The Evolution of Brand Preferences
Evidence from Consumer Migration

Bart J. Bronnenberg
CentER, Tilburg University

Jean-Pierre H. Dubé
University of Chicago and NBER

Matthew Gentzkow
University of Chicago and NBER

First version: April 26, 2010
This version: August 2, 2010

Abstract

We study the long-run evolution of brand preferences, using new data on consumers' life histories and purchases of consumer packaged goods. Variation in where consumers have lived in the past allows us to isolate the causal effect of past experiences on current purchases, holding constant contemporaneous supply-side factors. We show that brand preferences form endogenously, are highly

If an intelligent being from a remote planet was presented with certain facts about the trivial physical differences in brands and identical prices which exist in many product categories here on earth and asked to develop a model of consumer choice behavior for these conditions, he might assert with little hesitation that: consumers would be indifferent with respect to the available brands, choice would be a random process, and the market shares for the brands would be equal. (Bass 1974) been limited to estimating short-run effects using panel data spanning no more than 1 or 2 years (e.g., Erdem 1996, Keane 1997, Dubé, Hitsch and Rossi 2010).

In this paper, we study the long-run evolution of brand preferences, using a new dataset that combines Nielsen Homescan data on purchases of consumer packaged goods with details of consumers' life histories. Building on Bronnenberg, Dhar, and Dubé's (2007) finding that market shares of these goods with habit formation (Becker and Murphy 1988). Consumers in the model are myopic. Their choices in each period depend on the contemporaneous prices, availability, and other characteristics of the brands in their market, and on their stock of past consumption experiences, or "brand capital." The model has two key parameters: the weight on current product characteristics relative to the stock of past consumption (a), and the year-to-year persistence of brand capital (d).

We next present evidence for two key identifying assumptions. The first is that a consumer's micategories where advertising and visibility are low, suggesting that some element of habit formation is likely necessary to rationalize the data. We also assess how much of the geographic variation in shares not explained by brand capital can be attributed to variation in prices, display advertising, and feature advertising.

Our empirical strategy is closely related to work that uses migration patterns to study the formation of culture and preferences. Logan and Rhode (2010) show that nineteenth-century immigrants' expenditure shares for different types of food are predicted by past relative prices in their countries of origin. Luttmer and Singhal (2010) link immigrants' preferences for redistribution of wealth to the average preference for redistribution in their birth countries. Atkin (2010) shows that migrants within India are willing to pay higher prices to consume foods that are common in their state of origin. Our results also relate to the literature on the formation of preferences more broadly (Bowles 1998). Our work further relates to the broader literature on sources of entry barriers and incumbent advantages (e.g., Bain 1950, Williamson 1963). In particular, Foster, Haltiwanger, and Syverson (2010) show that the demand curves of manufacturing plants shift out over time, and that a model of endogenous demand-side capital formation similar to the one we develop herein can explain a significant share of older plants' size advantage relative to newer plants. Finally, our work relates to the conceptual literature on the long term effects of brand equity in marketing (e.g., Aaker 1991, Keller 1993).

Section 2 introduces our data. Section 3 presents descriptive evidence on the evolution of brand preferences. Section 4 introduces our model and estimation strategy. Section 5 presents evidence supporting our key identifying assumptions. Section 6 presents estimates of the model parameters, and derives implications for first-mover advantage and share stability. Section 7 presents evidence on mechanisms. Section 8 concludes.

2 Data

2.1 Purchases and Demographics

We use data from the Nielsen Homescan Panel on the purchases and demographic characteristics of 48,501 households. The panel is drawn from 50 regional markets throup15(anerlrns)(frowe2ynited)-owe2Statesvers drug stores, and so on. The data cover food, beverages, and many non-food items commonly found in

3 Descriptive Evidence

3.1 Measurement Approach

Index consumers by i, modules by j, and states by s. We focus on the top two brands in each category as defined above. Let i's observed *purchase share* in category j, $\hat{y}_{ij}$, be the number of purchases of brand 1 in category j divided by the total purchases of brands 1 and 2. Let $\hat{m}_{sj}$ be the mean of $\hat{y}_{ij}$ across all non-migrant households in state s.

For each migrant consumer i, we define the *relative share* in category j to be i's purchase share, scaled relative to the average purchase share of non-migrants in her current and birth states:

$$b_{ij} = \hat{y}_{ij}$$

3.2 Cross-Section

Table 2 summarizes variation in purchase shares. The average of the purchase share $\hat{y}_{ij}$ across all consumers and modules in our sample is 0.63. Conditional on purchasing at least one of the top two brands, consumers in the typical category make 3.0 purchases of the top brand and 1.7 purchases of the second-place brand. The cross-state standard deviation of the purchase share is 0.15. The absolute value of the gap between the purchase share in a migrant's current state and in her birth state is 0.11 on average. These geographic differences are broadly consistrcharoadlystasis64(pur213pattern(are)-2reportedpur2134(her)-2Bro

panel lines up with our inferences from the cross-section.

Restricting attention to those for whom the gap between leaving their state of birth and arriving in their current state is zero, we observe 115 consumers who report moving in the past year and 111 consumers who report moving between one and two years ago. Given that our survey was fielded in September 2008, we expect the first group to have moved between October 2007 and September 2008, and the second group to have moved between October 2006 and September 2007.

Figure 5 shows relative shares by month for those who report moving in the past year. Their relative shares for the months up to October 2007 are close to zero, indicating that their purchases before they move are similar to those of non-migrants in their states of birth. If moves are distributed uniformly within the October 2007 to September 2008 period, and if an individual's relative share jumps to 0.62 on moving, we should expect the points to increase linearly from zero to 0.62 in the second half of the figure. This pattern is exactly what we observe.

Figure 6 shows relative shares by month for those who report moving between one and two years ago. As we would expect based on the cross-sectional evidence, relative shares increase roughly linearly from October 2006 to September 2007 and then are flat at 0.62 or slightly increasing thereafter.

4 Model and Estimation

As a lens through which to interpret these results, we introduce a simple model of consumer demand with habit formation (Becker and Murphy 1988). The model serves two purposes. First, it allows us to quantify the preference persistence we observe in terms of an economically meaningful structural parameter: the rate at which the stock of preference "capital" derived from past experience decays. Second, it lets us consider the implications of our results for firms' short-run and long-run demand curves, the importance of first-mover advantage, and the stability of market shares over time.

4.1 Setup

We model a consumer deciding which of the top two brands to purchase in a particular module. We treat states as the relevant product market, assuming that supply-side characteristics of all brands are constant within state. We add subscripts for consumers, modules, and states when we turn to estimation in section 4.3 below.

The difference between the consumer’s indirect utility from the top brand and the second brand is

$$U = am(X;x) + (1 \quad a)k \quad n: \tag{3}$$

Here, $m(X;x)$ 2 (0

write y as a weighted average of past $m(X;x)$ plus a mean zero shock:

$$y_A = \sum_{a=1}^{A} w_a^A m(X;x_a) + e_A \tag{6}$$

where x_a is the vector of product characteristics the consumer faced at age a, $E_n(e_A) = 0$, w_a 2 $[0;1]$, and $\sum_{a=1}^{A} w_a = 1$.

Consider, now, the special case in which product characteristics, $x;$ vary across states but are constant over time. It is immediate that if the consumer has lived in the same state throughout her life, her expected purchase share is simply $y = m(X;x) + e;$ where x are the product characteristics in her current state. Suppose instead that the consumer has moved exactly once: she lived in a state with characteristics x until age a and then moved to a state with characteristics x^0. It is immediate from equation (6) that

$$y_A = b\,m\left(X;x^0\right) + (1 \quad b)\,m(X;x) + e_A: \tag{7}$$

where $b = \sum_{a=a\ +1}^{A} w_a^A$ and, hence, b 2 (

impact" effect of moving. The on-impact effect is the same regardless of the age at which the consumer moved. The remaining 1 a portion of the share gap closes gradually over time as her stock of brand capital adjusts. The adjustment is slower if d is close to one, and if the consumer was older when she moved (since in this case she has accumulated a larger stock of past brand experiences).

