ERIC WAARTS, MARTIN CARREE, and BEREND WIERENGA*

The authors build on the idea put forward by Shugan to infer product maps from scanning data. They demonstrate that the actual estimation procedure used by Shugan has several methodological problems and may yield unstable estimates. They propose an alternative estimation procedure, full-information maximum likelihood (FIML), which addresses the problems and yields significantly improved results. An important additional advantage of the procedure is that the parameters of the preference distribution can be estimated simultaneously with the brand coordinates. Hence, it is not necessary to assume a fixed (uniform) distribution of preferences. An empirical application is presented in which the outcomes obtained from Shugan's procedure are compared with those from the proposed procedure.

Full-Information Maximum Likelihood Estimation of Brand Positioning Maps Using Supermarket Scanning Data

The Defender model (Hauser and Shugan 1983) is a promising instrument for understanding the competitive positions of the various brands in a market and serves as a useful framework for testing the effects of alternative competitive strategies. Using the model, one can examine the implications of product policies as well as different pricing strategies for either current or new products. Results of the model show both the effects of actions on one's own market share and the effects on the market shares of competing brands.

One of the essential elements of the Defender model is the so-called "per dollar multidimensional brand map." The model needs information about the (perceived) locations of brands in a multidimensional product space to begin. The literature provides various ways for deriving per dollar multidimensional maps. Hauser and Gaskin (1984) demonstrate how such a map can be estimated by means of a perceptual mapping approach applied to cross-

sectional survey data. An interesting more recent development stems from work by Shugan (1986, 1987). Shugan proposes a procedure for direct estimation of the brand locations that uses a time series of aggregate price and market share data from supermarket scanning data. If one assumes that the Defender approach is valid, the idea is that only certain brand locations in a multidimensional space are consistent with price and share movements observed in the real market. The proposed method should enable one to infer the brand locations most likely to underlie the Defender mechanism that would explain the observed market dynamics.

Shugan's approach seems appealing. However, when applying the procedure to our own data, we noticed certain methodological problems that inhibited proper estimation. The problems are inconsistent estimators, recursive estimation where a simultaneous procedure making full use of the information in the data would be better, frequent occurence of dominated brands, and the exogeneous assumption of the preference distribution. In addition, our empirical application revealed problems with the stability of the inferred brand locations over time.

We first address the methodological issues and then present an alternative estimation procedure. The procedure solves the estimation problems to a large extent while allowing for relaxation of the (uniformity) assumption for the preference distribution. Finally, we report an em-
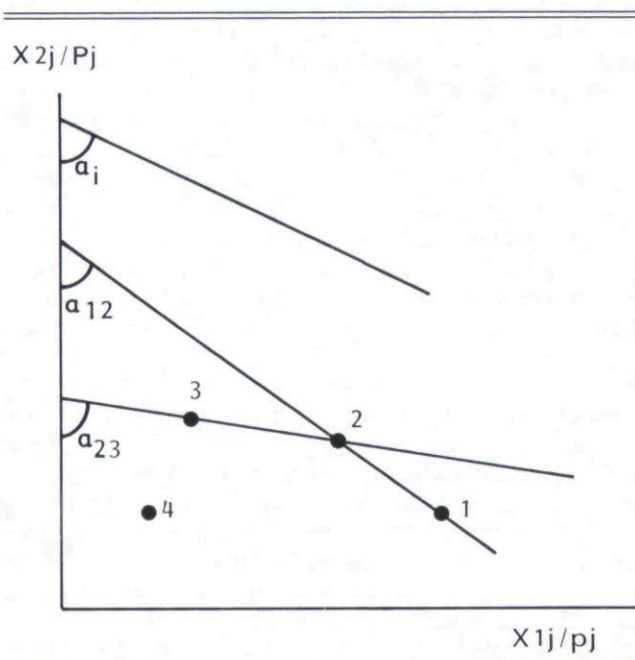
pirical application of the two estimation procedures on a particular product category and compare their results.

THE ORIGINAL RECURSIVE PROCEDURE

The Defender model infers market shares when given perceived brand locations, prices, and a preference distribution (corrected for awareness and distribution level, when appropriate). Shugan (1987), reasoning the other way around, states that one must be able to find the unknown perceived brand locations given brand prices, market shares, and preference distribution. The locations of the brands cannot be inferred from a single period, but must be estimated by using a time series of price and share data.

One can specify a precise aggregate relationship between brand locations, brand prices, and brand shares, which directly follows from the Defender assumptions about consumer choices. Consider a four-brand market as in Figure 1. The brands are ordered from the highest to the lowest value on the first per dollar dimension, x_1/p. The market share m_j of brand j depends on its per dollar position, the per dollar positions of the efficient adjacent brands, and the distribution of the consumer preference angles. Consumers with preference angles smaller than the angle α_{12} will choose brand 1 and consumers with angles larger than α_{23} will choose brand 3. Consumer i with $\alpha_{12} < i < \alpha_{23}$ in Figure 1 will choose brand 2. Note that brand 4 is dominated by brands 2 and 3 and therefore will not be preferred by any consumer.

