Mapping Consumers' Context-Dependent Consumption Preferences:

A Multidimensional Unfolding Approach

ABSTRACT

Marketers have long been interested in understanding how, and the extent to which, consumer choices may be influenced by the context in which a product is consumed. This study develops a context-dependent multidimensional unfolding (CDMDU) model to map consumers' context-specific behaviors via ideal points in multi-attribute space along with brand locations in that space. Unobserved heterogeneity in behavior is also accounted for. The model is empirically illustrated using panel data from consumers in the U.S. beer market on brand choices made in different contexts. Consumers are found to be more heterogeneous across at-home and out-of-home consumption than they are across social versus nonsocial contexts; nevertheless, there is still considerable heterogeneity in the data across consumption contexts. The model is then used to derive a firm's optimal direction of brand repositioning given its competitive landscape in the various consumption contexts as revealed by the map. A key observation when repositioning a brand is that consumer preferences can be correlated across contexts. As a result, a movement toward the ideal point in one particular context does not necessarily improve the firm's market competitiveness in other consumption contexts and can therefore hurt its overall performance in the market.

Keywords: Context-dependent preference, multiple ideal points, factor structure approach, context-dependent multidimensional unfolding (CDMDU) model, context-dependent iso-preference contours

1. INTRODUCTION

Marketing researchers have long been interested in understanding the potential influence of situations, circumstances, or contexts on consumer behavior, as consumers may have different preference judgments for the same brand across different situations and usage occasions (Belk 1974, 1975; Bettman, Luce, and Payne 1998; Dubow 1992; Fennel 1978; Laurent 1978; Srivastava, Alpert, and Shocker 1984, etc.). Lewin (1951) suggested that behavior may vary with contexts[1] and the consumers' overall attitudes. The interaction between the person and the contextual influences on behavior has been discussed in the consumer psychology literature (Belk 1974, 1975; Puto 1987). For instance, different consumption situations may have primed different attitudes and activated different goals for different individuals (Bargh and Tota 1988; Tversky and Kahneman 1981). As such, a consumer can have different decision-making goals across contexts; these different goals can result in different choices in different contexts, thereby illustrating the context-dependent nature of preferences.

The consumer behavior literature has also found support for the notion that a consumer's utility function can vary across the contexts in which he/she makes a decision (Carlson and Bond 2006; DeSarbo, Atalay, Lebaron, and Blanchard 2008; Simon 1955, 1990; Tversky and Kahneman 1991). One specific manifestation of a consumer's context-specific utility function is the presence of multiple "ideal points" or "ideal products" (i.e., ideal combinations of features or benefits) corresponding to different consumption contexts. Since the utility of a consumer for a brand (in a given context) depends on how far the brand is from the consumer's ideal product, variation in the ideal product across contexts would lead to different utilities for that brand in those contexts. Firms ignoring such multiple situation-specific ideal points are likely to end up with an inaccurate

---

[1] Context can be defined as the composition of the observable, aggregate characteristics of the time and place of observation that are not related to personal and stimulus attributes (Belk 1974; Tversky and Simonson 1993). Belk (1975) represented five groups of situational characteristics as general situations—physical surroundings, social surroundings, temporal perspective, task definition, and antecedent states. A consumer may have different preferences in these different contexts.

understanding of choice behavior and an incomplete knowledge of the competitive structure across brands in the market (DeSarbo et al. 2008; Lee, Sudhir, and Steckel 2002). Researchers have therefore emphasized the need to account for consumers' situation-specific ideal products (DeSarbo and Carroll 1981, 1985; Holbrook 1984).

The purpose of this study is to combine information on the context in which consumption takes place with the actual choices that marketers observe consumers making in those contexts. Doing so, as we show, allows us not only to understand how behavior changes across contexts but also to leverage that knowledge for firms interested in changing the perceptions of their brands in the marketplace. Although there exists a large stream of psychology literature on the role of context in consumer choices, relatively few papers in economics or quantitative marketing have empirically studied this issue (Thomadsen et al. 2018). One key constraint faced by researchers is the lack of comprehensive context and choice data. Furthermore, there are technical challenges associated with trying to accommodate contextual information in a parsimonious fashion within existing discrete choice models. In this paper we propose an empirically estimable, context-dependent multiple-ideal-point framework that allows us to incorporate situational factors in studying choice behavior. A potential contribution of our research, we believe, is thereby to create a link between econometric modelers and psychologists interested in this important issue of mutual interest.

Among various ideal point models, the multidimensional unfolding (MDU hereafter) approach has been one of the most popular statistical techniques to infer consumers' context-invariant ideal points (Green 1975). The MDU model has contributed greatly to the spatial depiction of brands for purposes such as market segmentation (Johnson 1971), positioning (DeSarbo and Rao 1986; Wind 1982), competitive market structure (Day, Shocker, and Srivastava 1979; Shocker and Stewart 1983; Srivastava et al. 1984), and consumer preferences/perceptions (DeSarbo and Carroll 1985; Green and Carmone 1970; Green and Rao 1972), etc. However, traditional MDU spatial representations were limited by typically assuming a single invariant ideal product (Lee et al. 2002). As part of efforts to incorporate multiple ideal points into spatial representations of consumer preferences, researchers have attempted to develop multiple-ideal-point spatial methodologies and joint space maps such as the multiple ideal point model (MIPM) (Lee et al. 2002) and the three-way clusterwise MDU model (DeSarbo et al. 2008; DeSarbo, Atalay, and Blanchard 2009). DeSarbo et al. (2008) note that there still exists room for further model extensions (e.g., inclusion of context and dimension weights, segment heterogeneity, confidence bands around estimated ideal points and brand locations, model selection heuristics) and for applications in consumer research involving actual consumer panel data.

In this study, we aim to contribute to the above literature by proposing a context-dependent MDU (CDMDU) model, which can address some of the model improvements that DeSarbo et al. (2008) call for. To develop a more flexible model with low computational costs, we leverage the stream of literature on the factor structure approach, which imposes a factor structure on the covariance matrix of consumer brand preferences to parsimoniously account for consumer heterogeneity (e.g., Chintagunta 1994; Elrod 1988; Elrod and Keane 1995; Kim and Chintagunta 2012). By incorporating such a modeling approach into the MDU setting, we derive a joint space map in which brands and context-dependent ideal points are located within the same (unobserved) attribute space along with varying preferences across consumers. An attractive feature of our model is that when there is only one consumption context, it simplifies to a traditional factor structure random coefficients model (e.g., the factor-analytic probit model proposed by Elrod and Keane 1995).

Previous context-dependent MDU studies[2] were typically based on surveys in which each participant was asked to report his or her preferences for each brand in each usage occasion. In other words, the number of observations across usage occasions is the same. Unlike such well-designed consumer surveys, however, longitudinal consumption data from diaries (or scanner-panel data) do not provide a sufficient number of observations for certain contexts at the individual level. So another feature of our CDMDU model is that it can flexibly handle data sparseness for specific brand-context combinations by sharing information on intra- and inter-person choice variation across multiple contexts. Thus, our research can be considered an extension of the MDU studies along two dimensions: model specification and accommodating (possibly) sparse data.

A third direction we aim to build on the previous literature is to use the joint space map to help firms reposition their products. Different consumption contexts represent distinct markets with different competitive landscapes since consumers' brand preferences and desired product attributes change as the environment changes (Yang et al. 2002). As the brands and ideal points for the contexts reside in a common (joint) space, the closer a brand gets to the ideal point for one context, the further it may move away from the ideal point for a different context. Consequently, the firm's movement towards the ideal point in one particular context does not necessarily improve its market competitiveness in other consumption contexts. Thus, firms need to devise different marketing strategies targeting consumers in different contexts or markets while accounting for such tradeoffs inherent in brand repositioning.

Given a specific objective function a firm wants to optimize, we then demonstrate how the CDMDU model can be used to derive the firm's optimal direction of brand repositioning, taking into account the competitive landscapes for all the consumption contexts. For the empirical

---

[2] DeSarbo et al. (2008) empirically illustrated their model with a student survey that collected information on the perceived effectiveness judgments for 11 over-the-counter analgesics in five situations (four different maladies and one overall [i.e., known discrete contexts]).

illustration of the CDMDU and to study a brand's potential repositioning decision, we use individual-level panel data on U.S. beer consumption in which each panelist's brand choice (e.g., Bud) on any given consumption occasion is uniquely associated with his or her consumption context (e.g., watching TV at home). To further enrich the substantive implications of our model, we conduct a discrete choice experiment on Amazon Mechanical Turk (MTurk hereafter) to validate our CDMDU-based brand repositioning.

The remainder of this paper is as follows: Section 2 briefly reviews the relevant literature on the methods. We then explain how we formulate the proposed CDMDU model by building on the extant literature. In Section 4, we empirically illustrate the CDMDU model along with marketing strategies using empirical results. Section 5 discusses the substantive implications of the CDMDU model vis-à-vis simple alternative models and offers conclusions.

2. LITERATURE REVIEW

2.1 Multidimensional unfolding (MDU) model

In Table 1, we provide a brief review of model developments in the MDU literature.

<Table 1 about here>

The simple MDU model (SMDU hereafter) was developed by Coombs (1950, unidimensional unfolding) and later generalized to the multidimensional case by Bennett and Hays (1960, multidimensional unfolding). In this SMDU, the latent utility for a brand is a negative function of the squared Euclidean distance between a brand location and a consumer's ideal point in the derived K-dimensional joint space, with the dimensions representing unobserved attributes or benefits based on which consumers assess the various products. Consumers (denoted by *i*) are assumed to be heterogeneous in their ideal points, IP_i (in Figure 1-1, which shows a *2*-dimensional map [i.e., *K=2*]), in this space and the weights (w_i) they place on the dimensions. The weights for a consumer are assumed to be the same across the *K* dimensions. Ideal points represent targets for marketers who attempt to position their brands near target segment ideal points (DeSarbo and Rao 1986). In SMDU, all brand points at a fixed radius around the consumer's ideal points reflect equal preferences since all dimensions are equally weighted (see Table 1). The group of brands with equal preferences on two dimensions can then be represented on a concentric circle around the particular ideal point.

<Figures 1-1 and 1-2 about here>

Subsequently, the SMDU was extended to the weighted MDU (WMDU hereafter) model (Carroll 1972; Carroll and Arabie 1980; DeSarbo and Hoffman 1986, 1987; DeSarbo and Rao 1984; Wedel and DeSarbo 1996), where the squared Euclidean distance is weighted by both dimension- and individual-specific scale parameters (w_i^k) (see the third column of Table 2). That is, unlike the SMDU, the WMDU can take the shape of either a vertical ellipse (i.e., the horizontal dimension is highly weighted) or a horizontal ellipse (i.e., the vertical dimension is highly weighted), depending on the weight for a particular dimension (see Figure 1-2). However, while the scale parameters (w_i^k) in the WMDU model allow for individual weights that differ across latent dimensions, the WMDU models were typically faced with an excessive number of parameters (Davison 1976, 1988). Context-dependent preferences were not considered in these models.

<Table 2 about here>

Given that consumers may have different preference judgments for the same brand across different situations and usage occasions, researchers extended the MDU model to cases where subjects have preferences for various stimuli (i.e., products or brands) over different contexts (e.g., times, consumption scenarios) (see the second column of Table 2). Here, context-dependent weights on each dimension of unfolding are required to estimate the respective perceptual space relevant to each particular context. Accordingly, the three-way metric unfolding model (DeSarbo and Carroll 1981, 1985) accounts for context-specific weights. That model, however, cannot portray context-dependent ideal points in K-dimensional joint space.

<Figure 1-3 about here>

To accommodate context-dependent ideal points, as seen in Figure 1-3, DeSarbo et al. (2008, 2009) proposed the three-way clusterwise MDU model as an extended version of the three-way metric unfolding model (DeSarbo and Carroll 1981, 1985). The main advantage of this model is that segment-level ideal points in a joint space map vary by context, usage occasion, and other factors (see the first column of Table 2). Another advantage is the simultaneous derivation of segment-level ideal points and brand positions as well as the classification of consumers/households into these derived market segments. Taken together, these models make important contributions to the consumer behavior and MDU literatures. However, the ability to accommodate individual differences across multiple contexts in a flexible and econometrically parsimonious manner remains an important challenge for the literature (see the third column in Table 2). In this paper, we propose one such modeling approach to address this challenge.

