Contextual and Temporal Components of Reference Price

An emerging consensus in marketing is that consumers respond to price relative to some standard or reference price. Most researchers modeling brand choice have reasoned that this standard is based on past prices of the brand. The authors argue that consumers do use reference prices, but one that is also based on context—other prices in the store—rather than on past prices alone. An analysis of households’ brand choices in two subcategories and over three cities supports this premise. Within context, the lowest price seems to be an important cue for reference price, whereas within time, a brand’s own past prices seem to be the most important cue. Households’ use of a contextual reference price also varies predictably across some consumer characteristics. Though their model can be applied to other categories, the findings have important managerial implications: Managerial focus on temporal reference prices could lead to an everyday high price, whereas focus on contextual reference prices could lead to an everyday low price. Only the inclusion of both contextual and temporal reference prices justifies variable pricing.

Reprinted with permission of the copyright owner. Further reproduction prohibited without permission.
which are increasingly and more easily available to practicing managers. The major limitations of analyzing reference prices by scanner data is that reference prices only can be inferred from behavior and correlations among variables need not imply causality.

In the interests of generalizability, we analyze our model over three cities and two subcategories of the product. We also analyze the model over different subgroups of consumers to determine how certain underlying characteristics, such as brand loyalty, brand sampling, and purchase frequency, influence the role of reference price. Though the model could have broad applications, the results are likely to apply primarily to grocery products. This class of goods makes up a sizable share of the consumer’s budget and is the most common type of consumer purchase. The following sections cover the literature, theory, model, empirical analysis, and managerial implications of the study.

Findings From Logit Models of Reference Price

Despite its intuitive appeal, reference price began to be formally modeled only in the last decade. In particular, most prior models of sales did not incorporate an explicit term for reference price (Tellis 1988a). Recently, several analyses of the role of reference prices in consumer choice have appeared in print (e.g., Kalwani et al. 1990; Lattin and Bucklin 1989), at conferences, and in manuscripts (e.g., Gurner and Little 1989; Rajendran 1988; Rajendran and Tellis 1990; Raman 1986). The underlying principle in all these models has been the same—that consumers’ brand choices depend on a comparison of actual price with reference price, in addition to other brand and consumer characteristics. However, the studies differed in their measures of reference price. In the interest of brevity, we do not review the theoretical basis and empirical results of each study, but merely summarize our conclusions.

First, empirical results generally support the inclusion of reference price in logit models of brand choice. Second, past studies differ in the specification of the choice model containing reference price. Though most include the difference between price and reference price, some explicitly include reference price and others the actual price of the brand. Proper model specification is critical to obtain and interpret correctly the signs of the coefficients, which are a problem in some studies. Third, most of the choice models measure reference price primarily as some average of past prices. Though some studies do consider the context indirectly (e.g., Lattin and Bucklin 1989; Raman 1986), they do not explicitly separate temporal and contextual reference prices. This separation is necessary to determine precisely the impact of context or time and price the brand optimally. Also, none of the studies identify which price within context or time defines reference price. Actually, none of the studies examine the importance of alternative cues for reference price. However, many behavioral theories and experimental studies underscore the importance of contextual reference prices, as the subsequent discussion illustrates.

The Nature of Consumer Response to Price

Role of Reference Price

Reference price applies only to those consumers who respond to differences in price. This fraction of consumers is substantial as we can infer from the frequent changes in weekly prices of grocery products and the explanatory power of price in brand choice models. The reference price also applies to the standard price consumers use for comparison when choosing among brands (or sizes) in real markets. We exclude reference prices elicited in surveys or laboratory experiments, called aspiration price by Klein and Oglethorpe (1987).

When choosing among brands, consumers evaluate prices not absolutely, but in comparison to some standard price, the reference price. The empirical studies reviewed previously support this premise, and at least two theoretical streams explain the logic. First, prospect theory (Kahneman and Tversky 1979), or mental accounting (Thaler 1985), suggests that consumers do not make decisions in terms of absolute wealth but of losses or gains relative to a reference point. In the current context, this would imply evaluating brands by comparing their price with a reference price. Second, several researchers have suggested that adaptation level theory (Helson 1964) and assimilation-contrast theory (Sherif 1963) support the role of reference price. According to these theories, consumers have a standard (reference price) by which they evaluate current information (new prices). New credible prices are assimilated and used to update the reference price. (Prices not deemed credible are contrasted and possibly rejected or affect the reference price to a lesser extent.)

Components of Reference Price

The key question we research is which sets of prices influence the reference price. We posit that reference price is based on two important components, contextual and temporal prices. These two components are parsimonious and encompass most of the relevant prices consumers use. Though the temporal component has been the focus of research using scanner data, some theories and scattered experimental studies do support the importance of contextual reference prices. For example, adaptation level theory suggests that consumers adapt to two sets of stimuli—focal and contextual. The focal stimulus is the target price, and the contextual stimuli are all other prices in the choice context. Prospect theory and mental accounting stress the importance of framing effects. The store environment is the most immediate and salient frame relevant to brand choice, so other prices in the store should be important determinants of reference price. Thaler (1985) in particular argues that when buyers do not know the seller’s costs, market prices determine reference price; for grocery products, market prices mainly reduce to store prices for brands in the product category.

How do consumers develop and use these two types of reference prices? Mazumdar and Monroe (1990) and Monroe, Powell, and Choudhury (1986) provide a frame-
work useful for this purpose. They suggest that consumers learn price information in two ways: intentional and incidental learning. Intentional price learning results from an active search and memorization of exact prices, typically for specific brands. Certain consumers who are very careful shoppers and strive to minimize costs adopt such behavior. Intentional learning includes explicit comparison of current prices with previous prices stored in memory. Thus, intentional learning seems to favor temporal reference prices.

In contrast, incidental price learning occurs when consumers compare prices across brands in the course of buying, without any explicit intention or effort to memorize them. These repetitive price comparisons over time do lead to low involvement learning, but such learning is likely to be of the relative price rank rather than of specific prices (Mazumdar and Monroe 1990). Because retailers frequently reduce the price of one brand in a choice set, incidental learning could soon be obsolete, requiring such consumers to keep comparing the prices across brands. Thus, incidental learning can be compatible with contextual reference prices.