Figure 1
GRAPHICAL REPRESENTATION OF THE DEFENDER MODEL

Such a brand is called "inefficient" or "dominated." Its market share will be zero (provided the other brands have full awareness and availability). The angle formed by the vertical axis and a line passing through both brands i and j (in time period t) is described as:

$$(1) \qquad \alpha_{ijt} = \arctan[(x_{1i}/p_{it} - x_{1j}/p_{jt})/(x_{2j}/p_{jt} - x_{2i}/p_{it})]$$

where:

i and j are efficient adjacent brands ($i < j$),
x_{dk} = the level of brand k on dimension d, and
p_{kt} = the price of brand k in time period t.

The mathematical relationship between (1) market share in period t and (2) per dollar brand locations in that period and the preference distribution is given by:

$$(2) \qquad m_{jt} = F(\alpha_{j,j+1,t}) - F(\alpha_{j-1,j,t})$$

where:

j = 1, 2, 3 ... J, ordered from high to low on the first dimension,
J = number of efficient brands in the market,
m_{jt} = market share of brand j in time period t,
α_{01t} = 0° for all t; $\alpha_{J,J+1,t}$ = 90° for all t, and
$F(\bar{\alpha})$ = the cumulative distribution of the consumer preference angle $\bar{\alpha}$.

The per dollar positions of the brands change over time as the brand prices vary. Combining equations 1 and 2 yields equation 3, which implies a procedure for inferring the positions of the brands from a series of prices and shares.

$$(3) \qquad \frac{p_{jt}}{p_{j-1,t}} = \frac{\left\{ x_{1j} + x_{2j} \tan F^{-1}\left[\left(\sum_{i=1}^{j} m_{it}\right)\right]\right\}}{\left\{ x_{1,j-1} + x_{2,j-1} \tan F^{-1}\left[\left(\sum_{i=1}^{j} m_{it}\right)\right]\right\}}$$

Shugan (1987) rewrites equation 3 into such a form that ordinary least squares regression analysis estimation is possible.

$$(4) \qquad \text{PRATIO}_{j-1,j,t} = \beta_{jx}v_{jt} + \beta_{jy}w_{jt}$$

where:

$$\text{PRATIO}_{j-1,j,t} = p_{jt}/p_{j-1,t}$$

$$\beta_{jx} = x_{1j}$$

$$\beta_{jy} = x_{2j}$$

$$v_{jt} = \left\{ x_{1,j-1} + x_{2,j-1} \tan\left[F^{-1}\left(\sum_{i=1}^{j-1} m_{it}\right)\right]\right\}^{-1}$$

$$w_{jt} = v_{jt} \tan\left[F^{-1}\left(\sum_{i=1}^{j-1} m_{it}\right)\right]$$

The procedure does not allow for simultaneous estimation of the preference function F; it must be measured separately in the marketplace, or a particular distribution

must be assumed. Shugan starts by assuming a uniform preference distribution F.

One of the coordinates to be estimated must be fixed: $x_{11} = 1$ (the scale of the map is arbitrary). For estimation of a two-brand map, equation 4, implying a constrained regression through the origin, is used with $j = 2$. The coordinates of the two brands then must be found by iteration over x_{21} and selection of the value of x_{21} that maximizes the R^2 of the regression.

Shugan estimates the multibrand map by using regression equation 4 recursively. This means that estimates of $x_{1,j-1}$ and $x_{2,j-1}$ are substituted into the jth regression. For every value of x_{21}, the coordinates of the other brands are estimated and the multiple R^2 is computed. The best solution is the set of coordinates that yields the highest multiple R^2 (which is equal to the minimal total sum of the squared errors).

The brands are assumed to be ordered from high to low on the first dimension. The ordering, however, is generally not known in advance. Therefore, regressions must be carried out for all possible orderings of the brands. The ordering that yields the best fit in terms of multiple R^2 is selected as the final best solution.

METHODOLOGICAL PROBLEMS

The following methodological problems are related to the procedure just presented.

Inverse Causality

Though with equation 4 Shugan does not suggest that prices are caused by market shares, note that in equation 4 price ratios are used as dependent variables and market shares as explanatory variables. Theoretically, one would expect that market shares are caused by price ratios and other factors (represented by the disturbance terms). Hence, market shares should be the dependent variables. If in reality market shares are caused by price ratios and disturbances, the use of market shares as independent variables is flawed and will lead to inconsistent estimators. The reason is that the market shares will be correlated with the disturbances (Maddala 1977, p. 153).

Recursive Use of a Single-Equation Model

The procedure estimates equation 4 recursively. As a consequence, the information that parameters arise in two consecutive equations is not used optimally. By estimating x_{1j} and x_{2j} ($j \neq 1, J$) in the $(j - 1)^{th}$ equation and substituting the estimates into the jth equation, we ignore the fact that the jth equation provides information about the values of x_{1j} and x_{2j}.

Another point should be noted. In estimating the equations separately, one implicitly assumes that the disturbances of the equations are mutually uncorrelated. That assumption is questionable, however, because variables such as in-store promotion of the jth brand in general influence all market shares. Especially the shares of the $(j - 1)^{th}$ and the $(j + 1)^{th}$ brand are expected to decrease as a result of promotional activities for the jth brand.

