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

Yan Xu† Bart J. Bronnenberg‡ Tobias J. Klein§

May 10, 2017

Abstract

We study consumer learning in a new consumer packaged goods category using purchasing data from a long balanced panel. The data are well-suited for this purpose because we observe consumers making their first purchases in the 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 several brands of an experience good. Their initial prior beliefs regarding the consumption utility they will experience when purchasing the products in the category 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. We allow all of these to be related to the time at which they first buy a product from the category. The value consumers attach to the product, together with the sensitivity to marketing variables, ultimately determines customer value. From a firm's perspective, it is important that promotions and product line length affect individual utility and thereby the inclination to buy a product, and thus also the speed at which consumers learn about their preference for the product. Estimating this structural learning model allows us to characterize learning effects and to perform counterfactual simulations. Based on those, we provide suggestions on optimal policy decisions.

Keywords: product adoption, new category growth, uncertainty, consumer learning.

---

*We are grateful to seminar/conference participants at the 2016 Marketing Science Conference, DIW Berlin, Erasmus University (Econometrics), Erasmus University (Marketing), Hong Kong Polytechnic University, Tilburg University, TU/e, and Toulouse Business School. We thank Aimark and GfK, and in particular Alfred Dijs and Bernadette van Ewijk, for providing the data. Funding by the Dutch Science Foundation (NWO) is gratefully acknowledged.

†PhD student, Department of Marketing, Tilburg University, Warandelaan 2, 5037 AB Tilburg, The Netherlands, y.xu_1@uvt.nl.

‡Professor, Department of Marketing, Tilburg School of Economics and Management, Warandelaan 2, 5037 AB Tilburg, The Netherlands, Bart.Bronnenberg@uvt.nl.

§Associate Professor, Department of Econometrics and OR, Tilburg School of Economics and Management, Warandelaan 2, 5037 AB Tilburg, The Netherlands, T.J.Klein@uvt.nl.

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 perspectives 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 be 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. Our balanced panel data are 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 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 the 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. Our results show that there are considerable differences both within and 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 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. Early contributions such as Stoneman (1981), Jovanovic (1982) and Meyer and Sathi (1985) model how past experiences can affect an agent's decision when he faces uncertainty. Erdem and Keane (1996) use a Bayesian learning model to characterize consumer brand choice under uncertainty, followed by Coscelli and Shum (2004), Crawford and Shum (2005), Israel (2005), Narayanan et al. (2005), Coscelli and Shum (2004); Crawford and Shum (2005); Israel (2005); Narayanan et al. (2005); Chintagunta et al. (2009), Narayanan and Manchanda (2009), Osborne (2011), Shin et al. (2012), Chintagunta et al. (2012), and Szymanowski and Gijsbrechts (2012). Ching et al. (2013) provide a recent review. Similar to Coscelli and Shum (2004), Israel (2005) and Shin et al. (2012), we model consumers as risk neutral Bayesian learners who may have biased perception of the product at the initial stage. We build a bridge between this 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 consumer's 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. Bronnenberg et al. (2015) document that consumers are willing to pay more for national brands. By estimating our structural learning model, we provide an alternative explanation: 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 the one of a competing product.

The remainder of the paper is structured as follows. Section 2 describes the data. In Section 3, we present model-free evidence that motivates our structural model. Section 4 describes our structural learning model. Section 5 provides details on the empirical implementation and a discussion of identification. Section 6 presents the estimation results. Model predictions, counterfactual experiments and implications are collected in Section 7. Section 8 concludes.

2 Data

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.[1] 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.

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

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 expansion in the extensive margin with a large group of new consumers adopting the category during our observation window.[2] 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 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.[3]

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

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

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

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 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 meal 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 meal every 17 weeks on average (the median is 7 weeks). We use 26 weeks as our the "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.[4]

We observe 1599 consumers who adopt the category between 2001 and 2008.[5] 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.[6]

[4] We did robustness checks by varying the filer rule between 20 and 36 weeks. The main patterns in our model-free description of the data (see Figure 1, Figure 2, and Figure 3) remain unchanged.

[5] 138 households are dropped because of the 26-week "filter rule".

[6] Based on the summary statistics in Table 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 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.

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. Table 1 presents summary statistics at cohort level.[7] Figure 2 in Section 3 below shows a distribution of consumers adoption timing and the thresholds of adopter cohorts.

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, 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 private labels. Knorr is manufactured by Unilever.

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,

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

Table 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.

only recently did the category develop into a major category.

Table 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.

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 panel, assuming, as is reasonable for the Netherlands, that a retailer charge common prices across outlets of the same chain.[8] 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

[8]We have verified and confirmed this for the top six national chains (Albert Heijn, C1000, Super De Boer, Plus, Jumbo, and Hoogvliet) that jointly take up about 70% market share.

Table 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.

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 UPC's 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 UPC's in a given week and the flavors offered for each brand are altered slowly over time. Thereby, we can capture the observed trend in variety over the years.

Table 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.[9] Product line length as measured by the available variety increase over time and at each point in time, the pioneer brand offers more variety.

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

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.

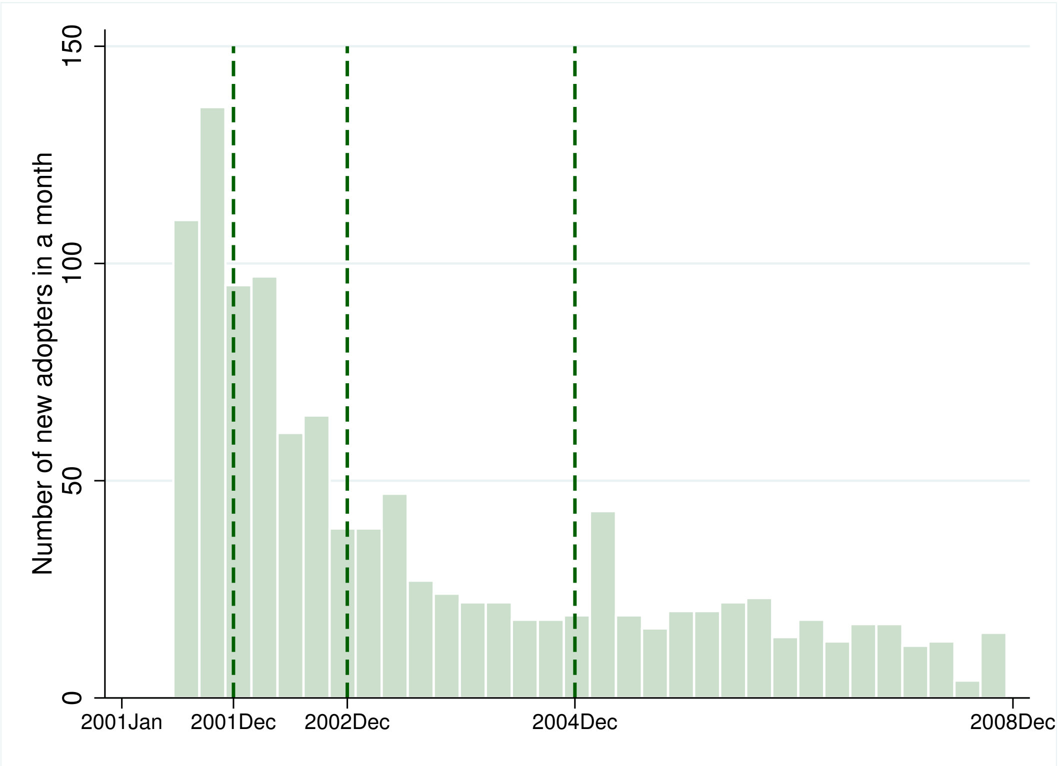
3.1 Market expansion and the evolution of brand sales

Figure 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 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, 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.

