). Time use during the great recession. *American Economic Review 103*(5), 1664–1696.

Alpert, F. H. and M. A. Kamins (1995, October). An empirical investigation of consumer memory, attitude, and perceptions toward pioneer and follower brands. *Journal of Marketing 59*(4), 34–45.

Arrow, K. J. (1971). The theory of risk aversion. *Essays in the theory of risk-bearing*, 90–120.

Bain, J. S. (1956). Barriers to new competition: Their character and consequences in manufacturing industries. *Cambridge (Mass.)*.

Baltas, G., P. C. Argouslidis, and D. Skarmeas (2010). The role of customer factors in multiple store patronage: A cost–benefit approach. *Journal of Retailing 86*(1), 37–50.

Bass, F. (1969). A new product growth model for consumer durables. *Management Science 15*(5), 215–227.

Baumgartner, H. and J.-B. E. Steenkamp (1996). Exploratory consumer buying behavior: Conceptualization and measurement. *International Journal of Research in Marketing 13*(2), 121–137.

Becker, G. S. (1965). A theory of the allocation of time. *The Economic Journal 75*(299), 493–517.

Benhabib, J., R. Rogerson, and R. Wright (1991). Homework in macroeconomics: Household production and aggregate fluctuations. *Journal of Political Economy 99*(6), 1166–1187.

Bertrand, M. and D. W. Schanzenbach (2009). Time use and food consumption. *The American Economic Review 99*(2), 170–176.

Biddle, J. E. and D. S. Hamermesh (1990). Sleep and the allocation of time. *Journal of Political Economy 98*(5), 922–943.

Bombardini, M. and F. Trebbi (2012). Risk aversion and expected utility theory: an experiment with large and small stakes. *Journal of the European Economic Association 10*(6), 1348–1399.

Bronnenberg, B. J. (2015, 46:3 (Fall), 480-498). The provision of convenience and variety by the market. *RAND Journal of Economics 46*(3), 480–498.

Bronnenberg, B. J. (2017). Innovation and distribution: A general equilibrium model of manufacturing and retailing. Working Paper, Tilburg University, Tilburg, NL.

Bronnenberg, B. J., S. K. Dhar, and J.-P. Dubé (2007). Consumer packaged goods in the united states: National brands, local branding. *Journal of Marketing Research 44*(1), 4–13.

Bronnenberg, B. J., J.-P. Dubé, M. Gentzkow, and J. M. Shapiro (2015). Do pharmacists buy bayer? informed shoppers and the brand premium. *Quarterly Journal of Economics 130*(4), 1669–1726.

Chan, T., C. Narasimhan, and Y. Xie (2013). Treatment effectiveness and side effects: A model of physician learning. *Management Science 59*(6), 1309–1325.

Che, H., T. Erdem, and T. S. Öncü (2015). Consumer learning and evolution of consumer brand preferences. *Quantitative Marketing and Economics 13*(3), 173–202.

Ching, A. and H. Lim (2016). A structural model of correlated learning and late-mover advantages: The case of statins. *working paper*.

Ching, A. T. (2010a). Consumer learning and heterogeneity: Dynamics of demand for prescription drugs after patent expiration. *International Journal of Industrial Organization 28*(6), 619 – 638.

Ching, A. T. (2010b). A dynamic oligopoly structural model for the prescription drug market after patent expiration. *International Economic Review 51*(4), 1175–1207.

Ching, A. T., T. Erdem, and M. P. Keane (2013). Invited paper-learning models: An assessment of progress, challenges, and new developments. *Marketing Science 32*(6), 913–938.

Ching, A. T. and M. Ishihara (2012). Measuring the informative and persuasive roles of detailing on prescribing decisions. *Management Science 58*(7), 1374–1387.

Chintagunta, P. K., R. L. Goettler, and M. Kim (2012). New drug diffusion when forward-looking physicians learn from patient feedback and detailing. *Journal of Marketing Research 49*(6), 807–821.

Chintagunta, P. K., R. Jiang, and G. Z. Jin (2009). Information, learning, and drug diffusion: The case of cox-2 inhibitors. *Quantitative Marketing and Economics 7*(4).

Colella, F. and A. van Soest (2013). Time use, consumption expenditures and employment status: Evidence from the liss panel. In *7th MESS Workshop*.

Coscelli, A. and M. Shum (2004). An empirical model of learning and patient spillovers in new drug entry. *Journal of Econometrics 122*(2), 213–246.

Cox, D. R. and E. J. Snell (1989). *Analysis of binary data*, Volume 32. CRC Press.

Crawford, G. S. and M. Shum (2005). Uncertainty and learning in pharmaceutical demand. *Econometrica 73*(4), 1137–1173.

Cutler, D. M., E. L. Glaeser, and J. M. Shapiro (2003). Why have americans become more obese? *The Journal of Economic Perspectives 17*(3), 93–118.

DeGroot, M. H. (1970). *Optimal statistical decisions.* New York: McGraw-Hill,.