Treatment of Dominated Brands

Shugan (1987, p. 4) states that his estimation procedure "requires the assumption that every brand obtains a positive market share during every time period." The reason is that equations 1 through 4 hold only for efficient brands. Though the assumption is necessary for accurate estimation, there is no guarantee that the estimated parameters produce predicted positive market shares for all brands in all the periods used in the estimation. Even with only positive market shares as data, the estimation procedure, searching for minimal error in price-ratio fitting, allows for inferred product locations that imply zero (or negative) calculated market shares (dominated brands). We performed an application on a real prices-shares time series of a four-brand market with positive true market shares only. We found that one of the brands (average true share of 9%) was predicted to be dominated (and to have negative market share) in 39% of the periods involved in the estimation. In case of dominance, the brand should be ignored in the prediction stage and its market share set to zero. Equations 1 and 2 should be applied to the remaining efficient brands only. However, the information that the market share of a dominated brand should be set to zero *ex-post* is not used in the least squares regression procedure. That drawback causes a bias in the estimation.

Assumption of a Fixed (Uniform) Preference Distribution

As Shugan (1987, p. 13–14) points out, the preference distribution assumption limits the applicability of the procedure. He shows that the shape of the preference distribution can have consequences for the derived locations of the brands and for the prediction of the market shares. The shape of the preference distribution therefore can be critical for policy recommendations following from a Defender application. With Shugan's method it is not possible to estimate the preference distribution directly, but a particular distribution must be assumed. One could, of course, estimate the model by using a large set of different assumptions about the shape of the preference distribution and then select the best one in terms of R^2. Shugan avoids this very cumbersome procedure and assumes a uniform distribution. Another possibility is to measure the preference distribution separately in the marketplace, but doing so would reduce the value of the proposed method.

ALTERNATIVE PROCEDURE

The Shugan approach is an important development in the estimation of brand positions from scanning data. By overcoming the problems just mentioned, further progress can be made. We demonstrate that this can be done by specifying the model as a simultaneous equations model with the proper directions of causalities. Reduction of the dominance problem, and the possibility of estimating the preference distribution in one run with the product

locations, improves the quality of the estimates and the applicability of the model.

A well-known econometric technique for estimating simultaneous equation models is the full-information maximum likelihood (FIML) method. It is a "system method" in which the parameters of all equations are estimated simultaneously, with all the information in the model (Maddala 1977, p. 486). We proceed by specifying a model that allows for estimation by the FIML technique. The model specification we propose is derived in the following steps.

Rewriting equation 3 as

$$(5) \quad \left\{ x_{1j} + x_{2j} \tan\left[F^{-1}\left(\sum_{i=1}^{j} m_{it} \right) \right] \right\} \frac{p_{j+1,t}}{p_{jt}}$$
$$= x_{1,j+1} + x_{2,j+1} \tan\left[F^{-1}\left(\sum_{i=1}^{j} m_{it} \right) \right],$$

and solving for $\sum_{i=1}^{j} m_{it}$, we get

$$(6) \quad \sum_{i=1}^{j} m_{it} = F\left(\arctan\left[\frac{x_{1,j+1} - x_{1j} \frac{p_{j+1,t}}{p_{jt}}}{x_{2j} \frac{p_{j+1,t}}{p_{jt}} - x_{2,j+1}} \right] \right),$$

$$j = 1, \ldots, J - 1; t = 1, \ldots, T,$$

$$\sum_{i=1}^{J} m_{it} = 1,$$

$$t = 1, \ldots, T.$$

This relation implies the following system of nonlinear equations.

$$m_{1t} = F\left(\arctan\left[\frac{x_{12} - x_{11} \frac{p_{2t}}{p_{1t}}}{x_{21} \frac{p_{2t}}{p_{1t}} - x_{22}} \right] \right)$$

$$m_{2t} = F\left(\arctan\left[\frac{x_{13} - x_{12} \frac{p_{3t}}{p_{2t}}}{x_{22} \frac{p_{3t}}{p_{2t}} - x_{23}} \right] \right)$$

$$- F\left(\arctan\left[\frac{x_{12} - x_{11} \frac{p_{2t}}{p_{1t}}}{x_{21} \frac{p_{2t}}{p_{1t}} - x_{22}} \right] \right)$$

$$(7)$$

$$m_{3t} = F\left(\arctan\left[\frac{x_{14} - x_{13} \frac{p_{4t}}{p_{3t}}}{x_{23} \frac{p_{4t}}{p_{3t}} - x_{24}} \right] \right)$$

$$- F\left(\arctan\left[\frac{x_{13} - x_{12} \frac{p_{3t}}{p_{2t}}}{x_{22} \frac{p_{3t}}{p_{2t}} - x_{23}} \right] \right)$$

$$m_{Jt} = 1 - F\left(\arctan\left[\frac{x_{1,J} - x_{1,J-1} \frac{p_{Jt}}{p_{J-1,t}}}{x_{2,J-1} \frac{p_{Jt}}{p_{J-1,t}} - x_{2,J}} \right] \right)$$

The x_{1j}'s and the x_{2j}'s are unknown parameters. It is clear that the m_{jt}'s do not change when both the x_{1j}'s and the x_{2j}'s are multiplied by some constant ξ. Therefore, following Shugan (1987, p. 6), we assume $x_{11} = 1$. Note that equations 5 through 7 are valid only if all J brands are efficient.

