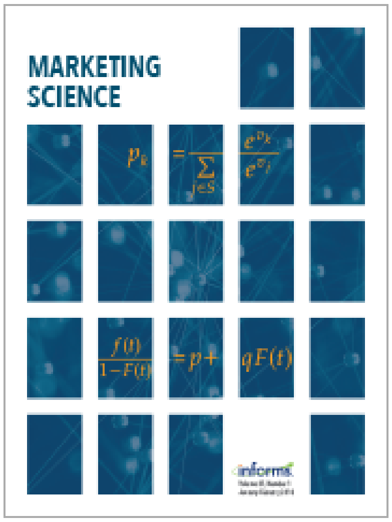


Marketing Science

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

Category Positioning and Store Choice: The Role of Destination Categories

Richard A. Briesch, William R. Dillon, Edward J. Fox,

MARKETING SCIENCE
Vol. 32, No. 3, May–June 2013, pp. 488–509
ISSN 0732-2399 (print) | ISSN 1526-548X (online)

http://dx.doi.org/10.1287/mksc.2013.0775

Category Positioning and Store Choice: The Role of Destination Categories

Richard A. Briesch, William R. Dillon, Edward J. Fox
Edwin L. Cox School of Business, Southern Methodist University, Dallas, Texas 75275
{rbriesch@cox.smu.edu, bdillon@cox.smu.edu, efox@cox.smu.edu}

We focus on *destination* categories, so named because they have the greatest impact on where households choose to shop and, more generally, on how category positioning affects which store a household chooses. We propose a reduced-form model-based analytical approach to identify categories that fill the destination role. Our approach determines which categories are most important to shoppers' store choice decisions and helps determine in which categories the retailer provides superior value. In addition, our approach allows us to understand the impact of the retailer's long-run merchandising policy decisions on the value it provides. Previous store choice research considered the effects of pricing, assortment and other merchandising decisions at the store level but did not focus on the effect of specific categories on store choice. This focus leads us to formulate a model that can (1) measure and explain the differential impact that specific categories have on shoppers' store choice decisions and (2) measure the relative value of retailers' category offerings, partitioning that value into the component resulting from retailer merchandising and the component that is nonmerchandising related. The model form captures differences in category value across stores (i.e., the store's category positioning) by specifying a spatial model for the store choice and category incidence intercepts. Our spatial model recognizes that stores position their offering vis-à-vis the category ideal based on long-run category merchandising decisions and that not all categories have the same importance in store choice decisions. We explore these issues for five retailers in the Charlotte, North Carolina market. We find that (1) category impact on store choice is highly skewed; (2) although categories with higher sales generally have a higher impact on store choice decisions, there are exceptions; (3) impact on store choice decisions does not vary systematically by the type of category (e.g., perishable versus dry grocery); and (4) our measure of category impact on store choice, although correlated with the category development index *between* retailers, is superior in that it provides a basis for comparing category impact *within* a retailer and how relative category value, based on long-run merchandising decisions, attracts shoppers to a store.

Key words: category positioning; category and store choice modeling; spatial modeling
History: Received: November 7, 2011; accepted: November 12, 2012; Preyas Desai served as the editor-in-chief and Gary Russell served as associate editor for this article. Published online in *Articles in Advance* March 21, 2013.

1. Introduction

A fundamental tenet of category management is that individual categories play different roles. The most widely used taxonomy identifies categories for the "destination," "routine," "occasional/seasonal" or "convenience" role. These *consumer-based* roles help retailers "understand how consumers view the category" (Blattberg et al. 1995, p. 22), which should enable retailers to manage "categories according to their importance to consumers" (ACNielsen 2006, p. 79). Surprisingly, these consumer-based category roles are described not from the consumer's perspective but by the "*positioning the retailer should take* based on the category's importance to the consumer" (Blattberg et al. 1995, p. 23, emphasis added). The common thinking is that categories should be managed according to their roles, thereby enabling retailers to use their product offering like a portfolio that attracts shoppers while profitably generating revenues.

Our focus is on *destination* categories, so named because they have the greatest impact on where households choose to shop and, more generally, on how category positioning (e.g., long-run merchandising policies) affects which store a household chooses. In other words, we are interested in which specific categories drive store choice decisions and how retailer merchandising influences those categories.

To our knowledge, no definitive definition of destination categories exists, yet destination categories have been discussed in the context of the different roles that categories play. For example, the Joint Industry Project on Efficient Consumer Reponse states in its original category management report that a destination category is "to be the primary category provider and help define the retailer as the store of choice by delivering consistent, superior target customer value" (Food Marketing Institute 1993, p. 26). Two implications of this statement help us understand the meaning of a destination category. First, being the primary provider suggests that the retailer is the store of choice, so destination categories should drive store traffic. Second, to be the store of choice requires that the retailer deliver more value in the category than its competitors do. Although retailers can allocate resources in an attempt to create destination categories, it is ultimately the shopper who determines whether the retailer is offering superior value in a particular category. We discuss the concept of destination categories in more depth in §3.

1.1. Objectives

In this paper, we propose a reduced-form model-based analytical approach to identifying categories that play the destination role. Because our model focuses on how shoppers respond to retailers' marketing mix decisions—not retailer decision making—it allows us to determine which categories are most important to shoppers' store choice decisions and which retailers provide superior value in different categories.[1] Our approach also allows us to understand the impact of a retailer's long-run merchandising policy decisions on the value it provides; specifically, we formulate a model that can (1) measure and explain the differential impact of specific categories on shoppers' store choice decisions and (2) measure the relative value of retailers' category offerings, partitioning that value into the component resulting from retailer merchandising and the component that is not merchandising related. This value partitioning clarifies the contribution of a retailer's merchandising to the value of its category offerings. It is also important to note that although previous store choice research focused on the effects of pricing, assortment, and other merchandising decisions at the store level, it did not consider the effect of *specific categories* on store choice.

To understand the role of category positioning in store choice, we formulate a mixed nested logit model that incorporates both store choice and category incidence decisions in a way that captures the effect of a store's weekly category merchandising decisions, as well as its long-run merchandising policies, on the probability of choosing that store. This reduced-form model of shopping behavior captures differences in category value across stores by specifying a spatial model for the store choice and category incidence intercepts. The store choice intercepts position each store in latent multiattribute space on the basis of store characteristics and the value that the store provides to consumers across *all* categories. The category incidence intercepts are parameterized in terms of the spatial distance between a household's perception of stores' offerings in a category and the household's ideal point for that category—in other words, the distance between what stores offer in a specific category and what the household wants in that category. Thus, our spatial model recognizes that stores can position their offerings vis-à-vis the category ideal based on long-run category merchandising decisions and that not all categories have the same importance in store choice decisions.

The remainder of this paper is organized as follows: Section 2 briefly reviews the extant literature and positions our work relative to this literature. The concept of a destination category is discussed in §3 and is specifically related to category development measures (i.e., category development indices, or CDIs). Section 4 presents two conceptual frameworks that explain how category incidence (conditional on the shopper visiting the store) is used to estimate category utility; these two conceptual frameworks guide our model specification. The store choice and category incidence component models are presented in §5, where we also explain how these model intercepts are parameterized and offer a baseline model that can be used to assess the usefulness of the proposed spatial representation. Section 6 discusses estimation issues. Our data set is described in §7. Modeling results follow in §8. Section 9 presents a policy analysis that investigates the impact of categories on store choice. We end with a discussion of limitations and avenues for future research.

2. Background and Contribution

Explaining store choice decisions has been of great interest to academics and practitioners alike. Researchers have studied a wide variety of factors that may influence a consumer's decision about where to shop including pricing, promotion, feature advertising, assortment, retail price format (HiLo versus everyday low pricing (EDLP)), shopping basket size and composition, travel distance/time, prior shopping experiences, the need for variety, shoppers' fixed and variable costs, cherry-picking, and household characteristics.[2]

This study contributes to the extant literature on store choice in a number of important ways.

• First, unlike previous store choice research, our primary interest is to understand the role that specific categories play in store choice decisions. To our

[1] For this reason we do not claim to recommend categories that the retailer could or should develop to serve the destination role. As we discuss in this section, and in more detail in §4, we specify a reduced-form model conditioned on current strategic retailer decisions, and consequently, we focus only on identifying categories that currently play the destination role.

[2] Representative papers include Reilly (1931), Baumol and Ide (1956), Huff (1964), Arnold et al. (1978, 1981), Arnold and Tigert (1982), Arnold et al. (1983), Broniarczyk et al. (2006), Chernev et al. (2003), Chernev and Hamilton (2009), Bell et al. (1998), Bell and Lattin (1998), Rhee and Bell (2002), Fox and Hoch (2005), Briesch et al. (2009), and Zhang et al. (2010).

knowledge, this is the first study to (1) address conceptually and empirically the issue of destination categories and (2) isolate the differential effects of category positioning decisions on store choice.

- Second, because we incorporate a spatial model of household preferences in which stores and categories are positioned in multiattribute space, we can assess the extent to which different categories drive store choice decisions. Furthermore, we show how the positions of a given store and category in multiattribute space reflect the value of the store's category offering.
- Third, in contrast to Bell et al. (1998) and Briesch et al. (2009), we do not adopt a shopping list metaphor or need-based approach to category incidence. By not conditioning on a shopping list or on category needs, we capture all purchase incidences and can therefore account for impulse purchases.[3]
- Fourth, we explicitly model both store choice and category incidence and consider a large number of categories (i.e., 80) in multiple retail formats (grocery and supercenter).
- Finally, in contrast to the earlier literature that either assumed that category value did not vary across consumers (e.g., Baumol and Ide 1956, Huff 1964) or allowed categories to have differential appeal across shoppers based on an unobserved shopping list or predetermined set of needs (e.g., Bell et al. 1998, Briesch et al. 2009), we allow stores to differentiate their category offerings based on assortment and other long-run merchandising policies. We find that categories do not necessarily have the same importance in store choice decisions. Holding price and promotion constant, the likelihood of category purchase depends on which store a shopper chooses.

3. Destination Categories and CDIs

3.1. Destination Categories

ACNielsen (2006) proposes an analytical basis for category role selection in which the first question to be answered is how important the category is to consumers. ACNielsen argues that high category importance is a necessary, but not sufficient, condition for a destination category. The retailer must also provide superior value in the category based on merchandising effort—in particular, product assortment.

Adopting this perspective, the retailer selects categories for the destination role.[4] That selection influences retailer actions, in particular merchandising (including product assortment) decisions. However, whether a category selected for the destination role achieves its objectives for the retailer is ultimately determined by consumers. They are the ones who determine the value a retailer is delivering in a specific category and where they will shop.[5] Although we can assume that all categories influence store choice to varying degrees, a destination category disproportionally increases the likelihood of a store being chosen because that store offers consumers superior value in the category. Following first-order principles, we capture value by estimating the utility that a shopper derives from a category at each store. As we discuss in §§4 and 5, the utility that a consumer derives from a category is determined by the attractiveness of that category at a specific store, along with other factors including category needs and in-store factors that influence the consumer while shopping.

Continuing with first-order principles, the more attractive the category is at a store, the higher the probability that a consumer chooses that store. Attractiveness is defined in terms of a long-run merchandising score that reflects how well positioned a category is relative to what consumers want in that category. The purpose of the merchandising score is to move the position of a store based on the perceived attractiveness of the store's long-run merchandising policies. The tacit assumption is that the smaller the distance between a store's location, adjusted for its merchandising attractiveness and what consumers want in a category, the higher the probability of shoppers choosing that store and the more that category plays the destination role for this retailer.

To fix ideas, in this paper we define a destination category as one that (1) substantially affects store choice decisions and (2) delivers superior value (at least in part) as a result of the retailers' long-run merchandising decisions. Thus, destination categories are jointly determined by the retailer and the consumer, in the sense that the retailer anticipates or responds to consumers' category preferences. Accordingly, the retailer does not *select* categories to serve in the destination role; rather, the retailer identifies categories that motivate and attract consumers. Note that we define a destination category as one that increases store traffic, not in terms of how much a consumer purchases while shopping. This perspective is consistent with the view taken by practitioners. Obviously,

[3] To be clear, although we capture multiple store visits, we do not explicitly model this behavior.

[4] Our focus is not on how a retailer selects a category for the destination role, but rather on identifying which categories have assumed that role.

[5] A brand manager can be thought of as playing a role similar to the retailer. The brand manager determines a brand's value proposition, one that they hope will resonate with consumers. However, it is consumers who ultimately determine whether the value proposition is relevant and believable and whether they will purchase the brand.

retailers welcome larger basket sizes and/or basket expenditures. However, interviews with category managers at large retailers revealed that higher category purchase quantities are associated with other category roles; these include commonly defined roles such as "routine" and "seasonal" as well as retailer-defined roles such as "basket builder" and "impulse." By contrast, the ability to drive store traffic was the most distinguishing feature of destination categories cited by the retail managers. Specific responses included that such categories "are trip drivers," "can be a traffic driver," and "can increase awareness and customer count."[6]

Our perspective that destination categories result in large measure from a retailer's long-run merchandising decisions has two salutary benefits. First, such decisions can be differentiated from short-run merchandising decisions (e.g., feature advertising) that may also increase store traffic. Second, this definition mitigates possible endogeneity concerns, as long-run decisions are, by definition, independent of short-run shocks to the system. In other words, strategic reactions by retailers to shopper behavior operate over a longer time frame, so it is possible to condition on current retailer behavior and then build a consumer model, relative to this conditioning.