Dodson, J. A., A. M. Tybout, and B. Sternthal (1978). Impact of deals and deal retraction on brand switching. *Journal of Marketing Research 15*(1), 72–81.

Duernecker, G. and B. Herrendorf (2015). On the Allocation of Time - A Quantitative Analysis of the U.S. and France. CESifo Working Paper Series 5475, CESifo Group Munich.

Duncan Herrington, J. and L. M. Capella (1995). Shopper reactions to perceived time pressure. *International Journal of Retail & Distribution Management 23*(12), 13–20.

Einav, L., E. Leibtag, and A. Nevo (2010). Recording discrepancies in nielsen homescan data: Are they present and do they matter? *Quantitative Marketing and Economics 8*(2), 207–239.

Eisner, R. (1988). Extended accounts for national income and product. *Journal of Economic Literature 26*(4), 1611–1684.

Erdem, T. (1998). An empirical analysis of umbrella branding. *Journal of Marketing Research 35*(3), 339–351.

Erdem, T. and M. P. Keane (1996). Decision-making under uncertainty: Capturing dynamic brand choice processes in turbulent consumer goods markets. *Marketing Science 15*(1), 1–20.

Erdem, T., M. P. Keane, and B. Sun (2008). A dynamic model of brand choice when price and advertising signal product quality. *Marketing Science 27*(6), 1111–1125.

Francis, N. and V. A. Ramey (2009). Measures of per capita hours and their implications for the technology-hours debate. *Journal of Money, Credit and Banking 41*(6), 1071–1097.

Gimenez-Nadal, J. I. and A. Sevilla (2012). Trends in time allocation: A cross-country analysis. *European Economic Review 56*(6), 1338–1359.

Golder, P. N. and G. J. Tellis (1993, May). Pioneer advantage: Marketing logic or marketing legend? *Journal of Marketing Research 30*(2), 158–170.

Greenwood, J., A. Seshadri, and M. Yorukoglu (2005). Engines of liberation. *The Review of Economic Studies 72*(1), 109–133.

Gronau, R. (1977). Leisure, home production, and work–the theory of the allocation of time revisited. *Journal of Political Economy 85*(6), 1099–1123.

Gronau, R. and D. S. Hamermesh (2008). The demand for variety: A household production perspective. *The Review of Economics and Statistics 90*(3), 562–572.

Gross, B. L. and J. N. Sheth (1989). Time-oriented advertising: a content analysis of united states magazine advertising, 1890-1988. *The Journal of Marketing*, 76–83.

Hamermesh, D. S. (2005). Routine. *European Economic Review 49*(1), 29 – 53.

Hamermesh, D. S., H. Frazis, and J. Stewart (2005). Data watch: The american time use survey. *The Journal of Economic Perspectives 19*(1), 221–232.

Hauser, J., G. J. Tellis, and A. Griffin (2006). Research on innovation: A review and agenda for marketing science. *Marketing Science 25*(6), 687–717.

Heerde, H. J. V., M. J. Gijsenberg, M. G. Dekimpe, and J.-B. E. Steenkamp (2013). Price and advertising effectiveness over the business cycle. *Journal of Marketing Research 50*(2), 177–193.

Hotelling, H. (1929). Stability in competition. *The Economic Journal 39*(153), 41–57.

Huang, Y. and B. Bronnenberg (2017). Pennies for your thoughts: Costly product consideration and purchase quantity thresholds. *working paper*.

Hurd, M. and S. Rohwedder (2005, May). Changes in Consumption and Activities at Retirement. DNB Working Papers 039, Netherlands Central Bank, Research Department.

Hurst, E. (2008). The retirement of a consumption puzzle. Technical report, National Bureau of Economic Research.

Iyengar, R., A. Ansari, and S. Gupta (2007). A model of consumer learning for service quality and usage. *Journal of Marketing Research 44*(4), 529–544.

Jacoby, J., G. J. Szybillo, and C. K. Berning (1976). Time and consumer behavior: An interdisciplinary overview. *Journal of Consumer Research 2*(4), 320–339.

Joachimsthaler, E. A. and J. L. Lastovicka (1984). Optimal stimulation level-exploratory behavior models. *Journal of Consumer Research*, 830–835.

Jovanovic, B. (1984). Matching, turnover, and unemployment. *Journal of Political Economy 92*(1), 108–122.

Kawaguchi, D., J. Lee, and D. S. Hamermesh (2013). A gift of time. *Labour Economics 24*, 205–216.

Keller, K. L. and D. R. Lehmann (2006). Brands and branding: Research findings and future priorities. *Marketing Science 25*(6), 740–759.

Kelley, E. J. (1958). The importance of convenience in consumer purchasing. *The Journal of Marketing*, 32–38.

Kenhove, P. V. and K. De Wulf (2000). Income and time pressure: a person-situation grocery retail typology. *The International Review of Retail, Distribution and Consumer Research 10*(2), 149–166.

Kerin, R. A., P. Varadarajan, and R. A. Peterson (1992, October). First-mover advantage: A synthesis, conceptual framework, and research propositions. *Journal of Marketing 56*(4), 33–52.