The model is restrictive in several important ways. First, we only model the relative utilities of the top two brands. We do not model the extensive margin of whether or not to make a purchase in a module at all, and we suppress substitution with other brands.

Second, we assume that the capital stock, $k;$ and the current demand characteristics, $m(X;x);$ are separable in the indirect utility function. The influence of prices or advertising on indirect utility, and hence on demand, will be the same regardless of a consumer's past experiences. The separability assumption delivers the prediction that the jump in relative share on moving (or "on-impact" effect) is the same regardless of the age at which a consumer moves. We make this assumption for tractability, and because it is consistent with the observed data, as seen in Figure 2.

Third, consumers in our model are myopic. We assume the consumer prefers the top brand to the second brand if and only if U 0. A sophisticated, forward-looking consumer would take account of the way purchases today will affect her capital stock, and thus her expected utility, tomorrow. Demand would therefore depend not only on current product characteristics, but also on expected future product characteristics.

Finally, we assume that the capital stock is a weighted average of past consumption. As discussed above, past experiences could affect present demand through other channels. Past consumption might matter because of learning, and so enter current demand through beliefs rather than preferences. Past exposure to advertising or past observation of peers might matter independently of the level of past consumption. We see our evidence as potentially consistent with all of these stories and our data do not allow us to distinguish them completely. We specialize to a habit model mainly because it is a simple way to capture the key facts. We consider evidence for advertising and peer effects in section 7 below.

4.3 Estimation

Index consumers by i, modules by j, and states by s as in section 3. Index years by t. For each consumer i, we observe a vector of purchase shares with typical element $\hat{y}_{ij}$, a vector of observables X_i, and a vector M_i which encodes i's history of migration—her current and birth state, the age at which she moved (a 11.9552 7 which pool these vectors across i.

We parametrize baseline demand $m()$ as:

$$m(X_i; x_{jst}) = g_{jst} + X_i I_j; \qquad (9)$$

where I is a vector of parameters and g_{jst} is shorthand for the value $g(x_{jst})$ of a function mapping the vector of product characteristics x_{jst} to a scalar. The vector X_i includes log income, as well as dummies for age, Hispanic identity, race, educational attainment, and employment status.

Our first identifying assumption is that there are no unobserved consumer characteristics correlated with both purchases and the exogenous variables M_i and X_i: $E\,($

5 Evidence on Identifying Assumptions

5.1 No Selection on Unobservables

Our first identifying assumption is that there are no unobserved consumer characteristics correlated with both purchase shares, $\hat{y}_{ij}$, and the observables, M_i and X_i.

Of particular concern is the possibility that migrants are selected to have unobserved brand preferences intermediate between the typical non-migrant in their state of birth and their current state of residence. It could also be the case that migrants who stay in a state for many years after moving have characteristics more similar to lifetime residents of that state than migrants who only stay for a few years.

The first test of our identifying assumption is the within-consumer analysis presented in Figures 5 and 6 and discussed in section 3 above. We see that the migrants look similar to non-migrants in their birth states in the months before they move. The mean relative share pooling months 10/06 to 9/07 for migrants living in their current state less than a year is 0.093, the 95 percent confidence interval is $(\ 0.025, 0.211)$, and we fail to reject $b = 0$ at the 10 percent level ($p = 0.12$). The data are also consistent with a discrete jump in migrant purchases on moving. Moreover, purchase shares for these consumers prior to moving are not significantly related to the age at which they moved ($p = 0.37$), providing no support for the hypothesis that the correlation between relative shares and age at move or years since moving in Figure 2 is primarily driven by selection on unobservables.

As a second test of our identifying assumption, we consider a sub-sample of brands that were introduced relatively recently. Under the assumptions of our model, a migrant who moved before either of two brands was introduced should have an expected purchase share no different from non-migrants in her current state of residence. If the identifying assumption was violated, where a consumer lived before the brands UnderasId no difcurrent state9 0 Td [u6639 where T_w is the number of years at least one brand in pair w has been available, t_i is the number of years since i moved, and $I()$ is the indicator function. We weight observations by $\left(\hat{m}_{s'j} - \hat{m}_{sj}\right)^2$ as in equation (2) above. Under our identifying assumption, we expect $w_1 > 0$, $w_2 = 1$, and $w_3 = 0$.

Table 4 presents the results. Consistent with our assumption, the coefficient on decades since moving is highly significant for those moving after the pair in question was introduced ($w_1 > 0$), but insignificant for those moving before the pair was introduced ($w_3 \approx 0$). Moreover, we cannot reject that the average shares of migrants who moved before the pair was introduced have the same average shares as non-migrants in their current state of residence ($w_2 \approx 1$). The results are robust to focusing on the complete set of pairs introduced since 1955, pairs introduced after 1975, and pairs introduced after 1985.

5.2 Expected Past Shares Equal Present Shares

Our second identifying assumption is that, conditional on observables, the expectation of baseline demand in a given module-state pair in any past year is equal to the expectation in the current year.

To test this assumption, we study the 27 modules for which we observe purchases of both current top-two brands in the historical CCA data. For each module-state pair, we compute the current purchase share in the Homescan data across both migrants and non-migrants. We then compare this share to the analogous share in the CCA data for the years 1948-1968, computed as described in section 2.3 above. Under our identifying assumption, we expect that the regression of past shares on current shares should have an intercept of zero and a slope of one.

Note that this prediction would only hold exactly if we compared past and current purchases of non-migrants. We cannot perform this test, because the CCA data do not report shares by migration status. The regression of past on current shares will still be informative, however, so long as migrants are a relatively small share of the population and/or migration patterns have been relatively stable over time.

Figure 7 presents a scatterplot of current versus past purchase shares. Each observation is a state-module pair. The diameters of the circles are proportional to the number of years of CCA data we have for the observation. The current and past shares are clearly not equal, possibly reflecting real changes in market structure over time as well as sampling variability. However, the fitted values, indicated by the dotted line, are very close to the 45-degree line.

Table 5 presents the corresponding regression of past shares on current shares, weighting by the number of years of CCA data, and clustering by module. The estimated constant is 0.084 and the estimated slope is 0.822. We cannot reject the joint hypothesis that the constant equals zero and the slope equals one ($p = 0.30$).

A possible concern is that the coefficient in this regression may be attenuated by measurement error in the current shares. Consistent with this hypothesis, restricting the regression to state-module pairs where

In Appendix C, we present auxiliary estimates using store-level price and quantity data from IRI that allow us to give an alternative interpretation of our shelf-space counterfactuals in terms of relative price changes. Pooling across 30 categories, we estimate an average demand elasticity of substitution of $\frac{\partial log(\frac{y_A}{y_B})}{\partial log(\frac{price_A}{price_B})} = -1.54$.

Suppose that A has a head start of 5 years. During this period, $y = 0$ as all consumers buy brand A. The accumulated capital stock at the end of those 5 years is $k = 0$. Brand B then enters and the two firms play a game that determines shelf space allocations. Abstracting from the details of this game, we know that if space allocations are equal ($x = 0.5$), we will have $y < 0.5$, and y will converge toward 0.5 but never reach it. Brand B will, thus, never achieve parity in the purchase share. If B has the majority of shelf space ($x > 0.5$), both y and k will reach 0.5 in some finite number of years. The larger is x, the faster the convergence. We can therefore ask how many years B would need to maintain a certain share of shelf space, x, to achieve purchase share parity.

More generally, we assume brand A's head start is $t \in \{1, 5, 10, 15, 25\}$ years and ask how fast the second firm achieves convergence using a level of $x \in \{0.55, 0.60, 0.65, 0.70, 0.75\}$. From the estimates in Appendix C, these shelf-space allocations are equivalent to price discounts of $1 - p_B/p_A \in \{0.08, 0.15, 0.22, 0.28, 0.34\}.$ Over sufficiently long horizons, it is important to account for the fact that some consumers will die (destroying some of A's capital) and others will be born (with much less of A's capital). We run the simulations assuming that the age distribution is stable over time and matches the empirical distribution we observe in our Homescan sample.