Our empirical analysis will consider only categories that are offered by the five retailers in our sample, but in principle, this need not be the case.[7] To identify a category as playing the destination role, we must demonstrate that the category influences store choice and, just as important, that the long-run merchandising policies have created value as evinced by the level of attractiveness consumers place on the category. As we demonstrate in §5, our modeling framework, which adopts a spatial parameterization for the store and category intercepts, allows us to estimate the utility that the shopper derives from a category and the impact that category utility—in particular, category merchandising—has on store choice decisions.

3.2. Category Development Index

One measure that has been proposed as an indicator of shoppers' preference for a specific category at a specific retailer is the category development index (CDI). At an aggregate level, the CDI reflects the preference of shoppers to purchase specific categories at one retailer (rather than at others) and is defined as the retailer's share in a particular category divided by its overall market share, multiplied by 100 (e.g., Dhar et al. 2001). Table 1 presents the CDIs of 80 product categories for five retail chains in the Charlotte,

[6] Information on our interviews with retail category managers can be obtained by contacting the third author.

[7] However, in such a case, the determination of value would be difficult to assess under quasi-monopolistic conditions.

Table 1 Category CDIs

Category	BI-LO	Food Lion	Harris Teeter	Winn-Dixie	Walmart
Carbonated beverages	127	102	115	121	70
Cigarettes	93	229	8	53	51
Cold cereal	105	96	128	98	84
FZ dinners/entrees	109	113	125	99	62
Fresh bread and rolls	101	98	140	111	65
Salty snacks	110	104	114	103	73
Beer/ale/alcoholic cider	97	158	72	90	64
Milk	127	105	108	116	68
Natural cheese	118	105	113	127	63
Cookies	94	96	126	101	85
Crackers	89	106	116	93	90
Luncheon meats	123	128	90	136	58
Breakfast meats	102	107	113	143	64
Total chocolate candy	61	65	82	64	187
Dog food	105	96	72	75	130
Ice cream/sherbet	100	114	160	116	34
FZ pizza	139	104	141	82	53
FZ poultry	83	73	140	93	68
Cat food	127	125	70	92	89
Soup	112	113	130	113	52
Coffee	95	96	94	120	81
RFG salad/coleslaw	134	97	129	133	55
Pet supplies	33	32	27	51	299
Processed cheese	136	116	88	122	61
Wine	54	127	117	63	59
Laundry detergent	89	96	101	70	94
Vegetables	138	129	97	130	43
Toilet tissue	96	96	94	97	99
FZ seafood	206	78	99	192	26
Snack bars/granola bars	98	78	87	78	124
Total nonchocolate candy	51	73	57	90	187
FZ novelties	118	123	129	109	42
Paper towels	83	103	94	115	99
Household cleaner	70	74	89	74	162
Dry packaged dinners	143	112	98	93	76
Internal analgesics	82	55	50	92	203
Dough/biscuit dough—RFG	104	102	136	142	54
Frankfurters	122	117	119	132	46
Vitamins	40	33	21	20	285
RFG juices/drinks	93	104	144	102	64
Yogurt	79	69	174	74	100
Bottled water	73	44	146	73	140
Soap	62	62	56	75	204
Toothpaste	82	50	94	92	186
Pastry/doughnuts	100	98	150	71	70
Cold/allergy/sinus tablets	69	47	59	53	237
Salad dressings—SS	133	110	115	114	57
FZ plain vegetables	136	95	160	100	42
Canned/bottled fruit	104	108	133	106	63
Snack nuts/seeds/corn nuts	69	93	85	80	122
Baking mixes	139	107	109	129	61
Bottled juices—SS	94	123	108	102	63
Skin care	34	20	43	34	321
FZ breakfast food	101	114	136	113	51
FZ bread/FZ dough	96	122	134	122	42
Canned meat	144	111	67	145	68
FZ meat	88	60	221	85	53
Dish detergent	69	78	126	91	101
Spices/seasonings	97	97	116	105	75
Cups and plates	77	78	83	113	123
FZ appetizers/snack rolls	88	96	106	162	60
RFG fresh eggs	115	95	123	134	70

Table 1 (Cont'd.)

Category	BI-LO	Food Lion	Harris Teeter	Winn-Dixie	Walmart
Shampoo	76	47	48	90	236
Batteries	66	36	58	75	252
Pickles/relish/olives	108	94	138	137	52
Margarine/spreads/butter	101	107	102	144	84
Dinner sausage	130	94	118	188	50
Deodorant	73	53	71	81	212
FZ desserts/topping	84	115	124	180	43
Shortening and oil	119	110	97	111	65
Baking needs	122	100	113	95	82
SS dinners	148	129	91	105	56
Toaster pastries/tarts	156	83	123	89	86
Air fresheners	80	70	71	84	191
Toothbrush/dental accessories	68	36	47	59	247
Food and trash bags	77	87	98	81	112
Spaghetti/Italian sauce	117	103	137	128	52
Sanitary napkins/tampons	78	72	85	49	180
Seafood—SS	127	107	99	114	60
Peanut butter	106	110	100	104	80

Note. FZ, frozen; RFG, refrigerated; SS, shelf stable.

North Carolina market. These chains and product categories will be the focus of the empirical analysis that follows in §8. CDIs can vary markedly across retail chains, as is the case for the five retail chains shown in Table 1. Although CDIs can undoubtedly identify store-by-category differences, they provide little insight into the reasons for, or consequences of, these differences for the following reasons.

- First, CDIs provide no information about the importance of different categories in store choice decisions, a key consideration in the identification of destination categories. For example, categories with high purchase frequency and/or that represent a high annual expenditure may have greater impact on store choice.
- Second, CDIs provide no insights as to the value added by the retailer's long-run merchandising policies. In contrast, our model framework separates factors that the shopper knows before choosing a store, which reflect a store's long-run category assortment and merchandising policies, from those that are only observable after choosing the store, i.e., while shopping.
- Third, CDIs have limited explanatory value; a high CDI (>100) might indicate that the category is a complement to other categories that actually influenced the store choice decision (see, for example, Manchanda et al. 1999), or it might indicate that the category is a complement to other purchases that were made because of advertised prices.
- Finally, because CDIs are neither household- nor time-specific, they are not at all informative as to which store a household may choose on any given shopping trip.

As we demonstrate in §9, our model form allows us to derive a metric that is highly correlated with CDIs across a store (within a category) but that remedies these limitations.

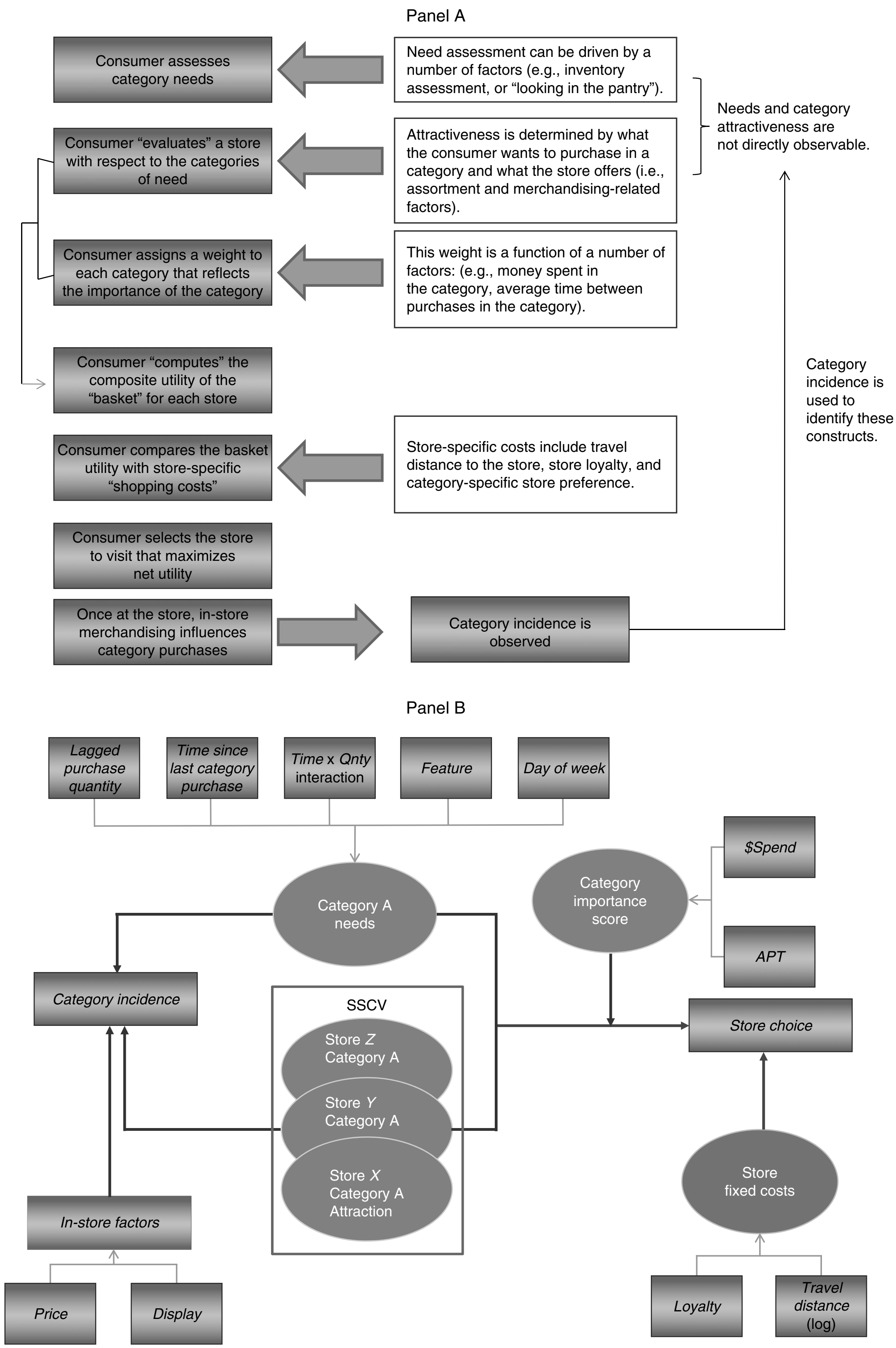
4. Conceptual Framework

It is well known that modeling store choice presents substantial challenges. Consider that the expected costs of shopping at each store, and hence the probability of choosing each store, depends on the shopper's intended purchases. Yet purchase intentions, reflected in either a shopping list (Bell et al. 1998) or a priori category needs (Briesch et al. 2009), are not observed—they must be inferred from actual purchases. Actual purchases, however, include both *intended* and *unplanned* purchases (see Bell et al. 1998, p. 354 for a detailed discussion). Unplanned purchases are made as a result of in-store displays and promotions (Kollat and Willett 1967), but shoppers are not exposed to these in-store stimuli until *after* they have chosen a store. Moreover, in-store stimuli are specific to a store, so unplanned purchases are dependent on which store is chosen.

To address this challenge, we will model observed purchases in such a way that intended and unplanned purchases are partitioned. Specifically, information that shoppers know prior to choosing a store will be reflected in the intended purchase probability, whereas in-store displays and promotions—to which shoppers are exposed only after choosing a store—will be reflected in the unplanned purchase probability. Note that these purchase probabilities will be modeled at the category, rather than the stock-keeping unit (SKU) level. In this way, we parsimoniously capture a store's diverse product offerings while avoiding the complexity of cross-product effects within categories and incorporate retailers' assortment policies, which are only relevant at the category level. Finally, we use intended category purchase probability and other relevant factors to estimate the utility that the shopper derives from each category, which is then incorporated directly into the store choice model. Note that the uncertainty of purchase intentions will be taken into account by estimating the category purchase (i.e., category incidence) and store choice models simultaneously.

To better illustrate our proposed approach and, specifically, how we will use category incidence (conditional on the shopper visiting the store) to estimate category utility, Figure 1 provides two conceptual frameworks. Figure 1, panel A, is representative of the mental model (we hypothesize) that shoppers use in deciding which store to visit, and panel B translates the shopper's conceptual framework into elements that are more closely aligned with the model proposed in the next section.

