The Effect of Reference Point Diagnosticity on Attractiveness and Intentions Ratings

Kwanho Suk

Song-Oh Yoon

Donald R. Lichtenstein

Sie Yeoun Song*

AUTHOR NOTE

Kwanho Suk is assistant professor of marketing at School of Business, Korea University, Seoul 136-701, Korea (phone: 822-3290-2612, fax: 822-3290-1307, e-mail: ksuk@korea.ac.kr). Song Oh Yoon is assistant professor of marketing at School of Business, Korea University, Seoul 136-701, Korea (Phone: 822-3290-2829, email: soyoon@korea.ac.kr). Donald R. Lichtenstein is professor of marketing, University of Colorado, Boulder (Phone: 303-492-8206; fax: 303-492-5962; email: Donald.Lichtenstein@Colorado.edu). Sie Yeoun Song is a doctoral student at School of Business, Korea University, Seoul 136-701, Korea (e-mail: hazmin@korea.ac.kr). Correspondence: Song-Oh Yoon. The authors express their deepest gratitude to John Lynch for his numerous insights on this research. We also thank Chris Janiszewski, Jaehwan Kim, Jong-Won Park, and members of the Korea University B.E.S.T. Marketing Group for their valuable suggestions. This research is supported by Korea University Grant to Kwanho Suk (K0517961) and to Song-Oh Yoon (K0714521).

Abstract

Researchers have long been interested in understanding cognitive processing differences across consumer judgments and choices. Despite representing a focal outcome in much research, less attention has been focused on the “choice-like” response of behavioral intentions. In this research we compare processing differences in the formulation of judgments of attractiveness and intentions. Based on the premise that different goals underlie these responses, we hypothesize that they differentially recruit alternative reference points due to differential reference point diagnosticity. We test this prediction in the domain of price attractiveness and purchase intentions ratings. Study 1 provides evidence that endpoints of the range of alternative prices are more predictive of ratings of price attractiveness than of purchase intentions, while price rank and distribution mean are more predictive of purchase intentions ratings than of price attractiveness ratings. Study 2 replicates this effect using a differing methodology. Finally, Study 3 provides a test of the external validity of these findings in a multi-cue setting.

KEY WORDS: Context Effects, Judgments and Intentions, Range Effects, Frequency Effects, Diagnosticity

An enduring theme in research on consumer behavior, social cognition, and behavioral decision-making has been that all consumer judgment is relative. Consumers rate or choose some brand or object relative to context that is externally available or recruited from memory (Hardie, Johnson, and Fader 1993; Herr 1986; Kahneman and Miller 1986; Lynch, Chakravarti, and Mitra 1991; Mayhew and Winer 1992; Miller and Prentice 1996; Mussweiler 2003; Schwarz and Bless 1992; Sherif and Hovland 1961). Which standards of comparison are recruited and how are they processed to evaluate some target?

The present manuscript investigates task effects on which elements of context are recruited and used to make judgments. In particular, we ask the question of whether the contextual elements that influence intention ratings differ from those that influence ratings of the attractiveness of a product or one of its features. An immense literature in consumer behavior and marketing has relied on two kinds of judgments: attractiveness ratings and ratings of behavioral intention. Some studies rely on one or the other, as in the large literature on forecasting purchase behavior by intentions (Alexander, Lynch, and Wang 2008; Clancy, Krieg, and Wolf 2006; Morwitz 2001; Morwitz and Schmittlein 1992). Other research treats intention ratings and attractiveness ratings as interchangeable, with similar predictions for effects of independent variables on these different measures (e.g., Howlett et al. 2009; Menon, Block, and Ramanathan 2002). In other cases both measures are collected in the same study and attractiveness perceptions are treated as antecedents of intentions (Anderson and Sullivan 1993; Schneider et al. 2005). Nothing in the literature to date has suggested that context effects might operate differently for these two classic marketing and consumer research dependent variables.

We develop and test a theory of why intentions and attractiveness ratings would be differentially susceptible to certain classic context effects in the literature, due to task differences in which reference points are recruited and perceived as diagnostic. In particular, many papers have pointed to effects of the average of a set of contextual stimuli (as in Helson's adaptation level theory and most work in marketing on reference prices); to the range produced by a set of context stimuli (e.g., Janiszewski and Lichtenstein 1999; Lynch, Chakravarti, and Mitra 1991; Hutchinson 1983; Volkman 1951), and the relative frequency or rank of a target within a context set (e.g., Parducci 1965; Niedrich, Sharma, and Wedell 2001). We argue that for attractiveness ratings, the highest and lowest stimuli seem most relevant, but that for intention ratings, the mean of the context set and the rank of the target in the context set will matter most.

The basis for our argument is that intention is more "choice-like" than attractiveness ratings. Typically, to intend to buy or select one alternative implies that one will not select other alternatives in the salient set. In contrast, for attractiveness ratings, there is no bar to giving the highest possible rating to more than one alternative. Payne (1982) has argued that for choice, an input is diagnostic if it separates the best alternative from the rest (See also Lynch, Marmorstein, and Weigold 1988). In the same way, we posit that comparison standards will seem diagnostic if they help one rating intentions ascertain whether a focal alternative is the best in the set. We argue that rank order information is particularly valuable for rating intentions, and that consumers rating intention will be particularly sensitive to rank order changes in a target when it is toward the "good" end of some distribution. Because it is not always obvious what the ordinal position of a target is in a set when stimuli are not sorted or when people must rely on memory, we also posit that consumers making intention judgments will be sensitive whether a target stimulus is better than average. We use simple algebraic models used in prior literature to determine conditions that make context mean, context range, and context relative frequency affect attractiveness ratings and purchase intention ratings for a set of target products.

In testing our predictions of differential recruitment of alternative reference points, a necessary condition is that a study context be chosen such that the focal attribute is salient enough to motivate consumers to expend the effort necessary to engage in information recruitment. Given that price is generally considered to be one of the most salient attributes across consumers, product categories, and purchase contexts (Briesch et al. 1997; Niedrich, Sharma, and Wedell 2001; Winer 1986), we test our hypotheses in the context of price distributions. It also happens to be the case that both price attractiveness (e.g., Adaval and Monroe 2002; Danziger and Segev 2006; Grewal, Marmorstein, and Sharma 1996; Grewal, Monroe, and Krishnan 1998; Inman, Peter, and Raghubir 1997; Janiszewski and Lichtenstein 1999; Lichtenstein and Bearden 1989; Niedrich, Sharma, and Wedell 2001; Urbany, Bearden, and Weilbaker 1988) and purchase intention (e.g., Baker et al. 2002; Dodds, Monroe, and Grewal 1991; Grewal, Monroe, and Krishnan 1998; Inman, Peter, and Raghubir 1997; Lalwani and Monroe 2005; Petroshius and Monroe 1987) are two of the most commonly assessed outcomes across pricing studies.[1]

In the balance of the paper, we first describe Range (Volkman 1951), Range-Frequency (Parducci 1965), and Adaptation-Level (Helson 1964) Theories as these theories are most often offered as theoretical bases for the use of distribution range endpoints, distribution rank, and distribution mean as reference points, respectively. Then based on Norm Theory (Kahneman and Miller 1986) and principles of diagnosticity (Feldman and Lynch 1988; Lynch, Marmorstein, and Weigold 1988), we offer rationale and hypotheses predicting: (1) stronger range effects on ratings of price attractiveness than on ratings of intentions, and (2) stronger rank and mean

[1] We should note that in addition to using price distributions, we also tested study hypotheses using miles per gallon for automobiles and selection level for movie rentals as the focal attributes. Results were largely directionally consistent with hypotheses, albeit also largely nonsignificant. While space limitations preclude a description of these studies, these studies are described, and the results are provided, in the Web Appendix. We speculate that the reason for the weaker results found for these two attributes and the stronger results found for price (reported subsequently) lies in the differential salience of the attributes to consumers.

effects on ratings of intentions than on ratings of price attractiveness. We then describe three experiments designed to test our hypotheses. We conclude with a discussion of the implications our findings.

THEORETICAL BACKGROUND

As noted above, the three most commonly offered theories for the use of alternative reference points are Range Theory, Range-Frequency Theory, and Adaptation-Level Theory. Applied to price perception, Range Theory (Volkman 1951) posits that there is a linear relationship between the product category price range and psychological perceptions of price. Consumers are hypothesized to assign the lower bound of the range of prices to the upper (i.e., favorable) anchor of their price perception scale, the upper bound of the range of prices to the lower (i.e., unfavorable) anchor of their price perception scale, and then do a linear mapping of observed prices to price perceptions. For example, consider that a consumer is asked to evaluate a price of $1.25 from a category with a price range of $0.75 to $1.75. As $1.25 lies equidistant between $0.75 and $1.75, it is mapped to the midpoint of the price perception scale (i.e., it is perceived as neutral). All else being equal, if the top end of the range were to be higher ($0.75 and $2.00), or the bottom end of the range were to be lower ($0.50 and $1.75), the $1.25 price would be judged as more and less favorably, respectively (Janiszewski and Lichtenstein 1999).

As evident from this illustration, applied to price perception, Range Theory predicts that only two prices, the upper and lower bounds of the range of encountered prices, are instrumental in affecting perceptions of a target price. However, according to Range-Frequency Theory (Parducci 1965), the perception of a price is influenced by a joint function of its distance to the highest and the lowest prices (i.e., range principle) and its rank within the given price set (i.e., frequency principle). Therefore, holding range constant, the frequency principle would posit that a target price will be perceived as lower/higher the more prices within the category fall above/below it. For example, in the example above, $1.25 will be judged less favorably within the range of market prices the more brand alternatives that fall on the low side of $1.25.

