Measuring Consumer Switching to a New Brand across Local Markets*

Paulo Albuquerque† Bart J. Bronnenberg‡

November 2, 2006

*We appreciate comments from Dan Ackerberg, Andrew Ainslie, Anand Bodapati, Charles Corbett, Jean-Pierre Dubé, Mike Duffy, Carl Mela, Sanjog Mishra, and Garrett Sonnier. We also acknowledge comments made by seminar participants at Dartmouth College, Duke University, Singapore Management University, University of Rochester, University of Science and Technology Hong Kong, University of Texas at Austin, Washington University St. Louis. We thank IRI and AC Nielsen for their generous contributions to the data sets used in this paper.

†Assistant Professor at the Simon Graduate Business School, University of Rochester, paulo.albuquerque@simon.rochester.edu. Financial support from the Portuguese Foundation for Science and Technology is gratefully acknowledged.

‡Professor at the UCLA Anderson School of Management, 110 Westwood Plaza, Ste. B-4.18, Los Angeles, CA 90095, bart.bronnenberg@anderson.ucla.edu

Measuring Consumer Switching to a New Brand across Local Markets

Abstract

We estimate local demand for new brands of consumer packaged goods (CPG), identify its underlying sources, and explore cross-market patterns in substitution to the new product using aggregate level data. We show that the standard choice-based demand model overestimates the degree of consumer switching to a new brand and underestimates consumer heterogeneity in a market. We introduce a model and an estimation procedure to avoid these problems. The approach augments the market-level time series with widely available summaries of household switching behavior. This data also has the benefit that it is informative about the size of the outside good. The latter allows us to estimate the size of the total market rather than assume it. Empirically, we use data from the Frozen Pizza Category to estimate our model. The data cover the launch of a new national brand, DiGiorno. We find that this new brand was very successful at targeting consumers from outside of the category and that cannibalization of existing brand share was largely avoided. We also find that substitution patterns and perceptions of brand similarity are different across markets. We explore the cross-market variation in switching patterns and find that local pre-entry share, and local branding explain a large part of how much an incumbent looses to the new entrant. In contrast, we find only very small national brand effects on switching to the new brand. Therefore, the origins of demand for the new brand are not common to markets.

Keywords: new products, perceptual maps, random coeffiicients logit model, national and local branding

1 Introduction

New products are pervasive in consumer packaged goods (CPG) industries (e.g., Kahn and McAlister 1997; MacLeod 1995). AC Nielsen estimates that there were more than 30,000 new products introduced in CPG industries in 2001 alone (AC Nielsen 2001). Whereas the majority of these introductions concerns me-too products, many mark the launch of true innovations. The prevalence of new product launches is also increasing over time (AC Nielsen 2001; Kahn and McAlister 1997). Finally, among several Western economies, the US market has the highest rate of the introductions of truly new products. Therefore, managerially, there is little disagreement about the long-term importance of successful new product introductions, especially in mature CPG industries that are often characterized by eroding profit-margins and escalating use of price-promotions. To estimate demand for a new brand, especially in an existing category, a manager needs to know with which existing brand(s) it substitutes, in each of several separate local markets.[1] In accordance with this, the aim of this paper is (1) to empirically evaluate local demand for a new CPG brand, and (2) to explore the geographic differences in new product substitution patterns.

In the marketing literature, brand switching has been studied mainly in the context of the firm's use of price and promotion instruments (e.g., Blattberg and Wisniewski 1989; Carpenter et al. 1988, Van Heerde, Gupta and Wittink 2003; Van Oest and Franses 2005). Yet surprisingly little is known empirically about consumer switching due to the firm's use of product innovations, i.e., a new CPG brand. Even less is known about how a new product substitutes differently with incumbent brands across local markets. We aim to fill this gap in the literature by studying substitution patterns across local markets using a combination of standard aggregate market level data sets that are widely available to managers. Thus, our focus is on the less-studied but strategically important case of brand switching as a consumer response to product innovation in mature CPG categories.

First, we are interested in the relative importance of competitive draw, cannibalization, and category expansion[2] in a local market that is due to the launch of the new CPG brand. A manager or other interested party generally needs to use aggregate level data for this task. The aggregate nature of the data presents a potential problem in identifying the sources of new product demand and in identifying the taste heterogeneity in the market. One contribution of our paper is that we

[1] With this terminology we mean markets that are demand-separated, i.e., where markets are far enough apart that consumers do not travel between them to benefit from lower prices.

[2] Competitive draw is defined as the fraction of demand for a new product that is caused by consumers switching away from competing brands. Cannibalization is defined as the fraction of demand that comes from consumers switching from the other brands marketed by the new brand's manufacturer. Finally, category expansion is defined as the fraction of demand for a new product that originates from consumers who bought other –indirect– substitutes before.

show how the integration of several widely available data sets solves this problem. Our random coefficients logit demand model identifies the sources of switching to a new brand and at the same time accounts for (1) temporary changes in other marketing activities such as price changes or promotional activities, (2) the heterogeneous preferences of consumers, (3) the endogenous nature of prices, and (4) the unobserved nature of the size of the outside good.

Second, whereas we have data covering the entire domestic industry, identification of the local demand system for new products permits the exploration of cross-market differences in demand. Therefore, we first estimate the local substitution patterns separately for each market and next relate those patterns to descriptors of national, regional and local marketing activity. A description of substitution patterns across markets may teach us about the long term impact of national and local shares of incumbents, of regional variables, and of marketing strategies prior to launch. This information can for instance be useful during national roll-outs to make predictions about switching patterns in unentered markets. The multimarket aspect of our paper is non-trivial. In most industries, market shares and switching patterns observed in a specific market can not be generalized to the entire population. Consumer perceptions, consumer characteristics, brand history, and firm marketing behavior all vary significantly across local U.S. markets (Bronnenberg, Dhar and Dubé, 2007).

Although our method uses market-level aggregate data, we are able to identify several aspects of individual level switching patterns. The intuition for our identification strategy and how it differs from other approaches is best explained imagining the existence in each market of an N (consumers) by T (time periods /markets/products) complete panel data set. One common approach to demand estimation (e.g., Nevo 2000) is to use aggregates across the first dimension of N consumers, i.e., use the $[T \times 1]$ market level cross-section or time series of shares to infer the mean and the dispersion (heterogeneity) in consumer tastes for a product and its characteristics. This method works well to identify mean preferences, but we show that it has considerable difficulty in recovering the dispersion of preferences across the consumer population (see also Bodapati and Gupta, 2004; Petrin, 2002). The method is generally also uninformative about the size of the outside good, for which therefore separate identification assumptions need to be made. Past studies warn that inferences about substitution and switching behavior are highly dependent on correct estimation of preference heterogeneity (Berry, Levinsohn, and Pakes 2004) and the size of the outside good (Nevo 2000).

We aim to redress these problems by adding distributions of consumer behavior (aggregates

across the T time periods) to better estimate the dispersion of preferences and the size of the outside good. Specifically, we add to our estimation information about the purchase set size. The purchase set is defined as the set of unique brands among which a consumer switches during a 12 month time span. Whereas we have no individual level data about the purchase set size S_i for a consumer i, data are readily available about the consumer distribution of the annual purchase set size, i.e., about $\Pr(S_i = 1)$, $\Pr(S_i = 2)$, etc.[3] We next impose that the logit demand system matches the correct distribution of purchase set sizes across consumers. This information on purchase set sizes is very informative about the dispersion of consumer preferences, because similar preferences across brands imply large purchase sets, while very dispersed preferences imply smaller purchase sets. Similarly, we can estimate the size of the outside good by matching the predicted probability that the consumers are in the market with readily available data on local penetration rates.

We estimate our model using data from the Frozen Pizza category and focus on the national launch of DiGiorno. Our findings are as follows. Methodologically, we show that adding data about the local category penetration to the local time series of market shares allows us to empirically identify the size of the outside good, and that adding data about the purchase set size distribution in the market helps us estimate taste variation in the market (brand and price heterogeneity). We show that ignoring this information leads to incorrect inferences about brand switching. A model without the additional information greatly overestimates the amount of switching that occurs in the market.

Substantively, we find that the new premium-priced DiGiorno brand was perceived in most markets to be closest of all brands to the outside good (e.g., delivery pizza). Accordingly, we infer (and model-free evidence supports) that Kraft was very successful at attracting new consumers from outside of the frozen pizza category with typically more than half of its local share originating from category expansion. We also find that competitive draw is the next largest source of sales for DiGiorno. Surprisingly, cannibalization of Kraft's incumbent brands was minor, even in markets where Kraft already had strong market share. Explaining switching behavior across brands and markets, we find that brand name and marketing mix variables, such as price and display, explain very little of the cross-market variation, with local aspects such as pre-entry share appearing as far more important. We this find that the sources of demand for the new brand are not common to markets.

[3] Both IRI and Nielsen have individual level panels that, while too small to cover all local markets exhaustively, are successfully used to compute marginals of consumer behavior such as category penetration, usage intensity, etc. These data are available as a standardized service in CPG industries.

In the next section, the paper discusses our approach in more detail. Section 3 presents our demand model. We describe both the Frozen Pizza industry and the data used to estimate the model in section 4. Section 5 describes the estimation algorithm and section 6 discusses the results. We conclude in section 7.

2 Our approach

To identify the demand for new products empirically, at least two different types of information can be used: individual-level household data and market-level (or sometimes store-level) aggregate data. Historically, household panel data have been used to study individual choice history and brand switching, while aggregate level data has been usually applied to monitor category and brand sales evolution, evaluate market shares and quantify the impact of marketing actions. Both types of data have their own advantages and disadvantages (Bodapati and Gupta 2004). Although the analysis in this paper could be done using individual level data, such information is at present either very sparse or simply unavailable to practitioners and analysts. Therefore, we are confined to using market-level data about sales and marketing strategies in CPG categories.

Two economic issues have become increasingly important when estimating demand models: endogeneity (usually of prices) and consumer heterogeneity (see, e.g., Chintagunta 2001). A feasible approach to solving the endogeneity problem has been suggested by Berry, Levinsohn and Pakes (1995) who propose an algorithm to identify the unobserved demand shocks that are taken into account by manufacturers when setting price. This in turn enables the use of instrumental variables estimators in determining consumer price effects.

Identifying heterogeneity in consumer tastes using aggregate level data is a harder problem to solve. According to Bodapati and Gupta (2004), heterogeneity is, in practice, hard to estimate because the only available information identifying it is the result of discrepancies in observed shares movements and the expected movements predicted by a homogeneous model. In cases where these discrepancies are small, heterogeneity parameters will only be weakly identified, if at all. Petrin (2002) points out the same predicament, defending that heterogeneity is identified only if some unusual substitution patterns not captured by the homogeneous model do occur and/or a change in the choice set is present, e.g., the introduction of a new brand. We show that even when such a condition is met, heterogeneity in preferences (or at least the rate of brand switching) is still difficult to estimate empirically.

