The authors propose a new approach to obtaining stockkeeping-unit (SKU)-level preferences and response sensitivities. The authors distinguish an attribute-level model, in which the unit of analysis is the market share for an alternative created by aggregation (e.g., Colgate toothpaste), from a truly disaggregate SKU-level model, and they establish an analytical relationship between parameters that they obtain from the two models. The authors show that SKU-level parameters can be recovered by calculation from estimated attribute-level parameters, circumventing the need for direct estimation of the more complex SKU-level model. They calibrate the store data market share model using 98 weeks of data for ten brands and 168 SKUs of toothpaste. Rather than estimate 168 preference parameters (when there is an “outside” alternative in addition to the 168 “inside” ones), it is only necessary to estimate ten brand-preference parameters from which the 168 parameters can be computed, as long as share and marketing-mix data are available at the SKU level. Covariate effects, such as marketing-mix response parameters, can be recovered in a similar fashion. Holdout tests demonstrate superior predictive performance, and the authors discuss implications for the derivation of elasticities for new SKU introductions.
and critique of this practice, see Zanutto and Bradlow (2004). A third option (and the one we pursue in this article) is to develop analytical relationships between the complex model and simpler aggregated models and to exploit them to obtain the required parameters.

SKUs AND PRODUCT ATTRIBUTES

Fader and Hardie (1996) represent consumer utility for a particular SKU as an additive combination of utilities for the attribute levels that define the alternative rather than as a single fixed effect. An important property of this model is that the number of parameters required does not increase with the number of alternatives but rather with the number of additional levels of underlying attributes. This procedure not only produces superior model fits in a sample but also enables researchers to create forecasts for new SKUs (provided that the individual elements that make up the new alternative are present in the market). Ho and Chong (2003) extend this idea and develop a model in which consumers “reinforce” chosen and nonchosen options in their patterns of SKU selection in a manner consistent with an experience-weighted attraction theory of consumer learning (Camerer and Ho 1999). In this model, the required number of parameters does not increase in either the number of SKUs or the number of attribute levels.

Both studies use household panel data, yet Bucklin and Gupta (1999) note that managers have a preference for store-level market share data, which is more readily available and believed to be more reliable (less subject to sampling biases). Moreover, academic researchers in marketing are showing considerable interest in market share models of the type advanced in the economics literature (e.g., Nevo 2001). In this article, we seek to bridge the gap between advances that are made in disaggregate choice modeling and the data preferences of retailers. We go beyond previous empirical studies and formalize the analytical relationships between the parameters that are estimated from simple “composite” demand models and the true SKU-level fixed effects and response parameters that can be derived from them, thereby circumventing the need for direct estimation.

A MOTIVATING EXAMPLE

To appreciate the task at hand, consider the following data from the toothpaste category, which we used in our empirical application. Table 1, Panel A, shows that there are 168 unique SKUs in the category. Furthermore, over a period of almost two years of observation, many SKUs are added, dropped, or stocked-out. Table 1, Panel B, indicates that each unique SKU can be described as a combination of single unique levels over the following five attributes: brand, flavor, form, function, and size. A “full assortment” that contains SKUs that span all possible combinations would yield $10 \times 14 \times 3 \times 7 \times 4 = 11,760$ SKUs.

Table 1 highlights several empirical observations, some of which (to our knowledge) appear to have been largely overlooked, yet seem highly germane to model building.

Panel A: The SKU choice set changes over time, with slightly fewer than one-third of the SKUs available for all weeks of the data set.

Panel B: There are many more possible SKU locations—that is, unique combinations of attribute levels (11,760)—than those that existing SKUs (168) occupy.

Table 1 highlights several empirical observations, some of which (to our knowledge) appear to have been largely overlooked, yet seem highly germane to model building.

Panel A: The SKU choice set changes over time, with slightly fewer than one-third of the SKUs available for all weeks of the data set.

Panel B: There are many more possible SKU locations—that is, unique combinations of attribute levels (11,760)—than those that existing SKUs (168) occupy.

A MOTIVATING EXAMPLE

To appreciate the task at hand, consider the following data from the toothpaste category, which we used in our empirical application. Table 1, Panel A, shows that there are 168 unique SKUs in the category. Furthermore, over a period of almost two years of observation, many SKUs are added, dropped, or stocked-out. Table 1, Panel B, indicates that each unique SKU can be described as a combination of single unique levels over the following five attributes: brand, flavor, form, function, and size. A “full assortment” that contains SKUs that span all possible combinations would yield $10 \times 14 \times 3 \times 7 \times 4 = 11,760$ SKUs.

Table 1 highlights several empirical observations, some of which (to our knowledge) appear to have been largely overlooked, yet seem highly germane to model building.

Panel A: The SKU choice set changes over time, with slightly fewer than one-third of the SKUs available for all weeks of the data set.

Panel B: There are many more possible SKU locations—that is, unique combinations of attribute levels (11,760)—than those that existing SKUs (168) occupy.

Panel B: There is heterogeneity across attributes in the extent to which brands “fill in” attribute levels. On the one hand, 65% of brand–size combinations (26 of 40) are occupied, suggesting that, on average, brands come in most sizes. On the other hand, only 20% of brand–flavor combinations (29 of 140) are occupied. Closer inspection of the data shows that the proliferation of levels of flavor occurs because there is relatively less overlap among brands on this dimension. Similarly, most brands offer most forms (18 of 30 possible locations are filled), whereas function has less overlap among brands.

Table 1

<table>
<thead>
<tr>
<th>Inventory Status</th>
<th>Number of SKUs</th>
<th>SKU (%)</th>
<th>Share (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Occasional stock-out</td>
<td>27</td>
<td>16</td>
<td>21</td>
</tr>
<tr>
<td>Established (always on shelf)</td>
<td>22</td>
<td>13</td>
<td>31</td>
</tr>
<tr>
<td>Added during observation period</td>
<td>41</td>
<td>25</td>
<td>20</td>
</tr>
<tr>
<td>Deleted during observation period</td>
<td>49</td>
<td>29</td>
<td>16</td>
</tr>
<tr>
<td>Added then deleted</td>
<td>29</td>
<td>17</td>
<td>12</td>
</tr>
<tr>
<td>Total</td>
<td>168</td>
<td>100</td>
<td>100</td>
</tr>
</tbody>
</table>

Panel B: There are many more possible SKU locations—that is, unique combinations of attribute levels (11,760)—than those that existing SKUs (168) occupy.

Panel B: There is heterogeneity across attributes in the extent to which brands “fill in” attribute levels. On the one hand, 65% of brand–size combinations (26 of 40) are occupied, suggesting that, on average, brands come in most sizes. On the other hand, only 20% of brand–flavor combinations (29 of 140) are occupied. Closer inspection of the data shows that the proliferation of levels of flavor occurs because there is relatively less overlap among brands on this dimension. Similarly, most brands offer most forms (18 of 30 possible locations are filled), whereas function has less overlap among brands.

Table 1

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Flavor</th>
<th>Form</th>
<th>Function</th>
<th>Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of levels</td>
<td>14</td>
<td>3</td>
<td>7</td>
<td>4</td>
</tr>
<tr>
<td>Total possible locations</td>
<td>140</td>
<td>30</td>
<td>70</td>
<td>40</td>
</tr>
<tr>
<td>Locations actually used</td>
<td>29</td>
<td>18</td>
<td>26</td>
<td>26</td>
</tr>
<tr>
<td>Locations used (%)</td>
<td>20</td>
<td>60</td>
<td>37</td>
<td>65</td>
</tr>
</tbody>
</table>

Notes: The total possible locations are equal to ten (number of brands) multiplied by the number of levels for the relevant attribute.
overlap. In summary, flavor and function (compared with form and size) appear to be attributes on which brands try to offer at least some SKUs that differ from those of competitors.

**Panel B:** In combination, the previous two observations suggest that the dimensionality of the estimation problem can be reduced considerably by focusing on attributes rather than on SKUs (see also Fader and Hardie 1996; Ho and Chong 2003).

Standard market share models estimated on the true alternatives dictate a large number of fixed effects. Moreover, the location of SKUs in the attribute space is clearly not random, implying that perhaps there is something to be learned from studying consumer sensitivity to changes in attribute levels and manufacturer choice of SKU locations. We do not address the latter endogeneity issue directly, but we note that it is an important area for future work (see Draganska and Jain 2005).