Finally, a third theory often used to explain how consumers form reference prices is Adaptation-Level Theory (Helson 1964). Applied to price perception, this theory posits that consumers evaluate target prices by comparing them to the arithmetic mean of previously encountered market prices, or a geometric mean where market prices encountered more recently are weighted more heavily. According to this theory then, target prices are perceived as lower/higher as they deviate more on low/high side of the mean of the price distribution.

The issue at hand relates to differential recruitment and use of references prices based on these three theories across response tasks of judgments versus intentions. We believe that Norm Theory (Kahneman and Miller 1986) in conjunction with principles of diagnosticity (Feldman and Lynch 1988; Lynch, Marmorstein, and Weigold 1988) provides a theoretical basis for offering predictions for such differential recruitment and use. Specifically, according to Norm Theory, each time a person is asked to evaluate a stimulus (e.g., product X at price Y), it acts as a probe and selectively recruits a context set of alternatives (e.g., alternative products at various prices) that serve as a frame of reference for the stimulus evaluation. These representations are generated in parallel and form a distribution with mean, mode, and range for each attribute that is relevant. Of critical importance, Kahneman and Miller (1986) emphasize that the evoked set of alternatives recruited by a probe is context-dependent. Based on the premise of differences in reference point diagnosticity in accomplishing attractiveness versus intention goals, we hypothesize that one such context-dependency relates to that of response task, attractiveness versus intention ratings.

Couched in the context of the present study, diagnosticity refers to the sufficiency or informativeness of the alternative price reference point(s) to accomplish the objectives for the decision task at hand (Lynch, Marmorstein, and Weigold 1988). We believe that range endpoints allow consumers to formulate judgments of price attractiveness more so than intentions, while stimulus rank and distribution mean allow consumers to formulate judgments of intentions more so than price attractiveness. Underlying this prediction that alternative reference points are differentially diagnostic for the two price ratings is the premise that prices may be evaluated independently of each other, but each relative to range endpoints, in forming price attractiveness judgments. However, as purchase intentions are "choice like," they necessarily require that the options be evaluated relative to each other in order for effective discrimination between options. For instance, based on proximity to range endpoints, we may simultaneously judge prices of many competing alternatives as all high or low or as all attractive or unattractive. Hence, that one price may be seen as attractive (relative to range endpoints) is not diagnostic that a second price is precluded from a similar evaluation. However, since an intention to purchase one item precludes the purchase of (or lowers the purchase probability for) competing items, there in essence is a comparative evaluation in that an intention to purchase will necessarily include evaluations of stimuli relative to each other. Hence, alternatives are evaluated more heavily relative to each other rather than relative to range endpoints. As such, the distribution endpoints become less diagnostic, and reference prices that are more indicative of how the target brand relates to the prices of alternatives (rank or mean) become more diagnostic, for intentions.

To illustrate our rationale way of an example, consider a distribution of fast food restaurants ranging in price from $4 to $11 for a meal. A price of $6 will be judged as relatively attractive due to a significant range effect, and less influenced by the number of options between

$4 and $6 (a less significant frequency effect). The reason is that ratings of judgments of attractiveness require no rating of one at the expense of others; numerous competing restaurants in the $4-$6 can also be seen as having attractive prices (range endpoints are more diagnostic for this judgment rating than is distribution rank or mean). Consequently, consistent with previous research supporting the recruitment of range endpoints for judgments of price attractiveness (Janiszewski and Lichtenstein 1999; Niedrich, Sharma, and Wedell 2001), for reasons of diagnosticity, we hypothesize that range endpoints will exert more influence on price attractiveness than intentions.

However, because forming an intention is "choice-like" in that an intention to patronize one restaurant lowers the probability of patronizing others, forming an intention is more likely to necessitate a process of comparing alternatives to each other. Because of this, a target alternative evaluated in the context of a purchase intention is more likely to recruit viable alternatives so that the comparisons can be made. Therefore, as the number of cheaper purchase options increases (affecting both distribution mean and target price rank), purchase intention for the target restaurant decreases. For this reason, the distribution of prices is predicted to be more diagnostic for purchase intentions than are the price endpoints, manifesting itself in either a frequency (rank) or adaptation-level (mean) effect. This prediction is consistent with previous research which suggests that, relative to a judgment task, an input's diagnosticity in a choice task is more likely to be a function of whether it enables a decision maker to discriminate among alternatives, that is, whether it separates "the best from the rest" (Lynch, Marmorstein, and Weigold 1988; Payne 1982).

H1: Price range will have a more pronounced effect on ratings of price attractiveness than on ratings of intentions.

H2: Price mean and rank of the target price will have a more pronounced effect on ratings of intentions than on ratings of price attractiveness.

STUDY 1

Method

Study 1 was designed to test H1 and H2. A total of 228 undergraduate students were randomly assigned to a 3 (price distribution: negative-skew low-mean vs. positive-skew low-mean vs. negative-skew high-mean) x 2 (response task: attractiveness rating vs. intention rating) between-subjects design. The study was conducted in a computer lab and 2 to 12 students participated in each session. Participants were told that we were interested in their responses to prices for a one-night stay at 25 different hotels (20 contexts and 5 targets) located on a popular beach (fictitious name) in Australia. They were told the prices would be presented sequentially on their monitor.

Participants were then exposed to one of the three distributions of 20 context prices, where prices were presented sequentially and in random order. The three distributions, shown in Appendix A, differed in range, mean price, and skewness (see Lim 1995; Niedrich, Sharma, and Wedell 2001; Smith, Diener, and Wedell 1989, for similar manipulations). Specifically, in addition to manipulating skewness (as indicated by distribution names), the endpoints of the positive-skew low-mean and negative-skew high-mean distributions were the same ($144-$240), but higher than in the negative-skew low-mean distribution ($96-$192). The mean prices of the negative-skew low-mean and positive-skew low-mean distributions were the same ($168) but lower than that of the negative-skew high-mean distribution ($216). The variation in the range, rank, and mean price across the three distributions allows for testing their influence on the target prices.

After rating each of the 20 context prices in terms of either attractiveness or purchase intentions, participants were exposed to prices for five target hotels (constant across conditions) that were presented sequentially and in random order. After exposure to each of these five hotels, participants again rated either the attractiveness of the hotel at the various prices (1 = unattractive, 9 = attractive) or their purchase intentions for booking a reservation (1 = do not want to purchase, 9 = want to purchase) for each. Manipulation checks followed assessment of these ratings by having respondents identify the lowest, highest, and mean prices in their context distributions from a list of nine prices ($96, $144, $156, $168, $180, $192, $216, $228, and $240).

Preliminary Checks and Manipulation Checks

Initial screening eliminated eight participants who did not complete the study or who pressed wrong keys. We also checked the correlation between response ratings and prices for each participant in order to identify participants who did not answer seriously or wrongly interpreted the scale anchors. We eliminated additional 12 participants whose correlation was either positive or zero, retaining 208 participants. The proportion of the removed participants is not significantly related to either the task ($\chi^2(1) = .33$, $p = .566$) or the distribution ($\chi^2(2) = 3.50$, $p = .174$). The average correlation of the retained participants was -.80.

Manipulation checks were conducted for the perceived lowest, highest, and mean prices of the presented prices. Perceived lowest price in the negative-skew low-mean (M = 113.63) was significantly lower than that in the positive-skew low-mean (M = 149.31; $F(1, 202) = 74.74$, $p < .001$, $\eta^2 = .25$) and the negative-skew high-mean (M = 153.86; $F(1, 202) = 97.31$, $p < .001$, $\eta^2 = .32$) conditions, with no significant difference between the latter two ($F(1, 202) = 1.16$, $p = .282$, $\eta^2 = .004$). Similarly, perceived highest price in the negative-skew low-mean condition (M = 192.68) was significantly lower than those in the positive-skew low-mean (M = 230.98;

$F(1, 202) = 201.21$, $p < .001$, $\eta^2 = .41$) and the negative-skew high-mean (M = 233.18; F(1, 202) = 230.43, $p < .001$, $\eta^2 = .46$) conditions. The latter two conditions did not differ (F(1, 202) < 1.0, $\eta^2 = .001$). These results indicate that the range manipulation was successful. Also, the perceived mean price in the negative-skew high-mean condition (M = 199.77) was significantly higher than those in the negative-skew low-mean (M = 167.67; F(1, 202) = 226.97, $p < .001$, $\eta^2 = .50$) and the positive-skew low-mean ($M = 173.27$; $F(1, 202) = 144.14$, $p < .001$, $\eta^2 = .31$) conditions. While the difference between the negative-skew low-mean and positive-skew low-mean conditions was also significant ($F(1, 202) = 6.74$, $p < .010$, $\eta^2 = .01$), this difference was significantly smaller compared with the differences from the negative-skew high-mean ($F(1, 202) = 68.68$, $p < .01$). Thus, the manipulation of the mean price was also successful.

Model Fit Tests

We tested the fits of the four alternative models: the Range, Frequency, Adaptation-Level, and Range-Frequency models (Model specifications are described in Appendix B). For each task-distribution condition, the average response rating for the five target prices was computed (see Table 1 and Figure 1) and used as input for nonlinear regression analyses. This yielded 15 aggregate mean ratings for each of the two rating task conditions that were used as dependent variables for model fitting. This tested the extent to which each model explains the patterns of the responses across the three distributions.[2] Table 2 presents the estimated coefficients and the goodness-of-fit indices (R^2) for each price model in different task conditions.

Insert Tables 1 and 2 and Figure 1 about here

----

[2] In addition to the reported results, we also conducted fit tests including two more comparable distributions. In a set of tests, we included the negative-skew low-mean and the positive-skew low-mean distributions that have the same mean price. In another set of tests, we included the positive-skew low-mean and the negative-skew high-mean that have the same range. The results are consistent with the reported analyses.

