Do Promotions Increase Store Expenditures?
A Descriptive Study of Household Shopping Behavior

Abstract

An important question that has been raised in supermarket retailing is whether weekly promotions induce households to increase their in-store expenditures or merely reallocate a predetermined spending amount in that week. That is, are households' grocery shopping expenditures preset before entering the store or are *flexible* and determined while in the store as a function of the specific store offerings encountered during the store visit? This is an important question for the retailer in light of the vast array of temporary promotions offered to consumers. Indeed, should expenditures be fixed before entering the store (for instance, as a function of the household's inventory and/or income), it is possible that retailers might decrease their profitability when running promotions by displacing expenditures from high margin items to lower margin products.

We claim that to answer this question meaningfully one must consider the totality of the household's within-store purchases (i.e., the market basket) and not just purchases of the promoted products. Using a rich database that contains the entire basket of goods bought over time by households from a given supermarket chain, we attempt to describe the drivers of both the level of expenditure and its allocation over the different groups of products. We use an extended version of the Almost Ideal Demand System (AIDS) for this purpose and our empirical results provide convincing evidence that while household expenditures do increase with promotions, there is also a significant reallocation of expenditures among the different groups of products. This implies that retailers have to choose carefully which items are promoted and to what depth, if promotions are also to increase profits, not merely store level expenditures.

Key Words: *Consumer Demand Theory, Market Basket and Household Expenditures, AIDS Model, Econometric Estimation*

Introduction

Temporary price reductions are widely used by grocery retailers as a promotional vehicle in order to induce shoppers to visit the promoting retailer's stores and purchase not only the promoted product(s), but also other regular-priced products. Temporary price reductions can also have an in-store effect whereby consumers may be induced to make unplanned purchases of the promoted products. It is claimed that almost sixty percent of household supermarket purchases are unplanned and the result of in-store decisions (Inman and Winer, 1999).

Hence, temporary price reductions can serve the dual roles of *attracting* shoppers to the retailer's store and inducing them to *increase* their total shopping expenditures. Bell, Ho, and Tang (1998) investigate the store choice decision. In this study, we examine the second effect and attempt to provide some insights as to how within-store household shopping expenditures are influenced by the retailer's pricing and promotion strategies.

There is a wealth of evidence to support the fact that price promotions do indeed increase the sales of the promoted products. There has also been an extensive amount of research done on what type of promotions should be offered, how retailers should time their various promotions, and by how much they should discount their products (see the comprehensive review in Blattberg and Neslin 1990). For example, Walters and MacKenzie (1988) have shown that promotions increase store traffic and have some impact on store sales. However, a question that has not yet been investigated is: where does the money consumers spend on buying promoted goods come from?

There are potentially two opposing answers to this question. First, it could be the case that households have a fixed or predetermined expenditure for their grocery shopping, for any given period (say, a week). In this case, should a household respond to a promotion and make an unplanned purchases, it will come at the expense of reduced spending on one or more other products. Alternatively, it could be the case that a household's expenditures are not fixed and any unplanned purchases of promoted items would be in addition to the planned purchases.

Why is that an important question? Profit margins are typically lower on promoted goods than on regular-priced items, unless they are offered as a result of steep manufacturer discounts (Drèze 1996). Hence, when retailers offer promotions on their own accord (e.g., on store brands or produce), this question is of critical importance to retailers. Indeed, if

1

promotional purchases come at the expense of other regular-priced products, the retailer could be decreasing her total profits when running temporary price reductions, despite an increase in the sales volume(s) and profit(s) of the promoted product(s). This would indicate that although promotions might be profitable at the category level, they might be costly at the store level.

Conversely, if promotions do increase household-level expenditures during a given shopping trip, they are then profitable both at the category and the store level. In such a case, it will be useful for the retailer to know how such expenditure effects vary across different products. For example, will promoting meat have a greater impact on store expenditures than promoting alcoholic beverages? Do temporary price reductions have a greater impact on increasing expenditures than in-store displays or feature advertisements, for a given product? How do household inventory levels affect expenditures? Clearly, knowing the answers to such questions will be of benefit to the retailer.

Related to the above issue of household expenditures, is the issue of how household expenditures are allocated across different product groups[1]. That is, if promotions induce households to change their total shopping expenditures, do they also bring about a re-allocation of those expenditures? For example, suppose a retailer has a temporary price reduction on meat, and further suppose that this induces a given household to increase its total weekly expenditures. The following questions are then of particular interest. (1) What fraction of that increase in expenditure goes to the product being promoted? (2) Is there, in addition, a re-allocation of expenditures that is brought about because of the nature of the relationship (substitutes or complements) between pairs of product in the household's shopping basket? That is, does the promotion of say, meat, result in an increase (or decrease) in the share of the budget allocated to bread (say)?

The objective of the proposed research is to address the above questions pertaining to household shopping behavior. We aim to provide insights that can be used by *retailers* in planning their pricing and promotional activities. The goal is not to provide a pricing decision support system but rather to provide qualitative insights for planning at the store level planning rather than at the product-category level. We analyze household shopping behavior for the entire basket of goods that are bought on visits to the store. We determine whether household expenditures are fixed or flexible, and how such expenditures are allocated across the basket of items that households buy. This analysis is done conditional on a shopping trip happening. We

do not attempt to explain why the shopping trip occurs or why the household chose the given store over others. That question has been addressed elsewhere (Bell *et al.* 1998).   Rather, we look at what happens in the store given that a customer has entered it. Prior to laying out our analytical framework and describing the data, we review the current literature as it relates to the above issues.

The analysis of household purchase behavior in the marketing research literature has largely focused on single-category purchase decisions. Although the economics literature on the neo-classical theory of household choice behavior has historically focused on household purchases or expenditures of a basket of goods (see Deaton and Muellbauer 1980a), this tradition has not permeated the marketing area in a significant way. Only recently have researchers in marketing begun to focus on the purchases of multiple categories (see for example Chintagunta and Haldar 1998; Ainslie and Rossi 1998; Bell and Lattin 1998; and Manchanda, Ansari, and Gupta 1999).

The findings from those studies provide useful information for decision making by both manufacturers and retailers. For example, knowing whether the sensitivity to price is category or household specific (Ainslie and Rossi 1998) has clear implications for the both manufacturers and retailers. Likewise, knowing the type of household an EDLP (or Hi-Lo) store will attract (Bell and Lattin 1998) is important in choosing a retail-pricing format. However, from the standpoint of the objectives of this paper, the preceding studies do not address the issue of the determinants of the total household expenditures on shopping trips and whether some products are more effective than others in increasing such expenditures. Further, they do not analyze the entire shopping basket of households and the allocation of the expenditure across the different items in the basket. Instead, the focus has been on analyzing the dependencies across a limited number of categories (two to six).[2]

Based on the preceding discussion, the objective of the current study is to address the question of whether consumers are "expenditure fixed" or "expenditure flexible" in relation to their purchases of a basket of goods or products, given the decision to visit a certain store / chain. To address this issue, we use a rich database that contains information on the purchases of entire basket of goods by a large sample of households. Our analysis is done in two stages. First, we attempt to get as much insights as possible by examining the descriptive statistics obtained from the data. Next, we attempt to go beyond the descriptive statistics by using a

demand model which allows for both category substitution as well as flexibility in total store level expenditures. Therefore, we first model the household shopping-trip expenditure as a function of various drivers such as prices, inventory levels, household characteristics, etc. Next, we use the neo-classical economic theory of consumer demand (Deaton and Muellbauer 1980a) to determine how the chosen level of expenditure is allocated across the different products in a shopping basket. For this, we use an extended version of the AIDS model of consumer demand developed by Deaton and Muellbauer (1980b). We estimate jointly the parameters of this system of demand and total store expenditure from an extensive database that has as records the entire basket of goods bought by a sample of over 25,000 households over a one-year period.

Our main empirical findings are as follows. (1) Household's within-store *total* expenditures are indeed influenced by the pricing and promotional activities of the retailer, although the impact varies across different product categories. This result extends previous findings of expenditure effects observed at the category level to the entire basket. The importance of this finding is that even though previous studies have shown expansion at the category level (Drèze and Hoch 1998), the expenditure impact on other categories was not known. From the retailer standpoint, as opposed to the manufacturer standpoint, it is the holistic question that needs to be addressed. This result also extends the findings of Block and Morwitz (1999) whose findings show that households make a large number of unplanned purchases i.e. buy items not prewritten on a shopping list. We note that their finding by itself does not imply flexible spending because a large number of unplanned purchases doesn't necessarily imply increased spending as the proportion of unplanned purchase might be relatively constant overtime. In addition, household inventory levels also influence their spending decisions and higher levels of inventory reduce total spending within the store. This result implies that retailers must consider the trade-off between increasing current expenditures and reducing future expenditures when promoting various products. (2) The allocation of the chosen level of expenditure across the different products is also influenced by the pricing and promotional activities of the retailer. This last result reveals some useful insights regarding the nature of substitution and complementarity between pairs of products. For example, we find that while promotion of alcoholic products increase store expenditures, it also brings about a reallocation of expenditures among the different groups of products. From the standpoint of the retailer, these findings

collectively offer insights into the selection of products that should be promoted to enhance market performance.

The rest of the paper is organized as follows. In the next section, we describe the data, the method of aggregation (over products), and the classification of household shopping behavior that allows for meaningful analysis. The following section describes the model that is used in the analysis and how the parameters of this model are econometrically estimated. Next, we report the results of the estimation and make inferences from them about household purchasing behavior. We also discuss the implications of those findings for the retailer. The final section concludes with a summary of our findings and a discussion of directions for future research.

Data Description, Product Aggregation, and
Classification of Shopping Behavior

1. Data Description

We use in our empirical analysis a panel data set that contains the purchase records (i.e., the shopping basket) for each household in the sample of around 25,000 households making purchases at the sponsoring supermarket chain. The sample spans a one-year period from March 1, 1996 to February 28, 1997. The data are from a European source and were made available through a private arrangement. There are four stores belonging to the same chain, vary in size from 20,000 to 60,000 square feet, with average annual sales from €20 to €70 million. The four stores are in four different geographic locations and less than 1% of the households shopped in more than one of the four stores. The shopping format is of a Hi-lo type with at least one competitor of the same format in each market. Customers were uniquely identified through their frequent shopper cards.

For each household and each shopping trip to the given retailer, the data (taken from the cash register receipts) contain information on the date and time of the purchase, the items (SKUs) bought, the quantities purchased, and the price paid for the item. In addition, a separate file describes the promotional activity of each SKU for the year of interest. Promotions last for a week and run from Thursday of one week to the Wednesday of the next. Households using the unique identification cards constitute about 69% of all visits to the four stores. Thus the sample is restricted to the households shopping at the given store (in each market) and who have joined

the frequent shopper program. Although this is a non-random sample which might suffer from selection bias, we feel that its large size make it representative of the type of shoppers who patron the chain we study. As to the larger question of the representativeness of the sample of all shoppers in a given market, we do not have access to sales data from the competing stores. However, we do expect our shoppers to be fairly representative of the overall market as the chain is the major player in the country (with a 39% market share) and as only one major competitor (25% market share). This competitor has the same Hi-Lo type of format.

