How Do Firms Build Market Share?*

David Argente†
Pennsylvania State University

Doireann Fitzgerald‡
Federal Reserve Bank of Minneapolis

Sara Moreira§
Northwestern University

Anthony Priolo¶
Lancaster University

March 2021

Abstract

We build a new data set to show that successful entrants in the consumer food sector build market share by adding new customers. They reach new customers by entering new geographical markets, placing their product in more stores in these markets, and by advertising direct to customers in markets where their product is available. We find no evidence that entrants manipulate markups to build market share. We estimate a structural model of endogenous customer base acquisition through marketing and advertising to match these facts. Our estimates suggest that the accumulation of customer base is subject to frictions which mean that entrant growth is a drawn-out process. This process generates market shares which are much more dispersed than underlying firm heterogeneity in cost or product appeal.

*We are grateful to the seminar and conference participants at SED, NYU, Columbia, Virtual International Trade and Macro Seminar, FRB STL Macro-Trade Workshop, University of Michigan, Johns Hopkins SAIS, Auburn University. The views expressed herein are those of the authors and not necessarily those of the Federal Reserve Bank of Minneapolis or the Federal Reserve System. Researcher(s) own analyses calculated (or derived) based in part on (i) retail measurement/consumer data from Nielsen Consumer LLC ("NielsenIQ"); (ii) media data from The Nielsen Company (US), LLC ("Nielsen"); and (iii) marketing databases provided through the respective NielsenIQ and the Nielsen Datasets at the Kilts Center for Marketing Data Center at The University of Chicago Booth School of Business. The conclusions drawn from the NielsenIQ and Nielsen data are those of the researcher(s) and do not reflect the views of Nielsen. Nielsen is not responsible for, had no role in, and was not involved in analyzing and preparing the results reported herein.

†Email: dargente@psu.edu. Address: 403 Kern Building, University Park, PA 16801.

‡Email: doireann.fitzgerald@gmail.com. Address: 90 Hennepin Ave, Minneapolis, MN 55401.

§Email: sara.moreira@kellogg.northwestern.edu. Address: 2211 Campus Drive, Evanston, IL 60208.

¶Email: a.priolo@lancaster.ac.uk. Address: Bailrigg, Lancaster, LA1 4YX, United Kingdom.

1 Introduction

Firms are mostly born small. Those which survive typically grow, initially fast, but then more slowly as they age. There is an extensive literature which relies on the productivity process and frictions in capital accumulation to explain these facts. However more recently, an empirical literature has pointed to an important role for customer base (also known as "demand" or "appeal")[1] in accounting for differences in firm size, both in the cross section, and within a given firm over time. Our goal in this paper is to understand how entrants accumulate customer base. To do this, we build a new data set on consumer food at the firm-product level. Although consumer food is a relatively narrow segment of the economy, accounting for around 5% of U.S. GDP, it has the advantage that we have access to high quality data on quantities sold, retail and wholesale prices, customer purchases, retail store presence, and advertising, all at the level of disaggregated geographies. These data are based on concording Nielsen Retail Scanner data, the Nielsen Household Panel, Nielsen Ad Intel data, and Nielsen Promo data.

The geographical dimension of the data is very useful, since the market for consumer food is geographically segmented. We first document that entering firms grow by entering new geographical markets as well as by expanding sales within their continuing markets. The new markets margin contributes on average 30% of sales growth over the first four years of firms which survive at least five years. Since markets are geographically segmented from the customer perspective, expanding the geographical reach of a firm's sales by definition involves reaching new customers. Of course, gradual expansion across markets could be driven by changes in productivity, or adjustment costs at the firm level such as financial frictions. But the fact that entry is staggered across segmented geographical markets within a firm allows us to isolate the relationship between within-market sales and a firm's history of participation in that particular market, conditioning on firm-level factors which are common across all markets irrespective of the timing of entry. We find that within-market sales growth is very fast just after entry into a market, and slows as the firm ages in the relevant market. This is consistent with an important role for accumulation of customers. Indeed, 2/3 of sales growth after entry into a market is due to the extensive margin of customers, with only 1/3 due to higher purchases per consumer.

Next, we examine what kinds of actions entrants take in order to accumulate customers. An extensive literature in macroeconomics emphasizes the potential role of dynamic pricing.In one version of this story, a firm charges low markups to new customers, and later raises markups once these same customers are "locked in." In the second version, entrants exploit

[1] See Foster, Haltiwanger and Syverson (2008) and Hottman, Redding and Weinstein (2016)

the fact that customers learn about new products by observing the purchases of others. By charging low markups on entry, they reduce initial profits, but increase sales, thereby shifting future demand and increasing future profits. Markups are raised as entrants age, and more and more customers become aware of their product.

Alternatively, a large literature in marketing finds that it is costly for manufacturers of consumer goods (such as food) to place their products in stores. Since the vast majority of sales of consumer food are mediated by retail stores, this implies that spending on store placement is key to reaching customers. Additionally, direct-to-customer advertising is common in the consumer food industry. If this advertising has an information component, it may be an important action through which entering firms build customer base. The data set we have developed allows us to investigate which of these actions entering firms take in order to reach customers. We first address the role of markups. Under the assumption that marginal cost of production is the same for a given firm in all markets, by comparing prices for the same firm across markets, we can make inferences about the behavior of relative markups. Based on this exercise, we find no evidence that markups change systematically with firm age in a market. This is true whether we use retail prices or wholesale prices. Meanwhile quantities grow systematically with firm age in a market.

Next, we examine marketing and advertising actions. We do not have direct evidence on expenditures related to store placement, but we observe store placement outcomes. Again, we use the strategy of differencing across markets within a firm to control for firm-level factors. Within-market growth is due to both expansion in number of stores, and growth in sales per store. While sales per store accounts for most of sales growth in the initial year in a market, growth in subsequent years is due mainly to expansion to more stores. We also estimate impulse-responses of quantities and prices to number of stores in the relevant market. Quantities sold increase one-for-one with number of stores, while prices are insensitive to the number of stores within a market that the firm sells to.

Finally, we investigate advertising actions. This leverages a unique feature of our data set: the fact that we have performed a global match of advertising at the brand-market level from the Nielsen Ad Intel data to the Nielsen Retail Scanner and Household Panel data on consumer food. While these data have been matched before for narrow product categories, or for large firms, we are the first to use these data to examine the role of advertising in entry for a wide selection of products. Using these data we document a positive association between survival and advertising behavior; new brands which ultimately survive are more likely to advertise in their entry year than those which exit within the next 4 years. For some media types, we observe advertising at the level of the geographic market. Within these media types, we focus on local TV advertising, for which we have the best geographical coverage. We use these data to estimate impulse-responses of quantities and prices to local TV advertising within a market. Quantities respond positively to advertising. This response is larger for entrants than for incumbent brands and comes mostly from increases in the number of customers rather than increases in sales per customer. Meanwhile prices do not co-move with local TV advertising. These facts on the relationship between advertising and sales are not evidence of causality in a particular direction. Nevertheless, it is noteworthy that we do find a positive association between advertising and sales, but no relationship between advertising and prices. This is consistent with the hypothesis that entrants use advertising to reach more customers, but not to affect the price elasticity of demand.

In the last part of the paper we build a structural model of market entry and customer acquisition. The model has two goals: i) quantify the contribution of intrinsic heterogeneity to firm size distribution and ii) quantify the contribution of the endogenous component of demand to firm growth. The model features monopolistic competition. Firms are heterogeneous in appeal, which is common across markets and face firm-market idiosyncratic demand. Entry at the market level is determined by stochastic sunk and fixed costs. In the model, firms accumulate costumers within markets through expenditures on marketing and advertising. We calibrate our model to replicate the within-market dynamics of quantities and prices after entry, as well as the geographic expansion of entrants. Our calibrated model replicates the shape of the impulse-response of sales to advertising we estimate in the data. Our estimates imply that the process of acquiring customer base magnifies cross-firm differences. We find that the variance in firm size on entry is six times the variance of intrinsic heterogeneity.

Our work is related to several literatures. Our finding that customer base plays a role in the slow growth of market share is consistent with Foster, Haltiwanger and Syverson (2008), who use U.S. Census of Manufactures data on a restricted set of industries for which it is possible to measure physical productivity. In contrast to Foster et al, we document this for a range of differentiated products, exploiting the fact that we observe staggered entry across multiple markets to control for common factors at the firm level. We also present evidence on the role of marketing and advertising in building market share. Our findings are similar to those of Fitzgerald, Haller and Yedid-Levi (2016) who show that demand plays a role in the slow growth of market share using customs data on the full range of merchandise exports for Ireland. Like us, they find no role for dynamic pricing in building market share, though they do not have data on marketing and advertising. Our work is also related to Hottman, Redding and Weinstein (2016), who use Nielsen Consumer Panel data to argue that product appeal plays a key role in explaining cross-sectional differences in firm size. We show that "product appeal" (i.e. the behavior of market share conditional on cost and markups) has a striking dynamic pattern, and provide evidence that it is endogenous to firm actions, namely marketing and advertising.

A large literature on "customer markets" posits that firms initially offer low markups to attract new customers, but later exploit customer lock-in due to e.g. search frictions to increase the markups they charge.[2] Some macroeconomists have argued that this behavior may contribute to countercyclical markups and hence to a countercyclical labor wedge. However, an alternative literature in marketing and international trade posits that firms instead use non-price actions such as marketing and advertising to build market share.[3] In the context of this literature, the procyclical behavior of advertising is a challenge to markup-based explanations of the labor wedge.[4] Our findings are also related to those of Argente, Lee and Moreira (2019). They use Nielsen Retail Scanner data to argue that the product cycle in consumer goods is very short, and they emphasize the importance of adding products for firm growth. In contrast to Argente, Lee and Moreira (2019), we focus on the *brand* rather than the UPC as the unit of analysis, and use the sequential roll-out of brands across markets to identify the dynamics of quantities and markups.

The rest of the paper is organized as follows. Section 2 describes the conceptual framework that guides our empirical analysis. In Section 3, we describe the data and provide relevant summary statistics. In Section 4 we document the asynchronous entry of the same brand across geographic markets and the relevance of this margin for growth. Section 5 documents the dynamics of prices and quantities within markets. Section 6 shows that marketing and advertising are used to acquire customers. In Section 7, we develop the model and calibration. Section 8 concludes.

2 Conceptual framework

We use the following conceptual framework to guide our empirical analysis. Let i index firms, and let m index markets. A firm's demand in a market depends on the price customers pay for the firm's good (P_t^{im}). It depends on customer base (D_t^{im}), which is a state variable of the firm's problem which depends actions taken by the firm as we presently describe. It also depends on a vector of exogenous variables ($\boldsymbol{\varepsilon}_t^{im}$). These exogenous variables may include market size, competitors' prices, and tastes for the the firm's good. These tastes may have a

[2]See, e.g. Phelps and Winter (1970), Bils (1989), Klemperer (1995), and more recently, Ravn, Schmitt-Grohé and Uribe (2006), Nakamura and Steinsson (2011), Gourio and Rudanko (2014), Gilchrist, Schoenle, Sim and Zakrajšek (2017) and Paciello, Pozzi and Trachter (2019).

[3]See, e.g. Arkolakis (2010), Eaton, Kortum and Kramarz (2011), Drozd and Nosal (2012), and Eaton, Eslava, Krizan, Kugler and Tybout (2014).

[4]See Hall (2014).

component that is common to all markets, and a component that is idiosyncratic to market m:

$$Q_t^{im} = d\left(P_t^{im}, D_t^{im}, \boldsymbol{\varepsilon}_t^{im}\right)$$

Note that markets are defined such that price discrimination across markets is possible. The assumption that market size and competitors' prices are exogenous to the firm is valid under monopolistic competition.

Customer base is something that firms may accumulate, subject to depreciation. Firms can invest in customer base by undertaking expenditures on marketing and advertising (A_t^{im}) specific to market m. Customer base may also depend on lagged sales, either because individual customers experience adjustment costs in switching between firms, or because new customers learn about the existence of the firm from observing the purchases of others. For incumbents, this relationship is given by:

$$D_t^{im} = i\left(D_{t-1}^{im}, A_{t-1}^{im}, Q_{t-1}^{im}\right)$$

while entrants may be endowed with an exogenous initial customer base $\underline{D}$. If future customer base depends on current expenditures on marketing and advertising, the decision of how much to spend is a dynamic one. If future customer base depends on current sales, this makes the firm's pricing decision dynamic. The importance of this customer base in demand, and the way in which firms accumulate it, are the key objects of interest in our paper.

Finally, the price customers pay for the firm's good in market m is the product of the manufacturer's markup (μ_t^{im}) over marginal cost of production (C_t^i), an iceberg cost of moving the good from the place of manufacture to market m (τ^{im}) and the distribution margin in market m (d_t^k):

$$P_t^{im} = d_t^k \tau^{im} \mu_t^{im} C_t^i$$

The distribution margin is assumed to be the same for all firms in market m, while the iceberg transportation cost is assumed fixed for a firm and a market. For a given firm, the markup may differ across markets (price discrimination) and over time, but the marginal cost of production is the same for all markets served by the firm. This latter assumption is crucial. To the extent that firm growth is constrained on the supply side, it will show up in marginal cost, and this affects all markets served by the firm.

We use this simple framework to guide the construction of a series of moments of the data, and to interpret their behavior. In the light of how these moments behave, we revisit this framework and make more specific assumptions about functional forms and distributions prior to structural estimation.

3 Data description

3.1 Retail sales data

Our primary source is the scanner data set from Nielsen Retail Measurement Services (RMS), provided by the Kilts-Nielsen Data Center at the University of Chicago Booth School of Business. This data set is collected from point-of-sale systems in grocery, drug, and general-merchandise stores. Each individual store reports weekly sales and the quantities of every barcode that had any sales volume during that week. We link firms and products with information obtained from GS1 US, which is the single official source of barcodes. Because the GS1 US data contains all of the company prefixes generated in the US, we combine these prefixes with the barcodes from the RMS. By linking firms to products, we are able to characterize the portfolio of every firm.

We select data covering the food sector during the period from 2006 to 2017. We focus on food products because the market for consumer food is more likely to be geographically segmented than the market for non-food consumer goods and because the coverage of the data is broader and more likely to be representative. The RMS covers about half of all U.S. food sales, and nearly the universe of firms and products in the sector. We determine the type of product by using the Nielsen product hierarchical structure. Barcodes are classified into very disaggregated product categories, modules, which are then aggregated into product groups, which are further aggregated product departments.[5] Our data comprises food departments – dry grocery, dairy, deli, packaged meat, frozen foods, and fresh produce – covering about 600 product modules. We characterize the type of products using product modules because they are sufficiently detailed, similar to NAICS 10-digit, an within modules the barcodes are measured using a common unit of quantity, which allow us to calculate unit values.

Throughout the paper, we refer to the combination of firm-brand-module as brands. The finest level of disaggregation of the RMS dataset is the barcode. We aggregate barcodes into

[5]For example, the barcode of a "Chobani Greek Yogurt Drink Peach 7 fl oz" is in the product module "yogurt-refrigerated-shakes & drinks", in the group "yogurt", and the department "dairy".

firm-brand-module. This strikes a balance: it allows us to aggregate quantities consistently, while ensuring that we do not have to deal with entry and exit of barcodes that may be due to minor product or packaging modifications. In addition, advertising takes place at the brand (rather than barcodes) level, and it is likely firms' internal organization aligns closely with their portfolio of brands Bronnenberg, Dhar and Dubé (2011). Brands are defined by RMS and are fairly detailed. There are unique 63,000 distinct brands, and roughly one quarter of brands have sales in multiple product modules. Given our focus on how firms building market share, this makes the firm-brand-module a suited baseline brand definition.

We combine all sales, quantities, and prices at the brand-market-year level (i.e. product firm-brand-module-market-year). Our baseline definition of market are the Nielsen Designated Market Area (DMA). While sales can be tracked at the store level, advertising is only available at the DMA level.[6] Besides allowing us to match the sales with the advertising data, DMAs are a convenient definition of market since they are large enough to be segmented from consumers' perspective and they align well with MSAs across the country. For the purposes of our empirical analysis, it is important to use a definition of market that is geographically distinct in the sense that customers are likely to purchase from firms participating in their local DMA, that they have access to all products sold in their local market, and that observing positive sales in the RMS database is a good measure of participation in a local DMA. There are 210 Nielsen-defined DMAs (approximately 14 counties per DMA), with coverage in nearly every major U.S. metropolitan area.

We aggregate from weekly data to the annual level, to avoid spurious entry and exit for seasonal products. We define sales of a brand as the total sales across all stores and weeks in the year and market. Likewise, we define quantity as the total quantity sold across all stores and weeks in the year-market, and price is the ratio of sales to quantity, which is equivalent to the quantity weighted average price. Appendix A provides additional details about our retail baseline dataset. Throughout the paper, we present evidence also using quarterly level aggregation and using other related data sets such as the IRI-Symphony data and the Nielsen Homescan Panel, also described in the appendix.

In addition to imperfect market penetration in the cross-section, our empirical analysis relies on the fact that the pattern of market participation is not static: we observe frequent episodes of brand entry at the market level. We say that a brand enters a market in year t if it has zero sales in that market in year $t-1$, and positive sales in year t (in a particular product module). Entry is left censored in 2006, and exit is right censored in 2017. Note

[6]DMAs are used by FCC as definition of markets. U.S. counties are uniquely assigned to a DMA based on historical viewing patterns. DMAs are usually centered around the largest metropolitan area in the region. Only seven counties are assigned to multiple DMAs.

that brands can (and do) enter a market multiple times during the sample. Table 1 reports average entry rates by year. It is also useful to define two other variables: market *age*, and completed spell *survival*. We define market *age* as the cumulative number of periods of continuous market participation at the brand-market level. Completed spell *survival* is the maximum market age achieved in a brand-market level sales spell, i.e. the market age on exit for that spell. In our implementation, age and survival are top-coded at 5 years. This allows us to assign a survival to (some) sales spells where entry is observed, but exit is censored by the end of the sample. Table A7 in Appendix illustrates these definitions for a hypothetical brand in a series of markets.

3.2 Advertising data

Our advertising data comes from comes from the Ad Intel database (ADI) provided by the Kilts-Nielsen Data Center at the University of Chicago Booth School of Business. The database provides occurrence-level advertising information such as time, duration, format, and estimated spending paid for each advertisement. For each occurrence there is very detailed information on the advertising brand, firm, and product type, using ADI own classification system. The data is available for ads featured on television, newspapers, coupons, digital, among other. A few of these media types are reported at the DMA level.

In our baseline analysis we use data for advertising on local television covering the period 2010-2016. Local television is commonly used advertising media used by consumer goods products firms and ADI provides unique data covering all DMAs.[7] We use information on the placement of an ad on a given channel, in a DMA, at a given day and time. Occurrences for each of these different media types can be matched with viewership data, which then yields an estimate of the number of impressions, or eyeballs, that viewed each ad (as in Shapiro, Hitsch and Tuchman (2020)). We summarize the information with a dummy variable capturing whether there was any ad occurrence, the number of ad occurrences, and the total impressions. We also use as robustness the gross rating points (GRPs), a frequently used measure of advertising exposure or intensity in the industry, calculated from the occurrence and impressions data.

Table 1: Summary Statistics

	Firms		Brands		Brands × market	
	All	Entrants	All	Entrants	All	Entrants
Total # unique	21,265	9,688	116,107	61,694	4,478,616	2,688,641
Survival (%)						
1 year	-	0.08	-	0.09	-	0.12
2 years	-	0.13	-	0.15	-	0.21
3 years	-	0.11	-	0.12	-	0.14
4 years	-	0.08	-	0.09	-	0.09
+5 years	-	0.59	-	0.53	-	0.42
Markets (#)						
mean	36	23	38	33	-	-
25th percentile	2	2	2	2	-	-
median	8	5	9	7	-	-
75th percentile	38	20	42	32	-	-
Sales ($1,000)						
mean	15,637	5,877	16,100	6,291	30,821	9,445
25th percentile	930	386	581	328	453	235
median	3,05	1,246	1,875	1,061	1,650	765
75th percentile	11,134	4,054	7,360	3,697	7,899	3,259

Notes: The table presents the summary statistics for the observations included in the baseline pooled sample for the period 2006-2017. For each of these categories, we report the total number of observations, statistics on survival, sales and market expansion. The statistics for sales are computed by determining the average annual sales (in thousands of dollars), deflated by the Consumer Price Index for All Urban Consumers.The table presents the average and distribution statistics of this variable.

3.3 Matching retail sales and advertising data

The main challenge of our data work is to create a dataset that includes both adverting and sales information at very detailed level. Our source data (RMS and ADI) use their own brand and product designations, and a simple fuzzy match of the two sources produces unsatisfactory results. We develop a matching algorithm that leverages the richness of the retail and adverting information using methods from the natural language processing literature to create systematic links between ADI and RMS observations. Appendix A.2 shows how we combine information on the product type descriptions, name of the firm, and name of the brand, to derive a criteria for a many-to-many positive match, and how we ensure that our matching algorithm produces reliable variation. While there are other papers that use retail and advertising data, there is not a dataset that does it across distinct types of products and for all firms/brands. For example, Shapiro, Hitsch and Tuchman (2020) were

[7] In Appendix A.5, we compare the importance of the different media types of the data and show that show that Local television is the dominant media type used firms and across all sectors. We also provide additional robustness evidence using coupon advertising covering about one third of DMAs.

able to match 288 of the top 500 brands in the RMS data. We developed a procedure that is capable of evaluate the match across all brands, small or large, entrants or incumbent. This is key to us since we are particularly interested in entrants.

With our matching algorithm at hand, we have a link between RMS and ADI identifiers and we put together a large scale dataset that has information on both sales and advertising intensity at the product-market-year level, where product is the combination of firm-brand-module as in the RMS.

4 Geographical expansion

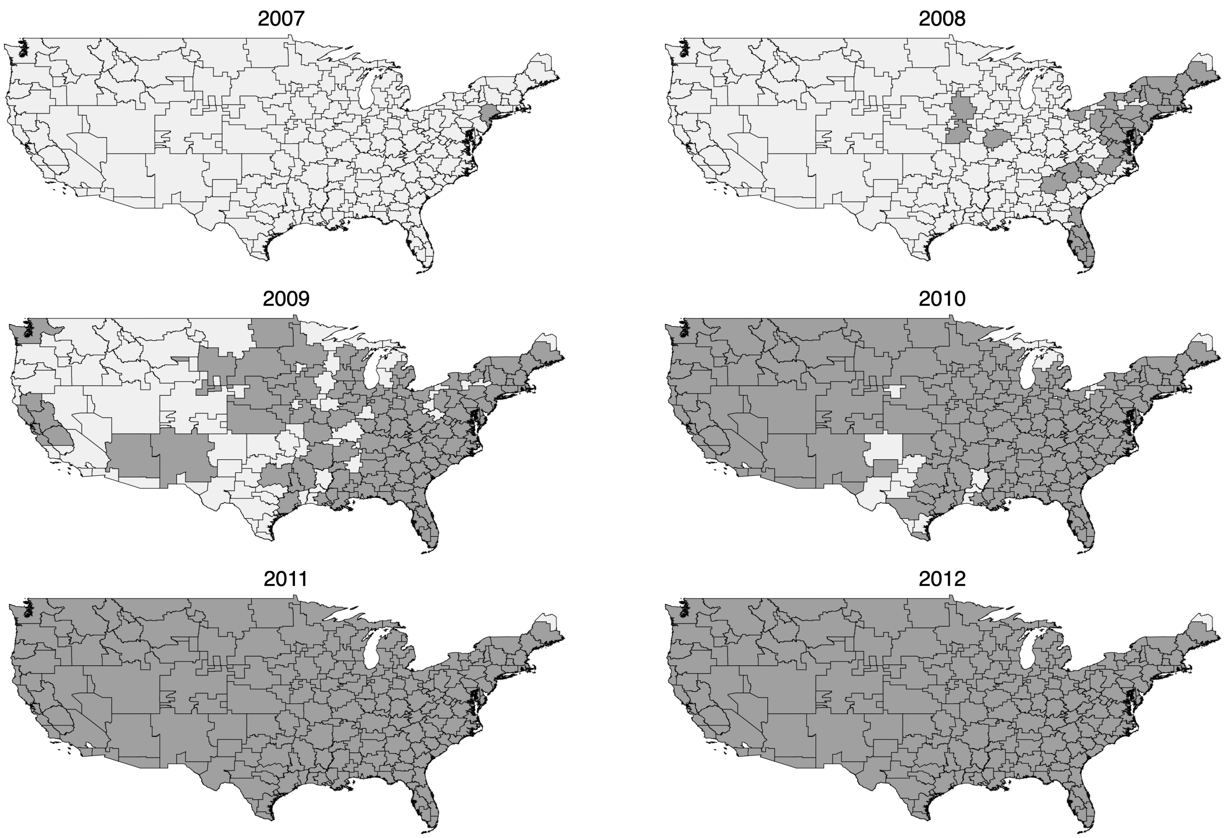
A key part of our strategy for investigating the demand side of firm growth relies on documenting how quantities, prices, etc. evolve with time since entry into a particular market, while differencing across markets to remove the first-order effect of dynamics in marginal cost. But if entrants enter all markets simultaneously, cross-market differencing will also absorb the effect of time since entry. So our strategy relies on entry at the market level being staggered. This is in fact an important feature of the data: entrants do not enter all geographic markets in the US simultaneously.

Figure 1 illustrates this pattern for one very successful firm, which starts selling in the U.S. in 2007. It sells in just one market in its entry year. By 2013 (not illustrated) it sells in all markets in the continental United States. The expansion path of this firm is extreme. Nevertheless, it illustrates two key facts about geographical expansion. (1) In their year of entry, most firms sell to few markets. (2) Conditional on survival, the number of markets a firm sells to tends to grow as the firm ages.

We now document these patterns more systematically. Figure 2 shows the distribution of number of markets in the year of entry for all entering firms in our data. The median number of markets on entry is 2, while the mean is 9.5 (out of 206 possible markets). The figure also shows the distribution of number of markets for all firms, both entrants and incumbents. The mass of this distribution is shifted to the right relative to that of entrants. However this figure cannot show whether the difference is due to selection (entrants which start in few markets are more likely to exit) or to growth (entrants gradually enter more markets). We now show that this is due to a combination of selection and growth, by summarizing the relationship between survival, age, and number of markets.

Let i index firms and let t index time. Let age_t^i be a vector of indicators for firm age (topcoded at 5). Let $survival^i$ be a vector of indicators for the total number of years the firm survives (also topcoded at 5). Let $cens^i$ be an indicator for right-censored survival (i.e. the firm has age less than 5 in the last year of the sample). We regress the log of the number

Figure 1: Example Geographic Expansion - Large Yogurt Brand



Notes: The figure shows the entry across markets of one of the largest brands of yogurt in the US. The darker regions depict the DMAs the brand entered in each of the years.

of markets (M_t^i) on year fixed effects, cohort-of-entry fixed effects, a full set of interactions between age_t^i and $survival^i$ (written $age_t^i \otimes survival^i$), and the indicator for right-censoring. The omitted category in the vector $age_t^i \otimes survival^i$ is the entry year for firms which survive at least 5 years:

$$\ln M_t^i = year_t + cohort^i + \beta' \left(age_t^i \otimes survival^i\right) + cens^i + \varepsilon_t^i \tag{1}$$

Figure 3 plots exponents of the vector of coefficients β against firm age, for firms surviving 1, 2, 3, 4 and 5+ years. Note that the number of markets in the entry year for firms which survive at least 5 years is normalized to 1. From this exercise we learn a number of things about the role of geographic expansion in the process of firm selection and growth. Although as we see from the figure that most firms sell in few markets initially, firms which survive longer sell in more markets in their entry year than firms which exit early. So

Figure 2: Distribution of number of markets: entrants and all firms

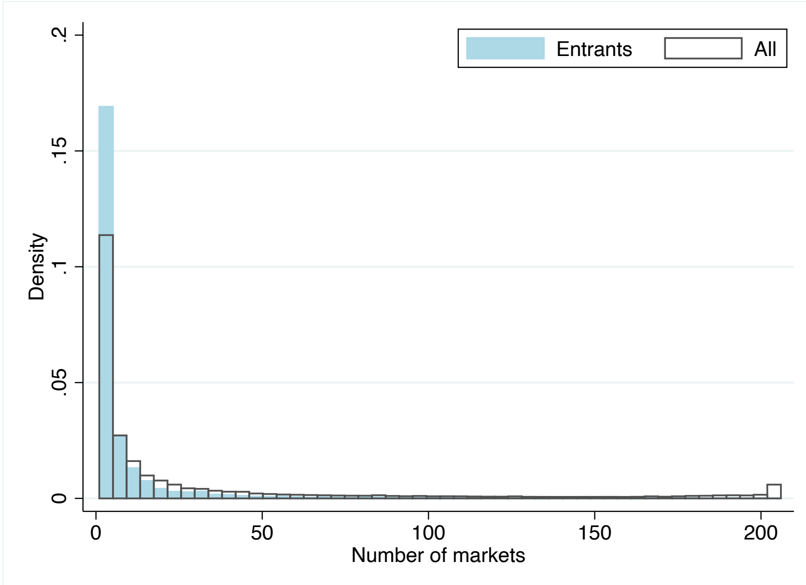
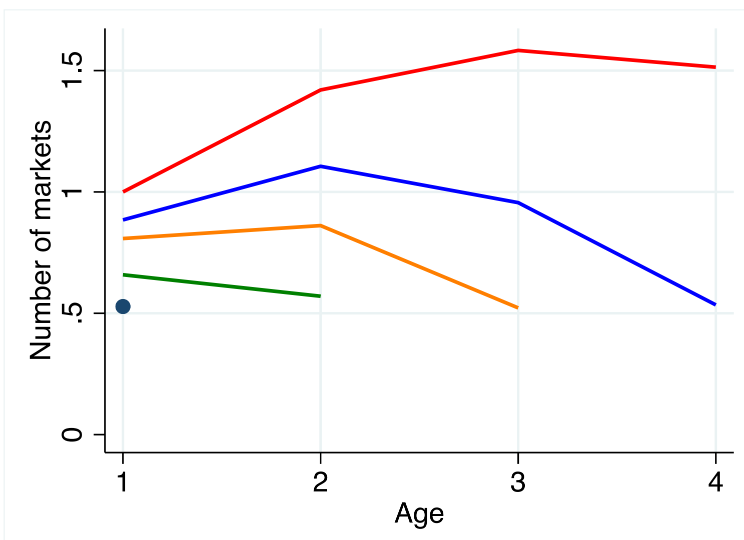


Figure 3: Geographic expansion



Notes: The figure plots the exponents of the vector of coefficients β against firm age, for firms surviving 1, 2, 3, 4 and 5+ years estimated in equation 1. Source: Calculated based on data from The Nielsen Company (US), LLC and marketing databases provided by the Kilts Center for Marketing Data Center at The University of Chicago Booth School of Business.

the difference between the distribution of number of markets for entrants and number of markets for all firms is due partially to selection. Second, for firms surviving 5+ years, the number of markets grows with age on average. Although the average pattern of growth is less spectacular than in the illustrative example, this shows that in addition to the role of selection, gradual geographic expansion of survivors also contributes to the difference between the distribution of number of markets for entrants and number of markets for all firms. This implies that there is cross-sectional variation across markets within a firm in the firm's market-specific age. This is precisely the variation that will allow us to separate demand dynamics from dynamics driven by the supply side.