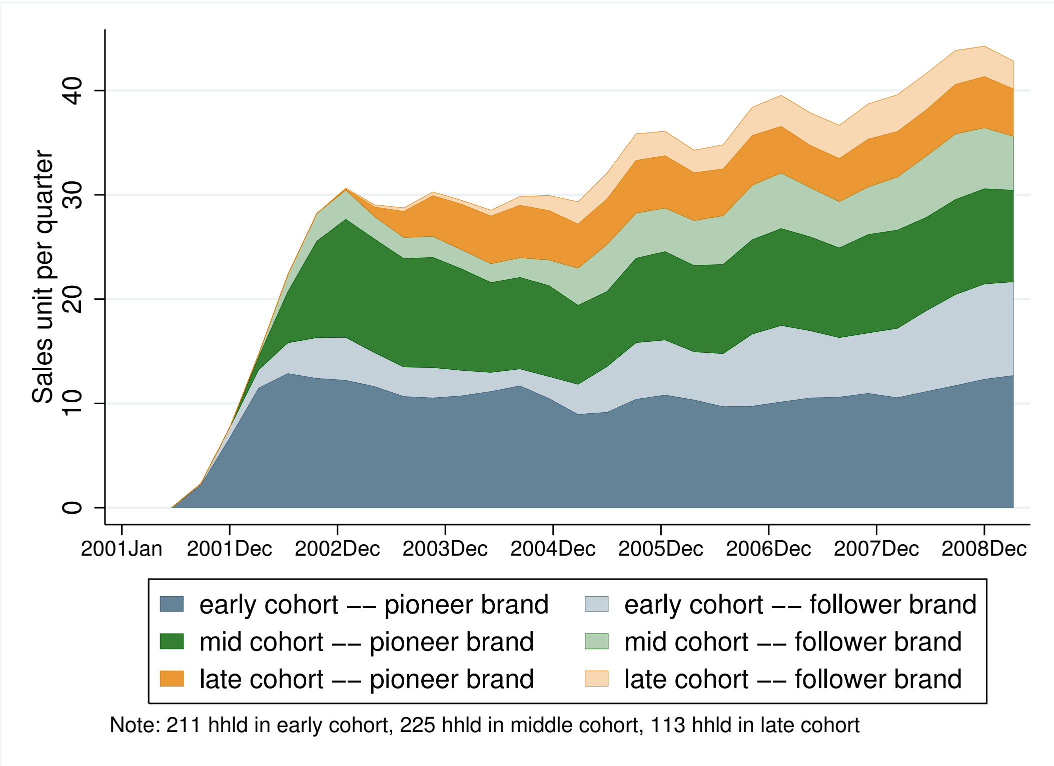
Next, in Figure 3, we plot the unconditional purchase shares of both brands in any given week and for each cohort. We call attention to three features of these plots. First, the unconditional shares of the pioneer brand among all the adopter cohorts decrease over time, while the market shares of the follower brand increase over time. Second, the average market shares differ across cohorts, with the early cohort having higher purchase

Figure 1: Extensive margin expansion



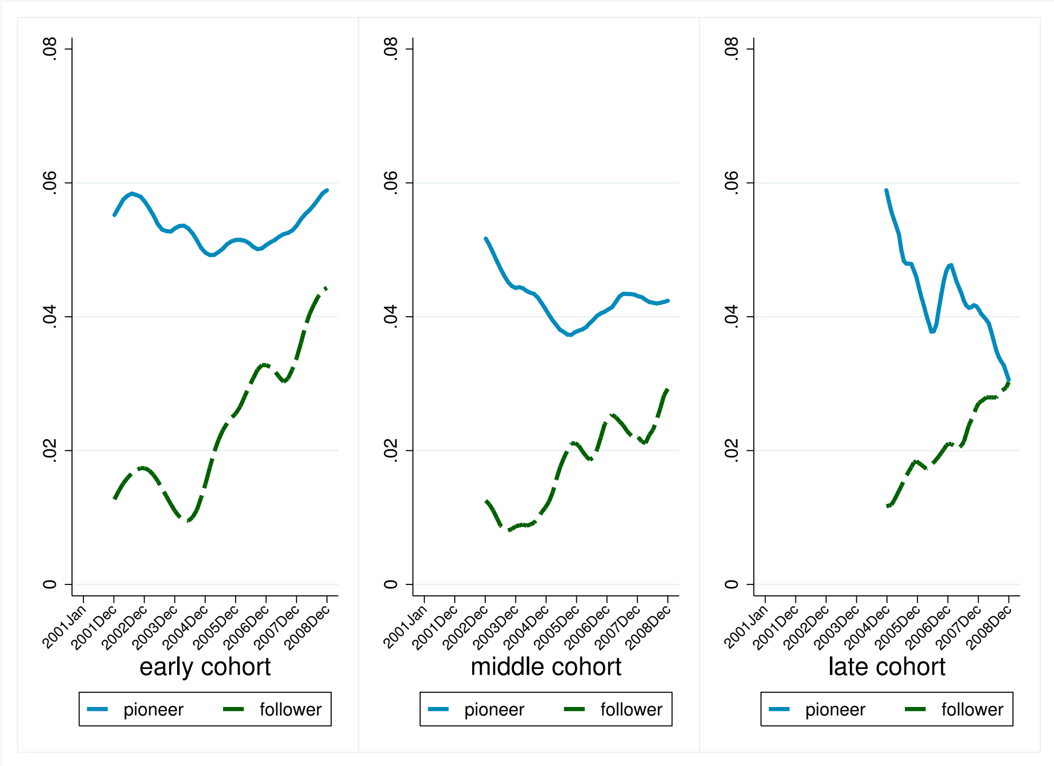
Notes: We grouped consumers who adopted the category between week 27 of 2001 and 2004 into three adopter cohorts and refer to them as early, middle and late cohort. Each cohort has a similar number of households.

Figure 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.

Figure 3: Purchase shares by cohort and time



Notes: Figure 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.

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 of 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 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.

4 The structural learning model

4.1 Brand choice decisions

The model introduced below is a Bayesian learning model of brand choice (see e.g., Erdem and Keane 1996). Consumers are assumed to have heterogeneous valuations 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,

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

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.

4.2 Consumption utility specification

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

$$u_{ijt} = \underbrace{q_{ijt} + \lambda_{ij}}_{\text{experienced match value}} + \alpha_i p_{ijt} + \omega_i x_{ijt} + \epsilon_{ijt}, \tag{2}$$

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

$$u_{i0t} = \epsilon_{i0t}. \tag{3}$$

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 consumer's brand preferences.

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 $q_{ijt'}$, with $t' < t$.

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.

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

$$I_{ij0} = \mathcal{N}\left(\mu_{ij0},\ \sigma^2_{ij0}\right) \tag{4}$$

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}. The standard deviation, σ^2_{ij0}, measures the accuracy of the consumer's prior belief. 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^2_{ν}. Further, the signals are independent and distributed normally across periods and individuals, i.e.,

$$q_{ijt} = \mu_{ij} + \nu_{ijt}; \quad \nu_{ijt} \sim \mathcal{N}\left(0,\ \sigma_\nu^2\right). \tag{5}$$

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

$$\mu_{ijt} = \begin{cases} \left(\frac{1}{\sigma_{ijt-1}^2} + \frac{1}{\sigma_\nu^2}\right)^{-1} \left(\frac{1}{\sigma_{ijt-1}^2}\mu_{ijt-1} + \frac{1}{\sigma_\nu^2} q_{ijt-1}\right) & \text{if } d_{it-1} = j \\ \mu_{ijt-1} & \text{if } d_{it-1} \neq j \end{cases} \tag{6}$$

and

$$\sigma_{ijt}^2 = \begin{cases} \left(\frac{1}{\sigma_{ijt-1}^2} + \frac{1}{\sigma_\nu^2}\right)^{-1} & \text{if } d_{it-1} = j \\ \sigma_{ijt-1}^2 & \text{if } d_{it-1} \neq j \end{cases}. \tag{7}$$

From the above updating equations, we see that the uncertainty about the true match value σ_{ijt}^2 diminishes from consumption as long as the signal variance σ_ν^2 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.