**Relative Importance of Components**

Which of these two components would be stronger? The context component could be stronger for several reasons. The main reason is that the contextual prices are more recent and salient than past prices at the point of purchase. Consumers must make a large number of brand choices in numerous categories. In addition, many of these brands within a category are similar and have fluctuating prices. One efficient choice strategy would be to choose a brand by comparing only the shelf prices before them without also trying to recall prices from past occasions.

Second, various surveys of consumers’ price knowledge suggest that consumers may not be sufficiently aware of prices to relate current prices to past prices of each brand in each category. For example, Dickson and Sawyer (1990, p. 47) find that “less than half (47.1%)” of the consumers they surveyed immediately following a selection of the correct price. Krishna, Currim, and Shoemaker (1990) find correct recall of regular prices of nine brands to vary from 0% to 32% of a sample of consumers. Urban and Dickson (1991, p. 50) find that even when provided with a range of prices, consumers in their sample are not “knowledgeable about the (past) prices of the 18 products” they study. So they conclude that current “market prices may serve as reasonable surrogates of consumers’ internal reference prices.”

Third, some recent studies provide indirect support that consumers could respond to price differences relative to current competitive rather than past prices. In a laboratory experiment, Jacobson and Obermiller (1990) found no support for adaptive, extrapopulation, and rational expectations that price expectations are formed from the series of past prices. On the contrary, they found that a brand’s expected price was determined primarily by its current price and those of other brands. For scanner data on coffee, Gupta (1988) found that price discounts had only a small effect on consumers’ timing and quantity of purchases but a strong effect on brand choices; therefore, price differences could lead consumers to switch brands more than change how much or when they buy. This behavior suggests interbrand price comparisons rather than intertemporal comparisons. Blattberg and Wisniewski (1989) found a systematic pattern of switching across national brands, private labels, and generics in response to price discounts. This response pattern again suggests interbrand price comparison rather than intertemporal comparisons. Though these three studies do not directly support a contextual reference price, they do suggest that consumers make interbrand price comparisons more than intertemporal price comparisons.

**Price Cues as Reference Points**

Which price cues or measures within context or time are most likely to affect the reference price? We test three measures for the context component: the highest, lowest, and simple mean of the prices of all brands in the store at the purchase occasion. The advantage of the mean is that it captures the effect of all the other brands; however, the end points of the price distribution, the highest and especially the lowest prices could influence consumers more than the mean, because of their salience and availability; the mean has to be computed. Indeed, the highest and lowest market prices have been researched as important reference points (Biswas and Blair 1991) or anchoring stimuli (e.g., Lynch, Chakravarti, and Mitra 1991). Furthermore, the lowest price could be more important than the highest, because it often is featured in local newspapers or displayed in the store.

We use two measures for the temporal component: (1) a single price (for all brands) based on some average of past prices paid on past purchases and (2) a single price for each brand based on some average of each brand’s past prices, assuming these prices could be observed by the consumer. Most past studies use the latter measure for the consumers’ reference price. However, the past prices paid could define the reference price better because the consumer has numerous occasions to elaborate on this price when choosing, paying for, and using the brand, whereas the consumer may not always observe the prices of each of the other brands.

**Variation Across Consumers**

Is the relative role of the two components of reference price likely to vary by consumer characteristics? We propose that the importance of these two components may vary especially by consumer purchase characteristics, three of which we can analyze with our data: brand preference strength, brand sampling, and purchase frequency. The first two are related.

The temporal component could be stronger among consumers who have a strong preference for one brand. We define brand preference as the consumer’s share of purchases devoted to his or her favorite brand. Consumers who have strong preferences for a brand are more likely to have strong beliefs and a well-developed cognitive structure about the brand (Johnson and Russo 1978). Such consumers are also likely to have better assimilation and recall of the prices of the favorite brand (Biehal and Chakravarti
in the form of a temporal component. These consumers might tend to be less price sensitive in general. However, to the extent that they are price sensitive, these consumers are more likely to use this temporal component than the context component to evaluate the current prices of a prospective brand.

The context component is likely to be stronger among consumers who sample several brands. Consumers could sample several brands either for variety, because of promotions, or to suit different members of the family (McAlister and Pessemier 1982). Such consumers could perceive a set of competing brands to be substitutes, and they could compare attributes with respect to this set. Such consumers are more likely to observe the prices of competitive brands on the shelf and develop a contextual reference price.

The importance of the two components also could vary by purchase frequency. Consumers who purchase the category more frequently could have better knowledge and recall of prices (Urban and Dickson 1991) and are more likely to use a temporal reference price. In contrast, infrequent buyers in the category are more likely to resort to a contextual reference price.

**Summary**

The previous discussion can be summarized into a set of seven hypotheses:

- **H1:** Consumer response to prices involves comparison with two standards: a contextual and a temporal reference price.
- **H2:** Price response could be stronger relative to a contextual reference price than to a temporal reference price.
- **H3:** The lowest competitive price would be a more important measure of the contextual reference price than the highest or the mean.
- **H4:** The mean of past prices paid would be a more important measure of temporal reference price than mean of past prices of each brand.
- **H5:** The contextual (temporal) reference price would be stronger (weaker) for consumers with lower brand preference.
- **H6:** The contextual (temporal) reference price would be stronger (weaker) for consumers sampling a wider number of brands.
- **H7:** The contextual (temporal) reference price would be stronger (weaker) for consumers with lower frequency of purchase.

**Modeling Consumer Response to Price**

The proper specification of a choice model including reference prices is critical to obtaining and interpreting the right signs of the various price terms. This section first develops a price response model from first principles and then extends it to a choice model.

**Modeling Price Response**

We initially assume that price is the only factor influencing brand choice and suppress all subscripts for brand, consumer, and time. Consumers can integrate the two components of reference price in one of two ways. First, they can evaluate the target brand's price against a single reference price (RF) formed as a weighted average of the temporal component (Rm) and context component (Rc), thus,

\[
RF = \alpha Rm + (1 - \alpha) Rc
\]

where \( \alpha \) (bounded by 0 and 1) is the subjective weight attached to each component. The utility (V) associated with buying the brand priced at P then can be represented as

\[
V = \beta RF - P
\]

where \( \beta \) is the weight attached to the price comparison assuming (initially) a linear utility function. (Later, we estimate the model by the logit function, which accommodates nonlinear utilities [McFadden 1974].) Because price is included as a negative in equation 2, we expect \( \beta \) to be positive. That is, the lower the brand's price or the higher a consumer's reference price, the more attractive the brand.