In addition to the purchase records, the data set contains some demographic information. Each household is associated with a given census tract code (roughly one city block), and for *each census tract* the data set contains information on average income, average number of persons per household, the social strata, and a measure of the food sales potential. Hence, the demographic information is limited because it is at the level of the census tract and not the household. Nevertheless, we do attempt to address the issue of whether there is any relationship between the demographic variables and household expenditures and allocation decisions.

2. Data Reduction

Data reduction is a thorny issue when dealing with large data sets like the one we analyze. Our data set describes the choices made by about 25,000 households, choosing any of 250,000 SKUs during over half a million shopping trips. Clearly, data reduction is in order and many approaches can be used to accomplish this task.

The size of the panel data set can be reduced based on three factors: (a) products, (b) shopping trips, and (c) households. Often, combinations of these dimensions are used and the choice of any particular approach will largely be governed by the objectives of the study. We describe below the approach taken in this study.

a) Products

One approach to reducing the dimensionality of the data set is to select a few product categories and then aggregate across the SKUs to the brand level within each of the chosen categories (e.g., Ainslie and Rossi 1998). However, given the objectives of our study, selecting a few product categories would be inappropriate, as we would not be able to analyze the complete basket of products that households buy. Hence, we choose to retain purchases of all

items, but aggregate them across SKUs into brands, then across brands to product categories, and from product categories to product groups. Hereinafter, we refer to the final level of aggregation as "products."

Invariably, the aggregation across products (or SKUs) involves judgment. Based on extensive discussions with the management of the retail chain, we identified 7 relevant final product groupings: meat, bakery goods, produce, dry goods, alcohol, household supplies (detergent, soap, insect repellent, etc.), and health and beauty-care. Each of the above seven product groupings contains different "product categories" aggregated within it. For example, the alcohol product grouping contains the purchases of beer, wine, and spirits product categories. The product category "beer" would contain the different brands (Chimay, Stella Artois, Jupiler, etc.). Likewise, the other product groupings (meat, bakery goods, etc.) are also composed of product categories and within them, the various brands. In table 1 we list the different product groups and their components.

Having defined the product groups, we now describe how we compute the price for each group on any given shopping trip. In computing product group prices, one can draw upon the extensive literature on "price indices" in the economics literature (see Deaton and Muellbauer 1980a). The basic idea here is to compute a 'group price' by using a weighted-average of the prices of the individual items in the group. The pertinent question is what items should be included and what should that weighting scheme be? On any given purchase occasion, should one consider only the items bought, or should one also consider items that were not purchased? Consistent with previous work on multi-category purchasing behavior, (Manchanda *et al* 1999; Krishnamurthi and Raj 1988) we construct and overall price for a category during a specific shopping occasion as the weighted average of product prices in effect that week where the weights are the long run share of each product bought by the household. We compute similar weighted averages for promotions. The exact formulation of the weighted prices and promotion variables are as follows:

Category Expenditure: $E_{gt}^{(h)} = \sum_{s=1}^{s} p_{st}.q_{st}^{(h)}$

Category Price: $P_{gt}^{(h)} = \sum_{s=1}^{s} p_{st}.w_{s}^{(h)}$

Category Promotion: $Pm_{gt}^h = \sum_{s=1}^{s} pm_{st}.w_s^{(h)}$

where:

p_{st} = price of SKU s during shopping trip t,

pm_{st} = 1 if SKU s was on promotion during trip t, 0 otherwise,

$q_{st}^{(h)}$ = quantity of SKU s bought by household h during shopping trip t,

S = number of product categories in group g,

T = number of shopping trips for household h, and

$w_s^{(h)} = \dfrac{\sum_{t=1}^{T} p_{st}.q_{st}^{(h)}}{\sum_{t=1}^{T}\sum_{z=1}^{S} p_{zt}.q_{st}^{(h)}}$ = weight associated with SKU s.

Hence, our category or product group price during a given trip by a given household reflects the price of that household's consideration set (all the products that the household might buy). The weight given to each item is the share of expenditure of that item over the whole span of the study for the given individual.

We do not explicitly define a "quantity" variable for each group of products. The reasons for this are twofold. First, it is not clear how a group quantity variable can be defined when aggregating across different types of products measured in different units to form the broad group. Second, as we shall see later it is not necessary to define a quantity variable because the two sets of dependent variables used in our analysis are the total expenditures and the expenditure shares of each group of goods, both of which are well defined and comparable across the groups. Having described how the data can be reduced via product aggregation, we next describe how data reduction can be accomplished by examining the type of shopping trips.

b) *Shopping Trips*

For a given household, not all shopping trips are the same. Kahn and Schmittlein (1989) classify household shopping trips into "regular" and "filler" trips. Drèze (1996) examines "regular" and "cherry-picking" trips. Hence, following those two studies, we use the data to classify household's shopping trips into three different categories: regular, filler, and cherry picking.

In figure 1(a), we show the relationship between the total amounts spent on a shopping trip (computed as an index) and the proportion of that expenditure spent on items that were on sale via temporary price reductions. Figure 1(b) depicts a similar relationship between the total number of items bought (as opposed to total expenditures in figure 1(a)) and the proportion of items bought on sale.[3] In each case, the vertical axis is an index that is normalized to 1 when the expenditure on items on sale is zero (or the proportion of number of items bought on sale is zero). Hence, the interpretation of the vertical axis in say figure 1(b) is that when 40% of the items are bought on promotion, then the total number of items bought is appropriately twice the number of items bought when the proportion of items bought on sale is zero.

Based on figures 1(a) and 1(b), we can classify household's shopping trips into the three categories: filler trips, regular trips, and cherry picking trips. *Filler* trips will have 10 or fewer items and less than 50% of these items are bought on promotion. *Cherry-picking* trips will also have 10 or fewer items, however at least 50% of these items will be bought on promotion. *Regular* trip represent baskets of 11 or more items, regardless of the proportion of products bought on sales.[4] Figure 2 shows the proportion of each type of trip in our database.

Of these three types of trips, regular trips are the most interesting. Indeed, filler trips are made with the express purpose of buying some missing products. In this case the budget for the shopping trip is most likely determined in the store and will equal the total purchase prices of these items. Any price reductions on the items that are sought will only result in lower expenditure without any budget reallocation. Conversely, cherry-picking trips are made to take advantage of specific price promotions. The budget for such trips is probably fixed before entering the store based on the advertised price and intended purchase quantities. We nevertheless retained for our analysis all three types of trips for the included households (see below).

Having described how the data can be reduced via product aggregation and classification of shopping trips, we next see if classifying households based on their shopping trips behavior can reduce the data.

c) *Households*

In figure 3, we classify households according to their shopping behavior. We see from figure 3 that a mere 1.5% of the households do only filler or cherry-picking shopping trips[5]. The

remainder (i.e. 98.5% of the households) has had at least one regular shopping trip. Further, 32.4% of households do all three types of shopping trips, while 62.2% do both regular and filler trips. Hence, an *a priori* classification of households as "large basket" shoppers or "small basket" shoppers a la Bell and Lattin (1998) would not be appropriate for this set of households.

d) *Final Dataset*

We retained all households in the sample except (1) those who made *only* filler and/or cherry-picking trips and (2) those households that made fewer than 2 trips (for reasons that will soon be apparent, we cannot compute the value of the inventory variable if only one shopping trip had taken place). Finally, for each household retained in the sample, we aggregated across the shopping trips made during a given week. This aggregation is necessary to account for the strong regularity in inter-shopping time observed by Kahn and Schmittlein (1989) and Jain and Vilcassim (1991). This tendency of shoppers to think in terms of weeks is also present in our data as is evidenced by figure 4 which shows a histogram of the number of elapsed days between shopping trips. There are indeed clear peaks on seven day multiples. Another justification for this is that the analysis of shopping trip expenditures from the retailer's standpoint is meaningful only at the weekly levels (i.e., the retailer's decision making time horizon).

After reducing the data based on the above criteria, our analysis focused on the purchases / expenditures of 7 different product groups by 23,635 households, making three-quarter million shopping trips over a one year period. The trips are aggregated into 528,207 week-trips (i.e., we have an average of 22.3 data points per household). These data are augmented with promotional and demographic data as described previously. We next describe the characteristics of the final data set and attempt to draw some insights about household expenditures and the effects of promotions.

Descriptive Analysis

Overall statistics for each product groups are given in Table 2. We see from this table that there are wide variations across categories. For instance, for both produce and dry goods, more than 50% of the shopping trips involved the purchase of promoted products (in these categories). In contrasts, only 10.3% of the shopping trips involved the purchase of promoted bakery products. Importantly, we see from Table 2 that for all product groups, the expenditures

during promoted trips are higher than on non-promoted trips (for example, the trip spend on meats during promoted weeks is €66.96 vs. €56.89 on non-promoted trips). This provides a clear indication that promotions can induce increases in expenditures. From Table 2, we also note that the percentage of expenditure allocated to a given product group increases during promotion trips. However, the effects vary across the different groups, with it being much smaller for bakery items than, say alcohol.

More importantly, we see that during trips that include purchases of promoted produce (or dry goods), households spend more on the non-promoted items in that same product group than is spent on the entire product group during non-promoted trips. Specifically, the total product group spend on produce during a non-promoted trip is €11.53, while during a promoted trip the expenditures on the non-promoted produce items is €15.89 (for dry goods the corresponding figures are €16.66 and €12.64, respectively). On the other hand, for alcohol the corresponding amounts are €9.47 and €14.06. This suggests that while overall expenditures increase with promotions in all cases, there are important differences among the different groups that the retailer must consider when deciding on the promotions.

To gain additional insights, we show in Table 3 the correlation between the spending levels on the different categories on a given shopping trip. All the correlation coefficients are positive, indicative of overall spending effects (i.e., an increase or decrease in spending is spread across categories) rather than substitution effects (i.e., an increase in one category is compensated by decrease in another category. The correlation analysis also shows that alcohol product sales are relatively independent from the sales of the other categories. In contrast, produce and dry goods expenditures are highly correlated with each other and with the other categories.

Going a step further, one can examine the correlation between the amounts spend on promotion in one category and the total expenditure in other categories. These correlation coefficients are shown in Table 4 where rows represent promoted sales, and columns the total category sales. As in the previous table, the absence of negative terms leads us to expect that the net effect of promotions will be to increase spending (both at the category level and through spill-over to other categories) rather than create substitution from one category to another.

Computing the temporal rather than the contemporaneous correlations shows negative first-order auto-correlations for both total spending (-0.03) and product group spending. At the

product level, we see higher auto-correlation for the non-perishable categories (-0.07 for HBC and HH Supplies, -0.05 for Dry Goods) than for perishable (-0.03 for Produce, -0.02 for Meat, -0.01 for Bakery). These results are what might be expected based on the ability to inventory products for future consumption. Interestingly, Alcohol, which is by and large a non-perishable product category, exhibits the auto-correlation of a perishable category (-0.01), suggesting some possible increases in consumption given that expenditures increase during promotions.

The above descriptive statistics indicate that promotions do affect spending. However the level of analysis is too crude to disentangle the effects in more detail. For example, some inferences were based on pairwise correlations and not on causal factors. Thus, in the next section, we model the household's expenditure and allocation decisions in order to draw more firm conclusions about the drivers of these decisions.