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

$$\begin{aligned} E(u_{ijt}|I_{ijt}) &= E(q_{ijt}|I_{ijt}) + \lambda_{ij} + \omega_i x_{ijt} + \alpha_i p_{ijt} + \epsilon_{ijt} \\ &= \mu_{ijt} + \lambda_{ij} + \omega_i x_{ijt} + \alpha_i p_{ijt} + \epsilon_{ijt} \end{aligned} \tag{8}$$

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, $\eta_{ij0} > \mu_{ij}$, 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, $\eta_{ij0} < \mu_{ij}$, 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:

$$\text{Prob}(d_{it} = j) = \frac{\exp(\mu_{ijt} + \lambda_{ij} + \omega_i x_{ijt} + \alpha_i p_{ijt})}{1 + \sum_{j=1,2} \exp(\mu_{ijt} + \lambda_{ij} + \omega_i x_{ijt} + \alpha_i p_{ijt})}. \tag{9}$$

5 Implementation

5.1 The empirical model

The goal of our empirical analysis is to characterize the evolution of brand preferences for each cohort, controlling for differences in permanent taste; and to measure how the response to marketing activities differs across cohorts.

Guided by 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.

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 consumer's initial belief of brand j; (2) σ^2_{cj0}, a cohort-specific variance of 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 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, parameter (3)-(7), the best data source is long term purchase 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 consumer's initial belief. The identification primarily comes from how 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 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.

5.3 Estimation

The primary complication in estimation is consumer heterogeneity and the number of unobservables (to the researchers). Besides the per-period logit errors, we do not observe a consumer's individual specific brand preference, λ_i, and his realizations of the non-deterministic part (ν_{ijt}) in the consumption experiences (q_{ijt}) at each purchase occasion. Using that the unobservables are assumed uncorrelated and are independent of the observables, we estimate our model by maximizing the simulated log likelihood, where the likelihood contributions are at the individual level and given by the probability to observe the entire sequence of choices, integrating over the persistent unobservables and the consumption signals of the entire sequence of a consumer's 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,

$$\text{Prob}\left(d_{it}=j|z_{it},\ I_{it};\ \theta\right)=\text{Prob}\left(d_{it}=j|d_{it}^{t-1},\ z_{it},\ q_{it}^{t-1};\ \theta\right), \tag{10}$$

where d_{it} is consumer i's observed choice in period t; $z_{it}=\{p_{it},\ x_{it}\}$, 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, the upper subscript on the right-hand side $t-1$ indicates histories up to the time period $t-1$ (e.g., $q_{it}^{t-1}=\{q_{i1},\ q_{i2},\ ...,\ q_{it-1}\}$ and $d_{it}^{t-1}=\{d_{i1},\ d_{i2},\ ...,\ d_{it-1}\}$). The analyst does not observe the consumption signals q_{ijt}. To account for their influence define ν_{ijt} as simulation draws of q_{ijt} which are i.i.d. across time periods, individuals and brands. Then we can define a draw from Equation (10) as $\text{Prob}\left(d_{it}=j|d_{it}^{t-1},z_{it},\nu_{it}^{T_i-1};\theta\right)$ and by integrating over all unknowns, we can define the likelihood contribution of one individual as:

$$\mathscr{L}_i\left(\theta|d_{it}^{T_i},\ z_{it}^{T_i}\right)=\int\left(\Pi_{t=1}^{T_i}\text{Prob}\left(d_{it}|d_{it}^{t-1},\ z_{it},\ \nu_{it}^{T_i-1};\ \theta\right)dF\left(\nu_{it}^{T_i-1},\ \alpha_i,\ \lambda_i\right)\right) \tag{11}$$

in which T_i is the total number of shopping trips made by consumer i. For each consumer, we observe a sequence of choices from $t=1$ to $t=T_i$. The draw of the consumption signals, $\nu_{it}^{T_i-1}$, individual price sensitivity, α_i, and the idiosyncratic brand taste, λ_{ij}, are consumers' private information that are not observed by the researchers. Therefore, the dimension of the integral in Equation (11) is T_i+1. We use simulation techniques to evaluate these integrals, and estimate our model using simulated maximum likelihood.

For the simulated maximum likelihood estimation, we employ a partly analytical approach. We first draw S vectors of the unobservables, $\left(\nu_{it}^{T_i-1},\ \alpha_i,\ \lambda_i\right)$, for each consumer

i from the distribution $F(\cdots)$. Then the simulated likelihood for consumer i is:

$$\mathscr{L}_i\left(\theta|d_{it}^{T_i},\ z_{it}^{T_i}\right)=\frac{1}{S}\sum_{s=1}^{S}\left[\left(\Pi_{t=0}^{T_i}\text{Prob}^s\left\{d_{it}|d_{it}^{t-1},\ z_{it}^{t}\right\}\right)|\left(\nu_{i1}^s,\ \nu_{i2}^s,\ ...,\ \nu_{iT_i}^s,\ \alpha_i^s,\ \lambda_i^s;\ \theta\right)\right] \tag{12}$$

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.

6 Estimation results

Table 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, $\hat{\mu}_{c10} > \hat{\mu}_{c1}$, 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

Table 4: Estimates (part 1 of 2): learning parameters

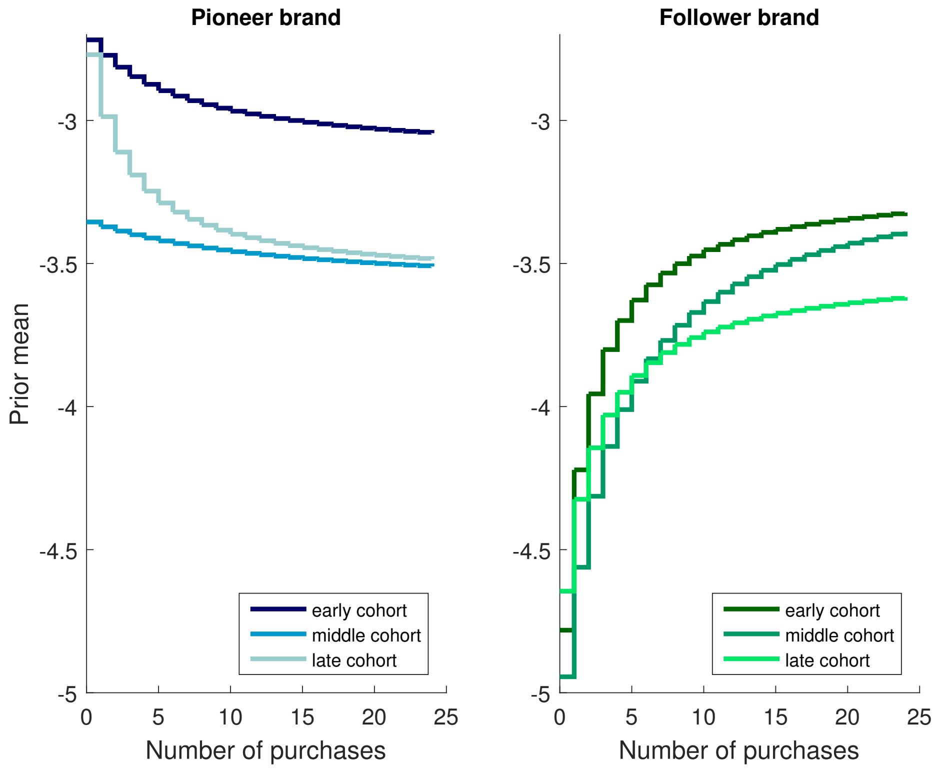
	par. est.	std. err.
initial brand-cohort belief mean:		
μ_{110}	-2.718	0.080
μ_{210}	-3.355	0.086
μ_{310}	-2.770	0.153
μ_{120}	-4.781	0.099
μ_{220}	-4.944	0.093
μ_{320}	-4.645	0.130
true brand-cohort match value:		
μ_{11}	-3.136	0.080
μ_{21}	-3.596	0.126
μ_{31}	-3.565	0.126
μ_{12}	-3.212	0.153
μ_{22}	-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_ν, to 0.5.