Comparisons of the fits of alternative models between the two task conditions show results supporting H1 and H2. The fit of Range model to the data was greater for attractiveness than intention (R^2 = .850 vs. .544). Conversely, both the Frequency and Adaptation-Level models performed better in explaining purchase intentions than attractiveness judgments (Frequency: R^2 = .943 vs. .760; Adaptation-Level: R^2 = .924 vs. .785). While the Range-Frequency model fits both the attractiveness (R^2 = .981) and purchase intention (R^2 = .952) ratings very well, the relative weight of the range (w) is substantially greater for the attractiveness than the purchase intention (w = .76 vs. .54) rating, showing a greater reliance on the range feature for the attractiveness rating and the greater dependence on the frequency feature for the purchase intention rating.

To assess the statistical significances of the observed model fit differences, the standardized beta coefficients (which are equivalent to the fit index, R, since there is only one predictor variable) of the Frequency and the Adaptation-Level models were transformed to Fisher's Z and were compared across the two rating tasks. As expected, the standardized beta coefficient for the Frequency model was statistically greater for the purchase intention rating than for the attractiveness rating (beta = -.97 vs. -.87; Fisher's Z = -2.11 vs. -1.35; z = 2.08, *p* = .038). Also, the standardized beta coefficient of the Adaptation-Level model was greater for the purchase intention rating than for the attractiveness rating (beta = -.96 vs. -.88; Fisher's Z = -1.98 vs. -1.38; z = 1.67, $p < .094$). These results support H2.

Since the Range model has four coefficients, it does not allow for direct comparisons of the fits using the standardized betas. Thus, the range effect was tested indirectly by adding an interaction term to Range-Frequency model: $R_{ik} = 9\text{-}8[(w + w_i\text{*task})(\$_{i,k} - \$_{min,k})/(\$_{max,k} - \$_{min,k}) + (1 – w - w_i\text{*task})F_{ik}]$, where w_i*task is the interaction between the strength of the range relative to the frequency effect (w) and the decision task dummy (0 = attractiveness rating, 1 = purchase intention rating). The results showed that the interaction was significant ($w_i = -.18$, SE = .04; t = 4.50, $p < .001$). The significant interaction with a negative sign means that the range effect is significantly stronger for attractiveness judgment than purchase intentions, further supporting H1.

ANOVA Tests

Although the results of model fit tests support our hypotheses, this method is based on aggregate data and thus does not reflect the variability of the participant in their responses to the price information. Thus, we test the models using a 2 (task) x 3 (distribution) x 5 (targets: $144, $156, $168, $180, $192) mixed ANOVA. The analyses allow for statistical tests recognizing within-participant variability in order to examine whether the patterns of the average ratings are consistent with model predictions. The influence of task and distribution on target price ratings were examined via the planned contrasts.[3]

Average favorableness of the targets. First, we compared the average attractiveness and intention ratings across distributions to examine whether the pattern of the means is as predicted by the Range, Frequency, and the Adaptation-Level models. We focus our analyses on the comparison between the negative-skew low-mean and the positive-skew low-mean and the comparison between the positive-skew low-mean and the negative-skew high-mean, because the three models predictions differ for these two sets of pairs.

First, the overall ratings of the negative-skew low-mean and the positive-skew low-mean distributions were compared. The two distributions have the same mean but they differ in their shape and range (see Appendix A). The Range model predicts that the overall rating should be

[3] The results of a 2 x 3 x 5 mixed ANOVA showed that the main effect of target prices ($F(4, 808) = 192.57$, $p < .001$) and distribution ($F(2, 202) = 16.77$, $p < .001$), the interaction between target and distribution ($F(4, 808) = 8.93$, $p < .001$), and the interaction between task and distribution ($F(2, 202) = 3.44$, $p = .034$) were significant. Although the three-way interaction was not significant ($F(4, 808) = .989$, $p = .413$), the expected three-way interaction involving quadratic trend (task x distribution x quadratic trend of the target) was significant ($F(2, 202) = 5.02$, $p = .007$).

more favorable in the positive-skew low-mean than the negative-skew low-mean distribution because targets in the former distribution are located near the highest end, and the targets in the latter are located near the lowest end. The Frequency model, on the contrary, predicts that the evaluation will be more favorable in the negative-skew low-mean (average rank of the target = 11.4) than in the positive-skew low-mean distribution (average rank of the target = 14.6), but the predicted difference is not big. On the other hand, the Adaptation-Level model predicts that there will be no difference between the two distributions because they have the same mean. Planned contrast within the attractiveness rating task showed that the average rating of the positive-skew low-mean ($M = 5.25$) is greater than that of the negative-skew low-mean ($M = 4.72$), as predicted by the range model, and the difference was marginally significant ($F(1, 202) = 3.15$, $p = .077$). In purchase intention rating, there was no significant difference between the negative-skew low-mean ($M = 4.94$) and the positive-skew low-mean ($M = 4.72$; $F < 1.0$). This moderating role of task was supported by a marginally significant interaction between task and distribution ($F(1, 202) = 2.94$, $p = .088$). Consistent with H1 and H2, these results again suggest that the Range effect is stronger for attractiveness rating, but the Adaptation-Level effect is more prominent for intention rating.

Next, the average ratings of the positive-skew low-mean and the negative-skew high-mean distributions were compared. The two distributions have the same price range, thus the Range model predicts that the ratings will not be different between the two distribution conditions. According to the Frequency and the Adapation-Level models, however, targets will be more favorably evaluated in the negative-skew high-mean condition as the mean context price is higher ($216 vs. $168) and the ranks of the targets are lower (5.5 vs. 14.6) in this distribution than the positive-skew low-mean distribution. Cell mean contrasts within the attractiveness rating task showed that the average rating in the negative-skew high-mean distribution (M = 5.99) was higher than that in the positive-skew low-mean distribution (M = 5.25; $F(1, 202) = 5.32$, $p = .023$). This pattern does not agree with the Range model. Instead, it is consistent with the Frequency and the Adaptation-Level models. In the purchase rating conditions, the rating was also significantly higher in the negative-skew high-mean than in the positive-skew low-mean (M = 6.49 vs. 4.72; $F(1, 202) = 29.15$, $p < .001$), supporting the Frequency and the Adaptation-Level models. Although, contrary to our predictions, the adaptation-level and the frequency effects were significant in both tasks, the significant task by distribution interaction ($F = 5.32$, $p = .022$) indicates that the relative impact of the frequency and the adaptation-level effects are stronger for the purchase rating than the attractiveness rating, as predicted by H2.

Trend contrast. To assess the differential impact of the frequency feature on rating tasks, trend contrasts were conducted. Contrary to the Range or the Adaptation-Level models, the Frequency model predicts that the relationship between the target prices and their ratings will be nonlinear for the negative-skew low-mean and the positive-skew low-mean distributions in which the changes in the rank of the target are not linear (see Appendix A). In the negative-skew low-mean distribution, the change in the rank of the targets is greater for the higher price range of the distribution. Therefore, the Frequency model predicts there will be bigger decreases in rating scores for the higher target prices, resulting in a (decreasing) concave pattern. On the contrary, in the positive-skew low-mean distribution, there are bigger changes in ranks for lower target prices. Thus, the frequency model predicts that there will be bigger changes in rating for lower price targets, resulting in a convex pattern. Such a nonlinear pattern is not expected for the negative-skew high-mean in which the ranks of the target changes in a linear fashion.

Figure 1 presents the patterns of the average ratings for each task and distribution condition. For the purchase rating task, a concave quadratic trend was significant for the negative-skew low-mean (F(1, 202) = 4.90, *p* = .028, $\hat{\psi}_{quadratic}$ = -.92) and a convex quadratic trend was significant for the positive-skew low-mean (F(1, 202) = 25.72, *p* < .001, $\hat{\psi}_{quadratic}$ = 2.49). However, as expected, the trend was not significant for the negative-skew high-mean (F(1, 202) = 1.25, *p* = .264). These results suggest that purchase intention was affected by the rank of the target within distribution. In the attractiveness rating task, however, the quadratic trend was not significant in all distributions (F's < 2.30, *p*'s> .131), showing a minimal impact of rank on attractiveness rating task. This was supported by a significant three-way interaction involving task, distribution, and quadratic trend (F(2, 202) = 5.02, *p* = .007).

Additional Tests on the Diagnosticity of the Rank Information

One additional prediction based on the hypothesized diagnosticity of rank for intention responses is the asymmetric impact of rank on purchase intention ratings. As intentions relates to "separating the best from the rest," the effect of changes in the rank of a target alternative on intentions ratings should be more pronounced when the changes occur at a portion of the distribution that is more favorable, hence more relevant for impacting intentions. For example, a change in rank from the 2nd lowest price to the absolute lowest price should have a more pronounced effect on purchase intentions for the target alternative than a change in rank from the 5th to the 4th. This prediction was tested by examining the sensitivity in actual ratings relative to the predicted ratings by the frequency model. The sensitivity score was calculated as the ratio of the differences in rating scores for the two adjacent target prices to those predicted by the estimated model. For example, the sensitivity score for the purchase intention rating of the $144 target and $156 target in the negative-skew low-mean distribution was computed by the ratio of the difference in purchase intention rating scores (i.e., 6.10 vs. 5.59) to the difference estimated by the frequency model (i.e., 5.93 vs. 5.68). Thus, the score greater than one indicates that the actual ratings are more responsive to the change in the target rank compared to those predicted by the model. If the rating of target price is more responsive to the change in its rank in a favorable region of the distribution, the sensitivity score should be negatively correlated with the price level of the target price. Also, this should be observed for the purchase intention ratings but not for the attractiveness ratings. In the analysis, negative-skew low-mean and positive-skew low-mean were included because the target prices of these distribution conditions include both positive and negative frequency indices. Correlation between the price level and the sensitivity was negative and significant (r = -.45, p = .077) for the purchase intention rating, but not significant (r = -.15, p = .571) for attractiveness rating. These results are consistent with our prediction that purchase intention ratings are influenced more by its rank when target price falls into a favorable (vs. unfavorable) part of the distribution. However, the fact that attractiveness rating shows no significant correlation with the price level is consistent with our assumption that judgments involve comparisons with both positive and negative ends of the context prices.