We also note that previous MDU studies were typically based on surveys, with participants asked to report their evaluations for each brand. For example, DeSarbo et al. (2008) asked respondents to report their effectiveness ratings for analgesics in four different contexts but not their relative preferences across the various products.[3] Providing guidance to firms on substantive issues such as repositioning requires translating these ratings into outcomes of interest to firms (e.g., market shares). Such a transformation is not straightforward, however. An additional consideration is that one can, when conducting surveys, always ensure an equal number of observations across usage occasions by eliciting consumers' preferences in all usage situations.

[3] Similarly, DeSarbo and Carroll (1985) used a survey questionnaire on students' purchase intentions for 10 brands of over-the-counter (OTC) medicines for three common maladies: headache, fever, and muscle aches. DeSarbo et al. (2009) used a similar strategy to DeSarbo et al. (2008) to measure breakfast/snack food evaluations.

With actual consumption choice data, one only records the actual occasions on which consumption occurs. This can lead to a data sparseness issue (i.e., there may be very few observations for specific brand–context combinations). The structure of our proposed model can deal with the issue of data sparseness (see our subsequent discussion on Table A-1 in Web Appendix A[4]).

Lee et al.'s (2002) MIPM accounts for the possibility of multiple preference contexts by exploiting household-level product-switching matrices obtained from scanner-panel data. In their MIPM framework, one of a consumer's multiple ideal points is activated with some probability at any given purchase. Accordingly, the likelihood of selecting a brand differs conditional on which ideal point is activated. Since the actual consumer motivations at each purchase are not observed, the data represent a mixture of the purchase behaviors corresponding to the different ideal points. Thus, the MIPM collapses into a single ideal point model when the estimated ideal points are empirically indistinguishable. However, estimating the model parameters requires a large number of purchase observations at the household-level and can be computationally demanding. Also, the estimation procedure involves maximum-likelihood estimation for inferring ideal points and a separate clustering procedure for the ideal points. DeSarbo et al. (2008, 2009) therefore proposed a new three-way clusterwise MDU model to simultaneously estimate a joint space map of brands and segment-level ideal points that vary across contexts and usage occasions. While their proposed model retains the spirit of DeSarbo and Carroll (1981, 1985) in the sense that consumers in a given consumption context have the highest utility at the derived ideal points, it is more parsimonious than the previous MDU models, with a smaller number of parameters required for segment-level estimation.

2.2 Factor structure approach

[4] In the previous MDU literature, the focus is on specific brands in the data collection. Researchers control the set of brands that are analyzed and can ensure responses for these brands across all contexts considered. Since our data are real consumer choices, we cannot control the set of products considered or ensure consumption in all contexts. In our empirical application, we therefore focus on the 10 biggest beer brands and collect the other brands chosen into a single composite 11th brand.

Consumer researchers have long recognized the importance of accounting for heterogeneity in consumer preferences. Early models that empirically studied choice behaviors of consumers looked at their decisions as to whether or not to purchase a product; this binary decision naturally fit into assuming a binomial distribution for the choice decision (Wadsworth 1960). To allow for heterogeneity across consumers in their probabilities of purchasing, researchers allowed the binomial parameter to vary across consumers following a beta distribution (Chatfield and Goodhardt 1970). With the advent of discrete choice models and panel data (with multiple observations per household), logit models became the basis for understanding choice behavior. To account for observable heterogeneity across consumers, researchers used demographic variables and lagged choices of panelists (Guadagni and Little 1983).

Next, researchers were interested in representing the *unobservable* differences across consumers (Gönül and Srinivasan 1993; Kamakura and Russell 1989; etc.)—that is, differences in consumer preferences and responsiveness that cannot be represented via observable variables of the consumers. These researchers assumed that the households' unobserved preferences and responsiveness followed a specific distribution (e.g., a multivariate normal distribution) and then proceeded to estimate the moments of the distribution. One issue with the normal distribution is that as the number of brands becomes large, the number of parameters in the covariance matrix explodes. With J (say, 10) brands, we will have $\frac{J\times(J-1)}{2}$ $(= \frac{10\times 9}{2} = 45)$ covariance matrix parameters. Now with S (say 5) contexts, these will lead to $\frac{J\times(J-1)\times S}{2}$ $(= \frac{10\times 9\times 5}{2} = 225)$ parameters just for the covariance matrices. Furthermore, there are not enough observations in some contexts to be able to reliably estimate these parameters.

To alleviate the parameter explosion problem associated with the covariance matrix, Elrod and Keane (1995) imposed a factor structure on the matrix that basically projects it onto a lower dimensional unobserved "attribute" space (while it is not possible to assign unique labels to these attributes, one can always correlate them to product characteristics ex post facto, as we do). The low-dimensional K (say, two dimensions or "factors" instead of the (J–1) = nine-dimensional space for 10 brands [i.e., K<<J]) space is characterized by (i) brand locations along the unobserved underlying attributes (a (J–1) × K = 9 × 2 matrix when there are 10 brands and two dimensions) that represent consumers perceptions of where the brands are located relative to one another; and (ii) household-specific "importance weights" for these attributes (a K × 1 = 2 × 1 vector for the two dimensions) that represent how important consumers perceive these attributes to be in driving their preferences. Since the distribution of the importance weights (assumed to be, e.g., multivariate normal) are of a much lower dimension than the number of brands, the number of covariance parameters to be estimated for this distribution is lower. Figure 2-1 provides an illustrative example with three brands in two-dimensional space.

<Figure 2-1 about here>

The factor structure on the covariance matrix results in model parsimony.[5] Extending the situation to one in which data are available across several contexts, Kim and Chintagunta (2012) allow (i) consumers in different contexts to have different vectors of importance weights for the unobserved attributes and (ii) locations of the brands in attribute space to also be heterogeneous across consumers following a discrete heterogeneity distribution with a finite number of support points. Figure 2-2 presents examples with three brands and two segments.

<Figure 2-2 about here>

While Kim and Chintagunta (2012) parsimoniously represent the heterogeneous competitive landscape in the market, the approach is not informative about context-dependent ideal points for consumers. The proposed CDMDU model utilizes earlier formulations of unfolding models

---

[5] While verifying how well this "approximation" does relative to the full model is not possible in our particular empirical context owing to the low number of observations in some contexts, other research in single-context situations has shown that it approximates the full distribution well (see, e.g., Chintagunta 1994).

discussed above for survey data and thereby incorporates multiple ideal points while fully leveraging the benefits of the factor structure approach for heterogeneity. Importantly, unlike the aforementioned factor-structure approaches in the literature, the CDMDU model can show how far each brand is located from the ideal point (similar in spirit to previous unfolding models.) To illustrate how our CDMDU model can be useful to understand context-dependent preferences, Figure 2-3 shows a simple example of the CDMDU with three brands, two segments, and two consumption contexts. Brands and ideal points for contexts have different locations in attribute space. Furthermore, the entire map is itself heterogenous across the two segments of consumers. Finally, while not visible on the maps, the importance weights that consumers associate with the attributes also vary across consumers.

<Figure 2-3 about here>

Another limitation of the Kim and Chintagunta (2012) approach is that as long as the importance weight for a dimension is positive (negative) the substantive recommendation to managers will be to increase (decrease) the level of the corresponding attribute, with the main constraint being the costs associated with accomplishing the increase (or decrease). However, in markets where such an assumption of "vector" preferences for the attributes would not be appropriate (e.g., sugar or alcohol content of a drink), our CDMDU approach would be more appropriate. Importantly, based on the estimates from our model, we are able to determine the optimal direction in which firms may want to reposition their products to achieve specific objectives. To better understand how our modeling approach draws from previous research streams, we provide a brief flow chart in Figure 3.

<Figure 3 about here>

3. MODEL FORMULATION

Our model formulation is predicated on the availability of panel data from consumers who make decisions in several different contexts. DeSarbo and Carroll (1985) define the consumer's "dispreference" or negative preference value as the distance between the context-dependent ideal point and a brand's location (DeSarbo and Carroll 1985; DeSarbo, De Soete, Carroll, and Ramaswamy 1988; DeSarbo, De Soete, and Eliashberg 1987; DeSarbo and Rao 1986). We specify the functional form of an individual i's ($i = 1, 2, \cdots, N$) indirect utility for brand j ($j = 1, 2, \cdots, J$) on consumption occasion t as follows:

$$U_{ijt} = \gamma_j^* + \sum_{s=1}^{S} dispreference_{ijs} B_{ist} + X_{ijt}\beta + \epsilon_{ijt} \qquad (1)$$

γ_j^* is the $J \times 1$ vector of intercept terms representing mean preferences (across consumers) for the brands; $dispreference_{ijs}$ denotes consumer i's time-invariant deviation from the mean preference for brand j in context s; B_{ist} is an indicator variable, which takes the value 1 if the context s is associated with consumer i's consumption occasion t; X_{ijt} is the $1 \times Q$ vector of marketing variables (e.g., advertising), associated with brand j and consumer i's consumption occasion t; β is the $Q \times 1$ vector of parameters, associated with the covariates X_{ijt}; and the error term ϵ_{ijt} is assumed to follow an IID Type I extreme-value distribution.

In Equation 1, we impose a factor structure (Chintagunta 1994; Elrod and Keane 1995; Kim and Chintagunta 2012) on individual deviations to recover brand locations and context-dependent ideal points in multi-attribute space. We generalize the WMDU model (Carroll 1972; Carroll and Arabie 1980; DeSarbo and Rao 1984) to estimate heterogeneous context-dependent ideal points $IP_{i,z}^{s}$ and brand locations z_{ij} for context s:

$$\begin{bmatrix} dispreference_{\,i1s} \\ dispreference_{\,i2s} \\ \vdots \\ dispreference_{\,ijs} \\ \vdots \\ dispreference_{\,iJs} \end{bmatrix} = \begin{bmatrix} \left(z_{i1}^{1} - IP_{i,z}^{s,1}\right)^2 & \left(z_{i1}^{2} - IP_{i,z}^{s,2}\right)^2 & \dots & \left(z_{i1}^{K} - IP_{i,z}^{s,K}\right)^2 \\ \left(z_{i2}^{1} - IP_{i,z}^{s,1}\right)^2 & \left(z_{i2}^{2} - IP_{i,z}^{s,2}\right)^2 & \cdots & \left(z_{i2}^{K} - IP_{i,z}^{s,K}\right)^2 \\ \vdots & \vdots & \vdots & \vdots \\ \left(z_{ij}^{1} - IP_{i,z}^{s,1}\right)^2 & \left(z_{ij}^{2} - IP_{i,z}^{s,2}\right)^2 & \cdots & \left(z_{ij}^{K} - IP_{i,z}^{s,K}\right)^2 \\ \vdots & \vdots & \cdots & \vdots \\ \left(z_{iJ}^{1} - IP_{i,z}^{s,1}\right)^2 & \left(z_{iJ}^{2} - IP_{i,z}^{s,2}\right)^2 & & \left(z_{iJ}^{K} - IP_{i,z}^{s,K}\right)^2 \end{bmatrix} \cdot \begin{bmatrix} -w_{is}^{1} \\ -w_{is}^{2} \\ \vdots \\ -w_{is}^{K} \end{bmatrix} \quad (2)$$

where z_{ij}^{k} denotes the kth coordinate of brand j for consumer i; $IP_{i,z}^{s,k}$ denotes the kth coordinate of ideal point given context s for consumer i; w_{is}^{k} is the kth coordinate of the individual's sensitivity to the distance between her/his ideal point and brand location in context *s*. Thus, w_{is}^{k} represents the importance of the kth dimension in context s for consumer i. Equation 1 can be written as follows[6]:

$$U_{ijt} = \gamma_j^* - \sum_{s=1}^{S} \sum_{k=1}^{K} \left(z_{ij}^{k} - IP_{i,z}^{s,k}\right)^2 w_{is}^{k} B_{ist} + X_{ijt}\beta + \epsilon_{ijt} \qquad (3)$$

