IMPACT OF SUPERMARKET PRODUCT DIFFERENTIATION WITH ASYMMETRIC INFORMATION ON CONSUMER BEHAVIOR

by

Suzanne L. Webster.*

Patrick Canning*

May 15, 2003

Paper prepared for presentation at the American Agricultural Economic Association Annual Meeting Montreal, Canada July 27-30, 2003

We are very grateful to Mark Denbaly for suggesting and supporting the project. We thank Charles Hallahan and J. Michael Harris for invaluable assistance with the production of the data.

**Suzanne L. Webster is an economist at the Economic Research Service at the U.S. Department of Agriculture, Washington, DC., and a doctoral candidate at the American University, Washington, DC. Patrick Canning is a senior economist at the Economic Research Service – U.S. Department of Agriculture, Washington, DC. Direct all correspondence to Suzanne Webster at swebster@ers.usda.gov*

INTRODUCTION:

The paper will examine the empirical evidence behind the significant growth in the sale of private label food products. For the most recent five years, supermarket private label products have gained in both dollar and unit share. The unit share of private label products in 1999 was 20.1%, increasing from 15% in 1995, while the dollar market share increased to 15.7% from 14.6% in 1992.[1] By linking individualized consumer food purchase data with advertising data for brand and private label food products the study will assess the price and information determinants of the consumers' demand for private label food products.

With the growth in two-worker households, increased time constraints have been placed on consumers. Consumers are now turning to convenience foods and eating out more often. As a result, supermarkets, a key segment in the food marketing system, have begun to compete both with restaurants and fast-food outlets, as well as with large discount retailers like Costco and Walmart. This has led, beginning in 1996, to their consolidating at a rapid pace. That in turn, resulted in changes in market structure and market share. Various economists offer different reasons for this supermarket consolidation, but that is beyond the scope of this research. What is relevant is that with a larger scale of market retailing a different kind of private label product has emerged. The question that arises is what role the consumer and private label promotion have played in driving these changes.

Advertising can take different forms: 1) national media, 2) trade promotion, or 3) coupons, or it can consist of a combination of any of these three. The nature of the way in which private label and national brand products are marketed may explain the gain in the private label share. The question examined in this paper is whether the consumers' response to the relative advertising intensities for brand and private label products affects the products' market share, or if only price matters. To examine this question, regression analysis is carried out to explain the

relationship between economic and demographic characteristics of the consumer, product specific data on prices and promotion, and their purchases of food products.

First, the paper will present a series of descriptive statistics based on race and life stage groupings from the linked dataset. Then, the paper will focus on how advertising influences the household based on Race and Life Stage groupings and explain what lies behind these statistics.

DATA

SOURCE OF DATA

The data set created for this paper required linking information taken from two databases, the AC Nielsen Homescan (Nielsen) database and the Leading National Advertiser's (LNA) database. First, information on private label food categories was taken from the *Private Label Manufacturers Association Yearbook (PLMA)*. Second, information on the leading brand products in the various food categories was taken from the Information Resources, Inc. (IRI) database. Then the LNA database was used to obtain the advertising data for the brand and private label food products and the Nielsen database was used to obtain information on the individual consumer food consumption and demographic composition of the households. This section describes each of these databases in detail as well as the process that was used to create the database.

The analysis uses new data produced by AC Nielsen and purchased for ERS-USDA use. The AC Nielsen Homescan is sampled from a 52,000-member panel, which is geographically dispersed and demographically representative. The sample is based on fifty-two (52) markets, four (4) U.S. census regions and eleven hundred and twenty seven (1127) rural households that

[1] Private Label Manufacturers Association, Executive Summary, 2000.

are not included in the fifty-two markets. It is balanced by household count, by county size within market, and remaining census regions.

The projection methodology incorporates county level demographics that were used to determine sample size targets. With robust sample representation of markets and the remaining U.S., consumer purchasing was captured.

The time frame is made up of weekly purchases for a 52-week period from January 3, 1999, to January 3, 2000. The purchase data includes information on weekly purchases by product category, brand/private label, product size, UPC code, UPC description, quantity, price-deal, price non-deal, coupon value and any common product attributes. Price-deal is the actual price after a sale price reduction. Coupon value represents the dollar amount the price would be reduced by presenting the coupon. Price non-deal is the actual shelf price, so the data has both prices before and after the sale.

CHOICE OF FOOD CATEGORIES

Various food categories that make up the database were chosen from the *Private Label Manufacturers Association (PLMA) Yearbook* for the years 2000 and 2001, based on data from Information Resources, Inc, for the years 1999 and 2000 respectively. The categories that were chosen were from among the top 20 Private Label categories, both in unit volume and dollar volume, as well as those where private label had been among the top 10 unit and dollar volume gainers. Commonly eaten categories were also chosen, such as Bread, Cereal, Soup, and Orange Juice. The following food categories were arrived at:

I. DAIRY section: a) Ice Cream/Sherbet/Frozen Yogurt,
   b) Refrigerated Yogurt, c) Cottage Cheese.

II. BAKERY section: a) Fresh Bread/Rolls, b) Pies and Cakes.

III. MEATS/FISH section: a) Frozen Poultry, b) Frozen Seafood, c) Dinner Sausage, (semi-prepared refrigerated and frozen, not fresh) d) Luncheon Meats.