5 Quantity and price dynamics within markets

There are three elements to our empirical strategy to document the dynamics of prices and quantities within markets. First, we purge the data of variation that is common to all markets served by a product-brand pair, and variation that is common to all brands selling in a particular product-market. In doing so, we remove the first-order effect of cost and quality heterogeneity: better brands may sell more and at lower (or higher) prices in all markets. We also remove the first-order effect of differences in market size and tastes: all brands sell more in bigger and richer markets, and all brands selling a particular product sell more (and potentially at higher prices) in markets with a strong taste for that particular product. This allows us to identify the behavior of market shares and markups,

Second, we focus on product-brand pairs which enter at least one market during our sample period and we examine how quantities, prices, number of stores and number of UPCs evolve with age in a market. This allows us to build an understanding of how quickly entrants reach their steady state market share, and what role (if any) is played by prices in this process.

Third, we allow the evolution of quantities, prices, etc. with age to differ across sales spells with different completed survival. There are two reasons to do this. Both rely on the assumption that the completed survival of a sales spell depends on the permanent components of brand-specific costs, product-market size, and brand-market-specific idiosyncratic demand. If this is the case, controlling for completed spell survival allows us to deal with potential bias in the relationship between quantity, price, etc. and age due to correlation between age and unobserved idiosyncratic demand. In addition, the efforts firms exert to build market share may to depend on the permanent components of costs, market size, and idiosyncratic demand. By allowing the relationship between quantities, prices, etc. and age to differ across brand-market observations based on completed survival, we can test whether this is the case.

More precisely, we simultaneously implement these three elements as follows. Let i index brands as defined above, let m index markets (DMAs), and let t index time (years). Let w_t^{im}

be log quantity or log price. We estimate:

$$w_t^{im} = \delta_t^m + c_t^i + \beta' \left(age_t^{im} \otimes survival_t^{im} \right) + \text{cens}^{im} + \varepsilon_t^{im} \tag{2}$$

where δ_t^m is a market-year-product module fixed effect that controls for demand side factors common to all firms in a market-year-product module. c_t^i is a brand-year fixed effect that controls for marginal cost and quality. age_t^{im} is a vector of indicator variables for brand i's (topcoded) age in market m, and $survival_t^{im}$ is a vector of indicators for the (topcoded) survival of the relevant sales spell. cens^{im} is a vector of indicator variables, for right- and left-censoring of the relevant sales spell.[8] Lastly, $\otimes$ denotes the Kronecker product. Naturally, we do not observe sales spells with market age greater than the completed survival so redundant interactions are dropped. Notice that linear combinations of the coefficients in β allow us to map out the average evolution of log quantities and log prices with age, for spells of different survival.

In Table 1 we report the distribution of sample value by sales spell survival and market age. Clearly, the bulk of the data is accounted for by observations where either age or survival, or both, are censored. Nevertheless, 15% of value is accounted for by observations where entry takes place in-sample. Similarly, the table reports the distribution of observations by sales spell survival and market age. In these terms, observations where entry takes place in-sample accounts for nearly half of the observations of the sample. Table B1 reports the OLS estimates of equation 2 with log revenue, log quantity, log price as dependent variables. The table report intercepts for spells of different survival, as well as the evolution of the dependent variable throughout a spell relative to its value on entry.

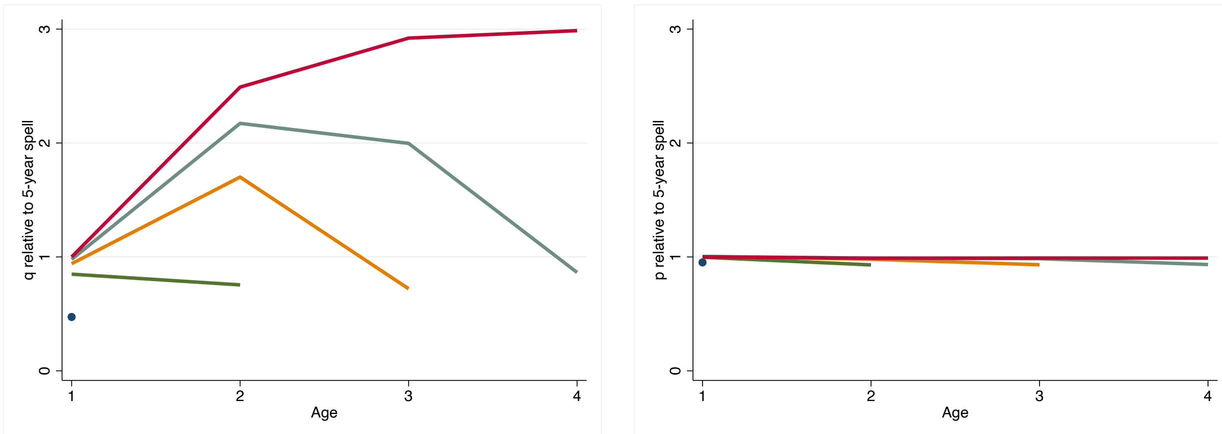
Panels (a) and (b) of Figure 4 illustrate the trajectories of quantity and prices. These trajectories are calculated by exponentiating the relevant sums of coefficients from Table 1.

There are three points to note about these results. First, conditional on costs, higher quantities on entry forecast longer survival, while quantities grow by more a factor of 3 between the entry year and year 4 in the longest sales spells. Meanwhile, there are hump-shaped dynamics of quantities in sales spells where exit is observed. As already noted, since we control throughout for product-market-year effects, these are the dynamics of market share. Since these dynamics of market share are conditional on product-brand-year effects, to the extent that marginal costs are similar across markets, they cannot be driven by costs. Instead, they must be due to movements *along* the demand curve through changing markups, or *shifts* in the demand curve faced by individual product-brand pairs in individual markets.

Second, since we condition on costs, we can infer the behavior of markups from the the

[8]Spells that are both left- and right-censored are classified as left-censored.

Figure 4: Quantity and Price Dynamics Within Markets



(a) Quantity (b) Price

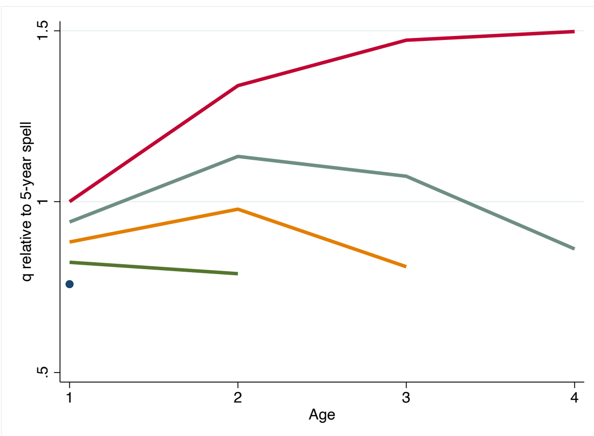
Notes: Based on exponentiating appropriate sums of coefficients from Column 2 of Table B1. Source: Calculated based on data from The Nielsen Company (US), LLC and marketing databases provided by the Kilts Center for Marketing Data Center at The University of Chicago Booth School of Business.

price equation. Higher markups on entry forecast survival longer than 1 year, though beyond that they have little predictive power. Note that this is based on within-brand cross-market variation after controlling for product-market effects. As such, this fact cannot be attributed to all firms having higher prices in some markets. As regards dynamics of markups, to the extent that prices vary with age, they *fall* with age. This is particularly pronounced in the year prior to exit, when prices are on average 9% below their level on entry. For the longest sales spells, prices are 1% lower in subsequent years than they are in the year of entry.

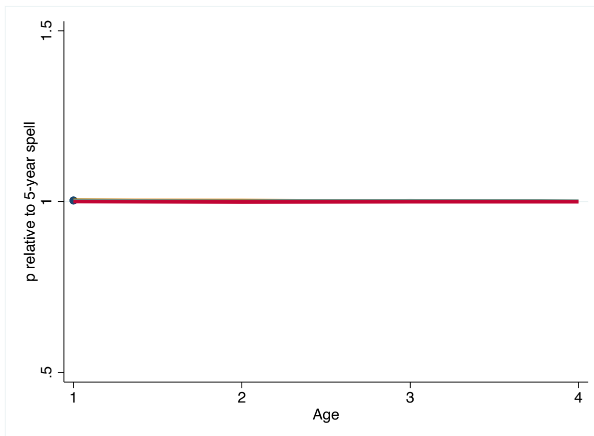
The joint behavior of market share and markups in the longest sales spells is contrary to the customer markets hypothesis, which posits that firms charge low markups to attract customers, and subsequently raise markups once customers are locked in. Upon entry into a market, all customers are new. Presumably the share of new customers falls as a brand ages in a market. But we do not find markups rising with age; on the contrary, they appear to fall with age. This is a feature of the longest sales spells, where presumably brands are likely to endure, and pricing is chosen to trade off current and future demand. It is also a feature of exiting spells, where rather than taking advantage of the fact that exit is imminent by raising prices to extract the maximum from remaining customers, prices actually fall immediately prior to exit. Moreover, given the modest reduction in markups with age in successful sales spells, it is unlikely that falling markups alone can account for growth in market share - this would require remarkably elastic demand.

Taken together, the behavior of market share and markups in successful spells suggest that firms build market share by *shifting* their demand curve rather than moving *along* their demand curve. This behavior is also confirmed when we use consumer level data from the Nielsen Homescan Panel.[9] We find similar hump-shaped dynamics for quantities as shown in Panel (a) of Figure 5. On the other hand, Panel(b) shows that the slight fall in markups prior to exit is not present in the consumer data suggesting that the fall is due to clearance sales. Panels (c) and (d) show that demand increases withing brands come mostly from new costumers.

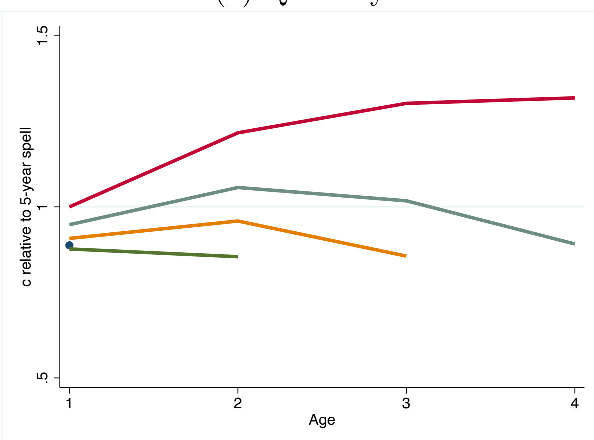
Figure 5: Quantity, Price, and Consumer Dynamics: Consumer Data



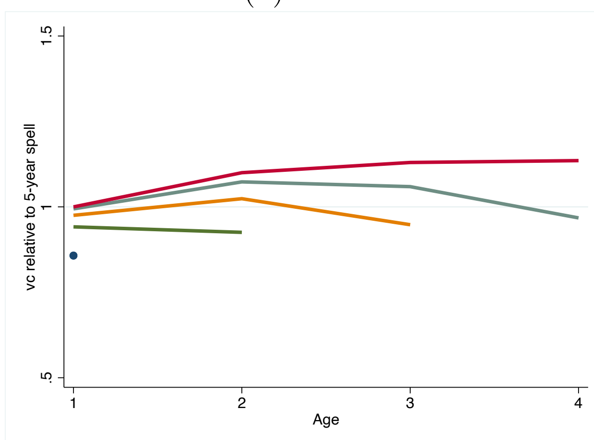
(a) Quantity



(b) Price



(c) Number of consumers



(d) Sales per consumer

Notes: Source: Calculated based on data from The Nielsen Company (US), LLC and marketing databases provided by the Kilts Center for Marketing Data Center at The University of Chicago Booth School of Business.

Robustness checks – Our findings are not sensitive to different time aggregation. Appendix B.1 shows that the dynamics of quantities and prices are similar when we use quarterly data.

[9]We provide a detailed description of these data in Appendix A.1.2.

They are also very similar when we use different brand aggregations. Figure B1 shows our findings when we aggregate across brands within firms and Figure B2 shows our results when we use a broader definitions of brands combining brands with similar names (e.g. Chobani Champions and Chobani Flip combined into Chobani). Aggregating into broad brands rules out the concern that our findings could be driven by the fact that Nielsen's definition of brands could change over time. Figure B3 shows similar findings when we use only the set of brands firms have at entry. Our results are also similar when we use a balanced panel of stores or when define markets as chains, when we defined them as chain × DMA instead of DMA, when the market is defined at the national level, or when we control add cohort controls. These results can be found in Figures B5- B9. Lastly, Figure shows that we find similar results when we use all categories available in the RMS, including non-food categories.

Other retail data – We also use the IRI Symphony data to provide further robustness to our results. This is useful because the data cover a different time period (i.e. 2001-2011) and also the data provides a sales flag which indicates when a product is on sale at a certain store. First, we find that the patterns of quantities and markups are very similar to those found in the RMS (see Figure B11). Second, we find evidence that the slight decline in markups prior to exit is due to clearance sales. We use quarterly data and the sales flag to document that the probability that a brand is on sale in its final quarter is 6-7% higher than the penultimate quarter. Furthermore, Figures B1 shows that the price of exiting brands is 6-7% lower than in quarter before exit. This is, both sales are more frequent and they are deeper the last quarter a brand is available at a store. Appendix B.2 describes this exercises in more detail.

Wholesale data – Lastly, we also use Nielsen PromoData, which collects information from one confidential grocery wholesaler in 47 markets for the period 2006-2012. This data set contains UPC-level wholesale prices for each date in each market. Panel (a) of Figure B12 shows the dynamics of wholesale prices when we aggregate barcodes to the brand level and estimate equation 2 at the quarterly level. Wholesale price dynamics are remarkably similar to retail price dynamics. If any, there is a modest reduction in markups particularly for successful brands. The data also allow us to separate wholesale prices with and without deals. Panel (b) shows that, once we examine prices inclusive of deals, the price level slightly declines prior to exit consistent with the presence of clearance sales just as in the retail data. Overall, in the wholesale level data we again find no evidence that markups rise with age.

6 Evidence on the role of non-price actions

In the previous section, we show that firms increase their market share upon entry, conditional on surviving, while prices do not exhibit any dynamics. Taking together, the evidence on quantity and price dynamics suggest that firms build their customer base by overcoming customer acquisition frictions, without using price actions. The model of customer market predicts that customer base depends on current sales, which would make firms to optimally have lower markups up on entry, and increase markups as it accumulates customer demand. The lack of evidence consistent with dynamic markups, is suggestive that the customer base of brands grows with non-price actions, such as investment on marketing and advertising. In this section, we provide direct evidence that firms invest on marketing and advertising, and that these investments indeed shift their market share.

Expenditures on marketing and advertising include any costly activity for creating, communicating, delivering, and exchanging offerings that have value for customers, such as promoting products to retailers, promoting products in stores, and advertising. We start by using evidence of investment in advertising to evaluate if exposure to advertising is associated with expansion of market share. The key advantage of using advertising data is that we can directly measure expenditures and make a direct association between those and sales performance. Advertising expenditures are, however, only one type of costly non-price actions. Therefore, we complement the analysis with evidence on the additional sources of marketing expenditures. In particular, we use information of expansion across stores within market to evaluate if growth in market share is consistent with marketing expenditures to access additional customers by being placed in additional stores.

6.1 Advertising expenditures

To evaluate the role of advertising expenditures in expanding customer base of entrants we start by assessing the prevalence of advertising among entering brands, specifically those that end up being long lasting, followed by evidence on the dynamics of advertising expenditures. Lastly, we study the joint distribution of sales and advertising by estimating impulse-responses of quantity and prices to advertising expenditures.

Advertising by entrants – Is advertising used by new brands? We study the prevalence of advertising using regression analysis by seniority of a brand. We use regression analysis to explore differences between short entering brands and incumbent brands:

$$Y_t^{im} = \alpha + \sum_{k=2}^{5} \beta_{E,k} \mathbb{1}[\text{Entrant } k]_t^{im} + \beta_I \mathbb{1}[\text{Incumbent}]_t^{im} + \theta_t^m + \varepsilon_t^{im} \quad (3)$$

where Y_t^{im} is either a dummy for having some advertising or a variable capturing intensity of adverting, and $\mathbb{1}[\text{Entrant } k]_t^{im}$ are dummy variables indicating if the brand i in market m is an entrant that lasted $k = 1, ..., 5$ years, and $\mathbb{1}[\text{Incumbent}]_t^{im}$ is a dummy for a brand older than four years old. We account for differences in advertising across types of products, markets, and time periods using fully interacted module-market-year fixed effects (θ_t^m). The coefficients $\beta_{E,2}$, $\beta_{E,3}$, $\beta_{E,4}$ and $\beta_{E,5}$ capture the incremental advertising (relative to entering brands that last only one period) of new brands that survive 2, 3, 4 and more than 5 years, respectively, and β_I captures the incremental advertising of incumbent brands.

Our results are presented in Table B2. Column (1) shows the estimated effects for the linear probability model of using local tv advertising pooling data for all years, where a firm is considered an entrant in the first 4 years old, and incumbent if older than 4 years old. The results show that prevalence of advertising varies substantially by type of brand. Entering brands that last 5 or more years advertise about 0.035 p.p. more than brands that do not perform well and exit within one year. Incumbent brands are likely to advertise about 0.030 p.p. more than entering brands that survive for a long period, indicating that surviving brands are more likely to use advertising. Column (2) presents equivalent results using only data in the entering year, and thus excluding variation coming from incumbent brands. The coefficients in column (2) capture advertising in the entry year, and the coefficients in column (1) capture the average advertising of these brands during their first years of activity (up to four years). For example, a brand that survives for 4 years is expected to advertise more 0.023 p.p. in the entering year than a brand that lasts only one year, and 0.033 p.p. over its life cycle than a brand that lasts only one year.

The results are qualitatively similar when we explore variation within markets and national wide variation. Columns (3) and (4) presents estimates using specification 3 for variation aggregated at the national level where the entry and incumbent information is defined at the national level. Likewise, we explore the evidence from other types of advertising and columns (5) and (6) presents the probability of using any media type covered by ADI at national level. The results indicate that the qualitative variation from local tv captures well the variation from other types of advertising media. Finally, column (7) and (8) shows estimated differences in local TV impressions across types of firms. The results indicate that on the intensive margin we also find that long lasting entering brands and incumbents have more ad impressions than short lasting brands.

Table 2: Advertising by entering brands

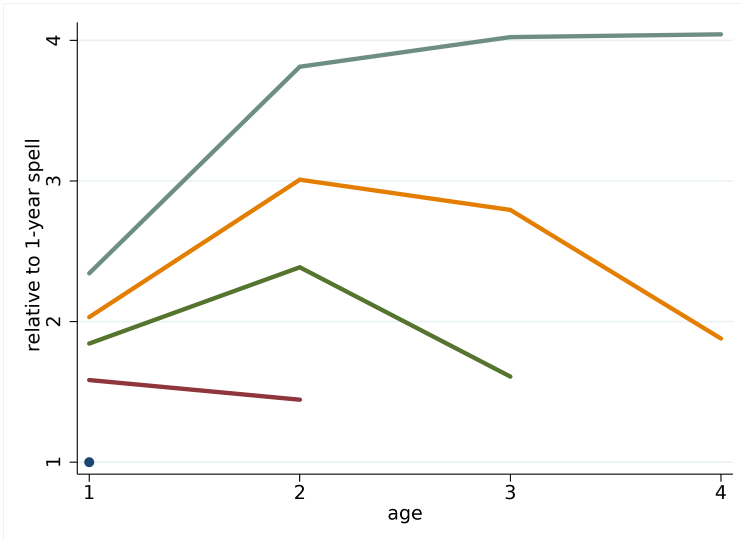
	$\mathbb{1}$[local tv > 0]		$\mathbb{1}$[local tv > 0]		$\mathbb{1}$[any media > 0]		IHS(local tv imp)	
	All	Entry	All	Entry	All	Entry	All	Entry
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Entrant $\beta_{E,2}$	0.003	0.003	0.007**	0.008***	0.002	0.003	0.020	0.032
	(0.006)	(0.004)	(0.003)	(0.003)	(0.006)	(0.006)	(0.087)	(0.060)
Entrant $\beta_{E,3}$	0.011*	0.014***	0.011***	0.009***	-0.007	-0.005	0.141	0.181***
	(0.007)	(0.005)	(0.003)	(0.003)	(0.006)	(0.007)	(0.097)	(0.070)
Entrant $\beta_{E,4}$	0.033***	0.023***	0.019***	0.015***	0.015**	0.022***	0.481***	0.333***
	(0.007)	(0.006)	(0.003)	(0.003)	(0.007)	(0.007)	(0.109)	(0.086)
Entrant $\beta_{E,5}$	0.035***	0.017***	0.016***	0.014***	0.022***	0.023***	0.523***	0.253***
	(0.007)	(0.006)	(0.003)	(0.003)	(0.007)	(0.007)	(0.111)	(0.084)
Incumbent β_I	0.066***		0.039***		0.054***		1.013***	
	(0.006)		(0.002)		(0.006)		(0.096)	
Observations	5,801,851	924,856	200,900	21,796	218,997	25,881	5,801,851	924,856
R-squared	0.179	0.285	0.051	0.147	0.067	0.137	0.178	0.278
Sample	market	market	national	national	national	national	market	market
Module-mkt-t	Y	Y	-	-	-	-	Y	Y
Module-t	-	-	Y	Y	Y	Y	-	-
Uncond. $\bar{Y}_{E,1}$	0.026	0.026	0.004	0.004	0.047	0.047	0.380	0.380

Notes: The table presents estimates of advertising by type of brand using specification in equation 3. We define survival based on first year and last year in the sample. Columns (1), (3), (5) and (7) use data from years 2010-2014. We do not use data for more recent periods because of right censoring problem of brands. Columns (2), (4), (6) and (8) use data from those same years but restricting to data in the first year(s) of activity. Columns (1), (2), (7) and (8) uses variation at the brand-market-year level and includes module-market-year fixed effects, while columns (3)-(6) use data at the brand-year level and includes module-year fixed effects. The dependent variables are a dummy for some local TV advertising in that market ($\mathbb{1}$[local tv > 0]), a dummy for some local TV advertising in any market ($\mathbb{1}$[local tv > 0]), a dummy for some advertising (any media) in any market ($\mathbb{1}$[local tv > 0]), and the inverse hyperbolic sine transformation of local tv impressions in that market (IHS(local tv imp)). The last line of the table reports the unconditional average of the dependent variable for the baseline group of entrants that survive only one year. The standard errors are clustered at the level of the brand and module-module-year (or module-year in columns (3)-(6)). The ***, **, and * represent statistical significance at 1%, 5%, and 10% levels, respectively..

Dynamics of advertising – How does advertising expenditures evolve over the life cycle? We follow specification 2 and estimate how expenditures with advertising evolve within market. Figure 6 plots the estimated trajectories of advertising measured using spending with local TV ads. As with quantities, higher advertising on entry forecast longer survival. For long lasting brands, spending with advertising grows between the entry year and first full year of activity, and remains fairly constant thereafter, and for mid-lasting brands, there are hump-shaped dynamics where exit is observed. This evidence is consistent with brands using advertising to attract additional consumers and expand their customer base within markets.

Sales and advertising – Does advertising affect sales of new brands? The evidence above indicates that entering brands use advertising, presumably because they believe that advertising affects their current or future sales positively. There is an extensive literature

Figure 6: Dynamics of advertising within market



Notes: Estimated for all brands in Ad Intel using specification 2. In Appendix Figure B1 we provide evidence for other measures of advertising and show that the patterns are similar. Source: Calculated based on data from The Nielsen Company (US), LLC and marketing databases provided by the Kilts Center for Marketing Data Center at The University of Chicago Booth School of Business.

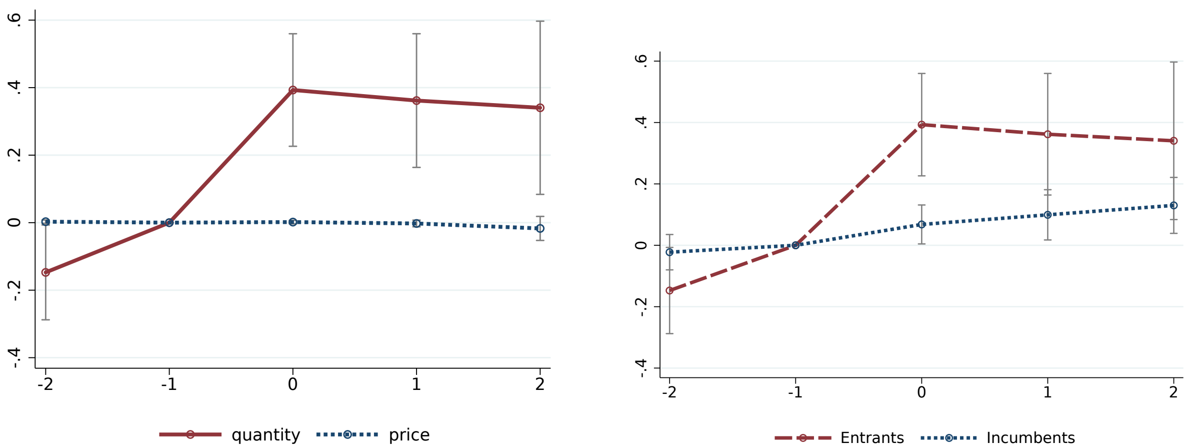
in the marketing literature estimating the causal effect of advertising on sales for specific products, and typically find that advertising affects sales positively. More recently, Shapiro, Hitsch and Tuchman (2020) using data similar to ours covering large established brands across many distinct types consumer good products finds that the elasticity of sales to advertising is small and highly heterogeneous with many non-significant effects.

In our paper we are mostly interested in studying the effects of advertising on the sales of entering brands (as opposed to incumbent established brands) and on decomposing the effect into effects on quantities and prices. To do so, we propose to quantify the association between quantity/prices and advertising using the following econometric specification:

$$\Delta w^{im}_{t+s,t-1} = \beta_s a^{im}_t + \gamma^{im} + \theta^m_t + \gamma_s X^{im}_t + \varepsilon^{im}_t, \qquad s = -2, ...2 \tag{4}$$

where $\Delta w^{im}_{t+s,t-1}$ is change in log quantity (or log prices) of brand i, in market m, in period $t+s$ relative to period $t-1$, a^{im}_t is a dummy variable indicating if the brand i, in market m, in period t, has some local TV adverting, and X^{im}_t is a set of controls. We control for module-market-time fixed effects that accounts for example, for market and product category shifts in demand. We also control for brand-market specific effects, thus filtering out, for instance, the time-invariant effects of brand-specific pricing power on the sales of specific markets. Importantly, this set of fixed effects ensures that the results are not driven by differences in advertising costs or attractiveness across markets or brand specific time-

Figure 7: Relationship between sales and advertising



(a) Quantity/Price for entrants

(b) Quantity for entrants/incumbents

Notes: Estimated using specification 4, including w_{t-1}^{im} as control. Source: Calculated based on data from The Nielsen Company (US), LLC and marketing databases provided by the Kilts Center for Marketing Data Center at The University of Chicago Booth School of Business.