Substituting equation 1 into equation 2 gives us

\[
V = \alpha \beta Rm + (1 - \alpha) \beta Rc - \beta P = \beta_1 Rm + \beta_2 Rc + \beta_3 P,
\]

where \( \beta_1 = \alpha \beta_2 \) and \( \beta_2 = (1 - \alpha) \beta_3 \).

The reduced form, equation 3, can be incorporated easily into a price response model such as the logit model. We refer to this as the *component form* of the price response model. The attractive feature of equation 3 is that it allows us to test explicitly whether consumers consider reference prices and how much weight they put on each component. Our previous discussion would suggest that \( \beta_1 \) should be negative, \( \beta_2 \) and \( \beta_3 \) positive, and, in general, \( \beta_2 \) should be smaller than \( \beta_3 \) (or \( \alpha < .5 \)). If \( \beta_1 \) and \( \beta_2 \) are both zero, then consumers respond to price without using reference prices at all. The unattractive feature of this form is that it shows brand choice depending directly on the consumer's reference price, which may not represent the true process.

Alternatively, one can argue that consumers make two price comparisons at the point of purchase, comparing the target brand's price against each of the components. Using the previous notation and logic, this argument would imply the following reduced form:

\[
V = \beta_1 (Rm - P) + \beta_2 (Rc - P).
\]

We refer to this equation as the *comparison form* of the price response model. Our hypotheses would suggest that \( \beta_1 \) and \( \beta_2 \) would be positive and \( \beta_1 \) should be smaller than \( \beta_2 \). The reason \( \beta_1 \) and \( \beta_2 \) are positive is that the brand becomes more attractive as its price decreases relative to either of its reference prices, but especially so relative to its contextual reference price. Our derivation of equation 4 indicates that it does not explicitly contain price. In other words, consumers do not respond to price relative to some reference point and at the same time respond to price directly. Mathematically, it is easy to see that even with a price term in equation 4, it would reduce to the second line of equation 3.

**Modeling Brand Choice**

**Comparison form.** In addition to price, past research indicates that features, displays, coupons, consumers' brand

---

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
loyalty, and prior choices influence consumers' current brand choice (Gupta 1988; Tellis 1988b). Equation 4 can then be relaxed to include these variables and allow for variation over brands, time, and consumers. Thus,

\[ V_{ijt} = \beta_0 (R_{mi} - P_{ij}) + \beta_1 (R_{mj} - P_{ij}) + \beta_2 L_{ijt} + \beta_3 C_{ij} (t-1) + \beta_4 D_{ijt} + \beta_5 F_{ijt} + \beta_6 N_{ijt}, \]

where the symbols L, C, D, F, and N stand for brand loyalty, brand choice, display, feature, and coupons respectively; and the subscripts i, j, and t stand for brand, consumer, and time. Assuming equation 5 can be estimated with errors that follow a Gumbel distribution, the probability of brand choice, \( P(C=1) \) then can be represented by the relatively simple multinomial logit model (McFadden 1974) as follows:

\[ P(C=1)_{ijt} = e^{V_{ijt}} \sum V_{ijk} \]

where \( k \) is an index for brands in the choice set. Note that equation 6 involves a nonlinear utility function and is similar to the approaches adopted by Lattin and Bucklin (1989) and Tellis (1988b). We define the base model as the one that includes all variables except the price terms. Three alternative models test the major hypotheses about the presence and importance of the two components of reference price: (1) the base model plus price (P) alone, (2) the base model plus "temporal reference price - price" (Rm - P) alone, and (3) the base model with both "temporal reference price - price" plus "contextual reference price - price" (Re - P). To test the hypotheses about the major mediators, we ran these models by appropriate subgroups for each mediator, as explained subsequently. To determine whether the reference price terms are relevant, we compare the model fits among the nested versions (1)–(3), or test the following component form.

Note that it is not theoretically meaningful to have the brand's price included along with the comparison terms in the same model run. Our theoretical exposition (equations 1–4) of the role of reference price clearly posits either the comparison form in which the brand's price already enters each of the two comparison terms (equation 4), or the component form in which it appears by itself alongside each component (equation 3).

Component form. Estimation of the component form of the model would provide further evidence of the role of reference prices, especially since the brand's price and each of the components enters the model separately. Because our empirical analysis is already complex, with four markets and three subsamples in each, we present this analysis in a technical appendix (1992) available from the authors. The results support the role of both components of reference price.

Data

The data are IRI scanner panel records for saltines from January 1984 to January 1986. We chose this category because it was the most detailed at the time we started this study (1980). The category is dominated by one or two brands, which can reduce interbrand price comparison and thus the importance of the context component. However, this influence would be in the direction opposite to our hypotheses. We classify this category into two subcategories, salted and unsalted. For health or taste reasons, most consumers are likely to have clear preferences for one or the other, and they probably do not perceive them as substitutes. Our analysis by subcategories ensures better definition of the relevant choice set and reference prices.

The data show 150 "offerings," covering 12 sizes. (An offering is a unique combination of vendor, item, size, and generation or version of UPC code.) Of these, 51 were purchased at least once during the two years covered by the data, of which 38 were of the 16-ounce size. We restrict the analysis to the 16-ounce size for the following reasons. First, this size dominates purchases in the category (93.6% of all purchase occasions). Second, modeling price comparisons over different sizes requires some form of unit pricing, but consumers do not seem to make use of unit price information (Dickson and Sawyer 1990). Third, focusing on one size gives a clearer and workable choice set. These decision rules yielded a choice set of four national brands in the salted subcategory and three in the unsalted subcategory. In addition, each store has a few private labels in each subcategory. Descriptive statistics are available in the technical appendix (1992).

The data on 16-ounce saltines cover three cities: Midland, Texas; Rome, Georgia; and Williamsport, Pennsylvania. Of the 42,180 purchases, the salted subcategory accounts for about 85% and the unsalted subcategory for the remaining 15%. To test the stability of our results, we analyze the model separately for each of the cities for the salted category, but pool observations across cities for the unsalted category because of inadequate observations.