The full-information maximum likelihood (FIML) method is used to estimate the parameters x_{1j} and x_{2j} of this model. Adding disturbances U_{jt} and using $x_{11} = 1$, we obtain the following model to be used for estimation.

$$(8) \quad m_{1t} = F\left(\arctan\left[\frac{x_{12} - \frac{p_{2t}}{p_{1t}}}{x_{21} \frac{p_{2t}}{p_{1t}} - x_{22}} \right] \right) + U_{1t}$$

$$m_{2t} = F\left(\arctan\left[\frac{x_{13} - x_{12} \frac{p_{3t}}{p_{2t}}}{x_{22} \frac{p_{3t}}{p_{2t}} - x_{23}} \right] \right)$$

$$- F\left(\arctan\left[\frac{x_{12} - \frac{p_{2t}}{p_{1t}}}{x_{21} \frac{p_{2t}}{p_{1t}} - x_{22}} \right] \right) + U_{2t}$$

$$m_{3t} = F\left(\arctan\left[\frac{x_{14} - x_{13} \frac{p_{4t}}{p_{3t}}}{x_{23} \frac{p_{4t}}{p_{3t}} - x_{24}} \right] \right)$$

$$- F\left(\arctan\left[\frac{x_{13} - x_{12} \frac{p_{3t}}{p_{2t}}}{x_{22} \frac{p_{3t}}{p_{2t}} - x_{23}} \right] \right) + U_{3t}$$

$$\vdots$$

$$m_{J-1,t} = F\left(\arctan\left[\frac{x_{1,J} - x_{1,J-1} \frac{p_{Jt}}{p_{J-1,t}}}{x_{2,J-1} \frac{p_{Jt}}{p_{J-1,t}} - x_{2,J}} \right] \right)$$

$$- F\left(\arctan\left[\frac{x_{1,J-1} - x_{1,J-2} \frac{p_{J-1,t}}{p_{J-2,t}}}{x_{2,J-2} \frac{p_{J-1,t}}{p_{J-2,t}} - x_{2,J-1}} \right] \right) + U_{J-1,t},$$

where the equation for m_{Jt} is left out because $\sum_{i=1}^{J} m_{it} = 1$, which implies $\sum_{i=1}^{J} U_{it} = 0$. The endogenous variables are m_{jt}, $j = 1, \ldots, J - 1$, and the exogenous variables are $p_{j+1,t}/p_{jt}$, $j = 1, \ldots, J - 1$. This

is in accordance with the causality assumptions. The parameters to be estimated are x_{1j} and x_{2j}, $j = 2, \ldots, J$, and x_{12}. To derive the likelihood function to be maximized, we introduce the following notation.

(9)
$$\mathbf{Y}_t = (m_{1t}\, m_{2t} \ldots m_{J-1,t})^T$$

$$\mathbf{z}_t = \left(\frac{p_{2t}}{p_{1t}}\frac{p_{3t}}{p_{2t}} \ldots \frac{p_{Jt}}{p_{J-1,t}}\right)^T$$

$$\boldsymbol{\beta} = (x_{21}\, x_{12}\, x_{22} \ldots x_{2J})^T$$

$$\mathbf{U}_t = (U_{1t}\, U_{2t} \ldots U_{J-1,t})^T$$

$$\mathbf{f}(\mathbf{z}_t, \boldsymbol{\beta}) = (f_1(\mathbf{z}_t, \boldsymbol{\beta})\, f_2(\mathbf{z}_t, \boldsymbol{\beta}) \ldots f_{J-1}(\mathbf{z}_t, \boldsymbol{\beta}))^T,$$

where $f_j(\mathbf{z}_t, \boldsymbol{\beta})$ stands for the systematic part of the right side of the jth equation in 8. The concentrated log-likelihood function to be maximized is (see Appendix)

(10)
$$\mathbf{L}^*(\boldsymbol{\beta}; \mathbf{y}, \mathbf{z}) = \text{const.} - \frac{1}{2}T \log \left| \frac{1}{T} \sum_{t=1}^{T} \right.$$

$$\left. [\mathbf{y}_t - \mathbf{f}(\mathbf{z}_t, \boldsymbol{\beta})][\mathbf{y}_t - \mathbf{f}(\mathbf{z}_t, \boldsymbol{\beta})]^T \right|.$$

This likelihood function is to be maximized with respect to $\boldsymbol{\beta}$, which means that the x_{1j}'s and the x_{2j}'s are estimated simultaneously. Numerical minimization of $-\log L^*$ with respect to $\boldsymbol{\beta}$ is performed by a comprehensive but standard quasi-Newton algorithm (routine E04JBF from the NAG Fortran Library), which yields an estimate $\hat{\boldsymbol{\beta}}$ of $\boldsymbol{\beta}$. The covariance matrix $\boldsymbol{\Omega}$ of the $\mathbf{U}_t$'s is estimated as

(11)
$$\hat{\boldsymbol{\Omega}} = \frac{1}{T} \sum_{t=1}^{T} [\mathbf{y}_t - \mathbf{f}(\mathbf{z}_t, \hat{\boldsymbol{\beta}})][\mathbf{y}_t - \mathbf{f}(\mathbf{z}_t, \hat{\boldsymbol{\beta}})]^T.$$