Properties of the proposed model specification

Since context-specific brand preferences can differ across users, we allow the sensitivity parameters w_{is} to vary across individuals. For the ideal point interpretation of the model, w_{is} should be non-negative for all *i* and *s*. Thus, we specify $w_{is}^{k} = exp\,(\tau_{is}^{k})$, where $\tau_{is} = (\tau_{is}^{1}, \tau_{is}^{2}, \dots, \tau_{is}^{K})$ is assumed to have a multivariate normal distribution across individuals with context-specific mean ($\boldsymbol{\tau_s}$) (represented by a $K \times 1$ vector) and a covariance matrix

$$\begin{bmatrix} (\sigma_s^1)^2 & 0 & 0 & 0 \\ 0 & (\sigma_s^2)^2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & (\sigma_s^K)^2 \end{bmatrix},$$

[6] By setting $\boldsymbol{\beta} = (\gamma_j^*, \beta)$ and $\boldsymbol{X} = (X_{1,j}, X_{2,ijt})$, we can write equation (3) as follows:

$$U_{ijt} = \boldsymbol{\beta}\boldsymbol{X}' - \sum_{s=1}^{S} \sum_{k=1}^{K} \left(z_{ij}^{k} - IP_{i,z}^{s,k}\right)^2 w_{is}^{k} B_{ist} + \epsilon_{ijt} \qquad (3')$$
$$\boldsymbol{\beta} = (\gamma_j^*, \beta)$$
$$\boldsymbol{X} = (X_{1,j}, X_{2,ijt})$$

where $X_{1,j}$ denotes the brand j's dummy variables; $X_{2,ijt}$ denotes marketing variables (e.g., advertising), associated with brand j and corresponding to consumption occasion t for consumer i; γ_j^* denotes an intrinsic preference for brand j, associated with the covariates $X_{1,j}$; and β is the $Q \times 1$ vector of parameters, associated with the covariates $X_{2,ijt}$.

which is the K-dimensional identity matrix scaled by a context-specific variance term $(\sigma_s^k)^2$. The diagonal matrix restriction stems from the identification conditions laid out in Elrod (1988) and Elrod and Keane (1995). Specifically,

$$\tau_{is}^k \equiv ln(w_{is}^k) \sim N(\tau_s^k, (\sigma_s^k)^2). \qquad (4)$$

Next, we allow the parameters in $-(z_{ij}^k - IP_{i,z}^{s,k})^2$ to vary across consumers following a discrete distribution with a finite number of supports or market "segments," Π. This will yield the following distribution for the preference deviation parameters:

$$dispreference_{ijs} = \sum_{k=1}^{K} -(z_{i\pi j}^k - IP_{i\pi,z}^{s,k})^2 w_{is}^k \qquad (5)$$
$$w_{is}^k = exp\ (\tau_{is}^k)$$
$$\tau_{is}^k \sim N(\tau_s^k, (\sigma_s^k)^2)$$
$$z_{i\pi j}^k = z_{\pi j}^k,\ \ IP_{i\pi,z}^{s,k} = IP_{\pi,z}^{s,k}$$
$$i \in \Lambda_\pi, \pi = 1, 2, \cdots, \Pi$$

where $i \in \Lambda_\pi$ if consumer i belongs to segment π $(\pi = 1, 2, \cdots, \Pi)$. Given this condition, the mean dispreference of brand j within a segment π given context s can be expressed as $\sum_{k=1}^{K} -(z_{i\pi j}^k - IP_{i\pi,z}^{s,k})^2 \times e^{\tau_s^k + \left(\frac{(\sigma_s^k)^2}{2}\right)}$ and the standard deviation can be written as $\sqrt{\sum_{k=1}^{K} (z_{i\pi j}^k - IP_{i\pi,z}^{s,k})^4 \times \left(e^{2(\tau_s^k + (\sigma_s^k)^2)} - e^{2\tau_s^k + (\sigma_s^k)^2}\right)}$ (see Web Appendix B for details). An identifying assumption we need to make is that the ideal point for a "base" context is normalized to $IP_{i,z}^s = (0, 0, \cdots, 0)$. In a case where there exists only one context, our model can be written as follows:

$$U_{ijt} = \gamma_j^* - \sum_{s=1}^{1} \sum_{k=1}^{K} (z_{ij}^k)^2 w_{is}^k B_{ist} + X_{ijt}\beta + \epsilon_{ijt}. \qquad (6)$$

In other words, our model simplifies to a version of the factor structure random coefficients model (e.g., the factor-analytic probit model in Elrod and Keane 1995).

Having data on consumers making choices in multiple contexts allows the proposed model to accomodate heterogeneity across consumers *within* a context. In particular, the model has both discrete and continuous components. The discrete heterogeneity component comes from (i) the $(J-1)\times K$ brand location matrix having a discrete distribution and (ii) multiple ideal points within the sth context having a discrete distribution—that is, the $(S-1)\times K$ location matrix having a finite number of supports. The continuous component of heterogeneity comes from allowing $\tau_{is}^{k} = ln\ (w_{is}^{k})$ for each context s to have a normal distribution across consumers (i.e., w_{is}^{k} follows a log-normal distribution). Our CDMDU model can handle data sparseness for specific consumer-brand-context combinations by sharing intra- and inter-person information across multiple consumption contexts. Specifically, projecting down from brand (*J*) space to attribute or factor (*K*) space, with brand locations and ideal points varying across consumers, with a finite number of supports, and with consumer- and context-specific attribute weights, facilitates model parsimony. We discuss the identification of the model parameters in extensive detail in the Web Appendix C.

Model estimation

We assume a Type I extreme-value distribution for the error term ϵ_{ijt} to obtain the multinomial logit model (McFadden 1974) for our choice probabilities. Based on the proposed model specification, the conditional choice probability P_{ijt} of consumer *i*'s preference for j associated with context s can be expressed as

$$P_{ijt} = \frac{exp(\gamma_j^* - \sum_{s=1}^{S}\sum_{k=1}^{K}\left(z_{ij}^{k} - IP_{i,z}^{s,k}\right)^2 w_{is}^{k}\, B_{ist} + X_{ijt}\beta)}{\sum_{l=1}^{J} exp(\gamma_l^* - \sum_{s=1}^{S}\sum_{k=1}^{K}\left(z_{il}^{k} - IP_{i,z}^{s,k}\right)^2 w_{is}^{k}\, B_{ist} + X_{ilt}\beta)}\,. \quad (7)$$

Since the parameters in the formulation are allowed to vary across users, which involves unobserved heterogeneity, we let the set of parameters $\{\theta_i, \psi_i\}$ vary across all consumers, where $\theta_i = \{\tau_{is}\}$ and $\psi_i = \left\{z_{ij}, IP_{i,z}^{s}\right\}$ for estimation. We assume that θ_i is a realization of the random variable θ that has a multivariate normal distribution $G(\theta)$; and ψ_i is a realization of the random vector ψ, which has a multivariate discrete distribution $H(\psi)$ across consumers with a finite number of supports Π, with each support associated with probability mass $\varphi(i \in \Lambda_\pi)$, $\pi = 1, 2, \cdots, \Pi$. So consumer i belongs to segment π with probability $\varphi(i \in \Lambda_\pi)$ such that $\sum_{\pi=1}^{\Pi} \varphi(i \in \Lambda_\pi) = 1$. Conditional on the set of parameters $\{\theta_i, \psi_i\}$, the likelihood function $L_{i|\theta_i,\psi_i}$ for consumer *i* is

$$L_{i|\theta_i,\psi_i} = \prod_{t=1}^{T_i} \left\{ \left\{ \prod_{j=1}^{J} \left(\frac{exp(\gamma_j^* - \sum_{s=1}^{S} \sum_{k=1}^{K} \left(z_{ij}^k - IP_{i,z}^{s,k}\right)^2 w_{is}^k B_{ist} + X_{ijt}\beta)}{\sum_{l=1}^{J} exp(\gamma_l^* - \sum_{s=1}^{S} \sum_{k=1}^{K} \left(z_{il}^k - IP_{i,z}^{s,k}\right)^2 w_{is}^k B_{ist} + X_{ilt}\beta)} \right)^{y_{ijt}} \right\} \right\} \quad (8)$$

where y_{ijt} is an indicator function, which takes a value of 1 if consumer i chooses brand j on occasion *t* and 0 otherwise. Each consumer *i* has a different total number of consumption occasions, denoted by T_i.

The above represents a version of the random effects model with a hybrid discrete-continuous heterogeneity distribution. After integrating the conditional likelihood function ($L_{i|\theta,\psi_i}$) over the multivariate normal distribution $G(\theta)$, the likelihood function for consumer i ($L_{i|\psi}$) is given by

$$L_{i|\psi} = \int_{\theta} L_{i|\theta,\psi} \, dG(\theta). \quad (9)$$

The individual likelihood in Equation 9 can be approximated through simulation as follows: (a) Draw a value of θ from a multivariate normal distribution function $G(\theta)$, θ^r, with the superscript r indicating the r^{th} draw. (b) Calculate the conditional likelihood function $L_{i|\theta^r,\psi}$ with θ^r. (c) Repeat steps (a) and (b) a large number, say *R,* of times and compute the average value. The simulated likelihood function can be expressed as below

$$\tilde{L}_{i|\psi} = \frac{1}{R} \sum_{r=1}^{R} L_{i|\theta^r,\psi}. \quad (10)$$

Note that $\tilde{L}_{i|\psi}$ is an unbiased estimator of $L_{i|\psi}$ by construction. Then, for the random vector ψ, we integrate over the discrete distribution H(ψ) with a finite number of supports, which leads to the following sample likelihood:

$$L = \prod_{i=1}^{N} \left\{ \sum_{\pi=1}^{\Pi} \tilde{L}_{i|\psi} \cdot \varphi(i \in \Lambda_\pi) \right\}. \tag{11}$$

4. EMPIRICAL ILLUSTIRATION[7]

4.1 Data

We use a subset of the individual-level panel data on U.S. beer consumption for the 24 months from January 2006 to December 2007 that was provided by Research International and also used in Kim and Chintagunta (2012). One of the useful characteristics of the data is that each panelist's brand choice (e.g., Bud Light) at any given consumption occasion is uniquely associated with his or her consumption context (e.g., watching TV at home). In Table 3 we provide descriptive statistics on the context-specific consumption frequencies.

<Table 3 about here>

As shown in Table 3, there are a total of 15,615 consumption occasions with 5 consumption contexts being defined and arranged in descending order of frequency in the data as follows: (1) relaxing, watching TV at home, (2) eating a meal at home, (3) working or doing a hobby at home, (4) at a dance club or sports bar, and (5) at a local pub. We focus on the 10 biggest beer brands by share in the sample, which account for 99.76% of all purchases by the consumers during the sample period. All the other brands chosen by these panelists are classified into a single composite 11th brand. Budweiser and Bud Light are the market share leaders (21.47% and 21.31%, respectively)

[7] To show that our proposed model can recover the true parameter values, we ran model simulations 100 times. We used a simple model setup with one segment for the top 10 brands and five contexts. In this setting, our estimation was able to recover the true parameters quite well. Details are available from the authors.

followed by Miller Lite (13.63%) and Coors Light (11.13%). As we can see, context 1 is the most frequent for all brands. Nevertheless, we find that preferences do vary across contexts. Further, the data show sparseness for certain brand consumption–context combinations, a feature our propsed model can handle.

We combine the above data with aggregate, market-level advertising information provided by TNS Media Intelligence (TNS MI) reports. These data are available at the designated-market-area (DMA) level and include weekly media advertising expenditures (such as those on cable TV, network TV, consumer magazines, newspaper, radio, outdoor advertising, internet, etc.). Since network TV advertising expenditures are common across consumers in all markets, we allocate total network TV dollars to each DMA in proportion to that DMA's television viewing households relative to the total number of television-viewing households in the United States. We construct a variable for the local advertising expenditures by combining the expenditures on local newspaper, radio, cable TV, internet, magazines, and outdoor advertising at the DMA level. We present descriptive statistics on the advertising expenditures, both national and local, in Table 4, below.