Household Expenditure the AIDS Model and Econometric Estimation

Based on the previous descriptive analysis, it is clear that when analyzing household shopping behavior we must not only allow for substitution between the different groups of products, but also allow the household's expenditure to be determined endogenously by the pricing and promotional activities undertaken by the retailer. Allowing for endogenously determined store expenditures indirectly allows for an outside good, which is required if the demand system is to be fully specified. Hence, we model this process in two stages.

a) Analysis Household Expenditures

In the first part of our analysis, we attempt to identify the impact that prices, promotions, and household inventory levels have on household shopping trip expenditures $(X_t^{(h)})$. Accordingly, we specify the following regression model:

$$\ln X_t^{(h)} = a_0 + \sum_{g=1}^{7} b_g \ln P_{gt}^{(h)} + \sum_{g=1}^{7} c_g Pm_{gt} + \sum_{g=1}^{7} d_g \ln I_{gt}^{(h)} + \sum_{l=1}^{L} f_l Z_l^h + r_t^{(h)} \qquad (1)$$

where $r_t^{(h)} = H^{(h)} + u_t$, $u_t \sim i.i.d.$ $N(o, \sigma_u^2)$ and for each product grouping g = 1,2,..7,

$P_{gt}^{(h)}, Pm_{gt}^{(h)}, I_{gt}^{(h)}$ and $Z_l^{(h)}$ are the price, promotion, inventory and l^{th} demographic variables for household h on purchase occasion t; a_0, b_g, c_g, d_g, and f_l are the regression coefficients, and 'ln'

denotes natural logarithm. We use a fixed-effects ($H^{(h)}$) specification to accommodate the unobserved heterogeneity across households.

An important variable in the above equation is the level of the household's inventory of the product group at the time of purchase. Since the household's inventory is unobserved, it has to be computed or inferred based on observed purchase quantities, the inter-purchase times, and assumptions about consumption rates. We operationalize the inventory measure as per Gupta (1988) with the exception that, instead of working with inventoried quantities, we work with the dollar value of the inventory. The rationale for this is simple. When working with high-level product aggregation (e.g., produce) it is not meaningful to track inventory in terms of quantities since the quantity numbers refer to different items that may have both different storage units (e.g., pounds, counts, ounces) and different usage rates (e.g., sixteen ounces of milk a day versus two ounces of cheese). An alternative approach would have been to compute a usage figure for every SKU (a liter of milk has four servings, Tide 64 ounces can be used for 18 washes). This would work well for all food and grocery items, but may not be appropriate for other products e.g., kitchen scrubbing pads. Hence, by keeping track of the inventory in terms of its monetary value we can ensure comparability across SKUs that comprise a given product group. Further, we express inventory using the same unit (Euros) as our other main variables of the model, price and expenditure. This helps in the interpretation of the results.[6]

In the spirit of Gupta (1988), we operationalize the inventory variable ($I_{gt}^{(h)}$) for household h and category g as:

$$I_{g,t}^{(h)} = I_{o,g}^{(h)} + \sum_{\tau=1}^{(t-1)} E_{g,t}^{(h)} - R_g^{(h)} \sum_{\tau=1}^{t-1} T_{g,t}^{(h)} \tag{2}$$

where:
$$I_{o,g}^{(h)} = \sum_{t=1}^{n_g} \frac{E_{g,t}^{(h)}}{n_g} \quad \text{(Initial inventory)} \tag{3a}$$

$$R_g^{(h)} = \frac{\sum_{t=1}^{n_g} E_{g,t}^{(h)}}{\sum_{t=1}^{n_g} T_{g,t}^{(h)}} \quad \text{(Rate of use)} \tag{3b}$$

and $T_{g,\tau}$ is the inter-purchase time between purchase occasion (τ-1) and (τ) and $E_{g,t}^{(h)}$ is the category expenditure for group g of household h during trip t as previously defined. One should

note that this model allows for different consumption rates for each category for each household[7].

Based on equation (1), we can form various hypotheses about household expenditures. For example, if:

i) household expenditures are exogenous and fixed:

$$\Rightarrow b_g = c_g = d_g = 0 \qquad\qquad \forall g = 1,2,...,7$$

ii) household expenditures are determined only by inventory levels:

$$\Rightarrow b_g = c_g = 0 \qquad\qquad \forall g = 1,2,...,7$$

$$d_g \neq 0 \text{ for at least one g.}$$

Likewise, various tests on the significance of the different regression coefficients will provide insights into the variables that drive household expenditures.

We note that in modeling the household expenditures, we have included only the drivers from the given store. If households comparison-shop *across* different stores (retailers) then the expenditure in the given store may also change because of price changes at other stores. We are not able to separate those two effects because of a lack of data on prices at other retailers. To the extent that price movements across stores are not highly correlated, we can still measure properly the impact of price changes within the given store on total household expenditures. But, as we have noted before we cannot verify that assumption and hence, we can provide only a partial explanation of the variation in total household expenditures.

b) Expenditure Allocation Decisions

We next model the allocation of the total expenditure $(X_t^{(h)})$ determined in (a) above across the seven different product groups: meats, bakery products, produce, dry goods, alcohol, household supplies, and health and beauty care. For that purpose, we use an extended version of the AIDS model developed by Deaton and Muellbauer (1980b). The AIDS model has been widely used for estimating demand systems in both economics and other areas. Since its original development by Deaton and Muellbauer (1980b), literally several hundreds of demand studies using the AIDS model have been reported in the research literature. We refer the reader to the extensive literature on the use of the AIDS model for further information (e.g., Molina 1997;

Labeaga and Lopez 1997; and references cited in both) and for marketing applications to the studies by Cotterill, Putsis, and Dhar (2000), Putsis and Cotterill (1999), and Cotterill and Putsis (1999).

The basic AIDS model that describes the allocation of expenditure across seven different groups of product by household h in week t can be expressed as:

$$w_{gt}^{(h)} = v_{gt}^{(h)} + \sum_{j=1}^{7} \xi_{gj} \ln P_{jt}^{(h)} + \theta_g \ln(\frac{X_t^{(h)}}{\overline{P}_t^{(h)}}) + \zeta_{gt}^{(h)} \qquad (4)$$

where $\zeta_{gt} \sim iidN(o, \tau_g^2)$, and $\sum_{g=1}^{7} w_{gt}^{(h)} = 1$.

In the above equation, for any week t, $w_{gt}^{(h)}$ is the expenditure share allocated to group g by household h, $P_{jt}^{(h)}$ are the prices encountered by household h for each of the seven groups ($j=1..7$), $X_t^{(h)}$ is the expenditure of household h and $\overline{P}_t^{(h)}$ is a price index. From the above equation we note that the ξ_{gj} coefficients capture the own- and cross-price effects, and θ_g measures the marginal effect of a change in total real expenditure on the share of allocation to group g, and $v_{gt}^{(h)}$ is a time varying household specific factor that captures the share of budget allocated to group g, above and beyond the price and real expenditure effects.

In the spirit of Blundell, Pashardes, and Weber (1993), we embellish the basic AIDS model to account for household inventory, promotional, and demographic effects. Specifically, we let:

$$v_{gt}^{(h)} = \delta_g + \pi_g^{(h)} + \sum_{j=1}^{7} \gamma_{gj} \ln I_{jt}^{(h)} + \sum_{j=1}^{7} \alpha_{gj} Pm_{jt}^{(h)} + \sum_{l=1}^{L} \eta_1 Z_l^h \qquad (5)$$

In the above equation, δ_g captures a household-invariant group specific effect. $\pi_g^{(h)}$ is a household effect specific to group g and the other coefficients capture the effect of inventory ($I_{jt}^{(h)}$), promotion ($Pm_{jt}^{(h)}$), and demographics (Z_l^h). Merging equations 4 and 5 we get:

$$w_{gt}^{(h)} = \delta_g + \sum_{j=1}^{7} \xi_{gj} \ln P_{jt}^{(h)} + \theta_g \ln(\frac{X_t^{(h)}}{\overline{P}_t^{(h)}}) + \sum_{j=1}^{7} \gamma_{gj} \ln I_{jt}^{(h)} + \sum_{j=1}^{7} \alpha_{gj} Pm_{jt}^{(h)} + \sum_{l=1}^{L} \eta_1 Z_l^h + \pi_g^{(h)} + \zeta_{gt}^{(h)}. \qquad (6)$$

As before, we use a fixed-effects $\left(\pi_g^{(h)}\right)$ specification to accommodate the unobserved heterogeneity across households. In addition, when the price index $\overline{P}_t^{(h)}$ is approximated by the

Stone price index $\left[i.e., \ln \overline{P}_t^{(h)} = \sum\limits_{g=1}^{7} w_{gt}^{(h)} \ln P_{gt} \right]$, the budget share equation (6) is linear in the

parameters.

The parameters of the AIDS model cannot be easily given a direct interpretation. However, when expressed in the form of elasticities, they do provide insights into the households' budget allocation decisions. The expenditure elasticity $\left(l_g \right)$ of product groups g evaluated at the given budget share w_g [8] is given by:

$$l_g = 1 + \frac{\theta_g}{w_g}. \tag{7}$$

We note that $l_g \geq 1$ if $\theta_g \geq 0$. Hence, if $\theta_g < 0$, the product group g can be treated as a "necessity," while if $\theta_g > 0$ the product group can be treated as a "discretionary" item. If $\theta_g = 0$ $\forall g$, then preferences are homothetic.

The uncompensated price elasticities of demand (Molina 1997) for two groups g and j evaluated at the average budget shares are given by:

$$e_{gj} = -\delta_{gj} + \frac{\xi_{gj}}{w_g} - \frac{\theta_g w_j}{w_g} \tag{8}$$

where δ_{gj} is the Kronecker delta (i.e., $\delta_{gj} = 1$ if $g=j$ and 0 otherwise).

The own-price $\left(e_{gg} \right)$ and cross-price $\left(e_{gj} \right)$ elasticities will provide insights into how demand is influenced by price changes. Of particular importance is that of determining the relative magnitudes of the own-price elasticities and cross-price elasticities. If the latter are high, then although price promotions may induce a household to buy more of the promoted goods, it also implies that there is substitution away from other product goods. Clearly, if such effects exist, it has implication for the retailer's pricing and promotion decisions.

d) Econometric Estimation

Equation (1) for the expenditure equation can be expressed as:

$$\ln X_{it}^{(h)} = a_0 + V_t^{1(h)} \beta + Z^{1(h)} f + r_t^{(h)} \tag{9}$$

where $r_t^{(h)} = w^{(h)} + u_t$, $u_t \sim i.i.d.$ $N(0, \tau_u^2)$ and

16

$$V_t^{1(h)} = \left[\ln P_{1t}^{(h)}, \ln P_{2t}^{(h)}, \ldots, \ln P_{gt}^{(h)}, Pm_{1t}^{(h)}, \ldots Pm_{gt}^{(h)}, \ln I_{1t}^{(h)}, \ldots, I_{gt}^{(h)} \right]$$

$$\beta^1 = \left[b_1, b_2, \ldots, b_g, c_1, c_2, \ldots, c_g, d_1, \ldots d_g \right]$$

The parameters of equation (9) can be estimated using the classic "within-group" estimation method for panel data (see Hausman and Taylor 1981, Hsiao 1992). That is, we compute the variables for each household h as deviation from the mean (i.e., $\widetilde{Y}_t^{(h)} = Y_t^{(h)} - \overline{Y}^{(h)}$ for any variable) and estimate using ordinary least squares (OLS) the following regression:

$$\ln \widetilde{X}_{it}^h = \widetilde{V}_t^{1(h)} \beta + \widetilde{u}_t^{(h)}. \tag{10}$$

OLS estimates of β are unbiased and consistent because the error term $\widetilde{u}_t^{(h)}$ is uncorrelated with the columns of $\widetilde{V}_t^{1(h)}$. Hence, estimating the coefficients associated with the price, promotion, and inventory variables are quite straightforward. These are called the within-group estimates ($\hat{\beta}_w$).