IV. FRUITS/VEGETABLES section: a) Canned Tomato Products, b) Frozen Vegetables, c) Canned/Bottle Fruit, d) Frozen Orange Juice.

V. OTHER section: a) Canned Soups, b) Cold Cereal, c) Pickles/Relish/Olives.

SELECTION OF LEADING BRANDS

From the various food categories that were chosen from the *PLMA Yearbook*, the leading brand products were developed in each of the seventeen categories using the IRI database housed at ERS. Tables on brand products were sorted by dollar sales to arrive at the top 5 or 6 brand products within each of the seventeen categories. Due to a lack of advertising data, only seventy-two (72) brand products made up the database to be analyzed.

ADVERTISING DATA FOR FOOD PRODUCTS

The Leading National Advertisers/Media Watch Multi-Media (LNA) database is published quarterly by Competitive Media Reporting and reports advertising expenditures in ten major media. The LNA database includes only companies spending $25,000 or more annually, to advertise a product in the combined ten media measured.

LNA measure and compile all paid advertising space and expenditures in member publications, currently 180+ consumer magazines and four Sunday magazines. They measure advertising space in 203 newspapers in the top 40 markets. Advertising expenditures are based on one-time open inch rates. Outdoor data includes billboard expenditures (poster and paint) in over 200 plant operator markets. Information on brand expenditure is provided by participating

plant operators and represents gross sales volume for participating plant operators only. Markets are generally defined according to Standard Metropolitan Statistical Areas (SMSA).

The MediaWatch section of LNA electronically monitors every broadcast minute on ABC, CBS, FOX and NBC television networks. MediaWatch's technique for estimating brand expenditures is to assign an estimated rate to each network program and to apply this rate to each brand commercial monitored in the program. Estimated brand expenditures are the total of each brand's monitored activity.

MediaWatch reports on national commercial television activity occurring in satellite distributed syndicated television. They gather program rate information each month from a large cross-section of advertisers and advertising agencies. This rate information is used to create representative rates for each program. These rates are then applied to each national commercial in the program.

MediaWatch electronically monitors seventeen cable television networks around the clock. They include, Arts & Entertainment, Black Entertainment Network, Cable News Network, Consumer News and Business Channel, Comedy Central, Discovery, Entertainment and Sports Programming Network, The Family Network, Headline News, Lifetime, Music Television, The Nashville Network, Nickelodeon, TBS SuperStation in Atlanta, Turner Network Television, USA Network, and Video Hits 1.

MediaWatch reports radio activity during news programming for 13 networks offered by three radio networks. These include five ABC radio sales networks, two CBS sales networks, and six Westwood radio air networks. Feature program activity is excluded. LNA provides national and regional advertising data from 3,500 stations in over 200 markets. This information is acquired through major national station representative organizations, and the dollars are based on station representative billing figures.

For each of the 72 brand Products, the LNA database was searched for the advertising expenditures in 1997, 1998 and 1999. The last step was to weight the brand advertising expenditure amounts based on the share of each brand product purchased in each census region. Multiplying the percentage of each brand product times the annual advertising budget for that product resulted in a weighted advertising average. These amounts were then summed by census region and imputed for brand advertising budgets by households in a particular census region. Each household in a given census region would receive the same advertising dollars, which resulted in variation across census regions but not within a census region.

A) PRIVATE LABEL ADVERTISING

Since the Nielsen Homescan database does not identify which supermarkets the households shopped in, the following process was derived as a proxy for the private label advertising that the households would experience. Private label advertising was constructed in three steps. In Step 1, a list of the top 50 supermarket chains (chains) for 1999 was taken from *Progressive Grocer's 2001 Marketing Guidebook (Guidebook).*[2] Advertising expenditures for 43 of the top 50 chains were obtained for the years 1999, 1998, and 1997 from the LNA database.

Step 2 involved using the 2000 *Market Scope, Desktop Guide to Supermarket Share (Market Scope)* book. The AC Nielsen Scantrack Market information contained in this book described data for each of their fifty (50) major Market areas.[3]

Each Market area contained information on the number and market share of all major supermarkets in that Market area, as well as their population and number of households. By dividing each chain's annual advertising budget by the total number of chain stores, a per-store-advertising amount was derived. The number of stores in that region was multiplied by the chain's per-store-advertising amount, which resulted in an advertising budget for each chain in

[2] Progressive Grocer's Guidebook, p.30.

that Market region. A similar process was performed for each of the four rural regions. Since each household in the Nielsen Homescan database is identified by a Market region code or rural, and a Census code identifier identifies all households, the private label advertising can now be linked directly to each household by means of their Market region code and/or Census code identifiers.