leads to interesting implications of price and promotion induced consumption experiences 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

Figure 4: Belief updating



category, with consumers who hold a high value adopting earlier and buying more.

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 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 match value, while the widths represent the consumer's uncertainty (variance) about variance. 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,

Table 5: Estimates (part 2 of 2): remaining parameters

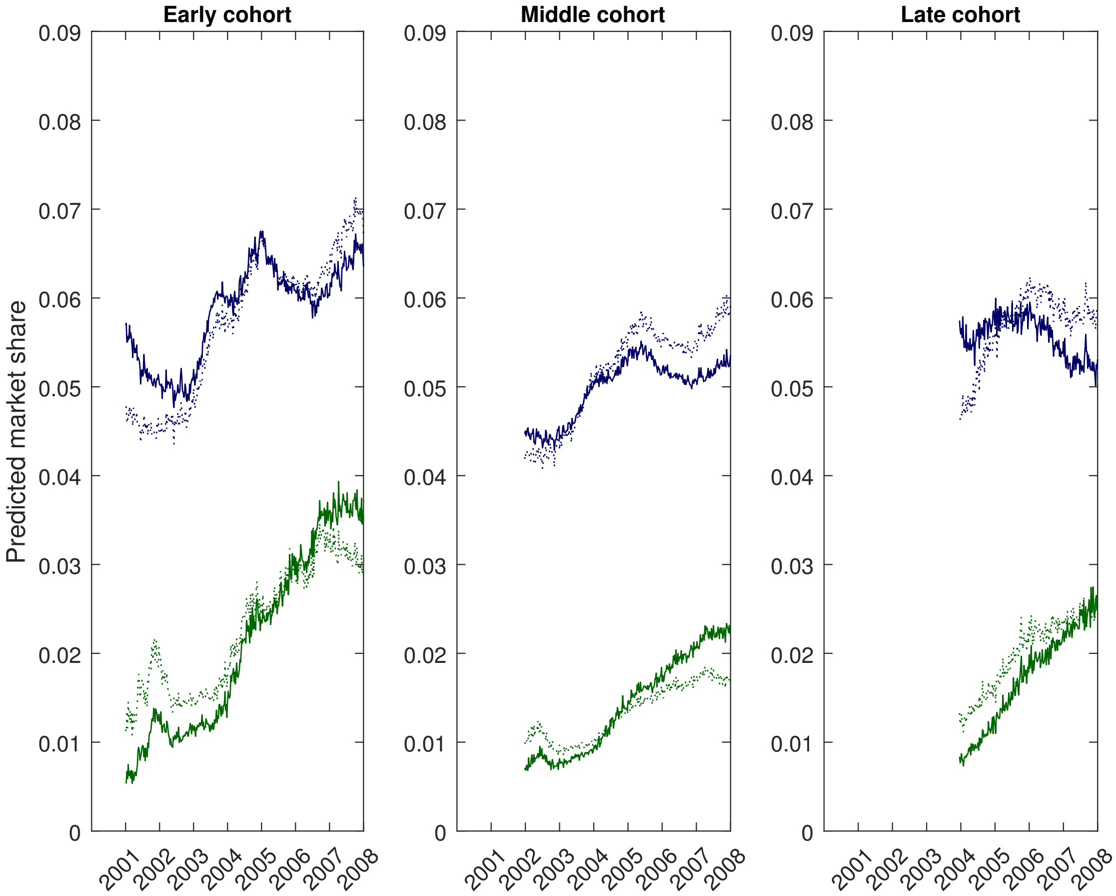
	learning model estimates	
	par. est.	std. err.
price coefficient:		
α_1	-1.082	0.068
α_2	-0.790	0.074
α_3	-1.238	0.115
$\sigma_{\alpha_1} = \sigma_{\alpha_2} = \sigma_{\alpha_3}$	0.871	0.048
variety coefficient:		
μ_{ω_1}	0.029	0.002
μ_{ω_2}	0.020	0.003
μ_{ω_3}	0.028	0.004
heterogeneity in permenant 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	NaN
number simulation draws	40.000	NaN

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 cohort (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). Estimates of these parameters are reported in Table 5. They show that the second cohort is the least price sensitive and also values variety the least. Moreover, within cohort heterogeneity in permanent tastes is lowest.

As a first way of illustrating the contribution 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

Figure 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 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.

that we (wrongly) impose that all learning has already taken place. Details are provided in Appendix B. Figure 5 shows that the static model will not be able to capture changes in market shares over time as well as the dynamic model does. More importantly, the static model will not allow us to conduct counterfactual experiments related to learning such as the ones presented in the following section.

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.

Market share dynamics Figure 6 shows the resulting evolution of market shares over time, by cohort.[10] 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 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 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.