In sum, past models which define utility at the individual level but use solely aggregate level data

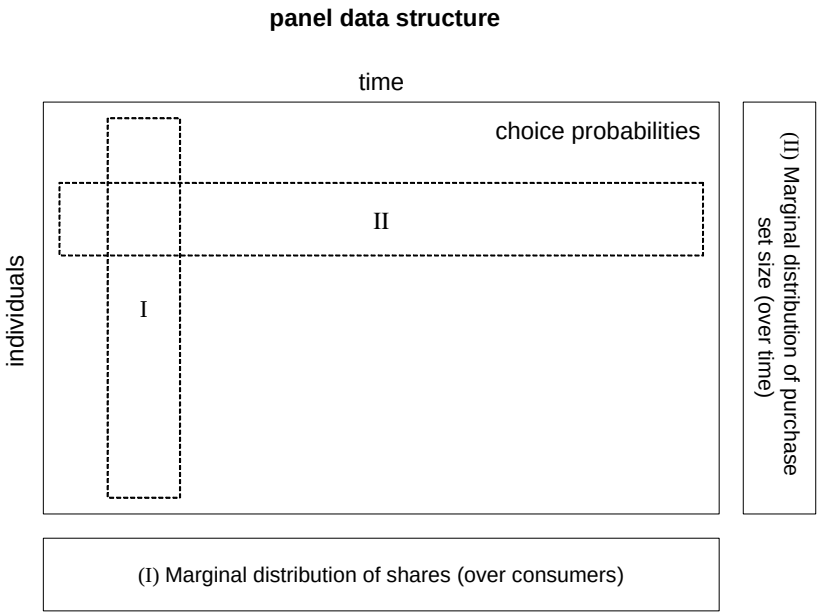

Figure 1: Two dimensions of data used in estimation

for estimation build their identification strategies on the fact that consumers leave a trace of their preferences and actions on the observed aggregate shares. The strength and level of identification of these individual signals using just aggregate level data has generated some debate in the past (Berry, Levinsohn and Pakes 2004; Nevo 2000; Petrin 2002). Our approach is to use market-level data about shares and the marketing mix, which is more accessible and less sparse, while simultaneously offsetting some of the shortcomings of these data with syndicated data about consumer behavior that validate observed patterns at the aggregate level. This additional information is also aggregate level information but it is aggregated in a different direction. Figure 1 illustrates what we mean by this. In an ideal situation, we would have the full –N consumers by T time periods– panel data of adoption behavior for the new brand. But because this information is not generally available or too expensive, we propose to use readily available information from the marginal summaries of both dimensions of the panel data (e.g., aggregates across time and consumers) rather than just one (the time series).[4] Using standard sampling theory, these summaries or aggregates are still highly accurate even if the panel data are sparse.

The combination of multiple sources of information to improve a demand model's accuracy is

[4] We note that the panel may include other dimensions such as markets (e.g., Nevo 2001), or a wide variety of products (e.g., Berry, Levinsohn, and Pakes 1995), in addition to or in lieu of the time dimension but the same argument about collapsing the consumer dimension applies in these cases.

not new in the context of durable goods (e.g., Berry, Levinsohn, and Pakes 2004; Petrin 2002). We add to this literature in several ways. First, we add information on consumer differences in purchase behavior instead of demographic consumer characteristics given purchase (Petrin 2002) to the estimation. This is particularly suited in a CPG context where such information is both easy to obtain and reliable. This information is useful to identify demand heterogeneity. Second, switching that occurs because of the introduction of a new brand is more influenced by the strength and dispersion of brand preferences. The dispersion of brand preferences therefore established a "base rate" of brand switching in a category that will likely be informative about the switching that should occur to a new brand. Finally, our approach also differs from past work, in that we show how information on distributions of purchase behavior (including statistical support for "no purchase" behavior) allows us to estimate the size of the outside good.

With respect to the exact choice of purchase information, we now present a straightforward example to motivate that the consumers' distribution of the purchase set size is uniquely qualified to identify taste variation. Suppose first that there is a market where all consumers have similar preferences for two hypothetical choice options. In this setting, consumers are practically indifferent between the two brands. Therefore, choice probabilities are very similar, close to 0.5 for each brand. The relative indifference causes consumers to frequently switch randomly between the two brands across time, possibly induced by promotion and price. As a result, over a time span that covers multiple purchases, the purchase set is expected to include both available brands. In a second market, half of the consumers prefer one brand, and half prefer the other. In the second case, consumers have strong preferences for one of the brands and a very low degree of switching. Over time, the purchase set - different brands bought - is likely to be composed of exclusively one of the brands.

Obviously, in practice we expect price variation to add variance to market shares over time, but these two scenarios illustrate the fact that aggregate level information about market shares is not enough to identify consumer heterogeneity and switching patterns. In both cases, we obtain market shares of 50% for each brand. However, consumer preferences and brand switching are dissimilar across the two cases.

Our approach involves adding the marginal distribution (over time) of the purchase set size. In our example, this information would adopt the following form. In the first scenario discussed above, the marginal distribution of the purchase set size would be 0% of individuals purchase one brand, whereas 100% of consumers buy two different brands time. For the second scenario, the

distribution is 100% of individuals buy one brand, 0% of individuals buy two brands. Thus, the two scenarios of preference dispersion across brands produce very distinct distributions of the purchase set size.

In addition to using information about the number of unique brands that a consumer switches among, we use the fraction of households who do not buy any brand from the category to help us estimate the outside good. This is defined as the fraction of households with a purchase set size of 0, or alternatively by 1 minus the penetration rate. Our estimation procedure imposes on the demand model that its predictions about the popularity of the outside good are consistent with the empirical data on penetration rates and probability that a household has purchase set size of 0 in the market.

3 Model

Since we wish to analyze the differences in substitution behavior across markets, our demand model is formulated at the market level. To avoid cluttered notation, we however suppress the market subscript. The model is the same in each market, but the parameter values and the data are market specific.

In each week $t = 1, ..., T$, the utility of brand $j = 0, ..., J$ for consumer $i = 1, ..., N$ is given by the following expression:

$$u_{ijt} = \alpha_{ij} + x_{jt}\beta_i + \xi_{jt} + \epsilon_{ijt}, \tag{1}$$

where α_{ij} is individual i's perception of brand j, x_{jt} is a K-dimensional row vector of observed marketing mix variables, β_i is a K-dimensional vector of individual specific marketing mix coefficients and ξ_{jt} includes demand shocks that are unobserved by the econometrician but considered by consumers in their purchase decisions and by manufacturers in their pricing decisions. Finally, ϵ_{ijt} is an i.i.d. mean-zero stochastic term with a Type-I Extreme Value distribution.

Consumers are allowed to be heterogenous in their preferences for brands and in their sensitivities to marketing mix variables. The vector of brand intercepts α_i has the following expression:

$$\alpha_i \sim N(\bar{\alpha}, \Sigma) \tag{2}$$

$\bar{\alpha}$ is a $[J \times 1]$ parameter vector that represents the mean-preferences for the brands. Both parameters $\bar{\alpha}$ and matrix Σ of size $[J \times J]$ can be estimated directly or can be represented using a factor structure (e.g., Chintagunta, Dubé, and Singh, 2002).

We chose to model random intercepts α_{ij} using a factor model for several reasons. First, the factor model produces a brand map, which plots brands that are perceived to be similar by consumers, close to each other. It also plots brands that are dissimilar in the eyes of consumers, far apart. A factor model is thus helpful in directly estimating brand similarity (e.g., Elrod 1988; Elrod and Keane 1995; Erdem 1996). Second, brand maps are of interest in the current paper, especially in the context of DiGiorno's advertising claim that it substitutes with delivery pizza as evidenced by the slogan "It's not delivery, it's DiGiorno!". From this claim, we could expect that DiGiorno substitutes with the outside good, a fact that is directly verifiable from how close DiGiorno is positioned to the outside good in the brand map. Finally, the factor model introduces correlation in the unobservable brand characteristics across brands, with few parameters. It thus presents the advantage of reducing the number of parameters required to estimate a full (in the sense of non-diagonal) heterogeneity matrix while remaining highly flexible.[5]

Using a two dimensional factor structure, the factor representation of household specific intercepts is

$$\alpha_{ij} = L_j \omega_i, \text{ with } z_i \sim N(w, I) \tag{3}$$

with L_j the coordinates of alternative j in the attribute space and ω_i the consumer's taste vector for these attributes. ω_i is assumed to be distributed Normal with mean w and variance covariance matrix I. These assumptions imply that $\bar{\alpha}$ in equation (2) is equal to Lw and that Σ is equal to LL'.[6]

The introduction of the mean brand effects, $\bar{\alpha} = Lw$, is akin to introducing brand dummies into the model. There are several reasons to include such brand level fixed effects. First, it is not certain that observed product characteristics capture all or much of the substitution patterns in the data. In such cases, "fixed effects should be included to improve the fit of the model" (Nevo 2000). Second, we are interested in the random shocks ξ_{jt} because they are economically related to prices. However, by accounting for the alternative specific mean utility components, we have also accounted for possible correlation between prices and the brand specific mean of unobserved quality. In turn, this has the advantage that we do not need an instrument for this correlation. Last,

[5] In contrast, while the random intercept model can be specified with a full variance covariance matrix, in practice it often is constrained it to be diagonal, especially with the use of aggregate data.

[6] The factor model cannot be estimated withouth several identification restrictions. Specifically, to cancel translation invariance, we fix the outside good to be placed in the origin of the attribute space. To cancel rotation invariance, we fix one alternative to be positioned on the horizontal axis. Finally, to cancel reflection invariance, we have set the mean preference for the first attribute to be positive. This model is still too flexible for our purpose. To be able to compare the positions of brands in the estimated brand maps L across markets, we have set w equal to a vector of ones. This is an overidentification, but it is practically necessary and does not empirically restrict the model as will become clear from the results later on.

accounting for the mean alternative specific utility means that the interpretation of the random shocks ξ_{jt} becomes more precise. Specifically, with the mean utility accounted for, the random shocks ξ_{jt} are zero in expectation at the brand level and represent seasonal variability in demand (e.g. calendar seasons, or special events such as Superbowl Sunday, etc.). It is likely that pricing depends on such seasonality in a different way than it does on brand-differences in unobserved attributes. Thus accounting for brand level mean utilities in each market, we can disentangle these two sources of endogeneity which otherwise would be left confounded.