The third and final step involved weighting all the final advertising expenditure amounts. A weighted-average was taken for each market region based on the number of stores each supermarket chain had in each market region. Multiplying the share of stores by the per-store advertising budget resulted in a weighted average for each supermarket chain in each market region that was then summed for a weighted-average by market region.

B) BRAND ADVERTISING

For each of the 72 brand Products, the LNA database was searched for the advertising expenditures in 1997, 1998 and 1999. Brand products for which there was no advertising data at all were dropped from the database.

The last step was to weight the brand advertising expenditure amounts based on the share of each brand product purchased in each census region in the dataset. This step was necessary in order to follow the same process used in the private label advertising data. To arrive at this amount, each food category file was first sorted by census region and then by brand product. The expenditure on each brand product was first summed and then the percent of total expenditure by region was calculated. The last step was to multiply the percentage of each brand product by the annual advertising budget for that product to arrive at a weighted-average advertising budget for that brand product by census region. These amounts were then summed by census region and imputed for brand advertising budgets by households in a particular census

[3] New York City had three market areas which were combined to coincide with the Market Scope data.

region. Each household in a given census region would receive the same advertising dollars, which resulted in variation across census regions but no variation within a census region.

Other data steps:

Purchases by Household ID were aggregated annually within the datasets. Where a household bought either brand or private label products, but not both, it was necessary to impute prices in order to calculate shares. First, average prices were calculated by region and food product. Total expenditure by region was divided by the total regional equivalent quantity to reach an average price for brand and private label food categories. These regional average prices by food category were then used as a proxy for prices for a household that did not purchase both brand and private label products.

A price index was calculated to represent a price deflator for total food expenditure. This variable will represent prices for "all other goods." The first step for the price index was to calculate the total expenditure by region on each of the sixteen food categories, both brand and private label. The second step involved calculating the share of each of the brand and private label food products by region, by dividing the total expenditure for the food product by the total expenditure for the entire region. For example, total expenditure for brand cereal in the East was divided by total expenditure for all brand food products in the East. Third, the share was then multiplied by the regional average price previously calculated to reach a weighted-average price. Then all prices for each census region for brand and private label were summed. The last step was to calculate the percentage of expenditure for brand and private label products in each region and multiply that by the previously calculated total price to reach a weighted-average price for each of the four census regions for "all other goods."

The last data step involved placing each of the 7200 individual households into nine "Life Stage" categories that are commonly used in industry standards.[4] The following are the nine household life stages and their definitions:

1) YOUNG SINGLES: Households with one member under 35 years of age. (172 or 2.39%)
2) CHILDLESS YOUNGER COUPLES: Households with 2+ members, under 35 years of age with no children under age 18. (211 or 2.93%)
3) NEW FAMILIES: Households with children under six years of age only. (335 or 4.66%)
4) MATURING FAMILIES: Households with children aged 6-12, or children under 6 and children 6-12, **or** children under 6 and children 13-17, **or** children 6-12 and children 13-17, **or** children under 6 and 6-12 and 13-17. (1226 or 17.04%)
5) ESTABLISHED FAMILIES: Households with children aged 13-17 years of age only. (622 or 8.64%)
6) MIDDLE-AGED SINGLES: Households with one member between 35-54 years of age. (722 or 10.03%)
7) MIDDLE-AGED CHILDLESS COUPLES: Households with 2+ members, between 35-54 years of age, with no children under 18. (1271 or 17.67%)
8) EMPTY NESTERS: Households with 2+ members, aged 55+ years, with no children under 18 years of age. (1956 or 27.19%)
9) OLDER SINGLES: Households with one member who is 55+ years of age. (680 or 9.45%)

[4] 2001 Product Preference Study developed by Progressive Grocer and AC Nielsen.

METHODOLOGICAL APPROACHES AND SOME ECONOMETRIC RESULTS:

There are at least two alternative models that account for the effects of relative prices and advertising on the share of private label food products. One is based on household production theory (Gorman, Becker) and the second is based on traditional consumer theory (Marshall). Briefly, these two models are as follows.

Model I:

Purchasing a product for consumption typically involves a decision to address an underlying desire to engage in a utility deriving activity and a decision to purchase specific product(s) that meet this underlying desire. These joint decisions could be explicit and direct, for example a desire to immediately consume a pint of fruit juice, and a decision to buy a pint of a nationally branded fresh orange juice product that directly fulfills that desire. Joint choices can also be complex and part of a larger decision process. For example, a desire to provide household members with a gourmet dessert cake, and a decision to purchase combined recipe ingredients (including a pint of a nationally branded fresh orange juice product), ones own cooking efforts, and household appliances, given a favorite recipe to produce such a cake.

Even with detailed market level or individual transactions data summarizing the outcomes of household purchasing decisions, it is extremely unlikely that all the elements that went into such decisions are observed by the data available to an analyst. For example, an unexpected seasonal surge in the supply of fresh blueberries that caused the market price to plummet may be the prime factor leading to the decision to make a cake, and thus, to buy fruit juice. Alternatively, unseasonably warm weather may be behind a surge in fruit juice sales. Advertising, such as from a retailer posting dessert recipes that include the blueberries, could also influence a surge in juice sales.