Table 2 reports Herfindahl indexes that are obtained from attribute-level shares and compares them with a base value of $1/L_a$, where $L_a$ is the number of levels for attribute $a$. The discrepancy between the Herfindahl index and the base value is the greatest for flavor and function. This finding and the data in Table 1 imply a relatively small number of “popular” levels for these two attributes and a relatively large number of levels that have smaller shares. Conversely, the distribution of market share over the different levels of size and form is quite close to the base $1/L_a$ value, which suggests parity across levels within these attributes.

**OVERVIEW AND CONTRIBUTION**

We derive a market share model that takes fundamental properties of the data (Tables 1 and 2) into account, and we provide the following contributions to the literature: First, we show how to recover underlying SKU-level fixed effects from a much more parsimonious share regression on attribute-level shares.\(^4\) For example, we assume that the researcher knows the fixed effects that arise from a market share model specified on brand-level data. In the toothpaste category, the fixed effects represent the mean utilities of, for example, Colgate, Crest, and Arm & Hammer brands. We show how these parameter estimates can be used to generate fixed-effect coefficients for each of the individual SKUs (e.g., medium-size, mint-flavored, gel-form Colgate toothpaste with tartar control) under the respective brand (i.e., Colgate). This is possible for any number of SKUs, and it works for any attribute.

Second, marketing-mix coefficients for an underlying SKU-level model can be recovered from a second-stage regression of transformed shares on appropriately mean-centered marketing-mix elements. Third, the attribute-level coefficients can be used to relate SKU-level response measures (e.g., elasticities) to higher-order attribute-level response measures. An important practical benefit of this approach is that the retailer gains insight into which attributes show the greatest sensitivity to change. We can provide forecasts for new SKUs that may be introduced to the assortment and can outperform current approaches with respect to predictive performance. More significantly, we forecast the price sensitivity of SKUs on the basis of their product characteristics. Thus, retailers and manufacturers can adopt our procedure to improve decisions in category assortment planning and new product design.

We organize the remainder of the article as follows: In the next section, we briefly review the relevant market share modeling literature and develop the specification and demand equations for our approach. Specifically, we explore the analytical relationship between an SKU-level model and an attribute-level model. Then, we describe the data that we used for the empirical application. We present empirical results and conclude with a discussion of the findings, which includes a comparison with an alternative “pure characteristics” modeling procedure.

**BACKGROUND AND MODEL**

Market share modeling has a long tradition that dates back to Cooper and Nakanishi’s (1974, 1982) early work on the multiplicative competitive interaction model and the forecasting models of Brodie and de Kluyver (1984).\(^5\) In recent years, researchers in marketing have focused attention on store-level market share data, in part because managers prefer this data to the household panel data that marketing academics often favor (Bucklin and Gupta 1999).

In some settings, it is reasonable to aggregate market share data to the level of brand and hold the set of brands constant over other dimensions of the data set (e.g., time and/or markets). However, if the goal is to better understand interrelationships among SKUs, it is necessary to take a more disaggregate approach. It is important not to be constrained by techniques that ex ante require aggregation. Zanutto and Bradlow (2004) show that “nonignorable” SKU selection schemes employed in most academic studies lead to severe and undocumented biases in fundamental model outputs, including share predictions, elasticities, and preference estimates.

Fortunately, the question of how SKUs can be represented by levels of their underlying attributes has been the

---

\(^4\)Data for the attribute-based model are generated by aggregation. In a brand-level model, for example, market share for the Colgate brand is defined by adding up the shares of all SKUs that have this brand name (and similarly for Crest, and so on). In a size-level model, shares for all SKUs under the small size, medium size, and so forth, are computed. Therefore, the number of equations in such a model is equal to the number of levels for the attribute in question (in our case, the brand model has ten levels and therefore ten equations per period).

\(^5\)For a more detailed review, see Leeﬂang and colleagues (2000, pp. 171–78).
subject of successful study in the household panel data literature (Fader and Hardie 1996; Ho and Chong 2003), and this work provides a starting point for our method.  Fader and Hardie (1996) stress that consumers choose SKUs—not ad hoc brand-level composites that researchers define—and that retailers and manufacturers also require SKU-level analysis. This is especially true in complex categories in which SKUs are clearly different in their fundamental characteristics (e.g., flavor, form, function, size). In the toothpaste category on which our model is calibrated, we observe considerable variation in SKUs on the shelf at any point in time but no variation over time in the number of brands that remains constant at ten. Figure 1 shows the time variation in the number of SKUs (previously summarized in Table 1).

We begin with a basic market share model that is commonly found in the literature and describe the relationship among fixed-effect parameters that are estimated at different levels of aggregation. In particular, we show how fixed-effect parameter estimates for SKU \(j\) (the most disaggregate unit of analysis), which are obtained from market share equations for all \(j = 1, \ldots, J\) SKUs, are related to fixed-effect estimates for level \(l\) of attribute \(a\) (\(a = 1, \ldots, A\), which is estimated from market share equations for the \(l = 1, \ldots, L_a\) levels of that attribute (where \(L_a\) is much smaller than \(J\); see Table 1). For ease of exposition, we focus on the brand attribute as the higher-level composite.

The following example previews our approach: Brand 1 comes in two SKUs of different sizes (small and medium), and Brand 2 comes in three unique SKUs of different sizes (small, medium, and large). We show how the SKU-level fixed effect for the first SKU, which is estimated from a five-equation system of SKU shares, is related to the brand-level fixed effect for Brand 1, which is estimated from a two-equation system in which the unit of analysis is brand share.

---

Figure 1
TIME SERIES OF NUMBER OF TOOTHPASTE SKUS AVAILABLE IN THE STORE

---

\(\text{Figure 1}\)
TIME SERIES OF NUMBER OF TOOTHPASTE SKUS AVAILABLE IN THE STORE

---

\(\text{Number of SKUs Over Time: Toothpaste Category}\)

---

\(^6\)Although progress has been made in the understanding of market structure and product hierarchies from store-level share data (e.g., Foekens, Leeflang, and Wittink 1997), FH’s (1996) and Ho and Chong’s (2003) household panel data-based methods are closer in spirit to our approach.

\(^7\)Although variation in the number of levels of an attribute may occur, it is relatively rare compared with variations in SKUs that are available in a choice set. The introduction of new products is one example in which this occurs (e.g., Vanilla Coke, where the vanilla flavor does not exist for any other soft drink).

---

\(^8\)Our derivation is based on the increasingly popular “outside good” formulation, but it also applies in standard models that do not use an outside good (e.g., Hardie et al. 1998). We discuss this issue subsequently and thank an anonymous reviewer for drawing it to our attention.

\(^9\)The choice of attribute is immaterial, and the researcher may choose any product attribute as the basis for aggregation. The results in Table 4 are based on the brand attribute, but the full set of identical results for other attributes is available from the authors on request.

---

\(^a\)Week is based on Information Resources Inc. data (http://www.infores.com/).
Fixed Effects

We first show how the parameters from two models—that is, the disaggregate SKU-level share model and the brand composite model—are related when each model contains only fixed effects. The goal is to demonstrate how SKU-level fixed effects may be recovered analytically when only the attribute-level model is estimated. As we noted previously, we focus on brand as the attribute over which aggregation takes place. This is for exposition only, and we do so without loss of generality. Subsequently, we extend the analysis and include the marketing-mix covariates price, feature, and display, and we discuss how to expand the set of independent variables further.