<Table 4 about here>

4.2 Joint space map of context-dependent multidimensional unfolding (CDMDU)

4.2.1 Parameters to estimate and their identification

Consider a dataset that has J brands and S contexts where we seek to project preferences onto a K-dimensional attribute space for each consumer segment π, $\pi = 1,2,\ldots, \Pi$. The number of parameters to be estimated would consist of the following categories: (i) $(J-1)\times K$ for the locations of brands ($z_{\pi j}$), with one brand located at the origin for each π, resulting in a total of $\Pi\times(J-1)\times K$; (ii) $\Pi\times(S-1)\times K$ parameters for the ideal point locations ($IP^{s}_{\pi,z}$) for each context ($s = 1,2,\ldots,S$), with the ideal point for one context constrained to lie at the origin; (iii) parameters corresponding to the (log-normal) distribution of consumer-specific attribute importance weights (w^{k}_{is} or τ^{k}_{is}) for each of the S contexts. From the identification in a single context case (e.g., Elrod and Keane 1995) we know that the mean vector in that case is zero. However, given the log-normal distribution in our case, we set $\tau^k_{s=\text{base}} = -\frac{ln2}{2}$ and $\sigma^k_{s=base} = \sqrt{ln2}$ for one context. This yields a mean $E\left(w^k_{is=base}\right) = exp\left(-\frac{ln2}{2} + \frac{ln2}{2}\right) = exp(0) = 1$ and a standard deviation, $\boldsymbol{sd}\left(w^k_{is=base}\right) = 1$. Thus, for this "base" context and for segment π, the intercept will be given by $\gamma^*_j - 1 \times \sum_{k=1}^{K}\left(z^k_{\pi j}\right)^2$ (see Equation 5 above and the subsequent discussion). We set context 5 as the base.

Since the set of parameters embedded in the distances from the ideal point, $\left(\left(z^k_{\pi j} - IP^{s,k}_{\pi,z}\right)^2\right)$—that is, the locations $(z^k_{\pi j})$ and the ideal point locations $(IP^{s,k}_{\pi,z})$—appear multiplicatively with the attribute weight parameters (w^k_{is}), one might be concerned about identifying them separately. Indeed, the previous literature needed to fix some of the distances or the weights to achieve identification. One of the key benefits of having data generated from multiple consumption contexts is that we are no longer restricted in this way. To see why, we could simply have first created five datasets, with each dataset containing the choices within one specific context. Each of these datasets would be like the ones used in the previous literature so we can get the mean preferences (γ^*_j) and brand locations $(z_{\pi j})$. Recall from Equation 6 that for the single-context case, we need to restrict the ideal point to lie at the origin and the distribution of attribute weights to have a mean of 1 and an identity covariance matrix. Now our approach pools data across S contexts but constrains γ^*_j and $z_{\pi j}$ to be the same across contexts. Doing so allows us to identify the ideal points for S–1 contexts (the base context still has it located at the origin) as well as the mean and variance of the distribution of weights for these S–1 contexts. Nevertheless, we need to fix some of the parameters to ensure identification. Furthermore, we are able to extract additional heterogeneity by allowing the locations and ideal points to be segment specific.

4.2.2 Model selection

We estimate the model parameters with the top 10 brands $(J-1=10)$, five consumption contexts $(S=5)$, and two marketing variables $(Q=2)$. Table 5 provides model fit statistics including AIC,[8] BIC,[9] and log-likelihood for several combinations of segments and dimensions (see DeSarbo et al. 2008, p. 152, on the need for model selection heuristics).

<Table 5 about here>

In the discussion of the model parameters, we focus on a two-dimensional map $(K=2)$ and two supports for the consumer segments ($\Pi=2$). Thus, the number of parameters is 83 (= $(J-1)$ parameters for mean preferences (γ_j^*); $(J-1)\times \mathrm{K}\times \Pi-1\times \Pi$ parameters for brand locations ($z_{\pi j}^k$); $(S-1)\times \mathrm{K}\times \Pi$ parameters for ideal point locations ($IP_{\pi,z}^{s,k}$); $(S-1)\times \mathrm{K}\times 2$ parameters for consumers' sensitivities to distance from ideal points; (w_{is}^k); 2 parameters for marketing mix variable effects; 1 parameter for the number of segments).

While the results in Table 6 suggest that including more segments and/or unobserved attribute dimensions improves fit, the nature of our substantive conclusions is largely unchanged. Furthermore, interpreting results with larger numbers of segments and attributes, while conceptually straightforward, negatively impacts interpretability. These detailed results are, however, available from the authors.

---

8 The AIC value of the model is

$$AIC = 2k - 2\,ln(\hat{L}),$$

where L denotes the maximized value of the likelihood function for the model and k is the number of estimated parameters in the model.

9 The BIC value of the model is

$$BIC = -2\,ln(\hat{L}) + k \cdot ln(n),$$

where $\hat{L}$ denotes the maximized value of the likelihood function of the model M (i.e., $\hat{L} = p(x|\hat{\theta}, M)$); x is the observed data; $\hat{\theta}$ is the parameter values that maximize the likelihood function; k is the number of estimated parameters in the model; and n is the number of observations.

4.2.3 Comparison with Kim and Chintagunta (2012)

To establish the incremental benefits of our proposed model vis-à-vis the Kim and Chintagunta (2012) specification, we begin with an empirical comparison of the fit and predictive ability of our model with those from that previous specification. We use the predictive log-likelihood criterion (e.g., Andrews, Ainslie, and Currim 2002; Varki and Chintagunta 2004) as the metric for comparison. We created a calibration sample and a holdout sample by randomly sampling half the panelists. After estimating the two models on the calibration sample and holding the parameters fixed at the optimal values, we then calculated the log-likelihood returned by the holdout sample. As seen in Table 6, we find that our proposed CDMDU model has the better predictive fit, with a lower value of AIC and BIC in both estimation and holdout samples.

<Table 6 about here>

4.2.4 Empirical results

In Table 7, we present our empirical results on the illustrative model with two segments and two unobserved attributes.

<Table 7 about here>

The parameter estimates in Table 7 correspond to Equation 3. For each of the two segments, we present brand locations ($z_{\pi j}$) along each dimension, context-dependent ideal points ($IP^{s}_{\pi,z}$) for each dimension, and the mean of the distribution of consumers' attribute weights ($w_{is} = exp\ (\tau_{is})$) and the estimated variances. We find that the two segments are quite balanced with, 59.31% of consumers (860) in Segment 1 and 40.69% (590) consumers in Segment 2.[10] To better understand

[10] We also estimated the model leaving w_{is} to be unconstrained (i.e., without imposing a non-negative restriction on w_{is}). This allows us to assess whether the data reveal households for which $w_{is} < 0$. We obtained –LL = 10,480.35, AIC = 21,126.71, and BIC = 22,563.60. Additionally, our estimates of SD (σ_s) are bounded away from 0, which indicates that there are a large number of customers for whom the model requires an anti-ideal point interpretation (Bordley 2011). See Web Appendix D for detailed results. As our qualitative results are similar, we focus on the ideal-point model since it allows easier interpretation. Also, see our follow-up discussion on negative ideal points in Web Appendix E.

the parameters reported in Table 7, we provide a pictorial representation of the brand locations and ideal points in pictorial form in Figure 4-1 for Segment 1.[11]

<Figure 4-1 about here>

As with typical perceptual maps, the closer two brands are in the space, the more similar are consumers' perceptions of these brands. By that token, focusing on the left panel in the figure, Budweiser and Heineken seem well differentiated in the market. The former is perceived to be similar to MGD and Natural Light, whereas Heineken is more closely associated with Miller Lite and Corona Extra. Turning to the locations of the ideal points, we see the ideal points for contexts 4 and 5 placed close together, with 1, 2, and 3 occupying proximal locations. Next we consider the distances of the brands from the ideal points for the contexts. Focusing on context 5, since consumers in context 5 give equal weights[12] to dimensions 1 and 2, we find that consumer preferences in the CDMDU model trail off uniformly in all directions away from that ideal point. These "iso-preference" curves for context 5 are represented by the concentric circles seen in Figure 4-1 (right panel).[13] The visualization on the right side of Figure 4-1 shows that the lowest dispreference for context 5 is for Busch, followed by Miller Lite, Corona Extra, Budweiser, Heineken, and so on. As we allow the competitive landscape among brands to differ across consumer segments, we also provide brand locations and ideal points for Segment 2 consumers along with the context-dependent iso-preference contours in Figure 4-2.

<Figure 4-2 about here>

[11] The ideal point for context 5 and the location of the composite brand are constrained to lie at the origin.

[12] As noted previously, for context 5, the weights are $w^1_{s=2} = 1$ and $w^2_{s=2} = 1$ and the square root of variance is normalized to 1 so that we can better discuss managerial insights from weights varying across contexts.

[13] In the literature on the general unfolding model (Carroll and Arabie 1980), the elliptical shapes centered on each context-dependent ideal point are referred to as "isopreference contours" since two points on an ellipse represent the same level of dispreference relative to that context. In Appendix E, we discuss how to interpret model parameters when the distance from the ideal point might have a positive impact on utility (i.e., an "anti"-ideal point or the signs of the effects of distance can vary across the attributes of the map).

There are important differences in the brand configurations for Segment 2 compared to Segment 1. First, Budweiser and Heineken are not as clearly differentiated in this segment; Heineken and Corona Extra continue to compete with each other. There are also differences vis-à-vis the locations of the ideal points for this segment. In particular, while those for contexts 1, 2, and 3 continue to be close, context 4 seems to be located farther away. On the other hand, the ideal points for contexts 4 and 5 are more proximal in this case. Next, in the right-hand panel of Figure 4-2 we show the iso-preference contours for context 5, and for comparison also show them for context 3. Unlike in context 5, consumers in context 3 put more weight on dimension 1 than on dimension 2—that is, $w_{s=3}^{1}$= 1.0202 and $w_{s=3}^{2} = 0.3238$ (see Table 7). Similarly, consumers in contexts 1, 2, and 4 also put more weight on dimension 1 than on dimension 2. So the iso-preference contours for context 3 (and 1, 2, and 4) are represented by the vertically compressed ellipses for context 3 in Figure 4-2. Even though the Euclidean distance between High Life and the ideal point is greater than the distance between Coors Light and the ideal point (i.e., 9.1745 > 8.3723), the dispreference for Coors Light is greater than that for High Life (i.e., 8.6220 > 3.9920). Hence, the iso-preference contours in a joint space map like Figure 4-2 can be helpful to marketers by enabling intuitive interpretations of the context-dependent competitive landscape among brands. In Web Appendix F we provide a detailed description of how to calculate (and draw) elliptical geometry to understand iso-preference contours in a joint space map and for illustrative examples.

It is worth noting that our joint space map contains only the dispreference terms (i.e., the relationship between the brand locations and context-dependent ideal points). To obtain the brand preference for any given brand j in context s for individual i, however, we have to consider both the mean preference (γ_j^*) for brand j and the individual i's deviations ($dispreference_{ijs} = \sum_{k=1}^{K} -\left(z_{i\pi j}^{k} - IP_{i\pi,z}^{s,k}\right)^2 w_{is}^{k}$) from that mean preference level. In Tables 8-1 and 8-2 we report these preferences for the two segments.

<Tables 8-1 and 8-2 about here>

In Figure 4-1, for example, we see that Busch is located closer to all the contexts than Budweiser, with its dispreference closer to 0, as seen in Table 8-1. In contrast, Budweiser's mean preference ($\gamma^*_{j=2}$) is 16.3280 whereas that for Busch is 3.0769. Thus, Budweiser is able to recover from its dispreference in all contexts (i.e., Budweiser's predicted market share is 33.3% on average for Segment 1, which is consistent with its 29. 5% of observed market share in Segment 1) but Busch ends up with a lower preference and hence a lower overall share (i.e., Busch's predicted market share is 9.1% on average for Segment 1, which is consistent with its 7.7% of observed market share in Segment 1). We find that (a) the overall brand preferences conform to the observed market shares, and (b) there is rich variation in preferences across brands, segments, and contexts, providing marketers with the important ability to differentiate their brands either to different consumer groups (i.e., segments) or different situations (i.e., contexts).

The final set of estimates concern the effects (β) of national and local advertising on the choices made by consumers. As noted previously, the estimation leverages variation in advertising spending across markets and over time. We report parameter estimates and standard errors for these marketing variables in Table 9.

<Table 9 about here>

The results in Table 9 show that both local and national brand advertising have a positive and statistically significant effect on consumers' brand choices. Our results for both local and national advertising taken together with the repositioning we suggest for brands below show this to be a potential vehicle for effecting positioning changes in this market.