Discussion

Results of the current study provide hypothesis-consistent evidence that range endpoints exert more influence on judgments than intentions (H1), while both rank and mean exert more influence on intentions than judgments (H2). Although our hypotheses were supported, the stimuli used in Study 1 have some limitations due to high correlations among the Range, Frequency, and Adaptation-Level models (*r* ranges between .52 and .90). Although it is natural that the predictions of the three models are correlated, it may make it difficult to pit one model against another. To overcome the potential problem, in Study 2, we independently manipulate one dimension while holding the other dimensions remain constant.

STUDY 2

In this study, each of the three features of the price distribution (i.e., mean, highest and lowest prices, and ordinal rank of the target price) was independently manipulated. Participants were asked to rate the target brand after examining prices of 10 brands in a hypothetical product category, which were presented either in ascending or random order. Depending on the task condition, they rated the attractiveness of the target price (1 = unattractive, 7 = attractive) or their intention to purchase the brand (1 = unlikely to purchase, 7 = likely to purchase).

Method

Participants and design. A total of 744 undergraduate students were randomly assigned to a 2 (task: price attractiveness rating vs. purchase intention rating) x 3 (manipulated dimension of the price distribution: range vs. rank vs. mean) x 2 (direction of the manipulation: high vs. low) x 2 (context price presentation order: ascending vs. random) between-subjects design.[4] Six sets of price distributions were created (see Appendix C) by shifting one of the three dimensions of the price distribution – range, rank, or mean – in an upward (high condition) or downward direction (low condition) from the standard price set consisting of 10 prices that served as a baseline. A Kolmogorov-Smirnov test indicated that the standard price set follows a normal distribution ($z = .32$, $p > .99$). The mean, median, and mid-point of the standard price set was $12.50 (SD = 1.44) and its prices ranged between $10.00 and $15.00. We set the target brand's price at $12.50, the neutral point in the standard price set.

In all six distributions, only the corresponding dimension was manipulated and the other two dimensions were held constant. For example, the price ranged between $8.00 and $15.00 in

[4]We had conjectured that when context prices were presented in systematic order, purchase intention ratings would be more sensitive to rank than to mean, but that when context prices were presented in random order, purchase intention ratings would be more sensitive to the mean than the rank. While results were directionally consistent with this hypothesized interaction, the interaction failed to reach the level of statistical significance. Therefore, due to space limitations and the desire to focus more exclusively on H1 and H2, this interaction hypothesis is not discussed in the paper. We believe this remains an interesting topic for future research.

the low range condition and between $10.00 and $17.00 in the high range condition. However, in both conditions the mean of the prices ($12.50) and the rank of the target price within the set (6th) remained the same. In the low and high rank (frequency) conditions, the target brand was the 7th and 5th least expensive, respectively, while the mean ($12.50) and the range ($10.00-$15.00) in these two groups were held constant. Similarly, the rank (7th and 5th inexpensive) in the high and low frequency conditions and the average of the context prices (low = $12.00 and high = $13.00) in the mean (adaptation-level) conditions were manipulated while other dimensions were held constant. Thus, a comparison between the low and high conditions allowed us to test the independent effect of the manipulated feature of the price distribution. Since the range, mean and the target price's rank are higher in the high (vs. low) condition in each price distribution, the target brand in this condition should be evaluated more favorably than in the low condition.

Procedure and measures. Participants randomly received one of the 24 versions of an experimental booklet. The first page instructed participants to examine a table containing market prices of 10 brands within the same product category. In the ascending order condition, the 10 prices were presented from lowest to highest. In the random order condition, the prices were presented randomly. After viewing these prices, participants were asked to turn to the next page and to evaluate the target price, $12.50, in terms of either price attractiveness or purchase intention. In evaluating the target, participants were allowed to turn to the previous page to examine the prices (i.e., stimulus-based evaluation). Subsequently, participants responded to manipulation checks designed to test how accurately they perceived highest and lowest prices, the mean price, and the rank of the target price within those brands.

Results

Manipulation checks. The manipulation check indicated that the recognized lowest price was lower in the low than the high range condition (M = 8.21 vs. 10.18; F(1, 732) = 466.20, *p* < .001, η^2 = .31) and the recognized highest price was higher in the high versus the low range condition (M = 16.80 vs. 15.17; F(1, 732) = 146.85, p < .001, η^2 = .14). Also, the frequency manipulation was successful since the perceived rank of the target price was significantly higher (i.e., less expensive) in the high rank than the low rank condition (M = 4.70 vs. M = 5.52; F(1, 732) = 37.74, *p* < .001, η^2 = .05). Finally, the mean price manipulation check revealed that the mean price was perceived to be lower in the low versus high mean condition (M = 12.26 vs. 12.82; F(1, 732) = 40.13, *p* < .001, η^2 = .05).

Hypothesis testing: Overview. The magnitude of range, frequency, and adaptation-level effects in each task and presentation order condition was tested by comparing the difference in the rating scores between high and low distributions. To test this, a series of planned partial interaction and cell mean contrast tests were conducted. Since there was no significant effects involving presentation order we collapsed the two order conditions. Thus, significance tests drew the omnibus 2 (task: attractiveness rating vs. purchase intention rating) x 3 (manipulated dimension of the price distribution: range vs. rank vs. mean) x 2 (direction of the manipulation: high vs. low) ANOVA (df = 732, MSE = 1.64).[5] Table 3 presents the average rating of each condition.

Insert Table 3 about here

Range effect. To test H1, a 2 (task: attractiveness vs. purchase intention) x 2 (direction of the manipulation: low vs. high) partial interaction test was run in the range distribution condition.

[5] The results of the ANOVA test showed that, as predicted, the main effect of direction of the manipulation (F(1, 732) = 17.79, p < .001; M_{low} = 3.86 vs. M_{high} = 4.25) and the three-way interaction (F(2, 732) = 5.26, p = .005) were significant.

The results indicated a significant task by manipulation direction interaction effect (F = 3.95, *p* = .047), showing that the range effect is stronger for the attractiveness judgment ($M_{difference}$ = .65) than for the purchase intention rating ($M_{difference}$ = .00). As predicted by H1, the range effect was significant for the attractiveness judgment rating (M_{low} = 3.83 vs. M_{high} = 4.48; F = 8.11, $p < .01$), but not for the purchase intention rating (M_{low} = 3.81 vs. M_{high} = 3.81; F < 1.0).

Frequency effect. To test H2, a 2 (task) x 2 (direction of the manipulation) partial interaction test was conducted for the frequency distribution condition. Results showed a significant task by manipulation direction interaction effect (F = 6.73, *p* = .010). Separate tests for each task condition showed that the frequency effect was significant for the purchase intention rating (M_{low} = 3.71 vs. M_{high} = 4.64; F = 16.20, $p < .001$), but not significant for the attractiveness rating ($_{low}$ = 4.03 vs. M_{high} = 4.11; F < 1.0). These findings support H2.

Adaptation-level effect. The above two-way partial interaction tests and cell-mean contrasts were also applied to the mean distribution in order to provide for the second test of H2. Although the two partial interaction test did not show a significant task by manipulation direction interaction effect (F < 1.0), the results of separate tests in each task condition were consistent with H2. The adaptation-level effect was significant for the purchase intention rating (M_{low} = 3.90 vs. M_{high} = 4.33; *F* = 3.32, *p* = .069), but not significant for the attractiveness rating (M_{low} = 3.85 vs. M_{high} = 4.14; *F* = 1.59, *p* = .208).

Discussion

Replicating the results of study 1, we again find evidence supportive of H1 and H2 that the influence of alternative reference points is affected by the decision task. As shown in Table 3, range endpoints have a stronger influence when the task is to judge the attractiveness of a given price as opposed to when the task is one of rating purchase intentions (H1), while rank and mean effects are more influential for purchase intention ratings than price attractiveness ratings (H2).

These hypothesis-consistent results recognized, both studies conducted thus far have in common that consumers focused exclusively on price information. Information relating to other product attributes was absent from both studies, a factor not representative of most market situations. Study 3 was conducted to address this void by testing the study hypotheses employing a methodology that recognizes the multiattribute nature of evaluations that consumers make every day in the marketplace.

STUDY 3

Method

A total of 185 undergraduate students participated in 2 (task: attractiveness rating vs. purchase intention rating) x 3 (price distribution: negative-skew low-mean, positive-skew low-mean, and negative-skew high-mean) between-subjects study. Each participant viewed descriptions for a set of seven 32-inch LCD TV options presented in ascending order in terms of price. Each option was defined on three attributes: brand name, picture resolution, and price. Brand name (JVC, Toshiba, and Panasonic) and picture resolution (1024 x 768, 1280 x 1024, and 1600 x 1200) had three levels, and price attribute had seven levels.[6] The price of the seven TV alternatives followed one of the three price distributions (see Appendix D): negative-skew low-mean, positive-skew low-mean, and negative-skew high-mean. As in Study 1, the three price distributions varied in the range ($1,100-$2,000 vs. $1,400-$2,300), skewness, and mean ($1,700 vs. $2,000).

The following procedure was employed to generate sets of product profiles (e.g., Lazari and Anderson 1994). First, seven TV profiles were created by combining the brand and picture

[6] The results of a pretest (n = 30) with independently recruited participants indicated that Panasonic was the most preferred and JVC was the least preferred brands.

resolution and then excluding two profiles with the best and worst combinations (i.e., Panasonic – 1600 x 1200 pixels and JVC – 1024 x 768 pixels). Then, these seven brand-resolution pairs were randomly assigned to the seven prices. Across the three price distributions, the prices of the same rank position were assigned to the same brand-resolution attribute levels. There were three different versions of random combinations.