Fixed effects: Attribute-level parameters from SKU-level parameters. Following the standard approach in the literature, market share for SKU \( j \) at time \( t \) is

\[
s_j(t) = \frac{\exp \mu_j}{1 + \sum_k \exp \mu_{kt}},
\]

where \( \mu_j \) is the SKU-specific utility relative to an outside good. In general, \( \mu_j \) can be a function of fixed effects for SKU \( j \) (\( \beta_j \)) and time-dependent covariates and their associated response parameters. The share of the so-called outside good is given by

\[
s_{0t} = \frac{1}{1 + \sum_k \exp \mu_{kt}}.
\]

Parameter estimation is facilitated by a linearizing transformation (Besanko, Gupta, and Jain 1998):

\[
\log(s_j(t)) - \log(s_{0t}) = \mu_j.
\]

The parameters of \( \mu_j \) can be estimated with ordinary least squares (OLS) under the assumption that the model also contains a structural error term \( \varepsilon_j \) that now takes on the role of the error in the regression equation. Level-specific intercepts for a particular attribute \( a \) and level \( l \) (\( \psi_l^a \)) can be represented in a similar fashion. To observe this, continue with the previous example in which SKUs 1 and 2 share the same brand name; that is, they both belong to the same level of the brand attribute (Brand 1). In general, \( S^b_l \) denotes the market share of level \( l \) of attribute \( a \) at time \( t \). For ease of exposition (and without loss of generality), we develop the model around the brand attribute, which facilitates simpler notation. Thus, \( S^b_1 \) is total market share of brand \( l \) at time \( t \). Continuing our illustrative example, the market share of Brand 1 (i.e., Level 1 of the brand attribute) is

\[
S^b_1 = s_{1t} + s_{2t} = \frac{\exp \mu_{1t} + \exp \mu_{2t}}{1 + \sum_k \exp \mu_{kt}}.
\]

The brand share specification and the SKU share specification use the same outside good value:

\[
S^b_{0t} = s_{0t} = \frac{1}{1 + \sum_k \exp \mu_{kt}}.
\]

Applying the log transformation to the brand-level model in Equation 5, we obtain the brand-specific mean utility value for Level 1 of the brand attribute (shared by SKUs 1 and 2):

\[
\psi^b_l = \log(S^b_{1t}) - \log(S^b_{0t}) = \log(\exp \mu_{1t} + \exp \mu_{2t}).
\]

The question of interest is, Can SKU-specific parameters be recovered from the attribute share equations? Before addressing this issue, we determine whether the reverse is possible: Can level-specific parameters for the brand attribute be recovered from the SKU share equations? Before answering in the affirmative, we assume that SKU-level utility is given by \( \mu_j = \beta_j + \varepsilon_j \). In our example, in which SKU 1 and SKU 2 belong to Brand 1, the following calculation recovers the brand-specific intercept for the attribute share model using only the parameters and error terms from the SKU-specific model as inputs

\[
(8) \quad \psi^b_l = \alpha^b_l + \nu^l_l = \log(\exp \beta_1 \exp \varepsilon_{1t} + \exp \beta_2 \exp \varepsilon_{2t}),
\]

where \( \nu^b_l \) is the level-specific error, and

\[
(9) \quad E(\psi^b_l) = \frac{1}{T} \sum_{t=1}^{T} (\alpha^b_l + \nu^l_l),
\]

\[
(10) \quad = \frac{1}{T} \sum_{t=1}^{T} \log(\exp \beta_1 \exp \varepsilon_{1t} + \exp \beta_2 \exp \varepsilon_{2t}),
\]

\[
= \alpha^b_1.
\]

To confirm that this relation holds, we obtain \( \beta_1 \) and \( \beta_2 \) and \( \alpha^b_l \) directly from OLS estimation on the SKU-level and brand-level market share equations, respectively. We also compute estimates of OLS residuals \( \exp(\varepsilon_1t) \) and \( \exp(\varepsilon_2t) \) from the item-level models. Using a simulated data set, we computed these values for the right-hand side of Equation 10, and we were able to recover perfectly (to four decimal places) the brand-level intercepts \( \alpha^b_1 \) that were estimated from the brand share model, thus demonstrating this link (we give details in the Appendix; see “Model Simulations”). Note that this procedure explicitly requires the SKU-level parameter estimates \( \beta_j \) and the residuals \( \varepsilon_j \) to recover the brand-level parameters.

Fixed effects: SKU-level parameters from attribute-level parameters. The next challenge is to recover SKU-specific intercepts by calculation, using only estimates from the attribute share model. For ease of exposition and without loss of generality, we focus on the brand attribute and assume all levels of this attribute are represented for all periods of the data.\(^{10}\) Moreover, we note that the following information is available after estimation of the attribute share models:

- The estimated attribute-model intercepts for the brand levels, \( \alpha^b_l, l = 1, ..., L^b; \)
- The design matrix of the SKUs (i.e., the description of each SKU as a combination of single unique levels for each attribute);
- The error components of the data estimated from the attribute share models;

\(^{10}\) This is much less stringent than requiring that all SKUs are available in each and every period, which is clearly not true in practice (see Table 1 and Figure 1).
• The outside good value; and
• The market share of each of the SKUs and the attribute share for each attribute level.

As we show subsequently, our method requires only the first and last point from the preceding list to recover the SKU-level fixed effects by calculation. Continuing with our example, we have
\[
\exp \beta_i^b = \exp \mu_i^b + \exp \mu_{z1}^b.
\]

In Equation 11, only two items enter the right-hand side sum, but in general, the number of elements on the right-hand side is equal to the number of SKUs that belong to the related level of a particular attribute (e.g., the number of toothpaste SKUs that share the Colgate brand). Let \( D_i = 1 + \sum_{j=1}^{N_i^b} \exp \mu_{ij} \) denote the denominator of the brand and SKU share models, and note that \( D_i \) can be rewritten as
\[
D_i = \frac{\exp \beta_i^b \exp \nu_{z1}^b}{S_h^b}, \text{ or in general as } D_i = \frac{\exp \beta_i^b \exp \nu_{z1}^b}{S_h^b}.
\]

Exponentiating the mean utility for any product (whether the outside good, a single SKU, or a composite such as a brand) and dividing it by the associated market share yields \( D_i \). That is, the numerator of this expression is a function of quantities that are obtained from estimation, and the denominator is simply observed in the data. Next, using the definition of SKU share \( s_{jt} \) from Equation 1, we recognize that the product of the exponentiated (and unobserved) SKU-level intercept and error component for SKU \( j \) under brand \( l \) is equal to the product of \( D_i \) and the SKU-level share
\[
\exp \beta_j \exp \epsilon_{jt} = D_i \times s_{jt}.
\]

Taking the log of both sides shows that the SKU-level intercept is related to the brand-level intercept as follows:
\[
\beta_j + \epsilon_{jt} = \alpha_l^b + \nu_l^b - \log S_h^b + \log s_{jt}.
\]

To obtain \( \beta_j \) by calculation, we take the average over \( t \), which yields
\[
\beta_j = \frac{1}{T} \sum_{t=1}^{T} \left( \alpha_l^b + \nu_l^b - \log S_h^b + \log s_{jt} \right).
\]

The last line of Equation 15 is simply the sum of the estimated brand intercept and the average log share of the SKU relative to the share of the parent brand. Because the denominator is weakly greater than the numerator, the log sum is always weakly negative. (Note that when brand \( l \) has only one SKU variant, the log sum is zero, and the brand-specific fixed effect \( \alpha_l^b \) equals the SKU-specific fixed effect \( \beta_j \).) Alternatively, we could rewrite Equation 15 as
\[
\beta_j = \alpha_l^b + \frac{1}{T} \sum_{t=1}^{T} \log \left( \frac{S_h^b}{s_{jt}} \right).
\]

Through the use of simulated data, the SKU intercepts calculated in this manner (see “Model Simulations” in the Appendix) reproduce the directly estimated SKU intercepts almost exactly.\(^{11}\)

**Adding Covariate Effects**

We now demonstrate how to recover covariate effects. Although we focus only on the usual marketing-mix effects (i.e., price, feature, and display), our model is easily extended to any number of covariates (e.g., interactions, lagged shares). As we show subsequently, all that is required is that the data for each covariate are mean centered at the appropriate level of aggregation. In circumstances in which this is not possible initially, it may still be accomplished by recoding the independent variable. For example, if the researcher wishes to specify an ordinal variable for a feature with \( k \) levels (e.g., Feature A, Feature B, Feature C), this can be achieved by creating \( k - 1 \) dummy variables. In the process of aggregation, the value of the independent variable for each SKU is share weighted to produce an attribute-level variable that is continuous between zero and one, so that mean centering is still possible in this scheme.