The asymptotic distribution of the maximum likelihood estimator $\hat{\boldsymbol{\beta}}$ is multivariate normal with mean $\boldsymbol{\beta}$ and covariance matrix $\boldsymbol{\Sigma}$. A consistent estimate of $\boldsymbol{\Sigma}$ is given by $\hat{\boldsymbol{\Sigma}}$, where

(12)
$$\hat{\boldsymbol{\Sigma}} = \left(\frac{-\partial^2 L^*}{\partial \boldsymbol{\beta}\, \partial \boldsymbol{\beta}^T}\right)^{-1} \text{ evaluated at } \boldsymbol{\beta} = \hat{\boldsymbol{\beta}}.$$

Note that this procedure allows for the introduction of a more general preference distribution. The parameters of this distribution function can be considered as additional elements of the parameter vector $\boldsymbol{\beta}$ and can be estimated simultaneously with the x_{1j}'s and the x_{2j}'s.

Local Optima and Starting Values

A problem with iterative estimation procedures such as FIML is the possible presence of local optima. Depending on the specific starting values of the parameters, the procedure could reach different "optima." Generally, this problem decreases with the number of observations, and therefore use of good starting points in combination with a fair number of observations is recommended.

Our applications of the FIML procedure to a real dataset confirm that local optima do occur in this model,

and that starting values are critical for the produced solution. Only an exhaustive search process starting with many possible parameter values and many possible orderings can more or less guarantee finding the globally best solution, but such a search is difficult to carry out in practice. One could perform an extensive search process using the FIML method, but the computing costs rise considerably as the number of brands increases. Hence, the aim is to find a promising starting point for the estimation procedure. We recommend making use of the work that has been done already, and using the outcomes of the "fast" Shugan procedure as "reasonable" starting input values for the FIML procedure. Though there is no guarantee that we will find the global optimum in terms of likelihood, this procedure ensures that the quality of the estimates will at least be equal to, but probably surpass, the quality of the initial configuration. Of course, one can do a sensitivity analysis afterward by taking alternative starting values and/or orderings.

If we want to estimate the preference distribution parameters simultaneously (e.g., using a beta distribution), we can start with uniformity values (α and β of the beta distribution both equal one). When the values of α and β are estimated, we can easily test whether the produced parameter values differ significantly from the starting values.

On balance, we think the proposed model specification and the use of the FIML method provide several advantages over the original procedure proposed by Shugan. Our procedure makes maximal use of the information available in the data, specifies causalities in the right direction, uses simultaneous equations, and also allows for estimation of the preference distribution in one estimation with the location parameters. The procedure does not theoretically solve the dominance problem. The estimation procedure still allows calculated shares to be negative, which of course should be set to zero afterward. However, as we show in the next section, our applications suggest that the new procedure circumvents the dominance problem to a large extent.

EMPIRICAL APPLICATION

In our empirical application of the estimation procedures, we used a time series of price and sales data obtained from a scanner-equipped supermarket in a town near the Dutch capital, Amsterdam. The data consisted of a two-year (1985–1986) series of weekly data on prices and sales of brands in a particular product class (frequently bought products in the food and beverage category). The exact product class and the brand names are disguised for confidentiality. Though the scanner database contains data on about 15 items in the product class, for our analysis we selected the two major national premium brands, a strong private label of the supermarket, and a cluster of smaller nationally distributed nonpremium brands (B-brands). Together these brands account for, on the average, an 84% share in the total product category.
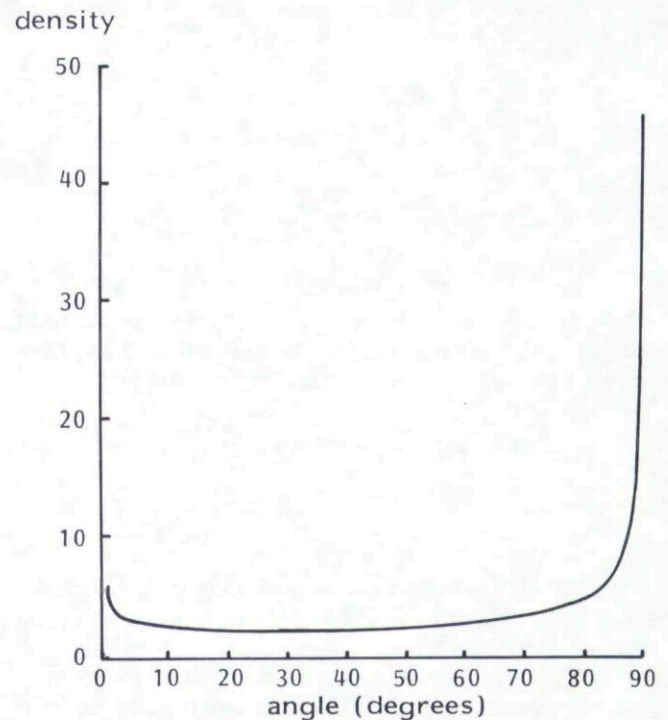
We estimated a four-brand map using Shugan's "re-