Table 7 shows the required number of years to catch up. The results show that at the estimated α and δ, equalizing shares in a reasonable amount of time requires significant investment. If A's head start is 5 years, B would need to hold 60 percent of shelf space (or discount its price by 15 percent) to reach market share parity in just more than a decade. To catch up in only 2 years, B would need to hold three quarters of shelf space or discount its price by more than 30 percent. If A's head start were 15 years, B would require 23 years at 60 percent of shelf space, or 3 years at 75 percent of shelf space, to reach market share parity.

6.4 Persistence under Market Shocks

Bronnenberg, Dhar, and Dubé (2009) show that regional share differences in consumer packaged goods industries persist over remarkably long periods of time. Current local shares are strongly predicted by who was the first entrant in a market, even when that entry happened a century ago, few consumers alive remember a time when both brands were not widely available, and the intervening years have seen large shocks to the economic environment such as the growth of supermarkets, changes in real income, wars, depression, and so on.

Our model does not predict how much persistence we should expect to see because it does not endogenize firm choices. The previous section showed that a second entrant would have to make large investments to catch up to the first entrant; it does not say anything about whether or not we will see those investments in equilibrium. In this section, we consider a specific assumption under which our model does have strong implications about persistence: complementarity between the stock of capital (k) and current investments in gaining market share (x).

In particular, extend the example of the previous section and suppose that supermarkets allocate shelf space in proportion to expected market share. That is, the shelf space allocation in period t is

$$x_t = \frac{1}{N_t}\sum_i y_{it}; \tag{12}$$

where N_t is the number of consumers in the market. Allocating shelf space proportional to market share is in fact a common rule of thumb for retailers, and one that some argue will be approximately optimal.[7] Such a rule will lead intuitively to persistence in shares because a brand that has a lead in the capital stock of experienced consumers will have a larger share of shelf space and consequently be purchased more often even by inexperienced consumers.

We ask how much persistence this dynamic can explain in the presence of shocks to the two brands' shares in each period. As above, we assume $m(X;x) = x$, where x is the share of shelf space allocated to brand B and is given by equation (12). Expected purchase shares are:

$$y_{it} = a x_t + (1 \quad a) k_{it} + k_t; \tag{13}$$

where k_t is an i.i.d. shock distributed uniformly on $[\ \bar{k};\bar{k}]$. Because of transmission through the capital stock, k_{it}, y_{it} depends on both past and present shocks.

We assume an existing market share for the leading brand of $0{:}75$, which has been in place for as long as consumers live. We fix $a = 0{:}62344$ 0 Td [(0)]TJ/F120 1,;,;

at the upper end of typical annual share movements in consumer packaged goods.[8] We then forward simulate 100 years of evolution for our hypothetical market.

Figure 10 plots the distribution of the market shares in the final year of the simulation across 1000 replications. The first panel shows that when we fix d at its estimated value ($0{:}974$), long-run market shares remain closely concentrated around their initial value of 0.75, even after 100 years of shocks. The probabilities that market shares are within 10 or 20 share points of their initial value after 100 years are 72 percent and 100 percent respectively. The mechanism generating the persistence is the recency-weighted window of past experiences in the consumer's brand capital stock. Within this window, shocks tend to cancel out over time. It is, thus, the stock of brand capital that buffers against the reinforcement of demand and supply shocks. The weaker the brand stock, the more market shares are subject to exogenous shocks that accumulate across time. Accordingly, the persistence weakens when we consider lower values for d and, effectively, shorten the relevant window of past experiences. The probability that market shares are within 10 share points of the initial values drops from 72% with $d = 0{:}974$, to 22% with $d = 0{:}224$, which is barely above the 20 percent one would expect if shares after 100 years attain a uniform distribution. As d decreases towards 0, historical advantages are all but erased.

From this simple simulation, we conclude that our estimates of preference persistence, combined with complementarity between current investment and brand capital, can rationalize stable market shares over long periods of time even in the presence of large shocks.

7 Mechanisms

7.1 Brand Capital

We estimate that 40 percent of current geographic variation in purchase shares is explained by variation in consumers' brand capital stocks. For tractability and ease of exposition, we have modeled brand capital formation in a habit framework, assuming the current capital stock is a function only of past consumption. As mentioned in the introduction, however, the brand capital stock may be partly a function of other variables, such as past exposure to advertising (Schmalensee 1983, Doraszelski and Markovich 2007), or past observations of consumption by peers (Ellison and Fudenberg 1995).

[8] Under the allocation in Equation (12), observe that equation (13) can be aggregated to $y_t = a y_t + (1 \quad a) \int_i k_{it} f(i) di + k_t$, where $f(i)$ is the age distribution in the population. Rearranging this aggregation, we obtain $y_t = \int_i k_{it} f(i) di + k_t / (1 \quad a)$. Hence, taking into account the allocation rule, the shocks on market shares are uniformly distributed on $[\;\bar{k}/(1 \quad a); \bar{k}/(1 \quad a)]$ t $[\;0{:}12; +0{:}12]$ at our estimated value for $a{:}$

To provide a first look at the mechanism behind brand capital, we ask how our parameter estimates depend on whether a category has high or low levels of advertising. Recall that we define a category to have high advertising if total expenditure by the top two brands is greater than the 75th percentile among all categories in our dataset. We re-estimate our main model allowing both the weight on brand capital $(1 \quad a)$ and the rate of persistence in brand capital d to differ by advertising intensity.

We also divide categories by the extent to which their consumption is socially visible. We code this measure subjectively. We judge products to be socially visible if (i) they are frequently consumed together with others in social situations, and (ii) they are frequently consumed or served directly from a package with the brand name visible. Products such as beer, soda, chips, ketchup, and cigarettes are therefore coded as socially visible. Products such as baby food, toothpaste, and cold remedies are not socially visible because they fail criterion (i). Products such as gravy mixes, frozen pasta, and shredded cheese are not socially visible because they fail criterion (ii). See Appendix Table 2 for the module-by-module coding.

As with advertising, we allow both $(1 \quad a)$ and d to differ by social visibility. Note that the correlation between the dummy for high advertising and the dummy for high visibility is low, so the sample splits by advertising and visibility should capture independent variation.

Table 8 presents the results. We find that advertising-intense categories have a significantly lower value of $a;$ and thus a significantly larger weight on the brand capital stock in utility. We cannot interpret this difference as causal, but it is consistent with the stock of past advertising exposure influencing current willingness to pay above and beyond the effect of past consumption. We find no significant differences in d, consistent with the influence of past consumption and past advertising decaying at a similar rate.

We see a similar pattern with social visibility. We find that categories with a high degree of social visibility have a smaller estimated a, implying greater weight on brand capital. This finding is consistent with past observations of peer consumption exerting an independent influence on current willingness to pay. We again find no significant difference in d.

7.2 Baseline Demand

The remaining 60 percent of geographic variation in purchase shares is driven by differences in baseline demand $m(X;x)$. Recall that the source of this result is the observation that when migrants move, their consumption shifts immediately toward the dominant brand in the destination market, closing 60 percent of the gap in purchase shares. It must be that migrants encounter some combination of lower prices, higher advertising, widespread availability, or other advantages of the dominant brand that lead to this jump in consumption. The results above do not speak to the role of specific supply-side variables, however.

We can use the aggregate IRI data to get some feel for the role of prices, display advertising, and feature advertising. Details of this exercise are provided in Appendix D. First, for each category, we compute the share of cross-market variation in the log difference in purchase shares explained by the following independent variables: (i) log relative prices, (ii) relative display intensity, (iii) relative feature intensity, and (iv) log relative prices, display intensity, and feature intensity together. We then compute the mean and standard deviation of these shares across categories.

We find that the cross-market correlation between relative shares and prices is 0.50 in the average category. The average share of variance explained by prices is 32 percent. Clearly, one reason migrants adjust their purchases immediately on moving is that they encounter lower prices. We find that the cross-market correlation of relative shares with feature and display advertising is 0.44 and 0.42 respectively, explaining 28 percent and 24 percent of cross-market variation on average. Migrants also encounter more features and displays for the dominant brand. Together, prices, feature, and display explain 49 percent of the cross-market variation in the average category.