To compute the coefficient (f) associated with the demographic variables, we do the following:

i)   We obtain an unbiased and consistent estimate of β as above,

ii) We use that estimate of β and compute a 'new' dependent variable

$$\overline{\ln X}_t^{(h)} - \overline{V}_t^{1(h)} \hat{\beta}_w, \text{ and}$$

iii) We estimate using OLS the following regression

$$\overline{\ln X}_t^{(h)} - \overline{V}_t^{1(h)} \hat{\beta}_w = a_0 + Z^{1(h)} f + w^{(h)} + \overline{u}_t^{(h)} \tag{11}$$

by treating ($w^h + \overline{u}_t^{(h)}$) as the random error term. This is the "between group" estimator of the parameter vector f. Note that in the between group regression, we only have one observation per household.

In principle, one could then obtain estimates of the parameter vector f associated with the demographic variables. There is however, one issue that must be considered. For the OLS estimates of f to be consistent, the random error $\left(\overline{r}_t^h = w^{(h)} + \overline{u}_t^h \right)$ must be uncorrelated with the demographic variables. This may not be the case if the unobserved heterogeneity component ($w^{(h)}$) is correlated with the observed (included) demographic variables ($Z^{(h)}$). If indeed they

are correlated, then OLS estimations should not be used to obtain an estimate of f. Hence, prior to estimation of the demographic effects, we perform a specification test as suggested by Hausman and Taylor (1981).

The specification test is as follows. As noted previously, the within group estimator of β ($\hat{\beta}_w$) is unbiased and consistent. We can also estimate β from the between group regression by running:

$$\overline{\ln X}_t^{(h)} = a_0 + \overline{V}_t^{1(h)}\beta + Z^{1(h)}f + w^{(h)} + \overline{u}_t^{(h)}. \tag{12}$$

This between group estimate of β ($\hat{\beta}_B$) will be consistent only if the random error $\overline{r}_t^{(h)}$ in the between group regression is uncorrelated with the right hand side variables. We then compute the following test statistics, which is Chi-squared distributed with degrees of freedom equal to the dimensions of β.

$$H = \left(\hat{\beta}_w - \hat{\beta}_B\right)'\left[Cov(\hat{\beta}_w) + Cov(\hat{\beta}_B)\right]^{-1}\left[\hat{\beta}_w - \hat{\beta}_B\right]$$

If we fail to reject the null hypothesis, then the OLS estimates of f would be consistent. If not, then the demographic effects are not easily estimated. Alternative estimation methods will have to be used (see Hausman and Taylor 1981).

The parameters of the AIDS models (equation 6) can also be estimated using the within-group and between-groups estimation methods described previously. The use of the Stone price index renders the models linear in the parameters. Hence, the estimating equations for the AIDS model pertaining to the budget allocation can be expressed in a fashion similar to equation (10). The procedure used to estimate the AIDS model is thus similar to the procedure used to estimate the expenditure model. We also test formally for the empirical validity of both the symmetry and homogeneity restrictions using tests of nested hypotheses. The adding-up restriction is always imposed as the budget shares sum to one.

We report next the results of the estimation and characterize the households' shopping behavior. We also discuss the implications of those results for the retailer in terms of possible pricing and promotional strategies. We also attempt to compare and contrast our results against those in the extant literature.

Empirical Analysis and Implications

a) Model Specification and Diagnostic Tests:

　　We first report the results of various diagnostics tests of our model specification. Having established the statistical validity of our estimation, we then report the results of the estimation of the model parameters and discuss their substantive implications.

　　When estimating the AIDS model, one must be careful to respect the adding-up constraint of the shares of expenditure (i.e., $\sum_{i=1}^{7} w_1 = 1$). Imposing the following constraints on the parameters does this: $\sum_{g=1}^{7} \delta_g = 1; \sum_{g=1}^{7} \xi_{gi} = 0, \forall i; \sum_{g=1}^{7} \theta_g = 0; \sum_{g=1}^{7} \gamma_{gi} = 0, \forall i; \sum_{g=1}^{7} \alpha_{gi} = 0, \forall i$. One can impose the constraints by either, specifying them explicitly in the estimation, or estimating only six of the seven equations and deriving the parameters for the seventh equation from the constraints. We chose the later approach, but note that the results of the estimation are invariant to the approach used.

i) <u>Hausman-Taylor Test</u>

　　When working with household level data, great care must be taken when specifying the demographic variables to be used in the model. Indeed, the purpose of these variables is to account for the heterogeneity in consumer response across households. As has been demonstrated by Hoch et al. (1995), constructs such as household "Size" or "Income" have a profound influence on estimated values for price elasticity of demand and product consumption. We had at our disposal 54 demographic variables. Using a stepwise variable selection procedure, we retained nine: household size, household income, a measure of the food consumption potential for the neighborhood in which the household lives, and 6 social class descriptors. In addition, we used 3 store dummies to account for differences in expenditure across stores. We ran the Hausman specification test on this model by estimating the within group regression (10), and the between group full model (12). The parameter estimates as well as some key regression statistics are shown in Table 5.

　　Looking at the reversals in the signs of the price coefficients of the within and the between group regression, it comes as no surprise that the results of the Hausman and Taylor test is that the between group estimator is inconsistent (H=38,890.175 > $\chi^2_{0.0001,21}$ = 53.96).

Therefore, we cannot obtain straightforward estimates of the effects of the demographic variables. As mentioned in the previous section, this indicates that our operationalization of the geo-demographic variables in not complete and the OLS parameter estimates obtained by estimating the full model (1) are inconsistent. This is a set back in that it does not allow us to make any inference regarding the impact of the demographic variables. However, we can still answer our research questions regarding the determinants and allocation of the within-store expenditures since the within group estimates (10) are consistent[9].

ii) Household Heterogeneity in Response Parameters

We used a fixed effect model specification to account for the unobserved heterogeneity across households included in our analysis. In addition, the results of the Hausman test imply that we cannot estimate the effects of demographic variables on the expenditure and budget allocation decisions. To ensure that our fixed effects specification is adequate (i.e. that the assumption of homogeneity across households in the response parameters is reasonable), we did a median-split of our data on the basis of household income. We then estimated our model separately on the low-income households and the high-income households, in addition to estimating a common model across all households. Although the estimated parameters for the high- and low-income households were somewhat different, they were substantively identical in that the signs and order of magnitudes of the parameters were similar across the three systems. For instance, the price coefficients for the Meat category in the expenditure equation are –0.879, -0.895, and –0.869 for the overall, low-income households, and high-income households respectively. Similar results were obtained for the other parameters. This result suggests that although there are differences across households, we can, as a first approximation, treat them as being identical in their response function without invalidating the substantive implications of our study. We should note however that use of a household specific usage rate for the inventory computations means that the model is not a pure fixed-effect model. However, in light of the median-split test, we believe that the specification is appropriate.

iii) Test for Heteroskedasicity

Another potential problem created by the study of a wide array of households is that of heteroskedasticity of the error terms. Heteroskedasticity does not affect the consistency of the

parameter estimates; however, it might affect their efficiency. To check for potential heteroskedasticity problems, we estimated the system of equations using the Generalized Method of Moment (GMM) estimation, in addition to using OLS regression. The results of the estimation showed that accounting for heteroskedastic errors did not change the results in any substantive manner. One plausible reason for this result is that the lack of efficiency of the parameter estimates is not likely to be a problem when dealing with very large data sets such as ours. It has been our experience that with such large data sets, the traditional 95% confidence interval in not restrictive enough. Hence we use a more stringent value of 0.0001 cut-off level for the p-values.

Having established the reasonableness of our model specification, we report next the results of the parameter estimation for the household within-store expenditure and allocation models. We also discuss the substantive implications of these estimates for the retailer's pricing and promotion decisions.

iv) Out-of-sample validation

To test for possible over-fitting of the model, we ran an out-of-sample validation. We re-estimated both the expenditure and the AIDS model using only the first 90% of the data set as calibration data and then used the parameter estimates to forecast expenditure and budget share for the remaining 10% of the data. Comparing the mean squared-errors both in–sample and out-of-sample, we find no cause for alarm. For instance, the MSE for the expenditure model was 0.32 in-sample and 0.36 out-of-sample. For the AIDS model, the difference in MSE ranged from 0.0317 (IS) vs. 0.0351 (OOS) for Produce to 0.0115 (IS) vs. 0.0211 (OOS) for Bakery.

b) *The Within-Store Household Expenditure Model*

As shown in Table 5 (the within-group estimates in column 2), the expenditure model was estimated using 528,206 observations and yielded a R^2 of 0.14.[10] In terms of our hypotheses, we see that all the price coefficients are negative and significant. This indicates that households finalize their budget decisions while at the store and that they are price sensitive. The positive promotion parameters indicate that households will increase their expenditures in response to promotions and that this increase is greater than what would be warranted by the

price decrease alone. Hence, *within-store expenditures are not fixed* and they do have a discretionary component.

We note that that as the expenditure model is expressed in *log-log* form, the coefficients of the price, promotion, and inventory variables represent "elasticities" (i.e., measure the percentage change in total expenditures change as a result of a 1% change in each of those variables). Thus, one can make meaningful direct comparisons across the different groups of goods and variables. Examining the price effects, we see that a 1% reduction in the prices of alcohol and dry goods have the largest impact on total expenditures (increase by 1.350% and 1.139%, respectively). Bakery products and produce items have the lowest group level elasticities (0.264% and 0.341%). Hence, promoting alcoholic beverages is the most attractive from the retailer's standpoint, if the goal is to increase total within store expenditures.

Examining the promotion effects, we see from Table 5 that as with the price effects, alcohol has the largest effect on expenditures (0.327%), followed household supplies (0.208%) and health and beauty care products (0.202%). As before, promoting produce items and bakery goods have lowest effects on store expenditures (0.097% and 0.131%, respectively). In general, the promotional effects are much smaller than the price effects.

The estimated inventory effects are all negative, except for bakery items. This implies that higher levels of inventory lead to reduced levels of in-store expenditures, with the effect being largest for alcohol (-0.029%), and least for meats and produce (-0.0031% and –0.008%, respectively). These results seem reasonable given that produce is a perishable item, while alcohol can easily be inventoried. The positive coefficient for bakery item inventory does seem odd however.