Figure 1 Conceptual Frameworks: (A) Conceptual Consumer Model and (B) Conceptual Modeling Framework



4.1. Conceptual Shopper Model

As depicted in panel A of Figure 1, the process begins with the shopper recognizing one or more category needs. He or she then, either consciously or unconsciously, evaluates each store's offering in the categories of need. In essence, this involves an evaluation of what the shopper wants to purchase and what the store offers. It is important to note that this evaluation reflects what the shopper knows before choosing a store—each store's long-run category assortment and merchandising policies, not the actual displays and (unadvertised) prices, which are only observable after choosing the store. In addition, the shopper weighs the importance of individual categories in satisfying his or her needs. These importance weights reflect the fact that certain categories may have a greater impact on store choice. At this point, the shopper is in a position to assess the *total basket utility* that he or she would derive from purchasing at each store under consideration. To determine which store to visit, the shopper compares the store-specific total basket utility with store-specific shopping costs. The shopper then chooses the store that maximizes his or her net utility, which in turn determines which in-store stimuli he or she sees and hence whether category purchases are consummated. Recall that only the household's category purchases (incidence) are observed, not the household's category needs or the store-specific category value. In our model, category needs and store-specific category value are identified by the household's category purchases.

4.2. Conceptual Modeling Framework

The modeling framework shown in panel B of Figure 1 adds further specificity to the shopper model described above. In the figure, ovals depict inherently unobservable factors (i.e., constructs) and squares depict observable variables. We see that factors hypothesized to affect *category incidence* are divided into those that influence the shopper prior to visiting a store (category needs and store-specific category value, labeled SSCV in the figure) and those that are experienced while shopping (*Price* and *Display*). We use inventory-related variables such as *Time since last category purchase* and *Lagged purchase quantity*, whether the households' favorite brands are advertised as *Features*, as well as the *Day of the week* the shopping trip took place to operationalize category needs. Inventory assessment and feature advertising occur prior to the store visit; the day of the week that the shopping trip took place provides information about the type of trip planned—stock-up/major versus fill-in (Kahn and Schmittlein 1989, Kim and Park 1997). And, as detailed in §5.1.1, we develop a spatial model to estimate the store-specific category value by reparameterizing the intercept term in the category incidence model. Store choice is hypothesized to be a function of category needs, store-specific category value, and store-specific fixed costs. However, the effects of category needs and store-specific category value are weighted by the importance of the category to the shopper, which, following ACNielsen (2006), is hypothesized to depend on the amount of money spent (*$Spend*) in the category and average time between category purchases (*APT*).

4.3. Comment

We acknowledge that there is a conditioning or *endogeneity* problem inherent in the notion of a destination category, given the dependency between observed purchases and store choice. Although we do not explicitly address the endogeneity issue, note the following. First, as discussed above we distinguish between factors that shoppers experience (or are exposed to) before choosing a store and after they are at the store, i.e., while shopping. Second, we adopt a *utility maximizing* approach as opposed to a *cost minimization* approach, as implemented by Bell et al. (1998) and Briesch et al. (2009).[8] Because we focus on utility maximization, we do not need to estimate the quantity that the household plans to buy because we do not calculate category costs, but rather we focus on category attractiveness.[9] Third, although our model form requires (via the conditioning argument) that a household be in the store to make a purchase, the conditioning is used only to estimate the attractiveness of the category; it is not assumed to be part of the household decision process. In our analysis, we consider the 80 largest of the 290 categories for which we have detailed price information. Consequently, there is no binding constraint that the household must purchase in one of the 80 largest categories given that they selected a store to visit. Finally, there is another reason to view this putative endogeneity problem as perhaps less severe. Although to some extent store positioning resulting from category merchandising may impact store choice, it is unlikely that retailers can respond to consumer purchasing patterns by changing product assortments in the short run, i.e., weekly. Rather, it is more likely that the majority of category merchandising and product assortment policies are fixed over short periods. In other words, it takes retailers longer to understand how category positioning might be changed to impact store choice. This argument is similar to the rationale for treating prices as exogenous. Retailers cannot detect consumers' responses to prices quickly enough to influence weekly pricing decisions (see Erdem et al. 2006).

[8] Utility maximization is the dual of cost minimization.

[9] As discussed in §2, this also relieves the need to adopt a shopping list metaphor (Bell et al. 1998) or need-based approach (Briesch et al. 2009).

5. Model Forms

5.1. Category Incidence Model

The indirect utility for household h purchasing category c on trip t (at store s) can be written as

$$U^C_{hsct} = V^C_{hsct} + \varepsilon^C_{hsct}, \tag{1}$$

where V^C_{hsct} denotes the deterministic component of utility. Consistent with our conceptual modeling framework, we partition the deterministic component of the utility of purchasing in the category in terms of category needs, in-store factors, and store-specific category value as follows:

$$\begin{aligned} V^C_{hsct} = {} & \varphi_{hsct} + \gamma_{0hc} + \gamma_{1hc} Time_{hsct} + \gamma_{2hc} Qnty_{hsct-1} \\ & + \gamma_{3hc} Time_{hsct} \times Qnty_{hsct-1} + \gamma_{4hc} FAdv_{hsct} \\ & + \gamma_{5hc} WKEnd_{hct} + \gamma_{6hc} Price_{sct} + \gamma_{7hc} Disp_{hsct}. \end{aligned} \tag{2}$$

In Equation (2), the first five covariates relate to factors that influence a household before shopping and which assist in determining the household's category needs:

Time = The number of days since the household last purchased in the category.
Qnty = Quantity purchased in a category on last visit.
Time × *Qnty* = The mean-centered interaction of number of days since the last purchase and the quantity last purchased.
FAdv = The proportion of each household's three favorite brands that are feature advertised at store s during trip t, weighted by the brands' share of household category purchases during the 26-week initialization period.
WKEnd = A dummy variable equal to 1 if the shopping trip occurred between Friday and Sunday; 0 otherwise.

The last two covariates in Equation (2) are in-store factors that influence the consumer while shopping:

Price = The average price of those brands in the category that are carried by all stores, weighted by each brand's long-run market share of the category (see Ainslie and Rossi 1998).
Disp = The proportion of each household's three favorite brands that are displayed at store s during trip t, weighted by the brands' share of household category purchases during the 26-week initialization period.

With respect to parameters, φ_{hsct} denotes store-specific category value (see Figure 1), which indicates how attractive category c at store s is for household h (how this parameter is estimated will be explained in the next section); γ_{0hc} represents the (estimated) baseline purchase frequency for the category, which accounts for differences in frequency of category purchases when estimating spatial parameters (spatial parameters will be discussed below); and, γ_{jhc}, $j = 1, 2, \ldots, 7$, reflect the impact of each covariate on the indirect utility of purchasing in category c, where we assume a variance components representation for these effects:

$$\gamma_{jhc} = \eta_j + \xi_c + \varepsilon_h. \tag{3}$$

In Equation (3), η_j gives the mean impact of covariate j, ξ_c gives the variance across categories, and ε_h gives the variance across households.

Note that we selected the covariates appearing in Equation (2) on the basis of our conceptual model framework, which partitions factors that could possibly influence category purchases and store choice based on whether the shopper experiences those factors before or after entering the store. There are, however, several covariates that warrant further discussion.

Because category needs depend on household inventory levels and the rate at which the inventory is consumed, we need to operationalize these variables; we use *lagged quantity* and *time since last category purchase* (and their interaction) as surrogates for these inventory levels, respectively, which is consistent with the extant literature (see, for example, Erdem et al. 2003, Briesch et al. 2009, Hendel and Nevo 2009). Accordingly, we expect that the probability of category purchase is negatively related to quantity last purchased and positively related to time since the most recent category purchase (Chib et al. 2004). We include *WKEnd* to account for the different shopping behaviors that are known to be associated with weekday and weekend trips; in particular, basket sizes on weekend trips are generally larger than basket sizes on weekday trips (see Kahn and Schmittlein 1989, McAlister et al. 2009). Thus, *WKEnd* is a plausible surrogate for both basket size and basket expenditure as well as other behaviors driven by the type of trip, i.e., stock up (major) versus fill in.

Finally, we chose not to model basket dollar expenditure directly for two reasons. First, our interviews with retail category managers indicated that destination categories are associated with driving traffic to the store as opposed to increasing basket size or basket dollar expenditure. Second, we performed a natural experiment in which we conditioned CDIs according to a household's basket expenditure; in other words, we recalculated the CDIs in Table 1 conditioning on high versus low dollar category spend (detailed results are in the Web appendix, available as supplemental material at http://dx.doi.org/

10.1287/mksc.2013.0775).[10] The differences are, for the most part, quite small, and there appears to be no systematic pattern to suggest that category dollar expenditures are driving category incidence. Although we do not explicitly include category dollar expenditure in the category incidence model, we recognize that category dollar spend may influence the importance of a category in driving store choice decisions. We discuss this issue in §5.2.

5.1.1. Category Attraction Parameter. To allow store choice to be influenced by what the shopper wants vis-à-vis what stores offer in a category, we adopt an ideal point representation for φ_{hsct}, the store-specific category value, which gives the attraction of category c at store s for household h. Accordingly, we define φ_{hsct} as

$$\varphi_{hsct} = \sum_d w_d(\alpha_{sd} + \Delta_{hscdt} - \alpha^I_{hcd})^2, \tag{4}$$

where

α_{sd} = the location of store s on latent attribute dimension d,
w_d = the importance weight of latent attribute dimension d,
α^I_{hcd} = household h's ideal point for category c on latent attribute dimension d, and
Δ_{hscdt} = long-run merchandising score for household h of category c at store s during trip t on latent attribute dimension d.

In the interest of parsimony, Δ_{hscdt} is defined as a linear composite of long-run merchandising covariates such that

$$\begin{aligned}\Delta_{hscdt} = {} & b_{1hd}Fav_{hsct} + b_{2hd}NUPC_{sct} + b_{3hd}NBrd_{sct} \\ & + b_{4hd}NSize_{sct} + b_{5hd}\%PL_{sct} + b_{6hd}\%Natl_{sct} \\ & + b_{7hd}AdvF_{sct} + b_{8hd}DispF_{sct} + b_{9hd}AvgPrice_{sct},\end{aligned} \tag{5}$$

where

Fav = Favorite brand availability—the summation of the purchase shares of the top three brands for the household if the brand is sold in store s during period t or is 0 otherwise;
$NUPC$ = The number of universal product codes (UPCs) per brand available at store s in category c during trip t divided by the average number of UPCs per brand carried by all stores in all periods;
$NBrd$ = The number of brands available at store s in category c during trip t divided by the average number of brands carried by all stores in category c over all periods;
$NSize$ = The number of sizes per brand available at store s during trip t divided by the average number of sizes per brand for all stores over all periods;
$\%PL$ = the percentage of private label UPCs available in category c at store s during trip t;
$\%Natl$ = the percentage of national brand UPCs available in category c at store s during trip t;
$AdvF$ = the frequency of brand advertisements—the number of weeks that at least one SKU in the brand was feature advertised divided by the total number of weeks;
$DispF$ = the frequency of displays—the number of weeks that at least one SKU in the category was on display divided by the total number of weeks;
$AvgPrice$ = the average volume-weighted prices across all SKUs in a category; and
b_{khd} = the component weight for long-run merchandising covariate k for household h on dimension d.

We use the first 26 weeks of data to initialize each covariate and then apply a 26-week window to assess availability. For example, if we observe that a respondents' favorite brand is available at store s, say, at week 2, we assume that the brand was also available in weeks 3–28. With the exception of *AvgPrice*, all of these merchandising covariates have been investigated in the context of the impact of assortments on store choice (see Briesch et al. 2009) and so are included in this study. However, in this study, we are interested in the longer-run impact of each covariate because we believe that consumers develop beliefs about store and category positioning based on their collective shopping experiences. We have included *AvgPrice* because the longer-run pricing practices of retailers play a role in shaping category perceptions as well; for example, EDLP versus HiLo.

Note that the role of Δ_{hscdt}, the long-run merchandising score, in our modeling framework is to move the position of store s on dimension d based on the perceived attractiveness of store s's long-run merchandising policies in category c. The tacit assumption is that the smaller the distance between the store's location (adjusted for merchandising attractiveness) and the category ideal point, the higher the probability of choosing that store. The magnitude and statistical significance of each component weight associated with a long-run merchandising covariate (b_{khd}) indicates the extent to which the covariate impacts the composite and, perhaps more importantly, the covariate's role in displacing a store along each latent attribute dimension. In other words, we will use the component weights to "adjust" the store's position on the latent

[10] We used median dollar category expenditure to form the high versus low household groups for each retailer.

attribute dimensions. Note also that (1) *AdvF* and *DispF* are incorporated in the merchandising score but only as long-term frequencies (see Ainslie and Rossi 1998), and (2) in deriving the store locations (α_{sd}) and category ideal points (α^I_{hcd}), we remove the baseline category purchase frequencies (γ_{0hc}) so that the impact of assortment and merchandising is independent of how often the household purchases in the category. We establish the conditions for identification of the spatial parameters below (proofs are available in the Web appendix).

