Title
Bridging the Two Traditions in Scanner-Data Research

Author
Cooper, Lee

Publication Date
1991-07-01

THE JOHN E. ANDERSON GRADUATE
SCHOOL OF MANAGEMENT AT UCLA

BRIDGING THE TWO TRADITIONS IN SCANNER-DATA RESEARCH

Lee G. Cooper

July 1991

Marketing Studies Center

John E. Anderson Graduate School of Management, UCLA

Working Paper No. 212

Bridging the Two Traditions in Scanner-Data Research

Lee G. Cooper
Anderson Graduate School of Management, UCLA

Center for Marketing Studies Working Paper No. 212
Anderson Graduate School of Management, UCLA
July, 1991

I would like to express my appreciation to M.J. Snyder for her assistance with the computations, and to A.C. Nielsen for the data used in this paper.

Abstract

This paper develops a method to bridge the two main streams of scanner-data research: individual choice analysis based on consumer panels, and market-response analysis based on aggregate (at least store-level) data. The advent of single-source data demands integrated analysis of both panel and store records. The approach proceeds by isolating the sources of information in the panel data that are missing from the record of the aggregate causal environment, and seeking a practical method for completing the picture. The regression-based method is illustrated using Nielsen single-source data on the catsup category in which the information on manufacturers' coupons from the panel data are missing from the causal data reflecting the in-store environment. The results show that diagnostically richer models can be obtained without a loss of forecast accuracy.

Introduction

The last decade has witnessed a tremendous explosion of research on how to use scanner data to learn about consumer and market response. The intense efforts at developing methods to deal with scanner data have fallen into two separate camps. On one hand, there has been extensive development of discrete choice (mainly logit-based) models and methods for understanding individual-level choice processes (Bawa and Ghosh 1991; Bawa and Shoemaker 1987; Bucklin and Lattin 1991; Carpenter and Lehmann 1985; Currim, Meyer, and Le 1988; Currim and Schneider 1991; Elrod 1988; Fader and McAllister 1990; Grover and Srinivasan 1989; Guadagni and Little 1983; Gupta 1988 and 1991; Jain and Vilcassim 1991; Kalwani et al. 1990; Kamakura and Russell 1989; Kannan and Wright 1991; Krishnamurthi and Raj 1991; Lattin and Bucklin 1989; Meyer and Cooper 1987; Moore and Winer 1987; Neslin 1990; Neslin, Henderson, and Quelch 1985; Pedrick and Zufryden 1991; Ramaswamy and DeSarbo 1990; Tellis 1988; Vilcassim and Jain 1991; Wheat and Morrison 1990a; Winer 1986 and 1989; and Zufryden 1984 and 1987). On the other hand, there have been substantial efforts to develop market-response models and methods that address both the vast size and detail of store-tracking data (Abraham and Lodish 1987; Allenby 1989 and 1990; Bemmaor and Mouchoux 1991; Blattberg and Wisniewski 1989; Bolton 1989; Bultez and Naert 1988; Cooper 1988; Cooper and Nakanishi 1988; Doyle and Saunders 1990; Kumar and Leone 1988; Russell and Bolton 1988; Shugan 1987; Vilcassim 1989; and Walters and MacKenzie 1988). The few efforts that have tried to use both panel and store data have been like Wheat and Morrison (1990b) who use the percentage of times deals occur (computed from store-level data) in there panel-data investigation of purchase-timing models, or like Pedrick and Zufryden (1991) where store-level causal variables are used in panel-level modeling efforts. In spite of the common modeling frameworks (i.e. discrete multinomial-logit [MNL] models and aggregate MNL models) used in some of these undertakings referenced above, no work has focussed on bridging between these two research streams.

There are very pragmatic reasons for attempting to merge the two traditions. Consumer panels contain information that is missing from the store-tracking data. Panel data tell for example about brand-usage frequency, about redemption of manufacturers' coupons, and about television-advertising exposures -- information that is simply not available in store-tracking databases. If such data are not incorporated into market-response models, the diagnostic value of the models is necessarily impaired, and the forecasting ability of the models may also be affected.

This paper proceeds by:

1. developing a segmentation scheme that divides the consumer panel into mutually exclusive and exhaustive groups based on usage frequency,

2. aggregating panel variables to the segment level,

3. estimating the store-level sales for each segment (for each brand),

4. calibrating an asymmetric market-share model for the store-level sales of each segment,

5. combining segment-level market-share models into a single forecast,

6. calibrating and forecasting from an aggregate market-share models, and

7. comparing forecasts and diagnostics.

Methods and Results

1. Developing a Segmentation Scheme

While there many useful bases for segmenting consumer panels (cf. Wind 1978), one of the most fruitful areas for investigation for market-share models concerns differences in usage frequency. As pointed out in Cooper and Nakanishi (1988, 43-44 and 89-93) parameters from aggregate market-share models will not reflect the individual choice processes if the heavy users in a category systematically purchase different brands than the light users. Readers are referred to Cooper and Nakanishi (1988) for an in-depth discussion of this issue.