While the traceability of all the determinants leading up to a decision to purchase a pint of fruit juice can be overwhelming, it is often possible to examine the determinants of certain stages in this decision. Household production theory (Gorman; Becker; Lancaster) is ideally suited for such cases. This theory of the household hypothesizes that utility deriving activities are the 'upper stage' of a household's decision process. In this upper stage, non-market goods, indexed over N utility deriving activities and denoted Z_n, are consumed in proportions that are most advantageous to the household given shadow prices, π_n. In the lower stages, each household will minimize the full cost of creating the non-market goods by combining market goods and services, X_m for $m \in M$, with a portion of their personal time endowment. Along with these market goods and time endowment, exogenous household capital, including human capital, are instrumental in determining the least cost outcome to the lower stage decision. As we demonstrate below, marketing information can interact with exogenous human capital in this framework to affect the decision outcome in both the lower and upper levels.

The theoretical model

Let 'h' denote one of H household groups, and let I^h denote the budget constraint of household h. The upper stage optimization problem for this household is stated in equation 1:

$$Max U^h(Z_1,...,Z_N) \ \ subject \ \ to \ \ \sum_{n=1}^{N} \pi_n Z_n = I^h \qquad (1)$$

Assuming the usual properties of the utility function in (1), an indirect utility function—$V^h(\boldsymbol{\pi},I^h)$—exists, and its properties are well documented (see Deaton and Muellbauer). One of these properties, known as Roy's Identity, states that the Marshallian demand function for any good or service, $n \in N$, is obtained from the indirect utility function as follows:

$$Z_n^h(\pi, \mathrm{I}^h) = \frac{\partial V^h(\pi^h, I^h)/\partial \pi_n^h}{\partial V^h(\pi^h, I^h)/\partial I^h} \tag{2}$$

The superscript (h) on the shadow price reflects the result that each household faces a unique composite unit cost for Z_n, as determined by cost minimization in the lower level problem.

Shadow prices, π, reflect marginal costs of producing non-market goods from purchased market goods and services, and household labor. Let $C(Y,l_Z,Z;K)$ denote the household production used to produce Z, where Y is a vector of markets, l_Z is household labor used to produce non-market goods, and K is a vector of exogenous household capital. Since Y is an index of commodity markets, not actual products, the market price vector, Q, depicts average market prices. If products in market Y_m are not differentiated, then the average price in market m is the same as the product price, and quantity used of the commodity equals the quantity used of the product in this market. By minimizing input costs, given market prices, Q and w, a cost function of the form C(Q,w,Z;K) exists and has the following properties:

$$\partial C/\partial Q_m = Y_m \qquad \partial C/\partial w = l_Z \qquad \partial C/\partial Z_n = \pi_n \tag{3}$$

Suppose that a subset of the non-market goods is produced by combinations of F different food market commodities and household labor. Further, suppose each food market has two types of products, X^1 and X^2 that the household can choose from. While these products are classified in the same commodity line, f, they are distinct and are marketed with different advertising programs, A^1 and A^2. It is assumed that this advertising interacts with household human capital to affect productivity.

For any market f, the household commodity production function combines products in market f with household labor and exogenous household capital to produce the desired quantity of commodity f as follows:

$$G\left(X_f^1, X_f^2; K(A_f^1, A_f^2)\right) \tag{4}$$

Product market demand expressions are obtained from solving the minimum cost solution to purchases of products 1 and 2 subject to G(…) = Y_f. This leads to a product market cost function with the following properties:

$$\partial G(P_f^1, P_f^2, \overline{Y}_f; A_f^1, A_f^2) \big/ \partial P_f^i \equiv G_i = X^i{}_f\,(P_f^1, P_f^2, \overline{Y}_f; A_f^1, A_f^2), \quad for\ i = 1,2 \tag{5}$$

Linear homogeneity of household production at all levels, including non-market and commodity market production, allows us to specify a numeraire good, such that prices in all markets are expressed relative to the price of the numeraire good. For the derivations that follow, it is convenient to assume that product 2 in market f is the numeraire good.

Since this research seeks to explain the recent trend of continuing gains in market shares among private label food products, the following expression is defined to depict private label market share, where a '1' superscript denotes the private label product and a '2' superscript denotes a national brand product:

$$S_f^0(p_f^1, \overline{Y}_f; A_f^1, A_f^2) = p_f^1 G_1 \big/ \left(p_f^1 G_1 + G_2\right) \tag{6}$$