Finally, we would argue that the merchandising and product assortment variables (and their definitions) that appear in Equation (5) reasonably capture a retailer's merchandising and product assortment decisions and are similar to the operationalizations used by others (e.g., Boatwright and Nunes 2001, Briesch et al. 2009); however, we recognize that these covariates may not be reflective of how households encode category information or form impressions about categories. For example, Hoch et al. (1999) present an interesting approach to capturing how households perceive the "variety" of an assortment. Their approach, however, is best suited to studies that consider a small number of items per category, unlike the present study, because it is based on computing the psychological distance between all items in a category (i.e., all pairwise comparisons).[11] We discuss this limitation of the current study further in §10.

5.1.2. Identification Conditions for Spatial Parameters. To identify the spatial parameters, i.e., store locations (α_{sd}) and category ideal points (α^I_{hcd}), we use category purchase incidence data along with a number of identifying constraints. In this section we provide general identification conditions. Conditions for identification of a k-dimensional solution rely on the identifying restrictions associated with the $k-1$ dimensional solution (see the Web appendix for proofs).

Condition 1. The weights for the dimensions are set to -1 for all dimensions. This identifies the scale of the map and ensures that all dimensions have the same scale.

Condition 2. One store is located at the origin (or the stores are centered at the origin; i.e., the sum of the store positions on each dimension add to 0). This restriction provides translational invariance for the stores and helps identify the category intercepts.

Condition 3. One category is located at the origin (or the categories are centered at the origin; i.e., the sum of the category positions on each dimension add to 0). This restriction provides translational invariance for the categories and helps identify the category intercepts and other store positions.

Condition 4. For each dimension d, $\alpha_{(s=d+1)d} > 0$. This restriction provides rotational invariance for the dimensions.

Condition 5. For each dimension d, $\alpha_{(s=d+1)j} = 0$ and $\alpha^I_{h(c=d+1)j} = 0$, $j = d+1, \ldots, D$.

Condition 6. The number of dimensions (D) is less than the number of stores (S), as $D+1$ stores are required to identify the category positions and intercepts.

5.2. Store Choice Model

The indirect utility for household h selecting store s on trip t is

$$U^S_{hst} = V^S_{hst} + \varepsilon^S_{hst}, \tag{6}$$

where V^S_{hst} denotes the deterministic component and ε^S_{hst} denotes the error term. Consistent with our conceptual modeling framework presented in §4, we parameterize the deterministic component of the indirect utility of store choice in terms of "store fixed costs" (e.g., travel distance/time), category needs, and store-specific category value. Recall that the latter two terms were used in specifying the deterministic component of the indirect utility of purchasing in the category shown in Equation (2). However, in-store factors relating to *Price* and *Display* should not be included because they are only observed by the shopper *after* selecting a store. Thus, for the purpose of specifying the store choice model, we define a reduced form of the deterministic component of the indirect utility of purchasing in a category previously introduced in Equation (2) as follows:

$$\begin{aligned} V^R_{hsct} = {} & \varphi_{hsct} + \gamma_{1hc} Time_{hsct} + \gamma_{2hc} Qnty_{hsct-1} + \gamma_{3hc} Time_{hsct} \\ & \times Qnty_{hsct-1} + \gamma_{4hc} FAdv_{hsct} + \gamma_{5hc} WKEnd_{hct}, \end{aligned} \tag{7}$$

where all terms have been previously defined.[12]

Whereas V^R_{hsct} gives the deterministic component of the indirect utility of a category, it is the total utility (i.e., value) of the basket that a household intends to purchase at a given store that should influence the household's decision as to which store to visit. To develop such a measure, we first consider the extent to which categories should be weighted differently in determining the impact that a category

[11] Another interesting approach is presented by Morales et al. (2005), who capture how a consumer organizes category assortment internally. We thank a reviewer for raising this issue and pointing us to the work of Hoch et al. (1999) and Morales et al. (2005).

[12] Notice that the parameter γ_{0hc} is also absent. Recall that γ_{0hc} represents the estimated baseline purchase frequency for the category, which is included to account for differences in frequency of category purchase when estimating spatial parameters. In the store choice model, adjusting for baseline purchase frequencies is not needed.

has on store choice. As discussed in §4, it is reasonable to expect that certain categories will have more weight in driving store choice; for example, categories such as carbonated beverages, which have high purchase frequency and relatively high dollar value are perhaps more likely to affect store choice decisions than categories such as salt, which have low purchase frequency and low dollar value (ACNielsen 2006). Letting κ_{hc} denote the weight that household h places on category c, we can write

$$\kappa_{hc} = \exp(u_c + \nu_{c1}\$Spend_{hc} + \nu_{c2}APT_{hc} + \nu_{c3}\$Spend_{hc} \times APT_{hc} + \zeta_{hc}),$$

where $*Spend*$ denotes the household's average dollar spend in the category and *APT* denotes the average time between purchases in the category. In estimating κ_{hc}, we mean center both covariates and therefore expect that $\nu_{c1} > 0$ and $\nu_{c2} < 0$, reflecting that, all else the same, the greater the dollar amount spent in a category and the shorter the time between purchases, the greater the category's weight in store choice decisions.

The category importance weight, κ_{hc}, is used to scale the value of the store's category offering for a given household, and aggregating over all categories that the household needs, we define a measure of the utility that the store offers for the household's entire basket of intended purchases. We denote the utility of the entire basket of intended purchases for household h on trip t as *BaskUtil*, where

$$BaskUtil_{hst} = \sum_{c \in S(C)} \kappa_{hc} \ln(1 + e^{V^R_{hsct}}). \tag{8}$$

In Equation (8), the term $\ln(1 + e^{V^R_{hsct}})$ is the *inclusive value* of a category, i.e., the maximum value of the utility for the category, excluding in-store factors. We can see from this specification that *BaskUtil* is an importance-weighted measure of utility that captures all categories the household intends to purchase at a given store on a particular trip.

Having defined all of the necessary covariates, we can express the deterministic component of the indirect utility of store s for household h on shopping trip t as

$$V^S_{hst} = \tau_{hs} + \beta_{1h}Loyal_{hs} + \beta_{2h}Dist_{hs} + BaskUtil_{hst}, \tag{9}$$

where

τ_{hs} = the intrinsic attraction of the store to the household;
Loyal = category-independent store loyalty measure—for the initialization period, the number of visits made by the household to a given store divided by the total number of stores visited during this same period;
Dist = natural logarithm of the travel time (in minutes) from the centroid of the household's zip + 4 to each store;
BaskUtil = importance-weighted measure of basket utility; and
β_{jh} = the parameters associated with the store and household-level covariates.

Household heterogeneity is captured by adopting a variance components representation for τ_{hs} and β_{jh} as follows:

$$\tau_{hs} = \phi_s + \sigma_h, \tag{10a}$$

$$\beta_{jh} = l_j + \zeta_h. \tag{10b}$$

In Equation (10a), ϕ_s denotes the mean attraction of store s, and σ_h gives the variance across households; in Equation (10b), l_j denotes the mean impact of covariate j, and ζ_h gives the variance across households.

Note that in deciding the specification for the store choice model, we had to make trade-offs as to the number and types of covariates to include to minimize the computational requirements to estimate a rather large set of free parameters with an objective function that requires approximating high-dimensional integration. For this reason, and because our primary focus is on the role that destination categories play in influencing where a consumer chooses to shop, we did not include household demographic variables in the store choice model. Rather, we focus on two household-level covariates that have been shown to influence store choice (see Bell et al. 1998). The *Loyal* covariate is intended to capture category-independent store loyalty, which reflects the household's intrinsic preference for a store. That is, the preference for a store that does not vary from trip to trip. This measure is analogous to the brand loyalty measure discussed in the brand choice literature (see, for example, Keane 1997) and is computed, as described above, in the same manner as brand loyalty (see Bucklin and Lattin 1992). The other household-level covariate, *Dist*, which measures distance traveled, uses information on the location of the household and the store visited to estimate the travel time from each household to each store. Travel times were estimated by geographic services provider ESRI using a proprietary algorithm that incorporates traffic and driving speeds. Finally, as stated above, we do not include household demographics in the model explicitly, but we do investigate the role that demographics play in the importance of a category to a household. To accomplish this, we follow a two-step process. First, for each household, we harvested the category importance parameter (κ_{hc}) using Bayesian posteriors from the estimated covariance matrix and parameters. Next, we regressed the interhousehold category importance parameters on the household's demographic covariates.[13] We report these results in §8.2.3.

5.3. Baseline Model

Central to the proposed model is the spatial representation in which an ideal-point specification (i.e., Euclidean distance) is used to define the attraction of a category. A tacit assumption in this formulation is that the attraction parameter plays a significant and informative role in defining the deterministic component of the indirect utility of purchasing in the category and ultimately in driving which store a household chooses.

Our baseline model tests this assumption by adopting a linear form for the deterministic component of the indirect utility of purchasing in a category originally defined in Equation (2). Specifically, for the purpose of estimating the baseline model, we define the deterministic component of the indirect utility of purchasing in a category as

$$\begin{aligned} V_{hsct}^{C*} = {} & \gamma_{0hc} + \gamma_{1hc} Time_{hsct} + \gamma_{2hc} Qnty_{hsct-1} + \gamma_{3hc} Time \\ & \times Qnty_{hsct-1} + \gamma_{4hc} FAdv_{hsct} + \gamma_{5hc} WKEnd_{hct} \\ & + \gamma_{6hc} Price_{sct} + \gamma_{7hc} Disp_{hsct} + \Delta_{hsct}, \end{aligned} \tag{11}$$

where all terms have been previously defined. There are several important differences to note. First, we have removed φ_{hsct}, the category attraction parameter, and as such moved away from the proposed spatial model. Second, we include Δ_{hsct}, the long-run merchandising score, directly in the deterministic component, but (contrary to Equation (5)) we redefine Δ_{hsct} as

$$\begin{aligned} \Delta_{hsct} = {} & b_{1hc} Fav_{hsct} + b_{2hc} NUPC_{sct} + b_{3hc} NBrd_{sct} \\ & + b_{4hc} NSize_{sct} + b_{5hc} \%PL_{sct} + b_{6hc} \%Natl_{sct} \\ & + b_{7hc} AdvF_{sct} + b_{8hc} DispF_{sct} + b_{9hc} AvgPrice_{sct}. \end{aligned}$$

Notice that the coefficients that give the impact of each long-run merchandising covariate do not vary by dimension, although we will still incorporate household and category heterogeneity:[14]

$$b_{jhc} = \pi_j + \zeta_c + \delta_h, \tag{12}$$

where π_j gives the mean impact for covariate j, ζ_c gives the variance across categories, and δ_h gives the variance across households. Thus, this baseline model not only tests the usefulness of the reduced-space spatial representation but also the nonlinear/distance specification for category attractiveness; in other words, it allows us to test the hypothesis that a simpler model—one that does not recognize that retailers can position their category offerings closer to what consumers want using long-run merchandising policies and that not all categories have the same importance in driving store choice—fits as well as our proposed model.

[13] We could have harvested all covariates used in the store choice model; however, our primary interest is the extent to which household demographics determine the importance of the category in driving store choice.

[14] Recall in our proposed model that category heterogeneity is parameterized in the spatial model specification.