If prices, feature, and display are correlated with other market-level product characteristics such as shelf space allocations, however, these regressions will overstate the share of variation explained. To address this issue, we exploit the panel structure of our data. For each category, we regress the log difference in purchase shares at the category-market-week level on market and week dummies, plus each of the independent variables above. From each of these regressions, we compute predicted values by multiplying the independent variable(s) of interest by their estimated coefficient(s). We estimate the share of variance explained by dividing the variance of the predicted value by the variance of the dependent variable. Finally, we compute the mean and standard deviation of the estimated shares across categories.

From these specifications, we estimate that variation in relative prices explains 20 percent of cross-market variation ($std.dev. = 13$ percentage points). Variation in relative feature intensity explains 7 percent ($std.dev. = 5$ percentage points), variation in relative display intensity explains 11 percent ($std.dev. = 9.8$ percentage points), and all three marketing variables together explain 21 percent ($std.dev. = 12$ percentage points).

A candidate variable we are unable to measure is shelf space allocation, or availability more broadly.

Marketing models used in practice to determine shelf space allocations often recommend that they be proportional to market share (Bultez and Naert 1988). To the extent that shelf space exerts a significant effect on consumption, shelf space could explain a significant share of the remaining variation.

Finally, it is possible that baseline demand depends in part on the observed consumption of others. This role for peer effects differs from the contribution to the brand capital stock discussed above. It would imply we might expect to see faster adjustment (higher a) for highly visible categories. As already discussed, Table 8 shows the opposite is true. This could mean that peer effects are not an important contributor to baseline demand, or that this effect is outweighed by their contribution to brand capital.

8 Conclusions

Our results suggest that much of consumers' observed willingness to pay for brands may reflect the influence of past experiences. We estimate that heterogeneity in brand capital explains a substantial share of geographic variation in purchases. Brand capital evolves endogenously as a function of consumers' life histories, and decays slowly once formed. Brand capital can explain large and long-lasting advantages to first movers. Finally, our results suggest that brand preferences play an especially important role in categories with high levels of advertising and social visibility.

[17] Doraszelski, Ulrich, and Sarit Markovich (2007), "Advertising Dynamics and Competitive Advantage," *Rand Journal of Economics*, 38(3), 557-92.

[18] Drolet, Aimee, Patrick Suppes, and Anand V. Bodapati (2008), "Habits and Free Associations: Free Your Mind but Mind Your Habits," Working Paper, University of California, Los Angeles.

[19] Dubé, Jean-Pierre H., Günter J. Hitsch, and Peter E. Rossi (2010), "State Dependence and Alternative Explanations for Consumer Inertia," forthcoming in the *RAND Journal of Economicse*

[32] Luttmer, Erzo F. P., and Monica Singhal (2010), "Culture, Context, and the Taste for Redistribution," forthcoming in the *American Economic Journal: Economic Policy*.

[33] Moore, Elizabeth S., William L. Wilkie, and Richard J. Lutz (2002), "Passing the Torch: Intergenerational Influences as a Source of Brand Equity," *Journal of Marketing*, 66(2), 17-37.

[34] Schmalensee, Richard (1982), "Product Differentiation Advantages of Pioneering Brands," *American Economic Review*, 72(3), 349-365.

[35] —- (1983), "Advertising and Entry Deterrence: An Exploratory Model," *Journal of Political Economy*, 91(4), 636-653.

[36] Smith, Michael D., and Erik Brynjolfsson (2001), "Consumer Decision-Making at an Internet Shopbot: Brand Still Matters," *Journal of Industrial Economics*, 49(4), 541-558.

[37] Thumin, Frederick J. (1962), "Identification of Cola Beverages," *Journal of Applied Psychology*, 36(5), 358-360.

[38] Williamson, Oliver E. (1963), "Selling Expense as a Barrier to Entry," *Quarterly Journal of Economics*, 77(1), 112-128.

Appendix

A Derivation of Equation (8)

We first write y_{A+1} recursively as a function of y_A. Define $z_a = \hat{y}_a \quad y_a$. For any $A > a\ ;$ we can expand equation (4) as:

$$y_A = a\, m(X;x^0) + (1 \quad a)\frac{\sum_{a=1}^{A\ 1} d^{A\ \ a}(y_a+z_a)}{\sum_{a=1}^{A\ 1} d^a} \tag{A.1}$$

Combining equation (A.1) with the analogous expression for y_{A+1} we can show that :

$$y_{A+1} = a\, m(X;x^0)\frac{d}{\sum_{a=1}^{A} d^a} + \left(1 \quad a\frac{d}{\sum_{a=1}^{A} d^a}\right) y_A + \frac{(1 \ \ a)}{\sum_{a=1}^{A} d^a} d z_A \tag{A.2}$$

Next, we write $b(a\ ;t\ \ +1)$ as a function of $b(a\ ;t\ )$. We know from equation (7) that for each a and t there exists $b(a\ ;t\ )$ such that

$$y_A = b(a\ ;t\ )\, m\left(X;x^0\right) + (1 \quad b(a\ ;t\ ))\, m(X;x) + e_A;$$

Using this fact along with equation (A.2), we can show that:

$$b(a\ ;t\ \ +1) = \quad \text{aa +x}$$

C Estimation of Elasticity of Substitution using IRI Data

We use aggregate store-level data on 2001-2005 purchases and prices from the IRI Marketing Data Set (Bronnenberg, Kruger, and Mela 2008) to estimate the average elasticity of substitution between the top two brands in a typical consumer packaged goods category. These data cover sales in 30 consumer packaged goods categories for 260 weeks across 47 markets. We use total volume by brand-market-week as our measure of purchases. We compute prices by dividing expenditure for each brand-market-week by volume. We focus on the top two brands in each category by total volume across all markets and weeks. For the top two brands in category j, P_{1jmt} and P_{2jmt} are prices, F_{1jmt} and F_{2jmt} are the feature advertising intensity levels, D_{1jmt} and D_{2jmt} are the display advertising intensity levels, and y_{jmt} is the

We first collapse the data to the category-market level by taking means across weeks of each variable. We then estimate the raw cross-market correlation in each category between the log ratio of shares and each marketing variable. We also run a regression in each category of the log ratio of shares on all three marketing variables jointly and compute the R^2. We report the mean and standard deviation of the correlation and R^2 across categories.

To address spurrious correlation between these marketing variables and time-constant unobservables, we also estimate panel regressions with market and week fixed effects for each category. From each of these regressions, we compute predicted values by multiplying the independent variable(s) of interest by their estimated coefficient. We estimate the share of variance explained by dividing the share of the predicted value by the total variance of the dependent variable. Finally, we compute the mean and standard deviation of the estimated shares across categories.

[redacted]