Without a consumer panel to go along with store-tracking data a model developer would not know if there was a correlation between brand choice and usage frequency. But with single-source databases come both the knowledge of the problem (if it exists) and the opportunity to address it through segmentation.

The problematic correlation between brand choice and purchase frequency can be detected with a simple χ^2 test on expected purchase frequencies. The theoretical market shares can be estimated from the long-term sales data in the store-tracking data or the panel (using the overall panel would eliminate any rejections due to the panel failing to track the store). The market shares times the total number of purchases in each panel segment provides the expected purchase frequency for each brand, while the observed purchase frequency is readily obtainable for each panel segment. The great size of the panels may make it too likely that trivial differences will be confirmed as statistically significant. Only experience over diverse categories will tell if extremely conservative testing procedure (i.e. extreme type-one error rates) should be invoked.

But before we can test for the correlation between brand choice and purchase frequency, we have to decide who are the heavy users and who are the light users. Cooper (1989a and b) split Nielsen's single-source panels into heavy- and light-user segments in 19 different ways (depending on different definitions of how heavy the heavy users had to be). All but the most extreme definitions lead to groups that were significantly discriminable on the basis of actionable variables such as income, number of members in the family, number of years at the same residence, number of TVs in the household, hours per week worked by the male head-of-household (HOH), highest degree obtained by male HOH, and number of hours per week worked by female HOH. Given this latitude, we decided to define heavy users as those panel households that purchased the top half of the total volume in the category. We know that in general it is tougher to predict more extreme events. This specification of segment membership means we will have ultimately to forecast less extreme volumes than with any other split.

In the catsup category 50% of volume is purchased by the top 20% of the households. In the yogurt category the top 10% of the households purchased 50% of the volume. We expect the concentration to vary from one category to another. Further, in any historical period the record of which households have purchased the most in a category is an imperfect indicator of who are the *real* heavy users. But since Schmittlein et al. (1991) have shown that the longer the term on which manifest concentration is estimated the closer the manifest concentration is to the *true* concentration, our approach is to use all available panel data to classify households into heavy- and light-users segments.

Using the approximately 50-50 split of volume to define segments in the catsup market, the market and segment shares for Heinz, Hunts, Del Monte, and a combination of the private-label and generic catsups appear in Table 1. The χ^2 for the independence of segment and share is 372 with 3 degrees of freedom. This extreme significance is, as speculated, due mainly to the large size of the panel, since the Φ coefficient is only .019 for this same cross tabulation.

Insert Table 1 about here.

2. Aggregate Panel Variables to the Segment Level.

The kinds of variables that would be useful to incorporate from the panel file into a market-response analysis are relatively obvious. We aren't interested in the number of toasters or dogs in the household not only because we have no hypotheses as to why certain brands would differentially benefit from the various levels, but more fundamentally because these kinds of variables do not vary across weeks to any appreciable degree. They would act like segment-specific constants or intercepts.

The obvious candidates for inclusion are those variables we wished were in the store-tracking database to begin with. Highest on this list would be information on manufacturers-coupon redemptions. While store-coupon redemptions have been reported, and used in market-share analysis (cf. Cooper and Nakanishi 1988), manufacturer-coupon redemptions have not. It may be that inclusion of this information will alter our understanding of price or store-coupon sensitivity.

A less obvious, but no less important, variable is commercial exposures. GRPs or TRPs could be recorded in a store-tracking database. But it is not obvious how commercial exposures would be routinely included in anything but a panel database.

In the catsup-category illustration that we develop in the next sections both manufacturer coupons and commercial exposures we initially aggregated for each segment. But in the catsup category too few brands engaged in TV advertising to obtain meaningful results. Thus only manufacturer coupons are used in this first application.

3. Estimating the Store-Level Sales for Each Segment (for Each Brand).

There are three basic assumptions that drive our method for estimating the store-level sales that correspond to each panel segment:

- The long-term average panel sales for a brand has an influence on contemporaneous store sales.

This implies that even if there are no sales in the panel segment for a particular brand in a given week that there still might be sales in the corresponding store segment. The higher the historic average sales the more likely it is that there will be sales in the store segment *ceteris paribus*, even in absence of sales in the panel segment.

- The contemporaneous panels sales for a brand has an influence on contemporaneous store sales.

Of course we expect the contemporaneous brand sales in a panel segment to be positively related to brand sales for the corresponding store segment. Each panel segment can have its unique sensitivity, i.e. the regression weight that relates contemporaneous panel sales for the heavy panel segment to contemporaneous store sales for the heavy store segment may be different from the regression weight that relates the corresponding sales figure for the light segments in the panel and in the store.

- The sensitivity of contemporaneous store sales to long-term panel sales is the same within brand across segments.