We used two purchase criteria to select the panelists for analysis. First, we included only those panelists who had records for all of the two years. (That is, we excluded panelists who joined late or left early). Second, we divided the two years of data into a calibration period of 35 weeks to develop measures of loyalty and reference price and an estimation period of 69 weeks to estimate the model. We included panelists only if they had at least 5 purchases in the first 35 weeks and 20 purchases during the whole two-year period. Thus, the samples are representative of those panelists who purchase frequently from the category. This fact could bias the results in favor of the temporal component in the direction opposite our hypothesis (H3). In addition, we assessed the impact of purchase frequency by split sample analysis of the model by the purchase frequency of buyers in the remaining sample.

Measures

All the measures were derived from scanner data. Purchases were recorded at the minute of occurrence, called a purchase occasion. Most of the promotional variables remained the same for the whole week. Yet, to avoid aggregation bias, we analyzed the data at the level of purchase occasions rather than aggregated to the week. We measured brand loyalty as the share of a household's purchases of each brand in the calibration period of 35 weeks. We meas-
ured lagged choice as a dummy variable, depending on whether a panelist bought the brand on the last purchase occasion. We measured price as the actual price paid by consumers in dollars, net of all discounts or deals offered in the store. We measured feature and display as dummy variables, which are 1 if the event was on at the purchase occasion and 0 otherwise. We measured coupons as a dummy variable, which is 1 if any household used coupons on that day and 0 otherwise. We did not incorporate coupons into the transaction price because coupons are very infrequent, vary widely in use, and cannot be measured very accurately.

We used two formulas to compute the temporal reference price: the simple mean and the weighted mean of prices on the past three purchase occasions. We used only three past periods because crackers are purchased quite infrequently (mean frequency = 1/month). Past studies have used from one (the most recent) to five past periods. The weighted mean uses declining weights of .571, .286, and .143 for older periods. This is an approximation of a geometric function with a common ratio of .5. We would try other weighting schemes if the results were sensitive to weights.

Our measures of the temporal and contextual components ensure maximal independence between the two constructs and minimize multicollinearity. Note that measures for the contextual component (the highest, lowest, or mean of current prices of brands) varies primarily by store, whereas the temporal component (past prices paid or past prices of each brand) varies primarily by household. Empirical analysis supports the independence of the two components. To determine how our results varied over subsamples of the panelists, we split the sample into two subsamples by brand preference, brand sampling, and purchase frequency. We then ran the logit model separately for each pair of subsamples. This approach is preferable to including the relevant interaction terms in the logit model because it is easier to execute and interpret. We measured brand preference by the panelist’s share of purchases over the two years for his or her most preferred brand and used .66 to separate panelists with strong and weak preference. We used the number of brands the panelists has bought even once over the two years as an index of brand sampling and used three (brands) to separate panelists with wide and narrow sampling width. We measured purchase frequency by the number of purchases in two years and 27 purchases to separate frequent from infrequent panelists.

Results and Discussion

We now discuss the results of the multinomial logit analysis (equations 5 and 6) for four “markets,” Rome, Midland, Williamsport, and the unsalted subcategory pooled over the three cities. We discuss results for only the comparison form of the model (equation 4) here because they are easier to interpret and quite consistent with the component form of the model (equation 3, technical appendix, 1992). Recall that in the comparison form, reference price is included as “reference price minus price”; so the correct sign for this term would be positive.

We evaluate alternative measures of reference price by the sign and the t-value of the coefficient and by the indices of model fit: (1) the “log-likelihood” and (2) the “−2 log-likelihood ratio” (−2 LLR) between two rival models; this statistic is distributed as x² with n₁ - n₂ degrees of freedom, where n₁ − n₂ is the difference in the number of variables in the rival models (Ben-Akiva and Lerman 1985, p. 166). We discuss the results in three parts: choice of measures, general results, and split-sample analysis.

Choice of Measures

As explained previously, we test three measures for the contextual reference price: the lowest, mean, and highest of the prices in the store. All three measures give significant coefficients with the right sign. However, the highest price as reference price gives the poorest results. Consistent with H₄, the lowest price as reference price gives the best results in terms of higher t-values and higher model x². As stated previously, consumers are also more likely to use the lowest price for the reference price than the mean or the highest of all prices, because the lowest price is more available or salient and often is advertised. So we proceed with the lowest price as the measure of contextual reference price. The lowest priced brand in our data is often a private label. Though it does not set the quality standard, these results seem to imply that it may well set the reference price.

We test two measures for the temporal component: the past prices paid by the panelist and past prices of each brand. Contrary to H₄, reference price defined on past prices paid is mostly insignificant. However, reference price defined on past prices of each brand is always significant with the right sign and higher log-likelihood values. Because this is also the measure used in the literature, we proceed with it. We also test two computational formulas for these reference prices: reference price defined as a simple average, or a weighted average of the last three periods. We obtain similar results with either formula, which suggests that the results may not be sensitive to the weighting scheme. Therefore, we do not try any other weights. Because the latter formula is more plausible, we proceed with that measure.

General Results

Table 1 presents the logit estimates of all variables with the previously mentioned measures for reference price. To show clearly the role of price, we present models partially nested for the price terms: first, the base model with all variables plus price by itself (model 1); then, the base model plus the temporal comparison (model 2); and last, the base model plus temporal and contextual comparisons (model 3). Because the existence of the context comparison is the key hypothesis of our article, this order ensures a more conservative test of our hypotheses.