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

Figure 6: Predicted market shares holding supply side unchanged over time

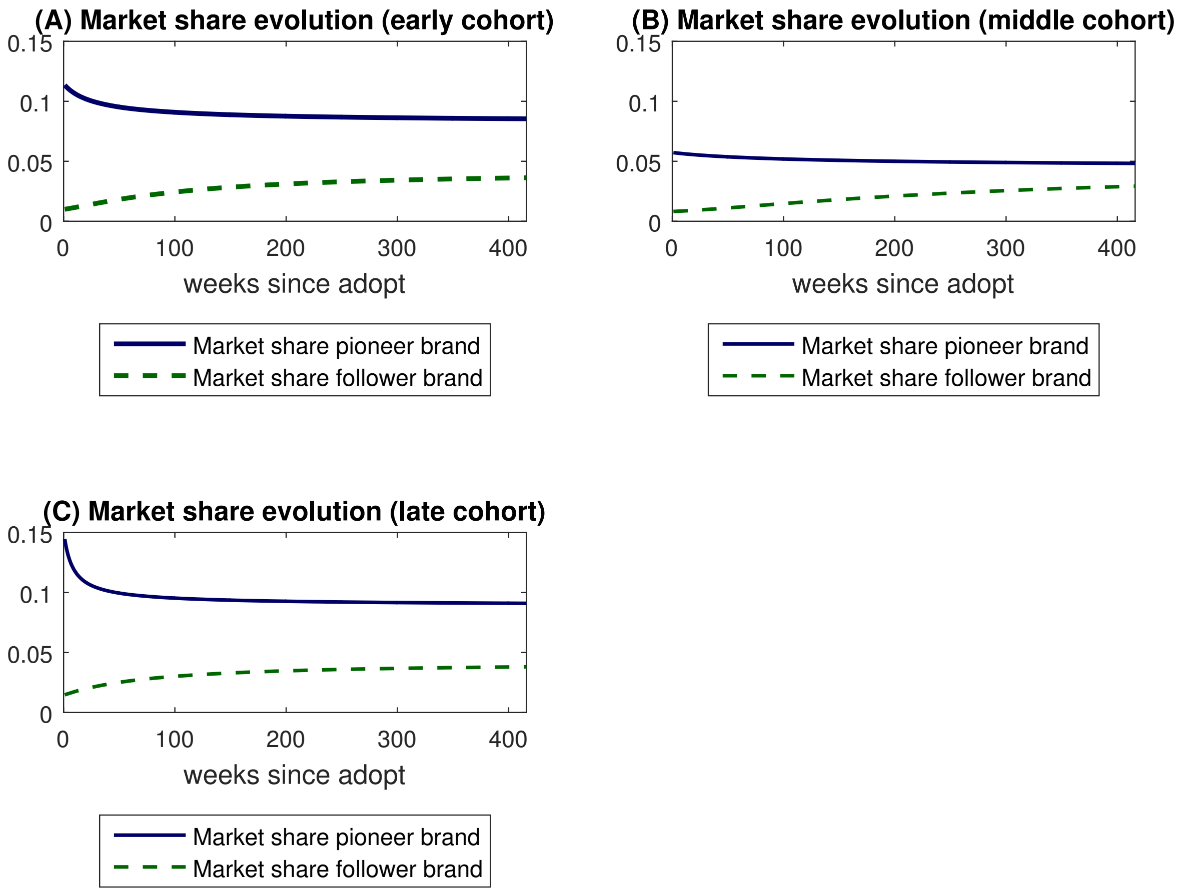
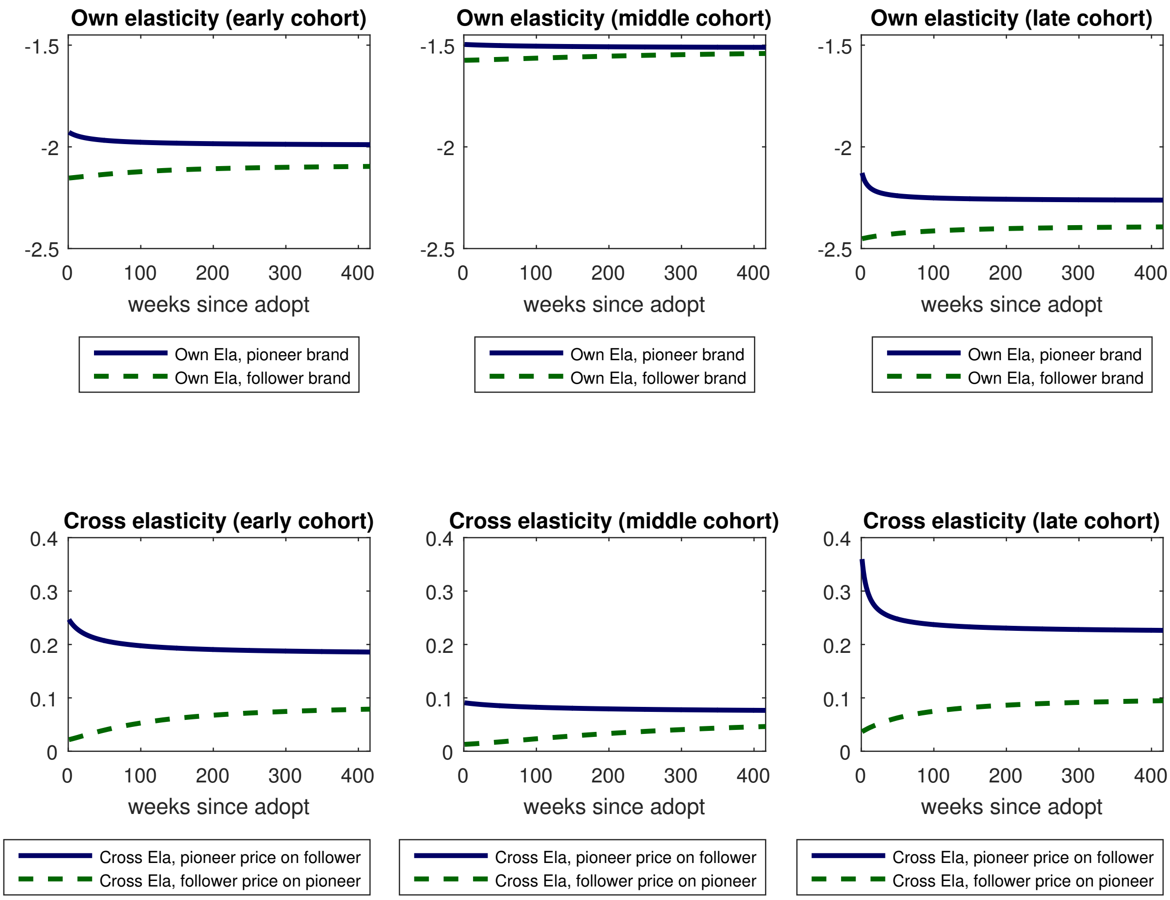


Figure 7: Predicted elasticities holding supply side unchanged over time



Elasticities and learning Segueing into the effect of marketing strategies and how they interact with consumer learning, we seek to understand the effects of price promotions and the implications of consumer learning on 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 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.

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.

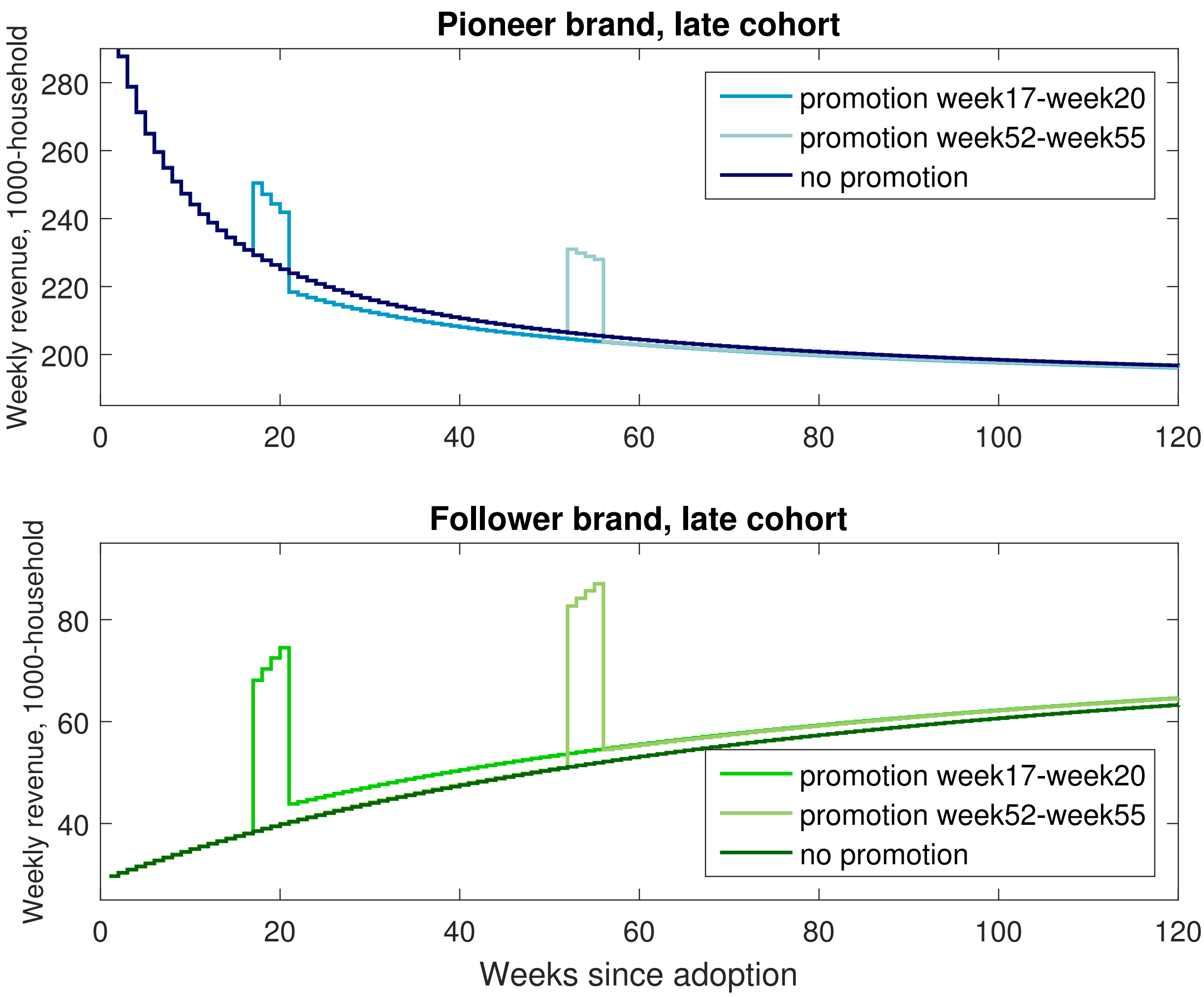
To see the long run effect of temporary price promotions, we calculate weekly revenue for the promoting brand from each cohort starting from the adoption week up to 250 weeks later.[11]