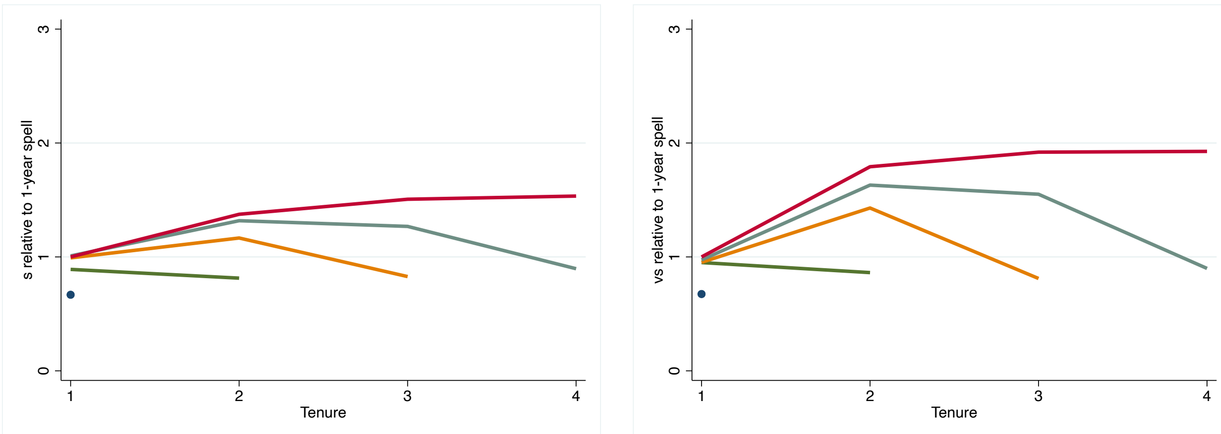
invariant predisposition to use advertising. Note that we evaluate the timing of the effects by running separate linear regressions with different lags and leads.[10]

We start by estimating the relationship between sales and advertising for entering firms. Figure 7 (a) plots the estimated β_s splitting for quantities and prices separately. We find a strong association between quantities and advertising and no association between prices and advertising. Our results indicate that quantities and sales spike at $s = 0$, and we do not find a similar relationship for s below zero. These dynamic specifications are useful for inferring the long-run elasticity of quantities to advertising, in contrast to the instantaneous elasticities. Under exogeneity assumptions in the context of linear local projections (**?**), the implicit long-run elasticity between quantities and advertising is the average of the β_k coefficients from $k = 0$ onward.

Figure 7 (b) plots the estimated β_s for quantities for two distinct samples: entrants and incumbents. Our results show that the association between advertising and sales is stronger for entrants than for incumbents firms, suggesting that the returns on advertising may be weaker for established brands. In Appendix B.3.2, we provide results using alternative specifications and samples.

[10]Furthermore, we consider an alternative specification that allows u to measure the instantaneous impact of advertising on sales. We use those results to evaluate the fit of our model.

Figure 8: Number of stores and sales per store dynamics within markets



(a) Number of Stores

(b) Sales per Store

Notes: Estimated using specification 2. Based on exponentiating appropriate sums of coefficients. Source: Calculated based on data from The Nielsen Company (US), LLC and marketing databases provided by the Kilts Center for Marketing Data Center at The University of Chicago Booth School of Business.

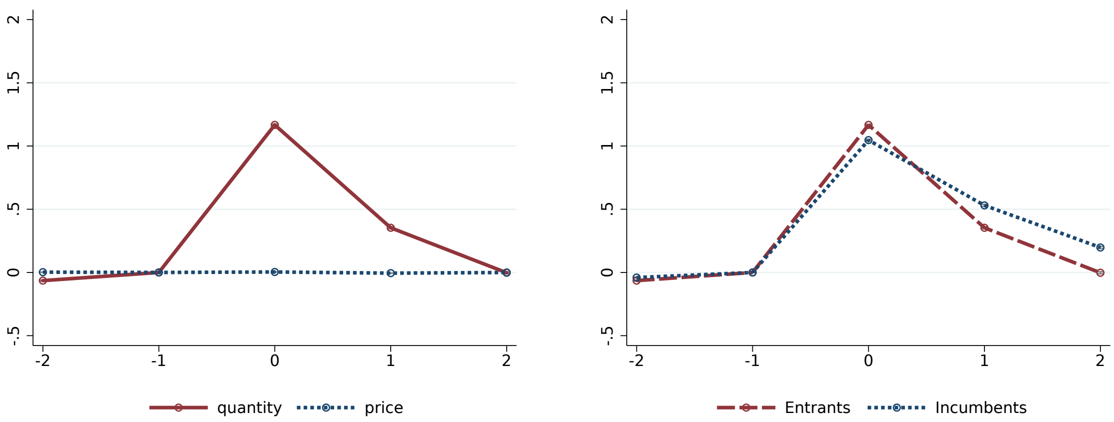
6.2 Other marketing expenditures: product placement in stores

Firms spend resources to enter into different stores through slotting fees – lump-sum payments made by manufacturers to retailers for stocking and shelf space for products – and other costs to establish relationships with retailers. While there is not comprehensive micro-level evidence of these investments, there is extensive evidence supporting the generalized practice of retailers to charge entrants with upfront fixed fees to obtain access to shelf space, defray upfront costs, and support downstream promotional activities (Sullivan, 1997; Klein and Wright, 2007; Marx and Shaffer, 2007; Hamilton and Innes, 2017).[11] We proxy the costs with these activities, by using information on the number of stores that contracted with the brands to provide shelf space in their stores.

We estimate the evolution of number of stores by using variation in brands across markets as specified in equation 2. Figure 8 (a) shows that brands that last longer in the market are sold in more stores and also generate more sales within store. Our interpretation of the evidence that brands reach more stores over time is consistent with brands spending resources with slotting fees that are crucial to reduce frictions to access to consumers and *shift* their demand curve. Furthermore, Figure 8 (b) shows that sales per store are also higher for successful brands, and explain a large amount of the sales growth within market.

[11] These payments are commonly observed not only in the food grocery industry but also in many other industries where firms need intermediaries to reach end consumers. While these fees are often associated with entry into retailers, there is also supporting evidence of fixed "pay-to-stay" fees that indicate that firms have operational costs to maintain access to stores.

Figure 9: Relationship between sales and number of stores



(a) Quantity/Price for entrants (b) Quantity for entrants/incumbents

Notes: Estimated using specification 4, including w_{t-1}^{im} and firm-year fixed effects as controls. Source: Calculated based on data from The Nielsen Company (US), LLC and marketing databases provided by the Kilts Center for Marketing Data Center at The University of Chicago Booth School of Business.

This suggests a role for non-pricing activities shifting the demand of brands within stores.

We further evaluate the role of marketing actions associated with access to stores by quantifying the association between quantity/prices and number of stores offering the brand in a market. We use the econometric specification 4, but for a_t^{im} measured as the log number of stores selling brand i, in market m, in period t. We control for fixed effects that ensure that the results are not driven by differences in marketing costs or attractiveness across markets or brand specific time-invariant predisposition to negotiate with stores or chains.

Figure 9 (a) shows the estimate coefficients β_s by running for entrants separate linear regressions with different lags and leads, for both quantities and prices. We estimate a very large elasticity between number of stores and quantities sold, and no association between prices and number of stores. Figure 9 (b) shows that the estimate coefficients of log quantity for entrants is very similar to those for incumbents, suggesting that the joint distribution of quantities and number of stores is similar across the two types of firms.

7 Model with Endogenous Customer Base

7.1 Setup

We now present a model of endogenous customer base acquisition through marketing and advertising expenditures. The model makes predictions about: (1) the evolution of the number of markets per firm, (2) the evolution of market shares and markups within firms and markets, (3) exit patterns at the firm and firm-market level.

We model the decision problem of a firm, indexed by i, which may participate in m distinct markets, indexed by m. We assume markets are segmented, so the firm is able to price discriminate across markets as well as target its marketing and advertising effort at the market level. We abstract from the product dimension, implicitly assuming that firms' choices are independent across products.

When firm i is born, it is endowed with product appeal χ^i, drawn from a known distribution. A firm's product appeal is *permanent* and is the same across all markets that the firm serves. We assume firm i dies with probability $p(\chi^i)$ each period, where $p(\cdot)$ is weakly decreasing.

Also at birth, firm i draws a vector of *permanent* market-specific idiosyncratic demand shocks, $\{\nu^{im}\}_k$. In addition, in each period, it draws a vector of *transitory* market-specific idiosyncratic demand shocks, $\{\eta_t^{im}\}_{kt}$. These latter shocks are independent and identically distributed over time and across markets.

Let $X_t^{im} \in \{0,1\}$ be an indicator for participation by firm i in market m at time t. Conditional on participation, demand faced by firm i in market m at time t is given by:

$$Q_t^{im} = M^k \left(P_t^{im}\right)^{-\theta} \left(D_t^{im}\right)^{\alpha} \chi^i \exp\left(\nu^{im} + \eta_t^{im}\right)$$

Here, M^k is the component of demand which is the same for all firms serving market m. It combines the impact of aggregate demand in market m with that of competitors' prices and the retail markup over the manufacturer's price. All of these components are assumed to be time-invariant, and identical for all firms, implicitly requiring monopolistic competition.

The firm i chooses its own price P_t^{im} (expressed in terms of an arbitrary numeraire) in each location m and each period t. This is the price it charges to retailers. The price customers face is $d^k \cdot P_t^{im}$, where d^k is the retail margin in market m, assumed constant over time, though potentially differing across markets. This retail margin is subsumed into M^k. We assume the price elasticity of demand is the same for all firms and markets. This will have strong implications for markups.

In addition to its own price, firm i's demand depends on its inherent product appeal χ^i and on the realizations of its permanent and transitory market-specific demand shocks in market m, $\nu^{im} + \eta_t^{im}$.

Finally, firm i's demand also depends on its market-specific customer base, D_t^{im}. Importantly, the firm can take actions to accumulate customer base. We assume that it does so by undertaking market-specific spending on marketing and advertising. We assume that a firm's market-specific customer base accumulates according to:

$$D_t^{im} = \left(1 - X_{t-1}^{im}\right) \underline{D}^k + X_{t-1}^{im} \left(\left(1-\delta\right) D_{t-1}^{im} + A_{t-1}^{im}\right)$$

Firms entering market m start with customer base $\underline{D}^k$. Indexing initial customer base by m allows us to capture the assumption that conditional on all sources of firm-specific heterogeneity, and on price, all entrants to a market start with the same market share. A_{t-1}^{im} is the increment to customer base at date t in market m that is due to marketing and advertising at date $t-1$. The depreciation rate of past customer base conditional on continued participation is δ. Customer base fully depreciates on exit from a market m.

The cost of marketing and advertising is given by $c\left(D_t^{im}, A_t^{im}\right)$, also expressed in terms of the numeraire. We assume this takes the particular functional form:

$$c\left(D_t^{im}, A_t^{im}\right) = \begin{cases} A_t^{im} + \phi\left(\frac{A_t^{im}}{D_t^{im}}\right)^2, & \text{if } A_t^{im} > 0 \\ 0, & \text{otherwise} \end{cases}$$

Since customer base is intangible, it is natural to assume irreversibility (i.e. $A_t^{im} \geq 0$).

At the level of an individual market, firm i faces sunk (S_t^{im}) and fixed (F_t^{im}) costs of participation, assumed to be independent and identically distributed, expressed in terms of the numeraire. Conditional on the firm's inherent appeal, χ^i, demand shocks $\{\nu^{im}\}_k$ and $\{\eta^{im}\}_{k,t}$ along with sunk (S_t^{im}) and fixed (F_t^{im}) costs allow for variation in the timing of entry and exit across markets. This is consistent with what we observe in the data.

In this model, current participation X_t^{im} and advertising efforts A_t^{im} affect both future and current payoffs. However, a firm's choice of price only affects current profits and not future profits. Therefore, the optimal price P_t^{im} is a standard constant markup over marginal cost, C_t^i, expressed in terms of the numeraire:

$$P_t^{im} = \frac{\theta}{\theta - 1} C_t^i$$

For simplicity, we assume that marginal cost is the same for all firms, normalized to 1, thus loading all firm-level heterogeneity into appeal.

The firm's dynamic problem is then as follows. At the beginning of period t, it observes the realization of the death shock. Conditional on survival, for each market, it observes $\left\{M^k, S_t^{im}, F_t^{im}, \nu^{im}, \eta_t^{im}\right\}$. Let $Z_t^{im} = \left\{M^k, \chi^i, S_t^{im}, F_t^{im}, \nu^{im}, \eta_t^{im}\right\}$ be a vector of exogenous states. Current net flow profit from market m conditional on participation is:

$$\begin{aligned}\pi\left(X_{t-1}^{im}, D_t^{im}, Z_t^{im}, A_t^{im}\right) = &\frac{(\theta-1)^{\theta-1}}{\theta^\theta} M^k \chi^i \left(D_t^{im}\right)^\alpha \exp\left(\nu^{im} + \eta_t^{im}\right) \\ &- c\left(D_t^{im}, A_t^{im}\right) - F_t^{im} - \left(1 - X_{t-1}^{im}\right) S_t^{im}\end{aligned}$$

The firm discounts future flows at rate β. The Bellman equation for the firm's market-m problem is:

$$\begin{aligned}V\left(X_{t-1}^{im}, D_t^{im}, Z_t^{im}\right) = \max_{X_t^{im}\in\{0,1\}, A_t^{im}\geq 0} &\left\{X_t^{im} \cdot \pi\left(X_{t-1}^{im}, D_t^{im}, Z_t^{im}, A_t^{im}\right) + \right. \\ &\left. \beta\left(1 - p\left(\chi^i\right)\right) \mathbb{E}\left\{V\left(X_t^{im}, D_{t+1}^{im}, Z_{t+1}^{im}\right) | Z_t^{im}\right\}\right\}\end{aligned}$$

subject to the evolution of the customer base.

7.2 Estimation

We estimate the model using simulated method of moments. The moments we use as targets are as follows: the median number of markets per firm on entry; the evolution of number of markets per firm with age conditional on survival; the exit hazard at the firm level; the evolution of quantities at the firm-market level with age conditional on survival; the exit hazard at the firm-market level. We use our advertising moments as non-targeted moments.

To estimate, we first need to make assumptions about distributions and functional forms. We assume that $\ln \chi^i \sim N\left(0, \sigma_\chi^2\right)$ and that $p\left(\chi^i\right) = \exp\left(-\psi \chi^i\right)$. We assume that $\nu^{im} \sim N\left(0, \sigma_\nu^2\right)$, and independently and identically distributed across firms and markets. We assume that $\eta_t^{im} \sim N\left(0, \sigma_\eta^2\right)$, and independently and identically distributed across firms, markets and over time. We discretize each of these three distributions using the Rouwenhorst method. We assume that the sunk cost of entry takes on three possible values:

$$S_t^{im} = \begin{cases} 0 & \text{with probability } \lambda\kappa \\ S & \text{with probability } \lambda(1-\kappa) \\ \infty & \text{with probability } 1-\lambda \end{cases}$$

Similarly, market-level fixed operating costs take on three possible values

$$F_t^{im} = \begin{cases} 0 & \text{with probability } (1-\omega)\gamma \\ F & \text{with probability } (1-\omega)(1-\gamma) \\ \infty & \text{with probability } \omega \end{cases}$$

Both sunk and fixed costs are independently and identically distributed across firms, markets and time.

The estimation then proceeds as follows. We choose the distribution of M^k to match the distribution of market size in the data. We set $\beta = (1/1.05)^{0.5}$, consistent with a period in the model corresponding to six months in the data. This allows us to model "part-year effects" which are part of the real world data generating process. Note that in our model, the price elasticity of demand θ is not identified by our target moments, and once the parameters $\{S, F, \underline{D}\}$ are appropriately scaled, the mapping between parameters and the moments of interest is independent of θ.

Then, given a vector of parameters $\mu = \left\{\sigma_\chi^2, \psi, \sigma_\nu^2, \sigma_\eta^2, \lambda, \kappa, S, \omega, \gamma, F, \underline{D}, \alpha, \delta, \phi\right\}$, we can solve for the optimal policy functions of the firm using value function iteration. Using these optimal policies, we simulate a cohort of 500 potential entrant firms in 200 markets, and follow them for 10 periods. We aggregate these data to the annual level and use the annual data to construct a vector of simulated moments, $\hat{m}(\mu)$, which correspond to the vector of data moments, $\bar{m}$. Notice, this allows the model to match the "part-year effects" in the data. We then use the simulated and data moments to calculate:

$$\Theta(\mu) = (\hat{m}(\mu) - \bar{m})' \, \Omega \, (\hat{m}(\mu) - \bar{m})$$

where Ω is a diagonal matrix with the inverse of the standard errors of the data moments on the diagonal. We then search over parameter vectors μ to minimize this criterion.

The mapping between moments and parameters is not one-to-one. But, broadly speaking, the firm-level moments identify the distribution of appeal and the process for sunk cost of market participation, the firm-market-level moments identify the $\{\alpha, \delta, \phi, \underline{D}\}$, and all

Table 3: Parameter estimates

σ_χ	ψ	σ_ν	σ_η	λ	κ	S	ω	γ	F	$\underline{D}$	α	δ	ϕ
1.54	36	0.37	1.31	0.10	0.90	2.17*	0.02	0.50	1.01*	0.02§	0.48	0.56	69

Notes: *Expressed as a multiple of flow profit in the entry period for entrants in the largest markets with mean χ, ν and η (drawn from their unconditional distributions). § Expressed as a share of deterministic steady state D for a firm with mean χ, ν and η. This share is the same for all markets.

moments jointly identify the process for fixed costs of market participation.

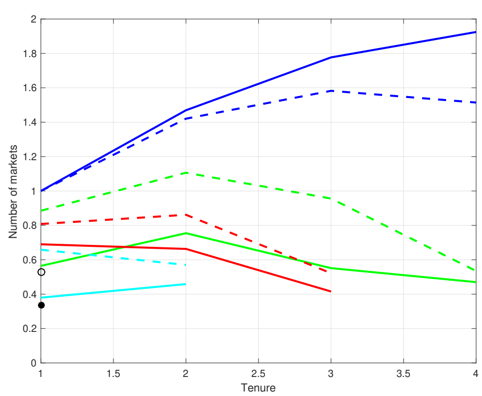
7.2.1 Estimates and fit

Table 3 reports our parameter estimates. Figure 10 shows how the model fits the evolution of the number of markets per firm with age conditional on survival. The model fits the market-level moments well. The fit of the firm-level moments is not as good. The model generates too much dispersion in number of markets per firm on entry, and growth in the average number of markets for surviving firms that is too fast.

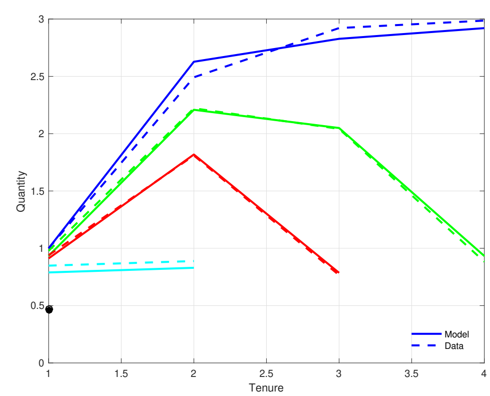
The targeted moments not reported in these figures are firm-level exit hazard and the median number of markets on entry. The exit patterns currently do not match well, and the median median number of markets in the data is 2, while its model counterpart in our baseline calibration is 3.

Figure 11 shows the impulse-responses of quantities and prices to advertising in the

Figure 10: Model vs Data



(a) Evolution of Number of Markets per Firm

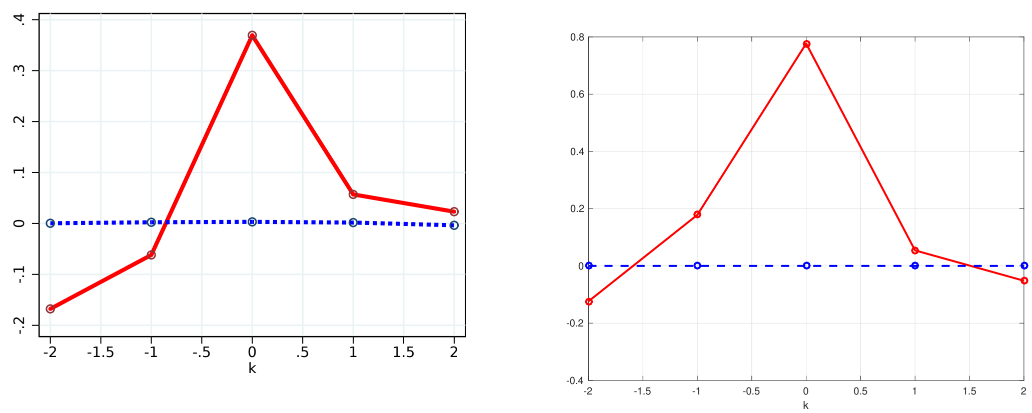


(b) Evolution of Firm-Market Quantity

Notes: The figure shows the simulation results of the model (solid lines) and compares them with their data counterpart (dashed lines).

data (left) and in the model (right), respectively. These are not explicitly targeted in the estimation. The qualitative pictures are very similar, though the response of quantities to advertising in the model is twice as big as the response in the data.

Figure 11: Model vs Data: Advertising



(a) Data (b) Model

Notes: Panel (a) the impulse response function in the data and panel (b) the analogue figure produced in the simulations of the model.

7.2.2 Contribution of exogenous appeal to the variance of revenue

Having estimated the model, we can use it to decompose the variance of revenue (our measure of firm size) across firms into the contribution of exogenous firm-level heterogeneity, χ^i, and the contribution of all other factors. Remember that:

$$R_t^i = \chi^i \sum_{k \in N_t^i} \left(\frac{\theta}{\theta - 1}\right)^{1-\theta} M^k \left(D_t^{im}\right)^{\alpha} \exp\left(\nu^{im} + \eta_t^{im}\right) = \chi^i \tilde{R}_t^i$$

The second term in this expression combines the number of markets firm i participates in, the size of those markets, idiosyncratic demand in those markets, and how much the firm has accumulated in the way of customer base in those markets. This combines the impact of exogenous factors such as the realizations of market-specific idiosyncratic demand, and of endogenous responses to those factors as well as to χ^i.

Table 4 reports the share of the variance of revenue which is accounted for by the variance of χ^i, by the covariance of χ^i and $\tilde{R}_t^i$, and by the variance of $\tilde{R}_t^i$. This is reported for the cohort of entrants in their first year, as well as for the cohort of survivors after four years.

Exogenous appeal accounts for only 1/6 of the variance in firm size in the year of entry. However the covariance of χ^i and $\tilde{R}_t^i$ accounts for over 1/3 of the variance in size. This is due to the fact that firms with greater appeal have a greater incentive to enter markets, and to invest in customer base in those markets.

Table 4: Variance decomposition of revenue in the model

	Share in variance of revenue		
	$Var\left(\chi^i\right)$	$2Cov\left(\chi^i, \tilde{R}_t^i\right)$	$Var\left(\tilde{R}_t^i\right)$
Year 1	0.17	0.37	0.46
Year 4	0.14	0.42	0.44

After four years, the variance of exogenous appeal accounts for only 1/7 of the variance in firm size, while the share accounted for by the covariance of χ^i and $\tilde{R}_t^i$ accounts for 3/7 of the variance. As firms have more opportunities to enter, and more time to invest in customer base, this endogenous contribution to the variance in firm size grows.

8 Discussion and conclusions

The customer markets literature argues that firms use markups as a tool to build market share. More precisely, this literature argues that firms attract new customers by offering them low markups, but then raise markups once customers are locked in. We show that the market for consumer food products does not exhibit these types of dynamics. While there is substantial growth in market share in the initial years of successful sales spells, these episodes are not accompanied by rising markups as the customer markets theory predicts. Instead, if anything, markups fall marginally as market share grows.

The patterns we document point toward a potentially important role for marketing and advertising activity in building market share. First, we show that the extensive margin of stores plays an important role in the growth of sales in a market. Since we know from the marketing literature that firms may incur direct costs in order to place their product in stores, this is suggestive of a role for marketing. Second, we find direct evidence supporting this using data on TV advertising.

To quantify the role of non-pricing activities magnifying differences in idiosyncratic heterogeneity across firms, we develop a dynamic structural model with accumulation of customers within and across markets. A calibrated version of our model indicates that the variance in firm size on entry is six times variance in intrinsic heterogeneity across firms suggesting an important role of non-pricing in the accumulation of customers over time.

References

Argente, David and Chen Yeh, "Product Life Cycle, Learning, and Nominal Shocks," *University of Chicago mimeo*, 2018.

—, Munseob Lee, and Sara Moreira, "The Life Cycle of Products: Evidence and Implications," *Available at SSRN 3163195*, 2019.

—, Salomé Baslandze, Douglas Hanley, and Sara Moreira, "Patents to Products: Product Innovation and Firm Dynamics," 2020.

Arkolakis, Costas, "Market penetration costs and the new consumers margin in international trade," *Journal of political economy*, 2010, *118* (6), 1151–1199.

Beraja, Martin, Erik Hurst, and Juan Ospina, "The aggregate implications of regional business cycles," *Econometrica*, 2019, *87* (6), 1789–1833.

Bils, Mark, "Pricing in a customer market," *The Quarterly Journal of Economics*, 1989, *104* (4), 699–718.

Bronnenberg, Bart J, Michael W Kruger, and Carl F Mela, "Database paper—The IRI marketing data set," *Marketing science*, 2008, *27* (4), 745–748.

—, Sanjay K Dhar, and Jean-Pierre Dubé, "Consumer packaged goods in the United States: National brands, local branding," *Journal of Marketing Research*, 2007, *44* (1), 4–13.

—, —, and Jean-Pierre H Dubé, "Endogenous sunk costs and the geographic differences in the market structures of CPG categories," *Quantitative Marketing and Economics*, 2011, *9* (1), 1–23.

Drozd, Lukasz A and Jaromir B Nosal, "Understanding international prices: Customers as capital," *American Economic Review*, 2012, *102* (1), 364–95.

Eaton, Jonathan, Marcela Eslava, Cornell J Krizan, Maurice Kugler, and James Tybout, "A search and learning model of export dynamics," *Unpublished manuscript*, 2014.

—, Samuel Kortum, and Francis Kramarz, "An anatomy of international trade: Evidence from French firms," *Econometrica*, 2011, *79* (5), 1453–1498.

Fally, Thibault and Benjamin Faber, "Firm Heterogeneity in Consumption Baskets: Evidence from Home and Store Scanner Data," *NBER Working Paper No. 23101*, 2017.

Fitzgerald, Doireann, Stefanie Haller, and Yaniv Yedid-Levi, "How exporters grow," Technical Report, National Bureau of Economic Research 2016.

Foster, Lucia, John Haltiwanger, and Chad Syverson, "Reallocation, firm turnover, and efficiency: Selection on productivity or profitability?," *American Economic Review*, 2008, *98* (1), 394–425.

Gagnon, Etienne, David Lopez-Salido, and Jason Sockin, "The Cyclicality of Sales, Regular and Effective Prices: Comment," *Available at SSRN 2631693*, 2015.

Gilchrist, Simon, Raphael Schoenle, Jae Sim, and Egon Zakrajšek, "Inflation dynamics during the financial crisis," *American Economic Review*, 2017, *107* (3), 785–823.

Gourio, Francois and Leena Rudanko, "Customer capital," *Review of Economic Studies*, 2014, *81* (3), 1102–1136.

Hall, Robert E, "What the cyclical response of advertising reveals about markups and other macroeconomic wedges," *NBER Working Paper*, 2014, *18370.*

Hamilton, Stephen and Robert Innes, "Slotting allowances and retail product variety under oligopoly," *Economics Letters*, 2017, *158*, 34–36.

Hottman, Colin J, Stephen J Redding, and David E Weinstein, "Quantifying the sources of firm heterogeneity," *The Quarterly Journal of Economics*, 2016, *131* (3), 1291–1364.

Jordà, Òscar, "Estimation and inference of impulse responses by local projections," *American economic review*, 2005, *95* (1), 161–182.

Klein, Benjamin and Joshua D Wright, "The economics of slotting contracts," *The Journal of Law and Economics*, 2007, *50* (3), 421–454.

Klemperer, Paul, "Competition when consumers have switching costs: An overview with applications to industrial organization, macroeconomics, and international trade," *The review of economic studies*, 1995, *62* (4), 515–539.

Luptáčik, Mikuláš, "Optimal price and advertising policy under atomistic competition," *Journal of economic dynamics and control*, 1982, *4*, 57–71.

Marx, Leslie M and Greg Shaffer, "Upfront payments and exclusion in downstream markets," *The RAND Journal of Economics*, 2007, *38* (3), 823–843.

Nakamura, Emi and Jón Steinsson, "Price setting in forward-looking customer markets," *Journal of Monetary Economics*, 2011, *58* (3), 220–233.

Paciello, Luigi, Andrea Pozzi, and Nicholas Trachter, "Price dynamics with customer markets," *International Economic Review*, 2019, *60* (1), 413–446.

Phelps, Edmund S and Sidney G Winter, "Optimal price policy under atomistic competition," *Microeconomic foundations of employment and inflation theory*, 1970, pp. 309–337.

Ravn, Morten, Stephanie Schmitt-Grohé, and Martin Uribe, "Deep habits," *The Review of Economic Studies*, 2006, *73* (1), 195–218.