For ease of exposition, we introduce only the price covariate; however, the empirical results also include feature and display. The introduction of covariates complicates matters, beginning with recovery by calculation of the SKU-level fixed effects \( \beta_j \), because the mean utilities for brand composite \( l \) and SKU \( j \) now become
\[
\psi_{jt}^b = \alpha_l^b + \gamma_l^b p_{jt}^b + \nu_l^b,
\]
and
\[
\mu_{jt} = \beta_j^b + \gamma j^b p_{jt} + \epsilon_{jt},
\]
where
\[
p_{jt}^b = \frac{1}{N_j^b} \sum_{t=1}^{N_j^b} \frac{s_{jt}^b}{S_h^b} p_{jt}.
\]

is the SKU-share-weighted price for the brand composite. The weighting is over the \( N_j^b \) SKUs that belong to brand \( l \).\(^{12}\) These changes affect the derivation that leads to Equation 15, which described recovery of fixed effects when both the SKU-level and attribute-level models did not contain covariates. The parameters \( \gamma \) and \( \gamma j \) are the (true) SKU and brand-level price response parameters, respectively; that is, \( \gamma \) is the parameter that would be obtained if an SKU-specific market share model were estimated. We now ask the following questions: What is the relationship between \( \beta_j \) and \( \alpha_l^b \) when both models include price covariates? What is

\(^{11}\)Calculated parameters are equal to estimated parameters to more than ten decimal places for all estimates. We also tested the robustness of this relation given the presence of interactions among the attributes (e.g., Colgate makes a better tartar control formulation than does Crest) and found that Equation 15 is still valid in this case.

\(^{12}\)A potential issue here is that of correlation among covariates (e.g., price, feature, display) and how aggregation might affect such correlations. Because all weighted average variables (e.g., the weighted average price shown previously) are computed with the same weights, the correlation among aggregates would equal the weighted sum of the correlations of items that constitute the aggregates. This value should be less than the highest correlation across items within the aggregate, so that if there is no problem with correlation at the SKU level, there will be no problem at the aggregate level. However, if correlation is to be mitigated, the method of defining new variables that Van Heerde, Leefflang, and Wittink (2005) describe can be applied. We thank an anonymous reviewer for drawing our attention to this point.
the relationship between $\gamma^b$ and $\gamma$? How do we recover the true $\gamma$ from the estimate of the attribute-level counterpart $\gamma^b$? Following the logic that leads to Equation 12, we can still write, for SKU $j$,

$$\exp \beta_j \exp \gamma^b \exp \epsilon_j = D_t \times s_j;$$

however, in this case, $D_t$ is given by

$$D_t = \frac{\exp \alpha_j^b \exp \gamma^b \exp \gamma \exp \epsilon_j}{S_k^b}.$$  

After taking logs and rearranging terms, we find that

$$\beta_j + \gamma p_j + \epsilon_j = \alpha_j^b + \gamma^b p^b_j + \epsilon_j + \log s_j - \log S_k^b,$$

so

$$\beta_j = \alpha_j^b + (\gamma^b p^b_j - \gamma p_j) + (\epsilon_j + \log s_j - \log S_k^b).$$

As we did previously, the fixed-effect values can be computed by taking averages over time:

$$\bar{\beta}_j = \frac{\alpha_j^b + \gamma^b p^b_j - \gamma p^*_j + 1}{T} \log s_j - \frac{T}{T} \log S_k^b,$$

where the double bar indicates the average of the share-weighted price $p^*_j$ taken over time $t = 1, \ldots, T$. We have been careful to distinguish what can be observed in the data, what can be estimated, and what quantities are unknown (and must be calculated). Unfortunately, Equation 21 shows that there is only one equation for each of the $j$ SKUs and two unknowns (price effect $\gamma$ and fixed effect $\beta_j$). We address this issue subsequently; however, note that the preceding expression simplifies to

$$\bar{\beta}_j = \alpha_j^b + \gamma^b p^b_j - \gamma p^*_j + \frac{1}{T} \sum_{t=1}^{T} \log \left( \frac{s^*_j}{S_k^b} \right).$$

To understand how the identification problem caused by the introduction of covariates complicates matters, consider the following example: Using simulated data, we again estimated the SKU-level and brand-level intercepts from the respective models, but this time, we included the marketing-mix variables. We found that calculating the SKU-level intercepts on the basis of the relationship we derived previously in Equation 15 (for models that do not have covariates) and ignoring the effect of price (and, in general, other covariates) that we specified in Equation 21 produces fixed-effect estimates that are consistently inaccurate. In a simulation of 1000 observations (see “Model Simulations” in the Appendix), the mean absolute difference between such estimates and the true estimates was approximately 1.95. In addition, Equation 21 is not directly useful, because the price and fixed effects cannot be separately identified. Fortunately, we now demonstrate that both problems can be circumvented such that the true SKU-level fixed effects $\bar{\beta}_j$ can be recovered from an appropriately reformulated model. In addition, we show that a second-stage regression can be used to recover the other parameter of interest—that is, the unknown price effect $\gamma$.

Recovering fixed effects: model reformulation with mean-centered covariates. Continuing our discussion with price as the covariate, we reformulate the mean utilities of the attribute and SKU-level models by including mean-centered prices rather than actual prices:

$$\psi^b_h = \alpha_j^b + \gamma (p^*_h - p_j^b) + \epsilon_j^b,$$

and

$$\mu_j = \beta_j + (p^*_j - p_j^b) + \epsilon_j.$$  

where $\alpha_j^b$ and $\beta_j$ are the fixed effects for each model. Mean centering occurs at the appropriate level of aggregation, with respect to the attribute-level average for each attribute level and the SKU-level average for the corresponding SKU. It is important to note that the covariate effect $\gamma^b$ is completely unaffected by this change (Raudenbush and Bryk 2002). This can be observed when Equation 23 is rewritten as follows:

$$\psi^b_h = \alpha_j^b - \gamma p^*_h + \gamma^b p^*_h + \epsilon_j^b,$$

where $\alpha_j^b$ is the fixed-effect estimate that results when the covariate has not been mean centered. The difference between $\alpha_j^b$ and $\alpha_j^b$ is the price estimate for the attribute multiplied by the average price for that level (and analogously for the SKU-level model). It is still the case that (1) the SKU-level and attribute-level price effects are not equal (i.e., $\gamma \neq \gamma^b$) and (2) the fixed effects $\bar{\beta}_j$ must be calculated from estimated values from the attribute-level models ($\alpha_j^b$, $\gamma^b$) and average log market shares and prices ($p^*_h$, $p^b_j$).

Using Equation 18, but with the mean-centered covariates, we obtain

$$\beta_j + (p^*_j - p_j^b) + \epsilon_j = \alpha_j^b + \gamma^b (p_j^b - p^*_h) + \epsilon_j^b + \log s_j - \log S_k^b.$$  

Taking expectations with respect to time, we observe

$$1/T \sum_{t=1}^{T} (p_j^b - p^*_h) = 0$$

which implies that SKU-level intercepts can be calculated from

$$E(\beta_j) = \alpha_j^b + \frac{1}{T} \sum_{t=1}^{T} \log \left( \frac{s_j}{S_k^b} \right).$$

The analytical form of Equation 25 is identical to that which we obtained for the model that has only fixed effects. However, we mean centered the included marketing-mix covariates before estimation. We tested this relation with simulated data (see “Model Simulations” in the Appendix), and we were able to recover the SKU-level intercepts from the brand-level intercepts, just as we did for the models without covariates. In addition, the true SKU-level fixed effects can be recovered with any attribute that the researcher desires to use as a basis for aggregation and with any number of appropriately mean-centered covariates in the model.

Recovering covariate effects: second-stage regression analysis. We have shown how true SKU-level fixed effects can be recovered by means of a parsimonious attribute-level model. However, recovery of the covariate effect ($\gamma$ in our...
example) is still complicated because though mean centering solved the identification problem for the $\beta$ parameters, it introduced an indeterminacy in process: $\gamma \times 0 = \gamma b \times 0$. Indeed, it was the introduction of this indeterminacy that made recovery of the $\beta$ possible. It is also the case that knowledge of $\gamma b$ (or any $\gamma a$ for attribute a, alone or in combination with estimates for other attributes) tells us nothing by itself about the true price effect $\gamma$.