Figure 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 literature on promotion attractions, some debate exists about whether such effects are positive or negative. In our learning framework, promotions stimulate consumption

[11]From our model-free evidence and estimation results, we see clear evidence that consumers are initially pessimistic about the follower brand, and a significant change in their purchase intention of the follower brand after becoming experienced. Hence, we choose to model consumer as if they have biased mean perception rather than have rational expectation. We aware of the possibility that new brand's price promotion could add extra "noise" to consumer's perception and may have negative effect when price signals quality (Erdem et al. (2008)). Erdem et al. (2008) models consumers as if they have rational expectation and quantifies the long term (negative) effect when price signals quality. We view our study as a complementary study of Erdem et al. (2008) on long term effect of temporary price promotion. For the follower brand, there is trade-off between "low price signal bad quality" and "stimulate pessimistic consumers' learning process"; which effect is the dominating one largely depends on consumer's initial belief. For the pioneer brand, those two effects are in the same direction—temporary price promotion may have negative long term effect.

Figure 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 D 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 6: Effects of price promotions

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

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.

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 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 term will lead to higher sales in the future.

8 Concluding remarks

In this paper, we study the relationship between adoption timing, consumer learning, and customer value in a new experience goods category, boxed meals.

Using 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 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.

In our current analysis, we took consumer's category adoption timing as given and estimate the cohort-specific primitives. In future work, we are interested in modeling the category adoption timing explicitly. Also, the boxed meal category is also one of the growing convenience goods categories. In future research, we will study which consumer characteristics, e.g., demographics, predict demand for such categories.

References

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.

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

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.

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.

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.

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. *QME 7*(4), 399–443.

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.

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

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.

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.

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

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.

Israel, M. (2005). Services as experience goods: An empirical examination of consumer learning in automobile insurance. *American Economic Review*, 1444–1463.

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

Jovanovic, B. (1982). Selection and the evolution of industry. *Econometrica: Journal of the Econometric Society*, 649–670.

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.

Meyer, R. J. and A. Sathi (1985). A multiattribute model of consumer choice during product learning. *Marketing Science 4*(1), 41–61.

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

Narayanan, S., P. Manchanda, and P. K. Chintagunta (2005). Temporal differences in the role of marketing communication in new product categories. *Journal of Marketing Research 42*(3), 278–290.

Osborne, M. (2011). Consumer learning, switching costs, and heterogeneity: A structural examination. *Quantitative Marketing and Economics 9*(1), 25–70.

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

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

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

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

Stoneman, P. (1981). Intra-firm diffusion, bayesian learning and profitability. *The Economic Journal 91*(362), 375–388.

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

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.

Appendices

A Additional summary statistics

Table 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 1 in the main text, but without dropping households that did not have more than two purchases.

B Static benchmark model

Figure (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 (13). Specifically, we use

$$u_{ijt} = \phi_{ij} + \alpha_i p_{ijt} + \omega_i x_{ijt} + \epsilon_{ijt} \tag{13}$$

and

$$\begin{cases} \phi_{ij} \sim \mathcal{N}\left(\mu_{\phi_{cj}}, \sigma^2_{\phi_j}\right) \\ \alpha_i = \alpha_c + a_i, \qquad\qquad a_i \sim \mathcal{N}\left(0, \sigma^2_{a_c}\right). \\ \omega_i = \omega_c \end{cases} \tag{14}$$

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 8 contains the resulting static model estimates.

Table 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	NaN
SimulationDrawNum	40.000	NaN

C Effect of purchase experiences on market shares and price elasticities

In Section 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 9 does so for the probability to buy either of the brands and Figures 10 through 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 7.