Shapiro, Bradley, Günter J Hitsch, and Anna Tuchman, "Generalizable and robust TV advertising effects," *Available at SSRN 3273476*, 2020.

Sullivan, Mary W, "Slotting allowances and the market for new products," *The Journal of Law and Economics*, 1997, *40* (2), 461–494.

APPENDIX

A Data Construction

A.1 Source datasets

A.1.1 RMS

Our primary data source is the Nielsen Retail Measurement Services (RMS) provided by the Kilts-Nielsen Data Center at the University of Chicago Booth School of Business. The data is generated by point-of-sale systems in retail stores. Each individual store reports weekly sales and the quantities of every barcode that had any sales volume during that week. We use data for the period from 2006 to 2017. The main advantage of this data set is its size and coverage. Overall, the RMS consists of more than 100 billion unique observations at the UPC $\times$ store $\times$ week level that cover approximately \$2 trillion in sales. This volume of sales represents about 53% of all sales in grocery stores, 55% in drug stores, 32% in mass merchandisers, 2% in convenience stores, and 1% in liquor stores. A key distinctive feature of this database is that the collection points include more than 40,000 distinct stores from around 90 retail chains and 2,500 counties. As a result, the data provide good coverage of the universe of products and of the full portfolio of firms in this sector,

We link firms and products with information obtained from GS1 US, which is the single official source of UPCs. Because the GS1 US data contains all of the company prefixes generated in the US, we combine these prefixes with the UPC codes from the RMS. By linking firms to products, we are able to characterize the portfolio of every firm with products in our sample. Furthermore, we can identify the sales, price, and quantity of each product belonging to every firm and compute these variables at the firm $\times$ market level. We mostly focus on measures of firm size such as as number of products, total sales, and number of markets. We also use this data set to identify the entry and exit of firms.

A.1.2 Nielsen Homescan Panel

We use the Nielse Homescan Panel data set from 2004 to 2016. The data set tracks the shopping behavior of 40,000–60,000 households every year covering 49 states and 2,967 counties in the United States. Each panelist uses in-home scanners to record their purchases. A twelve-digit universal product code (UPC) identifies the items the panelists purchase. The data contain around 3.27 million distinct UPCs grouped using the same hierarchical structure as the RMS. For each UPC, the data contain information on the brand and size. If the panelist purchases the good at a store covered by Nielsen, the price is automatically set to the average price of the good at the store during the week when the purchase was made. If not, the panelist directly enters the price. Nielsen reports detailed transaction information for each product purchased (e.g. UPC code, quantity, price, deals, and coupons). We combine this information with the weight and volume of the product to compute unit values. The data also contain information about each purchasing trip the panelist makes, such as the retailer, the location, and the date of the transaction.

A.1.3 IRI Symphony Data

We also use the IRI Symphony data, a data set very similar to the Nielsen RMS that includes prices and quantities for barcodes across the US. The data set also provides a sales flag which indicates when a product is on sale in a certain store. The data contain approximately 2.4 billion transactions from January 2001 to December 2011 which represents roughly 15 percent of household spending in the Consumer Expenditure Survey (CEX). Our sample contains approximately 170,000 products and 3,000 distinct stores across 43 metropolitan areas (MSA). The data covers 31 product categories which include: Beer, Carbonated Beverages, Coffee, Cold Cereal, Deodorant, Diapers, Facial Tissue, Photography Supplies, Frankfurters, Frozen Dinners, Frozen Pizza, Household Cleaners, Cigarettes, Mustard Ketchup, Mayonnaise, Laundry Detergent, Margarine Butter, Milk, Paper Towels, Peanut Butter, Razors, Blades, Salty Snacks, Shampoo, Soup, Spaghetti Sauce, Sugar Substitutes, Toilet Tissue, Toothbrushes, Toothpaste, and Yogurt. The dataset is discussed in more detail in Bronnenberg, Kruger and Mela (2008).

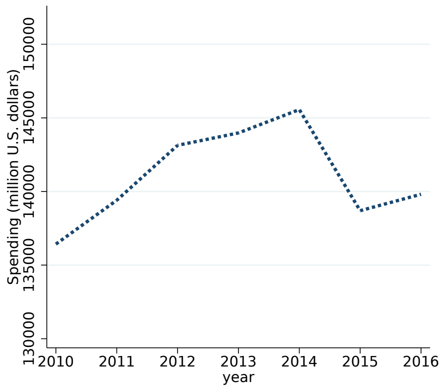
A.1.4 Ad Intel

Advertising data comes from the Nielsen's Ad Intel (ADI) database provided by the Kilts-Nielsen Data Center at the University of Chicago Booth School of Business. The data provides information for advertising starting in 2010, covering nearly 150 billion per year of spending in the U.S., and nearly 400 million observations per year (Figure A1). The database provides occurrence-level advertising information such as time, duration, format, a product type, and spending paid for each advertisement. The data is available for ads featured on television, newspaper and magazines, radio, cinema, coupons, outdoor, digital, among other (Table A1). A few of these media types are reported at the DMA level, and local television covers all 210 DMAs.

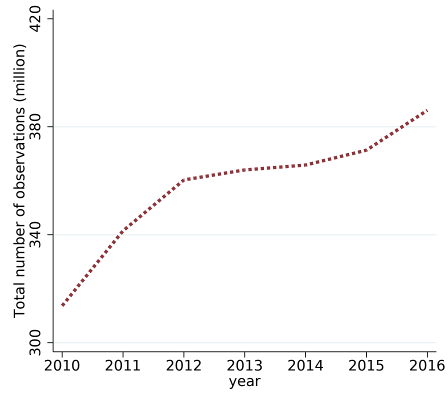
In our analysis we cover the period 2010-2016. As a baseline we use variation from local TV advertising (spot, network clearance and syndicated clearance), and coupon for

Figure A1: Overview Ad Intel

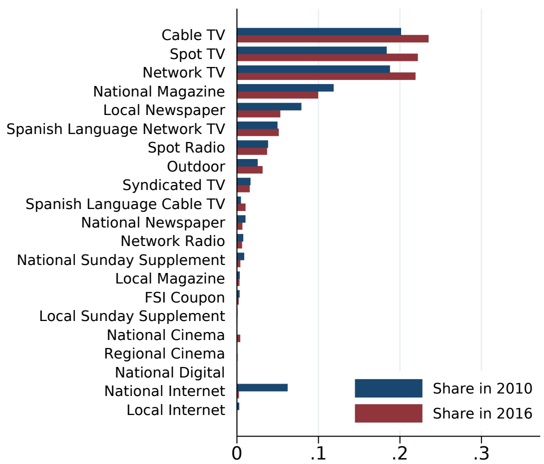
(a) Total spending



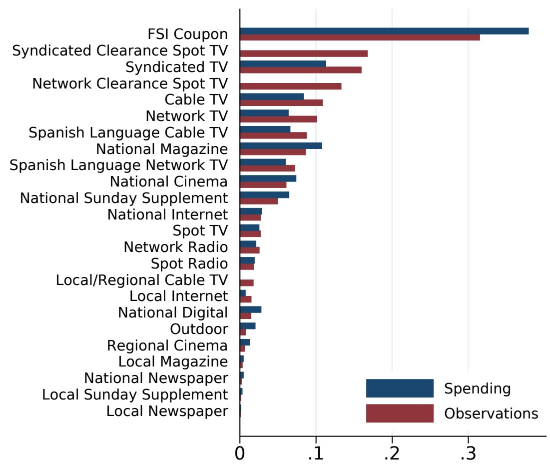
(b) Total number of observations



(c) Share of spending by media



(d) Share of food for each media



robustness. Spending with local TV advertising increased from 25 billion to nearly 31 billion.

While the ADI covers all sectors, we use the product classification system to select ads for food goods sold in grocery and drug stores. Figure A1 (d) shows that food products are advertised across all media types and are specially common in coupons, followed by local TV.

Table A1: List of Media Types covered by Ad Intel

Detailed Description	Our Description	Markets National/Local	Markets Number DMAs	Time period
Network TV	National TV	National	-	2010-2016
Spanish Language Network TV	National TV	National	-	2010-2016
Cable TV	National TV	National	-	2010-2016
Spanish Language Cable TV	National TV	National	-	2010-2017
Syndicated TV	National TV	National	-	2010-2016
Spot TV	LocalTV	Local	210	2010-2016
Network Clearance Spot TV	LocalTV	Local	210	2010-2016
Syndicated Clearance Spot TV	LocalTV	Local	210	2010-2016
Local/Regional Cable TV	LocalTV	Local	51	2010-2016
National Magazine	National Magazine	National	-	2010-2016
Local Magazine	Local Magazine	Local	31	2010-2017
FSI Coupon	Coupon	Local	78	2010-2017
National Newspaper	National Newspaper	National	-	2010-2016
National Sunday Supplement	National Newspaper	National	-	2010-2016
Local Newspaper	Local Newspaper	Local	76	2010-2016
Local Sunday Supplement	Local Newspaper	Local	5	2010-2016
Network Radio	National Radio	National	-	2010-2016
Spot Radio	Local Radio	Local	43	2010-2016
Outdoor	Outdoor	Local	164	2010-2016
National Internet	National Internet	National	-	2010-2016
Local Internet	Local Internet	Local	82	2010-2016
National Cinema	National Cinema	National	-	2013-2016
Regional Cinema	Local Cinema	Local	1	2013-2016

Notes: We do not cover data after 2016. National Internet and Local Internet has information until August 2017, and is then replaced by new Digital media type data. There is no spending for media types "Syndicated Clearance Spot TV". "Network Clearance Spot TV", and "Local/Regional Cable TV".

A.2 Algorithm to match retail sales and advertising data

The RMS and the ADI do not use the same identifiers and it is challenging to merge the two datasets. After studying different procedures, we developed a procedure that uses text similarity techniques for three distinct inputs that characterize identifiers across observations: product type descriptions, name of the firm, and name of the brand. Below we describe the algorithm (from the selection of these inputs and how we combine them to derive a criteria for a positive match), followed by alternative algorithms that we use for robustness analysis, and an extensive set of validation exercises.

Representative inputs – The retail sales data include three pieces of information that are key to distinguish across observations: a product module description (Nielsen RMS), a firm name text (GS1), and a brand description (Nielsen RMS). The advertising data is organized at a detailed brand code that uniquely distinguishes a product type, an advertiser parent/subsidiary, and a brand description.[12] Thus, we develop a crosswalk between datasets at the level of product module × firm × brand level (r) on the RMS, and at the level of product category × parent/subsidiary × brand (a) on the ADI.[13] None of these three dimensions – product type, firm, brand – is fully overlapping across the two data sets. For example, the brand identifiers in the ADI do not match with the brand descriptions in the RMS data. Either the advertised brand description is either more or less specific than the brand associated with each barcode. We also studied the possibility of using only brand descriptions in our match (as in Shapiro, Hitsch and Tuchman, 2020) and decided against it because there are several brand descriptions that are same or very similar and represent barcodes from distinct firms and product categories. By also using information on the nature of the product product advertised and the firm advertising we can reduce these sources of measurement error.

Measures of similarity for each input– We start by studying the product description on the ADI dataset to select the set of observations that are related to food consumer goods products. After studying the classification system of ADI, we select 271 distinct product categories, covering more than 5,000 distinct firms and almost 13,000 distinct brand descriptions. Using this selective sample, we proceed to compute measures of similarity across each of input.

We match the RMS **product modules** (in r) to the ADI product categories (in a) manually as a many-to-many crosswalk, with a few non-matched product types. We did it manually because there is a smaller number of doubtful matched that were better resolved by reading the barcode descriptions available on the RMS dataset and the product descriptions on ADI. In the baseline algorithm we attributed a similarity score $s_{ra}^{P} = 1$ if product category in a matches to module in r, and $s_{ra}^{P} = \frac{1}{2}$ if product category in a matches to other modules of group of module r but not to the module in r. We allow for this possibility because there are cases where the product category in ADI is too general and has too many close modules that constituted a good alternative match.

We determine the level of association of **firms** across datasets by using string matching algorithms applied to company name from GS1 and information on the advertiser from ADI. For most cases the advertiser parent and subsidiary coincide and when it does not coincide it is not clear to what level we should match to the GS1 names. Therefore, we use both variables in our algorithm. We start by running all names through a name-standardization

[12] ADI provides two distinct sources of firm information: advertiser parent and advertiser subsidiary.

[13] There are 147,665 unique RMS combinations, and 13,039 unique ADI combinations corresponding to 20,124 unique codes.

routine adapted from Argente, Baslandze, Hanley and Moreira (2020).[14] Using the name-standardization routine firm name we proceed to compute two measures of text similarity: a similarity score between firm GS1 in r and parent ADI in a (s_{ra}^{FP}) and another similarity score between firm GS1 in r and subsidiary ADI in a (s_{ra}^{FS}). We define the similarity scores of firm names as $s_{ra}^{F} = \max\{s_{ra}^{FP}, s_{ra}^{FS}\}$. The similarity scores are obtained using a token-based vectorial decomposition algorithm using log-weight to reduce false positive matches coming from words that are frequently found. The similarity is guaranteed to lie in the range $[0, 1]$, with zero corresponding to zero word overlap and one corresponding to the case in which the names are identical (or are multiples of one another). Almost 35% of parent/subsidiary on ADI have a similarity score of 1 which implies an exact match, and more than 60% of parent/subsidiary have a similarity score above 0.5.

We determine the level of association of **brands** across datasets by using string matching algorithms applied to brand descriptions from RMS and ADI. Merging brands warrants an additional challenge because brand variables in the RMS and ADI are not always specified at the same level. RMS assignes UPCs to brands, which are more aggregated than UPCs but are still typically disaggregated, with some exceptions. ADI assigns ads to brands and the ad itself may be specific to a reduced set of products or to the aggregated brand. For example, both datasets include brands like "Chobani" and specific sub-brands like "Chobani Simply 100", "Chobani Simply 100 Crush", or "Chobani Flip". After running all brand text through a name-standardization routine, we create two distinct brand definitions for each original brand b^0: a specific b^1 and a general brand b^2.[15] We compute measures of text similarity for each level of aggregation: a similarity score between original RMS brand r and original ADI brand a (s_{ra}^{B0}), similarity score between b^1 of RMS brand r and b^1 of ADI brand a (s_{ra}^{B1}), and similarity score between b^2 of RMS brand r and b^2 of ADI brand a (s_{ra}^{B2}). In our baseline specification, we use the similarity score $s_{ra}^{B} = s_{ra}^{B1}$, and used the other levels of aggregation for robustness exercises. As with the firm match, the similarity scores are obtained using a token-based vectorial decomposition algorithm using log-weight to reduce false positive matches coming from words that are frequently found. Almost 60% distinct ADI brands have a similarity score of 1 which implies an exact match, and only 10% have a similarity score below 0.5.

Criteria for match – The final step of the matching algorithm consists in using the similarity scores and determining the mapping between observations on the RMS and ADI.

[14] The routine handles capitalization, spaces, frequent abbreviations, common misspellings, among other.

[15] For example, for the original brand "Chobani Simply 100 Crush", b^1 becomes "Chobani Simply 100" and b_i^2 is "Chobani"; for the original brand "Chobani Simply 100", b^1 is the same and b^2 is "Chobani"; for the original "Chobani", both b^2 and b^1 are the same.

Every combination of the RMS product module × firm × brand level (r) with the ADI Intel product category × parent/subsidiary × brand (a) on the ADI is characterized by $\{s^F_{ra}, s^B_{ra}, s^P_{ra}\}$. After an extensive manual review of the datasets, we decided to treat every dimension equally and defined a systematic criterion that requires that at least one of the inputs needs to have a very high association (similarity score above 0.9), while still requiring that the other two dimensions have a sufficiently high degree of similarity (similarity above 0.7 and 0.5). More, specifically we define as a positive match when the following conditions are satisfied:

$$\text{Baseline} = \{ra \in \Omega \mid \max\{s^F_{ra}, s^B_{ra}, s^P_{ra}\} \geq 0.9 \ \wedge \ \text{median}\{s^F_{ra}, s^B_{ra}, s^P_{ra}\} \geq 0.7 \ \wedge \ \min\{s^F_{ra}, s^B_{ra}, s^P_{ra}\} \geq 0.5\} \tag{5}$$

There are clear trade-offs when setting these thresholds. Increasing these thresholds can potentially decrease the false positive matches but comes at the expense of largely increasing the false non-matches. Therefore, we consider decreasing the thresholds and evaluate its impact on the results:

$$\text{Match 2} = \{ra \in \Omega \mid \max\{s^F_{ra}, s^B_{ra}, s^P_{ra}\} \geq 0.8 \ \wedge \ \text{median}\{s^F_{ra}, s^B_{ra}, s^P_{ra}\} \geq 0.5 \ \wedge \ \min\{s^F_{ra}, s^B_{ra}, s^P_{ra}\} \geq 0.5\} \tag{6}$$

We also defined algorithms weighting differently the scores (s^P_{ra}, s^F_{ra}, and s^B_{ra}). More specifically, we define an alternative algorithm that satisfy the following conditions

$$\begin{aligned}
\text{Match 3} = \{ra \in \Omega \mid\ & ra \in B^2 \vee \\
& (s^F_{ra} \geq 0.9 \wedge s^{B2}_{ra} \geq 0.7 \wedge X_r = 1) \vee \\
& (s^F_{ra} \geq 0.9 \wedge s^{B2}_{ra} \geq 0.7 \wedge s^P_{ra} \geq 0.5 \wedge 2 \leq X_r \leq 5) \vee \\
& (s^F_{ra} \geq 0.9 \wedge s^{B2}_{ra} \geq 0.7 \wedge s^P_{ra} = 1 \wedge X_r \geq 6)\}
\end{aligned}$$

where X_a is the number of product modules in RMS of each firm-brand pair.

Level of aggregation – Using the matching algorithm, we produce a dataset at the level of product module × firm × brand level (referred to as *agg1*) after merging advertising into the RMS unit of analysis. When a RMS is matched with multiple ADI observations, we aggregate the advertising variables from all sources. In the case a single ADI merges into multiple RMS observations, we attribute to all RMS observations the same advertising variables. As discussed above, brand variables in the RMS and ADI are not always defined at the same level. Therefore, we also explored using two other distinct unit of analysis that ensure a smaller number of many-to-many. First, after applying the baseline matching, we created the dataset with retail and advertising at the product *group* × firm × brand level (referred to as *agg2*). Second, we created a dataset using a iterative procedure that aggregates the units of observation such that ensures a one-to-one matching. We use these two alternative aggregations of the baseline dataset in several robustness exercices throughout the paper (referred to as *agg3*).

Statistics of match – Table A2 provides statistics of the baseline dataset. Our baseline algorithm produces a many-to-many match that has 94, 364 total pairs, with 15, 742 distinct RMS observations and 11, 471 distinct ADI observations. This indicates that there is about six ADI observations per RMS observation matched, and eight observations per ADI unique code. To a large extent the many-to-many is a desirable feature because of the differences of aggregation across the two datasets, and the fact that often the ads do not target a specific brand or specific product modules. However, a large amount of many-to-many matches can also result of a datasets that does not produce a close link between retail and advertising information. Both alternative matching algorithms *match 2* and *match 3* produce many more matches and with relatively smaller increase in the total unique observations matched on the ADI.

We perform validation exercises to evaluate the robustness and quality of our match. We

Table A2: Match Statistics

		Baseline	Alternative	
	All	Match	Match 2	Match 3
RMS				
Unique codes	147,665	15,742	17,998	29,418
Observations r	147,665	15,742	17,998	29,418
Product Modules	603	580	583	596
Firm name	23,784	1,935	5,530	8,849
Brand Description	62,820	7,409	2,486	3,868
ADI				
Unique codes	20,124	11,471	12,203	13,793
Observations a	13,039	6,436	6,948	8,091
Product Categories	271	259	261	264
Parent Advertiser	5,070	1,722	1,842	2,186
Subsidiary Advertiser	5,562	2,009	2,139	2,518
Brand Description	10,138	4,631	5,022	5,921
Paires Matches		94,364	138,101	560,312
average obs/unique RMS		6	8	19
average obs/unique ADI		8	11	41

Note: The table reports the number of the distinct observations for each dataset and input of our algorithm in column "All". Column "Baseline Match" shows the statistics of the baseline match algorithm, and columns "Alternative" shows the statistics for the alternative algorithms.

use three main types of validation exercises: we evaluate manually the impact of each step of the algorithm (described above); an external validation using a sample dataset matching RMS and ADI; use statistics on cohorts and exit time for sales and advertising.

External Validation – For a random sample of 0.5% RMS observations, we manually find the best matched ADI observation. To ensure an independent selection of the best match, we assigned this task to independent readers that did not know any details on our algorithm. The external manual match found a match for less than 1/3 of the observations and the remaining were considered not matched. This manually checked data sample serves as our benchmark reference.

We start by comparing the comparing the many-to-many feature of our algorithm. Table A3 lists the matching pair distribution for the three criteria and compares it to the manually checked data. Among them, the *baseline match*, *match 2*, and *match 3* represent different selection rules (from most strict to least strict) with respect to firm name, brand description and the product category. The manual procedure generates more than half unique matches, while our baseline algorithm produces about 40% of unique matches, comparing with about 35% by *match 2* and 30% by *match 3*. As expected, allowing for lower thresholds of similarity scores of our inputs may increase the cases of many-to-many and may reduce the precision.

Next, assuming that the manual matching produces the true match, we evaluate the performance of our algorithm by categorizing each matched pair by the algorithms relative to the manual into five mutually exclusive cases. Starting with RMS observations that the algorithms matched, we can have the following cases:

- *True-positive* – The RMS-ADI matching pair in the algorithm is identical to the manually pairs;

Table A3: Distribution of number of ADI observations per RMS observations

Number	Baseline Match	Alternative Match 2	Alternative Match 3	Manual
1	6,225	6,601	9,013	108
2	2,173	2,297	2,849	25
3-5	3,016	3,389	5,014	35
6-10	1,860	2,150	3,714	18
11-20	1,489	1,897	2,783	1
21-50	811	1,298	2,827	2
+50	168	366	3,218	-
Total	15,742	17,998	29,418	189

Notes: The table reports the distribution of the number of ADI observations per RMS observation. For example, there are 2,173 unqiue RMS observations that match to two ADI observations, generating 4,346 matched pairs.

Table A4: Validation results for different matching creterion

	Matched						Not matched			
	True		False-absolute		False-redundant		true		False	
	pairs	*unique*	*pairs*	*unique*	*pairs*	*unique*	*pairs*	*unique*	*pairs*	*unique*
Baseline	288	99	45	10	43	2	383	383	77	77
Match 2	317	105	42	11	61	4	379	379	70	70
Match 3	398	142	276	6	97	2	362	362	38	38

Notes: The table reports the distribution of matches of the baseline and alternative matches when comparing the with the manual match. Each pair is classified according to 5 possible cases. The *pairs* represent the number of (r,a) combinations that are included for each case. The *unique* represent the number of RMS observations underlying the pairs.

- *False-positive absolute* – The RMS-ADI matching pairs from the algorithm is different from the manually checked data sample. The absolute means that the RMS shold have been matched (it was in the manual) but the matched ADI observations are different;

- *False-positive redundant*: The RMS-ADI matching pairs from the algorithm are different from the manually checked data sample. The redundant means that because of many-to-many matching issues, we have multiple matching pairs for the same RMS, and some are correct and some do not coincide with manually matching.

A natural outcome of our algorithm is that many firms that have sales do not advertize. In the cases our algorithm did not produce any match, we can have the following cases:

- *True-negative* – The RMS observation has not matched to any ADI observation, coinciding with the manual sample also does not have any matching observations.

- *False-negative* – The RMS observation has not matched to any ADI but the manual sample assigns ADI observations.

The main validation test results are summarized in the table A4. Regarding the three different selection criteria, the baseline match has the lowest true-positive matches while *match 3* has the highest true-positive matches. This is not surprise because the number of unique RMS observations matched in *match 3* is double that in the baseline match (see Table A2). But more true-positive matches comes with the cost of more false positive observations: 45 absolute false-positive matches in the baseline match versus 276 in *match 3*, and 43 redundant false-positive matches in the baseline versus 97 in *match 3*. Moreover, when evaluating with the negative matching outcomes, the baseline match outperforms *match 3* with 383 versus 362 true-negative matching results.

Our conclusion from this exercise is that the baseline match performs best. While, each selection criterion has its own advantage, our goal is to minimize false positives. A doubtful match with many-t0-many will reduce the precision of our results and contaminate our final estimation results.

Cohort and exit Year – As described above, our matching algorithm uses information on the product type, firm and brand. We do not use information on the location (dma) where the sales and advertising occurs and on the year where those two activities occur. Therefore, we use information on location and year to test our algorithm. More specifically, we expect that the majority of adverting does not occur more than a year before the product is introduced in the market and generates sales. To test the conjecture regarding the entry time and location, we compute the first time we observe any sale in a particular market (cohort RMS) and first time we observe any advertising in a particular market (cohort ADI), conditional on matching. We expect that the difference between the cohort of ADI and RMS to be larger than -1. Figure A2 presents the histogram of this difference for our baseline algorithm. The plot shows that the majority of product module × firm × exhibits advertising around or after entry into a market, and the results are even more striking when we use *Agg2* and *Agg3*.[16]

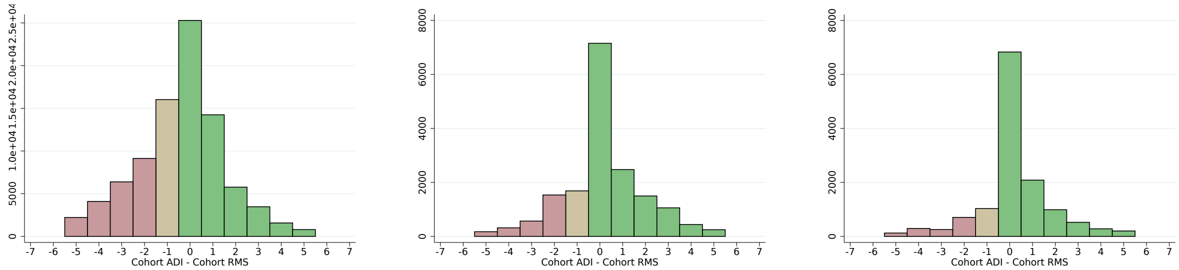
Likewise, we do not expect to have observations with advertising more than one year after the product no longer registers any sale. We compute the exit time in RMS and ADI and compare its difference. We conclude that, as expected, the majority of observations first ceases to advertise before stopping selling in that location, especially when we use *Agg3*.

We also compute similar statistics when we compute cohort and exit time in RMS and ADI across all markets. Our results indicate that as expected, the entry of advertising before entry of sales and the exit of advertising after exit of sales have a small probability.

[16]Note that this test is not fully accurate. The RMS dataset do not include all sales, and thus it is possible to have advertising before any sales in the RMS stores is there are sales in non-RMS stores.

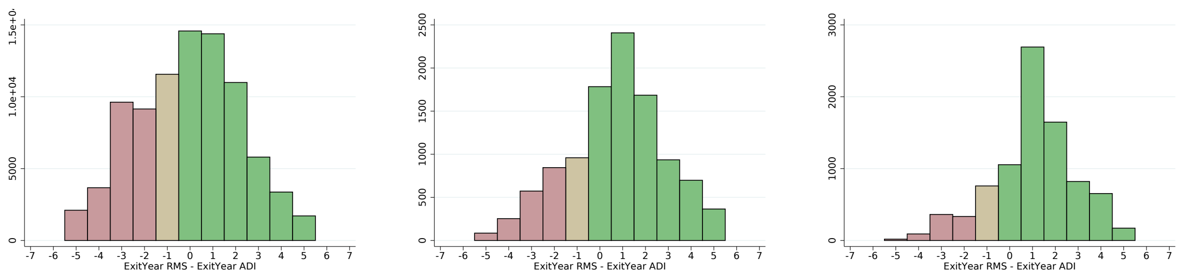
Figure A2: Differences in cohort and exit year

Cohorts DMAs



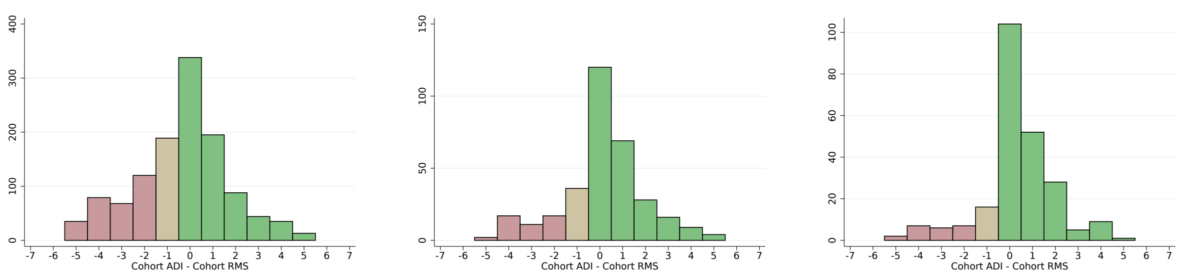
Baseline | Baseline Agg2 | Baseline Agg3

Exit Time DMAs



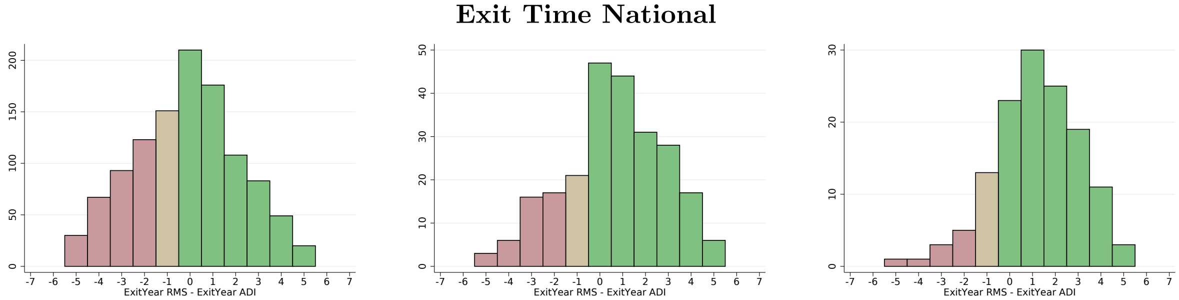
Baseline | Baseline Agg2 | Baseline Agg3

Cohorts National



Baseline | Baseline Agg2 | Baseline Agg3

Exit Time National



Baseline | Baseline Agg2 | Baseline Agg3

Notes: The plots show histograms of number of observations for the difference (first and third rows) between the cohort in ADI and cohort in RMS, (second and forth rows) between exit time in RMS and exit time in ADI. The first two rows compute cohort and exit time at the level of a market DMA, while the the last two rows at the national level. The red bars indicate possible measurement error due to matching procedure.