However, we can obtain an estimate of the true price effect without recourse to an SKU-level model through the following second-stage regression. First, having estimated the attribute-level model, we calculate the SKU-level fixed $\beta$ effects using Equation 25. Second, we subtract these calculated SKU-specific fixed effects from log-transformed and share-differenced information that enters the SKU-level regression model. We define $\gamma j_t = \log(s_{jt}) - \log(s_{0t})$ to be this variable. We then use SKU-level data for each SKU $j$ and time period $t$ and write the following regression equation:

$$
(\gamma j - \beta j) = \log(s_{jt}) - \log(s_{0t}) - \left( \alpha b + \frac{1}{T} \sum_{t=1}^{T} \log s_{jt} \right),
$$

or

$$
\gamma j = \gamma p_{mc} + \epsilon j,
$$

where $p_{mc}$ is the mean-centered price. The OLS estimation retrieves the same $\gamma$ as that obtained from a regression model based on the original SKU-level market shares.\(^\text{14}\)

**Summary**

In the preceding sections, we developed the analytical and empirical relationships that exist between market share models estimated at different levels of aggregation. Specifically, both the underlying SKU-specific fixed effects and the response parameters for covariates in an SKU-level model can be recovered by calculation by means of a parsimonious model estimated at a higher level of aggregation. The implications for researchers and practitioners are powerful: There is no need to resort to arbitrary aggregation schemes, to disregard data, or to focus only on categories with a relatively small number of alternatives. The analyst needs only to delineate the product category accurately and identify and describe the mutually exclusive and collectively exhaustive set of attributes and levels. Our formulation allows the researcher to be true to the properties of the underlying data and to implement models in a straightforward fashion with the use of standard software. In the next sections, we illustrate our procedure in an empirical analysis of the toothpaste category. We show that attribute-level equations are not only useful as a methodological shortcut but also have potential as tools for delivering substantive insights into consumer behavior.

**DATA**

We analyze toothpaste sales data from a single store in the Stanford Market Basket Database.\(^\text{15}\) It is well known that retailers use some categories to drive store choice decisions (Chen et al. 1999) and that toothpaste is a category that is typically not associated with traffic-building objectives (Drèze and Hoch 1998). Because we require an estimate of the number of visitors to the store to compute the outside good, it helps that the number of visitors coming to the store is not a function of activity in the toothpaste category itself. In addition, the toothpaste category exhibits a good amount of variety and large changes in the total number of SKUs available as a result of stock-outs, product introductions, and deletions (see Table 1 and Figure 1). Collectively, these features make it an ideal candidate category for calibrating our new approach. Our model analyzes market shares and is conditional on consumers’ decisions to choose the store and then the SKU offerings that are available in a particular week.

**Raw Data**

**SKU set and outside good.** In assembling the data for the study, we used information from Euromonitor, the IRI Marketing Fact Book, and the Stanford Market Basket Database.\(^\text{16}\) The toothpaste sales data come from a large supermarket retailer that operates in Chicago and cover 104 weeks (Week 1 is the week beginning June 1, 1991, and Week 104 is the week beginning May 22, 1993). All prices are in September 1991 dollars, and the original price series were deflated by means of the Bureau of Labor Statistics’ Consumer Price Index. Sales are in ounces, and prices are price per ounce of toothpaste.

As we noted previously, a key advantage of our approach is that we do not need to discard SKUs arbitrarily. However, we must be able to define the product category properly and to identify the mutually exclusive and collectively exhaustive group of attributes and levels that describe the individual SKUs. With this objective in mind, we must eliminate a small number of SKUs that exist in our raw data but do not fit the proper definition of “toothpaste.” The trimming process proceeded as follows: We began by selecting all 246 SKUs in the Information Resources Inc. “tooth cleanser” product category that sold in the store at some point during the observation period. From this group, we deleted 40 SKUs for which we had no attribute information, which left 206 SKUs.\(^\text{17}\) We removed another 35 SKUs for which we had category information but which were problematic because they were, for example, nontoothpaste tooth cleansers (e.g., liquid cleansers, teeth whiteners). Their combined category share was less than 2%. Finally, we removed the brand Topol, which is not a part of the regular toothpaste category. This small share brand is designed for smokers (and therefore contains a unique attribute—nicotine stain combatant—that is not shared by any other SKU). The end result is a choice set of 168 unique SKUs. On average, 95 of these are available on the shelf at any one time. This discrepancy speaks to the diversity in the category in terms of rate of entry and exit of SKUs and the prevalence of stock-outs.

We follow the standard approach (Nevo 2000) to estimate the value of outside good: Any shopper who visited the

\(^{14}\)This approach has implications for the standard errors (see “Standard Errors” in the Appendix).

\(^{15}\)These data have been used in several studies published in the marketing and economics literature (see Bell and Lattin 1998); for additional details and documentation, visit http://wrds.wharton.upenn.edu. The data set contains both household panel data and store-level data for each product category.

\(^{16}\)The full data set, documentation, and MATLAB code for estimation are available from the authors on request.

\(^{17}\)All but one of these SKUs had less than .3% category share overall.
store in a particular week is part of the “market potential.” Fortunately, the data set contains weekly information on the number of panelists visiting the store and total sales for each toothpaste SKU. We assume that the purchase behavior of the sample of panelists who visit the store represents that of the entire population of shoppers. We obtained the total amount spent by an average customer (average market basket dollar value per week) from panelist sample means. Dividing total weekly dollar sales by this value generates an estimate for the total number of visitors that week. We further assume that if a shopper decides to purchase, they purchase an average-sized pack. The estimated average number of visitors to the store is 10,288 (standard deviation = 1717), and the estimated average value for the outside good is 15,123 ounces (standard deviation = 1199).

**EMPIRICAL RESULTS**

Our analysis proceeds in two parts. First, we estimate the attribute share models for all five attributes (brand, flavor, form, function, and size) and report the model fits and parameter estimates for the fixed effects and the marketing-mix parameters. Second, we show how these estimates can be used to recover by calculation the underlying SKU-level model parameters. We discuss the potential implications of our empirical findings and provide possible substantive applications of our method in next section.

**Attribute Share Models**

Table 3 summarizes the findings from the five attribute share models. In accordance with our analytical development, we estimate the models with appropriately mean-centered marketing-mix variables. Because our panel data are collected over time, we mean center the time-specific marketing-mix variable for an attribute level with respect to the grand mean for that attribute level. If we do not mean center, we would be unable to recover the SKU-level parameters (we provide details subsequently). The numbers of parameters and observations differ across attribute-specific models in accordance with the number of levels, $L_a$, contained in each attribute a. For example, the brand model contains ten levels and therefore 980 observations (all models are estimated over 98 weeks of data) with a total of 13 parameters (10 fixed effects and 3 marketing-mix effects).

For ease of exposition, we do not report all attribute-level fixed effects (these are available on request); we simply summarize them according to their minimum, maximum, and median values. The model fit is good in all cases, with adjusted $R^2$ values ranging from .90 for the flavor model to .94 for the brand model, and all fixed effects are statistically significantly different from zero. The price response parameter varies across attributes and provides some indication of differential price sensitivity, with form appearing to be the most price-responsive attribute. The coefficients for feature and display have the expected positive signs. However, only the display parameters are significant. This is not surprising given that toothpaste is unlikely to be used as a traffic-building category and that some purchasing in the category is presumably spontaneous and opportunistic (Bucklin and Lattin 1991).

**Calculation of SKU-Level Effects**

Table 4 presents a comparison of the estimated SKU-level coefficients from the SKU-level model and the calculated SKU-level coefficients from the attribute-level model estimates in Table 3. For ease of illustration, we present the results from the brand attribute only (as we note at the bottom of Table 4, results for other attributes are identical). We find the following:

- **Fixed-effect recovery.** We summarize both the estimated and calculated 168 SKU-level intercepts using the minimum, maximum, median, and mean values (the full set of parameters are available on request). The average squared difference between SKU-specific estimates from the OLS SKU-level model and those generated by calculation from the attribute (brand) model is .00014, implying SKU fixed effects recovered by calculation are essentially identical to those obtained by direct estimation. The correlation between the two sets of parameters is close to one ($r = .99991$).

- **Marketing-mix effect recovery.** The attribute-level model recovers the price, feature, and display effects exactly. We recover the parameters in two steps. We enter calculated fixed effects into a second-stage regression of log-differenced SKU shares less calculated fixed effects on mean-centered SKU-level price, feature, and display variables (see Equation 26). Again, researchers can choose any attribute as the basis for aggregation and still recover the SKU-level parameters.