There is no reason that would force us to speculate that long-term historic sales for a brand affects one segment differently than another. We do of course expect different levels of average sales across segments, but this restriction is much more specific than that. It merely says that the sensitivity of contemporaneous sales to historic sales for a brand is the same for each panel segment. This amounts to a necessary restriction on the intercept for the regression model.

If we let s_{jt} be the store sales for brand j in period t, let p_{Ljt} and p_{Hjt} be the panel sales for the light-users segment and the heavy-users segment, respectively, and let s_{Ljt} and s_{Hjt} be the store sales we seek for the light segment and the heavy segment respectively, then these three assumptions can be represented in the following theoretical relations:

(1)

$$s_{Ljt} = \alpha_j \bar{p}_{Lj.} + \beta_{Lj} \, p_{Ljt}$$

(2)

$$s_{Hjt} = \alpha_j \overline{p}_{Hj.} + \beta_{Hj}\, p_{Hjt}$$

Estimates for α_j, β_{Lj}, and β_{Hj} may be obtained from the multiple-regression equation:

(3)

$$s_{jt} = \hat{\alpha}_j (\overline{p}_{Lj.} + \overline{p}_{Hj.}) + \hat{\beta}_{Lj}\, p_{Ljt} + \hat{\beta}_{Hj}\, p_{Hjt} + e_{jt}$$

Rather than merely estimate store segment sales from the theoretical relations in (1) and (2) using the parameter estimates in (3), advantage should be taken of the fact that these store segment estimate must sum up to a known total in each store. This can be achieved if we let w_{Ljt} and w_{Hjt} be weights such that:

(4)

$$w_{Ljt} + w_{Hjt} = 1$$

(5)

$$w_{Ljt} \times s_{jt} = s_{Ljt}$$

(6)

$$w_{Hjt} \times s_{jt} = s_{Hjt}$$

and obtain estimate of these desired weight by post hoc normalization of the sales estimates based on (1) and (2).

(7)

$$\hat{w}_{Ljt} = \hat{s}_{Ljt} / \hat{s}_{jt}$$

(8)

$$\hat{w}_{Hjt} = \hat{s}_{Hjt} / \hat{s}_{jt}$$

Plugging the weights estimated in (7) and (8) into (5) and (6) provides store-segment sales estimates that sum to the known total sales in the store.

This procedure was employed using the Nielsen single-source database tracking catsup sales in Sioux Falls, S.D. for 138 weeks. Table 2 presents the regression-model results for Heinz, Hunts, Del Monte, Control (Private-Label) Brands, generic catsup, and aggregation of all other branded catsups, and an aggregation of all catsup called "ALLBRAND" in the analysis. The top 20% of users in this market purchased approximately 53% of the catsup volume. These households constituted the heavy-users segment in the panel.

The models all have very high R-Square values, but one must be careful to recognize that since the normal intercept is suppressed, R-Square is more like a congruence coefficient that the coefficient of multiple determination we expect. All of these models are highly significant when compared to the simulation results reported by Korth and Tucker (1975). All of the parameters are also statistically significant. These tables also report the split of the baseline volume into the average historical weekly volume (in ounces) for each panel segment.

Insert Table 2 about here.

4. Calibrating an Asymmetric Market-Share Model for the Store-Level Sales of Each Segment.

Remember we have created a sales stream that we believe corresponds to the part of total brand sales that was purchased by the light-users segment, and a corresponding sales stream for the heavy-users segment. Except for the aggregated information from the panel file (which in this illustration concerned only the redemption of manufacturers coupons), the causal environment encountered by these segments is the same. So we calibrated an asymmetric MCI model using the procedures outlined in Cooper and Nakanishi (1988, p. 168). Basically these procedures estimate all differential effects and brand-specific intercepts, search among the residuals for potentially significant cross effects, and finally re-estimate all effects using weighted least squares. Given the tiny shares of the aggregate representing all other branded catsups, this aggregate was dropped from the analysis. The private-label and generic catsups were combined into an aggregate simply called Private Labels. Each brand was represented by a brand-specific intercept (Int) and a differential effect for line ads (Ad-L), major ads (Ad-M), end-of-aisle display (D-EA), front-aisle display (D-FA), in-aisle display (D-In), an aggregate of all other displays (D-AO), percent of brand volume sold on store coupon (Cp-S), and price (Price). Manufacturers coupon (Cp-M) was represented as the percent of panel-segment sales on which a manufacturers coupon was redeemed. All variables were represented as *exp(z-scores)* (cf. Cooper and Nakanishi 1988, pp. 68-78).