A.3 Retail data

RMS product classification scheme – The data are organized into a hierarchical structure. There are ten Nielsen-defined departments. Within each department, there are Nielsen-defined product groups, and within each product group, there are Nielsen-defined product modules. For example, in the Dairy department "YOGURT" is a product group, while "YOGURT-REFRIGERATED-SHAKES & DRINKS" is a product module within "YOGURT." There are more than 1,000 product modules in the full set of ten departments, and 600 modules in food departments.

Within the RMS, the finest level of disaggregation is the UPC. For each UPC, we observe the unit of quantity, e.g. "oz" , "quart", etc., and the number of units in a package. This allows us to compare quantities across UPCs within products by calculating unit values. For example, if Dannon brand yogurt is sold in 42 oz. tubs, while Chobani brand yogurt is sold in 4 oz. 6-packs, we can use ounces as the common unit of quantity, with Chobani being coded as 24 units and Dannon as 42 units. Nielsen also classifies UPCs into brands. Brands (e.g. Yoplait, Chobani) often cross products module. Firms (e.g. General Mills) have multiple brands. Table A5 provides more details of the hierarchical structure of the data. It reports the number of distinct markets, products, firms, firm-products, firm-brands, firm-brand-products, and firm-brand-products-markets as well of those present on average in a given year.

Store location – For each participating store, we know the parent chain, the 3-digit zip code, and the Nielsen Designated Market Area (DMA) where it operates. There are 210 Nielsen-defined DMAs (approximately 14 counties per DMA), with coverage in nearly every major U.S. metropolitan area. For example, the Philadelphia DMA includes eight surrounding counties in Pennsylvania, eight counties in New Jersey, and two in Delaware.

Table A5: Number of observations in different categories

	Avg yearly	Total distinct	Avg Entry	Total Entry
Baseline				
Product(firm-brand-module)	60,086	116,107	1,927	22,870
Firm-brand-module-markets	2,018,137	4,478,616	93,027	1,100,795
Other				
Modules	602	603	–	–
Firms	12,620	21,265	561	6,637
Firm-modules	41,087	72,500	1,024	1,2140
Firm-brands	32,354	63,230	1,079	12,818

Notes: Entry numbers exclude 2006.

Any cable provider serving a customer in one of these 18 counties is required to include local Philadelphia broadcast stations in the customer's cable package. We define markets as DMAs because i) they are large enough to be segmented from consumers' perspective, ii) they align well with MSAs across the country, and iii) this definition allows us to match the RMS with the ADI data.

DMAs versus retail chains – A potential concern of defining markets as DMAs is that entry of products occurs at the retail chain level instead of the local level (e.g. national retailers synchronizing entry of products across all their locations). Table A6 shows this is not the case. The top panel shows that the average new brand is introduced in 78% (median 0.86%) of the markets in retailers that operate in less than 5 markets, those that are mostly local. On the other hand, the average brand is introduced only in 15% (median 7%) of the markets where national retailers operate, those present in more than 150 markets in the US. These numbers are very similar for brands that have been in the market for at least 40 quarters. Brands are sold in 75% of the markets covered by local retailers and only in 18% of the markets covered by national retailers. Overall, we do not find strong evidence that national retailers synchronize the entry of brands across all their locations. Instead, we find that market penetration of a product-brand is far from complete. Our empirical analysis also relies on entry at the brand-product level that is non-synchronous across markets.

Table A6: Share of DMAs Covered by the Chain Where a Product-Brand is Present

Age of the Brand quarters)	Number of DMAs Covered by Chain	Mean	p25	p50	p75
1	<5	0.78	0.60	0.86	0.97
1	5-50	0.38	0.15	0.30	0.60
1	50-150	0.19	0.06	0.12	0.29
1	>150	0.15	0.02	0.07	0.23
40	<5	0.75	0.64	0.77	0.85
40	5-50	0.44	0.28	0.42	0.59
40	50-150	0.25	0.15	0.22	0.34
40	>150	0.18	0.11	0.16	0.24

Note: The table documents patterns of the geographic roll-out of brands. We consider brands at entry and brands that have been sold in the market for at least 10 years (40 quarters). We divide retailers into bins according to the number of markets (DMAs) in which they operate in order to distinguish between local, regional, and national retailers. The table reports the average, median, 25^th percentile and 75^th of the share of markets a brand is sold at.

Entry, age, survival – We say that a product enters in a new market in year t if it has zero sales in that market in year $t-1$, and positive sales in year t. Entry is censored in 2006, and exit is censored in 2017. Note that brands can (and do) enter a product-market multiple times during the sample. We define market *age* as the cumulative number of periods of continuous market participation at the product-brand-market level. Completed spell *survival* is the maximum market age achieved in a product-brand-market-level sales spell, i.e. the market age on exit for that spell. Table A7 illustrates these definitions for a hypothetical product-brand pair in a series of markets. In this table, as in our implementation, age and survival are top-coded at 5 years. This allows us to assign a survival to (some) sales spells where entry is observed, but exit is censored by the end of the sample.

Baseline sample selection – We select data covering the food sector over the period 2006-2017. We define products as the combination firm-brand-product. We make use of six of these departments in our analysis: Dry Grocery, Dairy, Deli, Packaged Meat, Frozen Foods, and Fresh Produce. We exclude the departments Alcoholic Beverages, Health & Beauty Care, Non-Food Grocery and General Merchandise.

Table A7: Illustrative example of definitions

Year	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016
Market	a. Participation										
A	X	X	X					X	X	X	X
B		X	X	X	X		X	X	X		
C	X	X			X	X	X	X	X	X	X
D		X	X			X	X	X	X	X	
E	X	X	X	X	X	X	X	X	X	X	X
Market	b. Market age, topcoded at 5										
A	cens	cens	cens					cens	cens	cens	cens
B		1	2	3	4		1	2	3		
C	cens	cens			1	2	3	4	5	5	5
D		1	2			1	2	3	4	5	
E	cens	cens	cens	cens	cens	cens	cens	cens	cens	cens	cens
Market	c. Completed spell survival, topcoded at 5										
A	cens	cens	cens					cens	cens	cens	cens
B		4	4	4	4		3	3	3		
C	cens	cens			5	5	5	5	5	5	5
D		2	2			5	5	5	5	5	
E	cens	cens	cens	cens	cens	cens	cens	cens	cens	cens	cens

Our baseline data set combines the sales of a product across all stores covered in the sample over a year. We aggregate from weekly data to the annual level, to avoid spurious entry and exit for seasonal products. We define sales of a product-brand as the total sales across all stores and weeks in the year and market. Likewise, we define quantity as the total quantity sold across all stores and weeks in the year-market, and price is the ratio of sales to quantity, which is equivalent to the quantity weighted average price. Throughout the paper, we present evidence also using quarterly level aggregation.

To minimize measurement error, we consider product-firm-brand pairs with an average revenue greater than 100 dollars in a given market over our sample period.

A.4 Adverting data

A.4.1 Local TV

Media types – Local TV advertising results from both Spot TV, Network Clearance Spot TV and Syndicated Clearance Spot TV.

There are the following types of TV advertising: (1) *Cable ads*, which are aired nationally, viewership data are available only at the national level; (ii) *Spot ads* are bought locally, and viewership measures are recorded locally, separately for each DMA; (iii) *Network and Syndicated Spot TV ads* are recorded in national occurrence files that can be matched with local measures of viewership in each DMA. The Network TV and Syndicated TV occurrence files record the ads at the national level (i.e. the date and time the ad is supposed to be broadcast at every local station). The Network Clearance Spot TV and Syndicated Clearance Spot TV occurrence files record them at local level (i.e. the date and time each ad is actually broadcast at every local station). The local channels have some authority to replace or move nationally scheduled ads. For each network program, ads from the local market are compared to commercials in the national database, thus identifying the local TV stations that did or did not air the "national" ad.

Variation in a brand's aggregate ad viewership across markets is due to both variation in occurrences across markets (more Spot ads were aired in market A than in market B) and variation in impressions (eyeballs) across markets (a Network or Syndicated ad aired in both markets A and B, but more people saw the ad in market A than in market B).

Collection method across DMAs – Markets are measured at the DMA level. For the largest 25 DMAs we use actual data recovered from set-top box devices, for the smaller DMAs we use data from diaries and imputation methods. Local viewership is measured differently across DMAs. In the 25 TV markets with the highest sales (e.g. New York, Los

Angeles, Chicago, Denver) the Local People Meter (LPM) is measured. Individuals register individually, the measurement is carried out on 365 days over 24 hours (and it is available monthly in our dataset). In 31 smaller markets (such as Nashville, Salt Lake City) they use the SET Meter (Diary and Electronic). In four sweeps in the months of February, May, July and November, target group data (that is, demographics information) are collected with the diary and validated with the data of the devices (TV set on/off) in the participating households. In the 154 TV markets with the lowest sales (e.g. Harrisburg, Honolulu), the use of TV is only recorded by means of a diary survey.

Each year, ADI processes approximately two million paper diaries from households across the country, for the months of November, February, May and July – also known as the "sweeps" rating periods. Seven-day diaries (or eight-day diaries in homes with DVRs) are mailed to homes to keep a tally of what is watched on each television set and by whom. Over the course of a sweeps period, diaries are mailed to a new panel of homes each week. At the end of the month, all of the viewing data from the individual weeks is aggregated. Because of the differences in measurement between markets, not all months have impression data available. This requires some method of imputing the data in missing months.

Variables – We are interested in the following measures of advertising: (i) indicator variable on whether there was some ad occurrence; (ii) number of ads occurrences; (iii) number of impressions. We also compute measures total duration, spending, and Gross Rating Points (GRPs).

GRP is measure of how many people within an intended audience might have seen the ad. It can take a value above 100. Say that an ad is seem by 40% population and it is aired 3 times, then we have 120. Rating is the percentage (0 to 100) of the Media Market that will likely be exposed to your advertisement (in this case 40%). In other words, it is used as a cumulative measure of the impressions an ad campaign can achieve.

The advertising information is recorded at the *occurrence* level, where an occurrence is the placement of an ad for a specific brand on a given channel, in a specific market, at a given day and time. Occurrences can then be linked to viewership/impressions using the natural keys described by Nielsen. Viewership data estimate of the number of *impressions*, or eyeballs, that viewed each ad. We use estimates of *universe* of TV audience in each market.

After merging merging the occurrences, impressions and universe files, we obtain a dataset organized in a per-occurrence manner, so the next step is to collapse occurrences into a unique time $\times$ dma $\times$ brand code (varies by product type, an advertiser parent/subsidiary, and a brand description). We create datasets collapsing at the quarterly and annual levels.[17]

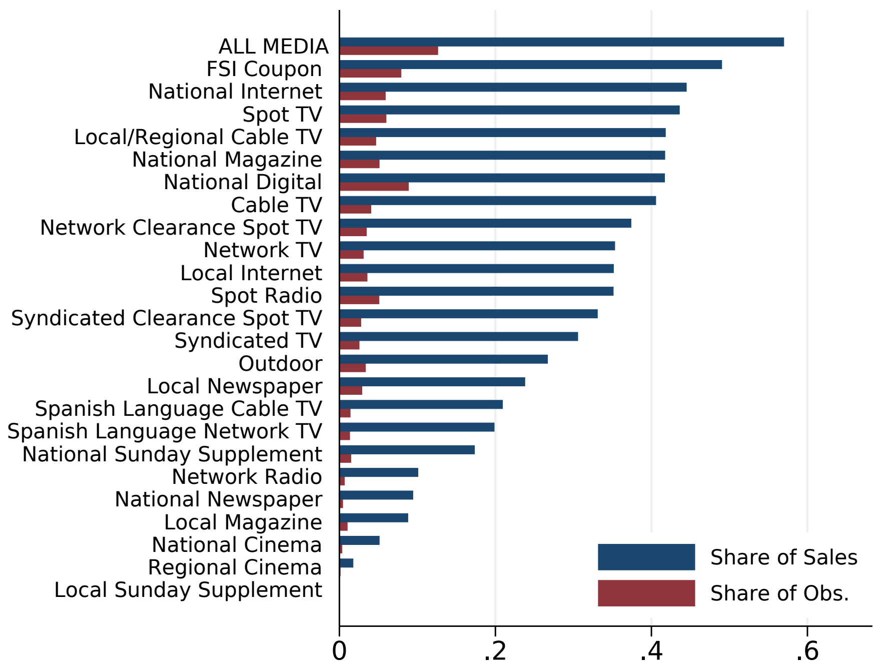
A.4.2 Coupon –

A.4.3 Other media types –

A.5 Matched retail and advertising dataset

A.5.1 Share of advertising products

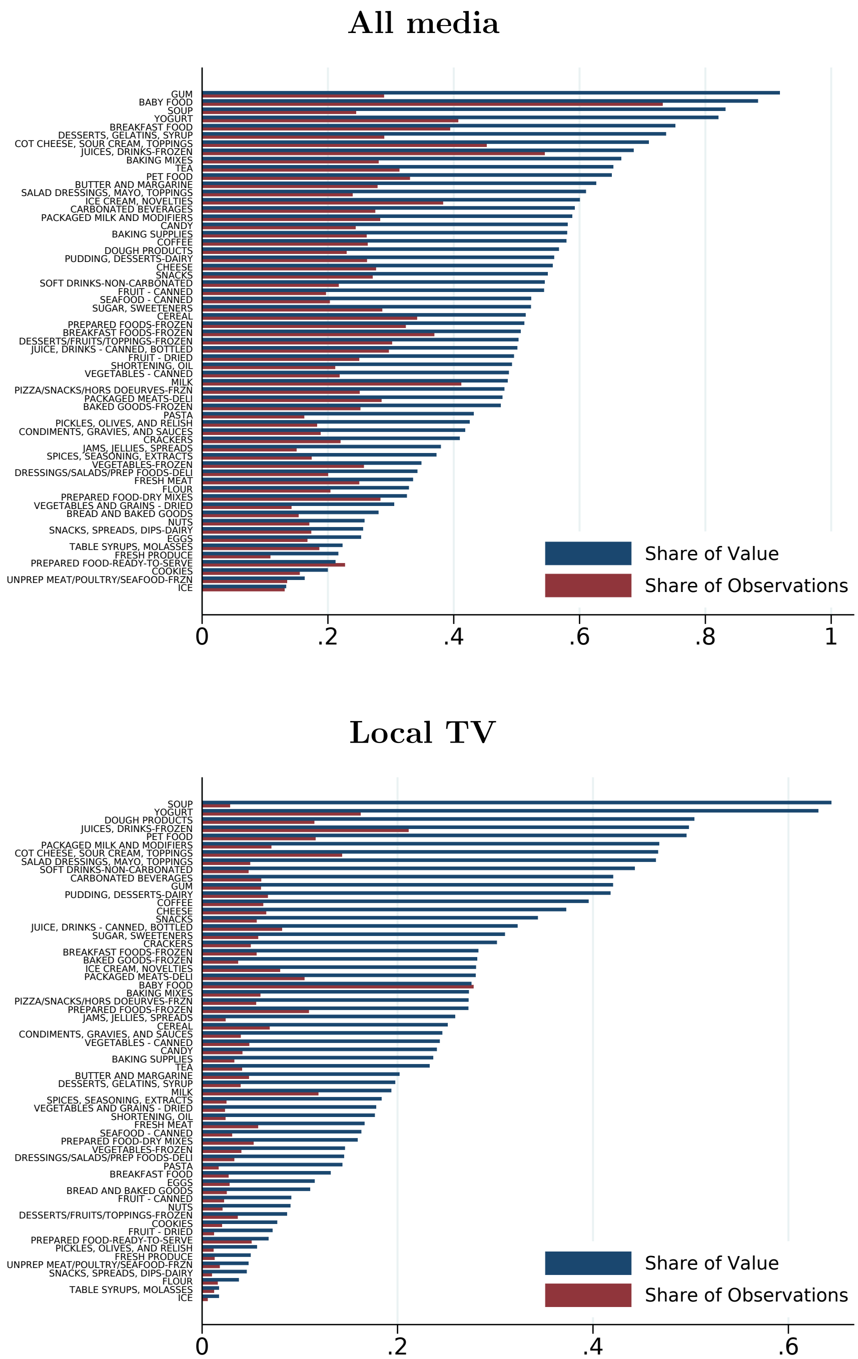
Figure A3: Share with some advertising by Media



Note: The figure presents the share of RMS observations/sales that used advertising in any year between 2010–2017 by media type. "ALL MEDIA" refers to share at any media across the distinct media types. Sales are measured as total sales in period 2010-2017.

[17] Because we have high time frequency data but not all markets are observed all the time, so we follow Shapiro, Hitsch and Tuchman (2020) imputation procedure for the missing observations.

Figure A4: Share with some advertising by product group - All media



Note: The figure presents the share of RMS observations/sales that used advertising in any year between 2010–2017 by media type. "ALL MEDIA" refers to share at any media across the distinct media types. Sales are measured as total sales in period 2010-2017.

A.5.2 Basic descriptive statistics

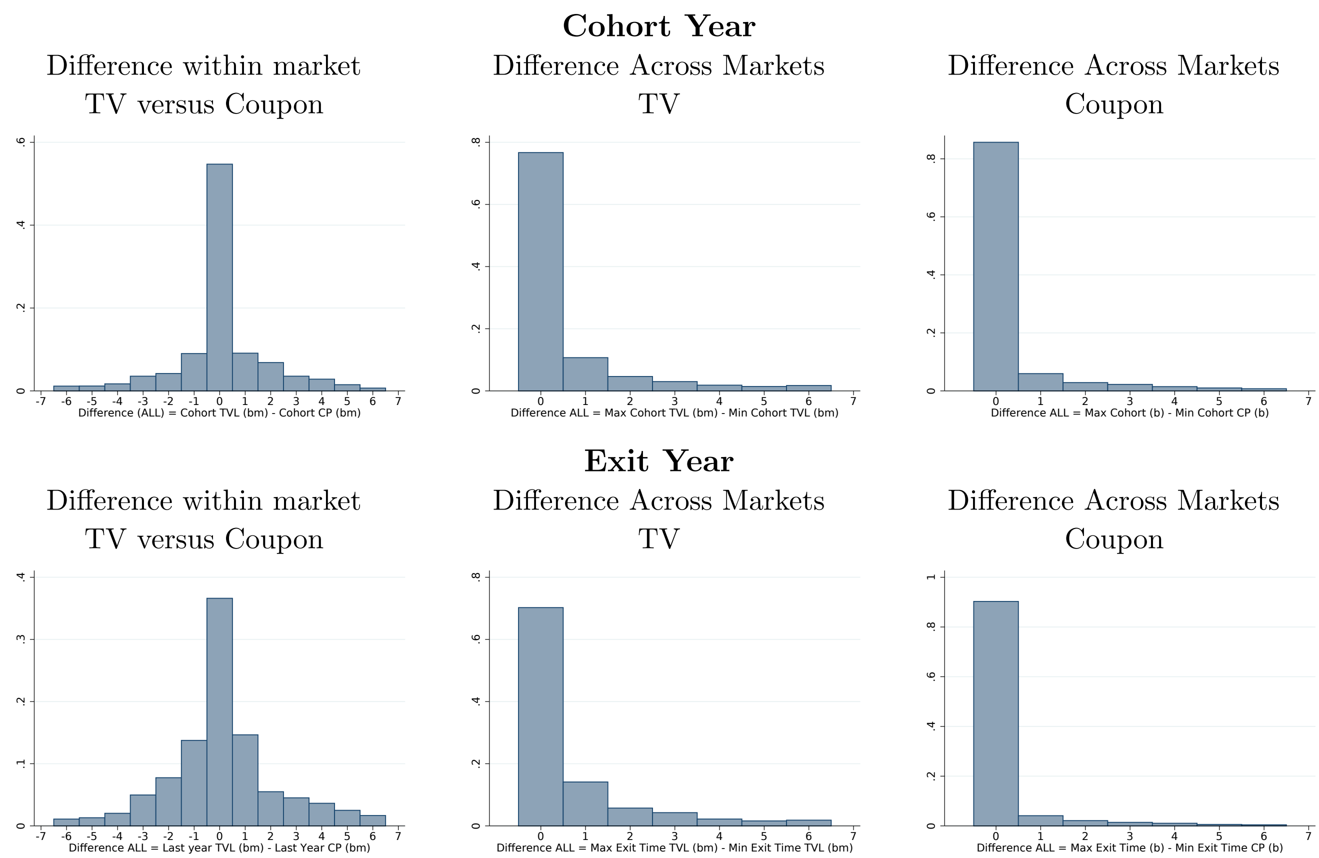
Table A8: Descriptive Statistics

	(1) All	(2) New	(3) Young	(4) Mature	(5) Small	(6) Medium	(7) Large
Observations	109918	48668	16765	44485	36842	36634	36442
Some Advertising (%)	22.8	19.5	21.0	27.1	15.5	20.4	32.5
(% sales)	67.5						
Some TV or CP Advertising (%)	7.7	5.1	6.3	10.9	2.8	4.9	15.3
(% sales)	52.9						
Some Local TV (%)	4.6	3.0	3.9	6.7	1.7	2.8	9.5
Average number of markets	73.2	87.2	71.7	66.5	14.5	41.8	93.0
Average share of markets (%)	55.1	63.1	59.9	50.0	72.4	44.2	55.2
Conditional on advertising on a market							
Share of years (%)	58.5	68.7	60.6	53.0	85.3	57.8	53.9
Average number Ads	363.8	371.1	325.8	368.5	707.9	286.2	325.0
Average GRPs	11.5	11.5	9.8	11.9	14.4	9.5	11.6
Average Impressions (thousands)	9723.9	8656.2	9666.2	10269.7	23306.1	8975.2	7506.5
Some Coupon (%)	6.4	4.3	5.2	9.1	2.2	3.9	13.2
Average number of markets	29.5	31.1	29.1	28.8	8.0	18.9	36.3
Average share of markets (%)	60.6	61.8	62.9	59.5	70.6	54.2	60.8
Conditional on advertising on a market							
Share of years	58.8	66.2	60.8	54.6	83.8	61.8	53.7
Average number Ads	4.5	4.1	4.8	4.6	6.0	4.1	4.3
Average spending	4.9	4.3	5.6	5.0	7.6	4.6	4.5
Both Local TV and Coupon (%)	3.4	2.2	2.8	4.9	1.1	1.7	7.3
Average number of markets	32.4	36.1	33.9	30.3	8.4	20.9	38.7
Average share of markets (%)	57.6	62.0	62.7	54.4	70.6	49.6	57.6
Conditional on advertising on a market							
Share of years (%)	48.0	58.0	50.2	42.7	82.8	50.5	42.3

Note: This table shows descriptive statistics for firm × brand1 × module (Aggregation 1). *Some Adverting* refers to the share of observations that are matched to ADI (in percent). Likewise, *Some Local TV*, *Some Coupon*, and *Both Local TV and Coupon*refers to the share of observations (in percent) that used local adverting, coupon or both, anytime in the period 2010-2016 in at least one market. For each type, and conditional on advertising in at least one market, we compute the average number of markets and the average share of markets with advertising relatively to the markets with positive sales. For each type, and conditional on advertising in that market, we compute the average number of ads, and other adverting variables. Column (1) presents the results for all observations, column (2)-(4) shows the results by age, column (5)-(7) shows the results by size tercile (within group). *New* is the set of observations with entry in the period 2010-2016, *Young* refers to observations that were born between 2007-2009, and *Mature* refers to observations that already existed in 2006.

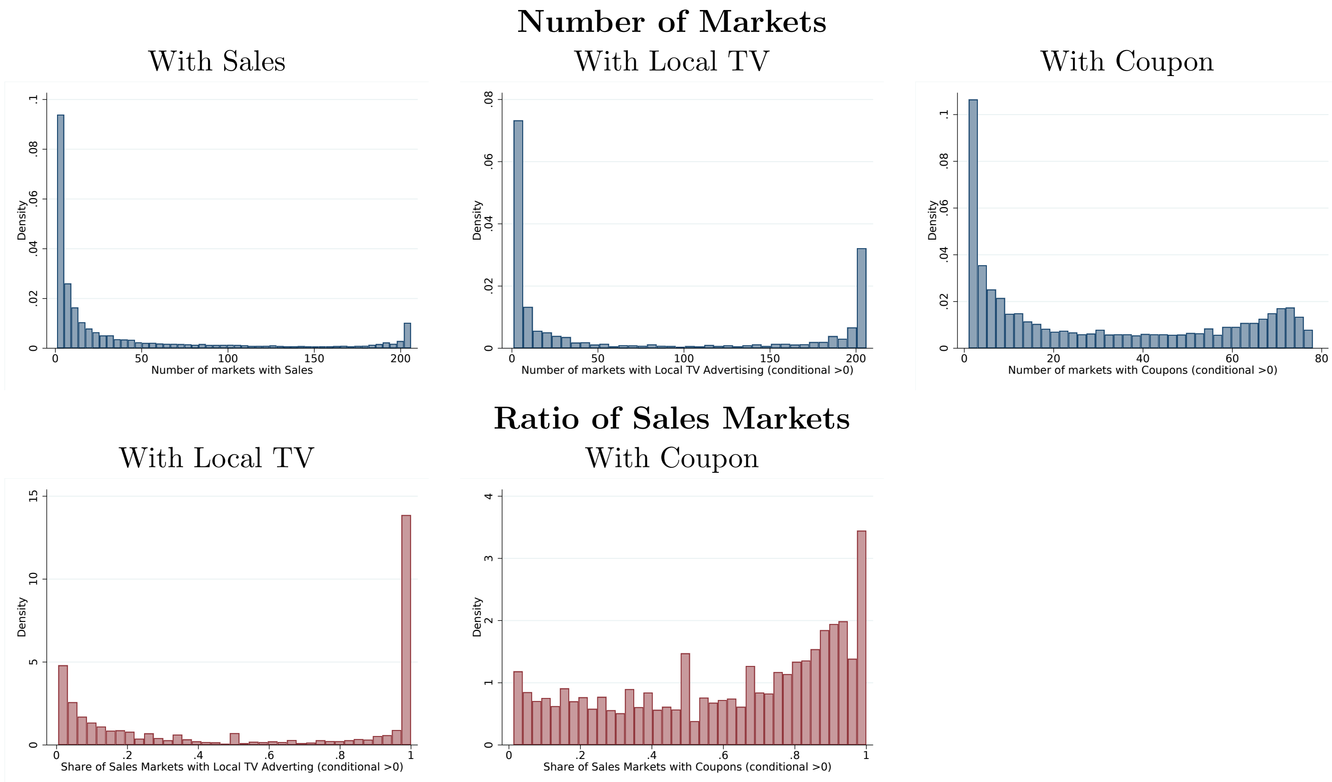
A.5.3 Synchronization of advertising across distinct media and markets

Figure A5: Differences in cohort and exit year



Notes: ...

Figure A6: Distribution of number of markets



Notes: ...