6. Estimation

Let θ_{V^S} denote the set of parameters specified in V_{hst}^S, the deterministic component of the indirect utility of household h choosing store s on trip t. Similarly, let θ_{V^C} denote the set of parameters specified in V_{hsct}^C, the deterministic component of the indirect utility of household h choosing to purchase in category c at store s on trip t. The probability that we observe household h ($h = 1, 2, \ldots, H$) purchasing a subset of categories c at store s ($s = 1, 2, \ldots, S$) on trip t ($t = 1, 2, \ldots, T$) can be written as

$$\begin{aligned} & \Pr(y_{hst} = 1 \cap C) \\ & = \Pr(y_{hst} = 1 \mid \theta_{V^S}) \times \prod_{c \in C} [(y_{hct} = 1 \mid y_{hst} = 1, \theta_{V^C})^{y_{hsct}} \\ & \qquad \times (1 - \Pr(y_{hct} = 1 \mid y_{hst} = 1, \theta_{V^C}))^{1 - y_{hsct}}], \end{aligned} \tag{13}$$

where $y_{hcst} = 1$ if household h purchases in category c at store s on trip t, 0 otherwise. The first leading term in Equation (13) gives the probability that household h will visit store s on trip t. Under the usual assumption that the error terms ε_{hst} have Gumbel distributions, the probability that household h will visit store s on trip t is given by

$$\Pr(y_{hst} = 1 \mid \theta_{V^S}) = \frac{e^{V_{hst}^S}}{\sum_{j=1}^S e^{V_{hjt}^S}}. \tag{14}$$

The remaining terms in Equation (13) give the probability that household h purchases in category c on trip t, conditional on shopping at store s. Assuming a binary logit model for the distribution of ε_{hsct}, the probability that household h purchases in category c on trip t conditional on choosing store s can be written as

$$\Pr(y_{hct} = 1 \mid s) = \frac{1}{1 + e^{-V_{hsct}^C}}. \tag{15}$$

Letting $\theta_h = [\theta_{V^S}, \theta_{V^C}]$, we can write the likelihood for the store choice and category incidence models as

$$\mathbf{L} = \prod_{h=1}^{H} \int_{-\infty}^{\infty} \prod_{t=1}^{T} \prod_{s=1}^{S} \Pr(y_{hst} = 1 \cap C \mid \theta_h)^{y_{hst}} f(\theta \mid \Sigma)\, \partial\theta, \tag{16}$$

where θ denotes the global set of store choice and category incidence parameters, Σ denotes the parameter covariance matrix, and $f(\theta \mid \Sigma)$ is the distribution of the parameter vector θ conditional on the covariance matrix Σ.[15] We assume that this distribution is multivariate normal. To account for heterogeneity across household purchase incidence and store choice decisions, we use a continuous distribution with the parameter covariance matrix Σ. To reduce the dimensionality of the covariance matrix Σ, we use a two-factor structure of the Cholesky, including parameters for unique components of the variance (see, e.g., Hansen et al. 2006, Briesch et al. 2009). The parameters in Equation (16) can be estimated using simulated maximum likelihood. Here, we use what is analogous to a mixed nested logit estimation procedure (see Train 2003, Chapter 6 and §7.6) with Halton sequences for the numerical integration implemented with a quasi-Newton algorithm and user-supplied (i.e., analytic) gradients.[16]

7. Data

We use a multioutlet panel data set from Charlotte, North Carolina that covers a 104-week period between September 2002 and September 2004. Panelists recorded all packaged and nonpackaged goods purchases using in-home scanning equipment; purchases made in all grocery and nongrocery stores are included so that the data are not limited to a small sample of grocery stores. This is important because packaged goods purchases are frequently made outside of grocery stores. Households are included in the sample if at least 80% of their purchases were made at the five store chains (four supermarkets and one mass merchandise supercenter) for which we have geolocation data and if they spent at least $10 every month.[17] The resulting data set includes 368 families with a total of almost 40,000 shopping trips. Descriptive statistics for these households are provided in Table 2, which provides information on household shopping behaviors and demographics. We use the first 26 weeks of data as an initialization period to identify categories purchased by each household as well as to identify the intertemporal variables for categories. We use the middle 52 weeks as an estimation sample and the final 26 weeks as a validation sample.

Table 2 Household Descriptive Statistics

	Mean	SD
Shopping behaviors		
Number of households	368	
Monthly spending	216.1	107.76
Average spend per trip	34.8	12.82
Number of trips	165.7	64.86
Number of stores visited	4.3	1.13
Trip share at favorite store (%)	57	18
Demographics		
HH income ('000) ($)	53.5	24.3
HH size	2.8	1.2
Children in HH (%)	34	47
Ethnicity (Caucasian) (%)	89	32
Elderly ($>$ 64 yrs. old) (%)	11	31
Education (college or above) (%)	38	49
Married (%)	81	39

Note. HH, household.

We have detailed price information for 290 categories. From those categories, we selected the top 80 based on total dollars spent in the category; only categories that were not substantial (i.e., in which fewer than 10% of the households purchased) were excluded. The top 80 categories comprise more than 75% of the average market basket (of products tracked by UPC). Table 3 presents category penetration rates and share of total grocery spending for each category, along with the average price for each retail chain. Store statistics are shown in Table 4; this table provides information about store loyalty, travel time from home to the closest store of the retail chain, spending per trip, and averages of the long-run merchandising covariates used in the spatial model, where five of the nine long-run merchandising covariates have been indexed to provide a relative measure. The long-run merchandising covariate statistics tells us something about the (relative) character of each retailer. For example, Walmart offers shoppers more national brands to choose from and, consistent with its EDLP format, offers fewer advertised items but uses more displays. Winn-Dixie, on the other hand, offers shoppers fewer brand choices but overindexes on the percentage of private label brands, whereas Food Lion overindexes on the number of UPCs per brand and BI-LO utilizes more frequent feature advertising than the other retailers. As we might expect, Walmart offers the lowest long-run average prices of any retailer.

It is worth highlighting two potentially interesting characteristics of the shopping behavior of our panelists. First, each household visits, on average, four different stores over the duration of the data and visits their favorite store on 57% of their trips. Thus, the data show substantial intertemporal variability in store choice.

[15] It is true that a store visit implies that the shopper made a purchase in at least one category. However, there are a total of 290 nonperishable categories (of which the 80 largest are modeled), along with perishable categories that are not included in the data set. Thus, it is possible that a shopping trip could be made without buying one of the 80 modeled categories. As a result, we do not have to impose the binding constraint that given a store choice decision, at least one modeled category must be purchased.

[16] To test the performance of our model, we performed a series of simulation experiments; details can be obtained by contacting the first author.

[17] The last criterion was used to ensure that panelists were faithful in recording their purchases and remained in the panel for the entire 104-week period.

Table 3 Category Descriptive Statistics

	Penetration rate	Share of spend	Average prices				
Category	(%)	(%)	BI-LO	Food Lion	Harris Teeter	Winn-Dixie	Walmart
Carbonated beverages	97.28	6.03	4.02	3.98	4.80	4.05	3.77
Cigarettes	23.10	4.46	23.24	22.50	24.32	23.48	24.36
Cold cereal	94.57	3.08	3.01	2.98	3.22	3.32	2.62
FZ dinners/entrees	80.71	3.01	3.01	2.95	3.53	3.03	2.78
Fresh bread and rolls	99.18	2.96	1.56	1.52	1.67	1.52	1.49
Salty snacks	96.47	2.75	3.37	3.33	3.65	3.51	3.05
Beer/ale/alcoholic cider	32.61	2.60	14.71	14.58	14.59	14.39	14.59
Milk	98.37	2.39	0.43	0.42	0.44	0.42	0.42
Natural cheese	92.12	2.12	4.30	4.13	4.74	4.59	3.79
Cookies	92.66	2.00	2.88	2.76	3.21	2.75	2.40
Crackers	95.65	2.08	3.19	3.07	3.55	3.16	2.65
Luncheon meats	87.77	1.98	3.29	3.39	3.80	3.38	3.07
Breakfast meats	84.78	2.02	3.60	3.45	3.89	3.77	3.18
Total chocolate candy	90.49	2.21	3.73	3.70	4.18	3.88	3.52
Dog food	50.82	1.44	0.90	0.81	1.01	0.95	0.77
Ice cream/sherbet	85.60	1.33	1.01	1.00	1.14	1.09	0.94
FZ pizza	57.61	1.28	2.99	2.87	3.32	2.93	2.61
FZ poultry	58.97	1.23	2.92	2.81	3.58	3.11	2.74
Cat food	34.51	1.14	1.12	1.05	1.30	1.29	1.03
Soup	96.20	1.89	1.45	1.39	1.49	1.48	1.28
Coffee	68.21	1.40	3.59	3.69	3.95	3.69	3.32
RFG salad/coleslaw	85.87	1.04	2.16	2.51	2.69	2.51	2.16
Pet supplies	50.00	1.44	0.38	0.86	1.40	0.21	0.28
Processed cheese	87.50	1.23	3.14	2.99	3.18	3.13	2.97
Wine	22.01	1.05	4.13	4.12	4.45	4.36	3.87
Laundry detergent	86.14	1.09	0.90	0.87	0.94	0.89	0.81
Vegetables	95.65	1.23	0.80	0.78	0.89	0.82	0.68
Toilet tissue	91.85	1.04	0.47	0.48	0.51	0.50	0.41
FZ seafood	48.37	1.03	4.70	4.02	4.67	4.14	3.54
Snack bars/granola bars	51.36	0.86	5.17	5.14	5.49	5.35	4.98
Total nonchocolate candy	83.15	1.07	3.58	3.42	3.75	3.41	2.80
FZ novelties	51.09	0.62	2.06	1.94	2.23	2.11	1.79
Paper towels	80.98	0.91	1.73	1.70	1.85	1.76	1.53
Household cleaner	79.89	0.92	1.93	1.84	2.04	1.93	1.67
Dry packaged dinners	72.28	0.89	2.59	2.44	2.76	2.72	2.25
Internal analgesics	61.14	0.91	1.56	1.37	1.65	1.65	0.85
Dough/biscuit dough—RFG	76.63	1.03	2.18	2.04	2.40	2.21	1.98
Frankfurters	67.39	0.74	2.72	2.60	3.07	2.73	2.31
Vitamins	42.66	0.87	0.74	0.80	1.03	0.93	0.61
RFG juices/drinks	75.54	0.89	0.60	0.58	0.67	0.59	0.51
Yogurt	55.71	0.61	1.58	1.44	1.65	1.57	1.49
Bottled water	44.29	0.55	2.29	2.02	2.55	2.47	1.99
Soap	76.90	0.81	2.78	2.68	2.81	2.86	2.42
Toothpaste	76.63	0.84	9.56	9.77	10.15	8.55	8.64
Pastry/doughnuts	58.70	0.71	2.94	2.95	3.12	2.93	2.55
Cold/allergy/sinus tablets	38.59	0.66	3.12	2.81	3.59	3.52	2.02
Salad dressings—SS	78.80	0.67	2.29	2.26	2.49	2.50	2.06
FZ plain vegetables	69.29	0.87	1.75	1.63	1.88	1.97	1.45
Canned/bottled fruit	84.51	0.76	1.28	1.24	1.38	1.29	1.14
Snack nuts/seeds/corn nuts	57.61	0.60	3.63	3.58	3.95	3.56	2.98
Baking mixes	82.88	0.71	1.36	1.32	1.49	1.46	1.17
Bottled juices—SS	78.53	0.68	0.60	0.58	0.63	0.61	0.54
Skin care	35.87	0.59	16.04	15.49	18.16	18.66	15.91
FZ breakfast food	50.27	0.61	2.69	2.66	2.79	2.70	2.32
FZ bread/FZ dough	61.41	0.68	2.22	2.19	2.31	2.28	1.92
Canned meat	56.79	0.58	2.66	2.64	2.78	2.73	2.49
FZ meat	42.39	0.53	3.56	2.68	3.24	3.06	2.68
Dish detergent	83.97	0.59	1.40	1.37	1.55	1.47	1.28
Spices/seasonings	80.98	0.65	4.97	4.92	5.92	4.79	4.89
Cups and plates	60.33	0.56	3.30	3.22	3.99	3.43	2.83
FZ appetizers/snack rolls	38.86	0.62	3.67	3.45	4.20	3.97	3.25
RFG fresh eggs	96.74	0.52	0.11	0.12	0.12	0.14	0.11
Shampoo	59.78	0.52	2.91	2.95	3.14	2.79	2.78
Batteries	55.43	0.70	0.91	0.95	1.00	0.97	0.88
Pickles/relish/olives	72.83	0.56	1.84	1.80	2.13	1.70	1.54
Margarine/spreads/butter	86.14	0.55	1.24	1.17	1.41	1.29	1.07
Dinner sausage	44.02	0.55	3.04	3.09	3.62	3.09	2.71
Deodorant	61.14	0.49	14.77	14.52	16.64	14.71	13.13

Table 3 (Cont'd.)

	Penetration rate	Share of spend	Average prices				
Category	(%)	(%)	BI-LO	Food Lion	Harris Teeter	Winn-Dixie	Walmart
FZ desserts/topping	60.33	0.54	2.73	2.64	2.98	2.78	2.26
Shortening and oil	80.43	0.55	1.29	1.33	1.45	1.37	1.24
Baking needs	72.01	0.64	2.65	2.39	2.78	2.69	2.33
SS dinners	61.14	0.55	1.39	1.36	1.49	1.38	1.24
Toaster pastries/tarts	41.30	0.44	1.94	1.89	2.02	2.11	1.74
Air fresheners	48.91	0.47	4.21	4.37	4.62	4.46	4.15
Toothbrush/dental accesories	47.01	0.51	27.72	23.66	20.69	21.49	22.15
Food and trash bags	84.51	0.50	1.18	1.10	1.20	1.29	0.90
Spaghetti/Italian sauce	71.47	0.54	1.04	1.03	1.20	1.10	0.98
Sanitary napkins/tampons	42.39	0.46	1.29	1.32	1.44	1.35	1.17
Seafood—SS	62.23	0.46	2.16	2.24	2.52	2.23	2.02
Peanut butter	64.40	0.46	1.73	1.70	1.85	1.83	1.61

Note. FZ, frozen; RFG, refrigerated; SS, shelf stable.