- **Fit and precision.** The adjusted $R^2$ value for the SKU-level model is .62. This is below that which we obtained for the

---

Table 3

**OLS ESTIMATION RESULTS FOR ATTRIBUTE-LEVEL SHARE MODELS**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Brand</th>
<th>Flavor</th>
<th>Form</th>
<th>Function</th>
<th>Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept (minimum)</td>
<td>-6.87</td>
<td>-7.74</td>
<td>-4.39</td>
<td>-8.07</td>
<td>-5.98</td>
</tr>
<tr>
<td></td>
<td>(.04)</td>
<td>(.04)</td>
<td>(.03)</td>
<td>(.05)</td>
<td>(.04)</td>
</tr>
<tr>
<td>Intercept (maximum)</td>
<td>-2.92</td>
<td>-2.42</td>
<td>-2.43</td>
<td>-2.72</td>
<td>-2.60</td>
</tr>
<tr>
<td></td>
<td>(.04)</td>
<td>(.04)</td>
<td>(.03)</td>
<td>(.05)</td>
<td>(.04)</td>
</tr>
<tr>
<td>Intercept (median)</td>
<td>-5.16</td>
<td>-5.54</td>
<td>-2.94</td>
<td>-4.83</td>
<td>-3.29</td>
</tr>
<tr>
<td></td>
<td>(.04)</td>
<td>(.04)</td>
<td>(.03)</td>
<td>(.05)</td>
<td>(.04)</td>
</tr>
<tr>
<td>Price</td>
<td>-6.18</td>
<td>-8.74</td>
<td>-11.50</td>
<td>-7.62</td>
<td>-2.17</td>
</tr>
<tr>
<td></td>
<td>(.35)</td>
<td>(.66)</td>
<td>(.83)</td>
<td>(.50)</td>
<td>(.38)</td>
</tr>
<tr>
<td>Feature</td>
<td>.38</td>
<td>.67</td>
<td>.59</td>
<td>.70</td>
<td>.88</td>
</tr>
<tr>
<td></td>
<td>(.38)</td>
<td>(.68)</td>
<td>(.32)</td>
<td>(.70)</td>
<td>(.36)</td>
</tr>
<tr>
<td>Display</td>
<td>2.45</td>
<td>3.51</td>
<td>1.59</td>
<td>3.09</td>
<td>1.96</td>
</tr>
<tr>
<td></td>
<td>(.33)</td>
<td>(.49)</td>
<td>(.74)</td>
<td>(.52)</td>
<td>(.26)</td>
</tr>
<tr>
<td>Adjusted $R^2$</td>
<td>.94</td>
<td>.90</td>
<td>.93</td>
<td>.94</td>
<td>.94</td>
</tr>
<tr>
<td>$N (L_a \times T)$</td>
<td>980</td>
<td>995</td>
<td>294</td>
<td>682</td>
<td>392</td>
</tr>
</tbody>
</table>

*a*All levels for all periods are present for brand, form and size.
attribute-level models (see Table 3), suggesting considerably more unexplained variation in SKUs than in attribute levels. The adjusted $R^2$ value for the regression we used to obtain the marketing-mix effects (Equation 26) is .21. The SKU-level model produces smaller standard errors, suggesting that even though parameters from our model are best linear unbiased predictors, they are less efficient (see “Standard Errors” in the Appendix).

Attribute-level price elasticities. Our model allows for the computation of price elasticities for levels of product attributes and for the attributes themselves. For example, given our model specification, the average own-price elasticities for product attributes are as follows: brand (–2.19), flavor (–2.73), form (–3.36), function (–3.03), and size (–.78). Because our focus is on parameter recovery (rather than on substantive insights), we do not pursue this further but defer such considerations to additional research. A more thorough investigation of own- and cross-price effects for levels of attributes should pay closer attention to model specification issues, such as lag and lead effects and other issues of dynamics (see Van Heerde, Lee, and Wittink 2003).

DISCUSSION AND CONCLUSION

Researchers and practitioners alike have devoted considerable time to the development of market share models. An important challenge in this effort is the identification of patterns of choice and competition that enables inferences to be drawn about market structure and market response. Unfortunately, in many cases, researchers are forced to make data-pruning or aggregation decisions that facilitate model estimation but are not necessarily true to the underlying structure of the data. We draw on FH’s (1996) and Ho and Chong’s (2003) insights of household panel data models—namely, several SKUs can be represented simply as combinations of a small number of product attributes. From this, we develop a model that exploits analytical relationships between disaggregate SKU-level parameters and composite attribute-level parameters.

We recover the SKU-level parameters using estimates from parsimonious attribute-level market share models in combination with market share data. Parameter recovery occurs without recourse to the estimation of the underlying SKU-level model. The primary contribution is methodological. However, the attribute-level results can further the understanding of how demand responds to category characteristics.

Attribute-Level and SKU-Level Price Elasticities

Russell and Bolton (1988) derive the relationship between disaggregate price elasticities and higher-level “submarket” price elasticities. They show analytically that given a proper definition of submarkets, a hierarchy of own- and cross-price elasticities can be derived. In our case, the large number of SKUs necessarily implies a large number of potential submarkets, so our approach to uncovering the relationship between attribute-level elasticities and SKU-level elasticities is empirical. We compute SKU-level elasticities on the basis of parameters we recovered by our analytical procedure, and we regress these on the appropriate attribute-level elasticities. That is, we regress the elasticity for the SKU with the features Colgate (brand), mint (flavor), gel (form), tartar control (function), and medium (size) on the attribute-level elasticities for Colgate, mint, gel, tartar control, and medium. The estimated coefficients from this regression represent the importance weights for the attributes in determining the overall SKU-level elasticity.

This procedure can be applied to estimate likely price response to new SKUs that are not currently part of the category. To demonstrate this approach, we selected eight new SKUs that were introduced in the last 17 weeks of the data, and we calibrated the model using the first 87 weeks of data. The new SKUs included five by Colgate, two by Crest, and one by Arm & Hammer. We are able to recover the SKU-level price elasticities in the holdout sample fairly well (a comparison of “actual” and predicted elasticities yields a mean absolute deviation [MAD] of .45 and a mean square error [MSE] of .05). The mean of the predicted elasticities is $-1.89$ (standard deviation = .34), whereas the mean and standard deviation for the estimation sample of 168 SKUs is $-1.95$ and .60, respectively.

Comparison with FH’s Pure Characteristics Approach (FHPC)

The FH household panel data model (a pure characteristics or “conjoint meets logit” approach) has generated a lot of interest among both academic researchers and practitioners. This is due to the parsimony of the formulation and the simplicity of estimation. Individual SKUs are expressed as additive linear functions of parameter values for levels of attributes that describe the SKU. A reasonable question is, Why not implement a direct replication of the pure characteristics model on market share data (i.e., regress log differentiated market shares on dummy variables for the levels of the different attributes)? In Table 5, we report the part-worths from this approach. Note that our model and the FHPC are comparable, and neither includes consumer behavior dynamics, such as state dependence, which are typically incorporated into household-level models esti-

---

Table 4

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Median</th>
<th>Mean</th>
<th>Price</th>
<th>Feature</th>
<th>Display</th>
</tr>
</thead>
<tbody>
<tr>
<td>Estimated</td>
<td>-9.08</td>
<td>-4.90</td>
<td>-7.09</td>
<td>-7.11</td>
<td>-5.47</td>
<td>7.94</td>
<td>28.58</td>
</tr>
<tr>
<td>Calculated</td>
<td>(0.08)</td>
<td>(0.12)</td>
<td>(0.94)</td>
<td>(0.99)</td>
<td>(2.50)</td>
<td>(15.66)</td>
<td>(7.83)</td>
</tr>
</tbody>
</table>

*Mean-square difference between parameters estimated from the SKU-level model and those calculated from the attribute-level model is .000144. As we noted in the text, using other attributes (i.e., flavor, form, function, and size) for the parameter recovery produces identical results.*
### Table 5
RESULTS FOR THE FHPC MODEL