B Additional Results

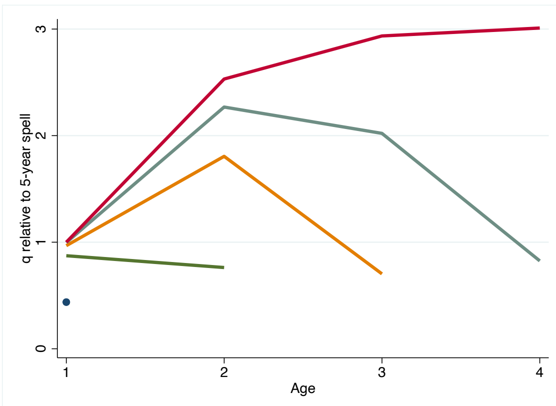
Table B1: Dynamics of Revenue, Quantity, and Price

Dependent Variable (Logs)	Revenue	Quantity	Price
Spell Length		Spell Intercept	
2 years	0.631***	0.0464***	0.585***
	(0.00456)	(0.000758)	(0.00458)
3 years	0.741***	0.0537***	0.687***
	(0.00510)	(0.000823)	(0.00513)
4 years	0.780***	0.0536***	0.727***
	(0.00575)	(0.000903)	(0.00579)
5+ years	0.798***	0.0492***	0.749***
	(0.00502)	(0.000813)	(0.00504)
Right cens.	1.060***	0.0379***	1.022***
	(0.00462)	(0.000761)	(0.00463)
Left cens.	2.539***	0.0313***	2.508***
	(0.00441)	(0.000743)	(0.00441)
Market Age		2-year Spells	
2 years	-0.187***	-0.0697***	-0.118***
	(0.00442)	(0.000719)	(0.00445)
Market Age		3-year Spells	
2 years	0.568***	-0.0247***	0.593***
	(0.00445)	(0.000676)	(0.00450)
3 years	-0.342***	-0.0757***	-0.266***
	(0.00544)	(0.000840)	(0.00546)
Market Age		4-year Spells	
2 years	0.783***	-0.0153***	0.798***
	(0.00541)	(0.000793)	(0.00547)
3 years	0.694***	-0.0200***	0.714***
	(0.00573)	(0.000843)	(0.00579)
4 years	-0.198***	-0.0734***	-0.124***
	(0.00681)	(0.00104)	(0.00683)
Market Age		5-year Spells	
2 years	0.901***	-0.0113***	0.913***
	(0.00356)	(0.000508)	(0.00360)
3 years	1.062***	-0.00984***	1.072***
	(0.00372)	(0.000534)	(0.00377)
4 years	1.084***	-0.0103***	1.094***
	(0.00381)	(0.000549)	(0.00385)
5 years	1.066***	-0.0167***	1.082***
	(0.00350)	(0.000502)	(0.00354)
Prod-brand-year	Yes	Yes	Yes
Prod-mkt-year	Yes	Yes	Yes
Observations	23,951,656	23,951,656	23,951,656
R-squared	0.762	0.967	0.803

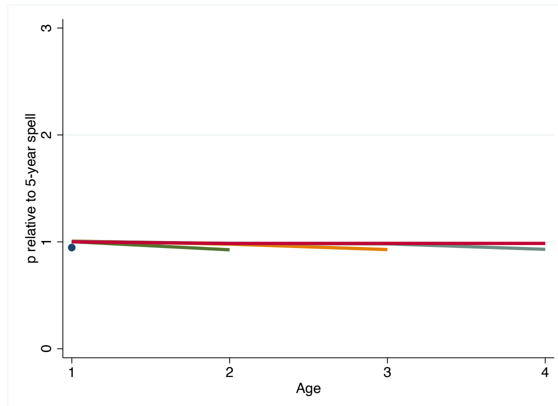
Notes: Dependent variable is an indicator for exit in the next period. The omitted category is market age equal to one year. *** significant at 1%. Source: Calculated based on data from The Nielsen Company (US), LLC and marketing databases provided by the Kilts Center for Marketing Data Center at The University of Chicago Booth School of Business.

B.1 Quantity and Price Dynamics Within Markets

Figure B1: Quantity and Price Dynamics Within Markets - Firms



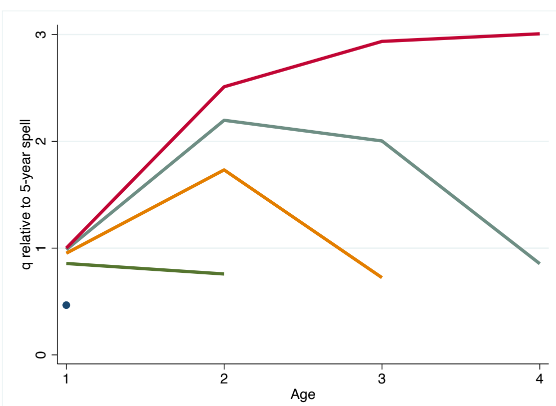
(a) Quantity



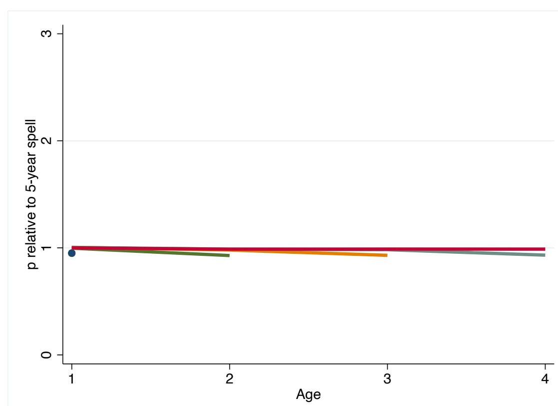
(b) Price

Notes: Based on exponentiating appropriate sums of coefficients after estimating equation 2 on data aggregated at the firm level. Source: Calculated based on data from The Nielsen Company (US), LLC and marketing databases provided by the Kilts Center for Marketing Data Center at The University of Chicago Booth School of Business.

Figure B2: Quantity and Price Dynamics Within Markets - Broad Brands



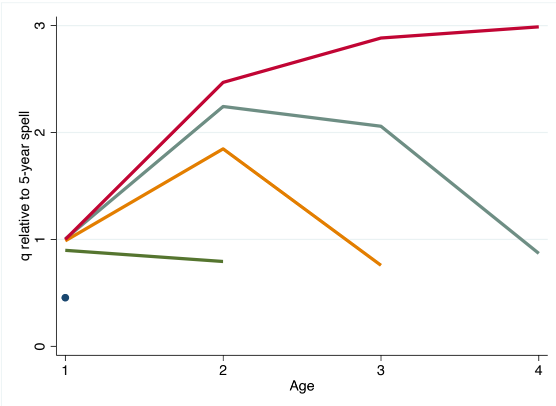
(a) Quantity



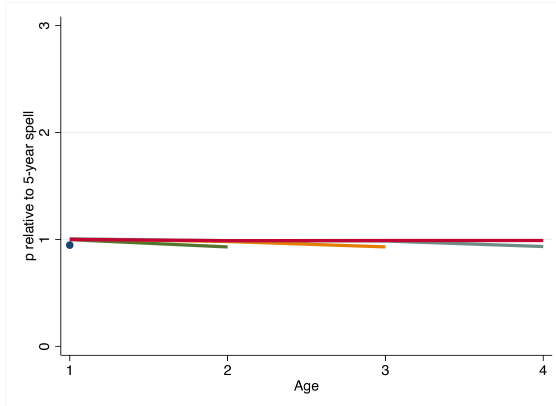
(b) Price

Notes: Based on exponentiating appropriate sums of coefficients after estimating equation 2 on data aggregated at using a definition of brands where we combine brands with similar names. Source: Calculated based on data from The Nielsen Company (US), LLC and marketing databases provided by the Kilts Center for Marketing Data Center at The University of Chicago Booth School of Business.

Figure B3: Quantity and Price Dynamics Within Markets - Only Original Brands



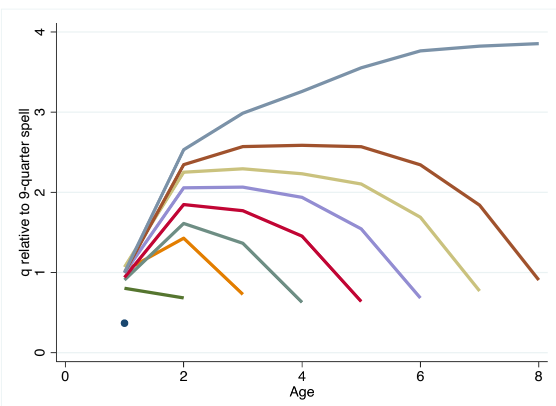
(a) Quantity



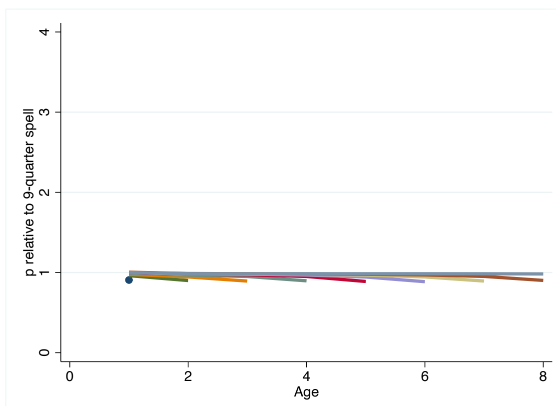
(b) Price

Notes: Based on exponentiating appropriate sums of coefficients after estimating equation 2 on data using only the brands each firm had at entry. Source: Calculated based on data from The Nielsen Company (US), LLC and marketing databases provided by the Kilts Center for Marketing Data Center at The University of Chicago Booth School of Business.

Figure B4: Quantity and Price Dynamics Within Markets - Quarterly



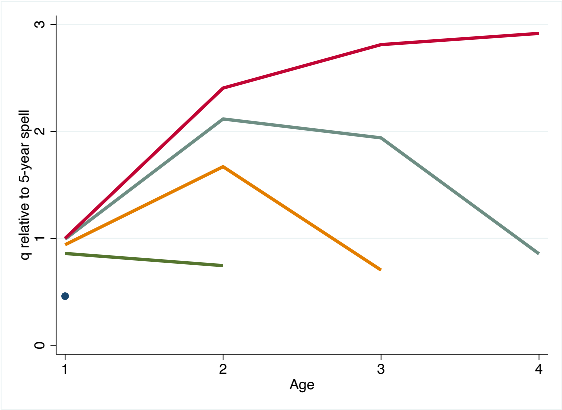
(a) Quantity



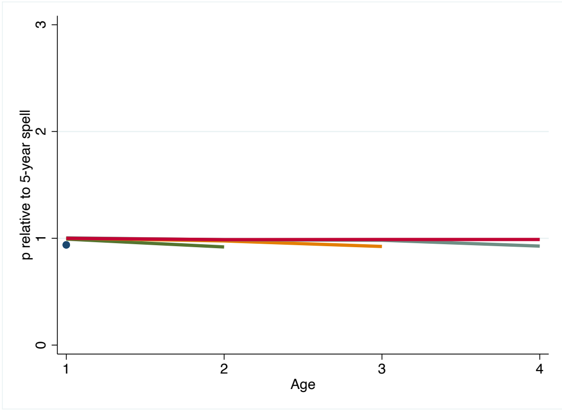
(b) Price

Notes: Based on exponentiating appropriate sums of coefficients after estimating equation 2 on data aggregated at the quarterly level. Source: Calculated based on data from The Nielsen Company (US), LLC and marketing databases provided by the Kilts Center for Marketing Data Center at The University of Chicago Booth School of Business.

Figure B5: Quantity and Price Dynamics Within Markets - Balanced Stores



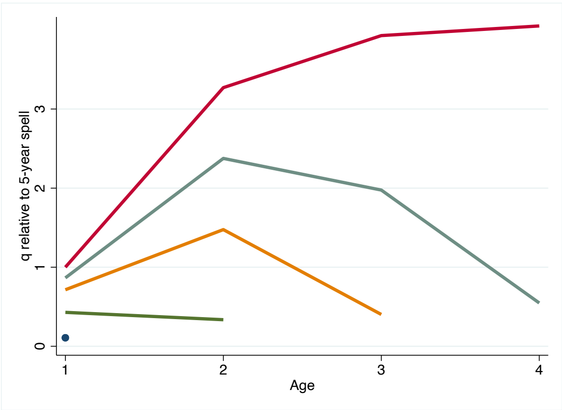
(a) Quantity



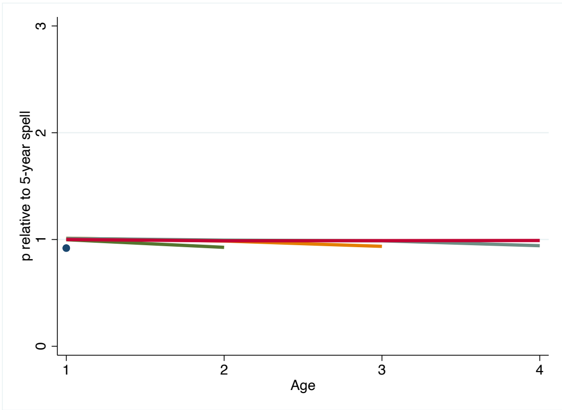
(b) Price

Notes: Based on exponentiating appropriate sums of coefficients after estimating equation 2 on data balanced at the store level. Source: Calculated based on data from The Nielsen Company (US), LLC and marketing databases provided by the Kilts Center for Marketing Data Center at The University of Chicago Booth School of Business.

Figure B6: Quantity and Price Dynamics Within Markets - Chains



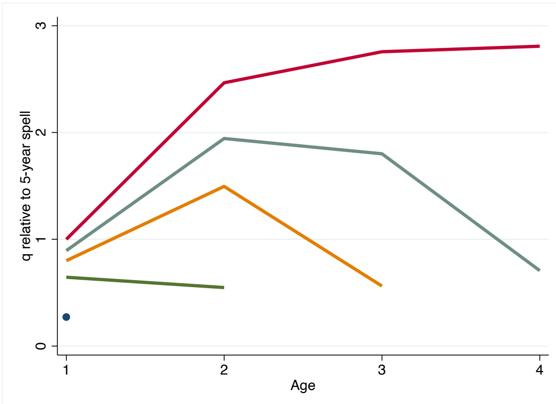
(a) Quantity



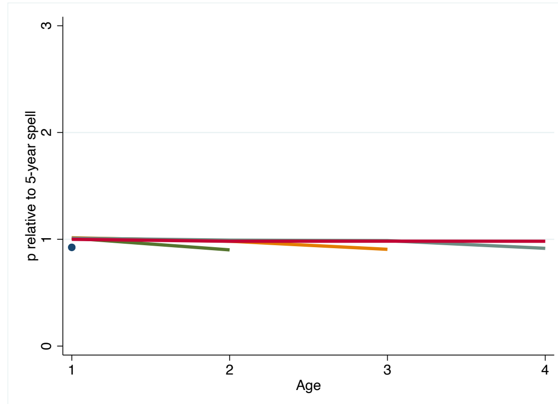
(b) Price

Notes: Based on exponentiating appropriate sums of coefficients after estimating equation 2 defining chains instead of DMAs as markets. Source: Calculated based on data from The Nielsen Company (US), LLC and marketing databases provided by the Kilts Center for Marketing Data Center at The University of Chicago Booth School of Business.

Figure B7: Quantity and Price Dynamics Within Markets - Chain × DMA



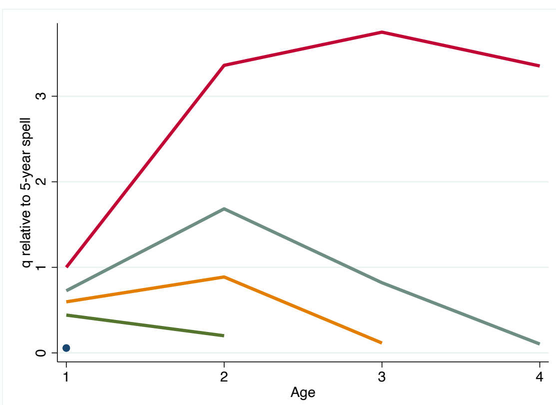
(a) Quantity



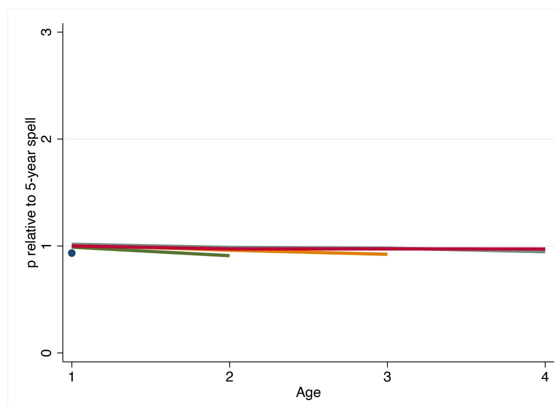
(b) Price

Notes: Based on exponentiating appropriate sums of coefficients after estimating equation 2 defining chain-DMA instead of DMAs as markets. Source: Calculated based on data from The Nielsen Company (US), LLC and marketing databases provided by the Kilts Center for Marketing Data Center at The University of Chicago Booth School of Business.

Figure B8: Quantity and Price Dynamics Within Markets - National



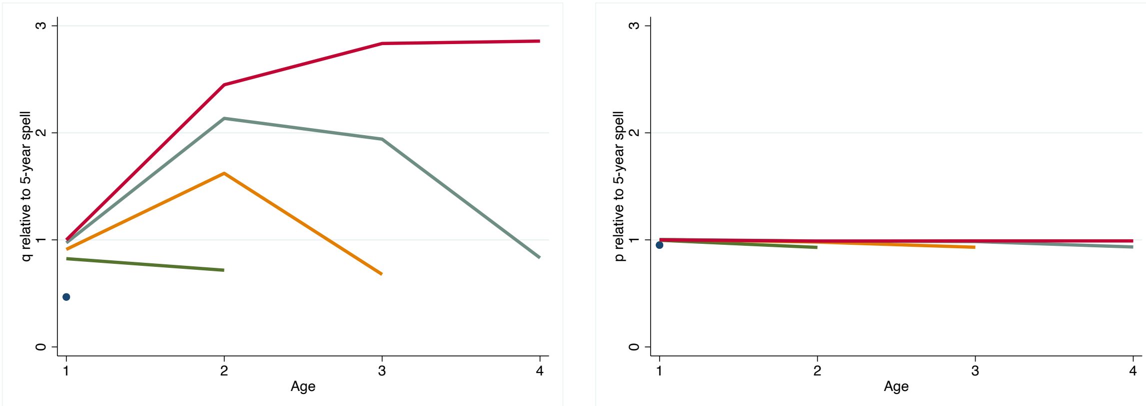
(a) Quantity



(b) Price

Notes: Based on exponentiating appropriate sums of coefficients after estimating equation 2 defining market at the national level. Source: Calculated based on data from The Nielsen Company (US), LLC and marketing databases provided by the Kilts Center for Marketing Data Center at The University of Chicago Booth School of Business.

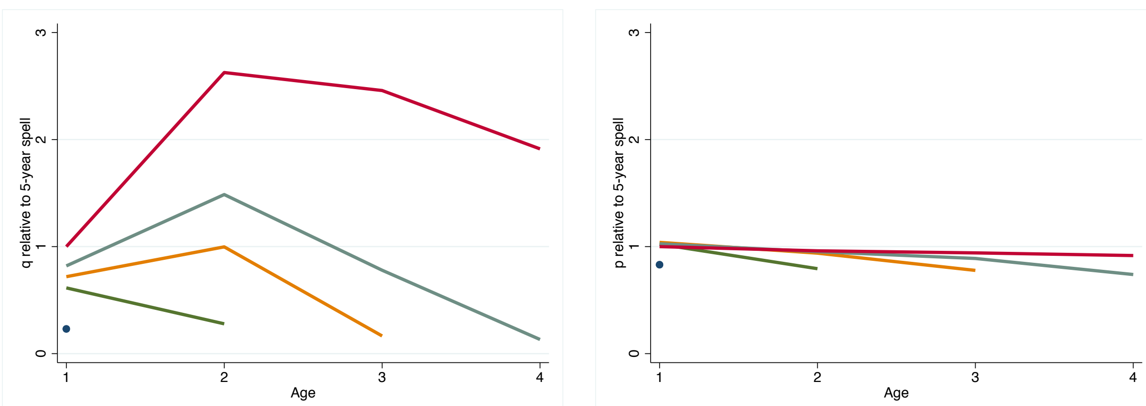
Figure B9: Quantity and Price Dynamics Within Markets - Cohort Control



(a) Quantity (b) Price

Notes: Based on exponentiating appropriate sums of coefficients after estimating equation 2 controlling for cohort effects. Source: Calculated based on data from The Nielsen Company (US), LLC and marketing databases provided by the Kilts Center for Marketing Data Center at The University of Chicago Booth School of Business.

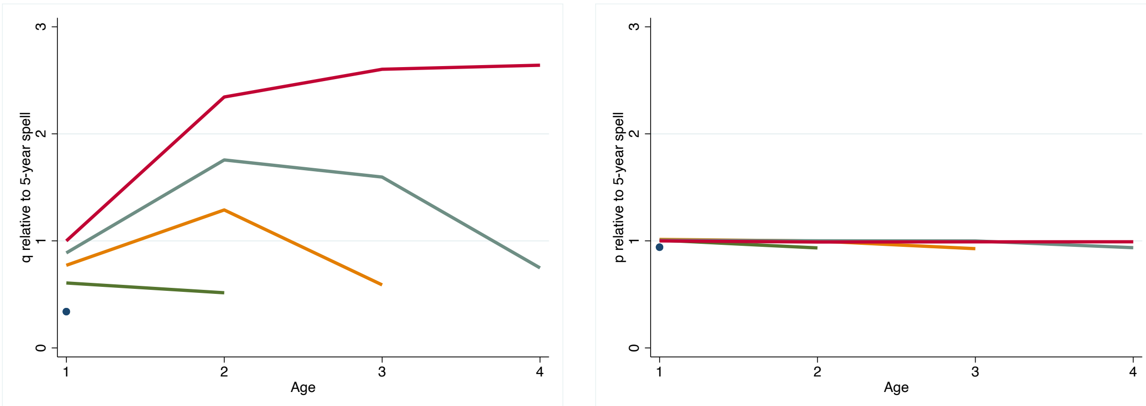
Figure B10: Quantity and Price Dynamics Within Markets - All Categories



(a) Quantity (b) Price

Notes: Based on exponentiating appropriate sums of coefficients after estimating equation 2 controlling for cohort effects. Source: Calculated based on data from The Nielsen Company (US), LLC and marketing databases provided by the Kilts Center for Marketing Data Center at The University of Chicago Booth School of Business.

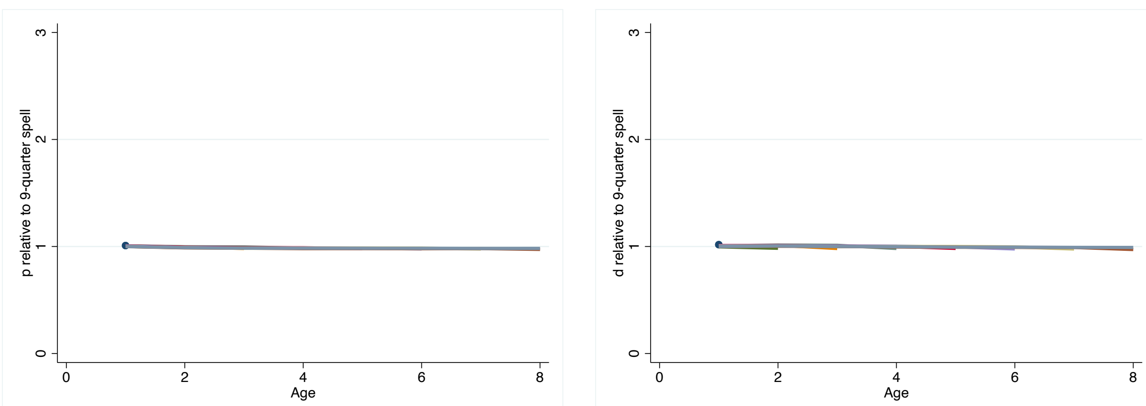
Figure B11: Quantity and Price Dynamics Within Markets - IRI Data



(a) Quantity (b) Price

Notes: Based on exponentiating appropriate sums of coefficients after estimating equation 2 using the IRI Symphony Data.

Figure B12: Wholesale Price Dynamics - PromoData



(a) Price (b) Price (Including Deals)

Notes: Based on exponentiating appropriate sums of coefficients after estimating equation 2 for wholesale prices and using the Promo Data.

B.2 Clearance Sales

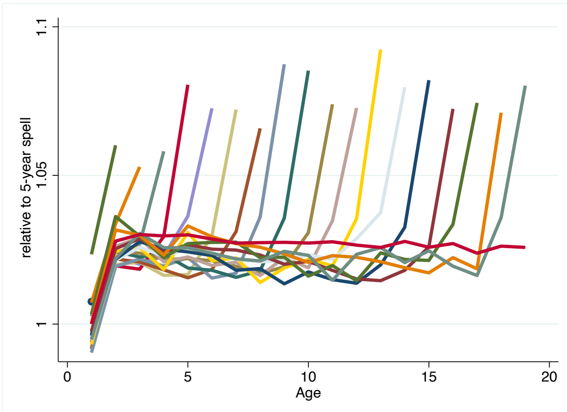
In this subsection we explore whether the decrease in price at the end of the life cycle of brands is due to clearance sales. This type of sales occur when retailers offer extra discounts on items about to permanently disappear from the shelves. The IRI Symphony data include a sales flags to detect temporary discounts.

Previous studies have use this flag to determine the prevalence of clearance sales. For instance, Gagnon, Lopez-Salido and Sockin (2015) report that, on average, the price of an exiting item is over 8 percent lower than the price that prevailed a quarter before the item's exit (that is, 14 to 26 weeks earlier), with a majority of product categories having a price drop of over 10 percent. In fact, the probability that an item is on sale in its final week in the sample is higher, at 30.5 percent, than for the typical item in the sample, at 23.4 percent, contributing to lower average prices at exit.[18]

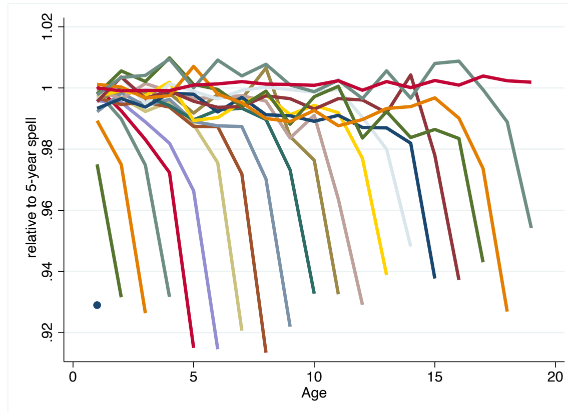
These patterns can also be seen in Figure B1, where we implement the specification in equation 2 using as dependent variables both the frequency of sales and the size of sales. Frequency of sales is measured as the fraction of barcodes-stores where the item was marked as being on sale within a brand-market in a quarter. Size of sales is measured as, conditional on being on sale, the percent deviation from the previous price. Panel (a) shows that the frequency of sales increases drastically the last quarter the brand is sold. Panel (b) shows that the size of sales also increases indicating that the reductions in prices at exit are larger than sales that take place at other stages of the life cycle. Both panels indicate the prevalence of clearance sales in the consumer goods sector.

[18] Argente and Yeh (2018) document that there is an increase in the frequency of sales and size of sales during the last weeks of the life cycle of the product defined as barcode.

Figure B1: Clearance Sales



(a) Frequency of Sales



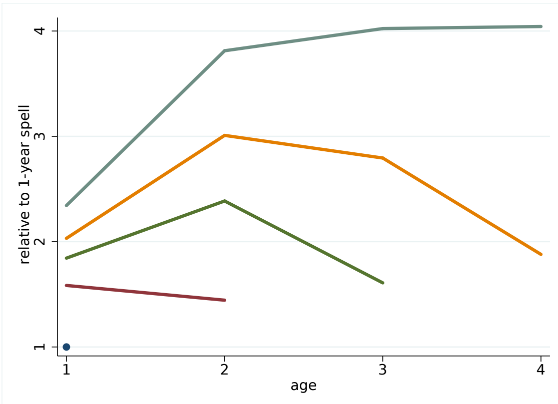
(b) Size of Sales

Note: Panel (a) shows the life cycle patterns of the frequency of sales estimated using equation 2. Frequency of sales is based on an indicator at the barcode-store level, which equals to one if the item is on sale, aggregated at the brand-market level. Panel (b) shows the life cycle patterns of the size of sales calculated as the average percent deviation from the price before the sale takes place. The data source is the IRI Symphony data.

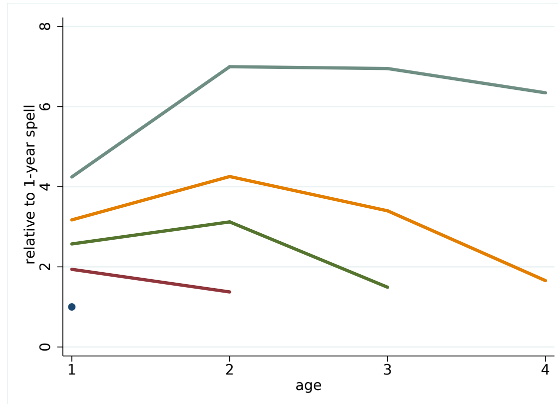
B.3 The role of advertising

B.3.1 Advertising by entrants

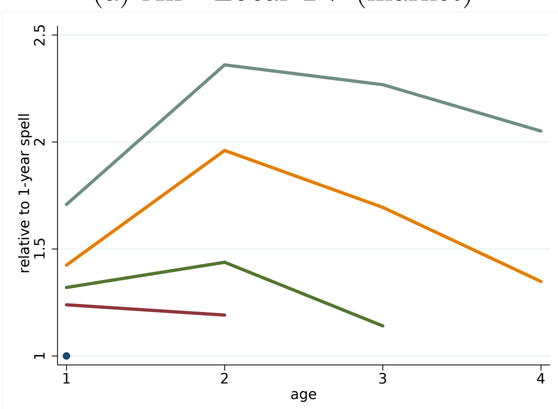
Figure B1: Dynamics of Advertising : additional results



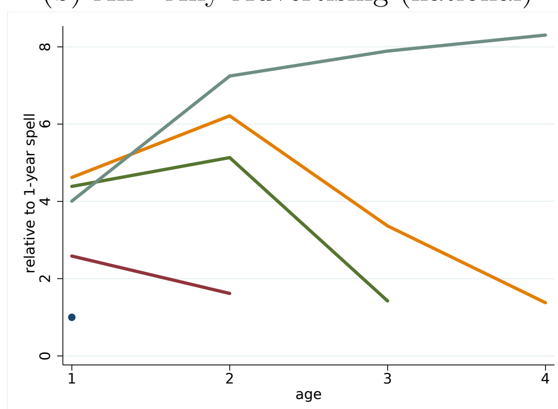
(a) All - Local TV (market)



(b) All - Any Advertising (national)



(a) Food - Local TV (market)



(b) Food - Any Advertising (national)

Notes: Source: Calculated based on data from The Nielsen Company (US), LLC and marketing databases provided by the Kilts Center for Marketing Data Center at The University of Chicago Booth School of Business.