Recall that the function G(…) describes the household process of assembling products within a common food commodity line, say bread, to fulfill an 'upper level' decision to provide the household with the necessary commodities to assemble the subset of Z non-market goods that require food commodities from commodity markets 1 to F. A popular functional form employed in applications of diet analysis using household production theory is known as the linear characteristics model (see Deaton and Muellbauer). This model hypothesizes a fixed coefficient (linear) function to describe the aggregation process. While such an assumption appears too restrictive for the present application, a popular, and somewhat generalized functional form known as the Constant Elasticity of Substitution function, or CES, is general enough to

encompass the linear characteristics function, as well as the Cobb-Douglas, and linear homogeneous CES functions. Aside from the linear homogeneity restriction that we have imposed on household production, the only other important restriction is that the elasticity of substitution between productions 1 and 2 remains constant within the area of observed behaviors (see Varian). The CES has the added benefit of expressing the exogenous advertising by the two marketing entities as a single ratio—A^1/A^2. For commodity market f, the CES function is as follows:

$$G\left(X_f^1, X_f^2; K(a_f^1)\right) = \left[K(a^1)\left(X^1\right)^{\frac{\sigma+1}{\sigma}} + K\left(X^2\right)^{\frac{\sigma+1}{\sigma}} \right]^{\frac{\sigma}{\sigma+1}} \qquad (7)$$

where a^1 is the ratio A^1/A^2 and σ is the constant elasticity of substitution between X^1 and X^2. It is left to the reader to confirm that substitution of expression (7) into expression (6) and total differentiation yields an expression whose only arguments on the right-hand side are expression in p^1 and a^1. Specifically:

$$\partial S_f^0(p_f^1; a_f^1) \big/ \partial p_f^1 = S_f^0(1 - S_f^0)(\sigma^{-1} + \hat{p}_f^1) \qquad (8)$$

where p hat refers to the percentage change in relative price. This result implies that market share of private label will decrease with its relative price in all but those cases where the household (σ>-1) considers the national brand alternative to be a weak substitute for the private label product. It is also straightforward to show that higher relative advertising increases the value market share of the private label product under normal assumption about the role of advertising in the model.

Model II:

Traditional consumer theory assumes that each household seeks to maximize the utility it derives from its consumption purchases in M separate markets. The welfare function is implicitly additive separable in consumption across the M markets subject to an exogenous

budget constraint. Equation 1) establishes the welfare of the household (W^h) as a function of the utility derived from the consumption of food and non-food items (X^M).

1) $W^h = \sum U^M [X^M (h)]$

Equation 2) decomposes Eq. 1 into its subcomponents, [non-food (X^N) and food (X^F)] where each market is allocated a separate budget. Leisure is also separable from consumption of food, which essentially means that work and income are exogenous with respect to the consumption of food.

2) $W^h = U^N [X^N (h)] + U^F [X^F (h)]$

The study will deal only with food products. Each market good has three arguments, X^F_1 (Brand label), X^F_2 (Private Label),and X^F_3 (Other food) and the household receives utility from each of these. Therefore, each household must allocate its budget across these three choices.

And $U^F [X^F (h)] = U^F (X^F_1, X^F_2, X^F_3,)$

The household's utility is a function of many things (i.e., food, non-food and leisure). The utility function is assumed to be separable in those items and therefore it is possible to assume that the household maximizes the utility received from food over choices of brand, private label and other food items, subject to a total food budget. The objective is to maximize the household utility derived from brand, private label and other food items subject to the household food budget constraint. Income allocated to food expenditure is totally consumed on brand, private label and other food items. All of these variables have prices.

3) Max. $U^F [X^F_1 (h), X^F_2 (h), X^F_3 (h); I^F]$

s.t. $Y^F = \sum P^F_i X^F_i (h)$

Income = $P_B X^F_1 + P_{PL} X^F_2 + P_O X^F_3$

Where:

Y^F = Food Budget

P^F = Price of food

i = B, PL, O,

$I^F = f(t, A)$

t = time

t = f (w)

w = Wage rate

(Opportunity cost of leisure)

A = Advertising

Utility is enhanced by Information (I^F) which is exogenous or part of the household's endowment. Information is a function of advertising and time. Time is the amount of time (t) used by the household to obtain the food attributes (shopping around). Advertising (A) is provided by the food industry through advertising or other promotional activities. If the cost of advertising is known and the cost of wages are known then Information is also known. If Information is known, then the optimal value of brand (B), private label (PL), other food products (O) are known. Additionally, time spent shopping around depends on wages/price of leisure and other demographics. Information, which allows the consumer to make a better choice (presumably less expensive), will improve the overall welfare of the household.

This results in the following **Demand Functions:**

4) $X^F_i(h) = X_i [Y^F, P^F, \omega (h); A]$

Define:

P^F_1 = Price of Brand product

P^F_2 = Price of Private label product

Y^F = Food Expenditure deflated by Price Index

ω = Demographics

A = Advertising

Since the demand function is homogenous of degree zero in prices, they can be normalized so that demand becomes a function of the relative prices of brand and private label

varieties and real food budget, along with relative advertising, and other demographics. The prices for brand and private label goods are observed prices, while actual food expenditure for the household deflated by a price index is used for all other food goods.