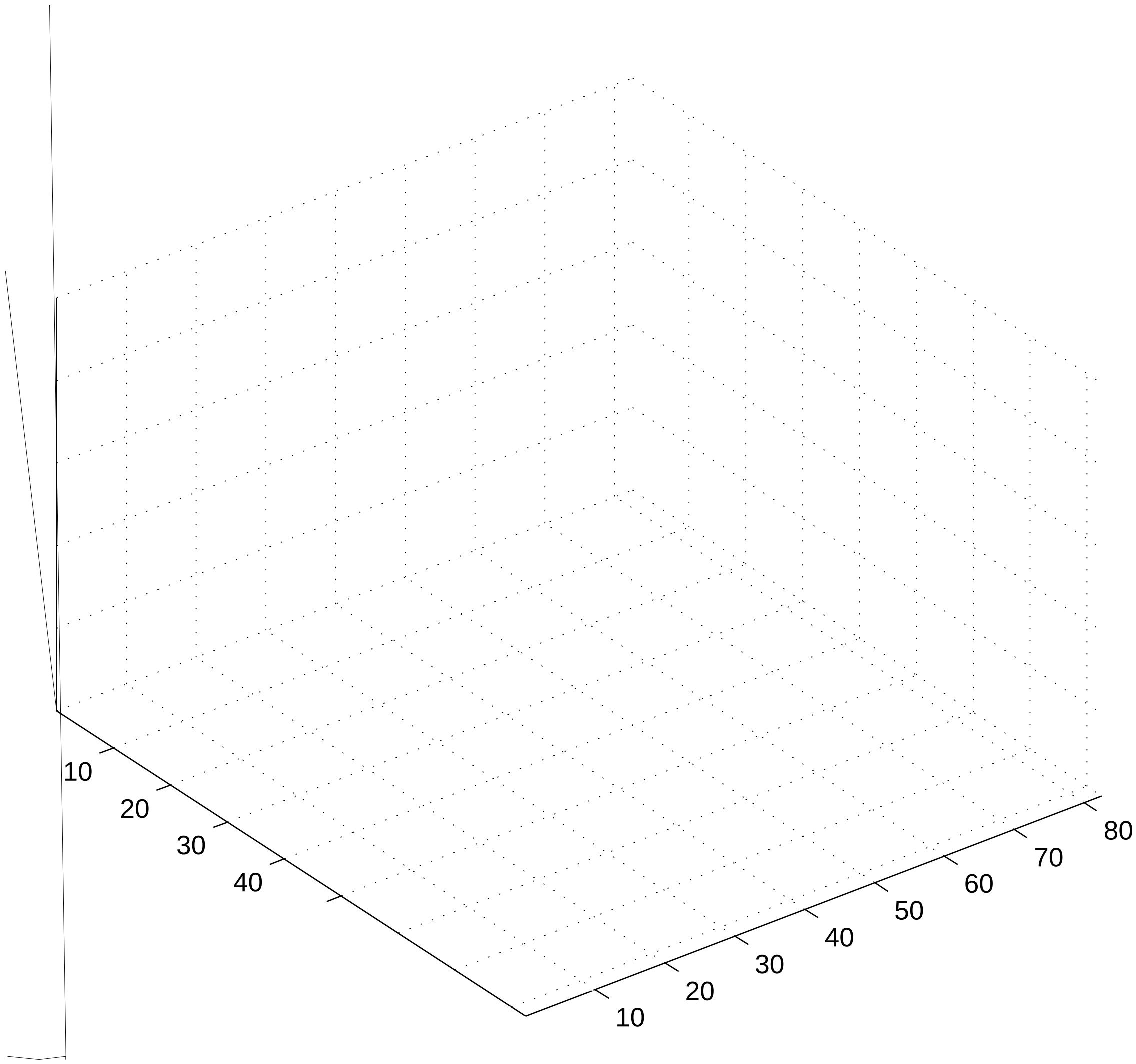

Figure 2: Relative Shares by Age at Move and Years Since Move

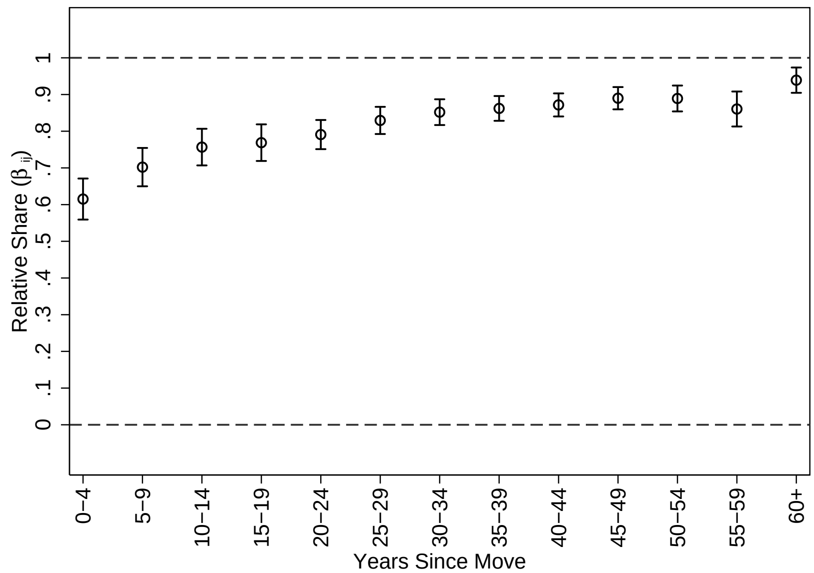

Figure 3: Relative Shares by Years Since Move

Notes: Whiskers indicate 95% confidence intervals. Standard errors clustered by module.

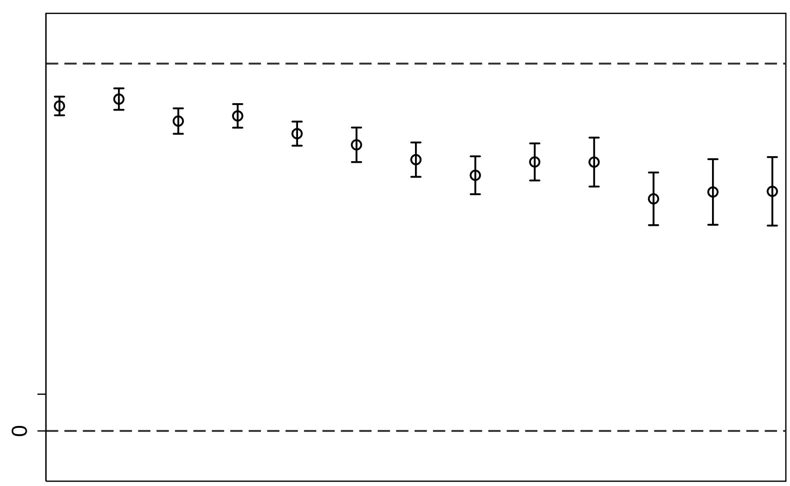

Figure 4: Relative Shares by Age at Move

Notes: Whiskers indicate 95% confidence intervals. Standard errors clustered by module.

Figure 5: Relative Shares by Month (Moved 10/07-9/08)

Notes: Whiskers indicate 95% confidence intervals. Standard errors are clustered by module. The sample consists of

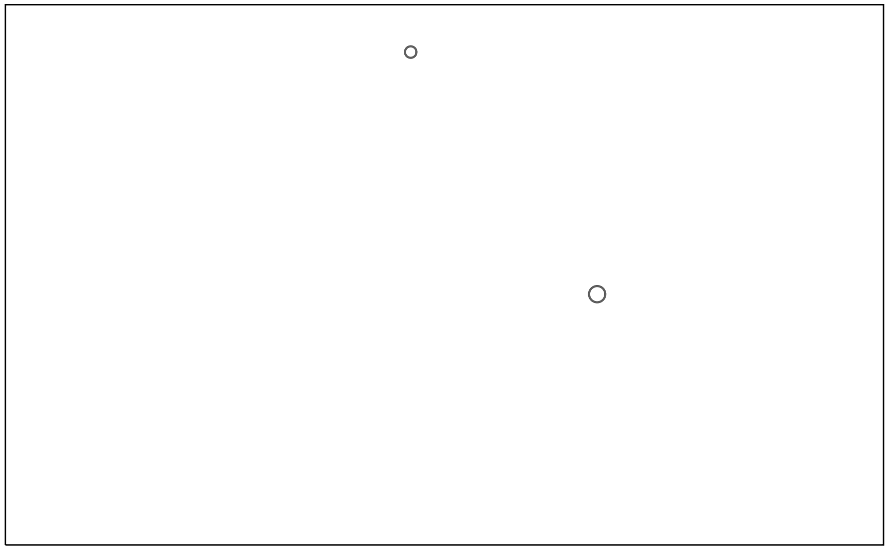

Figure 7: Historical and Current Purchase Shares

Notes: Each observation is a state-module pair. The y axis is average purchase share between 1948 and 1968, calculated using Consolidated Consumer Analysis. The x axis is the average purchase share in the 2006-2008 Homescan sample. The size of the circles indicates the number of years of CCA data used to calculate the historical purchase share. See section 5.2 for details.