In all four markets, both the temporal and the contextual reference price comparisons have the right sign and are significant. The t-values are generally large, ranging from 2.4 to 11. These results are consistent with H₄. Note that the log-likelihood value is highest for model 3 containing both
## TABLE 1
Coefficient Estimates from Multinomial Logit Models of Brand Choice
[t-statistics in parentheses]

<table>
<thead>
<tr>
<th></th>
<th>MIDLAND</th>
<th>ROME</th>
<th>WILLIAMSPORT</th>
<th>UNSALTED</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Model 1</td>
<td>Model 2</td>
<td>Model 3</td>
<td>Model 1</td>
</tr>
<tr>
<td>Loyalty</td>
<td>2.28*</td>
<td>2.07*</td>
<td>2.24*</td>
<td>2.53*</td>
</tr>
<tr>
<td></td>
<td>[19]</td>
<td>[18]</td>
<td>[18]</td>
<td>[30]</td>
</tr>
<tr>
<td>Lag. choice</td>
<td>1.59*</td>
<td>1.75*</td>
<td>1.63*</td>
<td>1.54*</td>
</tr>
<tr>
<td></td>
<td>[19]</td>
<td>[21]</td>
<td>[19]</td>
<td>[26]</td>
</tr>
<tr>
<td>Coupon</td>
<td>1.19'</td>
<td>1.03'</td>
<td>1.17'</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>[6]</td>
<td>[5]</td>
<td>[5]</td>
<td>[0]</td>
</tr>
<tr>
<td>Display</td>
<td>.52'</td>
<td>.70'</td>
<td>.54'</td>
<td>.79'</td>
</tr>
<tr>
<td>Feature</td>
<td>.98'</td>
<td>.87'</td>
<td>.79'</td>
<td>1.06'</td>
</tr>
<tr>
<td>Price</td>
<td>-.12'</td>
<td>-.94'</td>
<td>-.26'</td>
<td>-.17'</td>
</tr>
<tr>
<td>Temp.-pricea</td>
<td>1.77'</td>
<td>1.02'</td>
<td>3.46'</td>
<td>3.13'</td>
</tr>
<tr>
<td>Cont.-pricea</td>
<td>1.07'</td>
<td>.48'</td>
<td>.48'</td>
<td>.48'</td>
</tr>
<tr>
<td>Log-likelihood</td>
<td>957</td>
<td>978</td>
<td>953</td>
<td>1817</td>
</tr>
<tr>
<td>Schwarz’s BIC</td>
<td>984</td>
<td>1005</td>
<td>984</td>
<td>1841</td>
</tr>
<tr>
<td>-2 LLR relative to:</td>
<td>74*</td>
<td>32*</td>
<td>84*</td>
<td>62*</td>
</tr>
<tr>
<td>base model</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>model 1</td>
<td>10*</td>
<td>117*</td>
<td>41*</td>
<td>41*</td>
</tr>
<tr>
<td>model 2</td>
<td>52*</td>
<td>15*</td>
<td>273*</td>
<td>273*</td>
</tr>
<tr>
<td>Number of observations</td>
<td>6854</td>
<td>15,477</td>
<td>10,118</td>
<td>4004</td>
</tr>
</tbody>
</table>

*Significant at .05 or better

Temporal and contextual price comparisons, in all four markets, which supports the inclusion of two dimensions of reference price. In particular, the -2 log-likelihood ratio shows that model 3 with the two comparisons has a substantial and significant improvement over model 1 with price alone in all four markets. The -2 loglikelihood ratio also shows that the contextual comparison always provides a significant improvement to the model fit relative to the base model and to the model with price alone or with temporal comparison alone. Schwarz’s BIC, a more conservative information criterion shows that model 3 is better than model 1 in Rome and Williamsport, and as good in the other two markets; model 3 is much better than model 2 in all four markets. These results support H1.

H2 indicated that the contextual comparison would be stronger than the temporal comparison. The coefficients of the contextual comparison are similar to that of the temporal comparison in Midland and Williamsport, larger in the unsalted category, and much smaller in Rome. Similarly, the analysis of the log-likelihood ratios indicates that the contextual comparison is more important in the former two markets, whereas the temporal comparison is more important in the latter two markets. Thus, neither component dominates the formation of reference price.

The technical appendix (1992) presents the tests of the binary logit analysis of the component form of the price response model (equation 3). Besides confirming these findings, the results show that the temporal and contextual components do not substitute for the effect of price alone, but in every case enhance the size and significance of the price coefficient. The subsequent analysis sheds more light on the relative importance of these two components across various consumer characteristics.

### Analysis Across Split-Samples

#### Brand preference

Table 2 presents the multinomial logit results by panelists’ brand preference for only the price comparison terms in the model. We hypothesized (H4) that the context comparison would be stronger for panelists with weak preferences, and the reverse would hold for the temporal comparison. In each of the four comparisons, the coefficients for the context comparison are higher when preference strength is weaker, as expected. Moreover, one-tailed tests of these four differences are significant (at p < .01). In contrast, the temporal comparison is higher when preference strength is stronger in Williamsport and Midland, but not in the other two markets. Thus, the pattern is partly consistent with H5.

#### Brand sampling

Though brand preference and brand sampling are related by definition, they are not identical; their shared variance is 50% with our data and measures. Table 3 presents the results of the logit analysis by panelists’ brand sampling. We expected (H6) the context compar-
TABLE 2
Multinomial Logit Estimates of Price Coefficients by Brand Preference
[t-statistics in parentheses]

<table>
<thead>
<tr>
<th>PRICE</th>
<th>MIDLAND</th>
<th>ROME</th>
<th>WILLIAMSPORT</th>
<th>UNSALTED</th>
</tr>
</thead>
<tbody>
<tr>
<td>MEASURE</td>
<td>Strong</td>
<td>Weak</td>
<td>Strong</td>
<td>Weak</td>
</tr>
<tr>
<td>Temp.-price</td>
<td>1.11*</td>
<td>.84*</td>
<td>2.92*</td>
<td>3.38*</td>
</tr>
<tr>
<td></td>
<td>.35</td>
<td>.58*</td>
<td>.35</td>
<td>.4*</td>
</tr>
<tr>
<td>Cont.-price</td>
<td>.56*</td>
<td>1.50*</td>
<td>.19</td>
<td>.75*</td>
</tr>
<tr>
<td></td>
<td>[3]</td>
<td>[7]</td>
<td>[1]</td>
<td>[5]</td>
</tr>
<tr>
<td></td>
<td>1.53*</td>
<td>2.56*</td>
<td>1.53*</td>
<td>2.56*</td>
</tr>
<tr>
<td></td>
<td>[4]</td>
<td>[3]</td>
<td>[4]</td>
<td>[3]</td>
</tr>
<tr>
<td>Significance of test of difference</td>
<td>.01</td>
<td>.05</td>
<td>.01</td>
<td>.01</td>
</tr>
<tr>
<td># of observations</td>
<td>3903</td>
<td>2951</td>
<td>10100</td>
<td>5377</td>
</tr>
<tr>
<td></td>
<td>6725</td>
<td>3933</td>
<td>3662</td>
<td>342</td>
</tr>
</tbody>
</table>