In the experimental setting, each participant received a booklet containing instructions, product profiles, and evaluation questions. The cover page provided a brief description about the product and its attributes. The following page showed profiles of seven TV options with an evaluation scale for each option. These seven alternatives were presented in ascending order based on price. The evaluation task differed by condition. In the attractiveness rating condition, participants rated the attractiveness of each TV brand on a nine-point scale (1 = unattractive, 9 = attractive). In the purchase intention rating group, participants rated their willingness to buy each option on a nine-point scale (1 = unlikely to purchase, 9 = likely to purchase).

Results and Discussion

Using the following procedure, we assessed the relative impact of each of price models (i.e., Range, Frequency, Adaptation-Level, and Range-Frequency) on the estimated price coefficients (i.e., the impact of the price on ratings). First, price coefficients (i.e., part-worths) were obtained by running a regression analysis using evaluation rating as a dependent variable and the following variables as predictors: two brand name dummies, two resolution dummies, and six price dummies. These regressions were based on data pooled across respondents and run separately for each price distribution and task. Thus, the analyses yielded a total of 21 price coefficients for each task condition (i.e., six price dummy coefficients and a baseline (i.e., zero) coefficient for each of the three price distributions).

Next, as in study 1, the fits of the Range, Frequency, Adaptation-Level, and Range-Frequency models were estimated using nonlinear regressions with the 21 price coefficients obtained in the above analysis as a dependent variable and one of the following scores of the three distributions as an independent variable: the range index, the frequency index, the AL index, and actual price (see Appendix D).[7] As presented in the last two columns of Table 2, the fit of the Range model was higher for the attractiveness rating (R^2 = .822) than for the rating of purchase intention (R^2 = .659). On the contrary, the Frequency model and the Adaptation-Level model were better fitted for purchase intentions (R^2 = .888 and .879, respectively) than for price attractiveness ratings (R^2 = .533 and .560, respectively). The range weight (w) of the Range-Frequency model was also substantially greater for the attractiveness rating (w = .99) than for the purchase intention rating (w = .06).[8]

Further tests were conducted to examine the statistical significance of the observed differences. More central to our hypotheses, comparisons of the beta coefficients of alternative models in the two task conditions show that the fit index of the range model was greater for the attractiveness rating (beta = -.91) than for the purchase intention rating (beta = -.81), although the difference did not reach significance level (z = 1.30, *p* = .194). However, both frequency (beta = -.94 vs. -.73; z = 2.62, *p* = .009) and the adaptation-level coefficients (beta = -.94 vs. -.75; z = 2.48, *p* = .013) were statistically higher for the purchase intention rating than for the attractiveness rating. Thus, results show support for H2 and directional support for H1. We also examined the relative reliance on the range and rank feature in each rating task by testing the significance of the interaction between range weight and task (w_i) in the Range-Frequency model.

[7] For the estimation of the range model, we used the range index that is defined as: RA_{ik} = (target price – the lowest price in the distribution) / (the highest price in the distribution – the lowest price in the distribution). Using the method we used in Study 1 (Eq. A2) also yielded similar results.
[8] Range weight coefficients were constrained to range between 0 and 1.

Thus, the model input included the range (RA_{ik}) and frequency (F_{ik}) indices: $R_{ik} = a + b[(w + w_i*task)RA_{ik} + (1 - w - w_i*task)F_{ik}]$. The test showed that the interaction term was significant ($w_i = -1.08$, SE = .25, $p < .001$), indicating that the relative magnitude of the range effect is higher for the attractiveness rating and that the frequency effect is higher for the purchase intention rating.

GENERAL DISCUSSION

One issue that continues to be of focal attention in decision research is that of information processing differences across judgment and choice. Although much consumer research assesses behavioral intention in lieu of choice, little research attention has been focused on this “choice-like” response. The goal of the present research was to investigate differences between ratings of attractiveness and intentions in terms of their differential sensitivity to alternative reference points. Results across three studies provide consistent evidence that range effects are more diagnostic, hence relied upon more, for attractiveness than intention ratings. Alternatively, frequency (rank) and adaptation-level (mean) effects are more diagnostic, hence relied upon more, for intention than attractiveness ratings. These results underscore the notion that diagnosticity is goal-specific in that one reference point may be more diagnostic for one goal than another (Lynch, Marmorstein, and Weigold 1988), even in cases where goals are thought to be very closely related, in our studies, ratings of attractiveness and intentions.

We view our results as having important implications for researchers assessing attractiveness and/or intention responses, but who may consider the two responses to rely on common reference points, or perhaps more accurately, do not question that they do not. Perhaps one area where this is most common is in models where attractiveness (or similar affect-related judgments) is hypothesized to totally mediate effects on intentions and choice. To consider one such area, much research on consumer satisfaction has traditionally modeled this construct as the sole and total mediator of antecedent effects on customer purchase intentions (e.g., Anderson and Sullivan 1993). This practice extends to much recent research on the customer "value chain" models where all antecedent managerial and employee-driven effects on customer intentions and choice are totally mediated by customer satisfaction (Schneider et al. 2005). As satisfaction is a judgment and less "choice-like" than the two downstream variables of intentions and choice, our results suggest that different reference points may be differentially used in formulations of satisfaction versus intention and choice responses, hence a totally mediated model may not be justified.

As a second example of where attractiveness and intentions may be implicitly assumed to have common reference points, consider the domain of price perception in which our study was conducted. Different theories (e.g., Range, Range-Frequency, Adaptation Level Theory) are often applied to price perception in an attempt to understand and explain "consumer response to price." To our knowledge, no research has further delineated among alternative "consumer responses to price" based on the possibility of theoretical differences among these responses with respect to responsiveness to alternative reference prices. A contribution of the current research is that "consumer response to price" is too broad of a concept to consider when evaluating the relative predictive ability of alternative reference points. Rather, researchers need to differentiate between attractiveness and intention responses. As a case in point, Niedrich, Sharma, and Wedell (2001) found that range effects were more influential on price attractiveness judgments than were rank or mean effects, as did we. Given that many researchers might consider price attractiveness and purchase intentions interchangeable and reflective of "price responses" more generally, there may be a tendency to generalize the findings of Niedrich, Sharma, and Wedell (2001) to "price responses" more generally. Our results suggest that this might not be appropriate: Had Niedrich, Sharma, and Wedell (2001) assessed purchase intentions rather than price attractiveness, their effects may have been quite different.

These insights notwithstanding, we do believe that a boundary condition for our study relates to the fact that it was conducted in a stimulus-based, rather than memory-based, decision environment. This was necessary given the need to employ a between-subjects design where respondents in both task conditions had the same exposure to the stimuli for rating. That recognized, it is interesting to speculate what might occur in the rather common context where consumers perform both tasks, attractiveness judgments followed by intentions, separated in time and where intentions are memory-based. For example, in sitting in an office and deciding which restaurant to choose for lunch, what information gets retrieved from memory in forming an intention? Lynch, Marmorstein, and Weigold (1988) contend that in such a memory-based choice context, consumers may attempt to perform the choice task by recalling previously made judgments, or alternatively, they may attempt to retrieve attribute values across brands, or some combination of the two.

To the extent that our stimulus-based findings would generalize to a memory-based context, consistent with Norm Theory (Kahneman and Miller 1986), we would expect consumers to rely more heavily on the latter process (i.e., retrieve attribute values necessary for estimating rank or mean for use as reference values for intentions). However, such a process would seemingly be very taxing for consumers. For example, according to Norm Theory, such a process would entail consumers establishing a norm for each salient attribute (e.g., price, convenience, nutrition) across the context set of alternatives, and then judging the target restaurant (e.g., Burger King) relative to the norm for each of these salient attributes. What is the probability consumers would engage in such an extensive process? It probably is not high. Lynch, Marmorstein, and Weigold (1988) characterize consumers as cognitive misers and contend that they will be more likely to rely on attributes, and not their prior judgments, in making choice decisions only in cases where recall of diagnostic attributes is easy. In cases where recall of attributes is more difficult, consumers will be more likely to rely on previously formed judgments, and not their recall of attributes, as inputs to choice. Because of the retrieval task that would be involved in estimating mean and rank even for our simple example here would be taxing, we restrict the interpretation of our findings to a stimulus-based context where attribute recall is held constant.

One additional point worthy of consideration is the lack of attention of judgment-intention research relative to that of judgment-choice research, given the frequency with which intention is the ultimate outcome variable assessed. Moreover, in many "choice" studies, respondents are presented with a hypothetical choice. It would seem that on grounds of commitment being the primary underlying construct differentiating judgments and choice (Einhorn and Hogarth 1981), one could argue that unless a choice reflects a selection of an option with consequences equal to those incurred in a realistic setting (i.e., respondents actually realize the consequences of their choice), any "choice" measure actually reflects an intention of what they would do if faced with such a real choice in a real setting. In this sense, it may be more accurate to suggest that in many studies, choice measures are more "intention-like" than intention measures are "choice-like." In any event, we hope our research serves as an impetus for others to consider additional theoretical differences underlying judgments and intentions.

REFERENCES

Adaval, Rashmi and Kent B. Monroe (2002), “Automatic Construction and Use of Contextual Information for Product and Price Evaluations,” *Journal of Consumer Research*, 28 (March), 572-588.

Alexander, David L., John G. Lynch, Jr., and Qing Wang (2008), “As Time Goes By: Do Cold Feet Follow Warm Intentions for Really New Versus Incrementally New Products?” *Journal of Marketing Research*, 45 (June) 307-319.

Anderson, Eugene W. and Mary W. Sullivan (1993), “The Antecedents and Consequences of Customer Satisfaction for Firms,” *Marketing Science*, 12 (Spring), 125-143.