Table 3 and 4 report the parameters estimated for the market-share models for the Light and Heavy Segments, respectively. Both models fit very well in calibration (R-Squares of .824 and .829 respectively). One would hope that this would be the case since the cross-competitive effects are selected so that they are very likely to be significant in the calibration data set (approximately two-thirds of the store weeks are used for calibration). The models also fit quite well in cross calibration in which the specification developed in the calibration dataset is applied to the remain one-third of the store weeks). Here the R-Squares were .803 and .814 respectively. And finally the models also cross validated quite well. The cross validation used the parameter values form the calibration dataset to forecast market shares for the untouched data in the final one-third of the store weeks. Here we form a single composite for each segment and perform a simple regression on the appropriately transformed (log-centered) dependent measure. The R-Squares are .756 and .766 respectively for models with one predictor and 1607 degrees of freedom. The comparison to the aggregate model is discussed in Section 7. The cross effects are described verbally, by comparison with what would be expected from a symmetric market-share model (cf. Bell, Keeney, and Little 1975). In symmetric market-share models, when one brand loses share the other brands are represented as if they gain share strictly in proportion to their prior market shares. Further interpretation is deferred to section 7 where diagnostics are also compared.

Insert Tables 3 and 4 about here.

5. Combining Segment-Level Market-Share Models into a Single Forecast.

Combining the segment-level forecasts mentioned in the previous section into a single market-share forecast is relatively straightforward since we already have the weights from equation (7) and (8). Combined forecast of market share for each brand merely applies these weights to the estimated brand shares from the segment models.

The variance accounted for (VAF) in the forecast period of the combined forecast is over 71%. This is a strong indication of very good forecasting ability. The root mean squared error (RMSE) of .128 might seem high, but we must remember that there are tremendous swings in market share when looked at on a week-by-week and store-by-store basis.

6. Calibrating and Forecasting from an Aggregate Market-Share Models,

Excluding the manufacturers coupons, the procedures for specification of the aggregate market-share model were the same as those for the segment-level models. The R-Square in calibration was .845, in cross calibration was .816 and in cross validation was .772 -- just slightly higher in all cases than the corresponding values for the segment-level models.

The parameter values are reported in Table 5. Here too the cross effects are described verbally, in relation to what we expect from a symmetric market-share model. Further interpretation of the parameters is postponed to the next section.

Insert Table 5 about here.

7. Comparing Forecasts and Diagnostics.

Table 6 compares the summary statistics on model calibration, cross calibration, cross validation, and forecasts. The similarities are remarkable. Although a slight edge might seem to go to the aggregate model (there are no statistical tests to compare these R-Square values), when it comes to forecasting the models are identical to three decimal places.

Insert Table 6 about here.

The *condition index* is also reported for the models in each of the calibration datasets. This number is the ratio of the largest singular value (square root of the eigenvalue) to the smallest singular value in the sum-of-squares-and-crossproducts (SSCP) matrix for the (reduced-form) regression model. This index is discussed by Belsley, Kuh, and Welsch 1980) as an indicator of the degree of collinearity in the regression system. These authors indicate that condition indices over 100 can cause "... substantial variance inflation and great potential harm to regression estimate" (p. 153). Indices from 21 to 35 are moderate at worst and give further evidence that asymmetric market-share models that use exp(z-scores) (cf. Cooper and Nakanishi 1988, pp. 141-3) are exempt form the warnings concerning collinearity first given by Bultez and Naert (1975).

We may begin the comparison of the diagnostic value of this approach by looking at the parameter values listed in Table 7. This table shows only the differential-effect parameters for the three models calibrated earlier. There are no statistical tests for the equality of parameters across equations of this sort. What we are undertaking is not a statistical comparison, but rather a comparison of how differently these models would be interpreted by managers having to develop brand plans in a competitive environment.

Insert Table 7 about here.

For convenience the parameters have been grouped into four classes. Those with no border are relatively stable across analyses. Note that 28 of the 40 groups (70%) fall into this stable class, which is reassuring in many ways. The brand-specific intercepts, line-ad parameters, front-aisle display parameters, and store-coupon effects all seem stable within brands across analyses.

The solid, dark lines highlight parameter groups with major differences in levels (although no changes in the pattern of statistical significance across analyses). For Del Monte the aggregate price parameter is -.57, but the introduction of the manufacturers coupon and segmentation leads to a heavy-user segment that is very price sensitive ($\beta_{Hj\ Price}$ = -1.13) and somewhat coupon sensitive ($\beta_{Hj\ Coupon}$ = .26), while the light-user segment is much less price sensitive ($\beta_{Lj\ Price}$ = -.61) and more opportunistic users of manufacturers coupons ($\beta_{Lj\ Coupon}$ = .51). For Private-Label brands the manufacturers' coupon effect is not significant in either segment (as should be the case), but the segmentation still leads to major differences in the price parameter -- with the heavy-user segment being much more price sensitive than the light-user segment.

The dotted lines highlight groups in which the analyses reveal differences in the patterns of significance across analyses. In the aggregate analysis of Heinz there is a significant price effect. But segment-level analysis indicates that the heavy-user segment is not price sensitive, but is sensitive to manufacturers coupons, while the light-user segment is somewhat price sensitive and not effected by manufacturers coupons. For Hunts we see that both segments are somewhat sensitive to the presence of manufacturers coupons, but that all of the price sensitivity that appears in the aggregate is due to the sensitivity of the heavy-users segment.