*significant at .05 or better

Model Fit

Because of differences in theory, measures, and methods, we cannot directly compare our model as a nested version of previous logit models with reference price. However, all previous models defined reference price primarily on time (past prices). Our results show the substantial improvement in fit by including a contextual reference price. Some studies resulted in wrong signs with a temporal reference price. With proper model specification and proper measures for the contextual and temporal reference prices, we find that the price coefficients have the right sign in all runs of the model and all split samples. The coefficients are significantly different from 0 in almost every case. To put these results in perspective, we compare the explanatory power of our model versus those of prior studies, by the U^2, which ranges from 0 to 1 and measures the reduction in uncertainty. Our model, with an average adjusted U^2 of .65, compares very favorably, lending another measure of confi-

TABLE 3
Multinomial Logit Estimates of Price Coefficients by Brand Sampling
[t-statistics in parentheses]

<table>
<thead>
<tr>
<th>PRICE MEASURE</th>
<th>MIDLAND</th>
<th>ROME</th>
<th>WILLIAMSPORT</th>
<th>UNSALTED</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Wide</td>
<td>Nar</td>
<td>Wide</td>
<td>Nar</td>
</tr>
<tr>
<td>Temp.-price</td>
<td>.37</td>
<td>2.01*</td>
<td>3.59*</td>
<td>2.74*</td>
</tr>
<tr>
<td></td>
<td>[1]</td>
<td>[3]</td>
<td>[9]</td>
<td>[6]</td>
</tr>
<tr>
<td></td>
<td>2.44*</td>
<td>.58</td>
<td>2.44*</td>
<td>.58</td>
</tr>
<tr>
<td>Significance of test of difference</td>
<td>.01</td>
<td>.01</td>
<td>.01</td>
<td>.01</td>
</tr>
<tr>
<td>Cont.-price</td>
<td>1.80*</td>
<td>.09</td>
<td>.73*</td>
<td>.27</td>
</tr>
<tr>
<td></td>
<td>[8]</td>
<td>[3]</td>
<td>[4]</td>
<td>[1]</td>
</tr>
<tr>
<td></td>
<td>2.84*</td>
<td>1.71*</td>
<td>2.25*</td>
<td>1.60*</td>
</tr>
<tr>
<td></td>
<td>[14]</td>
<td>[6]</td>
<td>[5]</td>
<td>[3]</td>
</tr>
<tr>
<td>Significance of test of difference</td>
<td>.01</td>
<td>.01</td>
<td>.01</td>
<td>.05</td>
</tr>
<tr>
<td># of observations</td>
<td>3397</td>
<td>3457</td>
<td>4242</td>
<td>11235</td>
</tr>
<tr>
<td></td>
<td>4527</td>
<td>5591</td>
<td>1253</td>
<td>2751</td>
</tr>
</tbody>
</table>

*significant at .05 or better

Contextual and Temporal Components of Reference Price / 29

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
TABLE 4
Multinomial Logit Estimates of Price Coefficients By Purchase Frequency
[t-statistics in parentheses]

<table>
<thead>
<tr>
<th>PRICE MEASURE</th>
<th>MIDLAND</th>
<th>ROMA</th>
<th>WILLIAMSPORT</th>
<th>UNSALTED</th>
</tr>
</thead>
<tbody>
<tr>
<td>Temp.-price</td>
<td>1.15*</td>
<td>.95*</td>
<td>3.60*</td>
<td>2.41*</td>
</tr>
<tr>
<td>Significance of test of difference</td>
<td>not significant</td>
<td>.01</td>
<td>.01</td>
<td>.01</td>
</tr>
<tr>
<td>Cont.-price</td>
<td>1.16*</td>
<td>.85*</td>
<td>.28*</td>
<td>.90*</td>
</tr>
<tr>
<td>Significance of test of difference</td>
<td>not significant</td>
<td>.01</td>
<td>.01</td>
<td>not significant</td>
</tr>
<tr>
<td># of observations</td>
<td>4149</td>
<td>2705</td>
<td>9800</td>
<td>5677</td>
</tr>
</tbody>
</table>

*significant at .05 or better

Dence to our perspective (see technical appendix [1992] for details). In addition, across the four different cases, our model obtains successful positive prediction rates (brand choice) of 66% to 87% and total prediction rates (brand choice and nonchoices) of 90% to 92%.

Robustness of Results

The evidence overwhelmingly suggests that the contextual reference price is a significant determinant of price response and is at least as important as the temporal reference price. In addition, the lowest price could be an important measure of contextual prices, whereas the past price of each brand seems to be the best measure of temporal prices. Split sample comparison of the coefficients of the contextual reference price across brand preference and brand sampling provides further support of our interpretation of the structure of reference price. Our evidence comes from four different markets and three different split samples in each market. Thus, the findings are less likely to be caused by geographic, panelist, or random events. However, the results for Rome and the unsalted category are a little different from the other two cities, probably because of the greater popularity of national brands.

Some methodological considerations also support the robustness of the analysis. First, the partly nested models (Table 1) and correlation analysis of the logit coefficients indicates that collinearity between the contextual and temporal components is certainly not a problem. Second, we have included well-accepted measures for the promotional variables (Gupta 1988; Lattin and Bucklin 1989; Tellis 1988b). Third, the pattern of results remains the same with several alternate measures of the contextual and temporal reference prices and across the multinomial and binary logit models. Fourth, though our sample is biased toward heavier users, these users are more likely to use a temporal reference price; thus, our result is not in the direction of the bias.

On the basis of numerous logit analysis of household scanner data, we find that multicollinearity among the prototypical variables is not a serious problem. A key reason that multicollinearity is low could be because the analysis is at the individual household level; households buy at different times and from different stores. Serial patterns in promotions within the store level may not affect the correlation matrix at the household level, because each household shops randomly in nonsuccessive weeks. Also, these patterns become less dominant when mixed with other patterns of other stores. Finally, analyses at the individual household level yield a large number of observations. Mason and Perreault (1991) show that even if present, the problem of multicollinearity declines sharply with sample sizes beyond 100 and is small for samples above 300. For the main models, our observations range from 4000 to 15,000.