Hence, the retailer's strategy with respect to price promotions does involve a trade-off between increasing current in-store expenditures and reduced future expenditures because of inventory effects. This analysis has to be done category by category and the analysis done here can be used to guide these decisions.

c) Allocation of Total Expenditures AIDS Model Estimates

The previous section has shown us how shopping expenditures increase or decrease in response to price changes and promotional activity. The next step in the analysis is to study how the share of the budget allocated to each category varies in response to increases and decreases

in the total expenditures. To study this, we applied the AIDS model previously described to our seven categories. As with the expenditure model, we had to use the within-group estimation in order to obtain consistent estimates of the price, promotion and inventory variables. The tests of the symmetry and homogeneity restrictions resulted in rejecting both sets of hypotheses (F-statistic values of 408.9 and 136.6 respectively, both significant at the 0.0001 level). The unrestricted coefficients are reported in Table 6. To ease this interpretation of these coefficients, we computed the expenditure and price elasticities (using equations 7 and 8). They are reported in Tables 6 and 7, respectively.

Prima facia, the estimated parameter values that are reported in Table 6 seem reasonable. All the own-price coefficients are negative and significant, while the own-promotion variables are positive and significant. The own-inventory coefficients are negative, indicating that inventory levels do affect the budget allocation decisions. The cross-price coefficients are all positive (except for one) and significant, implying that although expenditures are flexible, there is also substitution at the product group level and less amounts are allocated when prices increase. Similarly, most cross-promotion coefficients are significant and negative, implying the existence of substitution effects.

The expenditure elasticities (Table 7) are less than one for four categories (Meat, Bakery, Produce, and Health and Beauty Care (HBC)), meaning that they are seen as "necessities." Three categories (Dry Goods, Alcohol, and HH Supplies) can be classified as "discretionary[11]" as their expenditure elasticities are greater than one. Figure 6 plots the share of each category as a function of expenditure. One can see that as expenditure increases, the four necessity categories see their share decline as the total expenditure increases (although HBC behaves in an almost homothetic fashion). The three discretionary categories see their share increase as budgets increase. The category that gains the most from an overall increase in expenditures is Dry Goods.

The own-price elasticities that are reported in Table 8 are all greater than one, implying that the demands for all product groups are elastic. In addition, the own-price elasticities are highest for HH Supplies (-6.13), Alcohol (-5.06), and HBC (-5.11). These estimates are consistent with the estimated expenditure elasticities, wherein HH Supplies and Alcohol were deemed discretionary items. The lowest own-price elasticity is that of Produce (-1.49), a result that seems consistent with the nature of the product (i.e., perishable).

Examining the cross-price elasticities reported in Table 8, we note that when the price of Alcohol is reduced, the impact on the demand for each of the other groups is relatively high (0.50, 0.31, 0.71, 0.96, 0.68, 0.39 for Meat, Bakery, Produce, Dry Goods, HH Supplies, and HBC respectively). However, reductions in prices of the other products have a relatively smaller impact as the demand for Alcohol (0.15, 0.03, 0.09, 0.20, 0.10, 0.09, respectively). In general, the elasticities exhibit an asymmetric structure that seems analogous to that observed at the brand level for differentiated products (see Blattberg and Wisniewski 1989). In this case, the more discretionary product groups (Alcohol, HH Supplies) seem to have the greater drawing power.

The results from the share allocation analysis reinforce the analysis done at the expenditure level. We see that the perishable categories are necessity goods. They are bought every week, but there is little room for category expansion through discretionary purchases. This makes them prime candidates as *loss leader* categories. Indeed, retailers need to be competitive on necessity categories as they are likely drivers of store traffic. At the same time, they can limit the loss they would incur from loss leader pricing since their demands are relatively capped by the perishability of the goods (i.e., sharp price decreases will not lead to stockpiling or purchase acceleration). Further, of the three perishable categories, meat is probably the best candidate as a loss leader since it has the largest elasticity.

Conversely, the discretionary categories should be used to generate profits. Alcohol, HH Supplies, and HBC are the most elastic product categories. Linked with their non-perishable nature, this will lead to stockpiling, inter-temporal switching and perhaps increased consumption. The implications of these results are that the retailer must be careful to keep margins high enough during promotions to make the promotions profitable on their own, regardless of any spillover or traffic generating effect.

Finally, we note that Dry Goods is the category that is the most discretionary purchase behavior. It has the most potential to increase in-store expenditures. Hence the retailer should pay special attention to the in-store promotions of this category. One caveat with respect to the above substantive implications is that we have discussed only the impact on expenditures and not retailer profits because of the unavailability of data on profit margins.

d) Clout vs. Vulnerability Analysis

Following Kamakura and Russell (1989), we performed a "Clout vs. Vulnerability" analysis of the cross-price elasticities (see Figure 6). When interpreting the results from this analysis, it is important to take the perspective of the retailer rather than the manufacturer. Whereas a manufacturer wants a products to have a high clout (and low vulnerability), a retailer must be careful when dealing with categories with high clout as their promotions will affect the sales of other categories, and thus affecting overall store-level profits.

Our analysis reveals that the Alcohol category has the most clout. This implies that the store managers must be careful when they plan their Alcohol promotions and measure the profitability of these promotions to take the potential negative effects on the other categories into consideration. One must note here that if one looked back at tables 2 and 3, one would see that a superficial analysis would have concluded that alcohol is actually the least influential category in terms of spending. This underscores the need to go beyond the simple descriptive statistics. We also see that Bakery and Produce can be treated independently from the other categories in that they have no clout and low vulnerability. Dry Goods, in contrast, is the most vulnerable category implying that its budget share drops as other products are price promoted.

Summary and Directions for Future Research

We set ourselves to investigate whether households enter grocery stores with preset expenditures in mind or whether they wait to be in the store, facing the products and their prices, to decide how much they will spend on that shopping trip. As noted, this is an important question as the answer to it greatly impacts how retailers should view promotions.

In the process of answering this question we raised a series of issues concerning the proper handling of large household panel data sets, wherein the entire basket of products purchased, not just a few select product categories, must be analyzed to provide meaningful insights. In the first step of our analysis, we used simple descriptive statistics generated from the rich data available to us to show that promotions do indeed induce households to spend more, but that the effects vary across different product groups. For example, promotions on alcoholic products and produce have greater expenditure effects than bakery goods. We also found evidence that this increase in spending is not limited to the promoted category but pertains to the overall basket with spillover effects to other categories.

25

To gain further insights, we used an extended version of the AIDS model and we showed that the expenditure decisions were both endogenous to the store and also affected by out-of-store constraints such as inventory levels. We also analyzed the allocation of the total expenditure to the different product groups constituting the shopping basket. Our analysis showed how products can be classified as *necessary*, *discretionary*, or *homothetic* goods, and how this classification affects the way retailers should think about these products when planning their pricing and promotional decisions. We showed why perishable products such as bakery items, produce, or meat are well suited for loss-leader promotions, while alcohol, household supplies, dry goods are well suited for in-store high margin promotions.

We also addressed a number of econometric estimation issues that must be considered when analyzing panel data. Specifically, we showed how to overcome any potential issues of endogeneity in prices and promotions by using the within group estimation method for panel data. We also highlighted the difficulty associated with estimating demographic effects (age, income, race, etc.) in the presence of unobserved heterogeneity.

One important concern about our results relates to the measurement of the inventory variable. As we only possess data for one store, we do not know about purchases made by our households to competing stores. Consequently, the inventory measure is inherently noisy which leads to bias and consistency problems (Neter, Wasserman, and Kutner 1983) with the inventory parameter estimates. Unfortunately, it is impossible for us to access competitor information. We do not either have a good instrument variable that can be used to improve on the consistency of the parameters.

In our analysis, we could not shed much light on how household demographic variables might influence households' expenditure and allocation decisions. From the standpoint of the retailer, this is an important consideration because demographic variables are actionable. Other papers (Hoch *et al* 1995) have shown that demographics are important drivers of price elasticity and other important parameters. We were not able to incorporate these measures in our analysis because we only possess demographic variables at the city block level which proved to be to high a level of aggregation to be useful. This does not mean that demographic variables are not important, rather than they are so important that they need to be known at the household level to be useful predictors.

Our analysis focused on the immediate and near term effects of price promotions. One issue that we have not explicitly addressed is that of the long-term effects of promotions. In a recent study, Nijs, Dekimpe, Steenkamp, and Hanssens (2000) analyze the long term effects of promotions on category sales for over 500 products and find that in over 95% of the cases there are no long-run effects. It thus seems that if there are benefits to promotions at the retail level, there are likely to be only in the short run. Here, as we have shown, the retailer must trade-off the immediate expenditure boost against any stockpiling effects and this can be achieved by an appropriate selection of product categories.

There are a number of issues that merit future research. First, more analysis is needed of households' entire basket of purchases to generalize the results we obtained from a single data set. We do not possess information about prices at competing stores. Although Walters (1991) finds little evidence a cross-store promotional effects, such effects might exits and impact household expenditures. Additionally, the data used in our analysis pertained only to the purchases made by the households in the stores of the given retailer. A more complete analysis should consider the purchases made by households across all retailers. That would require analyzing the household's store choice decision, as well as the decision of how much to spend, conditional on the store choice. An appropriate methodology for such an analysis would be to use a two-stage Tobit model, where in the first stage one analyzes the factors influencing the decision to visit a particular store, followed by an analysis of the expenditure decision (Maddala 1983).

Another extension to the analysis done here would be to model the type of shopping trip the household will undertake (regular, filler, or cherry-picking) as a function of covariates such as store prices, promotions, and household inventory. Such an analysis will provide additional insights to the retailer in terms in relation to how products should be priced and promoted. One could then analyze the impact (if any) of household stockpiling of products during promotions on their store visiting behavior.

Our analysis focused on the household's total expenditure decision and the allocation of that expenditure across broad groups of products. If one is interested in analyzing further the allocation of expenditures within a group of products (for example, within the alcohol group, how the expenditure is allocated across say, beer, wine, and spirits) then one invoke the concept of separability of preferences, two-stage budgeting and utility trees to further analyze household

choice behavior. (Deaton and Muellbauer 1980a, chapter 5). One can keep proceeding in a similar manner (for example, within beer, across the different brands) and thereby analyze the complete purchase patterns of a household down to the brand level.   There would be some challenging methodological and estimation issues here that, when addressed, will contribute significantly to the extant literature.

To summarize, our study is one of the first attempts to analyze the entire basket of household purchases when shopping for consumer non-durable products. Our results indicate that households are flexible with respect to their grocery expenditures and pricing and promotional decisions do indeed induce them to spend more. However, household inventory levels mitigate this effect and there is also a reallocation of expenditures across product groupings, which has implications for the retailer's pricing and promotional decisions. We believe that we have made an initial contribution to the study of household market basket purchase decisions. In addition, we provide insights to retailers to help them manage the relationship between product categories. We show how superficial correlation analyses fail to highlight the impact that a category such as Alcohol has on the expenditures made in other categories. Such analysis would lead to an overestimation of the profitability of alcohol promotions. We also show how perishable categories such as Produce or Bakery are ideal candidates for loss leader type promotions.

Future research that address the other issues we have identified will add to the knowledge base of the effects of the pricing and promotional decisions of the retailer on household store shopping behavior.

References

Ainslie, Andrew, and Peter E. Rossi (1998), "Similarities in Choice Behavior Across Product Categories," *Marketing Science*, 17 (2), 91-106.

Bell, David R., Teck-Hua Ho, and Christopher S. Tang (1998), "Determining Where to Shop: Fixed and Variable Costs of Shopping," *Journal of Marketing Research*, XXXV (August), 352-369.