B.3.2 Relationship between advertising and sales

Main specification We estimate the relationship between advertising and various measures of performance using regressions of the type

$$W_{t+s}^{firm^f brand^i product^m mkt^k} = \beta_s X_t^{fbmk} + \gamma^{fbmk} + \theta_t^{mk} + \varepsilon_t^{fbmk}, \qquad s = -2, ...2$$

where W_{t+s} is the outcome variable in period $t+s$ (IHS of quantity, price, sales, sales per store, number of stores, sales per barcodes, an number of barcodes), the advertising variable

Table B1: Descriptive statistics of advertising by entering brands

	Duration (years)				
	1	2	3	4	5
Entry year					
% advertising	0.004	0.012	0.022	0.040	0.125
% advertising (weighted sales)	x	x	x	x	x
# markets with advertising (average)	0.7	1.2	2.6	4.4	15.7
% markets with advertising (average)	X	X	X	X	X
Impressions (average)	X	X	X	X	X
Across all years					
% advertising	0.004	0.028	0.039	0.055	0.160
% advertising (weighted sales)	0.7	2.5	4.7	6.7	21.8
# markets with advertising (average)	12.4	15.2	20.8	27.0	50.3
% markets with advertising (average)	X	X	X	X	X
Impressions (average)	X	X	X	X	X

X_t (dummy for some advertising and IHS of number of ads or impressions). The standard errors are clustered at the level of firm $\times$ brand and product $\times$ market $\times$ year.

Alternative specification As a robustness, we also estimated the following regressions

$$\Delta W_{t+s,t-1}^{firm^f brand^i product^m mkt^k} = \beta_s X_t^{fbmk} + \alpha W_{t-1}^{fbmk} + \gamma^{fbmk} + \theta_t^{mk} + \varepsilon_t^{fbmk}, \qquad s = -2, ...2$$

where $\Delta W_{t+s,t-1}$ is the change in outcome variable in period $t+s$ relative to period $t-1$ (IHS of quantity, price, sales, sales per store, number of stores, sales per barcodes, an number of barcodes), the advertising variable X_t (dummy for some advertising and IHS of number of ads or impressions), and W_{t-1} is the level of the outcome variable in period $t-1$. The standard errors are clustered at the level of firm $\times$ brand and product $\times$ market $\times$ year.

Some Robustness Currently, our main results use annual data, cover sales and advertising in the period 2010-2016, excludes very small $firm \times brand \times product \times market$, and use agg1. We considered many robustness

- **Agg2** See below
- **Quarterly data** See below
- **HMS** See below
- **Only Long Lasting** See below

Table B2: Advertising by entering brands

	$\mathbb{1}$[local tv > 0]		$\mathbb{1}$[local tv > 0]		$\mathbb{1}$[any media > 0]		IHS(local tv imp)	
	All	Entry	All	Entry	All	Entry	All	Entry
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Entrant $\beta_{E,2}$	0.003	0.003	0.007**	0.008***	0.002	0.003	0.020	0.032
	(0.006)	(0.004)	(0.003)	(0.003)	(0.006)	(0.006)	(0.087)	(0.060)
Entrant $\beta_{E,3}$	0.011*	0.014***	0.011***	0.009***	-0.007	-0.005	0.141	0.181***
	(0.007)	(0.005)	(0.003)	(0.003)	(0.006)	(0.007)	(0.097)	(0.070)
Entrant $\beta_{E,4}$	0.033***	0.023***	0.019***	0.015***	0.015**	0.022***	0.481***	0.333***
	(0.007)	(0.006)	(0.003)	(0.003)	(0.007)	(0.007)	(0.109)	(0.086)
Entrant $\beta_{E,5}$	0.035***	0.017***	0.016***	0.014***	0.022***	0.023***	0.523***	0.253***
	(0.007)	(0.006)	(0.003)	(0.003)	(0.007)	(0.007)	(0.111)	(0.084)
Incumbent β_I	0.066***		0.039***		0.054***		1.013***	
	(0.006)		(0.002)		(0.006)		(0.096)	
Observations	5,801,851	924,856	200,900	21,796	218,997	25,881	5,801,851	924,856
R-squared	0.179	0.285	0.051	0.147	0.067	0.137	0.178	0.278
Sample	market	market	national	national	national	national	market	market
Module-mkt-t	Y	Y	-	-	-	-	Y	Y
Module-t	-	-	Y	Y	Y	Y	-	-
Uncond. $\bar{Y}_{E,1}$	0.026	0.026	0.004	0.004	0.047	0.047	0.380	0.380

Notes: The table presents estimates of advertising by type of brand using specification in equation 3. We define survival based on first year and last year in the sample. Columns (1), (3), (5) and (7) use data from years 2010-2014. We do not use data for more recent periods because of right censoring problem of brands. Columns (2), (4), (6) and (8) use data from those same years but restricting to data in the first year(s) of activity. Columns (1), (2), (7) and (8) uses variation at the brand-market-year level and includes module-market-year fixed effects, while columns (3)-(6) use data at the brand-year level and includes module-year fixed effects. The dependent variables are a dummy for some local TV advertising in that market ($\mathbb{1}$[local tv > 0]), a dummy for some local TV advertising in any market ($\mathbb{1}$[local tv > 0]), a dummy for some advertising (any media) in any market ($\mathbb{1}$[local tv > 0]), and the inverse hyperbolic sine transformation of local tv impressions in that market (IHS(local tv imp)). The last line of the table reports the unconditional average of the dependent variable for the baseline group of entrants that survive only one year. The standard errors are clustered at the level of the brand and module-module-year (or module-year in columns (3)-(6)). The ***, **, and * represent statistical significance at 1%, 5%, and 10% levels, respectively..

- **Entrants definition** See below
- **Period 2008-2016** Note presented. Because we can see sales before 2010, we can potentially use that variation for when $s = -2, -1$. The results are slightly worse when $s = -2, -1$.
- **Small observations** Not presented. Does not change much.

B.3.3 Main Annual RMS

Figure B2: Jorda Regressions in Level, Agg 1

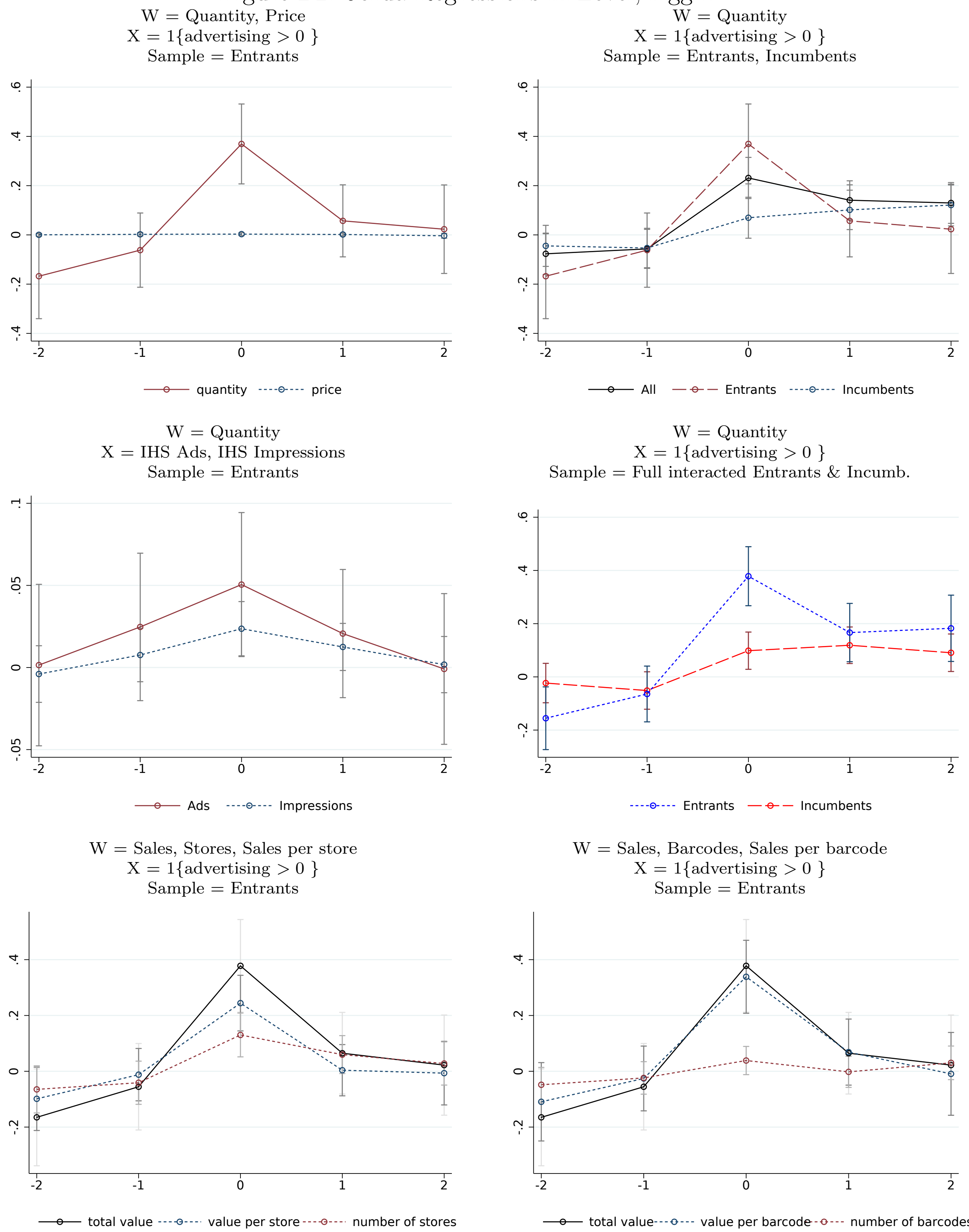


Figure B3: Jorda Regressions in Level, Agg 2

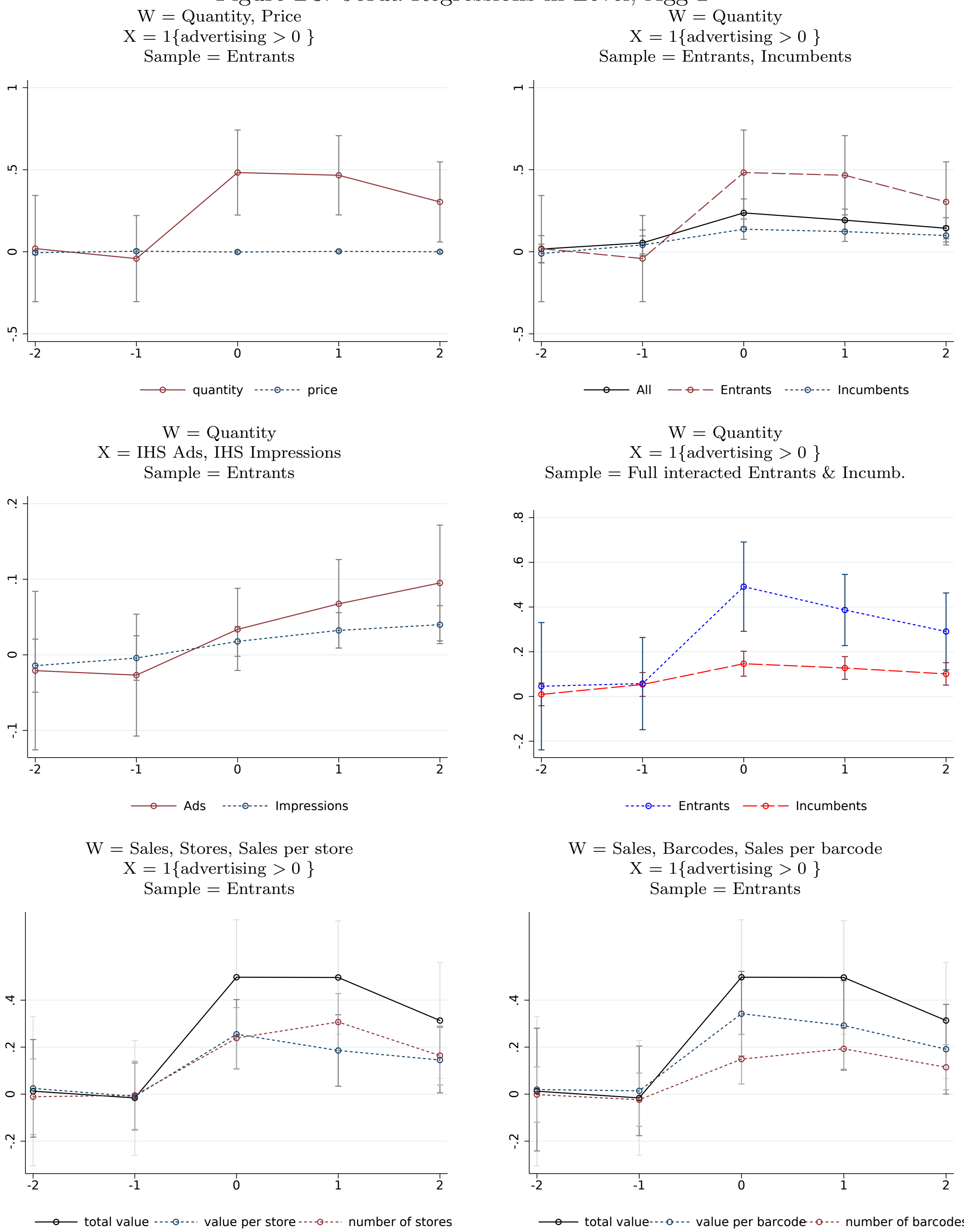


Figure B4: Jorda Regressions in Differences, Agg 1

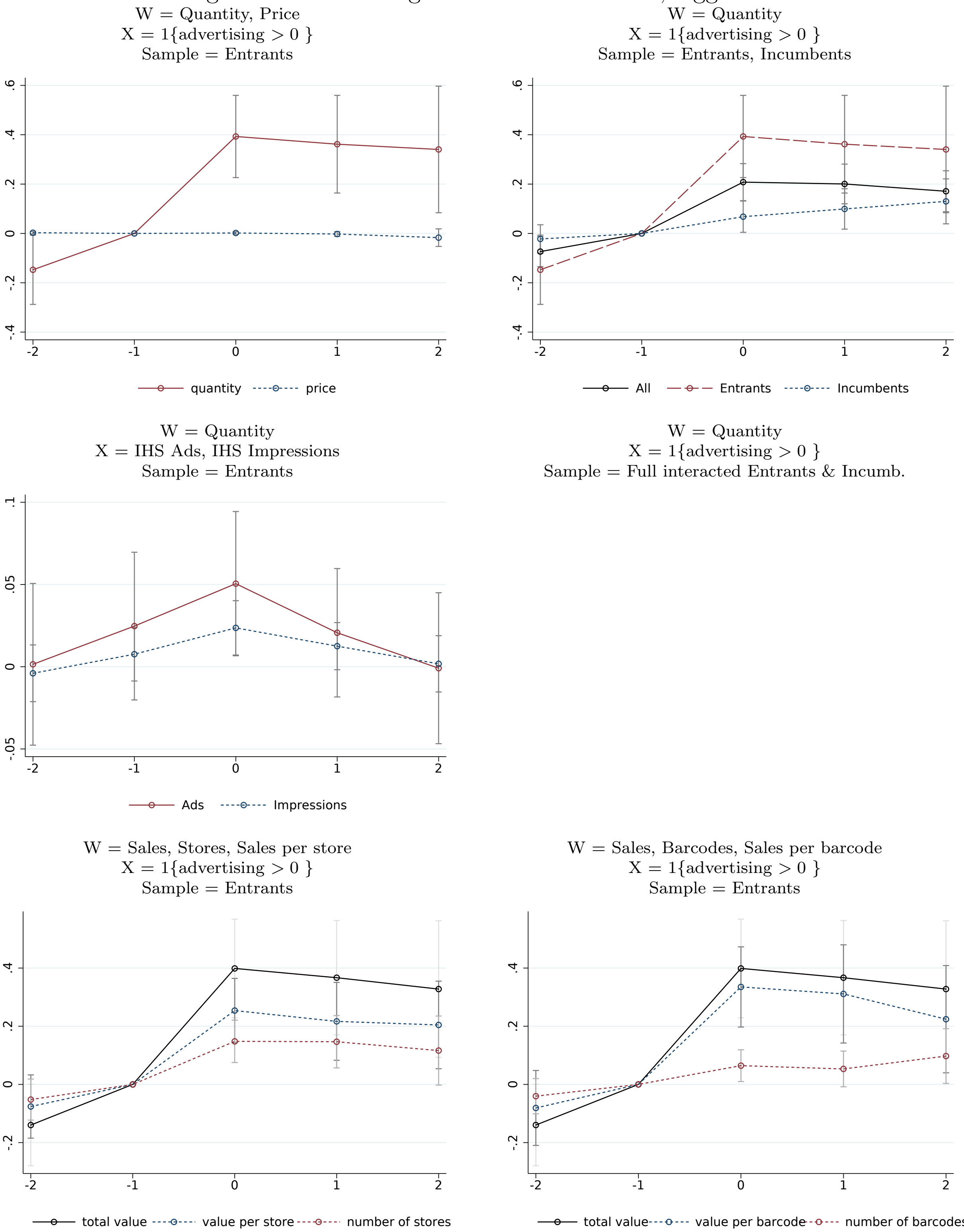
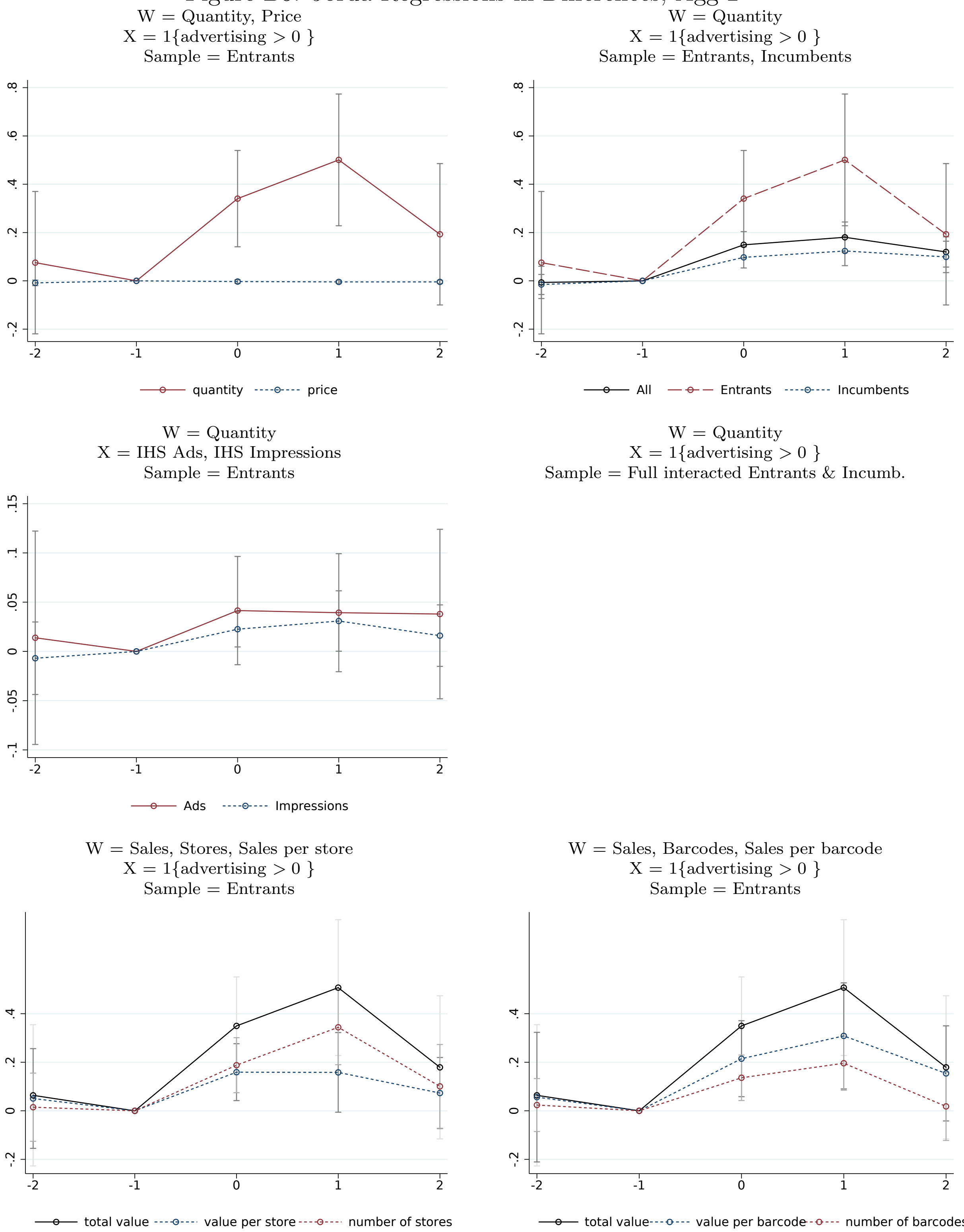


Figure B5: Jorda Regressions in Differences, Agg 2



B.3.4 Main Quarter RMS

Figure B6: Jorda Regressions in Level, Agg 1

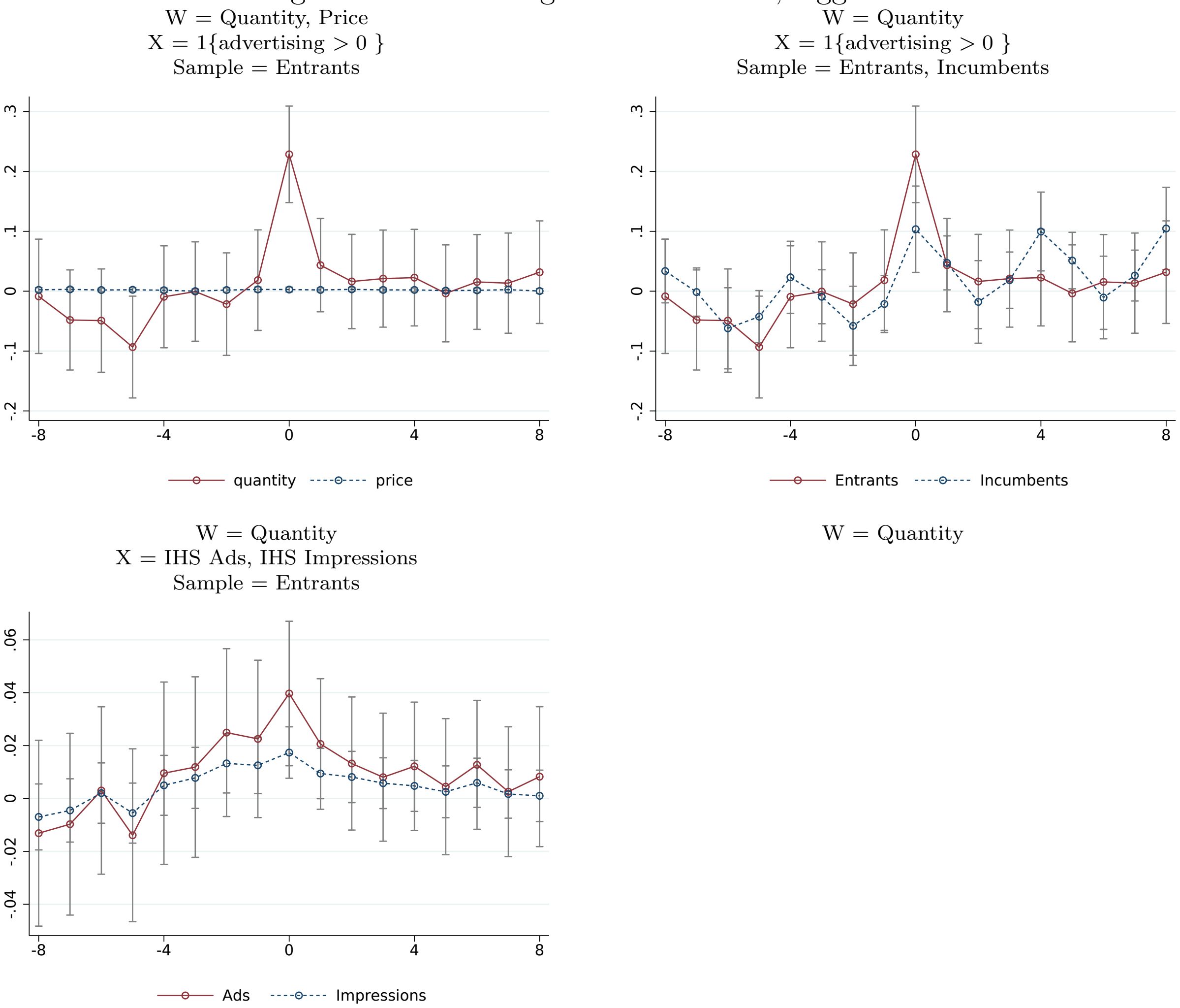
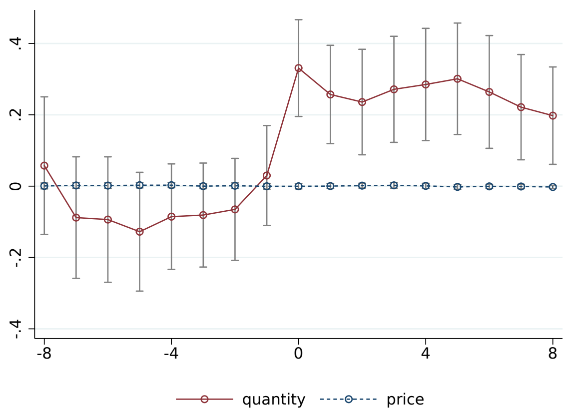
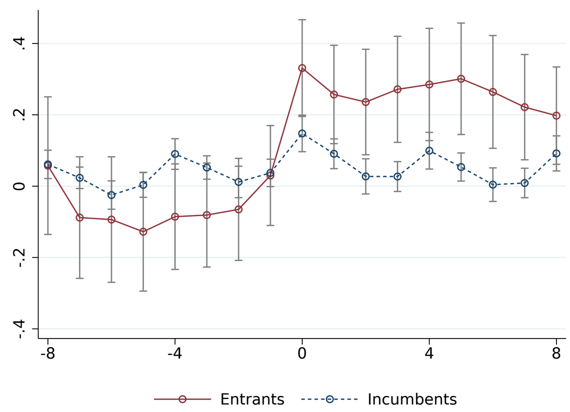


Figure B7: Jorda Regressions in Level, Agg 2

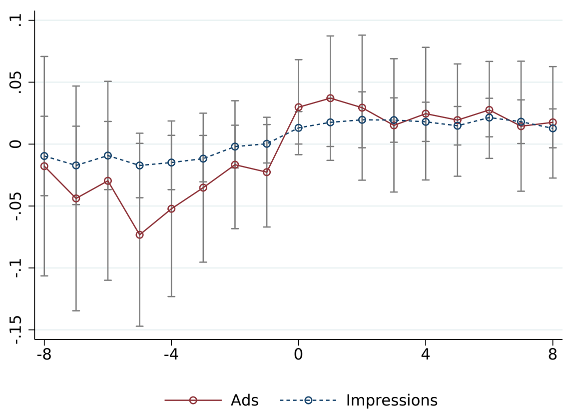
W = Quantity, Price
X = 1{advertising > 0 }
Sample = Entrants



W = Quantity
X = 1{advertising > 0 }
Sample = Entrants, Incumbents



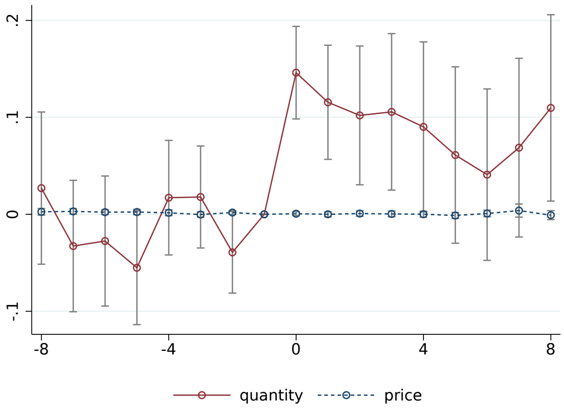
W = Quantity
X = IHS Ads, IHS Impressions
Sample = Entrants