Heterogeneity in response to marketing mix variables is modeled in the following way,

$$\beta_i = \beta_0 + \sigma' v_i, v_i \sim N(0, I) \tag{4}$$

In practice, consumers can choose among several choice options, the "inside goods," or decide to buy something else (including "nothing") in a given week, the so-called "outside good." Its utility, u_{i0t}, is fixed at zero for identification purposes. Given the extreme value distribution of ϵ_{ijt}, the probability of household i purchasing brand j at time t is given by:

$$\Pr(j_t \mid x_{jt}, i) = \frac{\exp\left(L_j \omega_i + x_{jt}\beta_i + \xi_{jt}\right)}{1 + \sum_{k=1}^{J} \exp\left(L_k \omega_i + x_{kt}\beta_i + \xi_{kt}\right)} \tag{5}$$

To obtain market level shares s_j, we integrate over the distribution of the characteristics of the individuals i (in our case, tastes for attributes and sensitivity to marketing mix effects):

$$\begin{aligned} \widehat{s}_{jt} &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \Pr(j_t \mid x_{jt}, i)\, \phi(v)\, \phi(\omega)\, \partial v \partial \omega \\ i &= 1, ..., I, \qquad j = 0, ..., J, \qquad t = 1, ..., T \end{aligned} \tag{6}$$

where $\phi(.)$ is the PDF of the normal distribution. With this set up, the parameters to be estimated are $\theta = [\beta_0, L, \sigma]$. The integral shown in equation 6 lacks an analytical form. We use simulation methods to estimate the shares and parameters, described in more detail in the estimation section.

4 Data

4.1 The Frozen Pizza industry

Our empirical analysis covers the Frozen Pizza category. Frozen pizza has become one of the most important categories among frozen food, accounting for about 19% of the its sales (Bronnenberg and Mela 2004; Van Heerde et al. 2004). According to industry experts and manufacturers, it represents almost 20% of the total pizza business, with delivery pizza being its main competitor outside of the category (*Pizza Marketing Quarterly*). During 1993-1995, the years preceding our

	1995	1996	1997	1998	1999
DiGiorno	0.01	0.04	0.10	0.13	0.13
Jack's	0.09	0.09	0.07	0.07	0.08
Red Baron	0.12	0.12	0.11	0.12	0.13
Tombstone	0.22	0.21	0.19	0.18	0.17
Tony's	0.12	0.12	0.10	0.09	0.08
Totinos	0.11	0.11	0.11	0.11	0.10
Fringe	0.33	0.31	0.31	0.30	0.31

Table 1: Evolution of average shares for some brands in the frozen pizza category in each of the years in the data set.

analysis, the category was viewed as largely stable and characterized by slow growth, with dollar sales marginally increasing from \$1.6 to \$1.7 billion. In 1995, Kraft launched a new brand into the market, DiGiorno. In late 1996, Schwan's followed by launching Freschetta. Both brands included a new feature, self-rising crust, which was considered a major development in the category. Combined with strong advertising, DiGiorno's introduction led to a fast increase in sales of frozen pizza with a sustained annual growth rate of approximately 12% through 1999 (Holcomb, 2000).

Two companies, Kraft and Schwan's Food Company, have a dominant place in the Frozen Pizza category, with a portfolio of multiple brands. Kraft's brands include DiGiorno, Tombstone and Jack's while Schwan's owns Tony's and Red Baron. Another national brand in this category is Totino's which is owned by Pillsbury. Our analysis of the introduction of DiGiorno will focus mainly on these six brands, which capture about 70% of the volume of the category.[7] All of them, except Jack's, are available nationally. Jack's distribution is limited to markets in the North-West and Mid-West region of the country but has a large share in those markets.

To avoid cannibalization of its existing brands, Tombstone and Jack's, Kraft exploited DiGiorno's rising crust attribute in its marketing. Specifically, because rising crust was associated with fresh baked or hand-tossed pizza, Kraft positioned DiGiorno as a close competitor of delivery pizza instead of traditional frozen pizza. This is most directly evidenced by the slogan "It's not delivery, it's DiGiorno!"

Average annual shares of brands from 1995 to 1999 are presented in Table 1. Nationally, the dynamics in DiGiorno's share reflect a roll-out that took three years. By 1999, DiGiorno had captured about 13% of the U.S. frozen pizza market.[8]

The launch of DiGiorno affected the incumbent brands differently in different markets. In Figure 2, we show the evolution of market shares in Pittsburgh and Oklahoma City, around the time of

[7] Freschetta is not included in our analysis of the introduction of DiGiorno, because it was introduced later.
[8] The same figure is independently reported in Holcomb (2000).

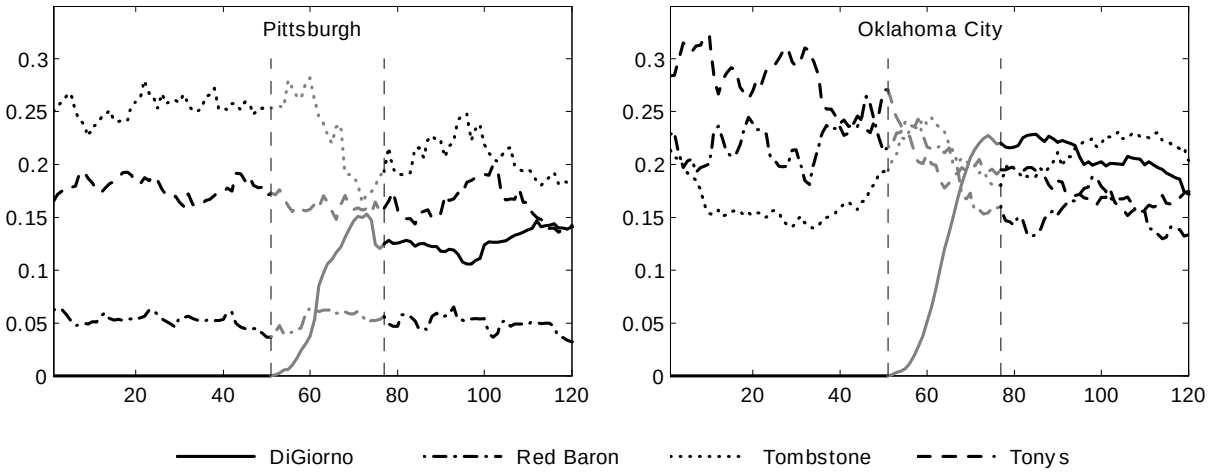

Figure 2: Evolution of market shares of the main brands in two markets at the time of DiGiorno's introduction

DiGiorno's introduction. In Pittsburgh, DiGiorno steals most share from Tombstone whereas the share of Red Barron is not at all affected by the launch of DiGiorno. Oklahoma City presents almost the opposite situation. Here, it is Red Barron who suffers the most severe share loss, but Tombstone's share remains unaffected. These differences suggest that substitution patterns are market specific. One expectation about local substitution patterns is that differences across markets arise because the incumbents' shares are different and new brands draw their share proportionally from those incumbents. But this proportional draw hypothesis does not accord well with the data in the Figure. For instance, in Oklahoma City, the share erosion for the incumbent brands is so unevenly concentrated in the Red Barron and Tony's brands that Tombstone, which was smaller prior to launch of DiGiorno, actually becomes larger than Red Barron and Tony's after it. These patterns are suggestive of more complex local substitution patterns.

4.2 Different data sources

Our analysis integrates four different data sets. The first data set contains market-level information on sales volume, price, local feature advertising and display levels. The data are indexed by week and market and cover 260 weeks, from January 1995 to December 1999, and 64 IRI markets. Markets are defined by IRI as metropolitan areas or as part of a state. The sample of 64 markets covers the lower 48 states of the US. The data are the result of aggregation over a sample of stores in each market. We use volume sales to compute the market shares of the inside goods. On average, there are about 10 brands of frozen pizza present in each market. For our empirical analysis, we

do not use a 26 week window immediately following launch of DiGiorno. The data in this window display dynamics of post launch sales that often reflect distribution rather than demand (see, e.g., Bronnenberg and Mela 2004). We are primarily interested in consumer substitution patterns that explain the differences between pre- and post-launch market shares given distribution. Thus, for the empirical analysis we censor the 26 week period after launch of DiGiorno (see Figure 2). This means that every market is represented by two time series, one representing the situation before launch and another showing the new market structure after DiGiorno comes into the market.

Second, we have weekly data on the local size of the frozen pizza category as a fraction of total store volume. This measure is informative about the dynamics in category volume (the total size of the inside goods) in a given market.

The third set of data consists of summary statistics about the purchase set size and penetration levels. These data are computed by the AC Nielsen company from its HomeScan panel. From these data, we use the local distribution of purchase set sizes, i.e., the percentage of consumers that buy 1, 2, 3, 4, or more unique brands over the duration of a year. We have these data for each of the 9 Census divisions[9] and for the year 2004. We also have annual penetration levels at the Census divisional level, for the years 2000 to 2003, measuring the percentage of people that have purchased pizza at least once during each year.[10] Each of the 64 markets maps into exactly one Census division.

Finally, from the Census 2000, we have obtained detailed demographic information for each market, which includes population size, average income, age, minorities importance in the population and family size.

5 Demand estimation

5.1 Overview

We estimate the demand model using the Generalized Methods of Moments (Hansen, 1982) combining three different moment conditions. First, we use similar moments as in Berry, Levinsohn, and Pakes (1995) and Nevo (2001). These moments require that the structural error terms, ξ_{jt}, are independent from the instrumental variables. In our application, the "BLP moments" mainly

[9] There are nine US Census divisions: New England, Middle Atlantic, Midwest, East North Central,West North Central, South Atlantic, East South Central, West South Central, Mountain, Pacific. For further definitions, see www.census.gov/geo/www/us_regdiv.pdf.

[10] A point of concern is that these data were collected a few years later than the sales data. However, the penetration measures are very stable across time. To obtain a best estimate for 1999, we "back-casted" the data by region to 1999 using cubic-spline extrapolation.

ensure that the mean market shares are recovered by the model.

Second, in many studies, the size of the outside good is based on an assumption that is more or less ad hoc. In this study, we define a moment using local category penetration data to estimate the size of the outside good empirically.

Finally, we use a third set of moments (the "heterogeneity moments") to identify taste variation. Whereas the "BLP moments" are found to be effective in capturing the mean utilities in the market, we find they underestimate consumer heterogeneity. We also find that this underestimation of taste variation strongly impacts determining the origins of the share for new entrants. Our proposed cure for this problem is to match the predicted distribution of purchase set sizes to local actual data about the number of brands that consumers switch among.

We show that these three sources of identification give us precise estimates of local demand. We do not use the supply side, i.e., prices, to estimate the demand parameters. In order for the observed shelf prices to be informative about the demand parameters in local markets, many assumptions are required about local pricing decisions by national multi-product firms and about the local category management strategies of retailers. The data offer little or no guidance in making such assumptions and wrong assumptions may deteriorate rather than improve our demand estimates.[11]

We now present the details for the implementation of each of these moments.

5.2 The BLP moments

In empirical studies of demand, the analyst often lacks observation of certain demand primitives that are observed and used by the manufacturer as inputs to the determination of price. This endogeneity of price, or the associated correlation between prices and unobserved attributes ξ_{jt} generally causes biases in the estimates of the demand parameters. Past literature has provided evidence of this endogeneity bias when using store-level data (Chintagunta, 2001; Villas-Boas and Winer, 1999).