Particularly for Heinz (the dominant brand in the category) we see diagnostically different patterns comparing the segment analyses and the aggregate analysis. Major ads and other displays appear effective in the aggregate but have no significant effect in either segment. The end-of-aisle displays that seem effective in the aggregate analyses influence the heavy users but not the light users. All of these difference lead us to believe that segment-based analyses provide a diagnostically richer picture of this brand and of the market -- without a sacrifice in forecast accuracy.

Conclusion

The goal of this effort was to propose and illustrate a method for bridging the too long separate traditions of individual-choice modeling and market-response modeling. We have shown that some kinds of information in panel databases can be used to segment otherwise aggregate market-response models, and other kinds of information can be integrated into these segment models to provide a diagnostically richer representation of market and competitive influences. And we have shown that these diagnostically rich, asymmetric market-share models can be estimated without the fear of collinearity.

This was a modest effort involving segmentation by usage frequency and the integration of a single variable from the panel database. But the underlying methods are so simple and robust that more venturesome applications seem readily doable.

References

Abraham, Magid M. and Leonard M. Lodish (1987), "PROMOTER: An Automated Promotion Evaluation System," *Marketing Science*, 6, 2 (Spring), 101-23.

Allenby, Greg M. (1989), " A Unified Approach to Identifying, Estimating and Testing Demand Structures with Aggregate Scanner Data," *Marketing Science*, 8, 3 (Summer), 265-80.

______________ (1990), " Hypothesis Testing with Scanner Data: The Advantage of Baysesian Methods," *Journal of Marketing Research*, XXVII (November), 379-89.

Bawa, Kapil and Avijit Ghosh (1991), "The Covariates of Regularity in Purchase Timing," *Marketing Letters*, 2, 2 (April), 147-57.

___________ and Robert W. Shoemaker (1987), "The Effects of Direct Mail Coupon on Brand Choice Behavior," *Journal of Marketing Research*, XXIV (November), 370-6.

Bell, David E., Ralph L. Keeney and John D. C. Little (1975), "A Market Share Theorem," *Journal of Marketing Research*, XII (May), 136-41.

Belsley, David A., Edwin Kuh, and Roy E. Welsch (1980), *Regression Diagnostics: Identifying Influential Data and Sources of Collinearity*, New York: John Wiley and Sons.

Bemmaor, Albert C. and Dominique Mouchoux (1991), "Measuring the Short-Term Effect of In-Store Promotion and Retail Advertising on Brand Sales: A Factorial Experiment," *Journal of Marketing Research*, XXVIII (May), 202-14.

Blattberg, Robert C. and Kenneth J. Wisniewski (1989), "Price-Induced Patterns of Competition," *Marketing Science*, 8, 4 (Fall), 291-309.

Bolton, Ruth N. (1989), "The Relationship Between Market Characteristics and Promotional Price Elasticities," *Marketing Science*, 8, 2 (Spring), 153-69.

Bucklin, Randolph E. and James M. Lattin (1991), "A Two-State Model of Purchase Incidence and Brand Choice," *Marketing Science*, 10, 1 (Winter), 24-39.

Bultez, Alain V. and Philippe A. Naert (1975), "Consistent Sum-Constrained Models," *Journal of the American Statistical Association*, 70 (September), 529-35.

____________ and ______________ (1988), "SH.A.R.P.: SHelf Allocation for Retailer's Profit," *Marketing Science*, 7, 3 (Summer), 211-31.

Carpenter, Gregory S. and Donald R. Lehmann (1985), "A Model of Marketing Mix, Brand Switching, and Competition," *Journal of Marketing Research*, XXII (August), 318-29.

Cooper, Lee G. (1988), "Competitive Maps: The Structure Underlying Asymmetric Cross Elasticities," *Management Science*, 34, 6 (June), 707-23.

_____________ (1989a), "Integrating Scanner-Panel and Store-Level Data," Seventh Marketing Science Conference, Duke, March.

_____________ (1989b), "Toward Integrating Scanner-Panel and Store-Level Data," Joint UCLA-USC Marketing Colloquium, UCS, May.

_____________ and Masao Nakanishi (1988), *Market-Share Analysis: Evaluating Competitive Marketing Effectiveness*, Boston: Kluwer Academic Publishers.

Currim, Imran S, Robert J. Meyer, and Nhan T. Le (1988), "Disaggregate Tree-Structured Modeling of Consumer Choice Data," *Journal of Marketing Research*, XXV (August), 253-65.

_____________ and Linda G. Schneider (1991), "A Taxonomy of Consumer Purchase Strategies in a Promotion Intense Environment," *Marketing Science*, 10, 2 (Spring), 91-110.

Doyle, Peter and John Saunders (1990), "Multiproduct Advertising Budgeting," *Marketing Science*, 9, 2 (Spring), 97-113.