Baker, Julie, A. Parasuraman, Dhruv Grewal, and Glenn B. Voss (2002), “The Influence of Multiple Store Environment Cues on Perceived Merchandise Value and Patronage Intentions,” *Journal of Marketing*, 66 (April), 120-141.

Briesch, Richard A., Lakshman Krishnamurthi, Tridib Mazumdar, and S. P. Raj (1997), “A Comparative Analysis of Reference Price Models,” *Journal of Consumer Research*, 24 (September), 202-214.

Clancy, Kevin J., Peter C. Krieg, and Marianne McGarry Wolf (2006), *Market New Products Successfully*, Lanham, MD: Lexington Books.

Danziger, Shai and Ruthie Segev (2006), “The Effects of Informative and Non-Informative Price Patterns on Consumer Price Judgments,” *Psychology & Marketing*, 23 (June), 535-553.

Dodds, William B., Kent B. Monroe, and Dhruv Grewal (1991), “Effects of Price, Brand, and Store Information on Buyers’ Product Evaluations,” *Journal of Marketing Research*, 28 (August), 307-319.

Einhorn, Hillel J. and Robin M. Hogarth (1981), “Behavioral Decision Theory: Processes of

Judgment and Choice, *Annual Review of Psychology*, 32, 53-88.

Feldman, Jack M. and John G. Lynch, Jr. (1988), "Self-Generated Validity and Other Effects of Measurement on Belief, Attitude, Intention, and Behavior," *Journal of Applied Psychology*, 73 (August), 421-435.

Grewal, Dhruv, Howard Marmorstein, and Arun Sharma (1996), "Communicating Price Information through Semantic Cues: The Moderating Effects of Situation and Discount Size," *Journal of Consumer Research*, 23 (September), 148-155.

———, Kent B. Monroe, and R. Krishnan (1998), "The Effects of Price-Comparison Advertising on Buyers' Perceptions of Acquisition Value, Transaction Value, and Behavioral Intentions," *Journal of Marketing*, 62 (April), 46-59.

Hardie, Bruce G. S., Eric J. Johnson, and Peter S. Fader (1993), "Modeling Loss Aversion and Reference Dependence Effects on Brand Choice," *Marketing Science*, 12 (Fall), 378-394.

Helson, Harry (1964), *Adaptation-Level Theory*, New York: Harper and Row.

Herr, Paul M. (1986), "Consequences of Priming: Judgment and Behavior," *Journal of Personality and Social Psychology*, 51 (December), 1106-1115.

Howlett, Elizabeth A., Scot Burton, Kenneth Bates, and Kyle Huggins (forthcoming, 2009), "Coming to a Restaurant Near You? Potential Consumer Responses to Nutrition Information Disclosure on Menus," *Journal of Consumer Research*, 36 (October).

Hutchinson, J. Wesley (1983), "Expertise and the Structure of Free Recall," in *Advances in Consumer Research*, Vol. 10, Richard P. Bagozzi and Alice M. Tybout, eds. Ann Arbor, MI: Association for Consumer Research, 585-589.

Inman, J. Jeffrey, Anil C. Peter, and Priya Raghubir (1997), "Framing of the Deal: The Role of Restrictions in Accentuating Deal Value," *Journal of Consumer Research*, 24 (June), 68-

79.

Janiszewski, Chris and Donald R. Lichtenstein (1999), “A Range Theory Account of Price Perception,” *Journal of Consumer Research*, 25 (March), 353-368.

Kahneman, Daniel and Dale T. Miller (1986), “Norm Theory: Comparing Reality to Its Alternatives,” *Psychological Review*, 93 (April), 136-153.

Lalwani, Ashok K. and Kent B. Monroe (2005), “A Reexamination of Frequency-Depth Effects in Consumer Price Judgments,” *Journal of Consumer Research*, 32 (December), 480-485.

Lazari, Andreas G. and Donald A. Anderson (1994), “Designs of Discrete Choice Set Experiments for Estimating both Attribute and Availability Cross Effects,” *Journal of Marketing Research*, 31 (August), 375-383.

Lichtenstein, Donald R. and William O. Bearden (1989), “Contextual Influences on Perceptions of Merchant-Supplied Reference Prices,” *Journal of Consumer Research*, 16 (June), 55-66.

Lim, Rodney G. (1995), “A Range-Frequency Explanation of Shifting Reference Points in Risky Decision Making,” *Organizational Behavior and Human Decision Processes*, 63 (July), 6-20.

Lynch, John G., Jr., Dipankar Chakravarti, and Anusree Mitra (1991), “Contrast Effects in Consumer Judgments: Changes in Mental Representations or in the Anchoring of Rating Scales,” *Journal of Consumer Research*, 18 (December), 284-297.

———, Howard Marmorstein, and Michael F. Weigold (1988), “Choices from Sets Including Remembered Brands: Use of Recalled Attributes and Prior Overall Evaluations,” *Journal of Consumer Research*, 15 (September), 169-184.

Mayhew, Glenn E. and Russell S. Winer (1992), “An Empirical Analysis of Internal and External

Reference Prices Using Scanner Data," *Journal of Consumer Research* 19 (June), 62-70.

Menon, Geeta, Lauren G. Block, and Suresh Ramanathan (2002), "We're at as Much Risk as We Are Led to Believe: Effects of Message Cues on Judgments of Health Risk," *Journal of Consumer Research*, 28 (March), 533-549.

Miller, Dale T. and Deborah A. Prentice (1996), "The Construction of Social Norms and Standards," in *Social Psychology: Handbook of Basic Principles,* E. Tory Higgins and Arie W. Kruglanski, eds. New York: Guilford, 799-829.

Morwitz, Vicki G. (2001), "Methods for Forecasting from Intentions Data," in *Principles of Forecasting: A Handbook for Researchers and Practitioners*, J. Scott Armstrong, ed. Boston: Kluwer Academic Publishers, 33-56.

——— and David Schmittlein (1992), "Using Segmentation to Improve Sales Forecasts Based on Purchase Intent: Which "Intenders" Actually Buy?" *Journal of Marketing Research*, 29 (November), 391-405.

Mussweiler, Thomas (2003), "Comparison Processes in Social Judgment: Mechanisms and Consequences," *Psychological Review*, 110 (July), 472-489.

Niedrich, Ronald W., Subhash Sharma, and Douglas H. Wedell (2001), "Reference Price and Price Perceptions: A Comparison of Alternative Models," *Journal of Consumer Research*, 28 (December), 339-354.

Parducci, Allen (1965), "Category Judgment: A Range-Frequency Model," *Psychological Review*, 72 (November), 407-418.

Payne, John W. (1982), "Contingent Decision Behavior," *Psychological Bulletin*, 92 (September), 382-402.

Petroshius, Susan M. and Kent B. Monroe (1987), "Effect of Product-Line Pricing

Characteristics on Product Evaluations," *Journal of Consumer Research,* 13 (March), 511-519.

Schneider, Benjamin, Mark G. Ehrhart, David M. Mayer, Jessica L. Saltz, and Kathryn Niles-Jolly (2005) "Understanding Organization-Customer Links in Service Settings," *Academy of Management Journal*, 48 (December) 1017-1032.

Schwarz, Norbert and Herbert Bless (1992), "Constructing Reality and Its Alternatives: An Inclusion/Exclusion Model of Assimilation and Contrast Effects in Social Judgment," in *The Construction of Social Judgments,* Leonard L. Martin and Abraham Tesser, eds. Hillsdale, NJ: Erlbaum, 217-245.

Sherif, Muzafer and Carl I. Hovland (1961), *Social Judgment*, New Haven: Yale University Press.

Smith, Richard H., Ed Diener, and Douglas H. Wedell (1989), "Intrapersonal and Social Comparison Determinants of Happiness: A Range-Frequency Analysis," *Journal of Personality and Social Psychology*, 56 (March), 317-325.

Urbany, Joel E., William O. Bearden, and Dan C. Weilbaker (1988), "The Effect of Plausible and Exaggerated Reference Prices on Consumer Perceptions and Price Search," *Journal of Consumer Research*, 15 (June), 95-110.

Volkmann, John (1951), "Scales of Judgment and Their Implications for Social Psychology," in *Social Psychology at the Crossroads*, John H. Rohrer and Muzafer Sherif, eds. New York: Harper, 273-296.

Winer, Russell S. (1986), "A Reference Price Model of Brand Choice for Frequently Purchased Products," *Journal of Consumer Research*, 13 (September), 250-256.

TABLE 1

STUDY 1: AVERAGE RATINGS OF THE TARGET PRICES

	Attractiveness Rating[a]			Purchase intention Rating[b]		
Target price	NSLM (n = 36)	PSLM (n = 37)	NSHM (n = 34)	NSLM (n = 39)	PSLM (n = 28)	NSHM (n = 34)
$144	5.94	6.51	6.59	6.10	6.25	7.16
$156	5.36	5.78	6.24	5.59	5.18	6.91
$168	4.81	5.20	6.00	5.23	4.43	6.56
$180	4.03	4.57	5.71	4.21	3.82	6.15
$192	3.44	4.19	5.41	3.56	3.93	5.68

NOTES: NSLM = negative-skew low-mean; PSLM = positive-skewed low-mean; NSHM = negative-skew high-mean.

[a] Attractiveness rating scale ranged between 1 (unattractive) and 9 (attractive).