Bell, David R. and James M. Lattin (1998), "Shopping Behavior and Consumer Preference for store Price format: Why 'Large Basket' Shoppers Prefer EDLP," *Marketing Science*, 17 (1), 66-88.

Blattberg, Robert C. and Scott A. Neslin (1990), Sales Promotion: Concepts, Methods and Strategies, Englewood Cliffs, New Jersey: Prentice Hall.

Blattberg, Robert C. and Ken Wisniewski (1989), "Price Induced Patterns of Competition," *Marketing Science*, Fall, 291-309.

Block, Lauren G. and Vicki G. Morwitz (1999), "Shopping Lists as an External Memory Aid for Grocery Shopping: Influences on List Writing and List Fulfillment," *Journal of Consumer Psychology*, 8(4), 343-375.

Blundell, Richard, Panos Pashardes, and Guglielmo Weber (1993), "What Do We Learn About Consumer Demand Patterns from Micro Data?," *The American Economic Review*, 83 (3), 570-597

Clements, K. W. and L. W. Johnson (1983), "The demand for Beer, Wine and Spirits: A System-Wide Analysis," *Journal of Business*, 3, 273-304.

Chintagunta, Pradeep K., and Sudeep Haldar (1998), "Investigating Purchase Timing Behavior in Two Related Categories," *Journal of Marketing Research*, XXXV (February), 43-53.

Cotterill, R., W. Putsis and R. Dhar (2000), "Assessing the Competitive Interaction Between Private Labels and National Brands," *Journal of Business*, forthcoming, January 2000.

Cotterill, R., W. Putsis (1999), "Market Share and Price Setting Behavior for Private Labels and National Brands," *Review of Industrial Organization*, forthcoming.

Deaton, Angus and John Muellbauer (1980a), Economics and Consumer Behavior, Cambridge: Cambridge University Press.

Deaton, Angus and John Muellbauer (1980b), "An Almost Ideal Demand System," *American Economic Review*, 70 (June), 312-326.

Drèze, Xavier (1996), "Rehabilitating Cherry Picking," Working paper, Marshall School of Business, University of Southern California, Los Angeles, CA.

Drèze, Xavier and Stephen J. Hoch (1998) "Exploiting the Installed Base Using Cross-Merchandising and Category Destination Programs," *International Journal of Research in Marketing*, Winter.

Fader, Peter and Leonard Lodish (1990), "A Cross-Category Analysis of Category Structure and Promotional Activity for Grocery Products," *Journal of Marketing*, 54 (October), 52-65.

Guadagni, Peter M., and John D. C. Little (1983), "A Logit Model of brand Choice Calibrated on Scanner Data," *Marketing Science*, 2 (Summer), 209-238.

Gupta, Sunil (1988), "Impact of Sales Promotions as When, What, and How Much to Buy," *Journal of Marketing Research*, 25, 342-355.

Hausman, Jerry A. and William E. Taylor (1981), "Panel Data and Unobservable Individual Effects," *Econometrica*, 49 (6), 1377-1398.

Hoch, Stephen J., Byung-Do Kim, Alan L. Montgomery, and Peter E. Rossi (1995), "Determinants of Store-Level Price Elasticity," *Journal of Marketing Research,* XXXII (February).

Hsiao, C. (1992), "Panel Data Analysis," Econometric Society Monographs, Cambridge University Press.

Inman, Jeffrey, and Russell Winer (1999), "Where the Rubber Meets the Road: A Model of In-store Consumer Decision Making," *Marketing Science Institute*, working paper # 98-132.

Kahn, Barbara E. and David C. Schmittlein (1989), "Shopping Trip Behavior: An Empirical Investigation," *Marketing Letters*, 1, December, 55-70.

Kamakura, W. A. and Gary J. Russell (1989), "A Probabilistic Choice Model for Market Segmentation and Elasticity Structure," *Journal of Marketing Research*, XXVI (November), 379-390

Krishnamurthi, L., S. P. Raj (1988), "A model of brand choice and purchase quantity price sensitivities," *Marketing Sciences*, 7, 1-20.

Labeaga, Jose and Angel Lopez (1997), "A Study of petrol consumption using Spanish panel data," *Applied Economics*, 29, 795-802.

Maddala, G. (1983). Limited-dependent and qualitative variables in economics, Cambridge: Cambridge University Press.

Manchanda, Puneet, Asim Ansari, and Sunil Gupta, (1999), "The 'Shopping Basket': A Model for Multi-Category Purchase Incidence Decisions," Marketing Science, 18(2), 95-114.

Molina, Jose Alberto (1997), "Two-Stage Budgeting as an Economic Decision-making Process for Spanish Consumers," *Managerial and Decision Economics*, 18, 27-31.

Narasimhan, C., Scott A. Neslin, and Subrata K. Sen (1996), "Promotion Elasticities and Category Characteristics," *Journal of Marketing*, 60 (April), 17-30.

Neslin, Scott, Caroline Henderson, and John Quelch (1985)"Consumer Promotions and the Acceleration of Product Purchases," *Marketing Science*, 4 (2), 147-165.

Nijs, Vincent R., Marnik G. Dekimpe, Jan-Benedict E.M. Steenkamp, Dominique M. Hanssens (2000), "The Category Demand Effects of Price Promotions," *Working Paper*.

Putsis, W. and R. Cotterill (1999),"Share, Price and Category Expenditure - Geographic Market Effects and Private Labels," *Managerial and Decision Economics*, forthcoming, 1999.

Vilcassim, Naufel J. (1993), "Extending the Rotterdam Model to Test Hierarchical Market Structures," *Marketing Science*, 8 (Spring), 181-90.

Vilcassim, Naufel J. and Dipak C. Jain (1991), "Investigating Household Purchase Timing Decisions: A Conditional Hazard Function Approach," *Marketing Science*, Winter,1-23.

Wales, T.J. and A.D. Woodland (1983), "Estimation of Consumer Demand Systems with Binding Non-negativity Constraints," *Journal of Econometrics*, 21, 263-285.

Walters, Rockney G., and Scott B. MacKenzie (1988), "A Structural Equations Analysis of the Impact of Price Promotions on Store Performance," *Journal of Marketing Research*, XXV (February), 51-63.

Walters, Rockney (1991), "Assessing the Impact of Retail Price Promotion on Product Substitution, Complementary Purchase, and Interstore Sales Displacement," *Journal of Marketing*, 55, pp. 17-28

Table 1: Category Composition

Meat	Bakery	Produce	Dry Goods	Alcohol	HH Supplies	HBC
Meat	Industrial Bread	Fruits	Frozen Food	Wine	Kitchenware	Perfumes
Deli	Specialty Bread	Vegetables	Package Goods	Liquor	Hardware	Hygiene
Fish	Danish	Cheese		Beer	Paper	Drugs
Sausages	Pastries	Salad Bar				Health Products

Table 2: Category spending

	Meat	Bakery	Produce	Dry Goods	Alcohol	HH Supplies	HBC
% Trips with Promotion	22.10%	10.30%	54.40%	50.10%	24.50%	15.40%	18.20%
Spending In Category on Non-Promo trips*	€8.23 (€7.49)	€2.90 (€2.78)	€11.53 (€9.99)	€12.64 (€11.95)	€14.06 (€27.17)	€6.52 (€6.32)	€6.64 (€6.52)
Spending in Category on Promo trips	€13.14 (€10.34)	€4.81 (€3.69)	€19.48 (€13.81)	€21.96 (€16.31)	€23.05 (€33.76)	€10.83 (€9.00)	€9.89 (€8.70)
Spending on Promo Items in Category in Promo Trips	€5.85 (€5.18)	€2.45 (€1.71)	€3.59 (€3.17)	€5.30 (€5.58)	€13.58 (€19.81)	€5.01 (€4.93)	€4.61 (€3.82)
Spending on Non-Promo Items in Category in Promo Trips	€7.29 (€8.75)	€2.35 (€3.15)	€15.89 (€12.74)	€16.66 (€14.45)	€9.47 (€26.28)	€5.83 (€7.54)	€5.28 (€7.59)
Trip Spending on Non-Promo Trips	€56.89 (€40.33)	€55.83 (€41.60)	€42.69 (€32.57)	€42.46 (€32.87)	€68.05 (€47.69)	€61.20 (€42.02)	€62.74 (€43.08)
Trip Spending on Promo Trips	€66.96 (€45.91)	€62.87 (€46.43)	€60.04 (€42.71)	€61.38 (€42.81)	€79.95 (€55.68)	€70.60 (€48.12)	€71.27 (€49.16)
Category Market Share (Non-Promo trips)	14.50%	5.20%	27.00%	29.80%	20.70%	10.70%	10.60%

* All spending amounts in Euros, Standard Deviations shown in parentheses.

Table 3: Category Spending Correlation Matrix

	Meat	Bakery	Produce	Dry Goods	Alcohol	HH Supplies	HBC
Meat	1.00	0.12	0.30	0.22	0.05	0.09	0.06
Bakery		1.00	0.20	0.15	0.03	0.06	0.05
Produce			1.00	0.43	0.07	0.18	0.13
Dry Goods				1.00	0.09	0.21	0.16
Alcohol					1.00	0.03	0.02
HH Supplies						1.00	0.13
HBC							1.00

Note: Spendings have been mean centered at the household level.

Table 4: Correlation Between Promoted Spending and Total Category Spending

	Meat	Bakery	Produce	Dry Goods	Alcohol	HH Supplies	HBC
Meat	*0.41*	0.04	0.08	0.07	0.02	0.03	0.01
Bakery	0.04	*0.31*	0.05	0.05	0.01	0.02	0.01
Produce	0.14	0.09	*0.42*	0.17	0.04	0.08	0.04
Dry Goods	0.09	0.07	0.16	*0.46*	0.04	0.08	0.05
Alcohol	0.03	0.01	0.03	0.05	*0.41*	0.01	0.01
HH Supplies	0.03	0.02	0.05	0.05	0.01	*0.39*	0.03
HBC	0.02	0.02	0.05	0.05	0.01	0.05	*0.37*

Note: Rows indicate promotions.