Elrod, Terry (1988), "Choice Map: Inferring a Product-Market Map from Panel Data," *Marketing Science*, 7, 1 (Winter), 21-40.

Fader, Peter S. and Leigh McAlister (1990), "An Elimination by Aspects Model of Consumer Response to Promotion Calibrated on UPC Scanner Data," *Journal of Marketing Research*, XXVII (August), 322-32.

Grover, Rajiv and V. Srinivasan (1989), "An Approach for Tracking Within-Segment Shifts in Market Shares," *Journal of Marketing Research*, XXVI (May), 230-6.

Guadagni, Peter M. and John D. C. Little (1983), "A Logit Model of Brand Choice Calibrated on Scanner Data," *Marketing Science*, 2, 3 (Summer), 203-38.

Gupta, Sunil (1988), "Impact of Sales Promotion on When, What, and How Much to Buy," *Journal of Marketing Research*, XXV (November), 342-56.

____________ (1991), "Stochastic Models of Interpurchase Time with Time Dependent Covariates," *Journal of Marketing Research*, XXVIII (February), 1-15.

Jain, Dipak C. and Naufel J. Vilcassim (1991), "Investigating Household Purchase Timing Decisions: A Conditional Hazard Function Approach," *Marketing Science*, 10, 1 (Winter), 1-23.

Kalwani, Manohar U., Chi Kin Yim, Heikke J. Rinne, and Yoshi Sugita (1990), "A Price Expectations Model of Customer Brand Choice," *Journal of Marketing Research*, XXVII (August), 251-62.

Kamakura, Wagner A. and Gary J. Russell (1989), "A Probabilistic Choice Model for Market Segmentation and Elasticity Structure," *Journal of Marketing Research*, XXVI (November), 379-90.

Kannan, P. K. and Gordon P. Wright (1991), "Modeling and Testing Structured Markets: A Nested Logit Approach," *Marketing Science*, 10, 1 (Winter), 58-82.

Krishnamurthi, Laksman and S. P. Raj (1991), "An Empirical Analysis of the Relationship Between Brand Loyalty and Consumer Price Elasticity," *Marketing Science*, 10, 2 (Spring), 172-83.

Korth, Bruce and Ledyard R Tucker (1975), "The Distribution of Chance Coefficients from Simulated Data," *Psychometrika,* 40, 3 (September), 361-72.

Kumar, V. and Robert P. Leone (1988), "Measuring the Effect of Retail Store Promotions on Brand and Store Substitution," *Journal of Marketing Research*, XXV (May), 178-85.

Lattin, James M. and Randolph E. Bucklin (1989), "Reference Effects of Price and Promotion on Brand Choice Behavior," *Journal of Marketing Research*, XXVI (August), 299-310.

Meyer, Robert J and Lee G. Cooper (1987), "A Longitudinal Choice Analysis of Consumer Response to a Product Innovation, " Chapter 20 pp. 424-51 in *Decision Making and Choice Behavior: Behavioural Modelling,* Golledge, R. and H. Timmermans (Editors), London: Croom-Helm.

Moore, William L and Russell S. Winer (1987), "A Panel-Based Method for Merging Joint Space and Market Response Function Estimation," *Marketing Science*, 6, 1 (Winter), 25-42.

Neslin, Scott A. (1990), "A Market Response Model for Coupon Promotions," *Marketing Science*, 9, 2 (Spring), 125-45.

___________, Caroline Henderson, and John Quelch (1985), "Consumer Promotions and the Acceleration of Product Purchase," *Marketing Science*, 4 (Spring), 147-65.

Pedrick, James H. and Fred S. Zufryden (1991), "Evaluating the Impact of Advertising Media Plans: A Model of Consumer Purchase Dynamics Using Single-Source Data," *Marketing Science*, 10, 2 (Spring) 111-30.

Ramaswamy, Venkatram and Wayne S. DeSarbo (1990), "SCULPTRE: A New Methodology for Deriving and Analyzing Hierarchical Product-Market Structure from Panel Data," *Journal of Marketing Research*, XXVII (November), 418-27.

Russell, Gary J. and Ruth N. Bolton (1988), "Implications of Market Structure for Elasticity Structure," *Journal of Marketing Research*, XXV (August), 229-41.

Schmittlein, David C., Lee G. Cooper, and Donald G. Morrison (1991), "Truth in Concentration in the Land of 80/20 Laws," Center for Marketing Studies Working Paper No. 210, Anderson Graduate School of Management, UCLA, July.

Shugan, Steven M. (1987), "Estimating Brand Positioning Maps from Supermarket Scanning Data," *Journal of Marketing Research*, XXIV (February), 1-18.

Tellis, Gerrard J (1988), "Advertising Exposure, Loyalty, and Brand Purchase: A Two-Stage Model of Choice," *Journal of Marketing Research*, XXV (May), 134-44.