→ Private Label expenditure share on a food item is:

$S^X_2 = X_2 /(X_1 + X_2)$

while the Brand expenditure share on a food item = $1- S^X_2$

Linearize the relationship of $S^X_2 = X_2 /(X_1 + X_2)$ by taking the natural log:

→ $\ln S^X_2 = \ln X_2 – \ln (X_1 + X_2)$

As a result, you can express the demand function from eq. 4 in terms of shares:

$$S^F_1 = S^F_1 (Y^F, P^F, \omega; A)$$

$$S^F_2 = S^F_2 (Y^F, P^F, \omega; A)$$

This will produce the following **linear regression equation** to be estimated:

$S^F_i = f (Y^F, P^F_i, \omega; A_{Fi}) + \varepsilon_{ti}$ Let: i = 1, 2, 3

1 = Brand

2 = Private Label

3 = Other Food

Since demand is unlikely to be exactly linear, an error term is needed to represent measurement error, misspecification, omitted variables, etc. The regression analysis will reflect 1) the effects of advertising and asymmetric information 2) the difference between brand expenditure shares and private label expenditure shares as well as transaction prices and 3) socioeconomic differences.

What is presented below is the empirical model and quantitative results obtained which adhere to methodology II. This methodology gives partially satisfactory results, at this point in time, and requires further refinement and research.

The research hypotheses will be tested using Tobit regression analysis. The model used has a limited dependent variable because, the consumer is faced with a decision of whether to consume or not. Because the dependent variable is a share, its value can be zero or one or some

fraction between zero and one. When many observations take on a single value (0 or 1) and the remaining observations follow the usual characteristics of a continuous variable (dollars spent) and the resulting dependent variable is part qualitative (to buy or not to buy) and part quantitative (amount bought), OLS estimates are biased.

The estimation results are based on Model II which was determined to better represent the current two-worker household that is faced with increased time constraints.

VARIABLES USED IN THE EMPIRICAL ANALYSIS

As mentioned earlier, the dependent variable is the share of private label food expenditure in each specific food category. The independent variables are REL_PRICE which is the relative prices for private label and brand food items, INCOME is the median income level in each of the 16 income categories. REL_EXP is the total food expenditure on each food category for the household deflated by the price index. In order to gather insight into the variability of the dependent variable, explanatory variables were chosen. Dummy variables were created for race of BLACK, ORIENTAL, HISPANIC, and OTHER, with white being omitted. Dummy variables were also created for census region of CENTRAL, SOUTH and WEST with East being omitted. WH_REL_ADV, BL_REL_ADV, OR_REL_ADV, OTH_REL_ADV and HSP_REL_ADV are interactive variables created by multiplying each race category by the relative advertising (private label/brand) variable. Life stage variables are YOUNG SINGLES, CHILDLESS YOUNG COUPLES, NEW FAMILIES, MATURING FAMILIES, ESTABLISHED FAMILIES, MID-AGE CHILDLESS COUPLES, EMPTY NESTERS and OLDERS SINGLES, with Mid-Age Singles being omitted.

EMPIRICAL MODEL

In order to use tobit analysis, the data must first be transformed so that each household has one record for the year of observation for each food category. In this case, the data was weekly household food purchases that were then aggregated annually.

SHARE = f (REL_PRICE REL_EXP INCOME BLACK ORIENTAL OTHER HISPANIC WH_REL_ADV BL_REL_ADV OR_REL_ADV OTH_REL_ADV HSP_REL_ADV CENTRAL SOUTH WEST YOUNGSINGLES CHILDLESSYOUNGCOUPLES NEWFAMILIES MATURINGFAMILIES ESTABLISHEDFAMILIES MIDAGECHILDLESSCOUPLES EMPTYNESTERS OLDERSINGLES)

REGRESSION RESULTS:

Overall negative coefficients were found for relative prices in Yogurt, Cottage Cheese, Frozen Poultry and Cereal. The remaining food categories had positive coefficients. The effect of the relative price on the private label budget share will be positive if the elasticity of substitution between private label and brand is <1 in absolute value (inelastic) and will be negative if the elasticity of substitution is >1 in absolute value (elastic).

The Food expenditure variable consisted of the total food expenditure for the household on each food category, deflated by the price index. Results produced negative coefficients across all food categories except Frozen Poultry, Frozen Seafood, Lunch Meats, Canned and Bottled Fruit and Cereal. They were significant in all cases except Yogurt, Frozen Poultry, Dinner Sausage, Tomato products, Frozen Plain Vegetables and Cereal. This indicates that people with more money to spend on the food category will purchase smaller shares of private label.

The coefficient on the Income variable was consistently negative and was significant in all categories except Frozen Poultry indicating greater brand purchase with affluence.

An interactive variable was created between relative advertising and each of the race categories, White, Black, Oriental, Other and Hispanic.

WHITES: Coefficients were positive in all food categories except for frozen plain vegetables. They were significant in all categories except for frozen orange juice and pickles relish and olives. This indicates that advertising matters significantly in all these categories.

BLACKS: Coefficients were positive in all food categories except frozen plain vegetables, frozen orange juice and pickles. Advertising was significant in all categories except frozen seafood, frozen orange juice and pickles indicating that advertising mattered in the purchase of private label goods in all food categories except the above mentioned.