[b] Purchase intention rating scale ranged between 1 (do not want to purchase) and 9 (want to purchase)

TABLE 2

STUDY 1 AND STUDY 3: MODEL FIT RESULTS

	Study 1		Study 3	
	Attractiveness rating	Purchase intention rating	Attractiveness rating	Purchase intention rating
Range model				
R^2	.850	.544	.822	.659
Frequency model				
R^2	.760	.943	.533	.888
Adaptation-Level model				
R^2	.785	.924	.560	.879
Range-Frequency model				
R^2	.981	.952	.823	.889
w	.76	.54	.99	.06

TABLE 3

STUDY 2: EVALUATION OF THE TARGET BY CONDITION

	Range Effect			Frequency Effect			Adaptation-Level Effect		
	Low	High	Diff.[a]	Low	High	Diff.[a]	Low	High	Diff.[a]
Attractiveness rating	3.83	4.48	.65**	4.03	4.11	.08	3.85	4.14	.29
	(65)	(63)		(62)	(62)		(61)	(63)	
Purchase intention rating	3.81	3.81	.00	3.71	4.64	.93**	3.90	4.33	.43*
	(62)	(64)		(63)	(61)		(57)	(61)	

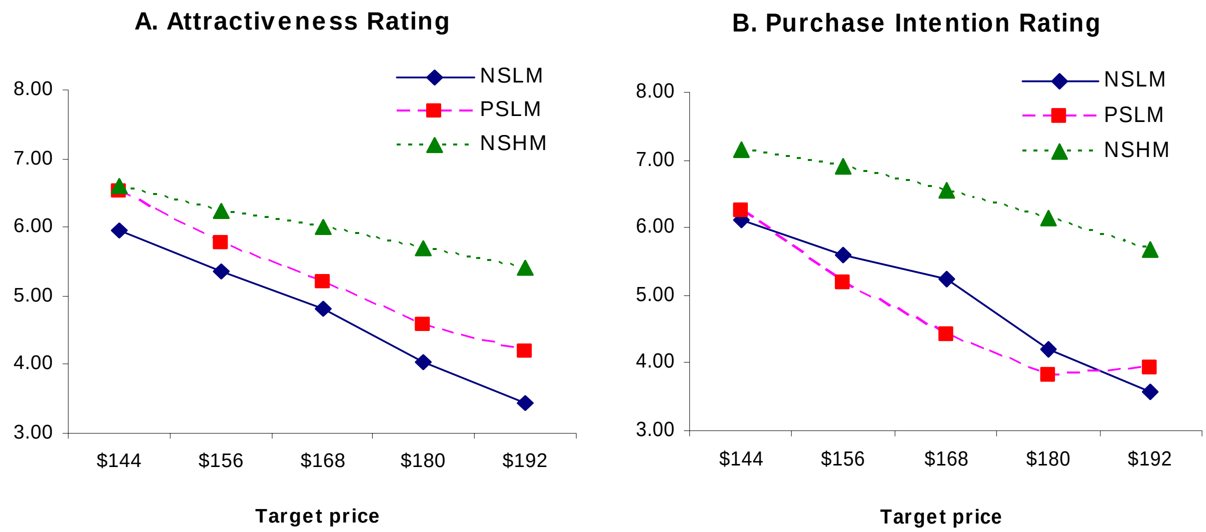
Note. Sample sizes are in the parentheses.

[a] Diff. presents the difference in ratings between high and low distribution conditions.

* $p < .10$, ** $p < .05$.

FIGURE 1

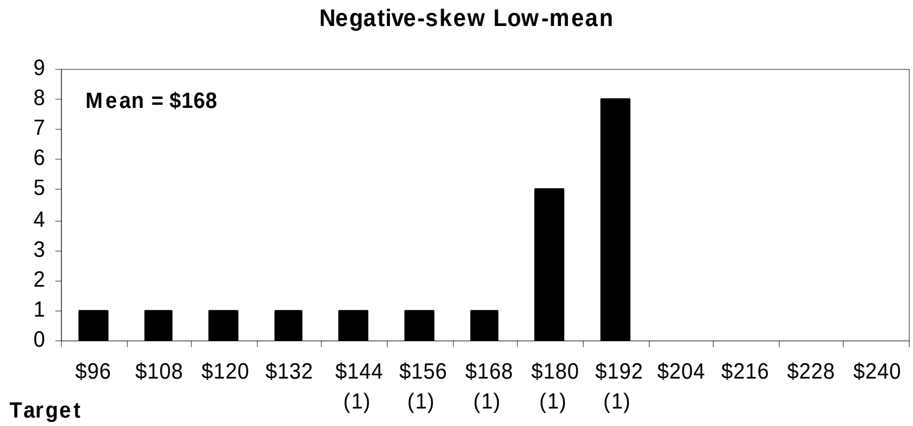
STUDY 1: AVERAGE RATINGS OF THE TARGET PRICES

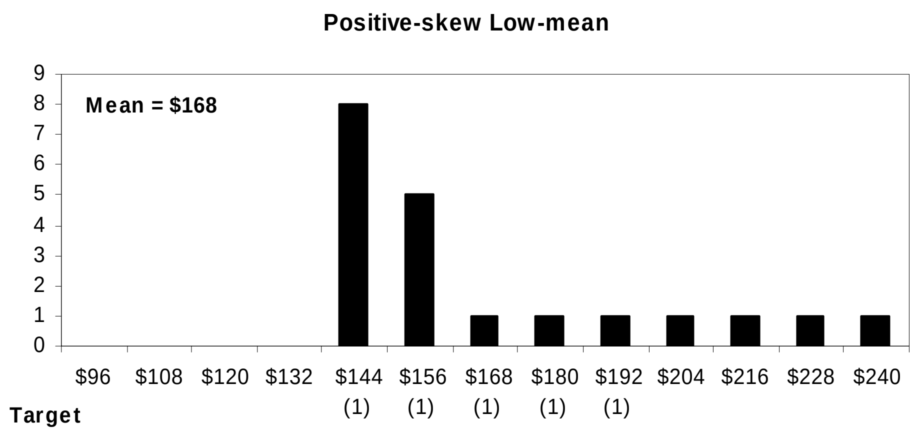


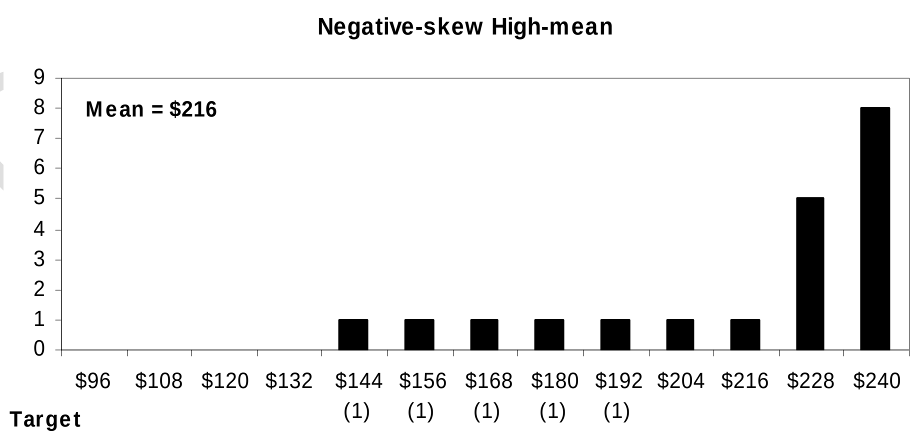
NOTES: NSLM = negative-skew low-mean; PSLM = positive-skewed low-mean; NSHM = negative-skew high-mean.

APPENDIX A

STUDY 1: FREQUENCY DISTRIBUTIONS OF PRICES BY DISTIBUTION







NOTES: The bars represent frequencies of the context prices in each distribution. The numbers in the parentheses are frequencies of the target price.

APPENDIX B

FITTING OF THE RANGE, THE FREQUENCY, THE ADAPTATION-LEVEL, AND THE RANGE-FREQUENCY MODELS

The fittings of the range, frequency, adaptation-level, and the range-frequency models were tested with the model indices and model specifications used in previous literature (e.g., Niedrich, Sharma, and Wedell 2001). According to the adaptation-level model, the rating of a target (R_{ik}) in distribution k is determined by its deviation from the distribution mean.

$$R_{ik} = a + b\,(\$_{ik} - \$_{al,k}), \qquad \text{(A1)}$$

where a and b denotes the parameters for the constant and the adaptation-level effect, respectively. The target price ($\$_{ik}$) and the average of the context prices given the distribution k ($\$_{al,k}$) are used as input.

The range model assumes that the evaluation of the target price ($\$_{ik}$) is relative to the lowest price ($\$_{min,k}$) and the highest price ($\$_{max,k}$) of the distribution. Because rating scales range between 1 (least favorable) and 9 (most favorable), the range score (ranging between 0 and 1) is mapped onto the rating scales as having the scale range of 8 and the highest rating of 9 (Study 1). The range model includes parameters for the extreme prices of the distributions included in the analyses.

$$R_{ik} = 9\text{-}8[(\$_{ik} - \$_{min,k})/(\$_{max,k} - \$_{min,k})]. \qquad \text{(A2)}$$

The frequency model used the frequency index that is defined as the rank of the target brand ($Rank_{ik}$) in inexpensiveness in the distribution including a total of N number of brands: $F_{ik} = (Rank_{ik} - 1)/(N_k-1)$. The frequency model is tested with two parameters (a and b) that reflect the constant and the frequency effect.

$$R_{ik} = a + b\,[(Rank_{ik} - 1)/(N_k-1)]. \tag{A3}$$

The range-frequency model assumes that the rating is affected by both the range effect and the frequency effect. The relative weight of the range effect over the frequency effect is captured by the range weight parameter, w.

$$R_{ik} = 9\text{-}8[w(\$_{i,k} - \$_{min,k})/(\$_{max,k} - \$_{min,k}) + (1 - w)F_{ik}]. \tag{A4}$$