In our analysis, we considered all brands in the store as constituting the panelist’s choice set. Inclusion of alternatives the panelist never considers introduces noise. We repeated the main analysis, taking into account consideration sets of the panelists. A consideration set was a set of brands that a panelist used at least once in the prior period of 35 weeks held out to estimate brand loyalty. This analysis accentuates our results without changing the basic pattern.

Limitations

Though our results demonstrate the importance of reference price in this category, its many limitations suggest the need for further research. The most obvious area is to confirm whether the reference price cues of this study extend to other contexts and categories. For example, consumers could use a median, mean, or some midpoint for the contextual reference price. When competition is between two major brands, such as Coke and Pepsi, the contextual reference price could be even more important than in crackers, with the discounted premium brand probably serving as the reference price. This difference would be captured by our model, though the importance and thus parameter estimates for the various measures of contextual reference price would be different. Similarly, if a category has a prototypi-
cal brand, its price may well be the best contextual reference price in our model.

If brand switching across sizes is common, we could not consider each size as a separate category. Choice among sizes could then be modeled as a logit either nested over brand and size or with each brand-size as an independent alternative (e.g., Guadagni and Little 1983). Our price specification then could be used with either form of the logit. If coupons constitute a major promotional tool, theory would need to specify whether loyal consumers use coupons more and how consumers in each group integrate coupons into their reference price. This issue gets more complex as one overlays interstore competition that doubles and triples coupons. Coupons also take many forms such as mail, in-pack, on-pack, and so on, each of which would have a different use with different implications for reference price (Dhar and Raju 1992).

The problem takes a different complexion when intercategory effects are influenced by price. For example, Mulhern and Leone (1991) have shown that sales of cake mix frostings are influenced by prices and sales of cake mixes. One approach to solve the problem would be to consider the effective reference price of a category as the selling price less discounts within the category and discounts in the related category to the proportion that they affect sales.

The formation of reference prices could be slightly different for durables. Though the contextual reference prices within the store could play the same role as for grocery products, the temporal (memory-based) reference prices would be influenced as much by past prices paid as by ads, phone calls, conversations with friends, and shopping between stores. For this reason, though scanner data for durables is available, it may have to be supplemented by surveys on these other dimensions. A good question is whether contextual reference prices would be as important as for grocery products.

In any case, research on reference prices can benefit from triangulation research using multiple methods. Surveys and experiments can shed valuable light on the process. The former would provide direct response from consumers about their uses of price. Careful design of such surveys can mitigate demand artifacts. Experiments provide rigorous tests of the causes of reference price and are especially useful in developing theory. Indeed, insight from scanner data can be enhanced greatly when combined with surveys and experiments.

Managerial Implications
To obtain meaningful and valid managerial implications, we carried out a simulation of our model that explored the optimum pricing to maximize profits given our findings on consumer response. Following Neslin, Powell, and Schneider’s (1992) and Tellis and Zufryden’s (1992) study, we used simulation rather than an analytic solution because our basic model involves a dynamic (three period lagged) multinomial logit, for which a closed form solution (if one exists) is very difficult. The setup, execution, and results of a detailed simulation is a study in itself. For brevity, we present here only the results of a very simple simulation (inspired by the Tellis and Zufryden 1992 approach), which used the following structure: get the best retail price for premium that maximizes manufacturer profits assuming (1) the manufacturer dictates prices to the retailer; (2) retail sales depend only on choice (equation 5), not store visits, purchase incidence, or quantity; (3) profits are margins times mean probability of brand choice; (4) prices of three other brands are known in advance (Table 5a); and (5) consider only the first 8 weeks of a 12-week horizon of constant margins.

This simple simulation still yields some interesting results, which can be viewed from three perspectives: (1) as propositions that can be subjected to a formal optimization that relaxes these assumptions; (2) as managerial implications that can be validated by more experimental or survey research; or (3) as reference price explanations for some retail pricing phenomena. In each case we give as precise a verbal explanation for the result as we can.

1. Wrongly including only price and ignoring the contextual and temporal reference prices leads to a suboptimal everyday low price. Including only a price term could be inadequate when temporal and contextual reference prices also are required; in this case, the manager would wrongly set a suboptimal everyday low price (EDLP, Table 5b, row 1). The low price is the result of responding to competing prices. The everyday low price is the result of ignoring the dynamics of price—the consumer recalls recent past prices and is less responsive to current low prices if they also occurred in the immediate past. By ignoring this behavior the manager does not increase the price after a discount to raise the consumer’s reference price. However, if only minimal sales occur at the high (list) price, firms can be restrained from maintaining or advertising such list prices by concern for being truthful or legal restrictions (Kaufman, Ortmeier, and Smith 1992; Tellis 1986). Also, low sales at the high price may not justify the cost and effort of price changes, so that firms can resort more profitably to EDLP, as some grocery retailers and manufacturers recently have.

However, firms could not raise price without risk and would not turn off regular buyers; and even high prices would not be credible. Our simulation merely assumes that the high price is the competitive list price.

2. Wrongly including only a temporal reference price and excluding the contextual reference price leads to a suboptimal everyday “high” price (EDHP, Table 5b, row 2). The high price is due to the lack of sensitivity to low competitive prices through the contextual reference price. The everyday high price is because the inclusion of the temporal reference price indicates that any discount would lead consumers to expect continued discounts and thus is unprofitable in the long run. Though everyday “high” prices are a rarity for most consumer products, it occurs for some products and many services (e.g., luxury cars, theater prices, restaurant prices, tuition). In these cases, managers fear that discounting leads to a decline in perceived quality and an inability to sustain demand at high prices.