ORIENTALS: Coefficients were consistently positive except for frozen plain vegetables. Coefficients were significant in all categories except for pickles.

OTHER: Positive coefficients were found except for frozen plain vegetables and pickles. Advertising was significant in ice cream, cakes and pies, frozen poultry, dinner sausage, lunchmeats, tomato products, frozen vegetables, soups and cereal.

HISPANICS: Positive coefficients were found on all categories except frozen vegetables and frozen orange juice. All food categories were significant except for bread, frozen seafood, frozen orange juice and pickles.

Overall, the relative advertising results were positive and very significant for all races and food categories except for a few minor exceptions. These results indicate that advertising for both private label and brand food products matters significantly.

Next, Race was examined in the various broad food categories of Dairy, Bakery Meats, Vegetables and Other.

DAIRY: Relative to whites, blacks spend between 17-20% less on Dairy products than do whites. Orientals spend between 74 to 77% less on dairy than whites. Hispanics spend between 10-33% less on dairy than do whites. The Other races spend 25% less on ice cream than do whites.

BAKERY: Relative to whites, blacks spend 8% less on bread. Orientals spend 2 ½ % less on cakes and pies. Hispanics spend 6% less on bread than whites

MEATS: Relative to whites, blacks spend between 20-42% less on dinner sausage and lunch meats and over 1 1/3 % less on frozen poultry. Orientals spend between 58% less on lunch meats and over double less for frozen poultry than do whites. Hispanics spend between 11-85% less on meats than whites except for frozen poultry which is almost double less than whites. The Other races spend more than 2 ½% less on frozen poultry than do whites.

VEGETABLES: Relative to whites, blacks spend 12% less on canned and bottled fruits than do whites. Orientals spend between 43–46% less on tomato products and canned/bottled fruit and over 1 1/6% less on frozen orange juice than do whites. There were no significant results for Hispanics or the Other race category for fruits and vegetables.

OTHER: Relative to whites, blacks spend 7% less on canned soups and 23% less on cereal but 10% more on pickles and relish. Orientals spend 71% less on cereal than do whites. Hispanics spend 10% less on canned soups than do whites. There were no significant results for the Other races category in this food section.

For each of the *race* categories, the largest share of private label food purchases was as follows:

WHITES:		
	Cottage Cheese	- 64%
	Pies and Cakes	- 70%
	Frozen Poultry	- 67%
	Frozen Seafood	- 24%
	Frozen Plain Vegetables	- 94%
	Frozen Orange Juice	- 51%
	Canned Soup	- 13%
	Cereal (RTE)	- 26%
BLACKS:		
	Frozen Plain Vegetables	- 94%
ORIENTALS:		
	Fresh Bread/Rolls	- 69%
	Tomato Products	- 62%
OTHER:		
	Yogurt, Rfg.	- 25%
	Dinner Sausage	- 19%
	Lunch Meats	- 35%
	Pickles, Relish & Olives	- 57%
HISPANICS:		
	Ice Cream, Sherbet & Frozen Yogurt	- 47%
	Yogurt, Rfg.	- 25%
	Lunch Meats	- 35%
	Canned/bottled Fruit	- 61%
	Pickles, Relish & Olives	- 57%

For each of the *Life Stages* categories the largest share of private label food purchases were as follows:

YOUNG SINGLES:

Cottage Cheese	-	83%
Pies and Cakes	-	78%
Frozen Seafood	-	54%
Dinner Sausage	-	15%

CHILDLESS YOUNG COUPLES:

Frozen Plain Vegetables	-	95%

NEW FAMILIES:

Tomato Products	-	60%
Canned/Bottled Fruit	-	66%

MATURE FAMILIES:

Tomato Products	-	60%

ESTABLISHED FAMILIES:

Ice Cream, Sherbet & Frozen Yogurt	-	46%
Fresh Bread & Rolls	-	68%

MID-AGE SINGLES:

Frozen Poultry	-	71%
Frozen Plain Vegetables	-	95%
Frozen Orange Juice	-	58%

OLDER SINGLES:

Yogurt, Rfg.	-	28%
Lunch Meats	-	30%
Cereal (RTE)	-	32%
Pickles, Relish & Olives	-	57%

Mid-Age Childless Couples and Empty Nesters did not have the largest share of private label food purchases in any food category.

For each of the *Region* categories, the largest share of private label food purchases was as follows.