Vilcassim, Naufel J. (1989), "Extending the Rotterdam Model to Test Hierarchical Market Structures," *Marketing Science*, 8, 2 (Spring), 181-90.

________________ and Dipak C. Jain (1991), "Modeling Purchase Timing and Brand Switching Behavior Incorporating Explanatory Variables and Unobserved Heterogeneity," *Journal of Marketing Research*, XXVIII (February), 29-41.

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

Wheat, Rita D. and Donald G. Morrison (1990a), "Estimating Purchase Regularity with Two Interpurchase Times," *Journal of Marketing Research*, XXVII (February), 87-93.

____________ and ________________ (1990b), "Assessing Purchase Timing Models: Whether or Not is Preferable to When," *Marketing Science*, 9, 2 (Spring), 162-70.

Wind, Yoram (1978), "Issues and Advances in Segmentation Research," *Journal of Marketing Research*, XV, 317-37.

Winer, Russell S. (1986), "A Reference Price Model of Brand Choice for Frequently Purchased Products," *Journal of Consumer Research*, 13, 2 (September), 250-71.

______________ (1989), "A Multi-Stage Model of Choice Incorporating Reference Prices," *Marketing Letters*, 1, 1 (December), 27-36.

Zufryden, Fred S. (1984), "Modeling Purchase Patterns on the Basis of Incomplete and Biased Consumer Purchase Diary and UPC Panel Data," *International Journal of Research in Marketing*, 1, 199-213.

_______________ (1987), " A Model for Relating Media Exposure to Purchase Incidence Behavior Patterns," *Management Science*, 33, 10 (October), 1253-66.

4

Table 1. Store and Segment Market Shares

Brand	**Store Share**	**Heavy Segment Share**	**Light Segment Share**
Heinz	67.40%	68.79%	68.25%
Hunts	15.80%	15.98%	15.31%
DelMonte	7.60%	8.15%	8.52%
PvtGeneric	9.20%	7.08%	7.91%

5

Table 2. Summary of Regression Analyses - Sioux Falls Panel

Brand	**R-Square**	**F (3,138)**	**Alpha Baseline**	**Beta Light**	**Beta Heavy**	**Average Volume Light**	**Average Volume Heavy**
Heinz	0.976	1827	8.596	15.137	10.562	2403	2784
(t-values)			10.811	5.248	3.429		
Hunts	0.965	1233	8.273	11.786	15.617	539	647
(t-values)			11.98	5.837	8.761		
DelMonte	0.923	539	5.024	14.287	13.298	300	330
(t-values)			6.095	6.746	7.314		
Ctl Brands	0.967	1315	13.381	8.623	14.585	233	243
(t-values)			17.993	5.173	10.519		
Generic	0.931	604	26.599	11.813	6.014	45	44
(t-values)			17.95	7.504	3.11		
All Others	0.578	62	10.496	4.737	6.843	3	3
(t-values)			6.401	4.301	6.283		
All Brands	0.983	2645	9.719	8.545	14.798	3524	4050
(t-values)			11.01	3.268	5.024		

6

Table 3. Summary of Market-Share Model -- Light-Users Segment

Brand	Int	Ad-L	Ad-M	D-EA	D-FA	D-In	D-AO	Cp-S	Cp-M	Price
Heinz	2.00	0.14	+NS	+NS	0.29	+NS	+NS	0.57	-NS	-0.49
Cross Effects	Heinz is less influenced by Del Monte's D-FA , Heinz is more influenced by Hunt's Cp-M, Heinz's D-AO has more influence on Private Labels than predicted by the symmetric model.									
Hunts	-NS	0.26	0.36	0.22	-NS	0.24	0.28	0.72	0.16	-NS
Cross Effects	Del Monte's price has more influence on Hunts, Heinz is more influenced by Hunt's Cp-M, Del Monte's D-EA has more influence on Hunts, Private Labels' D-EA has less influence on Hunts, Hunt's price has more influence in Private Labels, than predicted by the symmetric model.									
Del Monte	0.00	-NS	0.29	0.75	1.09	0.15	0.38	-NS	0.51	-0.61
Cross Effects	Heinz is less influenced by Del Monte's D-FA , Del Monte's price has more influence on Hunts, Del Monte's D-EA has more influence on Hunts, Private Labels' price, D-EA, and Cp-M have more influence on Del Monte, than predicted by the symmetric model.									
Private Labels	-0.62	-NS	0.17	0.35	+NS	0.27	+NS	0.68	+NS	-0.73
Cross Effects	Heinz's D-AO has more influence on Private Labels Private Labels' D-EA has less influence on Hunts, Hunt's price has more influence in Private Labels, Private Labels' price, D-EA, and Cp-M have more influence on Del Monte, than predicted by the symmetric model.									