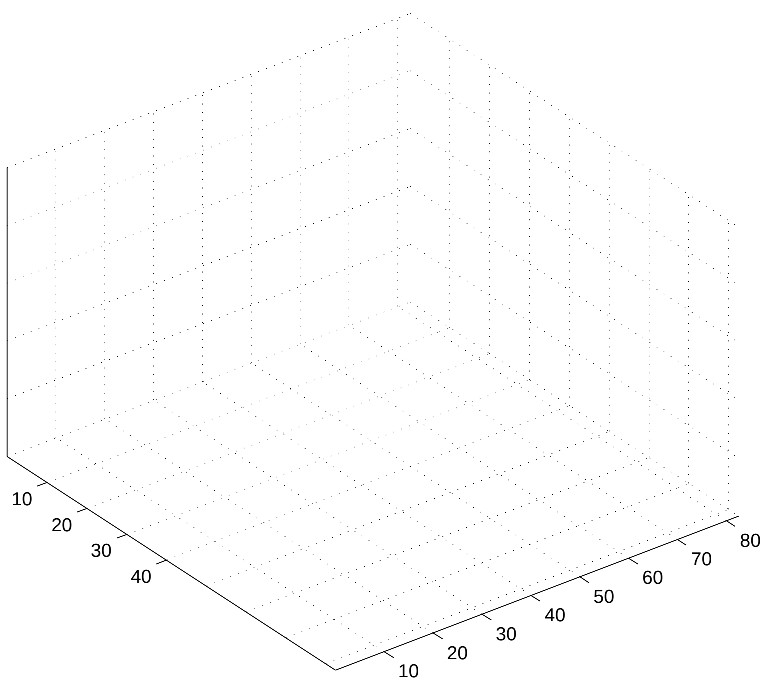

Figure 8: Relative Shares (Fitted Values)

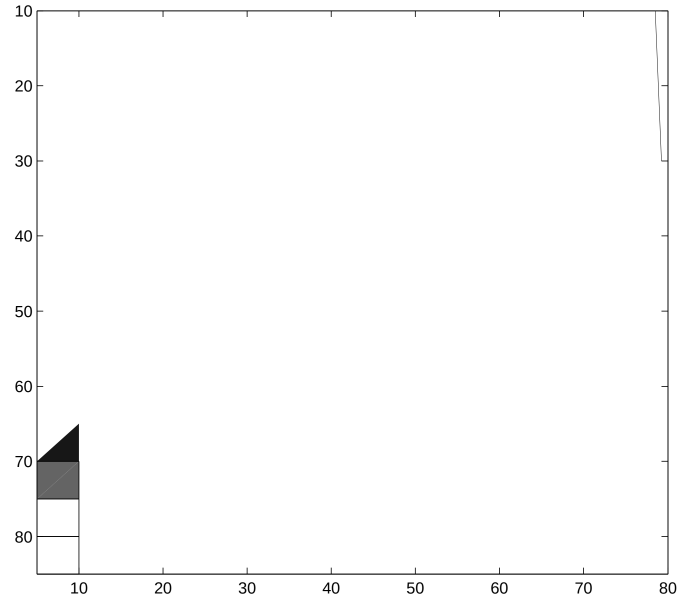

Figure 9: Relative Shares (Residuals)

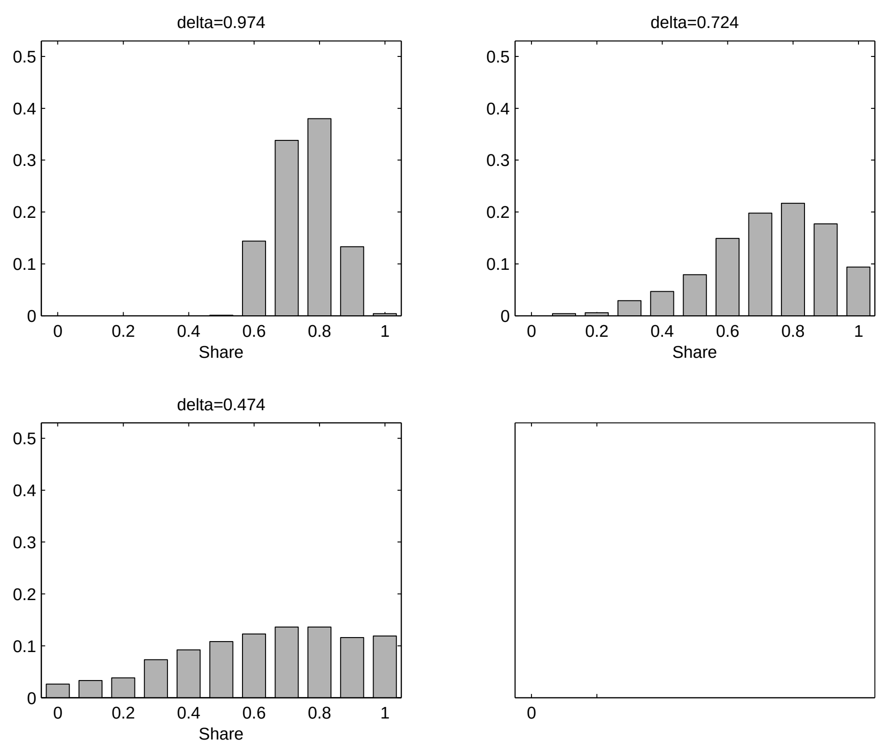

Figure 10: Persistence of Market Shares under Exogenous Shocks

Notes: Each panel contains the distribution of the long run predictions of market shares, $y_t;$ at $t = 100$. To initialize the stock of brand capital in each age cohort, market shares are assumed to have a life-time history of $y_t = 0{:}75$ for $100 < t < 1$, so that the stock of brand capital has formed at $0{:}75$ properly in each age cohort in our empirical sample. Distributions are computed across 1000 draws of the random shocks $k_t;$ $[t = 1; :::; 100]$. Comparisons across panels show the effect of the degree of persistence, $d;$ in brand capital on long run stability of market shares subject to demand and supply shocks.

Table 1: Migration Patterns

Region of birth	Region of residence			
	North East	Midwest	South	West
North East	6765	269	1539	448
Midwest	165	10654	1377	885
South	193	435	9725	292
West	56	214	341	4740

Notes: Table shows the number of households in the Nielsen Homescan sample by census region of birth and current residence.

Table 2: Summary Statistics for Final Sample

# Categories	238

Table 3: The Evolution of Brand Preferences for Migrants

Dependent variable:	Relative share (b_{ij})				
	(1)	(2)	(3)	(4)	(5)
Decades since move	0.098	0.079	0.075	-	0.092
	(0.009)	(0.009)	(0.010)	-	(0.016)
Decades since move	-0.009	-0.008	-0.007	-	-0.010
squared	(0.001)	(0.001)	(0.001)	-	(0.004)
Age (in decades)	-	-0.018	-	-0.019	-0.013
when moved	-	(0.005)	-	(0.005)	(0.008)
Constant	0.624	0.705	-	-	0.668
	(0.029)	(0.026)	-	-	(0.037)
Decades since move fixed effects	no	no	no	yes	no
Age when moved fixed effects	no	no	yes	no	no
Sample	all	all	all	all	age moved 25
# modules	238	238	238	238	238
# HH-module observations	528621	528621	528621	528621	212957

Notes: The dependent variable b_{ij} is the share of a migrant's top-two brand purchases going to the top brand, scaled relative to non-migrants in her current and birth states. $b_{ij} = 1$ implies her purchase share matches non-migrants in her current state. $b_{ij} = 0$ implies her purchase share matches non-migrants in her birth state. See section 3 for details.