To account for the endogeneity of price, the usual approach is to rely on instrumental variables and impose an orthogonality condition with the unobserved demand shocks ξ_{jt}. Berry, Levinsohn, and Pakes (1995) have proposed an algorithm to estimate the ξ_{jt}. In this algorithm, the indirect utility function $L_j\omega_i + x_{jt}\beta_i + \xi_{jt}$ is divided in an individual part $\mu_{ijt} = L_j\omega_i + x_{jt}\sigma'\nu_i$ and a mean utility for brand j at time t, $\delta_{jt} = x_{jt}\beta_0 + \xi_{jt}$. Next, given an initial value δ^0_{jt} and a set of parameter

[11] For instance, in our data, retail prices for the incumbents are similar before and after the launch of DiGiorno. Does this mean that the manufacturer charges the same price or the retailer absorbes the shock in whole-sale prices (e.g., Nelson, Siegfried, and Howell 1992)?

values, one can iterate the following expression until it converges[12]

$$\delta_{jt}^{n+1} = \delta_{jt}^{n} + \ln\left(s_{jt}\right) - \ln\left(\widehat{s}_{jt}\left(\delta_{jt}^{n}, \theta\right)\right), \tag{7}$$

where $\widehat{s}_{jt}$ is the predicted market share, which integrates (numerically) over all consumer types, s_{jt} is the actual share and n denotes the iterations in the BLP contraction mapping of equation (7). Once δ_{jt} are computed, we can estimate ξ_{jt} as the residual of the equation $\delta_{jt} = x_{jt}\hat{\beta}_0 + \xi_{jt}$, where in turn $\hat{\beta}_0$ can be estimated using an instrumental variables estimator.

In practice, the shares s_{jt} are often not data, i.e., are usually not directly observed. Instead, what is observed is the conditional shares of the inside goods, i.e., $\tilde{s}_{jt} = s_{jt}/\left(1 - s_{0t}\right)$. This issue is side-stepped by making an assumption about the total size of the market, thereby defining the outside good share, s_{0t}. Our implementation, in contrast to BLP (1995) or Nevo (2001), directly estimates the size of the outside good from data. It replaces s_{jt} in the estimation by the share among the inside goods, $\tilde{s}_{jt}$, which is data, multiplied by 1 minus the estimated share for the outside good (see next subsection).

$$s_{jt} = \tilde{s}_{jt} \times \left(1 - \widehat{s}_{0t}\right) \tag{8}$$

Given our additional moment restrictions below, this suffices for identification.[13]

As instruments, we use prices in other markets (as in Nevo 2001). A concern with such instruments is that they still correlate with the demand shocks if prices are set in clusters of markets rather than in each market individually. To alleviate this concern, we use as instruments prices in far away markets.[14] Related to the positions in the brand map, we impose that the positions L are chosen such that the ξ_{jt} are zero-mean for all brands before and after the launch by DiGiorno. Finally, the instruments for display and feature are the variables themselves. Arraying these instruments into the vector Z_{jt}, we write the "BLP moments" as:

$$G_1\left(\theta\right) : E\left[\xi_{jt}\left(\theta\right) \otimes Z\right] = 0, \tag{9}$$

where the expectation is taken over products and time.

[12] Convergence is obtained $\left|\delta_{jt}^{n+1} - \delta_{jt}^{n}\right| < \varepsilon$, for $\forall \delta_{jt}$, with ε very small ($-\exp\left(10\right)$ in our case).

[13] This modification in and by itself is close to the current practice in estimating demand models. Indeed, it may be realized that current practice also uses equation (8), however that it makes a priori choice about the quantity $\hat{s}_{0t}$ that is contained in it. In our case, we estimate this quantity.

[14] Specifically, for the market of interest m, we sort all other markets in terms of geographic distance to market m. Prices in markets distanced further away than the 10th closest market and on which same period prices are available are potential instruments. We use the prices in 3 far away markets as instruments.

5.3 The outside good moment

The size of the outside good is usually not observed, especially not in a CPG context. Nevo (2000, p. 527) notes there are generally two assumptions in setting the outside good. First, one should determine a variable to which the total size of the market is proportional and second one should determine the value of the proportionality factor. Nevo (2000) also observes that these choices strongly influence conclusions about demand systems and substitution effects. Instead of using an ad hoc assumption, we use data to inform us about the size of the market. In particular, we set the share of the inside goods proportional to total Frozen Pizza expenditure (FPE_t) in a market[15] and we then estimate the –non-structural– proportionality or scaling factor using data.

The proportionality factor is determined as follows. In equation (8), we replace $\widehat{s}_{0t}$ by $1 - \lambda \times \text{FPE}_t$, and we define a moment that chooses the scaling factor λ such that the model is consistent with observed category penetration.[16] The variable "Category Penetration" identifies λ, because –as in Nevo's observation above– different $\widehat{s}_{0t}$ will generate different parameters θ which in turn imply a certain penetration rate. Thus, a better notation would be to recognize that θ is a function of λ, i.e., write $\theta(\lambda)$.[17]

The predicted penetration rate is not a statistic that is readily produced by the model, but it can be computed efficiently as follows. For each simulated household i (resulting from a draw of υ_i and w_i), brand j, and week t, the model predicts a choice probability $\text{Pr}_{ijt}(\theta)$. The model's implied probability that the consumer chooses only the outside good over the last year included in the analysis is

$$\text{Pr}_i(\{\varnothing\}) = \prod_{t=T-51}^{T} \text{Pr}_{i0t}(\theta, \lambda, X_t). \tag{10}$$

This compound probability is smooth in θ and in λ. We require that the population mean of this probability, $E[\text{Pr}_i(\{\varnothing\})]$, is equal to 1 minus the observed penetration rate, π. Thus, we write the "outside good" moment as

$$G_2(\theta, \lambda) : E[\text{Pr}_i(\{\varnothing\})] = [1 - \pi], \tag{11}$$

where the expectation is taken over individuals i.

[15] Thus, category volume is measured as the share of Frozen Pizza expenditure (FPE_t) among all categories scanned in a given market. Alternatively, one can make the joint share of the inside goods proportional to category sales. However, total recorded category sales in our data is subject to dynamics in the IRI store sample. This makes that the dynamics of the Frozen Pizza share of store sales is better at capturing the dynamics of the joint sales of the inside goods.

[16] A notational distinction between the parameters θ and λ is made on the grounds that the former are the structural parameters of the demand system, while the latter is a –non-structural– scaling constant that translates data about category size into "data" about the outside good.

[17] Other papers have the same conditionality. Our paper differs from those in that we estimate λ rather than assume it.

In sum, this moment restriction requires of the demand system that the weekly size of the outside good is consistent with the actual fraction of households that do not buy Frozen Pizza, and that the size of the inside goods is consistent with the recorded weekly dynamics of category size.

5.4 The heterogeneity moments

As explained before, we use data on the distribution of purchase set sizes, S_i, to estimate the dispersion of preferences and, importantly in the evaluation of a new product introduction, the degree of switching in the Frozen Pizza category. Our data cover the empirical distribution of the purchase set size across households, $\Pr(S_i = 0)$, $\Pr(S_i = 1)$, $\Pr(S_i = 2)$, etc., in different regions in the United States. For example, in the Pacific Census division, 39% of households buy 0 Frozen Pizza brands in a year (penetration is 61%), 22% of households buys only 1 unique brand, 17% switch among 2 brands, 10% switch among 3 brands, and 5% switch among 4 brands (the remaining 6% of households switch among more than 4 brands).

To estimate the degree of preference dispersion in the market, we require that our estimated demand system matches the actual individual level distribution of purchase set sizes. For this purpose, we need to compute the implied purchase set size distribution of the model. The model's forecast of $S_i(\theta, X_t)$ can be computed recursively from the choice probabilities, $\Pr_{ijt}(\theta, X_t)$ in equation (5), for each simulated individual i in week t. As an example, we provide details on the predictions for $\Pr(S_i = 1)$ and $\Pr(S_i = 2)$.

Start with the joint probability that a weekly observed consumer buys brand j, nothing, or combinations thereof over the course of an entire year,

$$H_{ij} = \prod_{t=T-51}^{T} \left[\Pr_{ijt}(\theta, X_t) + \Pr_{i0t}(\theta, X_t)\right]. \tag{12}$$

This probability covers all purchase histories that combine any positive number of purchases of j with any positive number of purchases of the outside good. Therefore (using the notation in equation 10),

$$\Pr_i(\{j\}) = H_{ij} - \Pr_i(\{\varnothing\}) \tag{13}$$

is the probability that the purchase set is $\{j\}$ in a given year. Finally, the probability that the consumer has a purchase set size of exactly 1, is equal to the summation of $\Pr_i(\{j\})$ across choice options j.

$$\Pr(S_i = 1) = \sum_{j=1}^{J} \Pr_i(\{j\}). \tag{14}$$

Next, $\Pr(S_i = 2)$ can be computed from the probability that the consumer purchases j, k, nothing, or combinations thereof for an entire year,

$$H_{ijk} = \prod_{t=T-51}^{T} \left[\Pr_{ijt}(\theta, X_t) + \Pr_{ikt}(\theta, X_t) + \Pr_{i0t}(\theta, X_t)\right]. \tag{15}$$

This probability covers all purchase histories involving j, k, and the outside good. The probability $\Pr_i(\{j,k\})$ that consumer's purchase set is $\{j,k\}$, i.e. that the purchase set contains at least one j and one k but no other brands than the outside good is then (using equation 13),

$$\Pr_i(\{j,k\}) = H_{ijk} - \Pr_i(\{j\}) - \Pr_i(\{k\}) - \Pr_i(\{\varnothing\}). \tag{16}$$

As a final step, the probability that a purchase set of exactly size two is observed for a given individual is the sum of $\Pr_i(\{j,k\})$ across all unique combinations of j and k.

$$\Pr(S_i = 2) = \sum_{j=1}^{J} \sum_{k=j+1}^{J} \Pr_i(\{j,k\}) \tag{17}$$

The probabilities $\Pr(S_i = 3), \Pr(S_i = 4)$, etc., are recursively computed in a similar fashion.

Write the population values for the fractions $\Pr(S_i = n)$ as $\mathcal{F}_n$. Then, the final set of moments can be written as

$$G_3(\theta) : E\left[\Pr(S_i = n)\right] = \mathcal{F}_n, \; n = \{1, ..., 4\}, \tag{18}$$

where the expectation is again taken over households.

This set of moments ensures that the model parameters are chosen such that the implied amount of switching given prices, promotion, etc., matches the switching in the Frozen Pizza category observed during the introduction of DiGiorno.