cursive" procedure on the total 100-period time series. We then applied the FIML procedure, starting with Shugan's parameter values. We estimated the brand locations once assuming a uniform preference distribution and once assuming beta-distributed preferences. Results are reported in Table 1. As can be seen, our theoretical expectations are confirmed. In comparison with Shugan's solution, the FIML procedure produces different product locations that will affect predicted market shares under different price conditions. In the final solution, we find that both brand 1 and brand 4 are located on the axes. This finding is somewhat surprising because we only restrict the coordinates to be non-negative. However, this finding is not an artifact of the FIML procedure. In a study with test data derived from a hypothesized brand configuration with known positive coordinates, FIML reproduced these (nonzero) coordinates exactly. Table 1 also shows that the FIML procedure yields substantially lower squared errors, even assuming a uniform preference distribution. The resulting sum of squared errors is reduced by 47% (the sign test on the equality of the squared errors renders a p-value of 0.000). Results also show that the percentage of periods in which one or more brands are predicted to be dominated (denoted in the table by "% of dominance") decreases from 39% to 1%, which reduces the dominance problem to minor proportions.

The FIML solution can be improved further if we allow the preference distribution to be estimated freely. Our results show that for this application, the assumption of a uniform distribution cannot be confirmed (the likelihood-ratio test on $\alpha = \beta = 1$ renders a p-value of 0.000). Our predictions are additionally improved by an 11% sum-of-squared-errors reduction if we specify a beta distribution (the sign test renders a p-value of 0.000). Figure 2 shows the estimated density function of the preference angles. The figure indicates a high preference density for products that score high on the second dimension.

To compare the two estimation methods further, we explored the (in)stability of the estimated brand locations over "shifting" periods in our two-year time series. We

Figure 2

DISTRIBUTION OF PREFERENCE ANGLES

inferred a brand map for each of 50 consecutive 50-week periods (a map for period 1–50, one for period 2–51, and so on). Though the per dollar positions of a brand will vary from period to period, one expects the resulting "absolute" locations to be reasonably stable over time because the brand map should reflect the perceptions of customers. To predict future market shares accurately, stability of perceptions is a requirement. Naturally, some instability of the estimations can be expected, not only because of stochastics, but also because of the limita-

Table 1

RESULTS OF ALTERNATIVE PROCEDURES ON FOUR BRANDS/100 PERIODS

Brand	Recursive procedure		FIML/uniform procedure		FIML/beta procedure	
	Location	S.E.	Location	S.E.	Location	S.E.
1	(1.000, .070)	[—, —]	(1.000, .000)	[—, —]	(1.000, .000)	[—, —]
2	(.962, .745)	[.15, .10]	(.841, 1.179)	[.01, .10]	(.867, .812)	[.01, .18]
3	(.873, 1.063)	[.01, .03]	(.607, 1.962)	[.02, .11]	(.659, 1.262)	[.03, .26]
4	(.444, 1.195)	[.03, .02]	(.000, 2.170)	[—, .09]	(.000, 1.276)	[—, .26]

Results for parameters of the preference distribution, residual sum of squared errors, and percentage of dominance occurrence

Alpha					.734	[.08]
Beta					.386	[.04]
RSS	2.381		1.248		1.113	
% of dominance	39		1		1	

Figure 3
RSS FOR SHIFTING 50-WEEK PERIODS
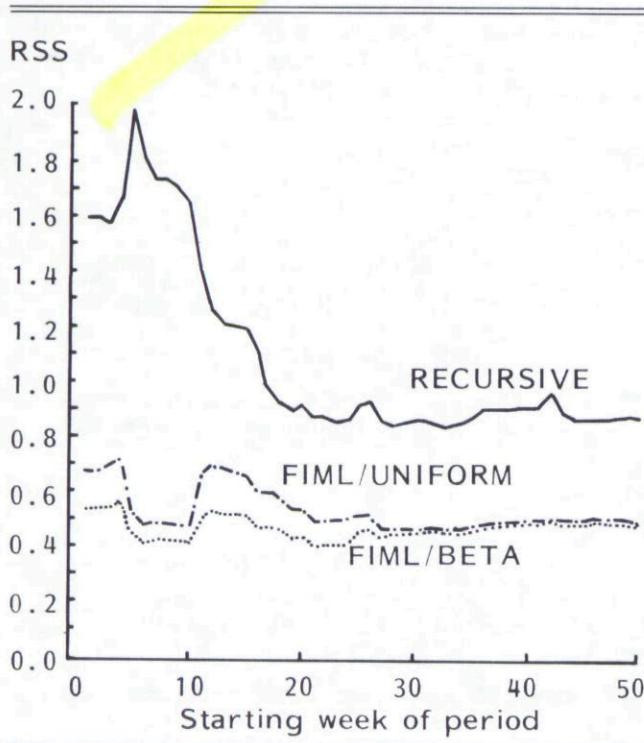