7

Table 4. Summary of Market-Share Model -- Heavy-Users Segment

Brand	Int	Ad-L	Ad-M	D-EA	D-FA	D-In	D-AO	Cp-S	Cp-M	Price
Heinz	2.30	0.15	+NS	0.07	0.22	0.03	+NS	0.61	0.11	--NS
Cross Effects	Heinz is less influenced by Del Monte's D-FA and Private Labels' Cp_M than predicted by the symmetric model.									
Hunts	-NS	0.21	0.34	0.21	+NS	0.24	0.25	0.71	0.16	-0.50
Cross Effects	Hunts is less influenced by Private Labels' price than predicted by the symmetric model.									
Del Monte	0.00	-NS	0.44	0.78	1.07	+NS	0.32	-NS	0.26	-1.13
Cross Effects	Del Monte's D-FA has less influence on Heinz, Private Labels' D-EA has more influence on Del Monte, and Del Monte's price has less influence on Private Labels than predicted by the symmetric model.									
Private Labels	-0.87	+NS	0.22	0.38	-NS	0.35	0.16	0.54	-NS	-1.30
Cross Effects	Heinz is less influenced by Private Labels' Cp_M, Hunts is less influenced by Private Labels' price, Private Labels' D-EA has more influence on Del Monte, and Del Monte's price has less influence on Private Labels than predicted by the symmetric model.									

8

Table 5. Summary of Market-Share Model -- Aggregate Market

Brand	Int	Ad-L	Ad-M	D-EA	D-FA	D-In	D-AO	Cp-S	Cp-M	Price
Heinz	2.06	0.17	0.07	0.08	0.20	+NS	0.13	0.60		-0.70
Cross Effects	Heinz is less influenced by Del Monte's D-FA , Heinz is less influenced by Hunt's price, than predicted by the symmetric model.									
Hunts	-NS	0.26	0.35	0.19	+NS	0.29	0.22	0.84		-0.40
Cross Effects	Heinz is less influenced by Hunt's price, Del Monte's price has more influence on Hunts, Private Labels' D-In has less influence on Hunts, than predicted by the symmetric model.									
Del Monte	0.00	-NS	0.38	0.75	1.08	+NS	0.33	-NS		-0.57
Cross Effects	Heinz is less influenced by Del Monte's D-FA , Del Monte's price has more influence on Hunts, Private Labels' price and D-EA have more influence on Del Monte, than predicted by the symmetric model.									
Private Labels	-0.79	-NS	0.19	0.35	_NS	0.39	0.17	0.61		-0.84
Cross Effects	Private Labels' D-EA has more influence on Hunts, Hunt's price has more influence in Private Labels, Private Labels' price, D-EA, and Cp-M have more influence on Del Monte, than predicted by the symmetric model.									

9

Table 6. Comparison of Market-Share Models

	Light-Users Segment R-Square	Heavy Users Segment R-Square	Aggregate Market R-Square
Calibration	0.824	0.829	0.845
D.F.	52 -- 3210	48 -- 3214	45 -- 3217
Condition Index	32.5	34.8	21.6
Cross Calibration	0.803	0.814	0.816
D.F.	51 -- 1558	47 -- 1562	44 -- 1565
Cross Validation	0.756	0.766	0.772
D.F.	1 -- 1607	1 -- 1607	1 -- 1607
Forecast Accuracy		**Combined**	**Aggregate**
	RMSE	0.128	0.128
	VAF	0.711	0.711

10

Table 7. Brand Summary -- Across Models

Heinz	Int	Ad-L	Ad-M	D-EA	D-FA	D-In	D-AO	Cp-S	Cp-M	Price
Aggregate	2.06	0.17	0.07	0.08	0.20	+NS	0.13	0.60		-0.70
Heavy Users	2.30	0.15	+NS	0.07	0.22	0.03	+NS	0.61	0.11	--NS
Light Users	2.00	0.14	+NS	+NS	0.29	+NS	+NS	0.57	-NS	-0.49
Hunts										
Aggregate	-NS	0.26	0.35	0.19	+NS	0.29	0.22	0.84		-0.40
Heavy Users	-NS	0.21	0.34	0.21	+NS	0.24	0.25	0.71	0.16	-0.50
Light Users	-NS	0.26	0.36	0.22	-NS	0.24	0.28	0.72	0.16	-NS
Del Monte										
Aggregate	0.00	-NS	0.38	0.75	1.08	+NS	0.33	-NS		-0.57
Heavy Users	0.00	-NS	0.44	0.78	1.07	+NS	0.32	-NS	0.26	-1.13
Light Users	0.00	-NS	0.29	0.75	1.09	0.15	0.38	-NS	0.51	-0.61
Private Labels										
Aggregate	-0.79	-NS	0.19	0.35	_NS	0.39	0.17	0.61		-0.84
Heavy Users	-0.87	+NS	0.22	0.38	-NS	0.35	0.16	0.54	-NS	-1.30
Light Users	-0.62	-NS	0.17	0.35	+NS	0.27	+NS	0.68	+NS	-0.73