5.5 Objective function and simulation

The objective function combines the three sets of moments previously described:

$$G(\theta) = \begin{bmatrix} G_1(\theta) \\ G_2(\theta) \\ G_3(\theta) \end{bmatrix}. \tag{19}$$

In order to compute the expectations in $G_1(\theta)$, $G_2(\theta)$, and $G_3(\theta)$ we need to use simulation. For instance, the expectation in equation (18)

$$E_i\left[\Pr(S_i = n)\right] = \int \Pr(S_i = n)\, \phi(\upsilon)\, \phi(\omega)\, \partial\upsilon \partial\omega \tag{20}$$

can not be computed analytically, but must be approximated. The simplest way of doing this, is to use the pseudo panel of (υ_i, ω_i) draws that is already in use for the approximation of the market share integrals in $G_1(\theta)$.

$$E_i\left[\Pr\left(S_i = n\right)\right] \approx \frac{1}{N}\sum_{i=1}^{N} \Pr\left(S_i = n|\theta, X_t, \upsilon_i, \omega_i\right) \tag{21}$$

This approximation is again smooth in the parameters θ.

Next, we use these approximations in a two-step GMM estimator (Hansen, 1982; Petrin, 2002).

$$\widehat{\theta} = \arg\min_{\theta\in\Theta}\left(\widehat{G}\left(\theta\right)' W\left(\widetilde{\theta}\right)' W\left(\widetilde{\theta}\right)\widehat{G}\left(\theta\right)\right) \tag{22}$$

where $\widehat{G}(\theta)$ is the sample analogue of $G(\theta)$ and $W\left(\widetilde{\theta}\right)$ is a weight matrix consisting of an estimate of the "square root" of the inverse of the variance-covariance matrix of the moments, obtained using $\widetilde{\theta}$, a preliminary consistent estimate of θ. For the first set of moments, $G_1(\theta)$, the weight matrix is given by

$$W_1\left(\widetilde{\theta}\right)' W_1\left(\widetilde{\theta}\right) = \left[\frac{1}{T}\sum_{t=1}^{T} g_{1t}\left(\widetilde{\theta}\right)\cdot g_{1t}\left(\widetilde{\theta}\right)'\right]^{-1} \tag{23}$$

where $g_{1t}\left(\widetilde{\theta}\right)$ are the moment values for each time period. The second and third moment is a result of a average of individual probabilities. The associated weight matrix is equal to the inverse of the variance, i.e.

$$W_2\left(\widetilde{\theta}\right)' W_2\left(\widetilde{\theta}\right) = \left[\frac{1}{N}\sum_{i=1}^{T} g_{2i}\left(\widetilde{\theta}\right)\cdot g_{2i}\left(\widetilde{\theta}\right)'\right]^{-1}, \tag{24}$$

where $g_{2i}\left(\widetilde{\theta}\right)$ is a vector moment values of $\left(\Pr\left(S_i = n\right) - \mathcal{F}_n\right)$, $n = 0, ..., 4$. Finally, $W\left(\tilde{\theta}\right)' W\left(\tilde{\theta}\right)$ is the block-diagonal combination of the two parts defined above.

5.6 Computing local switching

As stated in section 2, we compute local switching after estimating the demand system in each market separately. In order to obtain the switching from incumbent brands to DiGiorno, we compare two scenarios. The actual scenario, where DiGiorno was introduced in the market, and an alternative counterfactual case, where we remove DiGiorno from the market. We then compute the difference of shares of incumbent brands in the two scenarios. The idea behind this method is to identify who would have kept the share that was transferred to the introduced brand.

Formally, brand switching is computed using the following expression:

$$\begin{aligned} \Delta s_{jt} &= \frac{1}{N}\sum_{i=1}^{N}\left[\Pr_{ijt}\left(\theta, X_t, \text{DiGiorno in}\right) - \Pr_{ijt}\left(\theta, X_t, \text{DiGiorno out}\right)\right] \\ j &= 1,..,J, j \neq \text{DiGiorno} \qquad \forall t \text{ after DiGiorno's entry} \end{aligned} \tag{25}$$

Under the assumptions of the model this measure is less than or equal to 0 (incumbent brands will not gain share from the introduction of DiGiorno) and larger than minus the share of the incumbent brands prior to the launch of DiGiorno.

6 Results

6.1 Introduction

We present our results in the following order. First, we verify that the longitudinal and cross-sectional fit of the model is good. We also show that the standard model overestimates the degree of brand switching in a market. Second, we investigate the impact of the extra moments on the estimates of the model. Third, we discuss the structural parameter estimates and the implied brand switching. Fourth, we conduct a policy experiment in which we evaluate alternative positions for the new brand, DiGiorno. Finally, we explore the cross-market differences in substitution patterns by relating the market by brand substitution patterns to national branding, to regional branding, and local variables.

6.2 Model fit

We first evaluate how well the model explains market shares and the distribution of the purchase set size. In addition, to evaluate the improvement stemming from the additional moments, we compare our proposed model (the "augmented" model) with the model in which the additional moments are not included in the estimation (the "standard" model). Since the absence of these additional moments would not permit the estimation of the outside good, we use the estimated shares for this alternative from the augmented model as data.[18]

The model was estimated for the 64 markets in the data set individually. To illustrate the findings, we discuss the case of Chicago. Figure 3 displays the time series of actual and estimated shares for 2 brands in this market. The shares are estimated excluding the demand unobservables ξ_{jt} and the error term ϵ_{ijt}, as these are not observed by the analyst. Our model estimates market

[18] Because the standard model uses the outside good estimates from the augmented model, the standard model is likely to perform better than it would with an ad hoc assumption about the outside good. The contrast in relative improvement of model fit from the extra moments is therefore likely to be conservative.

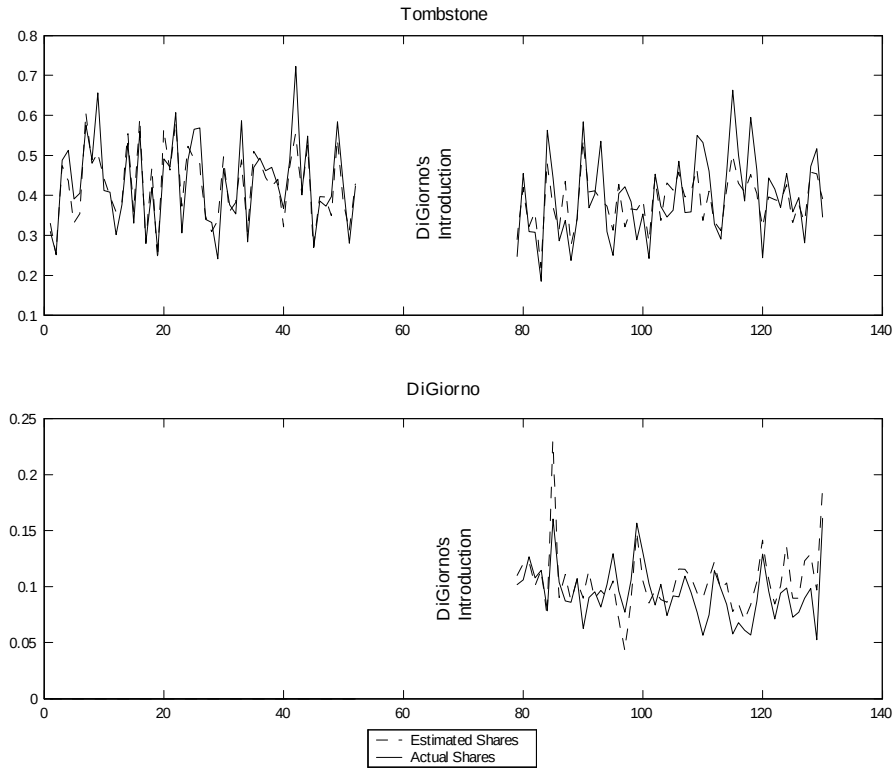

Figure 3: Actual and estimated shares for Tombstone and DiGiorno in Chicago.

shares that include the outside good, where as the actual data only considers the share among inside alternatives. In order to make the comparison possible, we compute the estimated share of each brand within the category.[19] It is clear that the model does very well at explaining the brand level variation in market shares. This is also true in the other markets. We note that the standard model does also well in recovering the market share time series.

Next, Table 2 shows that the augmented model is highly accurate at setting the demand parameters such that the model matches the purchase set sizes observed in the market. The augmented model implies the correct fraction of households that stay out of the market, are loyal to the same brand, or switch among 2, 3, or 4 brands. In contrast, the standard model does not, even when we use the information about the outside good borrowed from the augmented model. Specifically, the final column in the table shows that the information-poor model underestimates the fraction of households that are loyal to the outside good and that are loyal to a single brand. At the same time, it overestimates the number of households that switch among 2 or more brands. As we will show subsequently, this is because the standard approach underestimates consumer heterogeneity.

[19] We do this by dividing the estimated shares of each brand by $(1 - \hat{s}_{0t})$, in each time period.

purchase set size	actual	augmented model	standard model
0	0.238	0.261	0.039
1	0.214	0.216	0.158
2	0.170	0.175	0.307
3	0.132	0.124	0.250
4	0.096	0.089	0.106

Table 2: The implied purchase set size distribution for Chicago, using the augmented model and the standard model.

The patterns discussed here are not specific to Chicago but happen generally in the data.

6.3 The impact of the extra information

6.3.1 Estimating the outside good

The estimation of the outside good share is fundamental in the process of accurately identifying the switching between alternatives, since category expansion plays in many cases an important role in generating demand for a new brand, especially one that offers unique product features. In the case of the frozen pizza category, the outside alternative has additional relevance, since DiGiorno used a well documented media campaign that specifically targeted delivery pizza, a component of the outside good.

Our estimate of the outside good is based on market level penetration rates. Across all time periods in the analysis, we find an estimated weekly share of the outside good that ranges across markets from a minimum of 70% to a maximum of 95%. *Prima facie* these results are plausible and accord with the practitioner's estimate that the Frozen Pizza category constitutes 20% of the total Pizza Category (*Pizza Marketing Quarterly*). As an additional validity check, we computed the correlation between the estimated inside goods share $(1 - \hat{s}_{0t})$ and the average expenditures in frozen pizza, across the 64 markets. As expected we find a strong positive correlation across markets of 0.60. In general, markets where pizza is traditionally a popular choice, such as Chicago, display lower shares of the outside alternative.

We next analyze the evolution of the outside good, due to the introduction of DiGiorno, by comparing its average share before and after DiGiorno was launched. We plot the result in Figure 4(a) for each of the 64 markets. The size of the circles represent share losses suffered by the outside good, with larger circles reflecting more category expansion. The expansion shows a pronounced spatial concentration in the North-East part of the U.S. Figure 4 (b) shows the average share for DiGiorno. From these two graphs, it appears that the locations where DiGiorno achieves a higher

Figure 4: (a) Change in the share of the outside good before and after the launch of DiGiorno, (b) annual market share of DiGiorno starting 26 weeks after introduction

market share are those where we also see more reduction in the outside good. This finding suggests that a considerable percentage of DiGiorno's share came from the expansion of the category. We investigate this issue in more detail in the next subsections.