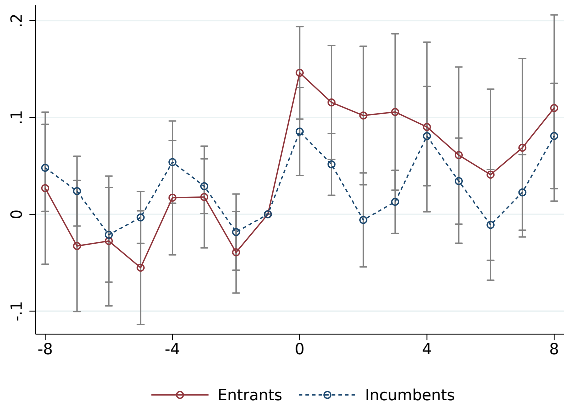
W = Quantity

Figure B8: Jorda Regressions in Diff, Agg 1

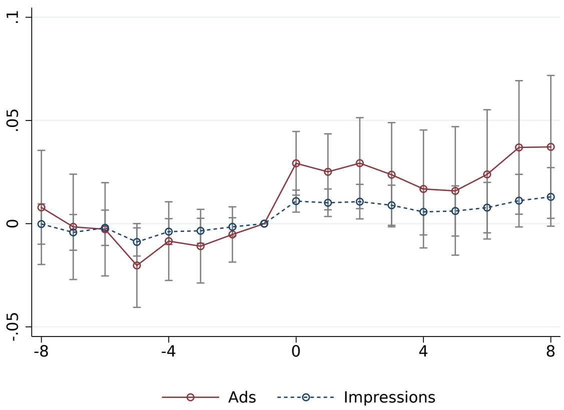
W = Quantity, Price
X = 1{advertising > 0 }
Sample = Entrants



W = Quantity
X = 1{advertising > 0 }
Sample = Entrants, Incumbents



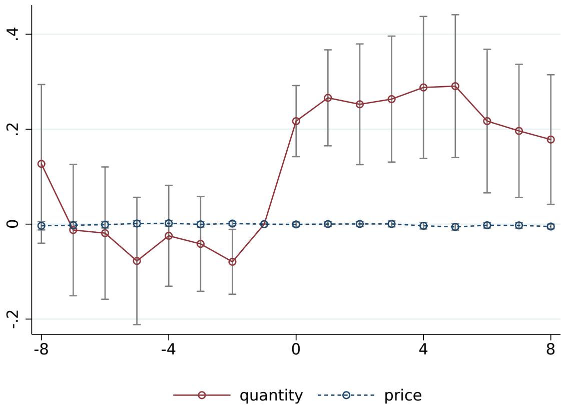
W = Quantity
X = IHS Ads, IHS Impressions
Sample = Entrants



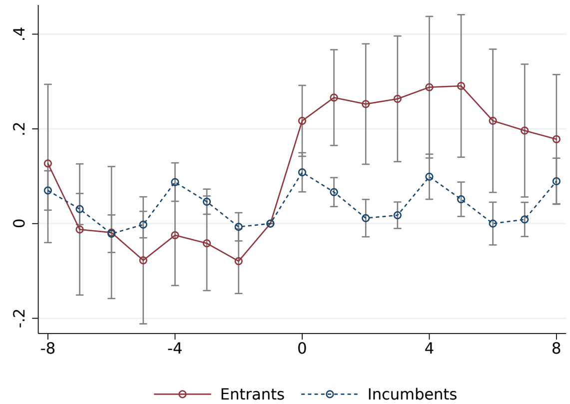
W = Quantity

Figure B9: Jorda Regressions in Diff, Agg 2

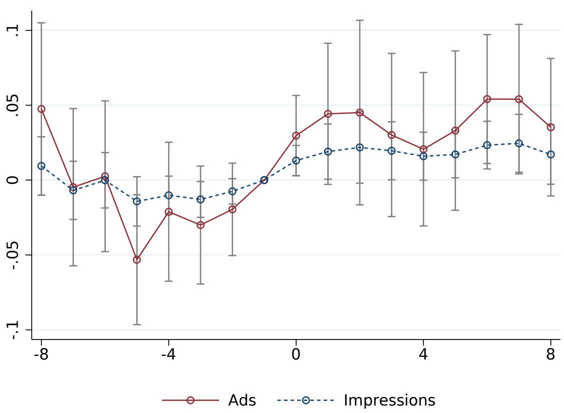
W = Quantity, Price
X = 1{advertising > 0 }
Sample = Entrants



W = Quantity
X = 1{advertising > 0 }
Sample = Entrants, Incumbents



W = Quantity
X = IHS Ads, IHS Impressions
Sample = Entrants



W = Quantity

B.3.5 Main Annual HMS

Figure B10: Jorda Regressions in Level, Agg 1

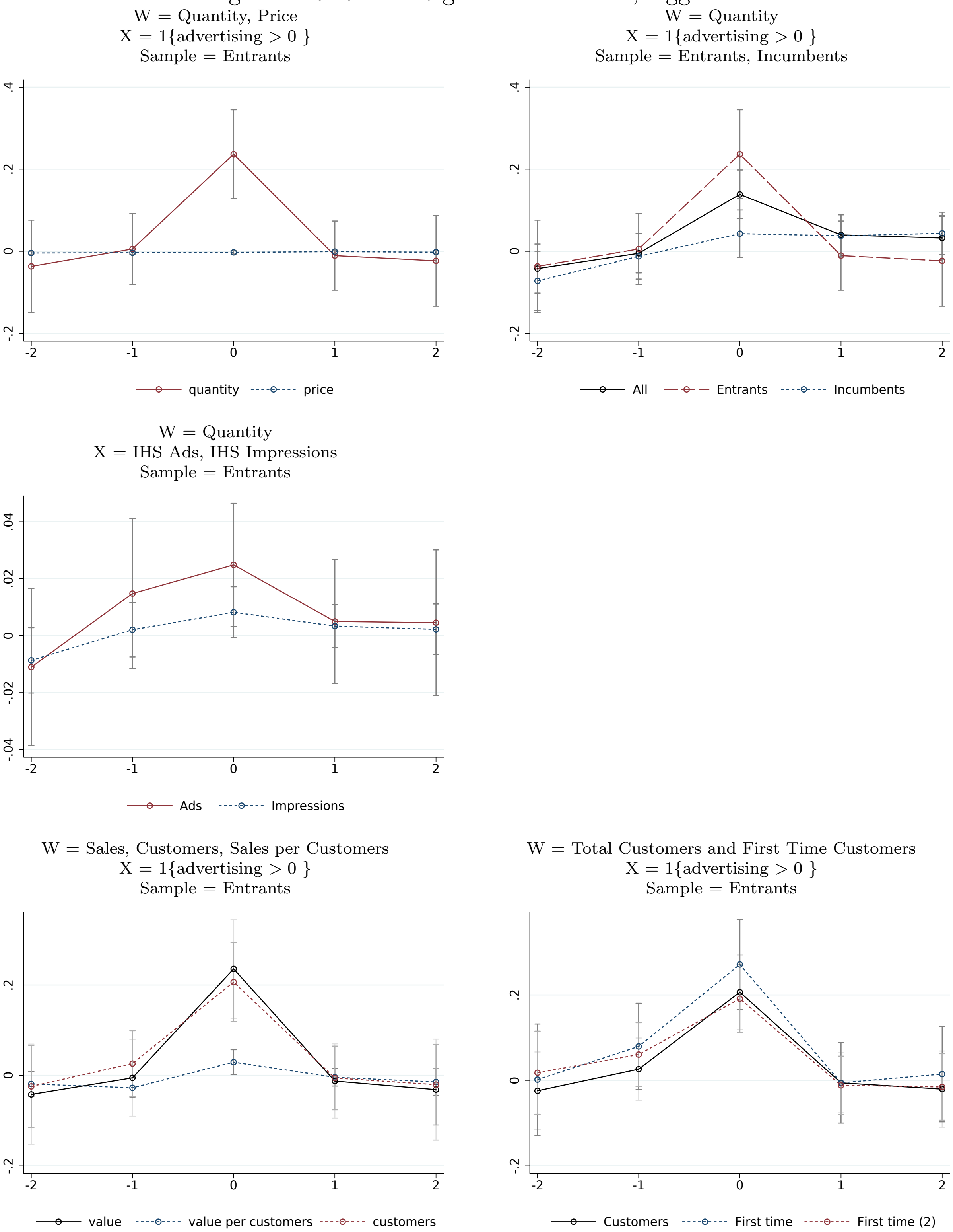
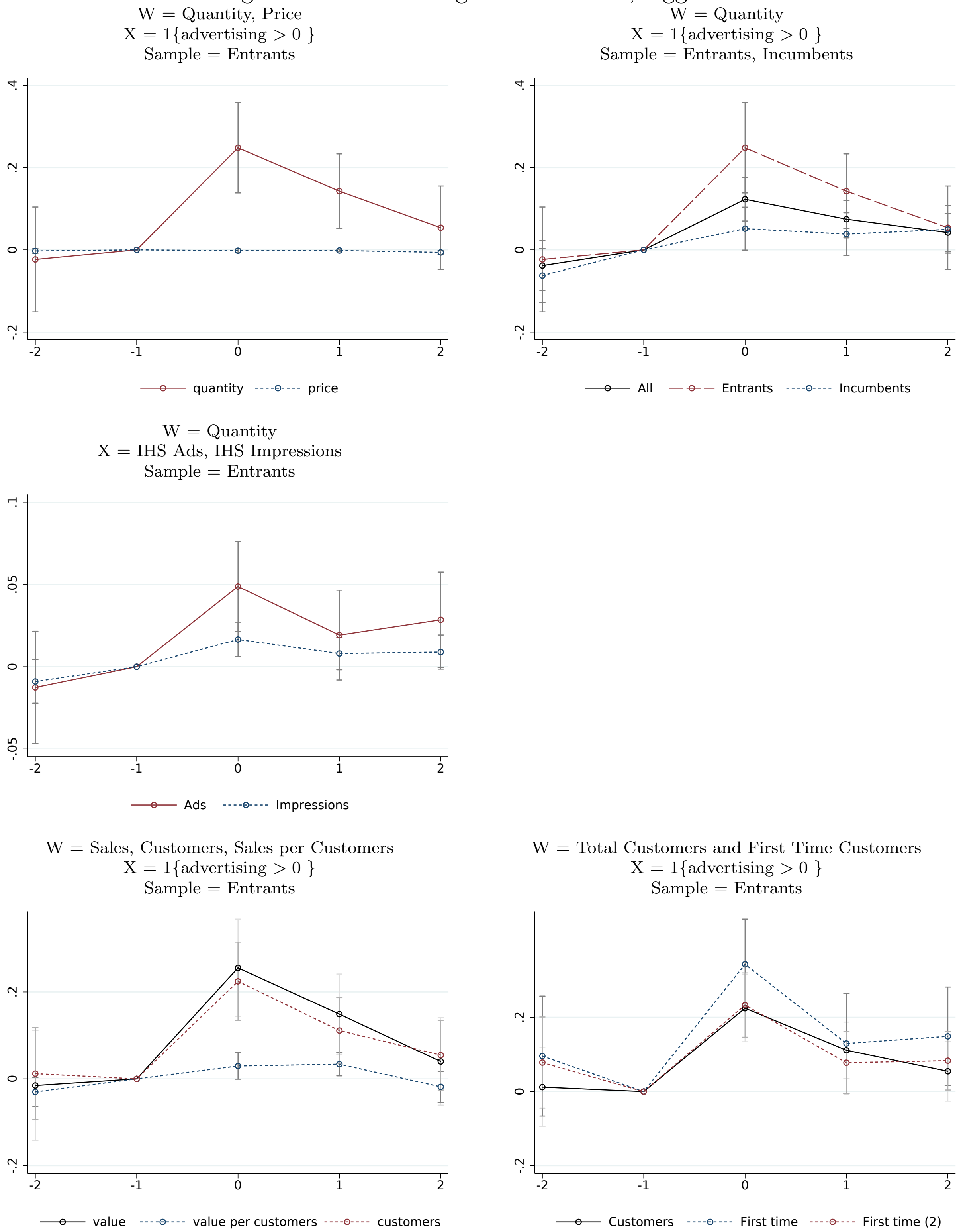


Figure B11: Jorda Regressions in Diff, Agg 1



B.3.6 Robustness - RMS, only long lasting (survival of at least 5 years)

Figure B12: Jorda Regressions in Level, Agg 1

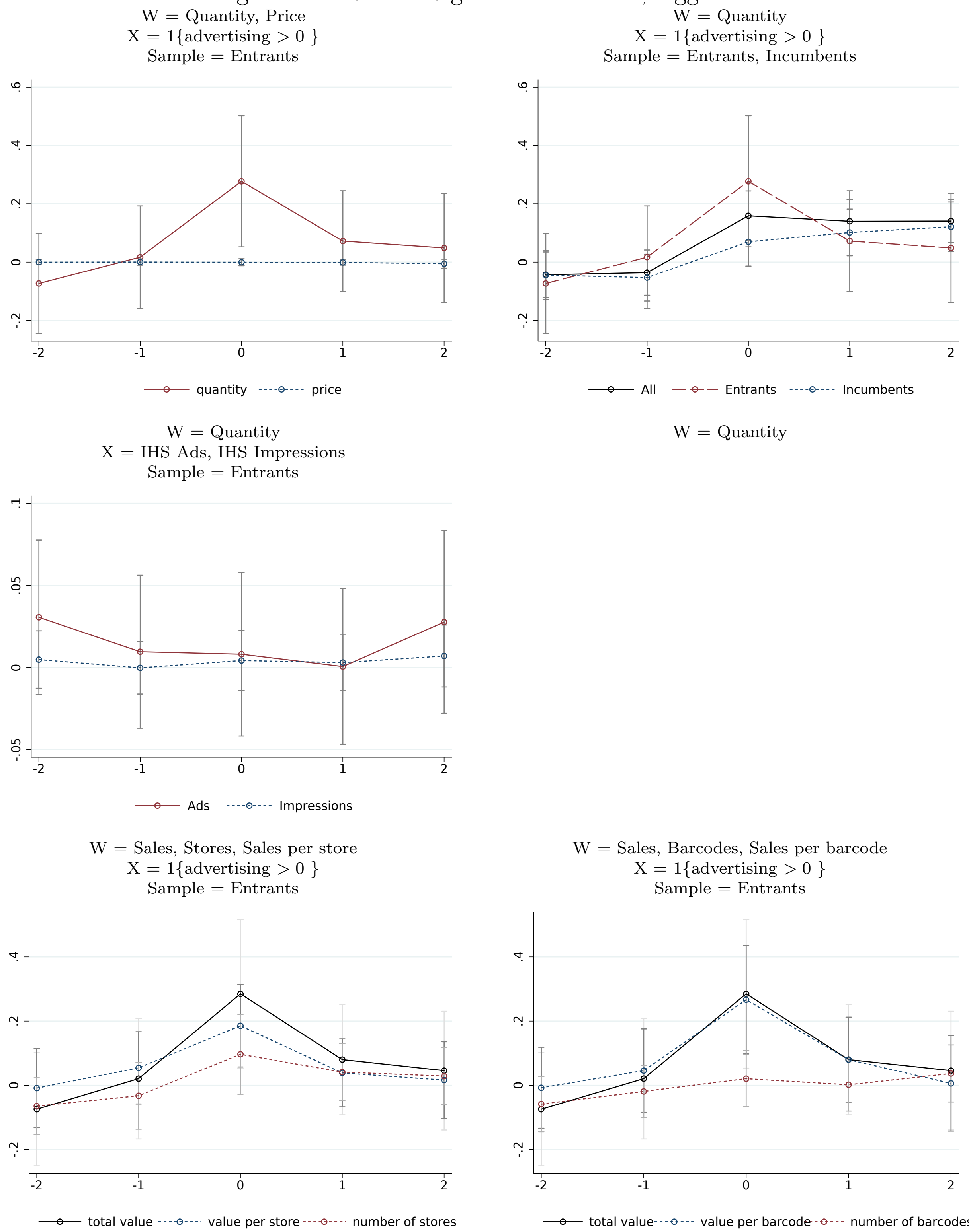
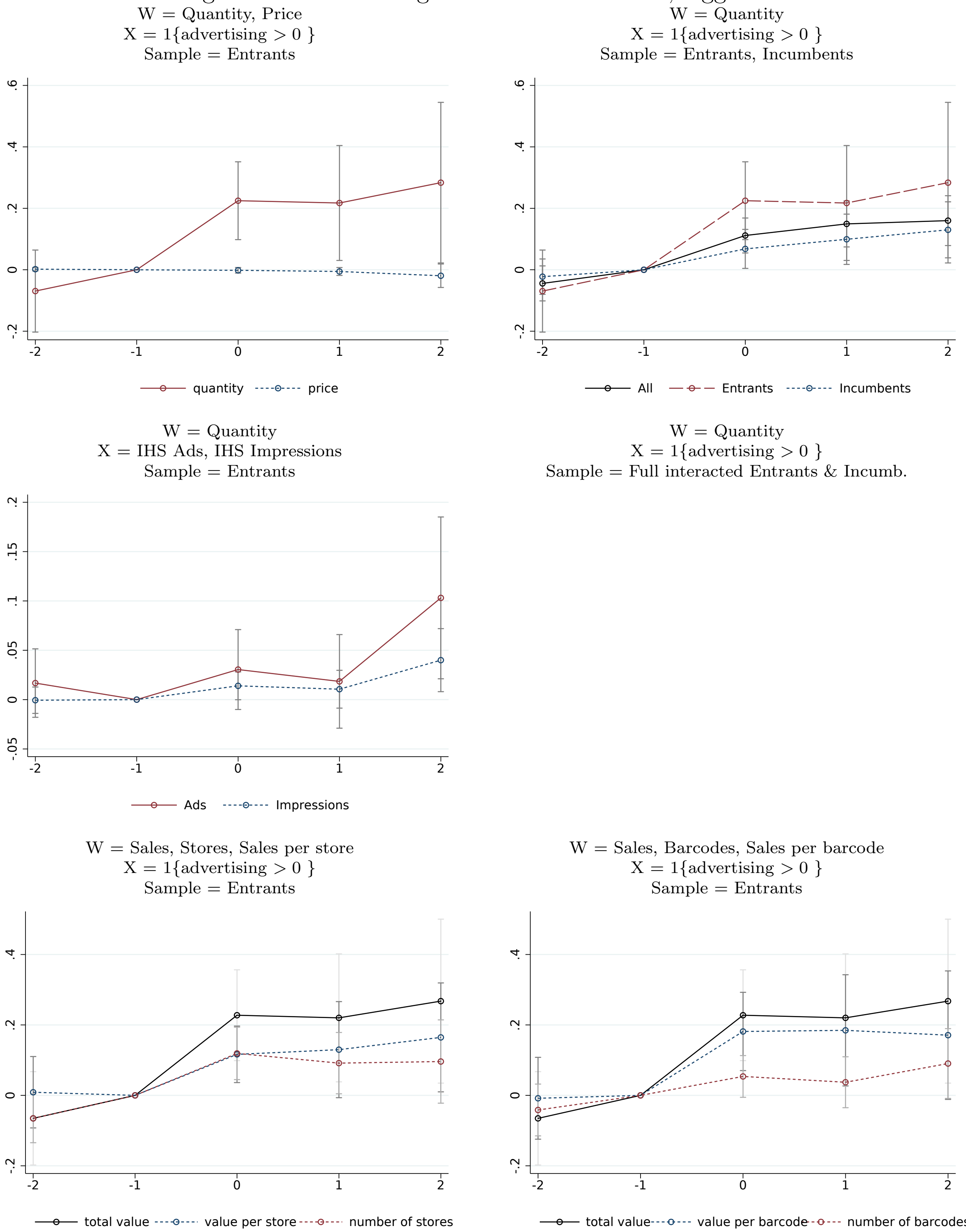


Figure B13: Jorda Regressions in Differences, Agg 1



B.3.7 Robustness - RMS, Alternative definition of Entrants (born after 2009)

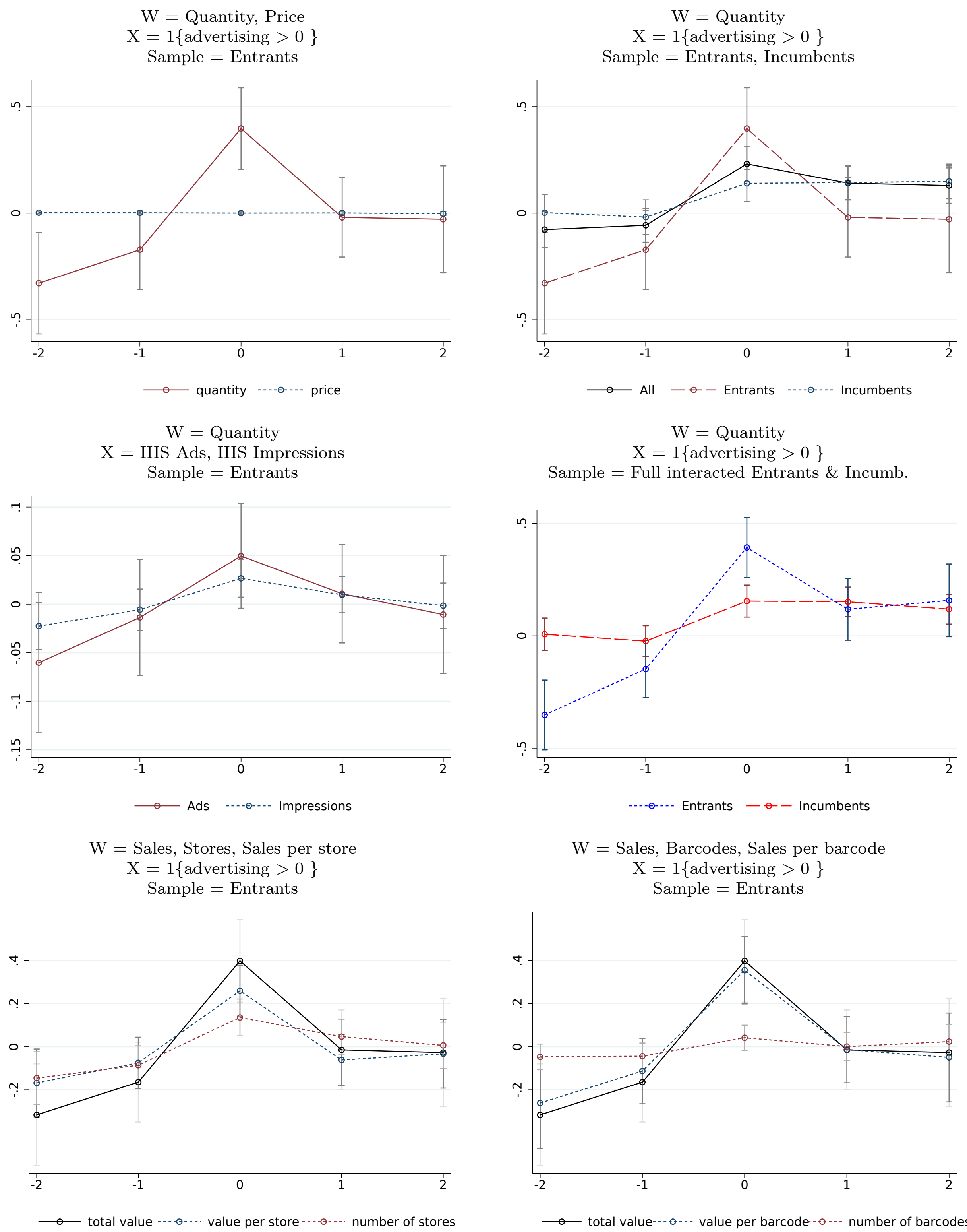
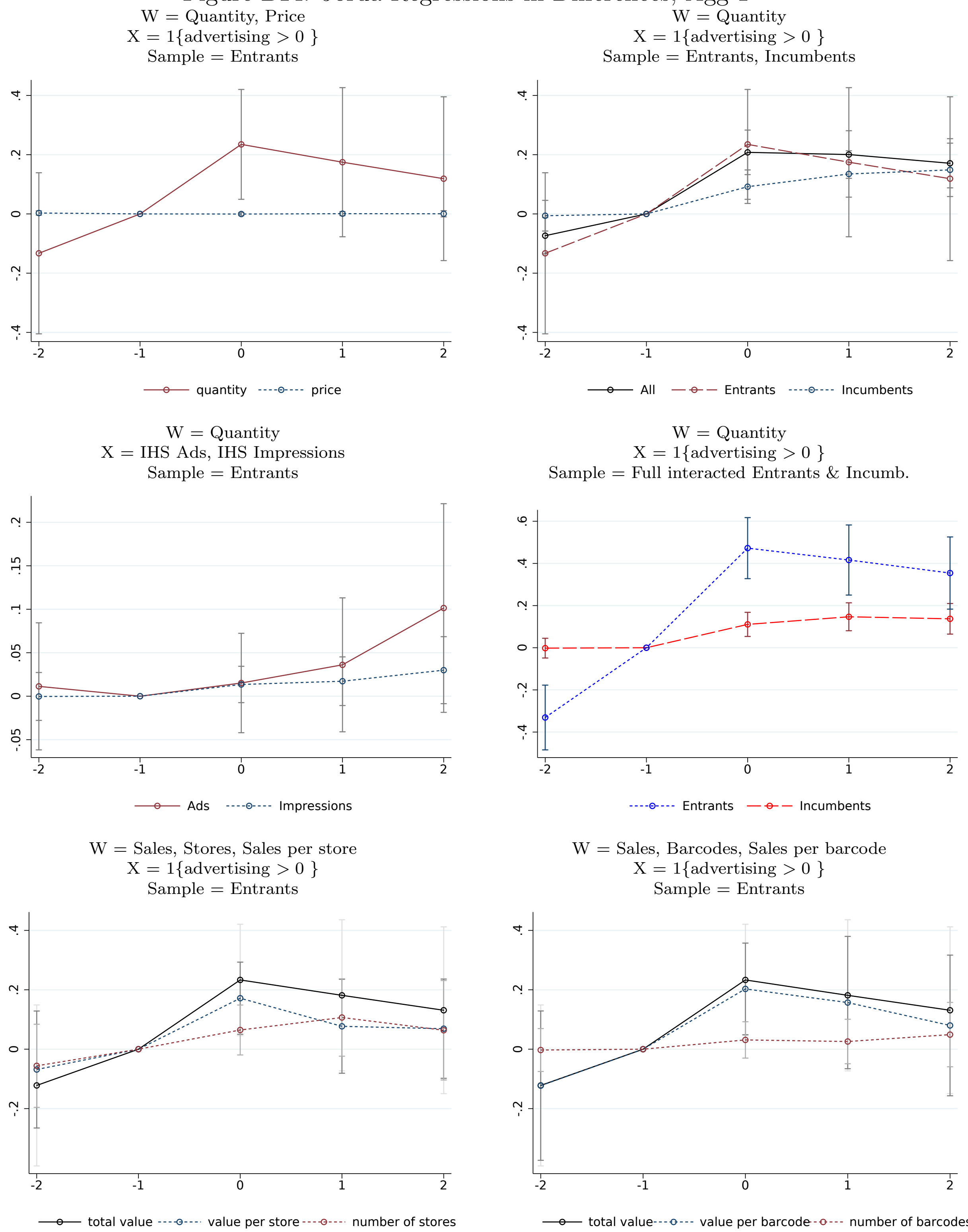


Figure B14: Jorda Regressions in Differences, Agg 1



C Elasticity of Substitution

In this section, we describe the procedure we use to estimate the elasticity of substitution θ. Specifically, we exploit the rich geographical variation available in our data. We estimate the following equation:

$$\Delta \log(S_{ijkt}) = \theta \Delta \log(P_{ijkt}) + \kappa_{ijt} + \lambda_{jkt} + \epsilon_{ijkt} \tag{7}$$

where i is a brand, j is product module, m is a DMA, and t is a quarter or year. S_{ijkt} refers to the share of sales of brand i in a given product-market-time period, θ is the elasticity of substitution, and P_{ijkt} is the product's unit price when sold by brand i. The terms κ_{ijt} and λ_{jkt} capture several fixed effects specifications.

Due to the likely simultaneity between the demand shocks in ϵ_{ijkt} and the products' prices, we require an additional identifying assumption to recover θ. Following beraja2019aggregate and fally2017firm, we use a Bartik-style instrument. The leave-out instrument is the average across markets excluding one product-market-time element: $\frac{1}{N-1}\sum_{j\neq m} \Delta \log(P_{ijkt})$. The instrument is used in a first-stage regression for the log change in unit price, $\Delta \log(P_{ijkt})$. The identifying assumption is that consumer taste shocks are idiosyncratic across DMAs whereas supply-side cost shocks are correlated across space.

We implement this strategy using quarterly data in order to have more variation. We use two types of instruments a national and state-level leave-out. We estimate θ under several restrictions which include: i) a balanced sample of observations over 48 quarters at the product-brand-market level, ii) considering only national brands (those present in at least 196 markets), and iii) considering markets that sell at least 9 brands within a product module. In all our specifications we obtain an estimate for θ between 1.5 and 2.1. When we use this procedure to estimate θ for each department separately, we find that "Dry Grocery" has the lowest elasticity of substitution (i.e. between 1.5 and 1.8) and "Packaged Meat" has the highest elasticity of substitution (i.e. between 1.8 and 2.1).