Table 4: Brand Pairs Introduced after 1954

Dependent variable: Relative share (b_{ij})	(1)	(2)	(3)
Moved after brand introduced:			
Decades since move (w_1)	0.007	0.007	0.018
	(0.002)	(0.003)	(0.005)
Constant (w_0)	0.657	0.701	0.693
	(0.055)	(0.075)	(0.090)
Moved before brand introduced:			
Decades since move (w_3)	0.001	0.001	0.001
	(0.002)	(0.002)	(0.002)
Constant (w_2)	0.854	0.852	0.880
	(0.100)	(0.101)	(0.101)
Only brand pairs introduced after	1954	1975	1985
# brand pairs	52	24	11
# HH-pair observations	86805	43083	22088

Notes: The dependent variable b_{ij} is the share of a migrant's top-two brand purchases going to the top brand, scaled relative to non-migrants in her current and birth states. $b_{ij} = 1$ implies her purchase share matches non-migrants in her current state. $b_{ij} = 0$ implies her purchase share matches non-migrants in her birth state. The sample includes purchases of brand pairs introduced in 1955 or later. The coefficients in the first two rows apply to migrants who moved after the first brand in the pair in question was introduced. The coefficients in the following two rows apply to migrants who moved before the first brand in the pair was introduced. See section 5.1 for details.

Table 5: Current and Historical Purchase Shares

Dependent variable: Purchase share 1948-1968	(1)	(2)	(3)
Current purchase share	0.822	0.926	1.039
	(0.119)	(0.105)	(0.089)
Constant	0.084	0.027	0.001
	(0.082)	(0.077)	(0.080)
Only include obs. if # Homescan HHs	0	200	500
p-value for (coeff=1) & (cons=0)	0.300	0.746	0.793
# Modules	27	25	21
# State-module obs.	325	188	115

Table 7: First Mover Advantage

Investment years to equate shares

	Shelf Space Investment (x) by Second Entrant				
First Entrant's Head Start (t)	0.55	0.60	0.65	0.70	0.75
1 year	10	4	2	1	1
5 years	27	12	6	3	2
10 years	34	19	10	5	2
15 years	36	23	13	7	3
25 years	38	26	17	9	4
Price Discount by Second Entrant Equivalent to this Shelf Space Investment (x)	8%	15%	22%	28%	34%

Appendix Table 1: Robustness of Structural Parameters

Appendix Table 2: Modules, Top Two Brands, and Selected Module Characteristics

Module	Brand 1	Brand 2	Aggregate Purch Share	Cross-State SD	Ad Intense	Socially Visible
Abrasive Clnsr-Liq	Soft Scrub	Comet	.90	.07	0	0
Abrasive Clnsr-Pwdr	Comet	Ajax	.78	.08	0	0
Adult Incont. Prod	Poise	Tena Serenity	.68	.15	0	0
Analgesic/Chest Rubs	Icy Hot	Vicks Vaporub	.55	.12	0	0
Antacids	Prilosec	Rolaids	.71	.08	1	0
Anti-Gas Products	Beano	Gas-X	.52	.13	0	0
Auto. Dishwshr Cmpnd	Cascade	Electrasol Jet-Dry	.73	.08	0	0
Baby Food-Strained	Gerber	Beechnut Stages	.70	.17	0	0
Bakery Bagels	Thomas'	Sara Lee	.74	.29	0	0
Bakery Bfast Rolls	Little Debbie	Entenmann's	.64	.24	0	0
Bakery Bread	Nature's Own	Sara Lee Soft & Smth	.50	.32	0	0
Bakery Buns	Sara Lee	Wonder	.61	.32	0	0
Bakery Cakes	Little Debbie	Hostess	.91	.07	0	0
Bakery Cheesecake	The Father's Table	Cheesecake Factory	.59	.24	0	0
Bakery Doughnuts	Hostess	Entenmann's	.52	.27	0	0
Bakery Misc.	Homestyle	Flatout	.51	.26	0	0
Bakery Pies	Little Debbie	JJ's	.52	.29	0	0
Bakery Rolls	King's Hawaiian	Martin's	.51	.36	0	0
Baking Cups & Liners	Reynolds	Wilton	.78	.07	0	0
Bath Additive-Liq	Lander	Mr. Bubble	.73	.20	0	0
Beer	Budweiser	Miller High Life	.64	.19	1	1
Bouillon	Wyler's	Knorr	.61	.25	0	0
Breath Sweetener	Tic Tac	Breath Savers	.72	.07	0	1
Butter	Land O Lakes	Challenge	.86	.27	0	0
Candy-Choc Minis	M&M Mars Snickers	Reese's Pnt Bttr Cup	.51	.07	0	1
Candy-Chocolate	M&M Mars M&M Plain	Reese's Pnt Bttr Cup	.52	.06	1	1
Candy-Diet. Non Choc	Life Savers	Baskin-Robbins	.68	.14	0	1
Candy-Dietetic Choc	Russell Stover	Whitman's Wgt Wtchrs	.81	.14	0	1
Candy-Hard Rolled	Pez	Smarties	.52	.11	0	1
Candy-Lollipops	Tootsie Roll Pops	Spangler Dum Dum Pop	.67	.11	0	1
Candy-Non Choc Minis	Tootsie Roll	M&M Mars Skittles	.76	.08	0	1
Candy-Non Chocolate	Y&S Twizzlers	Just Born	.51	.13	0	1
Candy-Special Choc	Hershey's Kisses	Russell Stover	.54	.07	0	1
Caramel Corn	Crunch 'n Munch	Cracker Jack	.71	.09	0	1
Cat Food-Dry	Meow Mix	Purina Cat Chow	.50	.07	0	0
Catsup	Heinz	Hunt's	.66	.13	0	1
Cereal-Dry	G M Cheerios	Post Hny Bnchs Oats	.54	.07	1	0
Cereal-Granola	Sunbelt	Nature Valley	.55	.16	0	0
Cheese-Amrcn Cheddar	Kraft	Cracker Barrel	.66	.33	0	0
Cheese-Amrcn Colby	Kraft	Crystal Farms	.81	.23	0	0
Cheese-Grated	Kraft	4C	.92	.06	0	0
Cheese-Misc.	Kraft	Sargento	.66	.12	0	0
Cheese-Mozzarella	Frigo Cheese Heads	Kraft Snkbls Polly-O	.68	.18	0	0
Cheese-Muenster	Sargento	Finlandia	.79	.22	0	0
Cheese-Shredded	Kraft	Sargento	.72	.15	0	0

Module	Brand 1	Brand 2	Aggregate Purch Share	Cross-State SD	Ad Intense	Socially Visible