6.3.2 Estimated heterogeneity and the origins of brand switching

The estimated degree of consumer preferences is very different between the standard model and the augmented model. For instance, the standard deviation of price responses in the Chicago population is estimated to be 0.73 when using the augmented model vs. 0.12 using the standard model. The positions L in the brand perceptions map are also different. Specifically, Figure (5) shows that DiGiorno, Red Baron and Jack's are all closer to the outside good in the standard model, making it more likely for people to change between those brands and the outside alternative.

The effect of this result on the estimated origins of DiGiorno's share is very substantial, as shown in Table 3.[20] The standard model displays a higher percentage of switchers from the outside good to DiGiorno then the proposed augmented model (66% vs. 41%), while the incumbent brands loose importance as a source of DiGiorno's share. Again, this can be explained from the differences in inferred brand positions and price heterogeneity. With the augmented model, DiGiorno is farther from the outside good alternative; in addition, a larger price heterogeneity estimated by this model implies that the most price sensitive consumers still stay out of the category when DiGiorno is

[20] The mean and standard deviation of the switching was obtained by computing the implied origins of switching across 100 parameter vectors, that were randomly drawn from the empirical distribution of the parameters in Chicago.

	augmented model		standard model	
	mean	st. dev.	mean	st. dev.
Tombstone	0.057	0.057	0.036	0.058
Jack's	0.197	0.080	0.116	0.081
Red Baron	0.122	0.123	0.075	0.056
Tony's	0.069	0.051	0.074	0.098
Totinos	0.144	0.124	0.035	0.086
Outside Good	0.411	0.149	0.664	0.163

Table 3: Estimates of the switching to DiGiorno, for the market of Chicago, using the augmented model and the standard model.

introduced.

6.4 Structural parameters and brand perceptions

After verifying model fit, and assessing the importance of the extra moments, we now turn our attention to the structural parameters and the sources of share of the new brand. We display parameter means and standard deviations in Table 4. Price, display and feature coefficients have expected signs, with the values of -1.25, 2.24 and 0.60 respectively. For Chicago, the parameters imply an average price elasticity of -1.13, while display and feature elasticities are 1.84 and 0.46.[21] Across all markets, these elasticities have an average of -1.9, 0.9 and 0.4 respectively.

The estimation of the proposed model delivers also the positions of the brands in a two-dimensional perceptual map. Figure 5 shows the locations of brand perceptions in Chicago. Several interesting findings are worth point out.

First, consumers perceive that DiGiorno is located closest to the outside good in this market. A more general analysis across all markets confirms this to be common. We computed the cross-market distribution of the "city-block" distance between the brand positions and the outside good, across the 64 markets, shown in Figure 6. DiGiorno is located closer to the outside good than any other brand, providing further confirmation that Kraft's strategy of positioning the brand as a substitute of the delivery pizza alternative (included in the outside good) was highly successful.

Second, the perceptual map shows all brands in the negative quadrant of the two attributes, when compared to the outside good. The two brands that are valued highest by consumers for the two perceived attributes (which have positive weights) are DiGiorno and Jack's, two of Kraft's brands. Given the brand positions, we can conclude that Kraft was able to avoid cannibalizing

[21] Feature and display elasticities are computed by increasing the percentage of sales featured or on display by 1% over all periods of time.

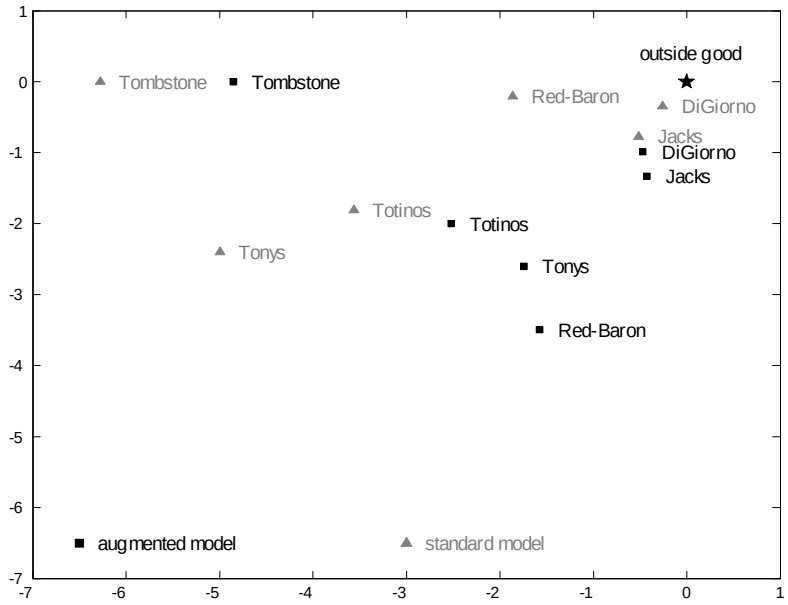

Figure 5: Perceptual maps for the main brands in Chicago, contrasting the augmented and the standard model.

their other main brand Tombstone. DiGiorno and Tombstone are usually located far away from each other and the switching from Tombstone represents a small percentage of DiGiorno's share. On the other hand, Jack's was more penalized by the entry of the new brand, since the brands were perceived as substitutes by consumers. It is also likely that the firm decided to give DiGiorno some of the resources assigned previously to Jack's (or were forced by retailers to do so), such as shelf and feature space, in the markets where this brand was present, leading to consumer switching to the newer brand.

Third, it is interesting to note that the two Schwan's brands are located in neighboring positions in Chicago, as it is so in the majority of the markets. In fact, there is a strong negative correlation between the shares of the two brands in those markets. Its reason arises from the fact that Schwan's uses a strategy of pulsing display and feature of these two brands alternatively, that is, when Tony's is featured, Red Baron is not, and vice-versa. This choice of policy is somewhat justified by restrictions imposed by retailers, for instance, on the space used for the category in inserts and advertisements.[22] Although the model accounts for such marketing mix variables in the utility function, brand positions indicate that the incentive to consumers to switch often between the two brands makes them closer substitutes.

[22] This information was provided to the authors in discussions with managers of the firms in the category.

	Mean	St. Dev.
1st Attribute		
Tombstone	-4.853	2.314
Jack's	-0.425	1.005
Red Baron	-1.575	1.710
Tony's	-1.743	1.056
Totinos	-2.521	2.524
DiGiorno	-0.467	0.325
2nd Attribute		
Jack's	-1.334	2.774
Red Baron	-3.497	3.338
Tony's	-2.602	2.799
Totinos	-2.004	6.292
DiGiorno	-0.986	0.890
Price	-1.253	0.015
Display	2.241	0.183
Feature	0.603	0.142
Price Heterogeneity	0.734	0.218
Outside Good Scale	0.266	0.064

Table 4: Estimation results for Chicago.

6.5 Evaluating repositioning strategies

Managers have considerable control over brand positioning (or repositioning) by changing its advertising message, its price or its promotion strategies. In the specific case of new products, it is also possible to test different positioning strategies in early markets and opt for the best policy for the remaining locations.[23] Our factor analytical approach allows us to measure the effects of alternative positioning strategies of the new brand.

In the policy simulations below, we assume different values for the attributes L for DiGiorno and, given the data and the parameters obtained by our model, we compute its average share and switching from incumbents. We continue to explore the Chicago market as illustration and present the results in Figure 7. We discretize the perceptual map space, in increments of 0.5 over the interval [-5, 1] for the two attributes, in order to evaluate the impact of alternative locations of the new brand.

Looking at the top left panel in Figure 7, switching from competitors is highest when DiGiorno has positions close to the three main competitors, i.e., when located around (-2,-4) for attribute

[23] For instance, DiGiorno was initially tested in four markets. A more widespread introduction across markets was done after a few months, and an even larger market introduction occurred a year later.

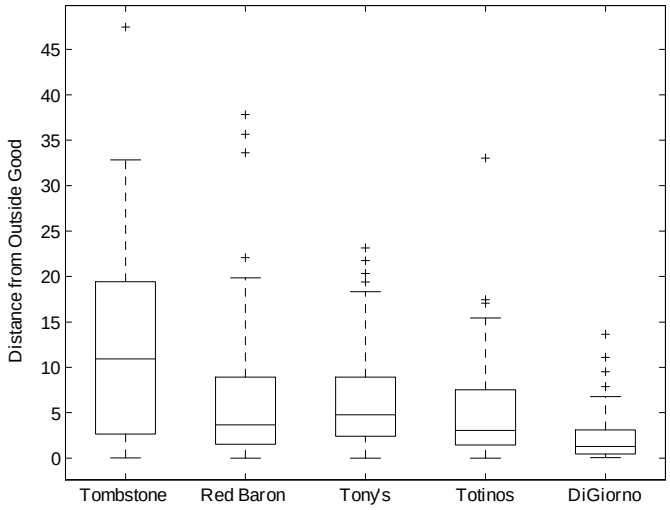

Figure 6: Boxplot of the city-block distance between the positions of the brands in the perceptual map and the outside good, for the 64 markets.

1 and 2 respectively. This area of the perceptual map appears, at first glance, to be attractive to DiGiorno, since a large percentage of its share, close to 70%, comes from competitor brands, with cannibalization below 20%. However, such positioning would place the brand in direct competition with Red Baron, Tony's and Totinos, in a very crowded area of the perceptual map, where differentiation between brands is low and price competition is likely to increase. The brand draws less share from competitors in regions close to the outside good (0,0), or close to Tombstone. Combining the cannibalization and category expansion panels in Figure 7 both locations are favorable in terms of allowing larger shares for DiGiorno (see also the totals in the lower right panel). However, while positioning closer to Tombstone causes most of DiGiorno's share to come from Kraft's extant brands, Tombstone and Jack's, positioning closer to the outside good attracts consumers from outside of the category to purchase the new brand.

The analysis of these simulations suggest that Kraft opted for the best strategy in creating share, i.e., one that allowed DiGiorno to become one of the main competitors in the market, while keeping cannibalization and direct competition within the category to a minimum.