Kimmel, J. (2008). *How do we spend our time?: evidence from the American time use survey*. WE Upjohn Institute.

Koopmans, T. C. (1949). Identification problems in economic model construction. *Econometrica, Journal of the Econometric Society*, 125–144.

Lee, J., D. Kawaguchi, and D. S. Hamermesh (2012). Aggregate impacts of a gift of time. *The American Economic Review 102*(3), 612–616.

Marcoul, P. and Q. Weninger (2008). Search and active learning with correlated information: Empirical evidence from mid-atlantic clam fishermen. *Journal of Economic Dynamics and Control 32*(6), 1921–1948.

Marinescu, I. (2016). Divorce: what does learning have to do with it? *Labour Economics 38*, 90–105.

Meyer, B. D. and J. X. Sullivan (2008). Changes in the consumption, income, and well-being of single mother headed families. *The American Economic Review 98*(5), 2221–2241.

Moscarini, G. (2005). Job matching and the wage distribution. *Econometrica 73*(2), 481–516.

Narayanan, S. and P. Manchanda (2009). Heterogeneous learning and the targeting of marketing communication for new products. *Marketing Science 28*(3), 424–441.

Nevo, A. and A. Wong (2015, July). The elasticity of substitution between time and market goods: Evidence from the great recession. Working Paper 21318, National Bureau of Economic Research.

Nowlis, S. M. (1995). The effect of time pressure on the choice between brands that differ in quality, price, and product features. *Marketing Letters 6*(4), 287–295.

Pashigian, B. P. and B. Bowen (1994). The rising cost of time of females, the growth of national brands and the supply of retail services. *Economic Inquiry 32*(1), 33–65.

Pozzi, A. (2012). Shopping cost and brand exploration in online grocery. *American Economic Journal: Microeconomics 4*(3), 96–120.

Rabin, M. (2000). Risk aversion and expected-utility theory: A calibration theorem. *Econometrica 68*(5), 1281–1292.

Raju, P. S. (1980). Optimum stimulation level: Its relationship to personality, demographics, and exploratory behavior. *Journal of Consumer Research 7*(3), 272–282.

Ramey, V. A. (2009). Time spent in home production in the twentieth-century united states: New estimates from old data. *The Journal of Economic History 69*(1), 1–47.

Ramey, V. A. and N. Francis (2009, July). A century of work and leisure. *American Economic Journal: Macroeconomics 1*(2), 189–224.

Roberts, J. H. and G. L. Urban (1988). Modeling multiattribute utility, risk, and belief dynamics for new consumer durable brand choice. *Management Science 34*(2), 167–185.

Rogers, E. M. et al. (1962). Diffusion of innovations. *Diffusion of innovations.*.

Salop, S. C. (1979). Monopolistic competition with outside goods. *The Bell Journal of Economics*, 141–156.

Shapiro, C. (1982). Advertising as a barrier to entry? *working paper*.

Shin, S., S. Misra, and D. Horsky (2012). Disentangling preferences and learning in brand choice models. *Marketing Science 31*(1), 115–137.

Solberg, E. J. and D. C. Wong (1992). Family time use: Leisure, home production, market work, and work related travel. *The Journal of Human Resources 27*(3), 485–510.

Spence, M. E. (1976, June). Product selection, fixed costs, and monopolistic competition. *The Review of Economic Studies 43*(2), 217–235.

Sridhar, K., R. Bezawada, and M. Trivedi (2012). Investigating the drivers of consumer cross-category learning for new products using multiple data sets. *Marketing Science 31*(4), 668–688.

Stancanelli, E. and A. van Soest (2012). Retirement and home production: A regression discontinuity approach. *American Economic Review 102*(3), 600–605.

Stratton, L. S. (2012). The role of preferences and opportunity costs in determining the time allocated to housework. *The American Economic Review 102*(3), 606–611.

Szymanowski, M. and E. Gijsbrechts (2012). Consumption-based cross-brand learning: Are private labels really private? *Journal of Marketing Research 49*(2), 231–246.

Szymanowski, M. and E. Gijsbrechts (2013). Patterns in consumption-based learning about brand quality for consumer packaged goods. *International Journal of Research in Marketing 30*(3), 219–235.

Thomas, J. (2017). The signal quality of grades across academic fields. *working paper*.

Venzon, D. and S. H. Moolgavkar (1988). A method for computing profile-likelihood-based confidence intervals. *Applied Statistics*, 87–94.

Vermeir, I. and P. Van Kenhove (2005). The influence of need for closure and perceived time pressure on search effort for price and promotional information in a grocery shopping context. *Psychology & Marketing 22*(1), 71–95.

Vroegrijk, M., E. Gijsbrechts, and K. Campo (2013). Close encounter with the hard discounter: A multiple-store shopping perspective on the impact of local hard-discounter entry. *Journal of Marketing Research 50*(5), 606–626.

Wathieu, L., A. Muthukrishnan, and B. J. Bronnenberg (2004). The asymmetric effect of discount retraction on subsequent choice. *Journal of Consumer Research 31*(3), 652–657.