4.3 Consumer behavior across different contexts

The interaction between the person and contextual influences on behavior has long been discussed in the consumer psychology literature (e.g., Belk 1974, 1975; Puto 1987). For instance, different consumption situations may have primed different attitudes and activated different goals for the different individuals (Bargh and Tota 1988; Tversky and Kahneman 1981). As such, motivating conditions (e.g., tiredness, thirst, boredom) and desired product attributes can change as the consumption environment or context (e.g., working around the house, hanging out at a party) changes (Yang et al. 2002). In particular, the consumption environment is associated with heterogeneous motivating conditions. That is, even if consumers face the same consumption context, different motivating conditions and brand preferences can arise. Since the benefits sought by the consumer change and these are related to preferences for specific product attributes, brand preferences change across consumption contexts (i.e., consumers behave differently across contexts). In our empirical illustration, specifically, the contexts within which the product category, beer, is consumed do matter—the same individual picks a different brand of beer depending on whether he or she is at home watching TV or hanging out at a local pub. While lab-based experiments have certainly found evidence consistent with some of our results, we emphasize here that our implications are obtained not in an experimental setting but from behavioral data.

In Figure 4-2, for example, consumers' ideal points are located close to one another for contexts 1, 2, and 3 and for contexts 4 and 5. Given that contexts such as consumption environments, locations, and activities can trigger automatic processes to impact consumer choice (Huang et al. 2015), what might drive such colocating patterns of context-specific ideal points? In the cases of contexts 1 (relaxing and watching TV at home), 2 (eating a meal at home), and 3 (working or doing a hobby at home), all involve activities consumers enjoy at home. While contexts 1 and 3 are associated with nonsocial consumption environments, activities in context 1 are more relaxing in nature (e.g., watching sports games, reading newspapers/books/magazines, internet surfing) than the activities in context 3 (e.g., fixing the car in the garage). On the other hand, while context 4 (dance club or sports bar) and 5 (local pub) can be seen as out-of-home activities in more noisy consumption environments, contexts 4, 5, and 2 (eating a meal at home) seem to be associated with a social consumption environment where people sit or hang out together, eat food, and engage in conversations.

If we simply categorize the environments as social or nonsocial and at home or out of the home (Yang et al. 2002), the colocations of ideal points in our empirical illustration appear to be driven more by the second of the two abovementioned environments. As discussed in Funnell (1988, 1997), both environments and motivations can be associated with consumers' unobserved heterogeneity for brand preferences. For instance, if consumers are enjoying beer together with food or snacks in such environments, they might seek common benefits of beer (i.e., specific product attributes) in contexts 2, 4, and 5. Context 4 (dance club or sports bar), however, is distinctly located from the other four contexts. This could be driven by unobserved consumer motivations because consumers may also enjoy listening to music and engaging in the more physical activities at a dance club or sports bar.

While we can similarly interpret the consumer behavior reflected in Figure 4-1, the colocating pattern of contexts is different between the two segments, especially for the location of context 4's ideal point. Compared to Figure 4-2, the ideal point for context 4 in Figure 4-1 is located closer to contexts 1, 2, and 3. Also, from the locations of brands, each of which has specific characteristics (e.g., bitterness, % alcohol, etc.), we can find differences in context-specific brand preferences between the two segments. This is even before taking into account the mean preference (γ_j^*) for brands—for example, whereas Segment 1 consumers prefer Busch the most and prefer Miller Light and Corona Extra almost equally in context 5, Segment 2 consumers prefer Miller Lite the most in context 5. As such, the CDMDU model can address such unobserved consumer heterogeneity in brand preferences across contexts, which might help in further theorizing by psychologists interested in these topics.

A related question that arises then is what beer brands consumers choose across different contexts. The utility depends both on the mean preference (γ_j^*) for brand j and individual i's deviations $\left(dispreference_{ijs} = \sum_{k=1}^{K} - \left(z_{i\pi j}^{k} - IP_{i\pi,z}^{s,k}\right)^2 w_{is}^{k}\right)$ from that mean preference level. In our empirical example, therefore, even if different consumption contexts can influence consumer behavior, consequent changes in consumers' brand choices are more likely to occur among minor brands since, as we saw previously in Tables 8-1 and 8-2, mean preferences play a major role in driving overall utility. This also indicates that minor brands with low γ_j^* can actively exploit brand repositioning strategies by reducing the distance between their locations and context-specific ideal points to create profitable niche markets for themselves. Later, we discuss how our analysis can aid in such repositioning.

4.4 Context-specific market structure

A useful metric when describing the nature of competitive market structure is to look at own-and cross-elasticities ($e_{jj}$and $e_{jj'}$ respectively) across brands (Russell 1992). (Own-) Cross-elasticity (e.g., Kamakura and Russell 1989) is defined as the percentage change in the market share of brand j, with respect to a 1% change in $(p_j)p_{j'}$, the price of brand $(j)j'$. Likewise, in our setting, cross elasticity $\eta_{jj'}$ can be defined as the percentage change in the market share of brand j, with respect to a 1% movement in the location of j' toward the ideal point. We use $\eta_{s,E,j'j'}^{\pi}$ $\left(\text{or } \eta_{s,E,jj'}^{\pi}\right)$ to denote the own- (or cross-) elasticities corresponding to a change in the K-dimensional Euclidean distance between $z_{\pi j}$ and $IP_{\pi,z}^{s}$ within a segment π given context s (see Web Appendix G for technical details). For better understanding, in Table 10 we compute the own- and cross-elasticities for the beer brands for Segment 2, below.

<Table 10 about here>

For example, the own-elasticity of Miller Lite $\eta^{\pi=2}_{s=3,E,j'=3,j'=3}$ is $0.00068\left(=\frac{0.33260-0.33237}{0.33237}\right)$. That is, if Miller Lite moves toward the ideal point of context 3 by 1%, Miller Lite's market share will increase from 33.237% to 33.260%. Similarly, we can interpret the cross-elasticity. Corresponding to Miller Lite's move toward context 3, the market share of MGD will decrease from 10.285% to 9.764%. Thus, the cross-elasiticy $\eta^{\pi=2}_{s=3,E,j=2,j'=3}$ is –0.05063 $\left(=\frac{0.09764-0.10285}{0.10285}\right)$. On the other hand, if MGD moves toward context 3, Miller Lite's market share will decrease from 33.237% to 32.301%, leading to a higher cross-elasticity, $\eta^{\pi=2}_{s=3,E,j=3,j'=2}$, which is –0.02819 $\left(=\frac{0.32301-0.33237}{0.33237}\right)$. Comparing these two cross-elasticities vis-à-vis Figure 4-2, we can say that Miller Lite has a higher impact on MGD than MGD has on Miller Lite.

Since cross-elasticities can reflect brand j''s power to steal market share from competitors, the asymmetrical cross-elasticities can show how conditions of the marketplace translate into changes in market shares. It appears that unpopular (e.g., Busch and Heineken) brands are more likely affected by competitors' movement than popular brands (such as Miller Lite). To clearly see why such a market structure exists, we utilize the concepts of *clout* and *vulnerability* (Cooper 1988; Kamakura and Russell 1989). While the original concepts were used to discuss the potential impact of changing prices, we apply them to the brand location shift in the product attribute space (e.g., innovation in a certain product attribute). Thus, we consider each brand's repositioning strategy by adjusting the distance between its location and consumers' ideal point associated with each different consumption context. We define summary measures of brand competition for each brand j' within a segment π ($i \in \Lambda_\pi$) given context s with competitive clout ($Competitive\ Clout^{\pi}_{s,j'}$) and vulnerability ($Vulnerability^{\pi}_{s,j'}$). These measures are computed by summing over all the cross-elasticities of brand j''s competitors (see Web Appendix G). Specifically, competitive clout[14] and vulnerability can be expressed as follows:

$$Competitive\ Clout^{\pi}_{s,j'} = \sum_{j' \neq j} (\eta^{\pi}_{s,j\,j'})^2 \qquad (14)$$
$$Vulnerability^{\pi}_{s,j'} = \sum_{j' \neq j} (\eta^{\pi}_{s,j'\,j})^2 \,.$$

The $Competitive\ Clout^{\pi}_{s,j'}$ shows what happens in the market when brand j' adjusts its distance closer to the ideal point for a context s within segment π. The distance reduction by brand j' with a high level of competitive clout will have a significant impact on the shares of other competing brands. On the other hand, the $Vulnerability^{\pi}_{s,j'}$ metric shows the extent to which the share of brand j' will be impacted by all the other brands' adjustments to reduce the distance between their locations and the ideal point. That is, a brand j' with high vulnerability will experience a significant reduction in market share as a result of its competitors' movement toward the ideal point.

Confining our attention to brands appearing in Figure 4-2 (right side), in Table 11 we compute both competitive clout and vulnerability below.

<Table 11 about here>

Among major brands, Miller Lite (Brand 3) appears to be in a good position because it has low vulnerability (and high relative competitive clout). Likewise, MGD (Brand 7) is also in a relatively good position among minor brands. On the other hand, Figure 4-2 indicates that if Miller Lite moves toward context 3, its distance to the ideal point for context 5 becomes longer. Specifically, Miller Lite's market share in context 5 will decrease from 34.063% to 33.961% (in Segment 2). Given different market structures across contexts, therefore, the question arises as to the possibility

---

[14] As discussed in Kamakura and Russell (1989), Cooper's (1988) measure of clout can be expressed as $Clout^{\pi}_{s,j'} = Competitive\ Clout^{\pi}_{s,j'} + \eta^{\pi}_{s,j'j'}$.

of correlations between consumers' preferences (i.e., location of context-specific ideal points) across contexts.

If a firm ignores such a possibility as shown in Miller Lite's case and relocates its brand toward a certain context-specific ideal point, it may end up in a worse position vis-à-vis other context-dependent ideal points (i.e., higher dispreference for those contexts). Since firms can devise marketing strategies targeting consumers in each different context or market, firms' marketing strategies (e.g., altering product features, product repositioning via advertising) should be constructed based on the competitive structure of each context-specific market. Therefore, in the presence of multiple context-dependent ideal points (each of which is uniquely associated with each different consumption context), a firm's movement toward the ideal point of one particular context does not necessarily improve its competitiveness in other consumption contexts but could instead hurt its overall performance in the market.

4.5 The optimal direction for brand repositioning

A typical discussion of repositioning, either in an academic or a corporate context, tends to be at a strategic level with suggestions made as to which direction a brand should move to improve its performance in the marketplace. The topic of competitive repositioning in a quantitative fashion has not been particularly widespread in the marketing literature. Here we propose an explicit, quantitative approach to understanding the effects of repositioning. We illustrate the tradeoff faced by brands when repositioning, and do so with two different brands. First, we consider repositioning strategies for Budweiser (Brand 2) in Segment 2. Since the ideal points for all five contexts are located on the lower side (see Figure 5), it would be ideal for Budweiser to move in that direction (i.e., downward).

<Figure 5 about here >

Note that $\overrightarrow{T_{s,J}^{\pi}}$ is a vector with a direction of displacement from brand location $(z_{\pi j})$ to context-dependent ideal points$(IP_{\pi,z}^{s})$ for segment π. $\overrightarrow{D_J^{\pi}}$ is defined as a vector of unit length in the direction in which we are considering repositioning the product. Thus, the vector reflects the difference between the brand location j and a new (but yet unknown) brand location such that $\left|\overrightarrow{D_J^{\pi}}\right| = 1$ for segment π. $\theta_{s,j}^{\pi}$ represents the angle between $\overrightarrow{T_{s,J}^{\pi}}$ and $\overrightarrow{D_J^{\pi}}$. Figure 5 displays a vector with displacement direction $(\overrightarrow{T_{s=2,J=2}^{\pi=2}})$ toward the ideal point for context 3 $(s = 3)$ from Budweiser $(j = 2)$'s current location (for Segment 2).