6.6 Multimarket exploration of the origins of DiGiorno's demand

Across all markets, Kraft's strategy of targeting the outside good (delivery pizza) to avoid cannibalizing its own brands appears to have been successful. The cross-market distribution of the switching from other brands owned by Kraft, from competitor brands and from the outside alternative to DiGiorno is shown in Figure 8. Only a small percentage of DiGiorno's share originated

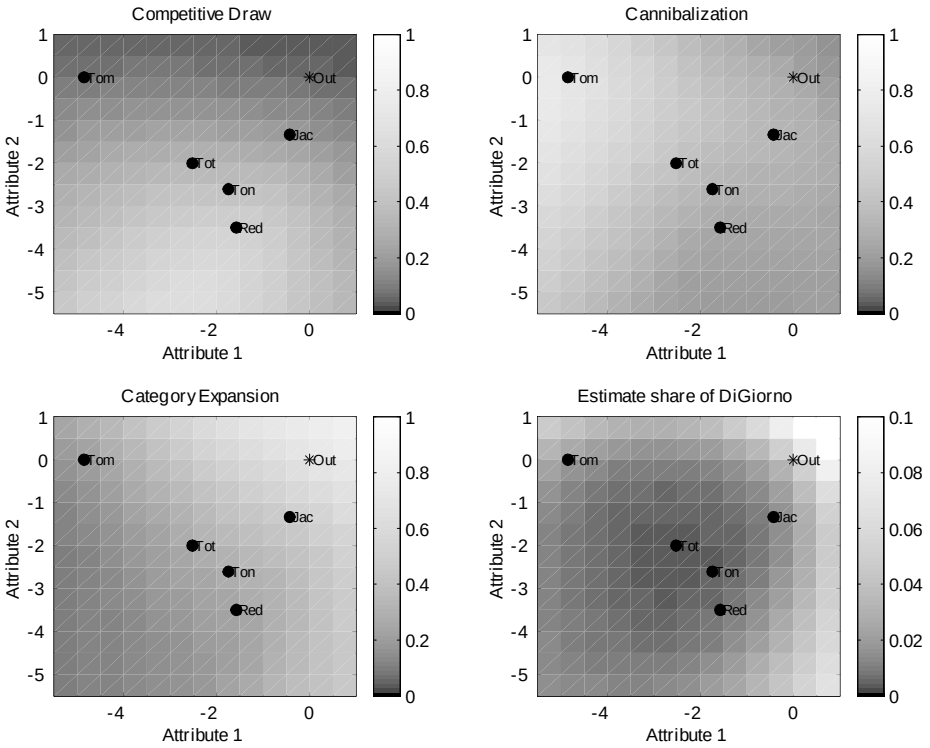

Figure 7: Cannibalization, Competitive Draw, Category Expansion and DiGiorno's Share for different positioning strategies of DiGiorno. Tom – Tombstone; Tot – Totinos; Ton – Tony's; Red – Red Barron; Jac – Jacks; Out – Outside good

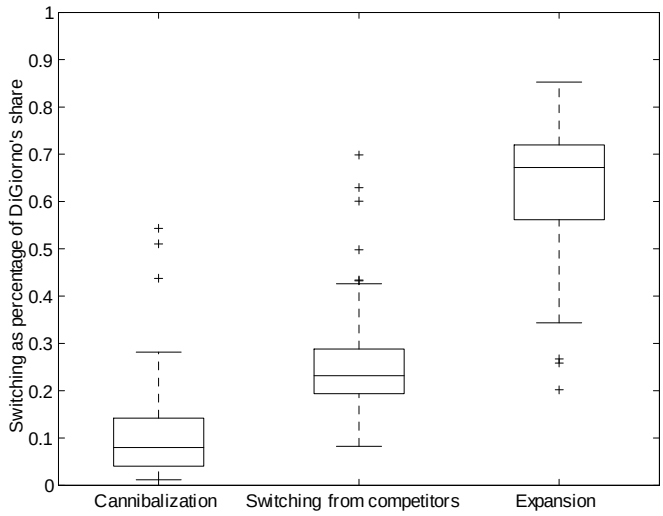

Figure 8: Percentage of DiGiorno's share originated from cannibalization, switching from competitors and category expansion.

from cannibalization of its own brands, on average about 11%. About 26% of the share came from competitors' brands and 63% from category expansion. This result provides evidence that an advertising campaign combined with innovative product characteristics[24] can be successful in positioning the brand as a direct competitor of a specific alternative, in this case from outside of the category.

As Figure (8) however shows, the relative importance of cannibalization, competitive draw and category expansion is not the same across markets and we observe considerable geographic variation. Figure 9 displays pie charts with the relative weights, for all locations in our data set. The three sources of switching are spatially dependent, e.g., there are sets of contiguous markets in which category expansion, cannibalization, or competitive draw, is similarly valued. The Mid-Western markets show the highest degree of cannibalization. These are markets where Jack's is an important competitor, and as seen before in the discussion of the case of Chicago, it was one of the brands that suffered most from the introduction of DiGiorno. Switching from competitor brands is larger in locations where prior to DiGiorno's introduction, Schwan's had a large share. The importance of category expansion as a source of DiGiorno's share is considerable across all markets. Although the attraction of new consumers to the category may be surprising at first glance, the data gives support to this result. The category revenue grew an average of 17%[25] across all markets during

[24] As mentioned before, DiGiorno had the advantage of having an innovative characteristic compared to competitors in the category - the rising crust.

[25] This growth is computed by comparing the expenditures on frozen pizza during the period under analysis. Prices of frozen pizza did not vary significantlly before or after the launch of DiGiorno.

Figure 9: Relative importance of cannibalization, switching from competitors and category expansion, across 64 markets.

the period in our analysis, which was enough to fuel a large portion of DiGiorno's volume share. We conclude from this analysis that the origins of DiGiorno's demand are (1) local and (2) spatially dependent.

Finally, we wish to explore the geographic differences of switching toward the new brand. Variability in the origins of demand may be related to local conditions in each market, such as pre-entry share, local prices or demographics. On the other hand, it is also possible that switching depends on regional factors, such as regional retail chains. Lastly, the origins of demand could be due to factors that are common across all markets, for instance, brand names. We intend to explore the primitives behind the switching behavior of consumers, by exploring the market-brand variation in the switching using regression analysis. The dependent variable in this regression is the $[((J-2)\times M)\times 1]$ column vector y_{jm} measuring the loss in shares of the inside good j to DiGiorno in market m as given by our model:

$$y_{jm} = -\frac{1}{T}\sum_t \Delta\widehat{s}_{jtm} = -\frac{1}{T}\sum_t \left[\hat{s}_{jtm}\left(\text{DiGiorno out}\right) - \hat{s}_{jtm}\left(\text{DiGiorno in}\right)\right] \tag{26}$$

(see equation 25). Since y_{jm} varies between zero and one (the share of an incumbent brand is always larger in the counterfactual scenario where the new brand is not introduced), we use the logit

transform on the dependent variable $\tilde{y}_{jm} = \log(y_{jm}/(1 - y_{jm}))$. We further group our explanatory variables into national, regional and local factors.

The national factors include brand dummies as a measure of any characteristic that varies across brands, but not across markets. It is possible that consumers end up perceiving DiGiorno to be a better substitute to some brands, which would lead these incumbents to loose a higher portion of their share to DiGiorno.[26].

Regional factors comprise variables specific to large regions in the U.S., such as retailer chains. We proxy this effects by including dummy variables for each Census division.

Local variables capture the impact of factors that vary across markets. During DiGiorno's introduction, each market presented its own set of initial conditions. We include in the analysis the average pre-launch share of each incumbent and the average expenditures on frozen pizza, measured as a percentage of the total basket size level, in each market, before DiGiorno was introduced. We next add relative price, feature and display of each brand, measured as the difference between the average level of these marketing mix variables for each incumbent brand and DiGiorno, in each market.[27] Finally, in terms of local consumer characteristics, we include several demographics, such as income per capita, population, family size, average age and percentage of whites in the population.

Our results are displayed in Table 5. We present six alternative models, which enable us to decompose switching variation into the previously presented factors. Model 1 focuses on the national factors - brand intercepts. Model 2 includes the regional dummies. Models 3 to 5 present alternative local variables, while model 6 includes all explanatory variables.

The first model with only national brand intercepts measures the impact of brand name on switching. For identification, one brand - Tombstone - is left out of the regression. The fit of this model is poor, with an R-square of 4%, suggesting that the brand names or any factors that are common across markets do not explain the origins of DiGiorno's share. This is remarkable, because it suggests that there are no brand fixed effects on substitution. In turn, this suggests that brand substitution is not related to brand differences in product characteristics that are common across markets. It is important to note that if the researcher observes just one market, the conclusion about the impact of brand effects may very different. For instance, in Chicago, Tony's contribution to DiGiorno's share is very small, which would at first glance imply that consumers from this brand are

[26] Instead of using brand dummies, we could instead include product attributes, such as calories per serving. This approach performed worse in explaining the switching variance.

[27] Averages are taken across time periods, over the first 52 weeks after the 26 week launch period for DiGiorno in each market.

	Model (1)			Model (2)			Model (3)			Model (4)			Model (5)			Model (6)		
	mean	std		mean	std		mean	std		mean	std		mean	std		mean	std	
Intercept	-6.68	0.12	*	-6.40	0.17	*	-7.13	0.17	*	-6.03	0.12		-13.65	2.71		-12.80	2.37	*
National Effects																		
Jack's	0.05	0.22														0.35	0.22	
Red Baron	0.27	0.17														0.57	0.15	*
Tony's	0.50	0.17	*													0.69	0.14	*
Totinos	0.19	0.18														0.53	0.19	*
Regional Effects																		
Region																		
New England				0.68	0.29	*										0.43	0.27	
Middle Atlantic				0.51	0.26	*										0.28	0.23	
East North Central				0.33	0.21											0.51	0.21	*
West North Central				-0.58	0.24	*										-0.14	0.24	
South Atlantic				-0.33	0.22											-0.53	0.21	*
East South Central				0.22	0.24											0.22	0.25	
West South Central				-0.55	0.23	*										-0.37	0.23	
Mountain				-0.64	0.26	*										-0.43	0.24	
Local Effects																		
Pre-entry conditions																		
Pre-entry share							2.81	0.48	*							3.71	0.44	*
Expenditures							0.01	0.03								0.00	0.04	
Firm controlled																		
Price[a]										0.26	0.15					-0.01	0.21	
Feature[a]										-3.33	0.81	*				-2.56	0.75	*
Display[a]										1.53	0.81	*				0.74	0.82	
Consumer effects																		
Population													0.01	0.03		-0.01	0.02	
Income													3.59	1.53	*	0.47	1.64	
Age													0.15	0.05	*	0.11	0.05	*
% White													-1.01	0.48	*	-1.09	0.44	*
Family Size													0.77	0.47		0.88	0.40	*
R-square	4%			19%			23%			7%			8%			43%		

[a]difference with DiGiorno's values

Table 5: Results of the regreession analysis on switching across markets

somewhat "insulated" from switching to the new brand. However, a cross-market approach shows that this is not the case, since in other markets, Tony's does show larger losses to DiGiorno. We conclude that in our application, the origins of switching have no meaningful relation to national brand characteristics including "brand label."

Model 2 includes regional dummies, where a positive significant effect in one of the regions represents higher switching to DiGiorno in that region than in the base region (Pacific Region). Regional factors explain about 19% of the variance in the switching, and capture factors such as the presence of the same retailers across regions. Two regions are significantly different from the base region. East North Central shows a higher switching to DiGiorno, while South Atlantic presents the opposite situation. In general, the North regions of the country were more favorable to DiGiorno, with higher shares located in those areas.