Table 5: Expenditure Model Parameters

Variable		Within Group		Between Group	
		Estimate	P value	Estimate	P value
Intercept		-2.1E-17	1.00	3.913	0.0001
Price	ln P_{Meat}	-0.879	0.0001	0.129	0.0001
	ln P_{Bakery}	-0.264	0.0001	0.078	0.0001
	ln $P_{Produce}$	-0.341	0.0001	-0.004	0.5912
	ln $P_{Dry\ Goods}$	-1.139	0.0001	0.185	0.0001
	ln $P_{Alchool}$	-1.350	0.0001	0.013	0.0139
	ln P_{HHSup}	-0.873	0.0001	0.070	0.0001
	ln P_{HBC}	-0.653	0.0001	0.042	0.0001
Promotion	Pm_{Meat}	0.180	0.0001	0.691	0.0001
	Pm_{Bakery}	0.131	0.0001	0.233	0.0002
	$Pm_{Produce}$	0.097	0.0001	-0.670	0.0001
	$Pm_{Dry\ Goods}$	0.141	0.0001	-0.293	0.0001
	$Pm_{Alchool}$	0.327	0.0001	1.375	0.0001
	Pm_{HHSup}	0.208	0.0001	1.200	0.0001
	Pm_{HBC}	0.202	0.0001	1.324	0.0001
Inventory	ln I_{Meat}	-0.003	0.0001	0.028	0.0014
	ln I_{Bakery}	0.009	0.0001	-0.027	0.0001
	ln $I_{Produce}$	-0.008	0.0001	0.041	0.0001
	ln $I_{Dry\ Goods}$	-0.020	0.0001	0.001	0.8002
	ln $I_{Alchool}$	-0.029	0.0001	0.021	0.0001
	ln I_{HHSup}	-0.013	0.0001	0.036	0.0001
	ln I_{HBC}	-0.013	0.0001	0.008	0.0042
Geo-Demographic	HH Size			-0.0001	0.3539
	Income			-9.18E-9	0.8170
	Food Pot			2.91E-6	0.5818
	Social 1			0.084	0.0001
	Social 2			0.046	0.0053
	Social 3			-0.225	0.6192
	Social 4			0.018	0.2519
	Social 5			0.008	0.6179
	Social 6			0.003	0.3480
	Store 140			-0.093	0.0001
	Store 146			0.001	0.8900
	Store 624			0.027	0.0002
	n	528,207		23,635	
	R^2	0.1398		0.3478	

Table 6: AIDS Model Parameters

	Variables	Meat	Bakery	Produce	Dry Goods	Alcohol	HH Supplies	HBC[+]
	Intercept	0.157926	0.082333	0.355045	0.239874	0.043905	0.051618	0.06930
Price	ln P_{Meat}	**-0.303965**	0.011354	0.065115	0.15214	0.018856	0.032705	0.02380
	ln P_{Bakery}	<u>0.001054</u>	**-0.05218**	<u>0.00411</u>	0.032648	<u>0.004238</u>	0.007701	<u>0.00243</u>
	ln $P_{Produce}$	<u>-0.001592</u>	-0.00428	**-0.10553**	0.074116	0.012646	0.015552	0.00909
	ln $P_{Dry Goods}$	0.022286	0.011626	0.078701	**-0.18936**	0.025665	0.032558	0.01852
	ln $P_{Alchool}$	0.04712	<u>0.008366</u>	0.144377	0.206107	**-0.47125**	0.040259	0.02502
	ln P_{HHSup}	<u>0.015046</u>	<u>0.015023</u>	0.114231	0.141603	<u>0.011643</u>	**-0.29798**	<u>0.00043</u>
	ln P_{HBC}	0.023757	0.018983	0.098864	0.109646	<u>0.01078</u>	0.008334	**-0.27036**
Promotion	Pm_{Meat}	**0.097955**	-0.00503	-0.02576	-0.04472	-0.00627	-0.00766	-0.00851
	Pm_{Bakery}	<u>-0.002311</u>	**0.057081**	-0.01407	-0.02807	<u>-0.00176</u>	-0.00703	-0.00385
	$Pm_{Produce}$	0.010164	-0.00713	**0.069483**	-0.04162	-0.00831	-0.00865	-0.01394
	$Pm_{Dry Goods}$	-0.012876	-0.00976	-0.04254	**0.095168**	-0.00721	-0.01018	-0.01261
	$Pm_{Alchool}$	-0.017862	-0.00679	-0.05597	-0.05799	**0.161788**	-0.01341	-0.00977
	Pm_{HHSup}	-0.009601	-0.00517	-0.03157	-0.04692	-0.00827	**0.10834**	-0.00681
	Pm_{HBC}	-0.010763	-0.00437	-0.03182	-0.0423	-0.00829	<u>-0.00274</u>	**0.10028**
Inventory	ln I_{Meat}	**-0.018742**	0.001813	0.004373	0.005336	0.002506	0.002317	0.00240
	ln I_{Bakery}	0.00166	**-0.0131**	0.001123	0.003308	0.002362	0.002331	0.00232
	ln $I_{Produce}$	0.003023	0.002063	**-0.02257**	0.009261	0.002201	0.003047	0.00298
	ln $I_{Dry Goods}$	0.002952	0.001618	0.009085	**-0.02378**	0.002833	0.003869	0.00342
	ln $I_{Alchool}$	0.00332	0.001528	0.006899	0.00711	**-0.02531**	0.003263	0.00319
	ln I_{HHSup}	0.00267	0.001353	0.00507	0.006531	0.002825	**-0.02159**	0.00314
	ln I_{HBC}	0.002822	0.001496	0.005956	0.006657	0.003054	0.003238	**-0.02322**
	ln (X/P)	-0.018079	-0.01537	-0.01683	0.040443	0.007043	0.007224	-0.00443

The own coefficients are in **bold**. Coefficients that are not significant at the 0.0001 level are <u>underlined</u>.

[+]The HBC equation was dropped from the estimation. The associated parameters are derived from the adding-up constraints. As a consistency check, we reran the analysis dropping HH Supplies rather than HBC. Consistent with the theory of demand systems that says that the parameter estimates are invariant to which equation is dropped, the results were identical to the 6th decimal. The Significance levels for this equation are drawn from this second analysis.

Table 7: Expenditure Elasticities

Meat	Bakery	Produce	Dry Goods	Alcohol	HH Supplies	HBC
0.813224	0.514852	0.918475	1.192075	1.060575	1.124174	0.932547
(0.0041)	(0.0088)	(0.0031)	(0.0034)	(0.0032)	(0.0753)	(0.0050)

Standard error for the elasticities are shown in parentheses. The standard error for the HBC elasticity was calculated by dropping HH Supplies from the model rather than HBC and rerunning the model.

Table 8: Price Elasticities

	Meat	Bakery	Produce	Dry Goods	Alcohol	HH Supplies	HBC
Meat	**-4.12229**	0.417369	0.32591	0.694387	0.154928	0.547968	0.370072
	(0.0271)	(0.0632)	(0.0196)	(0.0231)	(0.0231)	(0.0405)	(0.0372)
Bakery	0.017918	**-2.63158**	0.023387	0.145718	0.034521	0.127999	0.03917
	(0.0310)	**(0.0640)**	(0.0209)	(0.0252)	(0.0257)	(0.0515)	(0.0403)
Produce	0.042614	0.014602	**-1.4943**	0.28582	0.091477	0.229007	0.158363
	(0.0100)	(0.0217)	**(0.0077)**	(0.0081)	(0.0078)	(0.0374)	(0.0123)
Dry Goods	0.292412	0.524292	0.41087	**-1.93976**	0.200978	0.517676	0.304704
	(0.0754)	(0.1622)	(0.0566)	**(0.0606)**	(0.0578)	(0.1009)	(0.0913)
Alcohol	0.503515	0.308211	0.70788	0.959263	**-5.05989**	0.682034	0.386829
	(0.1643)	(0.3821)	(0.1286)	(0.1418)	**(0.0965)**	(0.2227)	(0.2030)
HH Supplies	0.168392	0.506334	0.559962	0.657566	0.096119	**-6.12937**	0.011526
	(0.1571)	(0.3435)	(0.1090)	(0.1066)	(0.1239)	**(0.1860)**	(0.1829)
HBC	0.256222	0.626501	0.484572	0.507296	0.089379	0.135253	**-5.11361**
	(0.1015)	(0.2247)	(0.0702)	(0.0721)	(0.0804)	(0.1276)	**(0.1058)**

Standard error for the elasticities are shown in parentheses. The standard error for the HBC elasticities were calculated by dropping HH Supplies from the model rather than HBC and rerunning the model.
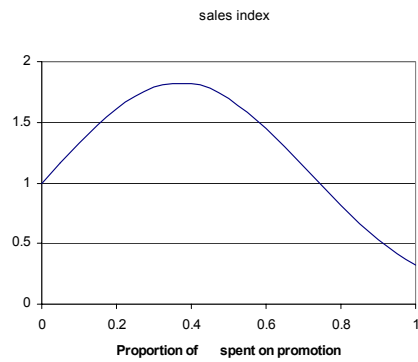
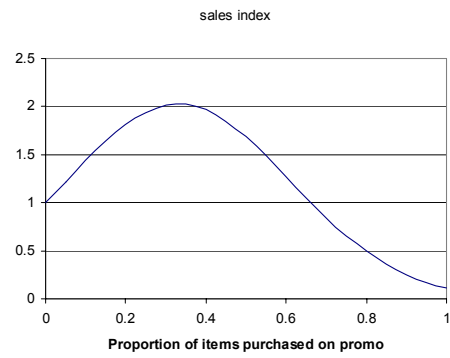**Figure 1(a) € Spend**   **Figure 1(b): Number of Item Bought**