Let $Proj_{\overrightarrow{T_{s,J}^{\pi}}}\overrightarrow{D_J^{\pi}}$ denote the orthogonal projection of vector $\overrightarrow{D_J^{\pi}}$ onto the (non-zero) vector $\overrightarrow{T_{s,J}^{\pi}}$. Then:

$$Proj_{\overrightarrow{T_{s,J}^{\pi}}}\overrightarrow{D_J^{\pi}} = \left|\overrightarrow{D_J^{\pi}}\right| cos\left(\theta_{s,j}^{\pi}\right) = \overrightarrow{D_J^{\pi}} \cdot \frac{\overrightarrow{T_{s,J}^{\pi}}}{\left|\overrightarrow{T_{s,J}^{\pi}}\right|} \qquad (15)$$

where the operator $\cdot$ denotes a dot product; $\left|\overrightarrow{T_{s,J}^{\pi}}\right|$ is the length of $\overrightarrow{T_{s,J}^{\pi}}$; $\theta_{s,j}^{\pi}$ is the angle between $\overrightarrow{D_J^{\pi}}$ and $\overrightarrow{T_{s,J}^{\pi}}$; and $\theta_{s,j}^{\pi} = cos^{-1}\left(\frac{\overrightarrow{D_J^{\pi}} \cdot \overrightarrow{T_{s,J}^{\pi}}}{\left|\overrightarrow{D_J^{\pi}}\right|\left|\overrightarrow{T_{s,J}^{\pi}}\right|}\right)$. This geometrical concept of projection helps us evaluate the influence of a brand's repositioning in a specific direction for each of the five contexts. Next, we consider Miller Lite's (Brand 3) optimal direction for repositioning. Here it is not clear which direction to move in since the brand is located between contexts (see Figure 6). For instance, we see an orthogonal projection of Miller Lite (Brand 3) for both context 3 and context 5, as the dashed and dotted vector, respectively.

<Figure 6 about here >

Briefly, as seen in Figure 6, the cosines of the angles show the degree of association between $\overrightarrow{D_J^{\pi}}$ and $\overrightarrow{T_{s,J}^{\pi}}$. For context 3, since $Proj_{\overrightarrow{T_{s=3,J=3}^{\pi=2}}}\overrightarrow{D_{J=3}^{\pi=2}}$ and $\overrightarrow{T_{s=3,J=3}^{\pi=2}}$ act in the same direction, the cosine of $\theta_{s=3,j=3}^{\pi=2}$ has a positive value. That is, $\overrightarrow{D_J^{\pi}}$ and $\overrightarrow{T_{s,J}^{\pi}}$ have the same direction if $0 \leq \theta <$

90 degrees. However, if $\overrightarrow{D_j^{\pi}}$ and $\overrightarrow{T_{s,j}^{\pi}}$ are in opposite directions, the vector projection has the same magnitude as $\left|Proj_{\overrightarrow{T_{s,j}^{\pi}}} \overrightarrow{D_j^{\pi}}\right|$ with a negative sign. As we can see from this figure for context 5, the cosine of $\theta_{s=5,j=3}^{\pi=2}$ has a negative value (i.e., $90 < \theta_{s=5,j=3}^{\pi=2} \leq 270$ degrees). In other words, Miller Lite's repositioning direction $\overrightarrow{D_{j=3}^{\pi=2}}$ has a strong association with context 3 rather than context 5. Therefore, based on the signs of brand*j*'s projected values, brand *j* can evaluate how its potential move might positively or negatively impact its market position in each consumption context.

As noted previously, to determine the direction of optimal repositioning, one would ideal want to represent the firm as a profit-maximizing entity and determine the equilibrium repositioning across firms. In our case, such an approach is not feasible owing to the lack of margin and repositioning cost data. Instead, we formulate the objective function in terms of the previously introduced concepts of elasticity (e.g., competitive clout and vulnerability). We acknowledge that each firm may have a different objective function when repositioning. Depending on the objective, a firm can set its own optimal repositioning path. Thus, given its flexibility, our model can be used by firms with different repositioning purposes. We believe that this enriches the managerial implications of our model. A simple objective is one in which a brand simply moves toward the ideal point of the largest context (or market) while ignoring the other contexts. Here, we discuss four additional types of objective functions and their corresponding optimal repositioning paths. Specifically, we look at the ceteris paribus impact of a potential repositioning by brand *j* for segment π in terms of *maximizing the objective across all contexts for that brand in that segment*. We represent the objective function as follows:

$$\sum_{s=1}^{S} (the\ specific\ objective) \times \left(Proj_{\overrightarrow{T_{s,j}^{\pi}}} \overrightarrow{D_j^{\pi}}\right) \qquad (16)$$

$$= \sum_{s=1}^{S} (the\ specific\ objective) \times cos\left(\theta_{s,j}^{\pi}\right)$$

where $the\ specific\ objective$ is

1) $f_s^{\pi} \times (OwnElasticity_{s,j}^{\pi})$
2) $f_s^{\pi} \times \left(Competitive\ Clout_{s,j}^{\pi}\right)$
3) $f_s^{\pi} \times (-Vulnerability_{s,j}^{\pi})$
4) $f_s^{\pi} \times (Competitive\ Clout_{s,j}^{\pi} - Vulnerability_{s,j}^{\pi})$

where the angle $\theta_{s,j}^{\pi}$ can be computed in terms of $\overrightarrow{D_j^{\pi}}$ and $\overrightarrow{T_{s,j}^{\pi}}$, such as $\theta_{s,j}^{\pi} = cos^{-1}\left(\frac{\overrightarrow{D_j^{\pi}} \cdot \overrightarrow{T_{s,J}^{\pi}}}{\left|\overrightarrow{D_j^{\pi}}\right|\left|\overrightarrow{T_{s,J}^{\pi}}\right|}\right)$;

f_s^{π} is the fraction of observations in context s for Segment π. Hence, the optimal direction of brand j's repositioning for Segment π can defined as the following $1 \times K$ vector:

$$optim\theta_{j,F_k}^{\pi} = cos^{-1}\left(\frac{\overrightarrow{optimD_j^{\pi}} \cdot \overrightarrow{e_{F_k}}}{\left|\overrightarrow{optimD_j^{\pi}}\right|\left|\overrightarrow{e_{F_k}}\right|}\right) \quad \text{where} \qquad (17)$$

$$\overrightarrow{optimD_j^{\pi}} = \underset{\substack{\overrightarrow{D_j^{\pi}} \\ i \in \Lambda_{\pi}}}{argmax}\left(\sum_{s=1}^{S} (the\ specific\ objective) \times cos\left(\theta_{s,j}^{\pi}\right)\right).$$

For illustrative purposes, we have computed the solutions for Miller Lite (Brand 3) for Segment 2 consumers and present all of the optimal directions for the brand with dotted arrows in Figure 7.

<Figure 7 and Table 12 about here>

We can intuitively understand the results in terms of Miller Lite's cosine ($\text{i.e.}, cos(optim\theta_{s,j=3}^{\pi=2})$) for all the contexts. If the cosine has a negative value for a context s ($cos(optim\theta_{s,j=3}^{\pi=2}) < 0$), the brand repositioning decision may come with a cost of losing market share in consumption context s. For example, Miller Lite's optimal direction for "net clout" (i.e., $\overrightarrow{optimD_{netclout}{}_{j=3}^{\pi=2}}$) leans slightly more toward context 1, context 2, context 3, and context 4 (i.e., $cos\left(optim\theta_{s,j=3}^{\pi=2}\right) > 0$, $s = 1, 2, 3, 4$). On the other hand, this comes with a negative cosine value for context 5 (i.e., $cos\left(optim\theta_{s,j=3}^{\pi=2}\right) < 0,\ s = 5$), indicating a possible decrease in Miller Lite's market shares for this context.

Table 12 presents Miller Lite's current market share and optimal repositioning paths for different types of objective functions along with market share changes for each context in response to a 1% change in Miller Lite's location in an optimal direction. The optimal directions are similar when concentrating on net clout or when maximizing competitive clout across all contexts; Miller Lite's market share decreases in context 5 but increases in the other contexts. However, different objectives can lead to different consequences in different contexts. In other words, depending on the firm's objective and corresponding optimal direction, tradeoffs can exist across contexts. Among the five different optimal directions for Miller Lite, $\overrightarrow{optimD_{own}}_{j=3}^{\pi=2}$ from maximizing its own elasticities leads to relatively high market share increases for all five contexts, with $optim\theta_{own,s,j=3}^{\pi=2}$ being its optimal direction and the five contexts being $0 \leq \theta < 90$ degrees (i.e., $cos(optim\theta_{own,s,j=3}^{\pi=2})$ > 0, s = 1, 2, 3, 4, 5).

5. DISCUSSION

A large stream of literature has shown the presence of context effects of various kinds. As Thomadsen et al. (2018) note, however, there are many different contexts in which consumers' choices can vary and it is important to consider those as well. In this paper we empirically show that the contexts within which a product category is consumed do matter—the same individual picks a different brand of beer depending on whether he or she is at home watching TV or hanging out at a local pub. Our proposed CDMDU model to understand such behavior showed the impact of contexts with "real" data. We then came up with specific ways in which managers can use the knowledge regarding choices in these different contexts to improve their market positions vis-à-vis other brands. This can provide a further impetus to study consumer choices in various contexts since it is this variation that allows firms to leverage context-specific preferences to improve their market positions.

To further emphasize the substantive implications of the model, we discuss how marketers can better understand what they need to do to move their brands toward an ideal point. Typically, repositioning is accomplished in two ways—the first is by changing the characteristics of the product (and communicating the change to consumers), and the second is by changing the messaging (i.e., the advertising toward the target consumer). To change product characteristics, one first needs to be able to translate the brand locations on the maps into levels of product characteristics. Second, if one chooses to reposition with advertising, it will be important to validate the ability of advertising to accomplish the repositioning. In this section we explore these two repositioning options by focusing on two aspects: (a) translating brand locations on the joint space map into product characteristics and (b) attempting to validate advertising-based respositioning via an MTurk exercise. In addition, there might be a concern regarding the complexity of the proposed CDMDU model, so we contrast our approach with simpler alternative approaches.

Product attributes and the dimensions of the joint space map

A challenge with the multi-attribute maps derived from our model is that the nomenclature for the attributes or dimensions of the maps cannot be obtained from the estimation. One approach to identifying the attributes is to correlate brand locations with physical and/or engineering characteristics of the product (Elrod 1988; Schiffman, Young, and Reynolds 1981). Alternatively, marketers can intuitively come up with the nomenclature for each dimension of the map (Chintagunta 1999). To this end we reviewed several studies (Lawless and Heymann 1998; Meilgaard, Dalgliesh, and Clapperton 1979; Schutz and Martens 2001) and derived the five main product characteristics associated with beer: price, calories,[15] alcohol content, color, and bitterness. International bitterness units (IBU) measure a compound in hops called alpha acids; this measurement is used to quantify bitterness and hence the taste of the beer. As a measure of

[15] Price ($) and calories/100 per 12 oz.

appearance, beer color is rated in units of the standard reference method (SRM), which has been used as the standard reference by the European Brewery Convention and the American Society of Brewing Chemists (Shellhammer 2009). For the 10 brands covered in this study, we provide information on these five product characteristics in Table 13 with abbreviations: price, calories, alcohol by volume (ABV), SRM,[16] and IBU.[17]

<Table 13 about here>

We can now explain locations along each dimension in the map as a linear-in-parameters function via regression on the product characteristics. Among all the possible combinations for the regression, we selected the subset of predictors based on R^2 and mean squared error of prediction (see Web Appendix H for the selection of characteristics with multiple regression models). As an illustration, consider the linear model formula for dimension 2 in Segment 2:

$$\begin{aligned} Dimension\ 2 = {} & -6.465 + (9.920) \times Price + (9.454) \times Calories + (-5.378) \times Color \\ & + (0.307) \times Bitterness \end{aligned}$$

<Table 14 about here>

Regression results identify the direction, size, and statistical significance of the relationship between a brand's location along each dimension and product characteristics. As seen from the *t*-values in Table 14, the coefficient of price is statistically significant at the 5% level. Although the other three characteristics missed the significance level (p = 0.05), the four predictors together explain 71.16% of the variance in a brand's location along dimension 2 (see Web Appendix H for detailed results).

The regression results show that we can move a brand in a positive direction along dimension 2 when we increase the price of the brand (+). The mean change in the movement along dimension 2 is expected to respond to a one-unit change in the number of calories by 9.454; of color by –

---

[16] A higher value of SRM corresponds to a darker color, with the values ranging from 1 to 50 SRM.
[17] A higher value of IBU indicates a more bitter taste, with the values ranging from 1 to 120 IBU.

5.378; and of the degree of bitterness by 0.307. To predict the brand's movement along dimension 2 for the following hypothetical changes in product characteristics—0.5 dollars, 0.1 ABV, 1 SRM, and 1 IBU—we calculate an expected change of 0.8344 along dimension 2: $(9.920 \times 0.5) + (9.454 \times 0.1) - (5.378 \times 1) + (0.307 \times 1) = 0.8344$. We caution the reader that while this gives us some basis for translating changes in characteristics to movements in the map and hence serves as a useful input in the repositioning exercise, the regression is based on a small number of characteristics and observations in the regression. Hence, our analysis here is suggestive of an attempt to translate locations along the map dimensions into physical characteristics that marketers can influence.