EAST:

Frozen Seafood	-	33%
Frozen Plain Vegetables	-	94%
Frozen Orange Juice	-	55%

CENTRAL:

Cottage Cheese	-	74%
Fresh Bread/Rolls	-	69%
Frozen Plain Vegetables	-	94%
Can/Bottled Fruit	-	65%
Canned Soup	-	15%
Cereal	-	27%
Pickles, Relish, Olives	-	57%

SOUTH:

Dinner Sausage	-	14%
Luncheon Meats	-	30%

WEST:

Ice Cream/Sherbet/Fz Yog.	-	60%
Yogurt, Rfg.	-	30%
Pies/Cakes	-	89%
Frozen Poultry	-	74%
Tomato Products	-	60%
Frozen Plain Vegetables	-	94%

The *education* categories were recoded as follows:

Female 1 = No Female Head

Female 2 = Some High School

Female 3 = Some College

Female 4 = College graduate or more

For each of the above education categories, the largest shares of private label food purchases were as follows:

FEMALE 1:

Yogurt, Rfg.	=	30%
Dinner Sausage	=	10%
Lunch Meats	=	31%
Frozen Plain Vegetables	=	96%
Frozen Orange Juice	=	52%
Cold Cereal	=	33%
Pickles,Relish,Olives	=	56%

FEMALE 2:

Ice Cream/Sherbet/FzYog.	=	49%
Cottage Cheese	=	74%
Fresh Bread	=	62%
Frozen Poultry	=	71%
Tomato Products	=	73%
Can/Bottled Fruit	=	65%
Canned Soup	=	16%

FEMALE 3:

Frozen Orange Juice	=	52%

FEMALE 4:

Pies/Cakes	=	71%
Frozen Seafood	=	25%

The *education* categories were recoded as follows:

Male 1 = No Male Head

Male 2 = Some High School

Male 3 = Some College

Male 4 = College graduate or more

For each of the above education categories, the largest shares of private label food purchases were as follows:

MALE 1:

Frozen Seafood	-	27%
Pickles, relish, olives	-	55%

MALE 2:

Ice Cream/Sherbet/FzYog.	-	52%
Yogurt, Rfg.	-	24%
Cottage Cheese	-	69%
Fresh bread/Rolls	-	64%
Dinner Sausage	-	13%
Lunch Meats	-	39%
Tomato Products	-	67%
Frozen Plain Vegetables	-	95%
Can/bottled Fruit	-	66%
Frozen Orange Juice	-	63%
Canned Soup	-	15%
Cold Cereal	-	29%

MALE 3:

Yogurt, Rfg.	-	24%
Pies/Cakes	-	72%
Frozen Poultry	-	67%

MALE 4:

There were no categories with the highest private label share.

ELASTICITIES:

RELATIVE PRICE: Relative price elasticity's are positive and less than one for all food categories except yogurt, frozen poultry and cereal. This indicates that if the price of private label goods goes up (down) faster than the price of brand products that households will make only limited substitutions toward or away from brands such that the budget share devoted to private label will go up or down. Intuitively, inelastic demand implies a price increase leading to expenditure increases.

FOOD EXPENDITURE: Elasticities on food expenditure for 'All Other Goods' are negative in all food categories except frozen poultry, frozen seafood, lunchmeats, can/bottled fruit and cereal where they are positive but less than 1. Negative elasticities on food expenditure indicate that as one spends more on bread or cereal, holding income constant, one might spend relatively less on private label. This is somewhat ambiguous and much more of an empirical question. In part it might reflect an element of increasing affluence that is just not captured in the income term.

INCOME: Income elasticities are negative for all food categories, indicating that with greater affluence there is greater brand purchase relative to private label purchases. (i.e., one could find greater purchase of both brand and private label, but would expect that there would be relatively less private label as income rises.

RELATIVE ADVERTISING: WHITE: Elasticities for relative advertising for the white race are positive and significant for most food categories except frozen plain vegetables which is negative and significant. This would indicate that relative advertising really doesn't affect the purchase of frozen plain vegetables. However, in all food categories where the elasticities are

positive and significant, when relative advertising is up then expenditure on private label goods will be up relative to expenditure on brand products.

BLACK: Elasticities for relative advertising for the black race are positive and significant for most food categories except for frozen vegetables, which is negative and significant.

ORIENTAL: Elasticities for relative advertising for the Oriental race are positive and significant in all food categories except for frozen vegetables which is negative and significant.

OTHER: Elasticities for relative advertising for the Other races are positive and significant for ice cream, pies and cakes, frozen poultry, dinner sausage, lunch meats, tomato products, soups and cereals. Elasticity was negative and significant for frozen plain vegetables.

HISPANIC: Lastly elasticities for relative advertising for the Hispanic race are positive and significant for most food categories except for frozen plain vegetables which is again negative and significant.

Positive elasticities mean that when relative advertising is up then expenditures on private label goods are up relative to expenditures on brand products.

Lastly, household income was divided into three groups, Low ($0 - $29,999), Middle ($30,000-$69,999) and High ($70,000-$100,000+). The data on the private label share is consistently over 50% private label purchases for the Middle income group, and the Low and High income groups are between 18-23% and 23 – 27% respectively. The following graph illustrates these percentages.

CONCLUSION:

Since this paper is a work in progress, the findings are tentative and subject to further refinement. That being said, we find overwhelmingly that advertising had a positive effect and was significant in almost every food category except a few indicating the significance of both private label and brand advertising. Additionally, elasticities for relative advertising by race categories are predominantly positive.