<table>
<thead>
<tr>
<th>Marketing Mix Covariate</th>
<th>Non-Mean-Centered Covariates</th>
<th>Mean-Centered Covariates</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Estimate</td>
<td>Standard Error</td>
</tr>
<tr>
<td>Marketing mix</td>
<td>Price</td>
<td>–5.45</td>
</tr>
<tr>
<td></td>
<td>Feature</td>
<td>7.50</td>
</tr>
<tr>
<td></td>
<td>Display</td>
<td>31.99</td>
</tr>
<tr>
<td>Attribute Level</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Brand</td>
<td>Aim</td>
<td>–.31</td>
</tr>
<tr>
<td></td>
<td>Aquafresh</td>
<td>–.95</td>
</tr>
<tr>
<td></td>
<td>Arm &amp; Hammer</td>
<td>.36</td>
</tr>
<tr>
<td></td>
<td>Close Up</td>
<td>–.09</td>
</tr>
<tr>
<td></td>
<td>Colgate</td>
<td>.55</td>
</tr>
<tr>
<td></td>
<td>Crest</td>
<td>.33</td>
</tr>
<tr>
<td></td>
<td>Gleem</td>
<td>–.50</td>
</tr>
<tr>
<td></td>
<td>Pepsodent</td>
<td>–.68</td>
</tr>
<tr>
<td>Flavor</td>
<td>Sensodyne</td>
<td>–.20</td>
</tr>
<tr>
<td></td>
<td>Bubble gum</td>
<td>–.11</td>
</tr>
<tr>
<td></td>
<td>Bubble mint</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td>Cool mint</td>
<td>–.47</td>
</tr>
<tr>
<td></td>
<td>Extra fresh</td>
<td>–.50</td>
</tr>
<tr>
<td></td>
<td>Fresh mint</td>
<td>.07</td>
</tr>
<tr>
<td></td>
<td>Mint</td>
<td>–.27</td>
</tr>
<tr>
<td></td>
<td>Original</td>
<td>–.10</td>
</tr>
<tr>
<td></td>
<td>Regular</td>
<td>–.18</td>
</tr>
<tr>
<td></td>
<td>Winter fresh</td>
<td>–.23</td>
</tr>
<tr>
<td></td>
<td>Clean mint</td>
<td>–.09</td>
</tr>
<tr>
<td></td>
<td>Natural mint</td>
<td>–.87</td>
</tr>
<tr>
<td></td>
<td>Hawaiian punch</td>
<td>–.33</td>
</tr>
<tr>
<td></td>
<td>Orange bubble gum</td>
<td>.02</td>
</tr>
<tr>
<td>Form</td>
<td>Gel</td>
<td>–.50</td>
</tr>
<tr>
<td></td>
<td>Paste</td>
<td>–.25</td>
</tr>
<tr>
<td>Function</td>
<td>Anticavity</td>
<td>–.49</td>
</tr>
<tr>
<td></td>
<td>Antiplaque</td>
<td>–.43</td>
</tr>
<tr>
<td></td>
<td>Regular</td>
<td>–.40</td>
</tr>
<tr>
<td></td>
<td>Tartar control</td>
<td>–.43</td>
</tr>
<tr>
<td></td>
<td>Triple protection</td>
<td>–.45</td>
</tr>
<tr>
<td></td>
<td>Antiplaque/tartar</td>
<td>–.427</td>
</tr>
<tr>
<td>Size</td>
<td>Small</td>
<td>–.77</td>
</tr>
<tr>
<td></td>
<td>Medium</td>
<td>–.65</td>
</tr>
<tr>
<td></td>
<td>Large</td>
<td>–.20</td>
</tr>
<tr>
<td></td>
<td>Adjusted R²</td>
<td>.48</td>
</tr>
</tbody>
</table>

Notes: Ultrabrite, smooth mint, gel and paste, sensitive, and extra large are normalized to zero.

The results based on data with mean-centered marketing-mix variables enable an exact replication of the price, feature, and display variables we obtained from our model (see Table 4).

Comparing the results from FHPC and those from our model, we find the following:

- **Fixed effects**: The attribute-level parameter estimates (or part-worths) are consistently overestimated when marketing-mix covariates are not mean centered. Furthermore, there is a relatively substantial difference between the true SKU fixed effects and those that the FHPC model imply. In Table 4, we show that our analytical recovery procedure reproduces the fixed effects exactly (MSE = .00014). Under the pure characteristics approach, the comparable MSE is .50. To be fair, the FHPC model imposes structure on the relationship between attribute levels and the SKU fixed effects, so this finding might be expected ex ante.

- **Attribute-level marketing-mix response**: As we show in Table 3, our model enables the examination of attribute-level marketing-mix sensitivity, whereas a pure characteristics approach does not enable direct computation of such effects. As an illustration, we generate forecasts of price sensitivity for new SKUs that are currently not part of the assortment by relating the SKU-level elasticities to their attribute-level counterparts.

More insight into the relative merits of such approaches can be demonstrated by an in-sample comparison of predicted SKU-level fixed effects for the ten top- and bottom-selling SKUs. We compare SKU-level fixed effects for three models: (1) estimated effects from the true SKU-level model, (2) calculated effects from the attribute-level model, and (3) estimated effects from the FHPC model. The MSE difference for the SKU-level and attribute-level models is .00014 (see Table 4). However, for the SKU-level and FHPC models, this is considerably larger, at .503. This dif-

18In an unpublished working paper, Hardie and colleagues (1998) extend FH to the market share setting and incorporate several new features (including hedonic price regressions and similarity variables).
ference is also reflected in the rank ordering of SKUs: Both the SKU-level and the attribute-level models produce almost identical utility rank orderings for SKUs, but the FHPC model produces markedly different rankings.

The superior performance of our attribute-level model is also evident in holdout tests in which we compute the out-of-sample mean utilities and market shares for the eight new SKUs we discussed previously. To generate forecasts for the new SKUs using the attribute-level model, we translate the recovered SKU-level intercepts to a set of part-worths for attribute levels. To accomplish this and to allow comparison with the FHPC model, we regress the recovered SKU-level intercepts onto the design matrix that the FHPC model uses.\footnote{In this regression, the number of observations is equal to the number of recovered fixed effects (i.e., 168). This step in our method is essentially identical to that which Nevo (2000) employs.} Using the recovered estimates, we forecast the mean utilities and the market shares of each SKU in the holdout sample. We constructed the calibration sample from the mean utilities and the market shares of each SKU in the holdout sample. For the FHPC model, the MAD and MSE values for the out-of-sample prediction of market share for the new SKUs are .0029 and .0000234, respectively. The attribute-level model produces slightly better out-of-sample forecasts, with values of .0024 and .0000184.\footnote{We thank two anonymous reviewers for prompting this discussion and helping us emphasize the flexibility of our procedure.}

Extensions

In this article, we develop and illustrate an analytical and empirical approach to connecting models that are estimated at different levels of aggregation. The model is surprisingly straightforward, yet it offers researchers a new and powerful way to analyze SKU-level issues using all the available data from a properly defined category with any number of SKUs. In line with FH’s (1996) work, the researcher simply needs to specify the mutually exclusive and collectively exhaustive list of attributes and levels that constitute the product category. For ease of exposition, we focused on the most “basic” version of our model, but it is important to note that several useful extensions can be handled with minor modifications of the standard setup. At least four issues warrant special mention:\footnote{In the regression, the number of observations is equal to the number of recovered fixed effects (i.e., 168). This step in our method is essentially identical to that which Nevo (2000) employs.}

- **Alternative market share models:** We have focused on the conditional logit model with an “outside good” (see Besanko, Gupta, and Jain 1998; Nevo 2000). However, our method is not restricted to this formulation and works equally well in a setup in which only conditional demand is modeled (i.e., there is no outside good in Equation 1). In this case, one SKU i is chosen as the reference SKU, and the same differencing procedure is applied to log market shares. The right-hand side of Equation 3 then becomes equal to \( \mu_{jt} - \mu_{it} \) so that all covariates must be differenced with respect to the values for SKU i before estimation. The fixed effects are interpreted relative to the reference SKU.
- **Expanded set of covariates:** In the empirical application, we included price, feature, and display. We also noted that it is straightforward to incorporate ordinal variables by recoding them as multiple dummy variables. In addition, time-specific (e.g., quarterly) fixed effects could be included, and/or the marketing-mix parameters could be interacted with these variables. Provided that the new covariates are appropriately mean centered, the underlying SKU-level fixed effects can be correctly recovered.
- **Multiple stores:** This logic also extends to a model that is estimated on data from multiple stores. In that case, each demand equation would reflect a store/attribute-level combination and would include store-level fixed effects. Mean centering for the covariates would be performed with respect to the store/attribute-level average.
- **Discrete heterogeneity:** Much of the panel data literature (including FH [1996]) has examined the issue of accounting for unobserved heterogeneity. Although we defer the estimation of such a model to further research, in principle, discrete heterogeneity could be incorporated by simply conditioning on, for example, two sets of market share equations and searching over the space of weights on the two segments to minimize the total squared error of the model. However, it is unclear as to how SKU-level heterogeneity might manifest when aggregated, and this issue warrants further investigation.