Brand repositioning via advertising

Here we attempt to validate advertising-based brand repositioning discussed above using data obtained from MTurk. The key idea here is to verify whether repositioning based on the brand and context locations translates, at least directionally, to the changes in market shares predicted by our model. To this end, we first recruited 230 participants to join a survey on beverages (e.g., coffee, soft drinks, and beer) and kept 217 participants who drink at least one beer per week. These qualified MTurk workers were presented all the information on the top 10 brands such as price, alcohol, color, and bitterness along with color photos as seen in Figure 8.

<Figure 8 about here>

Then, each participant was randomly presented a photo featuring each of the five consumption contexts. In total, 84 participants passed attention check questions and successfully completed the survey. These participants provided us with their brand choices in the five consumption contexts (i.e., their context-dependent preferences).[18] As the data showed evidence of both within and across variation in the participants' choices, we used this panel dataset to run our proposed

[18] All instruments and tables are reported in Web Appendix I.

CDMDU model. Based on the parameter estimates, we obtained both brand locations and context-dependent ideal points (see Web Appendix I). Since we are interested in a brand's repositioning and the consequent changes in consumers' brand choices, we re-contacted all of the 84 MTurk workers who had completed the survey (Part 1) and gave them the option of participating in a follow-up survey (Part 2). The time interval between Part 1 and Part 2 was about two weeks, as we had to estimate context-specific ideal points and brand locations from Part 1 and design Part 2 of the survey based on the estimation results. Among the 84 participants re-contacted, 52 participated in the follow-up survey. The main purpose of this part of the survey was to check whether advertising could effectively move a firm's position closer to the ideal point.

We focused on Corona Extra's (Brand 5) repositioning from its current position near context 4 to a position closer to the ideal point for context 2. To make survey participants perceive Corona Extra's repositioning, we showed the participants an advertisement (photo) of Corona Extra that emphasizes consumption context 2, eating a meal at home. Note that there might be a positive effect of advertising for Corona Extra (i.e., market shares of Corona Extra can increase regardless of contexts because of the advertising). Thus, for other brands, we also had the participants exposed to advertisements that emphasize a consumption context closely located to their current brand locations (i.e., without repositioning). New brand choices associated with each of the five consumption contexts were then obtained from the participants (see Web Appendix I for details).

The results of this exercise showed that advertisement appears to have a positive impact. For example, in the case of Budweiser (Brand 2), market shares increased in most contexts except for context 5. In particular, Heineken's advertising appears to be quite effective, with its market share increasing in all five contexts. Importantly, however, in the presence of other brands' advertising, Corona Extra's shares in contexts 3 and 4 decreased but its share in context 2 (i.e., targeted context) increased from 2.38% to 10.53%. While the results of this study are merely suggestive, we find them to be somewhat reassuring as to the potential ability to influence brand location by advertising differently across contexts.

Alternative Benchmark Models

As an alternative benchmark model, we consider correspondence analysis (CA hereafter), which enables researchers to easily detect relationships between variables with a seemingly easy-to-read visualization.[19] Compared to other familiar multivariate approaches to graphical data representation such as principal components analysis and canonical correlation analysis, CA offers the distinct advantage of producing two dual displays whose row and column geometries have similar interpretations. However, CA is not appropriate for hypothesis testing required for marketing research (Hoffman and Franke 1986). In fact, CA has the following well-known limitations: (1) Since the method requires some aggregation of information, using it results in the loss of information; furthermore, since there is no method for conclusively determining the appropriate number of and what specific combinations of dimensions to plot and inspect, there is an associated "curse of dimensionality." (2) It cannot address consumer heterogeneity (i.e., segmentation). (3) It is difficult to accommodate the roles of marketing variables (e.g., advertising) for marketing research. We conducted CA using a contingency table (see Table 3) with the row and column representing beer brand and consumption context, respectively. Here, we focus on the limitations in its visual interpretation: what comparisons can be made of distances—that is, between row points (i.e., beer brands), column points (i.e., contexts), and/or row and column points (i.e., beer brand-context)—in the geometric representation obtained from CA. In Web Appendix J, we provide the detailed procedures and results including tables and figures. We find that interpreting the between-set distances from correspondence analysis can potentially lead to inappropriate managerial decisions.

[19] We thank the review team for making this suggestion.

We also consider an alternative benchmark model that is much simpler than our proposed model. For instance, people may purchase expensive beer brands when out in public but purchase a relatively inexpensive brand when working around the home. Thus, we consider a model that simplifies consumption contexts into at-home and out-of-home environments.

<Table 15 about here >

In Table 15, we report in-sample and out-of-sample model fit statistics for both the proposed model and this alternative benchmark model. We find that the proposed model is better than this benchmark model in terms of both in-sample and out-of-sample fit. While acknowledging model complexity, we note that our approach also provides more specific recommendations on isues such as repositioning.

In summary, we have shown how, and the extent to which, consumer choices can be influenced by the "context" within which a product is consumed. In particular, by identifying multiple ideal points, the proposed model accommodates different preferences across contexts and segments. The major contributions of this paper are fourfold. First, we developed a new context-dependent MDU model that offers flexible and parsimonious spatial representation. It also simplifies to the factor structure random coefficients brand choice model when there is only one consumption context. Second, we applied the method to real purchase/consumption data where information on certain brand contexts was sparse. Third, given multiple context-dependent ideal points and dimension weights, we showed that a firm's movement toward the ideal point of one particular context does not necessarily improve a firm's market competitiveness in other consumption contexts but could hurt its overall performance in the market. Last, we demonstrated how a firm can optimally set the direction for its brand repositioning or product innovation based on various objective functions. Hence, our proposed framework can help firms prioritize their strategies for each consumer segment given various consumption contexts. Importantly, our study also reinforces the previous literature in consumer behavior that emphasizes that consumers may make different decision-making motivations across contexts. These motivations reflect the underlying heterogeneity across consumers that are then manifested as different brand choices under varying contexts.

Marketers have emphasized the view that choice is an interaction between the choice context and the choice process (Ben-Akiva et al. 2012). While context effects were considered as a purely heuristic, System 1 phenomenon equated with deviations from rationality (Wernetfelt 1995), higher-order cognitive processes (i.e., System 2) may result in less consistent consumer decisions than automatic responses (Lee, Amir, and Ariely 2009). Also, according to studies on the psychology of reasoning (e.g, Klauer, Musch, and Naumer 2000), context effects can occur in both System 1 and System 2 (Thomadsen et al. 2018). However, context effects in the literature have typically been understood as the way consumer choices deviate from the principles of utility maximization because of the concept's origins in the literature on judgment and decision-making (e.g., Chaiken 1980; Dhar and Gorlin 2013; Pocheptsova, Amir, Dhar, and Baumeister 2009), which accordingly limits our understanding of context effects on choice (Thomadsen et al. 2018). Given that observed choice behavior is influenced by both context effects and preferences, ignoring context effects can lead to a bias in preference measurement (Thomadsen et al. 2018). While the consumer behavior literature states that a consumer's utility function can vary across the contexts in which he/she makes a decision (e.g., Carlson and Bond 2006; DeSarbo et al. 2008; Simon 1955, 1990; Tversky and Kahneman 1991), a specific manifestation of a consumer's context-specific utility function is the presence of multiple "ideal points" corresponding to different consumption contexts. By incorporating multiple context effects into the choice in a formal mathematical manner, we separately construct dispreference terms from consumers' mean brand preferences in the utility function. Exploiting a discrete choice setting, we also generate a joint space map that allows us to numerically evaluate context effects solely associated with dispreference terms. Note that while both empirical researchers and behavioral researchers study the same phenomena, researchers have recently made efforts to build bridges between the two disciplines such as by expanding standard choice models to account for behavioral effects (Rooderkerk, van Heerde, and Bijmolt 2011). In the same vein, our modeling efforts can contribute to the behavioral literature as well.

REFERENCES

Andrews, R. L., Ainslie, A., & Currim, I. S. (2002). An empirical comparison of logit choice models with discrete versus continuous representations of heterogeneity. *Journal of Marketing Research, 39*(4), 479-487.

Bargh, J. A., & Tota, M. E. (1988). Context-dependent automatic processing in depression: Accessibility of negative constructs with regard to self but not others. *Journal of Personality and Social Psychology, 54*(6), 925.

Belk, R. W. (1974). An exploratory assessment of situational effects in buyer behavior. *Journal of Marketing Research*, 156-163.

______. (1975). Situational variables and consumer behavior. *Journal of Consumer Research* du mois de Décembre, 157-164.

Ben-Akiva, M. E., Gao, S., Wei, Z., & Wen, Y. (2012). A dynamic traffic assignment model for highly congested urban networks. *Transportation research part C: emerging technologies*, 24, 62-82.

Bennett, J. F., & Hays, W. L. (1960). Multidimensional unfolding: Determining the dimensionality of ranked preference data. *Psychometrika*, *25*(1), 27-43.

Bettman, J. R., Luce, M. F., & Payne, J. W. (1998). Constructive consumer choice processes. *Journal of Consumer Research*, *25*(3), 187-217.

Bordley, R. F. (2011). An Anti-Ideal Point Representation of Economic Discrete Choice Models. *Economics Letters*, *110*(1), 60-63.

Carlson, K. A., & Bond, S. D. (2006). Improving preference assessment: Limiting the effect of context through pre-exposure to attribute levels. *Management Science, 52*(3), 410-421.

Carroll, J. D. (1972). Individual differences and multidimensional scaling. Multidimensional scaling: Theory and applications in the behavioral sciences, 1, 105-155. *New York: Seminar Press, Inc.*

Carroll, J. D., & Arabie, P. (1980). Multidimensional scaling. *Annual Review of Psychology*, *31*(1), 607-649.

Carroll, J. D., Green, P. E., & Schaffer, C. M. (1986). Interpoint distance comparisons in correspondence analysis. *Journal of Marketing Research*, 271-280.

Chaiken, S. (1980). Heuristic versus systematic information processing and the use of source versus message cues in persuasion. *Journal of personality and social psychology*, 39(5), 752.

Chatfield, C. and Goodhardt, G. J. (1970). The Beta-Binomial model for consumer purchasing behaviour. *The Journal of the Royal Statistical Society Series C (Applied Statistics)*, *19*(3), 240-250.

Chintagunta, P. K. (1994). Heterogeneous logit model implications for brand positioning. *Journal of Marketing Research*, 304-311.

_______. (1999). Measuring the effects of new brand introduction on inter-brand strategic interaction. *European Journal of Operational Research, 118*(2), 315-331.

Clausen, S. E. (1998). Applied correspondence analysis: An introduction, 121. *Sage*.

Coombs, C. H. (1950). Psychological scaling without a unit of measurement. *Psychological review, 57*(3), 145.

Cooper, L. G. (1988). Competitive maps: The structure underlying asymmetric cross elasticities. *Management Science*, *34*(6), 707-723.

Darlington, R. B. (1968). Multiple regression in psychological research and practice. *Psychological Bulletin, 69*(3), 161.

Davison, M. L. (1976). Fitting and testing Carroll's weighted unfolding model for preferences. *Psychometrika*, *41*(2), 233-247.

______. (1988). A reformulation of the general Euclidean model for the external analysis of preference data. *Psychometrika*, *53*(3), 305-320.

Day, G. S., Shocker, A. D., & Srivastava, R. K. (1979). Customer-oriented approaches to identifying product-markets. *The Journal of Marketing*, 8-19.

DeSarbo, W. S., Atalay, A. S., LeBaron, D., & Blanchard, S. J. (2008). Estimating multiple consumer segment ideal points from context-dependent survey data. *Journal of Consumer Research*, *35*(1), 142-153.

DeSarbo, W. S., Atalay, A. S., & Blanchard, S. J. (2009). A three-way clusterwise multidimensional unfolding procedure for the spatial representation of context dependent preferences. *Computational Statistics & Data Analysis*, *53*(8), 3217-3230.

DeSarbo, W. S., & Carroll, J. D. (1981). Three-way metric unfolding. In *Proceedings of the Orsa-Tims Market Measurement Conference*.