Table 4 Store Descriptive Statistics

Variable	BI-LO	Food Lion	Harris Teeter	Winn-Dixie	Walmart
Loyalty (%)	13.54	32.01	19.43	11.68	23.34
	(0.19)	(0.27)	(0.27)	(0.18)	(0.20)
Travel time	20.33	9.33	17.51	17.45	42.01
	(20.38)	(7.37)	(13.21)	(17.27)	(27.10)
Spent/trip ($)	28.87	26.89	33.98	24.67	22.96
	(21.79)	(18.87)	(22.56)	(17.64)	(19.86)
Fav (%)	38.96	50.46	42.13	33.65	45.98
	(0.09)	(0.11)	(0.08)	(0.09)	(0.09)
NUPC[a] (%)	92.24	119.47	99.76	79.67	108.86
	(0.03)	(0.07)	(0.06)	(0.04)	(0.04)
NBrd[a] (%)	98.78	101.85	99.57	82.49	117.31
	(0.04)	(0.11)	(0.08)	(0.05)	(0.12)
NSize[a] (%)	98.21	105.34	97.47	95.94	103.05
	(0.02)	(0.03)	(0.03)	(0.02)	(0.03)
%*PL*[a] (%)	99.23	101.96	94.41	128.27	76.14
	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)
%*Natl*[a] (%)	88.37	96.70	101.91	71.17	141.85
	(0.01)	(0.05)	(0.02)	(0.02)	(0.03)
AdvF (%)	52.90	34.85	38.01	30.90	0.18
	(0.04)	(0.03)	(0.04)	(0.02)	(0.03)
DispF (%)	28.65	27.71	26.51	20.20	74.56
	(0.02)	(0.02)	(0.02)	(0.02)	(0.04)
AvgPrice ($)	0.99	0.98	1.11	1.02	0.90
	(0.003)	(0.003)	(0.003)	(0.002)	(0.002)

Note. Standard deviations are in parentheses.
[a]Numbers are indexed to provide a relative measure.

It is also useful to examine single-category shopping trips. Our data set includes 573 trips during which only a single category (one out of the 290 categories[18] with UPCs captured in our data set) was purchased. On 36% of these single-category trips, the shopper chose the store he or she visited most often. On 69% of single-category trips, however, the shopper

[18] To provide a more extensive examination of single purchase shopping trips we considered all 290 categories in our database as opposed to the 80 categories we selected for modeling purposes.

chose the store with the highest CDI (recall that CDI measures the extent to which retailers get more or less than their fair share of category sales). Thus, without the confounding effects of other categories in the market basket, we find that shoppers were almost twice as likely to choose the store "specializing" in the category as their favorite store, which strongly suggests that specific categories do indeed affect store choice.[19]

8. Results

We began by fitting several models and progressively increasing the number of dimensions specified for the latent multiattribute space. We stopped increasing the number of latent attribute dimensions when the BIC and CAIC information-theoretic statistics indicated that the improvement in log likelihood from adding an additional dimension did not compensate for the increase in model complexity.

8.1. Model Fit

Table 5 provides goodness-of-fit statistics for both in-sample and out-of-sample results. First, we see that all of the proposed models fit better than the baseline model. Recall that these proposed models include a spatial representation for the attraction parameter. Thus, the superior fit of the proposed models suggests that the relative positions of stores and category ideals in perceptual space may provide insights into the role of categories and category merchandising in store choice decisions and enable retailers to make their stores more attractive by better accommodating shoppers' preferences.

In terms of in-sample fit, the three-dimensional solution has lower BIC and CAIC information-theoretic statistics than either the two- or four-dimensional solutions. Table 5 also reports in- and out-of-sample log likelihoods along with hit rates. In terms of hit rates, all of the models perform equally well, although the three-dimensional solution provides the highest hit rates in and out of sample. The three-dimensional solution also yields the lowest out-of-sample log likelihood. Given that the three-dimensional solution fits the data best, we will focus on this solution in the remainder of our analyses and discussion.[20]

Table 5 Model Fit Summary

	Baseline	Two-dimensional	Three-dimensional	Four-dimensional
Estimation sample				
Households	368	368	368	368
No. of trips	25,472	25,472	25,472	2,572
Log likelihood	−552,084	−517,947	−514,035	−513,513
No. of parameters	194	387	503	617
Choice hit rate (%)	40.9	40.9	41.1	41.0
Incidence hit rate	87.4%	87.3%	89.4%	87.3
BIC	1,106,927	1,041,394	1,035,220	1,035,796
CAIC	1,107,121	1,041,781	1,035,723	1,036,413
Hold out sample				
Households	368	368	368	368
No. of trips	12,016	12,016	12,016	12,016
Log likelihood	−258,465	−242,439	−240,748	−240,765
Choice hit rate (%)	43.3	43.4	44.6	43.6
Incidence hit rate (%)	87.6	87.7	88.8	87.7

8.2. Parameter Estimates

Table 6 presents estimates of the store choice and category incidence parameters. The first set of columns pertains to the mean parameter estimates and the second set of columns to the heterogeneity standard errors.

8.2.1. Store Choice Equation. Focusing first on the store choice model, we see that all of the store intercept mean estimates, which reflect the intrinsic attraction of a store, are statistically significant.[21] The *Loyal* mean parameter estimate is also statistically significant and positive, consistent with previous findings of inertial behavior in store choice (Rhee and Bell 2002). The *Dist* mean parameter estimate is statistically significant and negative; i.e., all else the same, households prefer to shop at stores that are closer to them.

We see that almost all of the heterogeneity standard deviations in the rightmost panel of Table 6 are statistically significant. The exception is *Dist,* which indicates that households are homogeneous in their preference to shop at stores closer to them. It appears that households are heterogeneous in their inertial behavior (*Loyal*) and are slightly more heterogeneous in their intrinsic preference for Walmart compared with the other retailers.[22]

8.2.2. Category Incidence Equation. Turning to the category incidence model, we see that all of the

[19] This analysis, similar to the CDI metric, is not informative about the *relative* impact of different categories on store choice, however.

[20] In response to one anonymous reviewer, we also compared the fit of the three-dimensional model to a simple model that includes only two covariates: *Dist* and *Loyal*. Whereas the simple model has only a slightly lower store-choice likelihood, it is less attractive than the proposed model for a number of reasons. One way to think about the proposed model is that it corrects bias in the simple model estimates arising from omitted variables and correlation with the error term—note that we observed large differences not only in the intercepts in the two models but also in the distance and loyalty parameters. Perhaps more importantly, the simple model has far less diagnostic or informative value to retailers; it does not include any of the category data. We believe that the alternative model described in §5.3 stands as the "theoretically relevant" baseline model because, by not including the category attraction parameter, it addresses the key hypothesis of the proposed model that category positioning does indeed matter.

[21] Unless otherwise noted, all parameters are statistically significant at the $p < 0.05$ level.

[22] Dividing a parameter mean value by its heterogeneity value gives a relative measure of response heterogeneity.

Table 6 Model Structural Parameters

	Mean			Household heterogeneity			Category heterogeneity		
	Value	SE	*t*-Value	Value	SE	*t*-Value	Value	SE	*t*-Value
Store choice									
τ-Food Lion	0.469	0.040	11.67	0.130	0.027	4.84			
τ-Harris Teeter	−0.382	0.050	−7.69	0.220	0.026	8.52			
τ-Winn Dixie	−0.131	0.059	−2.24	0.345	0.029	11.76			
τ-Walmart	1.501	0.130	11.64	0.225	0.042	5.39			
Loyal	4.050	0.041	97.65	0.189	0.040	4.46			
Dist	−0.261	0.020	−13.05	0.006	0.010	0.53			
Category incidence									
γ-Intercept	2.370	0.093	25.51	0.131	0.003	40.88	0.049	0.004	12.35
Time	11.624	0.264	44.05	1.080	0.293	3.68	0.757	0.269	2.82
Qnty	−0.649	0.028	−23.41	0.035	0.029	1.20	0.128	0.029	4.37
Time × *Qnty*	37.650	2.644	14.24	0.846	3.126	0.27	0.585	3.286	0.18
WKEnd	0.081	0.100	14.85	0.034	0.004	8.00	0.010	0.004	2.46
Price	−2.370	0.091	−25.93	0.369	0.004	102.47	0.033	0.004	9.08
Fadv	−0.081	0.130	−0.63	0.065	0.141	0.46	0.133	0.122	1.09
Disp	0.326	0.0144	22.65	0.0207	0.014	1.46	0.022	0.015	1.42
K-category importance									
$*Spnd*	4.805	2.521	1.99						
APT	−1.444	0.341	−4.23						
$*Spnd* × *APT*	−0.384	4.254	−0.09						

covariates are statistically significant with one exception: whether items in the category were feature advertised (*Fadv*). The *Fadv* result is consistent with the findings of Bodapati and Srinivasan (2006) and suggests that, after controlling for other covariates, shoppers are not significantly more likely to purchase in a category because of feature advertising. The *Price* and *Disp* parameter mean values have the expected algebraic signs—lower weekly prices and more frequent displays in the category increase the probability of category incidence. The *Time* parameter's mean estimate is positive, suggesting that the more time since the last category purchase, the more likely the household will purchase in the category on the current trip. The *Qnty* parameter's mean estimate is negative, suggesting that the greater the quantity of the last category purchase, the less likely the household is to purchase in the category on the current trip. The *Time* × *Qnty* interaction's positive mean estimate suggests that shoppers feel significant stock pressure, consistent with the findings of Assunção and Meyer (1993). Finally, perhaps not too surprising, making the shopping trip on the weekend increases the probability that the household will purchase in a category, as weekend trips may reflect stock-up trips that result in larger baskets.

Turning next to the household heterogeneity parameters, we see that four of the eight heterogeneity standard deviations in the first right panel of Table 6 are statistically significant. It appears that households are homogeneous in their response to *Qnty*, *Time* × *Qnty*, *Fadv*, and *Disp* but are heterogeneous with respect to category purchase incidence (γ-intercept), days since they last purchased in the category (*Time*), shopping day preferences (*WKEnd*), and weekly prices (*Price*). It is interesting to note that households appear to be somewhat more heterogeneous in their responses to *Time*, *Qnty*, and *Disp* compared with *Price*. Turning next to category heterogeneity, five of the eight heterogeneity standard deviations are statistically significant. Categories appear to be homogeneous with respect to *Time* × *Qnty*, *Fadv*, and *Disp*. Finally, there appears to be much more heterogeneity in category price response as compared to the other covariates.

8.2.3. Category Weights. As shown in Table 6, both *APT* and $*Spnd* parameter estimates are statistically significant and have the expected algebraic sign, although their interaction is nonsignificant. As expected, it appears that the impact of a category on store choice is greater if there is less time between category purchases (i.e., the category is purchased more frequently) and if the household's spending in the category is greater.

Table 7 summarizes the influence of household demographic characteristics on the importance that shoppers place on a category. For each household demographic variable shown in Table 2, Table 7 shows where we found a significant relationship between interhousehold variation in κ_{hc} and the households' demographic characteristic. We use a "S+" or "S−" in Table 7 to denote a positive or negative relationship. We see from the table that in 38 of the 80 categories, interhousehold variation in category importance is associated with one or more of the household demographic variables under consideration. Among the statistically significant relationships, we find that

Table 7 **Impact of Household (HH) Demographics on the Importance of a Category**

Category	Elderly (65+)	HH size	HH income	College or above	Married	Children in HH	Ethnicity (Caucasian)	No. of significant parameters
1 Carbonated beverages							S−	1
3 Cold cereal							S−	1
6 Salty snacks				S−		S+		2
11 Crackers	S+							1
12 Luncheon meats				S−				1
14 Total chocolate candy					S+			1
15 Dog food	S−							1
17 FZ pizza	S+							1
21 Coffee	S+			S+				2
22 RFG salad/coleslaw				S+				1
23 Pet supplies						S+		1
25 Wine	S−					S−	S−	3
27 Vegetables					S+			1
28 Toilet tissue			S−					1
31 Total nonchocolate candy	S−							1
33 Paper towels				S−		S+		2
34 Household cleaner			S+		S−			2
36 Internal analgesics			S−					1
37 Dough/biscuit dough—RFG			S+		S−			2
38 Frankfurters		S+	S−			S+		3
39 Vitamins	S+							1
41 Yogurt		S−		S+		S+		3
42 Bottled water							S−	1
45 Pastry/doughnuts			S−					1
48 FZ plain vegetables					S−			1
50 Snack nuts/seeds/corn nuts		S+						1
51 Baking mixes			S−		S+			2
52 Bottled juices—SS				S+				1
53 Skin care			S+			S−		2
56 Canned meat			S−	S+		S+		3
63 Shampoo	S+						S+	2
67 Dinner sausage				S−				1
68 Deodorant		S+				S−		2
73 Toaster pastries/tarts							S+	1
74 Air fresheners		S−				S+		2
76 Food and trash bags	S+							1
77 Spaghetti/Italian sauce						S+		1
79 Seafood—SS		S−				S+		2

Notes. FZ, frozen; RFG, refrigerated; SS, shelf stable. "S+" indicates a statistically significant positive relationship; "S−" indicates a statistically significant negative relationship.

elderly households (65 years of age or older) place greater importance on crackers, breakfast meats, coffee, and vitamins, for example, whereas households with children place more importance on yogurt, paper towels, frankfurters, salty snacks, spaghetti, and Italian sauce. In general, although we find significant covariation between interhousehold category importance and household demographics, the relationships were for the most part weak; household demographics accounted for less than 2% of the variation in category importance and, across all comparisons, about 20% of the possible relationships were statistically significant at the $p < 0.10$ level and less than 10% at the $p < 0.05$ level.