Figure 4
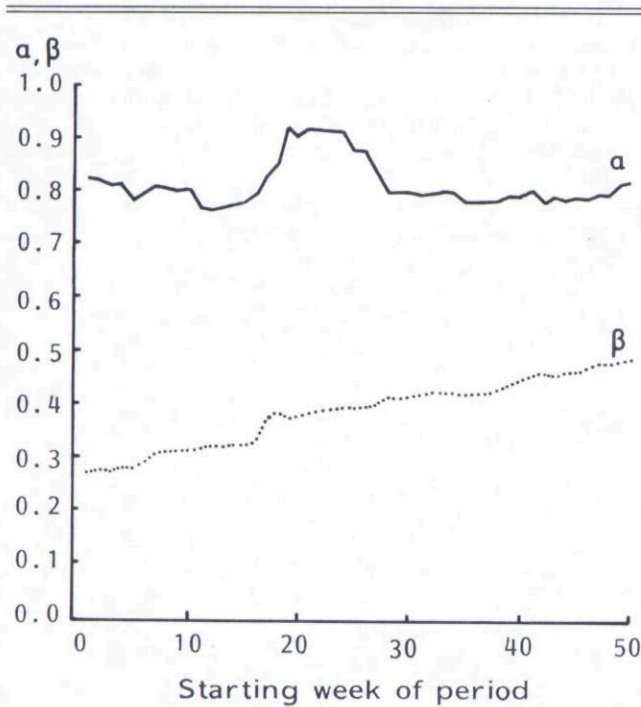### BETA-DISTRIBUTED PREFERENCES: α AND β FOR SHIFTING 50-WEEK PERIODS

Table 2
PRODUCT LOCATIONS AND PREFERENCE DISTRIBUTION ESTIMATIONS OVER 50 CONSECUTIVE 50-WEEK PERIODS (means and variances)

Param-eter	Recursive procedure		FIML/uniform procedure		FIML/beta procedure	
	Mean	Variance	Mean	Variance	Mean	Variance
x_{12}	.95	.0001	.85	.0002	.86	.0001
x_{13}	.89	.0016	.62	.0005	.65	.0002
x_{14}	.48	.0057	0	0	0	0
x_{21}	.19	.0236	0	0	0	0
x_{22}	.53	.0646	1.12	.0196	.65	.0414
x_{23}	.78	.1052	1.89	.0145	1.03	.0671
x_{24}	.93	.1022	2.11	.0098	1.05	.0744
Alpha	—	—	—	—	.81	.0018
Beta	—	—	—	—	.38	.0042

tions of the model itself, which only uses information about prices and shares.

Results of the analysis are shown in Figures 3 and 4 and Table 2. Figure 3 shows the quality of the predictions in terms of resulting squared errors (RSS) for each of the procedures. The FIML procedures yield stable and high quality predictions in comparison with the recursive method used by Shugan. Estimation of the (beta) preference distribution slightly improves the quality of the solutions. Table 2 summarizes the means and variances of the estimated product locations and preference distribution parameters over the moving 50 periods. Because the periods overlap 98%, the variances should be used for comparison only. Note that the stability of the estimated locations improves with use of the FIML procedure; the variance of each of the estimated brand coordinates diminishes. Estimates of the (beta) preference parameters reveal that the distribution parameters are very stable and significantly different from one (uniformity). The results in Figure 4 show a slight upward trend in beta, meaning that the relative preference for the second dimension is decreasing over time. The stability of the coordinates on the first dimension increases when the beta distribution is used, and the stability on the second dimension seems to decrease somewhat.

CONCLUSIONS

Our study demonstrates that the recursive regression procedure, proposed by Shugan to estimate brand locations from scanning data, is a good starting point but has several problems—inconsistent estimators, no simultaneous estimation, dominated brands, and unstable estimates. Those problems can be solved (the dominance problem only partially) by using the full-information maximum likelihood method. In addition, that method makes it possible to estimate the parameters of the preference distribution simultaneously with the brand locations and to test for a specific distribution. Results of an empirical comparison of the two methods show signifi-

cantly improved predictions in terms of resulting errors and stability of the parameters by using the proposed method.

APPENDIX

Model 8 can be written as

$$(13) \qquad \mathbf{Y}_t = \mathbf{f}(\mathbf{z}_t, \boldsymbol{\beta}) + \mathbf{U}_t, \qquad t = 1, \ldots, T.$$

Assuming that the $(J - 1)$-dimensional random vector $\mathbf{U}_t$ has a multivariate normal distribution with mean vector $\mathbf{0} = (0\ 0\ \ldots\ 0)^T$ and covariance matrix $\boldsymbol{\Omega}$, it follows that the joint density function of $\mathbf{U}_t$ is

$$(14) \qquad h(\mathbf{u}_t) = (2\pi)^{-(J-1)/2} |\boldsymbol{\Omega}|^{-1/2} \exp\left\{-\frac{1}{2} \mathbf{u}_t^T \boldsymbol{\Omega}^{-1} \mathbf{u}_t\right\}.$$

Making use of the fact that the Jacobian of the transformation $\mathbf{U}_t \rightarrow \mathbf{Y}_t$ equals one, the joint density function of $\mathbf{Y}_t$ is derived as

$$(15) \qquad g(\mathbf{y}_t) = (2\pi)^{-(J-1)/2} |\boldsymbol{\Omega}|^{-1/2}$$

$$\cdot \exp\left\{-\frac{1}{2}[\mathbf{y}_t - \mathbf{f}(\mathbf{z}_t, \boldsymbol{\beta})]^T \boldsymbol{\Omega}^{-1}[\mathbf{y}_t - \mathbf{f}(\mathbf{z}_t, \boldsymbol{\beta})]\right\}.$$