3. Only the inclusion of both temporal and contextual reference prices leads to an optimal pricing scheme of alternating high (list) and low (discount) prices (Table 5b, row 3). The contextual reference price prompts a response to competitive low prices by requiring a low price for one’s brand at least some of the time. The temporal reference price prompts a high price for at least some of the time to raise

Contextual and Temporal Components of Reference Price / 31
TABLE 5
Description and Results of Simulation for Premium

5a. Prices ($) of Competing Brands by Week

<table>
<thead>
<tr>
<th>Other Brands</th>
<th>Week 1</th>
<th>Week 2</th>
<th>Week 3</th>
<th>Week 4</th>
<th>Week 5</th>
<th>Week 6</th>
<th>Week 7</th>
<th>Week 8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Zesta</td>
<td>1.09</td>
<td>1.09</td>
<td>1.27</td>
<td>1.27</td>
<td>1.27</td>
<td>1.27</td>
<td>1.09</td>
<td>1.09</td>
</tr>
<tr>
<td>Sunshine</td>
<td>1.27</td>
<td>1.27</td>
<td>.89</td>
<td>1.27</td>
<td>1.27</td>
<td>1.27</td>
<td>1.27</td>
<td>1.27</td>
</tr>
<tr>
<td>Pvt Label</td>
<td>.69</td>
<td>.69</td>
<td>.69</td>
<td>.69</td>
<td>.69</td>
<td>.59</td>
<td>.69</td>
<td>.69</td>
</tr>
</tbody>
</table>

5b. Optimum Prices for Premium by Model and Week for Retail Margin = 53%

<table>
<thead>
<tr>
<th>Model</th>
<th>Time In Weeks</th>
<th>Price Strategy</th>
<th>Sale Unit</th>
<th>Profits $</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1  2  3  4  5  6  7  8</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Price Only</td>
<td>1.12 1.12 1.11 1.12 1.11 1.12 1.12</td>
<td>EDLP</td>
<td>7416 14337</td>
<td></td>
</tr>
<tr>
<td>Temporal Only</td>
<td>1.27 1.27 1.27 1.27 1.27 1.27 1.27</td>
<td>EDHP</td>
<td>6665 9947</td>
<td></td>
</tr>
<tr>
<td>Contextual + Temporal</td>
<td>1.07 1.03 1.27 1.05 1.27 1.09 1.05</td>
<td>Variable</td>
<td>7795 16779</td>
<td></td>
</tr>
</tbody>
</table>

5c. Increase in Sales and Profits With Proposed Model (temporal and contextual price comparison)

<table>
<thead>
<tr>
<th>Over Base Models (only price, or only temporal price comparison)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Base Model</td>
</tr>
<tr>
<td>Retail Margins % of Selling Price</td>
</tr>
<tr>
<td>%</td>
</tr>
<tr>
<td>----</td>
</tr>
<tr>
<td>37</td>
</tr>
<tr>
<td>45</td>
</tr>
<tr>
<td>53</td>
</tr>
<tr>
<td>60</td>
</tr>
</tbody>
</table>

the consumers' temporal reference price, so that any subsequent low price then can be perceived as a discount. The combination of temporal and contextual reference prices leads to an optimal scheme of alternating high and low prices, which is typical for most consumer products.

This explanation for alternating high and low prices is independent of the price discrimination theory, which also has been proposed to explain the temporal variation in retail prices (see Tellis 1986 for a review). According to this view, the low price is meant to appeal to price-sensitive or informed consumers who would be willing to wait for a discount or switch to a discounted brand, whereas the high (list price) is maintained to capture consumer surplus from uninformed, impatient, or loyal consumers (Varian 1980; Sobel 1984).

The two explanations are not contradictory and actually could be complementary. For example, the price-sensitive consumers (of the price discrimination theory) could be the ones that respond to price via temporal and contextual reference prices (of the reference price theory). The price discrimination theory focuses on the impact of multiple consumer segments on temporal price variation. The reference price theory focuses on how consumer response in one or more segments affects price variation.

Both explanations require manufacturers to provide retailers with a sufficiently large margin so the retailers can vary their prices by discounting to consumers from time to time. Thus, Procter & Gamble's shift to an EDLP to retailers could prevent retailers from offering discounts because they do not have the margin with which to play. But neither explanation requires that manufacturers themselves vary prices by offering periodic discounts to the trade.

4. Wrongly ignoring the contextual reference price leads to substantially lower sales and profits. If the true response model should contain both contextual and temporal reference price terms (as the statistical analysis shows), then using one that contains only price or temporal price terms leads to suboptimal profits (Table 5b, last three columns). The reason is that the manager sets the same everyday price that is either too high or too low; these suboptimal higher or lower prices lead to substantially lower sales and profits.

5. The gain from including contextual reference prices increases with higher retail margins (Table 5c). The reason for this gain is that the profit curve gets steeper at higher margins, and optimal prices lie farther apart under each of the rival price response models. So the gain in sales by including contextual reference prices instead of price alone is higher as margins increase, ranging from 14% to 19% for retail margins of 37% to 60% (Table 5c). The corresponding sales increase from using a contextual and temporal reference price instead of temporal reference price alone ranges from 14% to 98%. These proportions are massive because contextual reference prices are more critical in our data. Because margins for grocery products are small, the corresponding increase in profit are not as dramatic. Nevertheless, they are by no means negligible. In general, as retail margins per unit increase, managers have more to gain by properly specifying the price response curve, especially in terms of sales and when reference prices are important.

Conclusion

Though a consensus is developing on the importance of the consumer's reference price in brand choice, empirical research has focused primarily on the temporal reference price. We argue that consumers' reference price could be affected by both a temporal and a contextual component; the contextual component could be even more important because of the primacy of contextual effects at the point of pur-
cause of the primacy of contextual effects at the point of purchase. We tested this proposition with an empirical analysis across four markets of the cracker category, using three conditions within each market and different measures for reference price. We also explored the managerial implications of the results by simulation.

The key empirical results are as follows:

1. Both the contextual and temporal reference prices are significant predictors of consumer choice in all four markets.
2. The contextual component is at least as strong as the temporal component in general, but it is stronger when brand preference is weaker, brand sampling is wider, and shopping is infrequent.
3. The low-priced brand tested as the most important measure for the contextual reference price. A moving average of past prices of each brand tested as the most important measure for the temporal reference price.

The key managerial implications are as follows:

4. Managerial focus on price alone, while ignoring the contextual and temporal reference prices, could lead to a suboptimal everyday low price.
5. In contrast, a focus on temporal reference price alone (excluding the contextual reference price) could lead to a suboptimal everyday high price.
6. Only the inclusion of both temporal and contextual reference price leads to an optimal pricing scheme of alternating high (list) and low (discount) price. This explanation for variable pricing is quite independent from that based on segmented markets that appears in the literature.
7. Wrongly ignoring the contextual reference prices could lead to substantially lower sales and lower profits especially at higher retail margins.

Though researchers must explore the generalizability of these results across different categories, competitive scenarios, and measures, the method itself could be immediately applicable to retail managers who seek to price scientifically.

REFERENCES