Figure 9: Dependence on choice probability on purchase experience

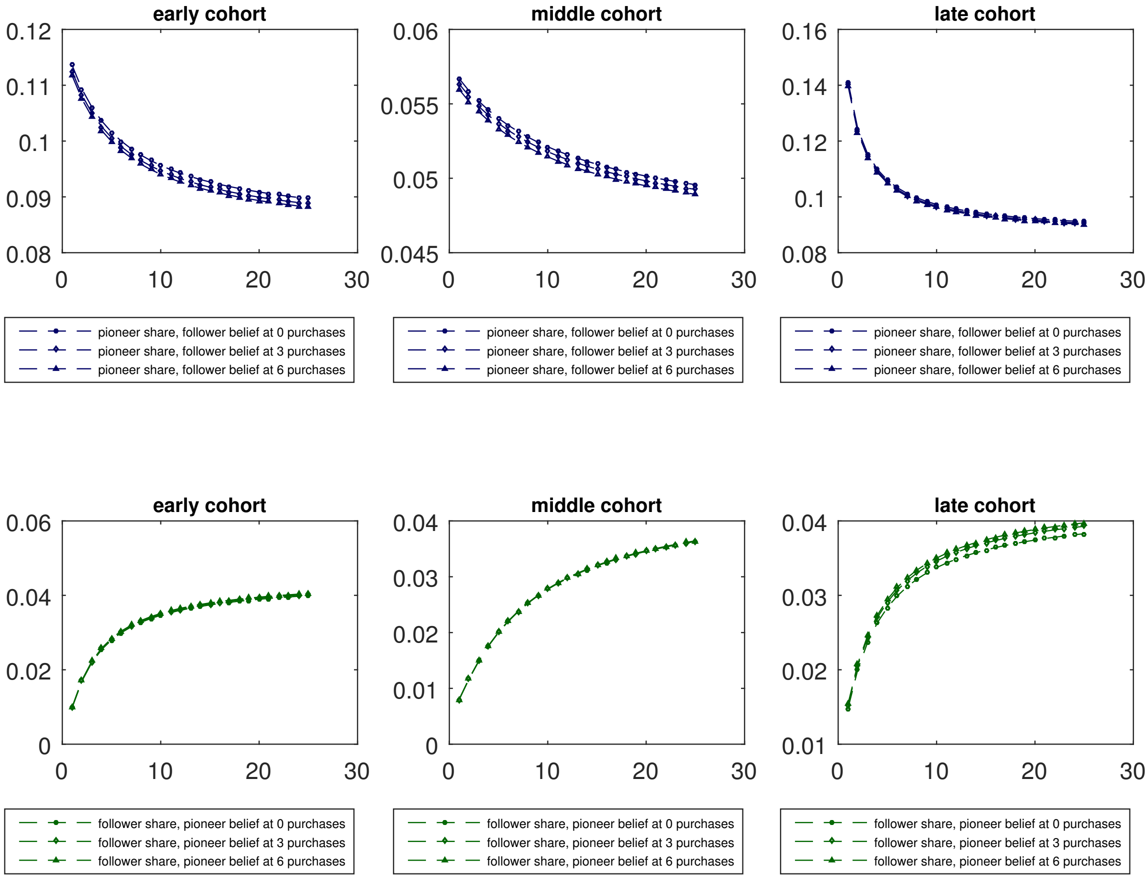


Figure 10: Dependence on elasticity on purchase experience—early cohort

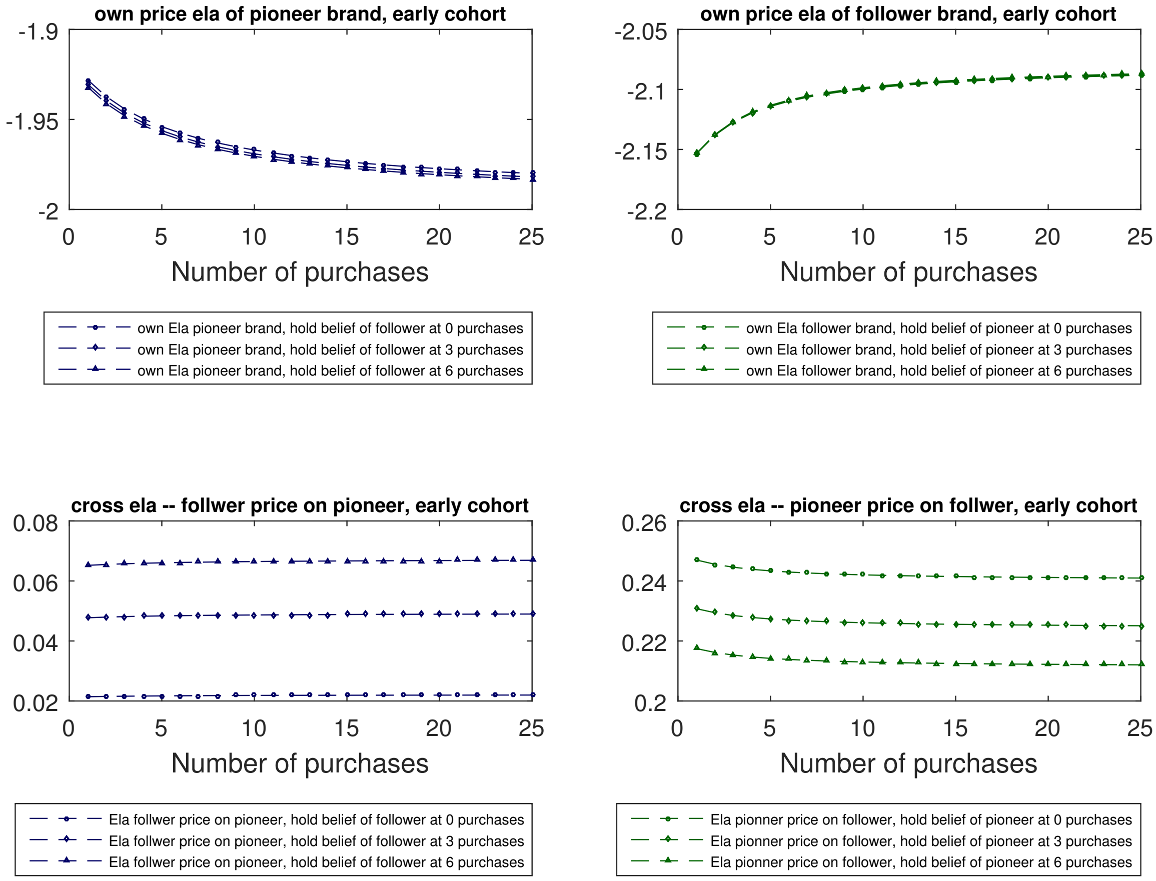


Figure 11: Dependence on elasticity on purchase experience—middle cohort