8.2.4. Store Positions and Long-Run Merchandising Parameters. We find that all of the store position parameters are statistically significant across all three latent attribute dimensions. Interestingly, although there is a demonstrable relationship between physical geography and the derived perceptual store distance (i.e., we find that approximately 22% of the variation in the perceptual distances between stores is explained by median travel time) nearly four-fifths of the variation in perceptual store distances is not explained by the geographic location of the stores.

The long-run merchandising parameters were estimated for each of the three latent attribute dimensions. We find that 20 of the 27 mean parameter estimates are significant. There is not much intuition in the store positioning or long-run merchandising parameters, however, because of the dimensionality of our model. As a consequence, we have left a detailed discussion of these parameter estimates for the Web appendix.

9. Policy Analysis

Though our reduced-form model does not allow us to recommend categories that could *potentially* serve in the destination role for a retailer, it does allow us to identify categories that *actually* serve in the destination role. In this section, we investigate the influence of *BaskUtil* and its components, as well as long-run merchandising policies in influencing store choice decisions.

9.1. Measuring Category Impact on Store Choice Decisions

Recall from Equation (8) that *BaskUtil* measures a household's utility for the entire basket of intended category purchases on a particular trip if the retailer's store were chosen. *BaskUtil* incorporates the effects of retailers' long-run merchandising policies and category positions, weighting each category in the basket by its importance to the household's store choice decision. By decomposing *BaskUtil* into its category components, we can acquire a more nuanced view of the impact of a store's merchandising decisions and identify which stores have been delivering more value in the category. The decomposition is accomplished by averaging the category's contribution to *BaskUtil* for a given store across households and trips, then scaling by household penetration (percent households) and average purchase frequency (number per year). Scaling by household penetration and purchase frequency adjusts the utility of the category on an "average" trip to reflect the number of trips on which the category is sought.

Table 8 presents the decomposed *BaskUtil* values. Categories are arranged in descending order of utility. The table also provides the average category utilities as well as the sales rank of each category. We begin by looking across categories to assess differences in category utility. Three interesting results are evident:

1. Category utility is highly skewed—the top four categories have nearly the same collective impact on store choice decisions as the other 76 categories.

2. There appears to be a relationship between average category utility and category sales—the top ten categories in terms of utility include five of the top ten categories in sales ranking.[23] That categories with higher sales generally have a greater impact on store choice decisions may not be too surprising;[24] however, there are categories with relatively low sales such as yogurt, refrigerated salad/coleslaw, and coffee that have a high impact on store choice decisions.

3. We find no systematic pattern in average category utility rankings in terms of category type (e.g., perishable, dry grocery, nonfood). Across the top quartile of category utilities, we find nearly an even split between perishable/fresh and dry grocery categories with somewhat fewer nonfood categories. If anything, dry grocery categories are slightly more prominent among the top quartile of categories than other types.

We now turn our attention to differences in category utilities across retailers. Overall, we find that Harris Teeter has the highest average category utility (0.378) followed in order by BI-LO (0.333), Winn-Dixie (0.311), Food Lion (0.285), and Walmart (0.153). These average utilities reflect the percentage of categories in which each retailer offers shoppers higher-than-average utility: Harris Teeter (85% of categories), BI-LO (68% of categories), Winn-Dixie (51% of categories), Food Lion (36% of categories), and Walmart (19% of categories). Relative to the other retailers, Harris Teeter appears to be providing relatively high value for shoppers in more categories than any retailer, whereas Walmart provides relatively high value in the fewest categories. In which categories does Walmart offer high value? All are nonfoods—deodorant, cold/allergy/sinus, batteries, household cleaners, shampoo, soap, toothbrushes, toothpaste, skin care, vitamins, pet supplies and air fresheners—with the exception of chocolate and nonchocolate candies.[25] Walmart is the only mass merchandiser among the retailers in our data set, so our finding that it offers higher-than-average value in non-food categories provides face validity for the *BaskUtil* measure. In light of Walmart's strong competitive position, it may seem counterintuitive that it does not offer high value in food categories compared with grocery retailers. Yet Walmart has by far the largest store choice intercept (see Table 6), suggesting that the retailer is generally preferred to its grocery retail competitors.[26] In addition, Walmart's low average prices (see Table 4) together with the large negative price parameter (see Table 6) imply that shoppers are drawn to Walmart for its low prices, not necessarily because of its expertise in specific food categories.

Earlier, we argued that CDIs lack diagnostic value in terms of category importance in store choice decisions. Indeed, our primary motivation for developing *BaskUtil* and its decomposition was to create a

[23] Category sales explain 48.1% of the variation in average category utility values.

[24] *BaskUtil* depends on κ, which is a linear function of household spending in the category, and our decomposed category utility value is adjusted to reflect the number of trips. Both category spending and number of trips play a role in determining category sales.

[25] An analysis of retailer merchandising for total chocolate and nonchocolate candy categories shows that Walmart offers far more brands and UPCs per brand and is much more likely to carry shoppers' favorite brands than any other retailer.

[26] As a mass merchandiser, Walmart offers many categories that the grocery retailers do not; this is also likely reflected in its store choice intercept.

Table 8 Decomposition of *BaskUtil*: Category Utilities

Rank	Category	BI-LO	Food Lion	Harris Teeter	Winn-Dixie	Walmart	Average	Sales rank
1	Carbonated beverages	7.717	6.029	6.946	7.115	2.854	6.132	1
2	Salty snacks	3.245	2.865	3.443	2.917	1.481	2.790	6
3	Fresh bread and rolls	1.922	1.578	2.151	1.758	0.747	1.631	5
4	RFG salad/coleslaw	0.897	1.117	1.734	1.363	0.406	1.104	22
5	Crackers	0.984	0.939	1.525	0.791	0.537	0.955	11
6	Beer/ale/alcoholic cider	1.045	1.246	1.069	0.784	0.210	0.871	7
7	Yogurt	0.941	0.476	1.389	0.690	0.408	0.781	41
8	Cold cereal	0.905	0.690	0.920	0.879	0.379	0.754	3
9	Coffee	0.598	0.612	0.920	0.710	0.340	0.636	21
10	FZ breakfast food	0.536	0.551	0.736	0.504	0.160	0.498	54
11	Toilet tissue	0.538	0.446	0.533	0.651	0.282	0.490	28
12	Cups and plates	0.495	0.283	0.711	0.508	0.297	0.459	60
13	Milk	0.564	0.450	0.573	0.504	0.202	0.459	8
14	Dough/biscuit dough—RFG	0.541	0.376	0.450	0.539	0.168	0.415	37
15	FZ dinners/entrees	0.475	0.443	0.610	0.368	0.174	0.414	4
16	Deodorant	0.368	0.257	0.351	0.210	0.708	0.379	68
17	Pastry/doughnuts	0.312	0.423	0.617	0.352	0.189	0.378	45
18	Wine	0.375	0.444	0.304	0.310	0.088	0.304	25
19	Bottled water	0.290	0.127	0.523	0.181	0.272	0.279	42
20	Laundry detergent	0.261	0.286	0.337	0.217	0.224	0.265	26
21	Ice cream/sherbet	0.243	0.260	0.392	0.260	0.064	0.244	16
22	Dish detergent	0.217	0.187	0.286	0.305	0.212	0.241	58
23	Toaster pastries/tarts	0.262	0.178	0.262	0.350	0.114	0.233	73
24	Canned meat	0.328	0.251	0.186	0.285	0.112	0.233	56
25	FZ desserts/topping	0.196	0.242	0.324	0.295	0.058	0.223	69
26	Internal analgesics	0.190	0.194	0.282	0.213	0.177	0.211	36
27	Snack nuts/seeds/corn nuts	0.175	0.131	0.298	0.154	0.160	0.183	50
28	FZ novelties	0.230	0.189	0.260	0.170	0.037	0.177	32
29	Baking mixes	0.190	0.173	0.190	0.213	0.053	0.164	51
30	Cold/allergy/sinus tablets	0.147	0.103	0.182	0.126	0.247	0.161	46
31	Seafood—SS	0.164	0.142	0.218	0.140	0.055	0.144	79
32	Bottled juices—SS	0.157	0.144	0.163	0.137	0.062	0.132	52
33	FZ appetizers/snack rolls	0.139	0.118	0.225	0.102	0.051	0.127	61
34	FZ meat	0.115	0.075	0.228	0.082	0.038	0.108	57
35	Batteries	0.051	0.043	0.060	0.031	0.164	0.070	64
36	Soup	0.080	0.069	0.090	0.074	0.025	0.068	20
37	Cookies	0.074	0.069	0.100	0.051	0.040	0.067	10
38	Total nonchocolate candy	0.046	0.047	0.060	0.040	0.066	0.052	31
39	Vegetables	0.059	0.051	0.056	0.065	0.015	0.049	27
40	Paper towels	0.048	0.029	0.044	0.049	0.025	0.039	33
41	Pickles/relish/olives	0.041	0.036	0.059	0.045	0.011	0.039	65
42	Cigarettes	0.031	0.053	0.014	0.025	0.008	0.026	2
43	Household cleaner	0.023	0.018	0.030	0.021	0.033	0.025	34
44	Shampoo	0.019	0.016	0.018	0.016	0.047	0.023	63
45	Natural cheese	0.027	0.019	0.029	0.027	0.008	0.022	9
46	Dry packaged dinners	0.024	0.019	0.023	0.018	0.009	0.019	35
47	Margarine/spreads/butter	0.021	0.016	0.022	0.025	0.007	0.018	66
48	Soap	0.013	0.015	0.014	0.013	0.032	0.017	43
49	Shortening and oil	0.018	0.016	0.023	0.021	0.006	0.017	70
50	SS dinners	0.020	0.023	0.016	0.015	0.007	0.017	72
51	Dog food	0.019	0.015	0.019	0.017	0.013	0.016	15
52	Luncheon meats	0.021	0.017	0.019	0.021	0.005	0.016	12
53	FZ pizza	0.020	0.015	0.026	0.011	0.007	0.016	17
54	Processed cheese	0.019	0.014	0.016	0.016	0.006	0.014	24
55	FZ seafood	0.025	0.015	0.015	0.012	0.004	0.014	29
56	Toothbrush/dental accesories	0.014	0.013	0.011	0.006	0.025	0.014	75
57	RFG fresh eggs	0.013	0.012	0.016	0.015	0.004	0.012	62
58	Toothpaste	0.008	0.007	0.015	0.005	0.016	0.010	44
59	Cat food	0.013	0.009	0.013	0.009	0.007	0.010	19
60	RFG juices/drinks	0.013	0.010	0.014	0.010	0.003	0.010	40
61	Frankfurters	0.010	0.008	0.011	0.011	0.003	0.009	38
62	Breakfast meats	0.009	0.009	0.010	0.010	0.003	0.008	13

Table 8 **(Cont'd.)**

Rank	Category	BiLo	FoodLion	Harris Teeter	Winn-Dixie	Walmart	Average	Sales rank
63	Spaghetti/Italian sauce	0.007	0.009	0.013	0.008	0.003	0.008	77
64	Skin care	0.002	0.003	0.004	0.003	0.023	0.007	53
65	Peanut butter	0.009	0.007	0.008	0.007	0.004	0.007	80
66	Canned/bottled fruit	0.008	0.005	0.009	0.007	0.002	0.006	49
67	Spices/seasonings	0.005	0.006	0.008	0.004	0.003	0.005	59
68	Total chocolate candy	0.005	0.004	0.007	0.003	0.007	0.005	14
69	Baking needs	0.005	0.004	0.007	0.005	0.003	0.005	71
70	FZ bread/FZ dough	0.004	0.003	0.005	0.004	0.001	0.004	55
71	Snack bars/granola bars	0.004	0.003	0.003	0.003	0.003	0.003	30
72	Dinner sausage	0.003	0.003	0.003	0.003	0.001	0.002	67
73	Food and trash bags	0.002	0.002	0.002	0.002	0.002	0.002	76
74	Vitamins	0.001	0.001	0.001	0.001	0.004	0.001	39
75	FZ poultry	0.001	0.001	0.003	0.002	0.001	0.001	18
76	FZ plain vegetables	0.001	0.001	0.002	0.001	0.000	0.001	48
77	Pet supplies	0.000	0.001	0.002	0.000	0.001	0.001	23
78	Sanitary napkins/tampons	0.001	0.001	0.001	0.000	0.001	0.001	78
79	Salad dressings—SS	0.001	0.000	0.001	0.001	0.000	0.000	47
80	Air fresheners	0.000	0.000	0.001	0.000	0.001	0.000	74

Note. FZ, frozen; RFG, refrigerated; SS, shelf stable.

metric that shows which categories are most important to shoppers when choosing a store while also capturing the relative value that different stores provide in each category. Analyzing category utilities across retailers *within a category* enables us to determine the relative value provided by a retailer in a specific category. If a retailer provides more value in a specific category than competing retailers, we would expect shoppers to be more likely to choose that retailer when they intend to purchase in the category and, as a result, be more likely to actually purchase in the category at that retailer. Accordingly, we calculated within-category correlations (i.e., across stores) between the category utility values shown in Table 8 and their corresponding CDI values, which appear in Table 1. Across the 80 categories in our data set, we generally find strong positive correlations: the mean correlation is 0.684, 81.3% of the correlations are greater than 0.5, and 65.0% of the correlations are greater than 0.75.[27] Interestingly, computing correlations *within-store* (i.e., across category) for the five retailers reveals small and nonsignificant relationships (the mean correlation is 0.105). Thus, although the category utilities obtained from decomposing *BaskUtil* are correlated with a store's *relative* category development, they are more informative about which categories have greater impact on store choice for a given retailer. This is important in understanding how relative category value, based in part on long-run merchandising policies, attracts shoppers to a retailer's stores.