APPENDIX C

STUDY 2: DISTRIBUTION OF THE PRICES

	Range		Rank (Frequency)		Mean (Adaptation-Level)	
	Low	High	Low	High	Low	High
Brand A	8.00	10.00	10.00	10.00	10.00	10.00
Brand B	9.00	10.25	11.00	10.50	10.25	11.50
Brand C	11.50	10.50	11.50	11.00	10.50	11.75
Brand D	12.00	10.75	11.75	12.00	10.75	12.00
Brand E	12.25	11.25	12.00	12.75	11.00	12.25
Brand F	13.75	12.75	12.25	13.00	12.75	14.00
Brand G	14.25	13.00	13.00	13.25	13.00	14.25
Brand H	14.50	13.50	14.00	13.50	13.25	14.50
Brand I	14.75	16.00	14.50	14.00	13.50	14.75
Brand J	15.00	17.00	15.00	15.00	15.00	15.00
Range						
Lowest	8.00	10.00	10.00	10.00	10.00	10.00
Highest	15.00	17.00	15.00	15.00	15.00	15.00
Rank of the target	6	6	7	5	6	6
Mean	12.50	12.50	12.50	12.50	12.00	13.00

APPENDIX D

STUDY 3: PRICE DISTRIBUTION AND THE RANGE, FREQUENCY, AND ADAPTATION-LEVEL (AL) INDICES

Price Distribution	Brand 1	Brand 2	Brand 3	Brand 4	Brand 5	Brand 6	Brand 7
Negative-skew							
Low-mean							
Price	$1,100	$1,450	$1,700	$1,800	$1,900	$1,950	$2,000
Range index	0.00	0.39	0.67	0.78	0.89	0.94	1.00
Frequency index	0.00	0.17	0.33	0.50	0.67	0.83	1.00
AL index	-600	-250	0	100	200	250	300
Positive-skew							
Low-mean							
Price	$1,400	$1,450	$1,500	$1,600	$1,700	$1,950	$2,300
Range index	0.00	0.06	0.11	0.22	0.33	0.61	1.00
Frequency index	0.00	0.17	0.33	0.50	0.67	0.83	1.00
AL index	-300	-250	-200	-100	0	250	600
Negative-skew							
High-mean							
Price	$1,400	$1,750	$2,000	$2,100	$2,200	$2,250	$2,300
Range index	0.00	0.39	0.67	0.78	0.89	0.94	1.00
Frequency index	0.00	0.17	0.33	0.50	0.67	0.83	1.00
AL index	-600	-250	0	100	200	250	300

The Effect of Reference Point Diagnosticity on Attractiveness and Intentions Ratings

Kwanho Suk, Song-Oh Yoon, Donald R. Lichtenstein, and Sie Yeoun Song

WEB APPENDIX

EXTENSIONS TO NON-PRICE ATTRIBUTES

Method

The stimuli and procedure for this study were similar to those used for study 3 except that two product categories were employed and they were defined on non-price attributes. A total of 115 undergraduate students participated in a 2 (task: attractiveness rating vs. purchase intention rating) x 3 (focal attribute distribution: negative-skew low-mean vs. positive-skew low-mean vs. negative-skew high-mean) x 2 (product category: cars and movie download website) mixed-design study. Task and attribute distribution were between-subjects and product was within-subjects.

Cars were defined on gas mileage, acceleration, and warranty. Gas mileage, which was the focal attribute, had seven levels, and both acceleration (10.9, 12.0, 13.4 seconds) and warranty (65k, 70k, 80k miles) had three levels. Movie download websites were described on the number of movie selections (focal attributes with 7 levels), download speed (3500, 3800, 4200 kb/s), and movie information quality (3.5, 4.0, 4.6 points). These focal attribute levels varied across the three distribution conditions (see Appendix). For each target category, seven profiles were generated using the same procedure as in study 3.

Participants received a booklet containing instructions, profiles of cars, profiles of movie websites, and evaluation questions. The presentation order of cars and movie websites was randomized. The seven profiles of a category (i.e., car or movie website) were presented in ascending order in terms of the focal attribute values. Participants in the attractiveness rating condition rated the attractiveness of each alternative on a nine-point scale (1 = unattractive, 9 = attractive). In the purchase intention rating group, the intention to purchase each option was expressed on a nine-point scale (1 = unlikely to purchase, 9 = likely to purchase).

Results and Discussion

Like study 3, the Range, Frequency, Adaptation-Level, and Range-Frequency models were fitted. Table W1 presents fits statistics for both cars and movie websites. For car evaluations, the range effect is stronger for attractiveness rating than purchase intention rating (R^2 = .930 vs. 866). Unexpectedly, the fits of the both Frequency and Adaptation-Level were slightly higher for attractiveness rating than purchase intention ratings (Frequency model R^2 = .897 vs. .891; Adaptation-Level model R^2 = .897 vs. .851). For movie websites, as expected the fit of Range model was higher for attractiveness rating (R^2 = .947 vs. .861) and Frequency and Adaptation-Level model fits were higher for purchase intention rating (Frequency model R^2 = .857 vs. .779; Adaptation-Level model R^2 = .860 vs. .805).

Statistical tests between the two tasks were also conducted by comparing standardized betas (Table W2). For cars, there were no significant differences between the attractiveness ratings and purchase intention ratings across any of models (*p*s > .14). For movie websites, the direction of effects was consistent with hypotheses, but in terms of statistical significance, results were only partially supportive. The difference was significant only for the Range model (beta = .96 vs. .87; $p < .05$) but not for other models (*p*s > .43).

We also tested whether the range weight (w) is stronger for the attractiveness rating task by adding the interaction between range weight and task (w_i). For cars, the interaction term did not differ from zero ($w_i = -.30$, SE = .25; t = 1.21, $p > .23$), providing no support for our hypotheses. However, evaluations for the movie websites showed that the interaction was significant ($w_i = -.60$, SE = .19; t = 2.99, $p < .01$), showing the range (frequency) has a stronger influence for the attractiveness (purchase intention) evaluation, partially supporting H1 and H2.

We also compared across different models including the influence of raw attribute levels in the same decision task. First, the Range model showed the best fit for both cars and movie websites in attractiveness rating task. For car category, the coefficient of the Range model (beta = .95) was not different from that of the Adaptation-Level (beta = .94). For movie websites, the fit of the Range model (beta = .96) was significantly higher than Frequency (beta = .88, $p < .10$) and directionally higher than the Adaptation-Level model (beta = .90, $p = .12$), Next, in the purchase intention rating, the Frequency model and the Adaptation-Level model showed improved fit over the Range model for both product categories. For cars, the fits of the Frequency (beta = .94) and the Adaptation-Level (beta = .94) were better than that of the Range model (beta = .89), but the differences were not significant statistically ($p > .22$). A similar pattern was observed for movie websites. While the Frequency (beta = .93) and the Adaptation-Level (beta = .93) showed better fit than that of the Range (beta = .87),the differences did not reach the statistical significance (ps > .32). Overall, the consistency of results with respect to directionality of effects, coupled with the finding that most of the hypothesized differences were not statistically significant, is consistent with prior theorizing regarding the applicability of the theories to non-price attributes, yet at the same time, the non-price attributes lacking the general salience of price.

TABLE W1

MODEL FIT RESULTS

	Cars		Websites	
	Attractive judgment rating	Purchase intention rating	Attractive judgment rating	Purchase intention rating
Range				
R^2	.930	.866	.947	.861
Frequency				
R^2	.897	.891	.779	.857
Adaptation-Level				
R^2	.897	.851	.805	.860
Range-Frequency				
R^2	.972	.937	.954	.910
w	.54	.42	.84	.63

APPENDIX

DISTIRIBUTION OF THE FOCAL ATTRIBUTE LEVELS

A. Gas Mileage (km/liter) of Cars

Price Distribution	Brand 1	Brand 2	Brand 3	Brand 4	Brand 5	Brand 6	Brand 7
NSLM							
Attribute level	11.5	13.5	15.0	16.0	17.0	17.5	18.0
Range index	0.00	0.31	0.54	0.69	0.85	0.92	1.00
Frequency index	0.00	0.17	0.33	0.50	0.67	0.83	1.00
AL index	-4.0	-2.0	-0.5	0.5	1.5	2.0	2.5
PSLM							
Attribute level	13.0	13.5	14.0	15.0	16.0	17.5	19.5
Range index	0.00	0.08	0.15	0.31	0.46	0.69	1.00
Frequency index	0.00	0.17	0.33	0.50	0.67	0.83	1.00
AL index	-2.5	-2.0	-1.5	-0.5	0.5	2.0	4.0
NSHM							
Attribute level	13.0	15.0	16.5	17.5	18.5	19.0	19.5
Range index	0.00	0.31	0.54	0.69	0.85	0.92	1.00
Frequency index	0.00	0.17	0.33	0.50	0.67	0.83	1.00
AL index	-2.5	-0.5	1.0	2.0	3.0	3.5	4.0

B. Number of Movie Titles of Online Movie Websites

Distribution	Brand 1	Brand 2	Brand 3	Brand 4	Brand 5	Brand 6	Brand 7
NSLM							
Attribute level	350	600	900	1,000	1,100	1,150	1,200
Range index	0.00	0.29	0.65	0.76	0.88	0.94	1.00
Frequency index	0.00	0.17	0.33	0.50	0.67	0.83	1.00
AL index	-550	-300	0	100	200	250	300
PSLM							
Attribute level	600	650	700	800	900	1,200	1,450
Range index	0.00	0.06	0.12	0.24	0.35	0.71	1.00
Frequency index	0.00	0.17	0.33	0.50	0.67	0.83	1.00
AL index	-300	-250	-200	-100	0	300	550
NSHM							
Attribute level	600	850	1,150	1,250	1,350	1,400	1,450
Range index	0.00	0.29	0.65	0.76	0.88	0.94	1.00
Frequency index	0.00	0.17	0.33	0.50	0.67	0.83	1.00
AL index	-550	-300	0	100	200	250	300

NOTE.—NSLM = negative-skew low-mean; PSLM = positive-skew low-mean; NSHM = negative-skew high-mean