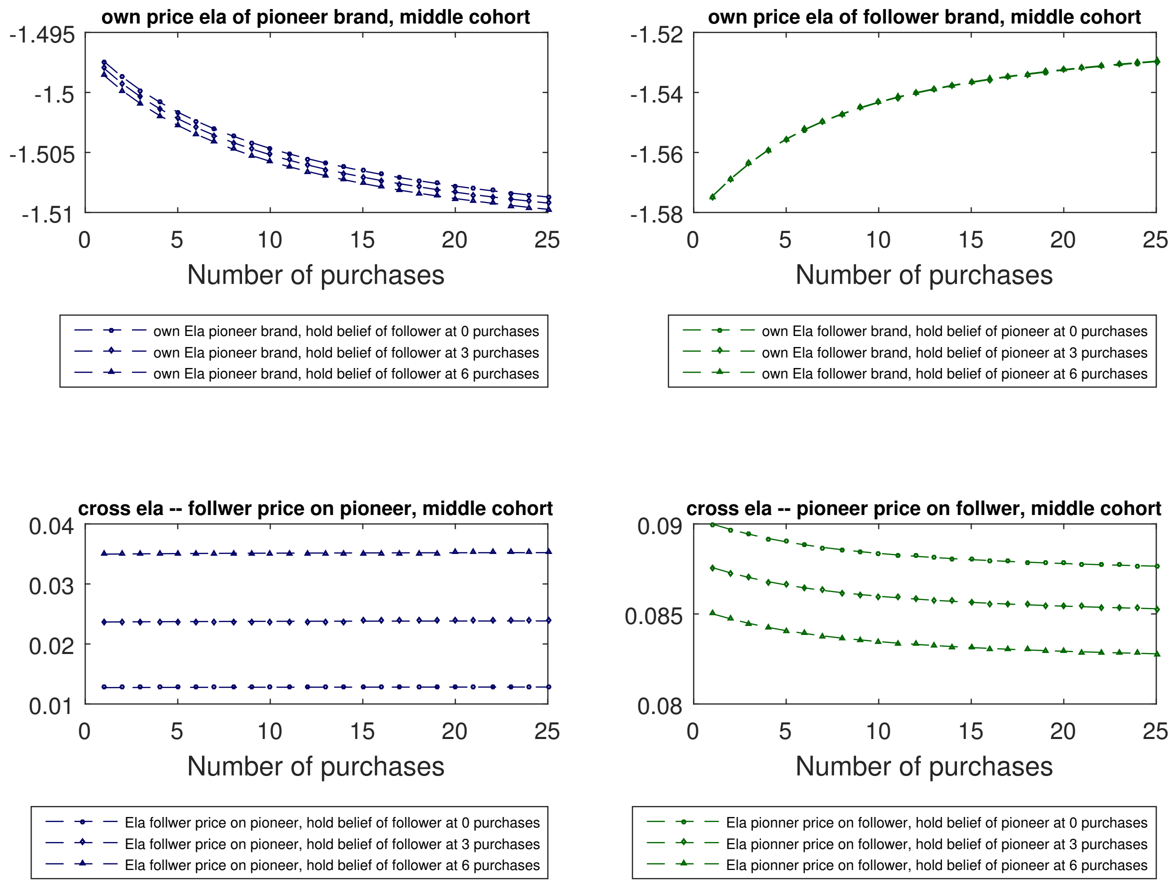
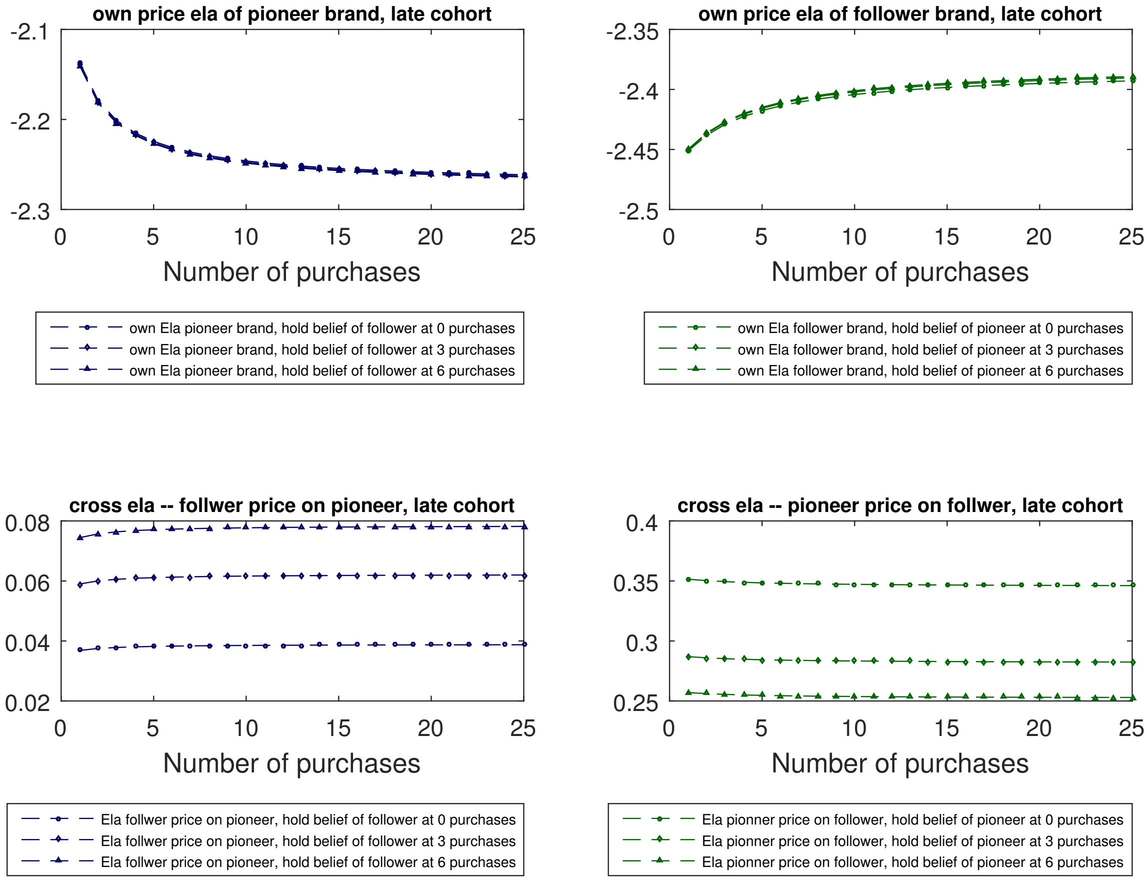
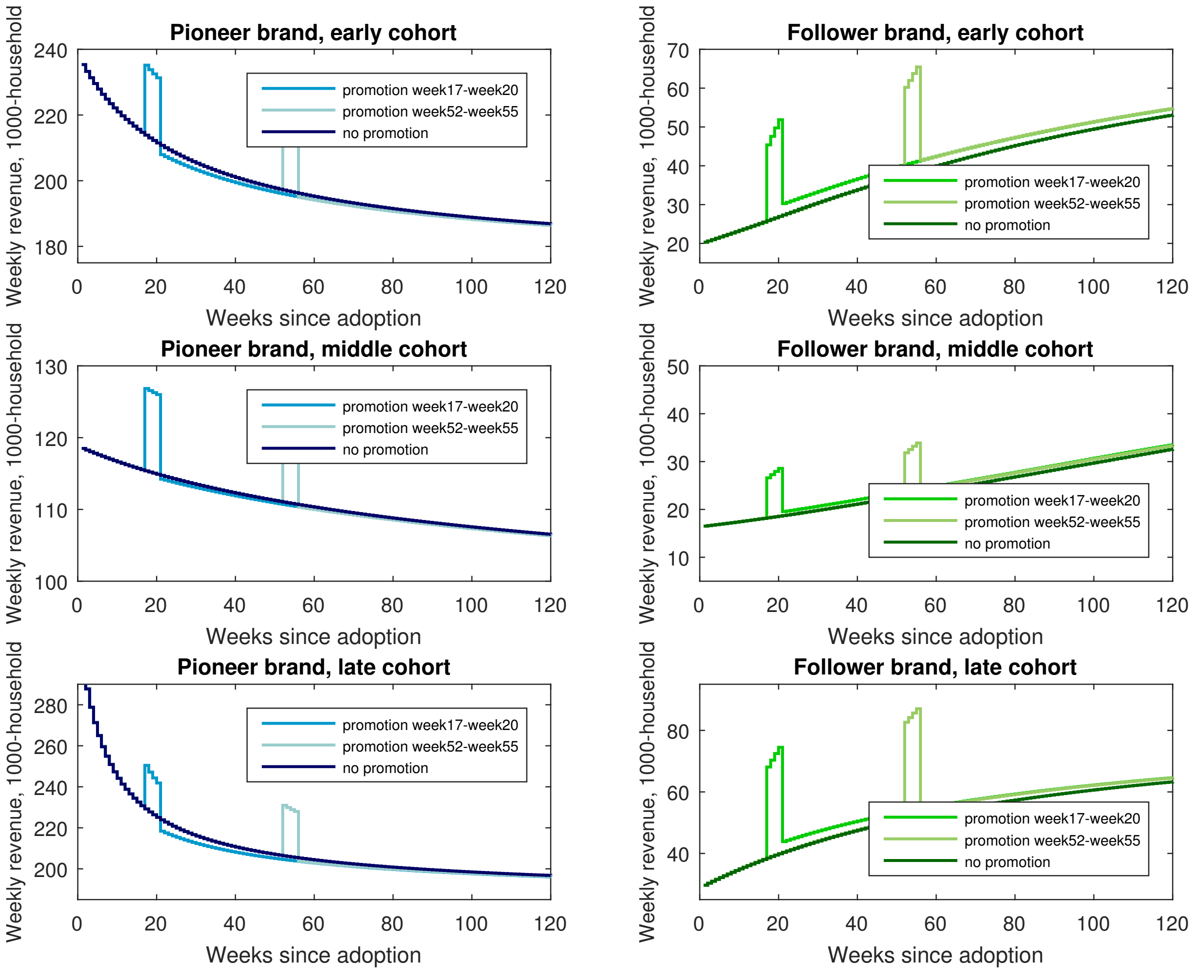


Figure 12: Dependence on elasticity on purchase experience—late cohort



D Price promotion experiment

Figure 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.