Module	Brand 1	Brand 2	Aggregate Purch Share	Cross-State SD	Ad Intense	Socially Visible
Eye Drops & Lotions	Visine	Alcon Systane	.53	.16	0	0
Facial Tissue	Kleenex	Puffs	.63	.07	1	0
Floor Care Cleaner	Swiffer Wet Jet	Clorox Ready Mop	.87	.11	1	0
Foot Cmfrt Products	Gold Bond	Dr Scholl’s	.63	.18	0	0
Foot Prepn-Athlts Ft	Lamisil AT	Tinactin	.54	.21	0	0
Foot Prepn-Misc.	Dr Scholl’s	Pro Foot	.85	.07	0	0
Frozen Dinners	Banquet	Healthy Chc Cmpt Slc	.67	.09	1	0
Frozen Pot Pies	Banquet	Marie Callender’s	.52	.11	0	0
Frozen Snacks	Totino’s	Superpretzel	.76	.12	0	0
Fruit Drinks-Misc.	Minute Maid	Tropicana	.65	.18	0	1
Fruit Juice-Misc.	Dole	Tropicana	.78	.14	0	1
Fruit Juice-Orange	Tropicana	Minute Maid	.67	.16	0	1
Fruit Spread	Smucker’s Simply Frt	Polaner	.52	.27	0	1
Frzn Asian Entrees-1	Weight Watchers	Tai Pei	.59	.17	0	0
Frzn Asian Entrees-2	Lean Csn Cafe Clsscs	Banquet	.56	.13	0	0
Frzn Italn Entrees-1	Weight Watchers	Bertolli	.63	.08	1	0
Frzn Italn Entrees-2	Weight Watchers	Healthy Chc Simp Slc	.51	.17	1	0
Frzn Meat Entrees-1	Banquet	On-Cor	.56	.22	0	0
Frzn Meat Entrees-2	Lean Csn Cafe Clsscs	Boston Market	.51	.13	0	0
Frzn Mexcn Entrees-1	El Monterey	Jose Ole	.67	.16	0	0
Frzn Mexcn Entrees-2	Weight Watchers	Banquet	.60	.18	0	0
Frzn Misc. Entrees-1	Stouffer’s	Mrs. T’s	.57	.18	1	0
Frzn Pltry Entrees-1	Tyson	Banquet	.68	.10	0	0
Frzn Pltry Entrees-2	Weight Watchers	Boston Market	.62	.15	0	0
Frzn Seafd Entrees-1	Gorton’s	Weight Watchers	.64	.16	0	0
Gelatin Salad-Refrig	Jell-O Ref	Winky Ref	.89	.09	0	0
Gravy Mix	McCormick	Pioneer	.74	.15	0	0
Gravy-Canned	Heinz Homestyle	Campbell’s	.56	.12	0	0
Gum-Bubble	Dubble Bubble	Adams Bubblicious	.73	.11	0	1
Hair Color-Women’s	Clairol Nice ’n Easy	Revlon Colorsilk	.55	.08	1	0
Hair Prepn-Women’s	Sunsilk	Pantene Pro-V	.54	.18	0	0
Hair Spray-Women’s	Suave	White Rain	.55	.10	0	0
Hand Sanitizer	Germ-X	Purell	.52	.13	0	0
Health Bars/Sticks	Zone Perfect	Clif	.52	.20	0	1
Hominy Grits	Quaker	Jim Dandy	.88	.11	0	0
Honey	Sue Bee	Golden Nectar	.68	.24	0	1
Horseradish	Silver Spring	Gold’s	.59	.38	0	0
Ice Cream Cones	Joy	Keebler	.53	.13	0	1
Ice Cream-Bulk	Breyers	Dreyer/Edy’s Slw Chn	.64	.12	0	1
Ice Milk & Sherbet	Dreyer’s/Edy’s	Blue Bell	.66	.36	0	1
Insoles	Dr Scholl’s	Pro Foot	.77	.10	1	0
Jam	Smucker’s	Welch’s	.76	.10	0	1
Jelly	Welch’s	Smucker’s	.59	.12	0	1
Laxatives	Metamucil	Benefiber	.56	.18	1	0
Lemon/Lime-Diet	Sprite Zero	Diet Seven Up	.51	.16	0	1
Lemon/Lime-Regular	0	1				

Module	Brand 1	Brand 2	Aggregate Purch Share	Cross-State SD	Ad Intense	Socially Visible
Lighters	Bic	Scripto	.78	.07	0	1
Lip Remedies-Misc.	Carmex	Blistex	.70	.16	0	0
Lip Remedies-Solid	Chap Stick	Blistex	.76	.05	0	0
Lunches-Refrig	Osc Mayer Lunchables	Armour Lunch Makers	.85	.09	1	0
Margarine & Spreads	Shedd's	Blue Bonnet	.51	.12	0	0
Marshmallows	Kraft Jet Puffed	Campfire	.94	.05	0	1

Module	Brand 1	Brand 2	Aggregate Purch Share	Cross-State SD	Ad Intense	Socially Visible
Salad Dressing-Liq	Kraft	Ken's Steak House	.64	.17	0	1
Salad Dressing-Refrg	Marie's	Marzetti	.56	.31	0	1
Salad Toppings-Dry	Hormel	Oscar Mayer	.67	.13	0	1
Salads-Misc.	Reser's	Ready Pac Bistro Sld	.63	.24	0	0
Sandwiches-Frzn/Ref	Lean Pockets	Hot Pockets	.52	.07	0	0
Sauce Mix-Taco	Old El Paso	McCormick	.54	.21	0	0
Sauce-Asian	Kikkoman	La Choy	.70	.11	0	1
Sauce-Barbecue	Kraft	Sweet Baby Ray's	.61	.17	0	1
Sauce-Chili	Heinz	Tuong Ot Sriracha	.81	.18	0	0
Sauce-Cocktail	Kraft	McCormick	.64	.26	0	1
Sauce-Cooking	Hunt's Manwich	Del Monte	.92	.06	0	0
Sauce-Dipping	Marzetti	Litehouse	.81	.28	0	1
Sauce-Hot	Louisiana	Texas Pete	.59	.34	0	1
Sauce-Marinara	Prego	Hunt's	.52	.08	0	0
Sauce-Meat	A.1.	Heinz 57	.80	.15	0	0
Sauce-Mexican	Pace	Tostitos	.53	.19	1	1
Sauce-Misc.	Prego	Kraft	.59	.20	0	1
Sauce-Pepper	Tabasco	Frank's Redhot	.57	.19	0	1
Sauce-Pizza	Ragu	Contadina	.70	.18	0	0
Sauce-Worcestershire	Lea & Perrins	French's	.69	.15	0	1
Sauces & Gravies	Buitoni	Garden Fresh Gourmet	.61	.26	0	0
Seasoning Mix-Chili	McCormick	Carroll Shelby's	.84	.12	0	0
Seasoning Mix-Misc.	McCormick	Sun Bird	.54	.13	0	0
Shampoo	Suave Naturals	Pantene Pro-V	.53	.07	1	0
Shave Cream-Men's	Edge Advanced	Barbasol	.51	.10	0	0
Shave Cream-Women's	Skintimate	Gillette Satin Care	.65	.07	0	0
Sinus Remedies	Tylenol Sinus	Sudafed PE	.66	.14	0	0
Snacks-Misc.	SunChips	GM Chex Mix	.52	.05	0	1
Snacks-Variety Pk	Frito-Lay	Wise	.98	.04	0	1
Soap-Bar	Dove	Dial	.53	.09	0	0
Soap-Liq	Softsoap	Dial	.77	.06	0	0
Soap-Specialty	Suave Naturals	Dove	.52	.11	1	0
Soda Straws	Forster	Glad	.75	.19	0	0
Soup Mix-Dry/Bases	Maruchan	Lipton	.61	.11	0	0
Soup-Canned	Campbell's	Progresso	.80	.06	1	0
Soup-Frzn/Refrig	Tabatchnick	Skyline	.57	.32	0	0
Throat Lozenges	Ricola	Halls Breezers	.64	.12	0	0
Toast/Breadsticks	Old London	Wasa	.51	.16	0	0
Toilet Bowl Cleaner	Lysol	Clorox	.52	.06	0	0
Toilet Tissue	Charmin	Angel Soft	.54	.07	1	0
Toothbrushes	Colgate 360	Oral-B Indicator	.55	.11	1	0
Tortilla Chips	Doritos	Tostitos	.64	.06	0	1
Trail Mix	Planters	GM Chex Mix	.79	.13	0	1
Vinegar	Heinz	Pompeian	.73	.15	0	0
Vitamins-Children	Flintstones	L'il Crttrs Gummy Vt	.71	.13	0	0
Vitamins-Misc.	Nature Made	Nature's Bounty	.71	.13	0	0
Vitamins-Multi	One A Day	Centrum Silver	.60	.08	1	0

Module	Brand 1	Brand 2	Aggregate Purch Share	Cross-State SD	Ad Intense	Socially Visible
Water-Sparkling	Vintage	Perrier	.62	.30	0	1
Water-Still	Glaceau Vitmn Water	Nestle Pure Life	.52	.13	1	1
Wave Setting Product	Garnier Fructis Styl	Pantene Pro-V Style	.66	.11	0	0
Yogurt-Frozen	Turkey Hill	Wells Blue Bunny	.57	.37	0	1
Yogurt-Refrig	Yoplait	Dannon	.62	.10	1	0

Notes: Brand 1 and brand 2 in each module defined by total purchases. Aggregate purchase share for a given module is total purchases of brand 1 / (total purchases of brand 1 + total purchases of brand 2), and is calculated using all households in the Nielsen Homescan data. Cross-state standard deviation of the average purchase share for non-migrants is computed by averaging purchase share within each state-module pair, and then taking the mean of the standard deviation across states for each module. Cross-state standard deviation is calculated using the final sample as described in section 2.4.