[27] A nonparametric sign test for correlations greater than 0.75 yields less than a 0.01 probability that this would occur by chance.

9.2. Capturing the Impact of Long-Run Merchandising on Store Choice

The category utilities shown in Table 8 depend on the store-specific category attraction parameter (see Equation (4)) and the underlying multidimensional spatial model. In the perceptual model, the distance between a store's position and the shopper's category ideal is adjusted based on the effectiveness of the retailer's long-run category merchandising policies (see Equation (5)). However, because our best-fitting model includes three latent dimensions, and because store positions, category ideal points, and the algebraic signs of merchandising parameters all differ across dimensions, specific merchandising recommendations would be idiosyncratic and difficult to explain. On the other hand, the long-run category merchandising variables (e.g., favorite brands carried, number of brands, number of UPCs per brand, number of sizes per brand) are all mean-centered. By replacing the estimated parameters for these variables with zeros, we can thus determine what the category utility would have been if the retailer's long-run merchandising policies had been "average." By comparing these hypothetical values with the estimated category utilities in Table 8, we can parsimoniously evaluate the impact of retailers' long-run merchandising policies on store choice decisions.

Table 9 investigates the extent to which actual long-run merchandising policies adopted by each retailer resulted in higher category utilities than average policies would have produced for the top quartile of categories shown in Table 8. In the table, the value 1 indicates that the retailer's category utility is higher than that which would have been obtained under average merchandising, and the value 0 indicates that the retailer's category utility is lower than

Table 9 **Effective Merchandising Findings**

Utility rank	Category	BI-LO	Food Lion	Harris Teeter	Winn-Dixie	Walmart	% effectively merchandising
1	Carbonated beverages	0	0	0	1	0	20
2	Salty snacks	0	0	0	1	0	20
3	Fresh bread and rolls	1	0	0	1	0	40
4	RFG salad/coleslaw	1	0	0	1	0	40
5	Crackers	1	0	0	0	0	20
6	Beer/ale/alcoholic cider	0	0	0	1	1	40
7	Yogurt	0	0	0	0	0	0
8	Cold cereal	0	0	0	0	0	0
9	Coffee	0	0	1	0	0	20
10	FZ breakfast food	1	0	1	0	0	40
11	Toilet tissue	1	0	0	0	0	20
12	Cups and plates	1	0	1	0	0	40
13	Milk	0	0	0	1	0	20
14	Dough/biscuit dough—RFG	1	0	1	1	0	60
15	FZ dinners/entrees	0	0	0	0	0	0
16	Deodorant	0	0	0	0	0	0
17	Pastry/doughnuts	0	1	0	0	0	20
18	Wine	1	0	0	1	1	60
19	Bottled water	0	0	0	0	0	0
20	Laundry detergent	0	0	0	0	0	0
	Effectively merchandising (%)	40	5	20	0	10	23

Note. FZ, frozen; RFG, refrigerated.

what average merchandising policies would have produced. Interestingly, in 14 of the top 20 categories, at least one retailer merchandizes the category relatively effectively, while in 7 of the top 20 categories, a single retailer is merchandising effectively.[28] BI-LO and Winn-Dixie both merchandise 3 of the top 5 and 8 of the 20 most high-impact categories effectively. However, Winn-Dixie is the only retailer that is merchandising effectively in carbonated beverages, salty snacks, and milk—three categories that rank in the top 10 in terms of sales. In contrast, Food Lion merchandises only 1 of the top 20 categories (pastry and doughnuts) effectively, whereas Walmart merchandises only 2 of the top 20 categories (beer/ale/alcoholic cider and wine) effectively.

10. Concluding Remarks

Our study has focused on how individual categories impact store choice decisions and on developing a method for identifying destination categories. We model only consumer shopping behavior in response to retailers' marketing mix decisions, not retailers' response to consumer demand. This reduced-form modeling approach precludes us from recommending specific actions that retailers could implement to develop new destination categories.[29] Our data set includes the 80 largest of 290 total categories that are tracked by universal product code. Unfortunately, perishable categories are not included in our data set, so their impact on store choice is not incorporated in our model. This is a common shortcoming of panel data from syndicated data providers but is nevertheless a limitation of our analysis. Our analysis is also limited by geographic information. Store choices are recorded at the trip level, but we do not know where each trip originated and terminated. We therefore have assumed that each trip begins and ends at the panelist's home. Again, this is a common characteristic of panel data, but it does introduce measurement error into our model and underestimates the effect of geographic convenience on store choice. A final limitation of our study is that causal data for category purchases are pooled across stores. We assume that, if any store in a retail chain displays a particular SKU, then all stores in that chain display the SKU. In other words, we assume uniform implementation of display within a retail chain. This assumption also results in measurement error, underestimating the effect of display on category purchases.

Our research could be extended in the future in different ways. First, to recommend specific actions that retailers could implement to develop destination categories, a structural model that comprehensively captures the relationship between consumer shopping behavior and retailers' strategic marketing mix decisions (including merchandising) incorporating equilibrium behavior should be developed. Second, future researchers can develop more direct analytic methods to address the inherent endogeneity of category purchase and store choice. Third, our use of perceptional

[28] This is not to say that retailers may have other goals (e.g., profitability) that are not considered in our analysis.

[29] Our discussion of these issues has benefited from the comments of the associate editor and two anonymous reviewers.

distance mapping could be used to evaluate merchandising effectiveness at the category level and to suggest approaches to improving merchandising effectiveness category by category. Fourth, the efficacy of our model for selecting destination categories could be tested experimentally, either by matching stores of a given retailer within a geographic market or by comparing stores across geographic markets for the same retailer. Finally, the framework we have developed could be extended to address the question of how much shoppers buy; i.e., purchase quantity.[30]

Acknowledgments
The authors thank Information Resources, Inc. for providing the data used in this paper. Any mistakes or omissions are the sole responsibility of the authors.

Supplemental Material
Supplemental material to this paper is available at http://dx.doi.org/10.1287/mksc.2013.0775.

References
ACNielsen (2006) *Consumer-Centric Category Management* (John Wiley & Sons, Hoboken, NJ).

Ainslie A, Rossi PE (1998), Similarities in choice behavior across product categories. *Marketing Sci.* 17(2):91–106.

Arnold SJ, Tigert DJ (1982) Comparative analysis of determinants of patronage. Lusch RF, Darden WR, eds. *Retail Patronage Theory: 1981 Workshop Proc.* (Center for Management and Economic Research, University of Oklahoma, Tulsa), 118–124.

Arnold SJ, Ma S, Tigert DJ (1978) A comparative analysis of determinant attributes in retail store selection. Hunt HK, ed. *Advances in Consumer Research*, Vol. 5 (Association for Consumer Research, Ann Arbor, MI), 663–667.

Arnold SJ, Roth V, Tigert DJ (1981) Conditional logit versus MDA in the prediction of store choice. Monroe KB, ed. *Advances in Consumer Research*, Vol. 9 (Association for Consumer Research, Washington, DC), 665–670.

Arnold SJ, Oum TH, Tigert DJ (1983) Determinant attributes in retail patronage: Seasonal, temporal, regional and international comparisons. *J. Marketing Res.* 20(2):149–157.

Assunção JL, Meyer RJ (1993) The rational effect of price promotions on sales and consumption. *Management Sci.* 39(5):517–535.

Baumol WJ, Ide EA (1956) Variety in retailing. *Management Sci.* 3(1):93–101.

Bell DR, Lattin JM (1998) Shopping behavior and consumer preference for store price format: Why "large basket" shoppers prefer EDLP. *Marketing Sci.* 17(1):66–88.

Bell DR, Ho TH, Tang CS (1998) Determining where to shop: Fixed and variable costs of shopping. *J. Marketing Res.* 35(3):352–369.

Blattberg RC, Fox EJ, Purk ME (1995) *Category Management: A Series of Implementation Guides*, Vols. I–IV (Food Marketing Institute, Washington, DC).

Bodapati A, Srinivasan V (2006) The impact of feature advertising on customer store choice. Working paper, Stanford University, Palo Alto, CA.

Boatwright P, Nunes JC (2001) Reducing assortment: An attribute-based approach. *J. Marketing* 65(3):50–63.

Briesch RA, Chintagunta P, Fox E (2009) How does assortment affect grocery store choice? *J. Marketing Res.* 46(2):177–189.

Broniarczyk SM, Hoyer WD, McAlister L (2006) Retail assortment: More $\neq$ better. Krafft M, Mantrala MK, eds. *Retailing in the 21st Century: Current and Future Trends* (Springer, Berlin), 225–238.

Bucklin RE, Lattin JM (1992) A model of product category competition among grocery retailers. *J. Retailing* 68(3):271–293.

Chernev A, Hamilton R (2009) Assortment size and option attractiveness in consumer choice among retailers. *J. Marketing Res.* 46(3):410–420.

Chernev A, Mick DG, Johnson MD (2003) When more is less and less is more: The role of ideal point availability and assortment in consumer choice. *J. Consumer Res.* 30(2):170–183.

Chib S, Seetharaman PB, Strijnev A (2004) Model of brand choice with a NO-purchase option calibrated to scanner panel data. *J. Marketing Res.* 41(2):184–196.

Dhar S, Hoch SJ, Kumar N (2001) Affective category management depends on the role of the category. *J. Retailing* 77(2):165–184.

Erdem T, Imai S, Keane MP (2003) Brand and quantity choice dynamics under price uncertainty. *Quant. Marketing Econom.* 1(1):5–64.

Erdem T, Keane M, Sun B (2006) The impact of advertising on consumer price sensitivity in experienced goods markets. Working paper, New York University, New York.

Food Marketing Institute (1993) *Efficient Consumer Response: Enhancing Consumer Value in the Grocery Industry* (Food Marketing Institute, Washington, DC).

Fox EJ, Hoch SJ (2005) Cherry-picking. *J. Marketing* 69(1):46–62.

Hansen K, Singh V, Chintagunta P (2006) Understanding store-brand purchase behavior across categories. *Marketing Sci.* 25(1):75–90.

Hendel I, Nevo A (2009) A simple model of demand accumulation. Working paper, Northwestern University, Evanston, IL.

Hoch SJ, Bradlow ET, Wansink B (1999) The variety of an assortment. *Marketing Sci.* 18(4):527–546.

Huff D (1964) Redefining and estimating a trading area. *J. Marketing* 28(1):34–38.

Kahn B, Schmittlein D (1989) Shopping trip behavior: An empirical investigation. *Marketing Lett.* 1(1):55–69.

Keane MP (1997) Modeling heterogeneity and state dependence in consumer choice behavior. *J. Bus. Econom. Statist.* 15(3):310–327.

Kim B-D, Park K (1997) Studying patterns of consumer's grocery shopping trip. *J. Retailing* 73(4):501–517.

Kollat DT, Willett RP (1967) Consumer impulse purchasing behavior. *J. Marketing Res.* 4(1):21–31.

Manchanda P, Ansari A, Gupta S (1999) The "shopping basket": A model for multicategory purchase incidence decisions. *Marketing Sci.* 18(2):95–114.

McAlister L, George EI, Chien Y-H (2009) A basket-mix model to identify cherry-picked brands. *J. Retailing* 85(4):425–436.

Morales A, Kahn BE, McAlister L, Broniarczyk SM (2005) Perceptions of assortment variety: The effects of congruency between consumers' internal and retailers' external organization. *J. Retailing* 81(2):159–169.

Reilly WJ (1931) *The Law of Retail Gravitation* (Pillsbury, New York).

Rhee H, Bell D (2002) The inter-store mobility of supermarket shoppers. *J. Retailing* 78(4):225–237.

Train KE (2003) *Discrete Choice Methods with Simulations* (Cambridge University Press, Cambridge, UK).

Zhang Q, Gangwar M, Seetharaman PB (2010) Store loyalty as a category-specific trait: What drives it? Working paper, University of Iowa, Iowa City.

[30] This extension was suggested by an anonymous reviewer.