Conclusion

We show analytically that a market share model specified at the SKU level is related to another market share model aggregated to the attribute level and that this relationship holds for any attribute that is part of the product category description. This knowledge facilitates the construction of SKU-level parameter values by calculation from estimates obtained from the simpler aggregated model, and it avoids the need for direct estimation of the disaggregate model. We illustrate this approach empirically using store-level scanner data from the toothpaste category.

Although the number of SKUs changes dramatically from week to week and there is the possibility of interaction among attributes (e.g., Colgate’s tartar control toothpaste is viewed more favorably than the alternative that Crest produces), our methodology is able to recover parameters of the SKU-level market share model very well. Out-of-sample forecasts reveal that the attribute-level model produces slightly better predictions than the pure characteristics approach. The attribute-level model’s simplicity, parsimony, and reliance on analytical relationships make it an appealing method for both researchers and practitioners. In subsequent work, we plan to apply this new method to substantive problems that require the use and understanding of data from all SKUs in a product category.

APPENDIX

Model Simulations

We generated market shares for SKUs \( j = 1, \ldots, J \) over time \( t = 1, \ldots, T \) on the basis of an additive specification for the indirect utility as a function of attribute utilities. The attribute utilities and the simulated SKU-level utilities are set relative to a hypothetical outside good: If consumer utility for an SKU is less than zero, no purchase occurs. We used combinations of brand and size and had a complete design of two sizes and three brands (for a total of six SKUs). We used a set of prices with average values that were unrelated to the mean utilities for the attribute levels. Therefore, indirect utilities are

\[
U_{jt} = \beta_j + \gamma p_{jt},
\]

where \( \beta_j = f(b_j, s_j) \) is the SKU-level intercept—a linear additive function of brand and size attributes, respectively.
The terms \( \gamma \) and \( \beta_j \) are parameters to be recovered in the simulation. A standard normal residual is added to \( U_{jt} \) and the data are transformed to obtain regression equations for SKU-level shares (see Equation 3). To obtain the simulated attribute share values, we add up market shares of SKUs that belong to a common level (same brand or same size) of an attribute. Formally, we do this using

\[
\begin{align*}
Bx &= \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \\
Sx &= \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \\ 1 \\ 0 \\ 1 \end{bmatrix}.
\end{align*}
\]

We use the matrix \( Bx \) to produce attribute share equations for brands and the \( Sx \) matrix to produce attribute shares for sizes. A Kronecker delta operator multiplies the transformation matrix by an identity matrix \( I_T \) of dimension \( (T \times T) \) so that we have the mapping of SKU shares to brand shares \( S^b = (I_T \otimes Bx)'s \) and SKU shares to size shares \( S^s = (I_T \otimes Sx)'s \), where \( s \) denotes the SKU share vector of dimension \( (J \times T) \times 1 \). The design matrix for brand and size regressions can be created using \( Bx \) and \( Sx \) stacked \( T \) times. For the SKU market share regressions, we stack the identity matrix \( I_T \) \( T \) times. With the simulated matrix of market shares, we build the estimation equations using the transformations \( \log s_{jt} - \log s_{0t} \), where both the outside good value for the attribute share and the SKU share models have the same values for each observation \( t = 1, \ldots, T \) (see Equation 6). The estimation procedure is the same as in Equations 15 and 26.

**Standard Errors**

The standard errors of the SKU level intercepts are obtained from

\[
\text{Var}(\beta_j) = E(\beta_j^2) - E(\beta_j)^2,
\]

which can be rewritten as

\[
\begin{align*}
(\text{Al}) \quad \text{Var}(\beta_j) &= E\left\{ \alpha_h^2 + (\nu_h^2 - \nu_j^2) + \frac{1}{T} \sum_{t=1}^{T} \log \frac{s_h}{s^b} \right\}^2 \\
&- E\left\{ \alpha_h^2 + \frac{1}{T} \sum_{t=1}^{T} \log \frac{s_h}{s^b} \right\}^2 \\
&= E\left[ (\nu_h^2 - \nu_j^2)^2 + 2E\left[ (\nu_h^2 - \nu_j^2) \log \frac{s_h}{s^b} \right] \right] \\
&+ E\left[ \left( \log \frac{s_h}{s^b} \right)^2 \right] - \left[ \frac{1}{T} \log \frac{s_h}{s^b} \right]^2 \\
&= E\left[ (\nu_h^2 - \nu_j^2)^2 \right] + \frac{1}{T} \sum_{t=1}^{T} \log \frac{s_h}{s^b} \right]^2 \\
&- \left[ \frac{1}{T} \sum_{t=1}^{T} \log \frac{s_h}{s^b} \right]^2.
\end{align*}
\]

where the second step is a result of expanding the terms in brackets and taking expectations. This is a decomposition of the standard error into (1) a component due to the SKU, (2) a component due to brand \( l \), and (3) a component due to a mean-centered log ratio of the SKU share to the attribute share. For \( \epsilon_{jt} \), we can use the estimated OLS residual from Equation 26. Similarly, for \( \nu_{jt}^b \), we use the residual of the OLS brand share regression. The standard errors for parameters associated with model covariates \( (\gamma) \) is \( \sqrt{s^2(XX')^{-1}} \), where \( X \) is the mean-centered covariate, and \( s^2 \) is the estimated variance, \( \epsilon^2/\left(N - k\right) \).

**Price Elasticities**

The usual calculation of point elasticities for price in the market share model is

\[
\eta_{jk} = \frac{\partial s_j}{\partial p^k} \frac{\hat{p}_k}{s_j},
\]

where \( \eta_{jk} \) is the price elasticity for price of option \( k \) on the share of option \( j \), and \( s_j \) and \( \hat{p}_j \) are the mean share and price, respectively, for option \( j \). In the specification of market share that we give in Equation 1, this becomes

\[
\eta_{jk} = \gamma \hat{p}_j (1 - \hat{s}_j) \forall j = k,
\]

\[
= -\gamma \hat{s}_j \hat{p}_j \forall j \neq k.
\]

In the case of the mean-centered market share specification, the price “point” at which we take this elasticity is zero (recall that the price coefficient for the mean-centered is identical to the price coefficient for the non-mean-centered price market share model). The mean utility of choice \( j \) in the non-mean-centered model is related to the attribute utility \( \beta_j^{mc} \) of the mean-centered model as follows:

\[
\mu_j = \beta_j^{mc} + \gamma \hat{p}_j + \epsilon_j.
\]

Therefore, the elasticity should take into account the indirect effect of a price change on mean utility by the average of the covariate. Therefore, the own-price elasticity is as follows:

\[
\eta_{jj} = \frac{\partial s_j}{\partial p^j} \frac{\hat{p}_j}{s_j}.
\]

\[
= \gamma \hat{p}_j (1 - \hat{s}_j) \hat{p}_j
\]

\[
= \gamma \left( 1 - \frac{\partial \hat{p}_j}{\partial p^j} \right) \hat{p}_j.
\]

implying that we need to consider the impact of \( -\gamma (\partial \hat{p}_j/\partial p^j) \) in calculating the elasticity. A similar effect is present in the cross-price elasticity:

\[
\eta_{jk} = -\gamma \left( 1 - \frac{\partial \hat{p}_k}{\partial p^j} \right) \hat{s}_k \hat{p}_k.
\]

If these are temporary price changes, the impact is likely to be small. If the change is more permanent, the cross-price elasticity must take into account its impact on the average price.
REFERENCES