sales index

Proportion of   spent on promotion

sales index

Proportion of items purchased on promo

Figure 2: Classification of shopping trips

No. of item purchased

**Regular
58.7%**

10

**Filler
39.6%**

**Cherry-Picking
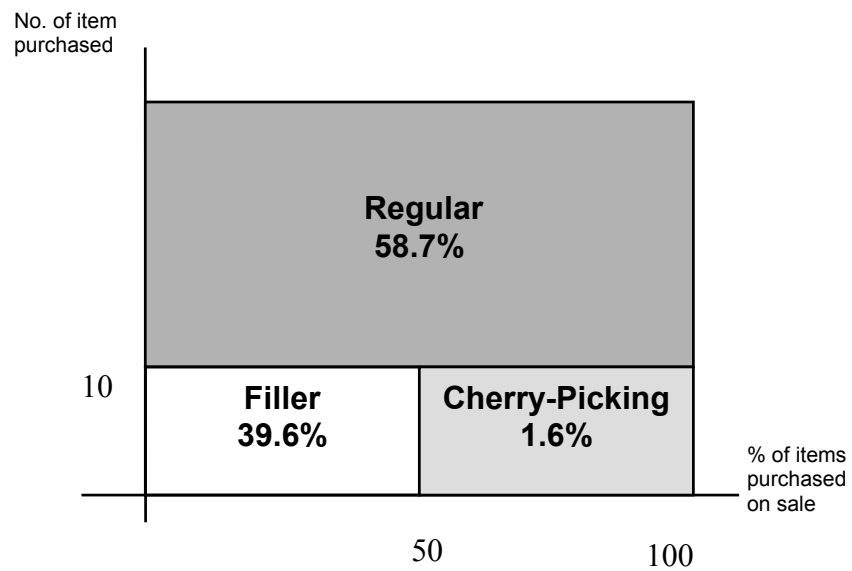1.6%**

% of items purchased on sale

50   100

Figure 3: Classification of Households
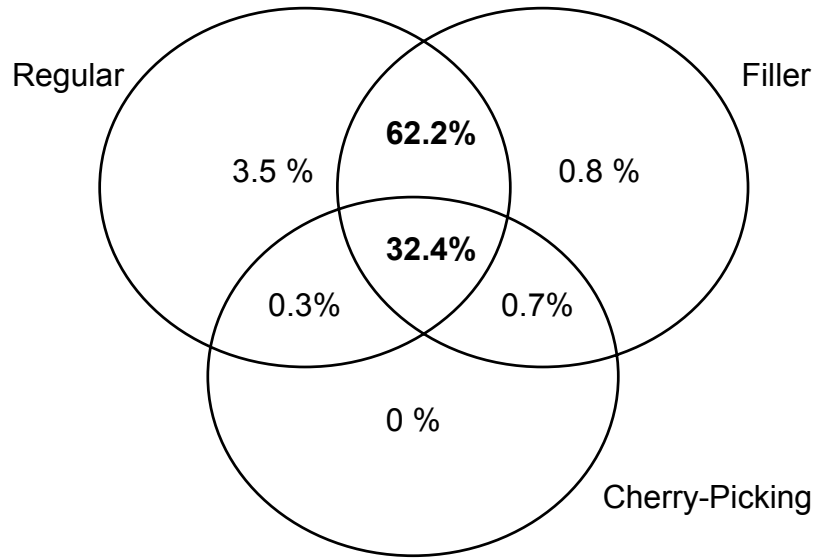
Figure 4: Inter-shopping Time (in Days)
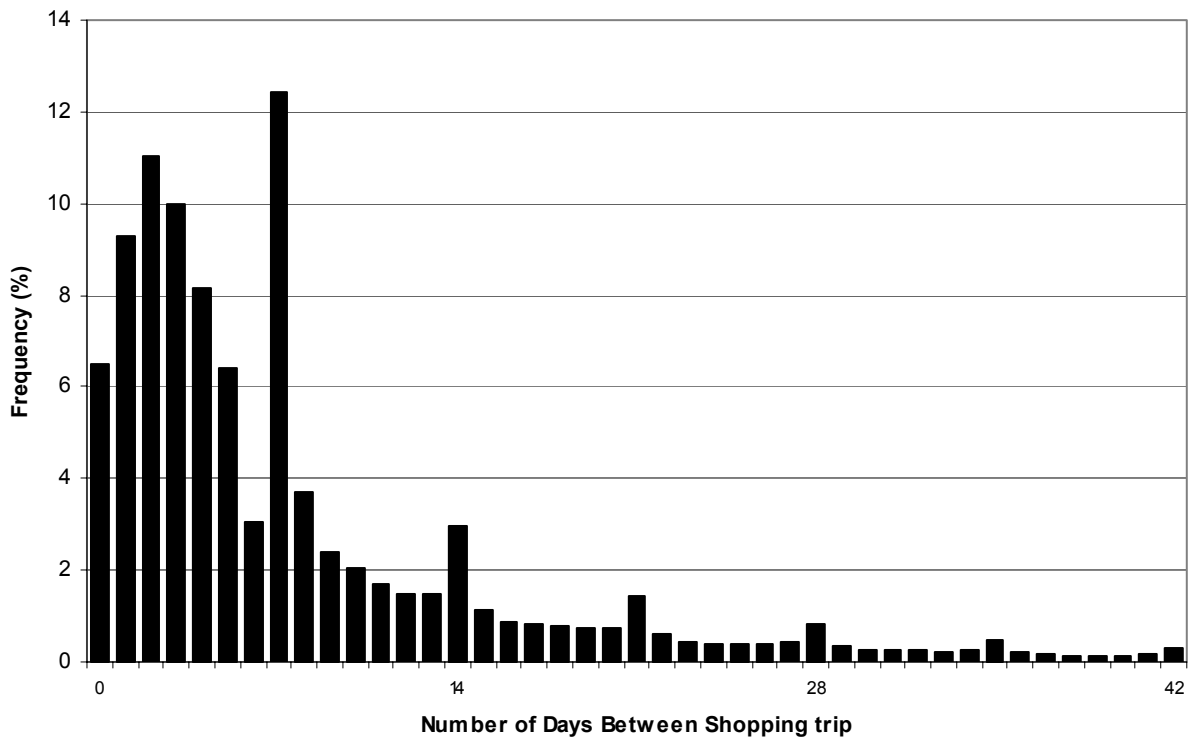
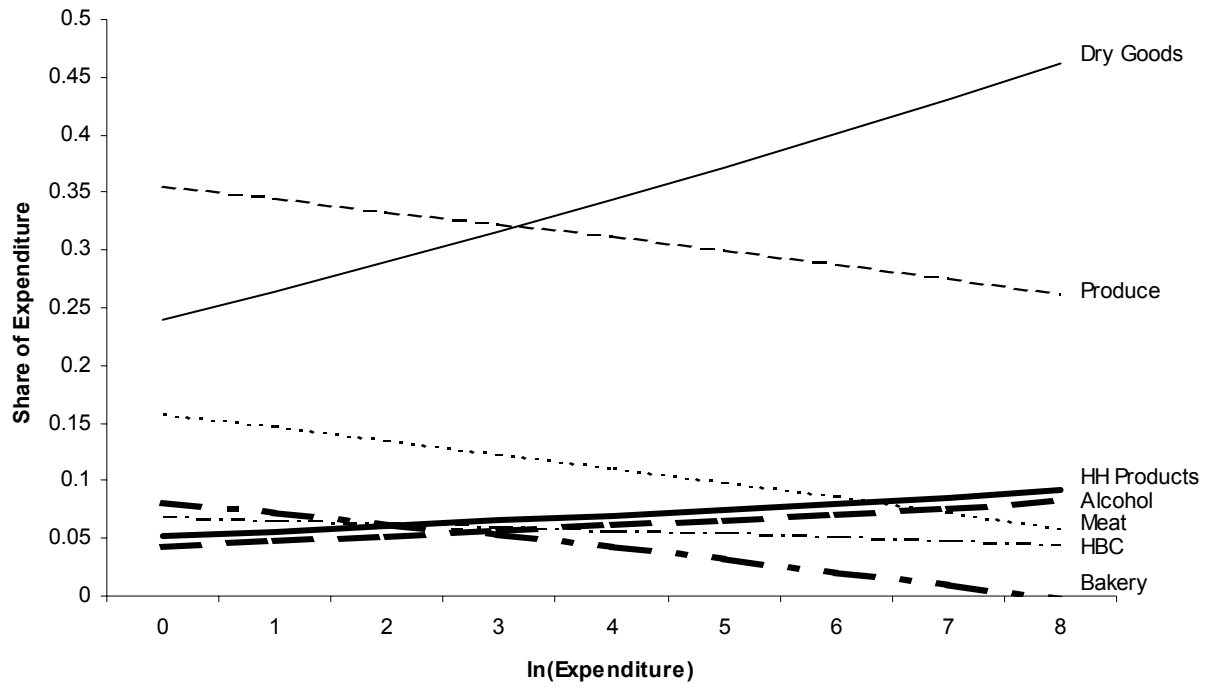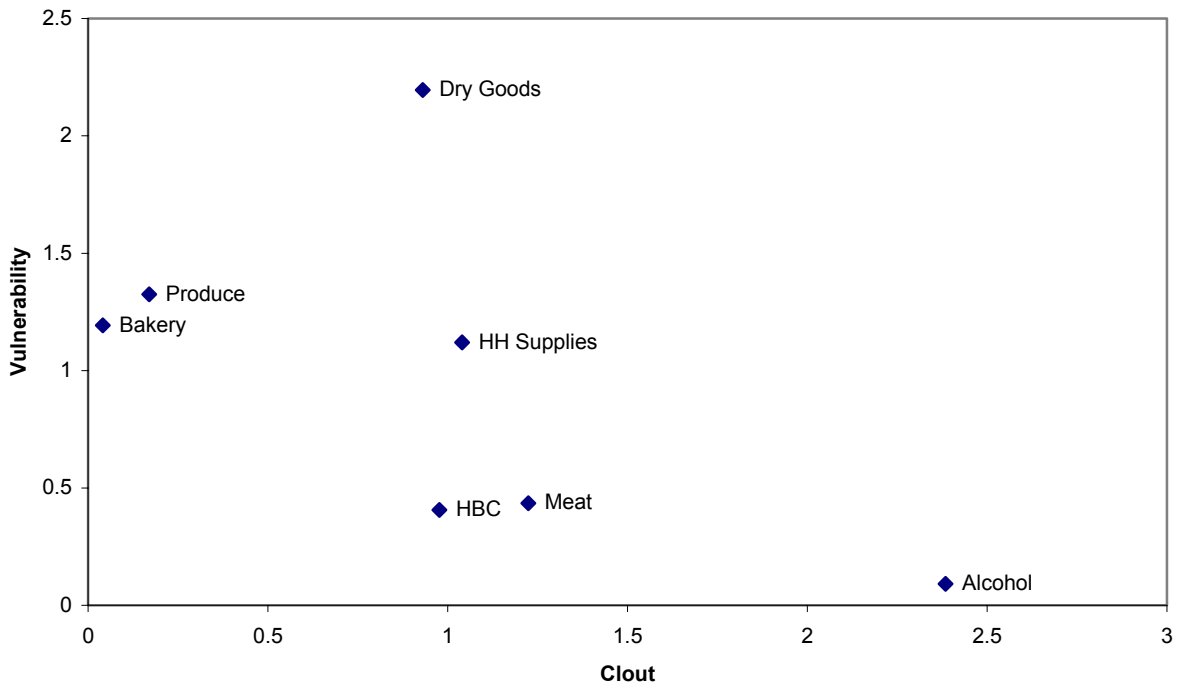**Figure 5: Share of Expenditure**

Figure 6: Clout vs. Vulnerability

[1] We will describe more formally the different product groups in the next section. Also, we use the terms "product," or "product groups" interchangeably.

[2] Another stream of research (e.g., Fader and Lodish 1990; Narasimhan, Neslin and Sen 1996; Hoch, Kim, Montgomery, and Rossi 1995) has used cross-sectional analysis to explain variations in factors of interest (e.g., price elasticity) as a function of common product markets characteristics. However, those studies too do not address the issue of the determinants of the level of household expenditures or the allocation of that expenditure across different products.

[3] In both cases, the figures were drawn based on the estimates obtained from regressing the total expenditure (total # of items) against a quadratic function of the proportion of total expenditures on sale items (or proportion of number of items on sale), after controlling via dummy variables for store effects.

[4] We chose 10 items as the cut-off point based on discussions with the management of the retail chain. It is also the maximum number of items that can be purchased at the "express" checkout counters.

[5] Presumably, they do their regular shopping trips at another store.

[6] We recognize that all operationalizations of inferred levels of inventory are flawed. We have attempted to eliminate some of the obvious errors, given the level of product aggregation.

[7] We have tried other inventory formulations. We have used a straight Gupta model using $Q_{g,t}^{(h)} = E_{gt}^{(h)} / P_{gt}^{(h)}$ as a quantity measure. We also used a non-constant usage-rate formulation derived from Neslin, Henderson, and Quelch (1985). The results from our analyses show that the price and promotion coefficients are not affected (to the second decimal) by the specification of the inventory formulation. However, the inventory coefficients have more face validity under the current specification.

[8] Because the AIDS model is non-linear, overall elasticities are computed by averaging the point estimate elasticities for each observation in the data set.

[9] Hausman and Taylor (1981) suggest an alternative estimation procedure to obtain consistent estimates of the demographic variables. The procedure would involve partitioning the set of variables $\mathbf{V}$ (equation 12) into two, one set of which is orthogonal to the error term and hence, can be used as instruments. In our case, this is not possible because all the variables of $\mathbf{V}$ are household specific and are likely to be correlated with the composite error term.

[10] Note that the dependent variable is expressed as deviations from the mean and hence, it tends to lower the estimated R^2 value.

[11] The economic literature traditionally classifies products based on their expenditure elasticities as *necessity* and *luxury* goods. However, given the nature of our products, *necessity* and *discretionary* products seem more appropriate while retaining the same connotation.