Model 3 includes two variables that measure several aspects of the market structure prior to DiGiorno's introduction: pre-entry shares of incumbent brands and the percentage of the total basket size spent on frozen pizza by consumers. Pre-entry share is by itself the factor that explains the highest part of variation, with higher share brands contributing significantly more to DiGiorno's share than smaller brands. However, from the still low R^2 of this model (0.23), it is clear that switching is not explained from merely a proportional draw or IIA hypothesis, where each brand would contribute to the share of DiGiorno in proportion to its previous shares.

Marketing mix variables are included in Model 4. We include the relative price, feature and display, by computing the difference between prices, feature and display of each incumbent brand and DiGiorno. These variables explain about 7% of the switching variation. The coefficients show that once accounting for all other effects (Model 6), feature is the only variable that has a significant impact: the more an incumbent brand is featured when compared to DiGiorno, the less it looses to the new brand.

Finally, demographic characteristics are explored in Model 5. Consumer demographics capture differences in tastes, attitudes towards frozen pizza or new brands, etc.. They explain about 8% of the variation. Switching to DiGiorno was higher in markets where the average age is higher and family size is larger.

Aggregating all factors, we are able to explain 43% of the total variation of the switching to DiGiorno.

From this analysis, we find that the origins of DiGiorno's share differ strongly across markets. Surprisingly, we find almost no brand main effects in the switching patterns. National brand

differences in substitution are not a main factor in explaining switching to DiGiorno. Second, switching to DiGiorno follows a regional pattern. Whereas we find regional fixed effects, we did not explore the causes of this regional distribution. It may be related to the presence of retailers across markets or differences in perception for the new and the incumbent brands across markets. Finally, local variables do the best in explaining differences in brand switching. These variables include pre-entry conditions, which we found to be the most important factor, local marketing mix policies and local demographics. We also find that an incumbent brand manager can, up to a point, influence the switching by having different levels of marketing mix variables for its incumbent brand (in our case, feature).

7 Conclusion

In this study, we provided an in-depth analysis of local demand for a new national CPG product. Methodologically, we aimed to add insight into resolving a recurring dilemma in the estimation of demand. On one hand, access to an optimal level of information, i.e., very rich data on every individual (or at least large samples of individuals) for all markets/stores under analysis, is almost impossible to obtain. On the other hand, market level data are easier and cheaper to obtain, but are considerably less informative about the individual behavior, as details are lost due to aggregation. Our analysis combines multiple sources of information, all readily available to the marketing manager, or the interested analyst to offer a feasible solution to this conundrum. Specifically, we propose to augment the time series on sales or market share, which is a marginal of the individual level data across households, with summary statistics of consumer purchase behavior, which is a marginal of the individual level data across time periods.

We have shown that using this combination of data produces markedly different estimates of consumer demand for a new product. Specifically, the traditional model that relies on the time series only, gravely overestimates the degree of brand switching that takes place in the market. This is because such models tend to underestimate the degree of preference heterogeneity in the market place. We identified the household purchase set size as a variable that is very informative of dispersion of preferences. Our proposed approach determines the level of preference heterogeneity that reproduces almost exactly the correct number of brands among which a consumer switches, thereby aiding in the identification of brand switching and in predicting new product demand.

Another methodological contribution of our approach is that we estimate the overall size of the outside good by relating popularity of the outside good to observed local consumer tendencies to

stay out of the market. Therefore, our demand model can be used to identify the three sources of market share of a new brand - cannibalization, competitive draw and category expansion - through the use of easy to obtain market level data.

Substantively, the empirical illustration used data from the Frozen Pizza category and focused on the launch of a highly successful CPG brand. From company interviews and from its well known media campaign, Kraft was focusing its attention on the delivery market as one of their main targets. Our estimates of the introduction of DiGiorno confirms that the outside good was the main source of DiGiorno's demand. However, there are large differences across markets of how successful DiGiorno was at stealing share from the outside good. Similar to what we observe in the outside good, switching from the inside goods is also very different across markets.

We finally demonstrated that, in our application, the cross-market variation in the origins of new brand demand are related to several local factors, but not to national branding. Thus, the substitution among brands is not a national phenomenon. Of the factors that could explain the variation in substitution patterns, we find that local factors such as pre-entry share and regional brand popularity are both important.

We note several avenues for future research. The differences across markets in local brand similarity are important drivers of new product demand. Yet, little is known about why brand similarity is different across markets. In other words, we do not know whether the primitives of brand similarity are consumer related, retailer related, or related to the marketing actions of the manufacturers. Another avenue for future research is to investigate the broader role of advertising content in shaping the substitution patterns during the new product launch.

Finally, in practical terms, we believe our model is helpful to managers in evaluating the impact of new product introductions in local markets. To the extent that this impact is spatially dependent across markets, our model can be used in phased national roll-outs, such as the one used by Kraft to launch DiGiorno, to forecast new product switching at a pre-entry stage, based on post-entry data from nearby markets.

References

[1] AC Nielsen (2001), *Consumer Insight: New Product Introductions*, Consumer and Market Trends Report, AC Nielsen, Chicago, IL

[2] Berry, S.T. (1994), "Estimating Discrete-Choice Models of Product Differentiation," *RAND Journal of Economics*, 25:2, pages 242-262

[3] Berry, S.T. , J. Levinsohn and A. Pakes (1995), "Automobile Prices in Market Equilibrium", *Econometrica*, 63:4, 841-890

[4] Berry, S.T., J. Levinsohn and A. Pakes (1998). "Differentiated Products Demand Systems from a Combination of Micro and Macro Data: The New Car Market," *NBER Working Papers 6481*, National Bureau of Economic Research, Inc.

[5] Blattberg, R. and K. Wisniewski (1989), "Priced-Induced Patterns of Competition," *Marketing Science*, 8:4, 291-309

[6] Bodapati, A. and S. Gupta (2004). "The Recoverability of Segmentation Structure from Store-Level Aggregate Data," *Journal of Marketing Research*, 41:3, 351-364

[7] Bronnenberg, B.J., S.K. Dhar, and J.P. Dubé (2007), *Market Structure and the Geographic Distribution of Brand Shares in Consumer Package Goods Industries*, working paper, UCLA

[8] Bronnenberg, B.J., and C. Sismeiro (2002), "Using Multimarket Data to Predict Brand Performance in Markets for Which No or Poor Data Exist," *Journal of Marketing Research*, 39:1, February, 1-17

[9] Bronnenberg, B.J. and C.F. Mela (2004), "Market Roll-out and Retailer Adoption for New Brands," *Marketing Science*, 23:4, 500-518

[10] Brownstone, D. and K. Train (1999), "Forecasting New Product Penetration With Flexible Substitution Patterns," *Journal of Econometrics*, 89, 109-129

[11] Carpenter, G.S., L. Cooper, D.M. Hanssens, and D. Midgley (1988), "Modeling asymmetric competition," *Marketing Scienc*e, 7:4, 393-412

[12] Carpenter, G.S. and Lehmann, D.R. (1985), "A Model of Marketing Mix, Brand Switching, and Competition," *Journal of Marketing Research*, 22:3, 318-329

[13] Chintagunta, P.K. (2001), "Endogeneity and Heterogeneity in a Probit Demand Model: Estimation Using Aggregate Data," *Marketing Science*, 20:4, 442-56

[14] Chintagunta, P.K., Dube, J. P. and V. Singh (2002). "Market Structure Across Stores: An Application of a Random Coefficients Logit Model with Store Level Data," *Econometric Models in Marketing,* 16, 191-221

[15] Elrod, T. (1988), "Choice Map: Inferring a Product Market Map From Panel Data," *Marketing Science*, 7, 21-40

[16] Elrod, T., and M. P. Keane (1995), "A Factor-Analytic Probit Model for Representing the Market Structure in Panel Data," *Journal of Marketing Research*, 32, 1-16

[17] Erdem, T. (1996), "A Dynamic Analysis of Market Structure Based on Panel Data," *Marketing Science*, 15:4, 359-378

[18] Dekimpe, M.G., D.M. Hanssens (1999), "Sustained spending and persistent response: A new look at long-term marketing profitability," *Journal of Marketing Research*, 36(4), 397-412

[19] Fader, Peter S. and Bruce G.S. Hardie (1996), "Modeling Consumer ChoiceAmong SKUs," *Journal of Marketing Research*, 33:4, 442-452

[20] Hansen, L.P. (1982), "Large Sample Properties of Generalized Method of Moments Estimators," *Econometrica*, 50:4, 1029-1054

[21] Holcomb, R. (2000), *Food Industry Overview: Frozen Pizza,* Report FAPC-103, Food and Agricultural Products Center, Oklahoma State University

[22] Imbens, G.W. and T. Lancaster (1994), "Combining Micro and Macro Data in Microeconometric Models," *Review of Economic Studies*, Blackwell Publishing, 61:4, 655-80

[23] Kahn, B. E and L. McAllister (1997), *The Grocery Revolution*, Reading: Mass Addison Wesley

[24] Kalyanaram, G., and G.L. Urban (1992), "Dynamic Effects of the Order of Entry on Market Share, Trial Penetration, and Repeat Purchases for Frequently Purchased Consumer Goods," *Marketing Science*, 11:3, 235-250

[25] MacLeod, W.C. (1995), *Statement of Grocery Manufacturers of America, Inc.*, FTC Hearings on Global and Innovation-Based Competition, http://www.ftc.gov/opp/global/macleod.htm

[26] Nelson, P., J.J. Siegfried, and J. Howell (1992), "A Simultaneous Equations Model of Coffee Brand Pricing and Advertising," *The Review of Economics and Statistics*, MIT Press, vol. 74 (1), 54-63

[27] Nevo, A., (2000), "A Practitioner's Guide to Estimation of Random-Coefficients Logit Models of Demand," *Journal of Economics & Management Strategy*, 9:4, 513–548

[28] Nevo, A. (2001), "Measuring Market Power in the Ready-to-Eat Cereal Industry.", *Econometrica*, 69, March, 2001, 307-342

[29] Petrin, A. (2002), "Quantifying the Benefits of New Products: The Case of the Minivan." *Journal of Political Economy*, 110(4), pp. 705-29

[30] Pizza Marketing Quarterly, *Rising Crust Pizza, Should You Worry?*, http://www.pmq.com/digiorno.shtml.

[31] Van Heerde, H., C.F. Mela and P. Manchanda (2004), "The Dynamic Effect of Innovation on Market Structure," *Journal of Marketing Research,* 41:2, 166-183.

[32] Van Heerde, H., S. Gupta, and D.R. Wittink (2003), "Is 75% of the Sales Promotion Bump Due to Brand Switching? No, Only 33% Is," *Journal of Marketing Research*, 40:4, 481-491

[33] Van Oest, R. and P.H. Franses (2005), "Which Brands Gain Share from Which Brands? Inference from Store-Level Scanner Data," *Quantitative Marketing and Economics*, 3, 281-304