Hence, the log-likelihood function of the $\mathbf{y}_t$'s equals

$$(16) \quad L(\boldsymbol{\beta}, \boldsymbol{\Omega}; \mathbf{y}, \mathbf{z}) = \text{const.} - \frac{1}{2} T \log|\boldsymbol{\Omega}| - \frac{1}{2} \sum_{t=1}^{T}$$

$$\cdot [\mathbf{y}_t - \mathbf{f}(\mathbf{z}_t, \boldsymbol{\beta})]^T \boldsymbol{\Omega}^{-1}[\mathbf{y}_t - \mathbf{f}(\mathbf{z}_t, \boldsymbol{\beta})]$$

$$= \text{const.} + \frac{1}{2} T \log|\boldsymbol{\Omega}^{-1}|$$

$$- \frac{1}{2} \text{tr}\{([\mathbf{y}_1 - \mathbf{f}(\mathbf{z}_1, \boldsymbol{\beta})] \ldots$$

$$\cdot [\mathbf{y}_T - \mathbf{f}(\mathbf{z}_T, \boldsymbol{\beta})])^T \boldsymbol{\Omega}^{-1}([\mathbf{y}_1 - \mathbf{f}(\mathbf{z}_1, \boldsymbol{\beta})] \ldots$$

$$\cdot [\mathbf{y}_T - \mathbf{f}(\mathbf{z}_T, \boldsymbol{\beta})])\}$$

$$= \text{const.} + \frac{1}{2} T \log|\boldsymbol{\Omega}^{-1}|$$

$$- \frac{1}{2} \text{tr}\{\boldsymbol{\Omega}^{-1}([\mathbf{y}_1 - \mathbf{f}(\mathbf{z}_1, \boldsymbol{\beta})] \ldots$$

$$\cdot [\mathbf{y}_T - \mathbf{f}(\mathbf{z}_T, \boldsymbol{\beta})])([\mathbf{y}_1 - \mathbf{f}(\mathbf{z}_1, \boldsymbol{\beta})] \ldots$$

$$\cdot [\mathbf{y}_T - \mathbf{f}(\mathbf{z}_T, \boldsymbol{\beta})])^T\}.$$

Differentiating with respect to $\boldsymbol{\Omega}^{-1}$ (using the results $\partial \log|\mathbf{A}|/\partial\mathbf{A} = (\mathbf{A}^T)^{-1}$ and $\partial\ \text{tr}(\mathbf{AC})/\partial\mathbf{A} = \mathbf{C}^T$, where $\mathbf{C}$ does not involve elements of $\mathbf{A}$) yields

$$(17) \qquad \frac{\partial L}{\partial\boldsymbol{\Omega}^{-1}} = \mathbf{O} \Rightarrow \frac{1}{2} T((\boldsymbol{\Omega}^{-1})^T)^{-1}$$

$$-\frac{1}{2}([\mathbf{y}_1 - \mathbf{f}(\mathbf{z}_1, \boldsymbol{\beta})] \ldots$$

$$\cdot [\mathbf{y}_T - \mathbf{f}(\mathbf{z}_T, \boldsymbol{\beta})])([\mathbf{y}_1 - \mathbf{f}(\mathbf{z}_1, \boldsymbol{\beta})] \ldots$$

$$\cdot [\mathbf{y}_T - \mathbf{f}(\mathbf{z}_T, \boldsymbol{\beta})])^T = \mathbf{O}.$$

The $(J - 1) \times (J - 1)$ matrix $\mathbf{O}$ is defined to have all elements equal to zero. Equation 17 gives

$$(18) \qquad \hat{\boldsymbol{\Omega}} = \frac{1}{T}([\mathbf{y}_1 - \mathbf{f}(\mathbf{z}_1, \boldsymbol{\beta})] \ldots [\mathbf{y}_T - \mathbf{f}(\mathbf{z}_T, \boldsymbol{\beta})])$$

$$\cdot ([\mathbf{y}_1 - \mathbf{f}(\mathbf{z}_1, \boldsymbol{\beta})] \ldots [\mathbf{y}_T - \mathbf{f}(\mathbf{z}_T, \boldsymbol{\beta})])^T$$

$$= \frac{1}{T} \sum_{t=1}^{T} [\mathbf{y}_t - \mathbf{f}(\mathbf{z}_t, \boldsymbol{\beta})][\mathbf{y}_t - \mathbf{f}(\mathbf{z}_t, \boldsymbol{\beta})]^T.$$

Substituting 18 into 16, we get the concentrated log-likelihood function as in equation 10.

REFERENCES

Hauser, John R. and Steven P. Gaskin (1984), "Application of the 'Defender' Consumer Model," *Marketing Science*, 3 (Fall), 327–51.

——— and Steven M. Shugan (1983), "Defensive Marketing Strategies," *Marketing Science*, 2 (Fall), 319–60.

Maddala, G. S. (1977), *Econometrics*. New York: McGraw-Hill Book Company.

Shugan, Steven M. (1986), "Inferring Brand Positioning Maps From Price/Share Data: The Case of Bathroom Tissue," working paper, Graduate School of Business, University of Chicago, nr. 15.

——— (1987), "Estimating Brand Position Maps Using Supermarket Scanning Data," *Journal of Marketing Research*, 24 (February), 1–8.