_______. (1985). Three-way metric unfolding via alternating weighted least squares. *Psychometrika*, *50*(3), 275-300.

DeSarbo, W. S., De Soete, G., & Eliashberg, J. (1987). A new stochastic multidimensional unfolding model for the investigation of paired comparison consumer preference/choice data. *Journal of Economic Psychology, 8*(3), 357-384.

DeSarbo, W. S., De Soete, G., Carroll, J. D., & Ramaswamy, V. (1988). A new stochastic ultrametric tree unfolding methodology for assessing competitive market structure and deriving market segments. *Applied Stochastic Models and Data Analysis*, *4*(3), 185-204.

DeSarbo, W. S., & Hoffman, D. L. (1986). Simple and weighted unfolding threshold models for the spatial representation of binary choice data. *Applied Psychological Measurement*, *10*(3), 247-264.

_______. (1987). Constructing MDS Joint Spaces from Binary Choice Data a Multidimensional Unfolding Threshold Model for Marketing Research. *Journal of Marketing Research*, *24*(1), 40-54.

DeSarbo, W. S., & Kim, S. (2013). A review of the major multidimensional scaling models for the analysis of preference/dominance data in marketing. In *Quantitative Modelling in Marketing and Management* (pp. 3-25).

DeSarbo, W. S., & Rao, V. R. (1984). GENFOLD2: A set of models and algorithms for the general unfolding analysis of preference/dominance data. *Journal of Classification*, *1*(1), 147-186.

_______. (1986). A constrained unfolding methodology for product positioning. *Marketing Science*, *5*(1), 1-19.

Dhar, R., & Gorlin, M. (2013). A dual-system framework to understand preference construction processes in choice. *Journal of Consumer Psychology*, 23(4), 528-542.

Dubow, J. S. (1992). Occasion-based vs. user-based benefit segmentation: A case study. *Journal of Advertising Research*, *32*, 11-18.

Elrod, T. (1988). Choice map: Inferring a product-market map from panel data. *Marketing Science*, *7*(1), 21–40.

Elrod, T., & Keane, M. P. (1995). A factor-analytic probit model for representing the market structure in panel data. *Journal of Marketing Research*, 1-16.

Erdem, T. (1996). A dynamic analysis of market structure based on panel data. *Marketing Science*, *15*(4), 359-378.

Fennell, G. (1978). Consumers' perceptions of the product use situation. *The Journal of Marketing, 42*, 38-47.

_______. (1988). Action as counterchange: Identifying antecedents of the domain and goal of action. L. Alwitt, ed. Proceedings 95th Annual Convention of the American Psychological Association. *American Psychological Association*, Washington, DC.

_______. (1997). Value and values: Relevance to advertising. L. Kahle, L. Chiagouris, eds. *Values, Lifestyles, and Psychographics*. Erlbaum, Hillsdale, NJ.

Gönül, F., & Srinivasan, K. (1993). Modeling multiple sources of heterogeneity in multinomial logit models: Methodological and managerial issues. *Marketing Science*, *12*(3), 213-229.

Green, P. E. (1975). Marketing applications of MDS: Assessment and outlook. *The Journal of Marketing*, 24-31.

Green, P. E., & Carmone, F. J. (1970). Multidimensional scaling and related techniques in marketing analysis. *Boston: Allyn and Bacon.*

Green, P. E., & Rao, V. R. (1972). Applied multidimensional scaling. *Holt, Rinehart and Winston, New York.*

Guadagni, P. M., & Little, J. D. (1983). A logit model of brand choice calibrated on scanner data. *Marketing Science, 2*(3), 203-238.

Hoffman, D. L., & Franke, G. R. (1986). Correspondence analysis: graphical representation of categorical data in marketing research. *Journal of marketing Research*, 213-227.

Holbrook, M. B. (1984). Situation-Specific Ideal Points and Usage of Multiple Dissimilar Brands. in *Research in Marketing*, *7*, ed. Jagdish N. Sheth, Greenwich, CT: JAI Press, 93–131.

Huang, G., Khwaja, A., & Sudhir, K. (2015). Short-run needs and long-term goals: a dynamic model of thirst management. *Marking Science,* 34(5), 702–721.

Johnson, R.M. (1971). Market segmentation: a strategic management tool. *Journal of Marketing Research*, 13-18.

Kamakura, W. A., & Russell, G. (1989). A probabilistic choice model for market segmentation and elasticity structure. *Journal of Marketing Research*, *26*, 379-390.

Kim, M., & Chintagunta, P. K. (2012). Investigating brand preferences across social groups and consumption contexts. *Quantitative Marketing and Economics*, *10*(3), 305-333.

Klauer, K. C., Musch, J., & Naumer, B. (2000). On belief bias in syllogistic reasoning. *Psychological review,* 107(4), 852.

Kolpin, K. M., & Shellhammer, T. H. (2009). The human bitterness detection threshold of iso-α-acids and tetrahydro-iso-α-acids in lager beer. *Journal of the American Society of Brewing Chemists*, 67(4), 200-205.

Lattin, J. M., & McAlister, L. (1985). Using a variety-seeking model to identify substitute and complementary relationships among competing products. *Journal of Marketing Research*, 330-339.

Laurent, G. (1978), "A Study of Multiple Variant Consumption for Frequently Purchased Consumer Products," doctoral dissertation, *Sloan School of Management, Massachusetts Institute of Technology.*

Lawless, H. T., & Heymann, H. (1998). Sensory evaluation of food: practices and principles. *New York: Chapman and Hall.*

Lewin, K. (1951). Field theory in social science.

Lee, L., Amir, O., & Ariely, D. (2009). In search of homo economicus: Cognitive noise and the role of emotion in preference consistency. *Journal of consumer research*, 36(2), 173-187.

Lee, J. K., Sudhir, K., & Steckel, J. H. (2002). A multiple ideal point model: Capturing multiple preference effects from within an ideal point framework. *Journal of Marketing Research*, *39*(1), 73-86.

McAlister, L. (1982). A dynamic attribute satiation model of variety-seeking behavior. *Journal of Consumer Research, 9*(2), 141-150.

McFadden, D. (1974), 'Conditional logit analysis of qualitative choice behavior', in P. Zarembka, ed., *Frontiers in Econometrics*, Academic Press, New York, 105–142.

Meilgaard, M. C., Dalgliesh, C. E., & Clapperton, J. F. (1979). Beer flavour terminology. *Journal of the Institute of Brewing*, *85*(1), 38-42.

Pocheptsova, A., Amir, O., Dhar, R., & Baumeister, R. F. (2009). Deciding without resources: Resource depletion and choice in context. *Journal of Marketing Research*, 46(3), 344-355.

Puto, C. P. (1987). The framing of buying decisions. *Journal of Consumer Research, 14*(3), 301-315.

Russell, G. J. (1992). A model of latent symmetry in cross price elasticities. *Marketing Letters*, 3(2), 157-169.

Rooderkerk, R. P., Van Heerde, H. J., & Bijmolt, T. H. (2011). Incorporating context effects into a choice model. *Journal of Marketing Research*, 48(4), 767-780.

Schiffman, S. S., Young, F. W., & Reynolds, M. L. (1981). Introduction to multidimensional scaling: Theory, methods, and applications. San Diego, CA: Academic Press.

Schutz, H. G., & Martens, M. (2001). Appropriateness as a cognitive-contextual measure of food attitudes. In *Food, People and Society*. Springer Berlin Heidelberg., 247-266

Shellhammer, T. H. (2009). Beer color. In *Beer: A Quality Perspective*. Elsevier Inc., 213-227

Shocker, A. D., & Stewart, D. W. (1983). Strategic Marketing Decision Making and Perceptual Mapping. *Marketing Science*, 224-239.

Simon, H. A. (1955). A behavioral model of rational choice. *The Quarterly Journal of Economics, 69*(1), 99-118.

_______. (1990). Invariants of human behavior. *Annual Review of Psychology, 41*(1), 1-20.

Srinivasan, V., & Shocker, A. D. (1973). Estimating the weights for multiple attributes in a composite criterion using pairwise judgments. *Psychometrika, 38*(4), 473-493.

Srivastava, R. K., Alpert, M. I., & Shocker, A. D. (1984). A customer-oriented approach for determining market structures. *The Journal of Marketing*, 32-45.

Stein, C. (1960) "Multiple Regression", Contributions to Probability and Statistics: Essays in Honor of Harold Hotelling, I Olkin, ed., 424–443. *Stanford Univ. Press, Palo Alto*.

Thomadsen, R., Rooderkerk, R. P., Amir, O., Arora, N., Bollinger, B., Hansen, K., John, L., Liu, W., Sela, A., Singh, V., Sudhir, K., & Wood, W (2018). How Context Affects Choice. *Customer Needs and Solutions, 5*(1-2), 3-14.

Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science*, *211*(4481), 453-458.

_______. (1991). Loss aversion in riskless choice: A reference-dependent model. *The Quarterly Journal of Economics, 106*(4), 1039-1061.

Varki, S., & Chintagunta, P. K. (2004). The augmented latent class model: Incorporating additional heterogeneity in the latent class model for panel data. *Journal of Marketing Research, 41*(2), 226-233.

Wadsworth, G. P. (1960). Introduction to probability and random variables (No. 04; QA273, W3.).

Wedel, M., & DeSarbo, W. S. (1996). An exponential-family multidimensional scaling mixture methodology. *Journal of Business & Economic Statistics*, *14*(4), 447-459.

Wernerfelt, B. (1995). A rational reconstruction of the compromise effect: using market data to infer utilities, *Journal of Consumer Research*, 21 627–633.

Wind, Y. J. (1982). Product policy: concepts, methods, and strategy*, 8*. Reading, *MA: Addison-Wesley*.

Yang, S., Allenby, G. M., & Fennel, G. (2002). Modeling variation in brand preference: The roles of objective environment and motivating conditions. *Marketing Science, 21*(1), 14-31

TABLES AND FIGURES

Table 1. Literature review on the MDU models

	Model	Nature of advance
Simple MDU (SMDU) model (Coombs 1950; Bennet and Hays 1960)	$\Delta_{ij}= -w_i \sum_{k=1}^{K} (z_j^k - IP_i^k)^2 + b_i + \epsilon_{ij}$	Basic model for MDU
Weighted MDU (WMDU) model (Carroll 1972; Carroll and Arabie 1980; DeSarbo and Rao 1984; DeSarbo and Hoffman 1986, 1987; Wedel and DeSarbo 1996)	$\Delta_{ij}= -\sum_{k=1}^{K} w_i^k (z_j^k - IP_i^k)^2 + b_i + \epsilon_{ij}$	Each dimension k on the map is weighted differentially.
Three-way metric unfolding model (DeSarbo and Carroll 1981, 1985)	$\Delta_{ijs}= -\sum_{k=1}^{K} w_s^k (z_j^k - IP_i^k)^2 + b_s + \epsilon_{ijs}$	Each dimension gets a different weight for each context in which consumption occurs.
Three-way clusterwise MDU model (DeSarbo et al. 2008; DeSarbo et al. 2009)	$\Delta_{ijs}= -\sum_{\pi=1}^{\Pi} p_{i\pi} \sum_{k=1}^{K} w_s^k (z_j^k - IP_\pi^{s,k})^2 + b_s + \epsilon_{ijs}$	Segment-level ideal points can be obtained across multiple contexts.
Context-dependent MDU (CDMDU) model (This paper)	$U_{ijt} = \gamma_j^* - \sum_{s=1}^{S} \sum_{k=1}^{K} (z_{ij}^k - IP_{i,z}^{s,k})^2 w_{is}^k B_{ist} + \epsilon_{ijt}$	Retaining the spirit of MDU, the CDMDU model is applicable to any transactional (choice) data. It can also estimate both dimension- and individual-specific scale parameters by accommodating individual-level ideal points across multiple contexts in a parsimonious manner.

π $(\pi = 1, 2, \cdots, \Pi)$ is the market segment/cluster (unknown); $i \in \Lambda_\pi$ if consumer i belongs to segment π; k $(= 1, \cdots, K)$ is the dimension (unknown);
Δ_{ij} is the dispreference for brand j by consumer i; Δ_{ijs} is the dispreference for brand j by consumer i in the sth context;
z_j^k denotes the (perceived) kth coordinate of the location of the brand j; z_{ij}^k denotes the (perceived) kth coordinate of the location of brand j by consumer i;
IP_i^k denotes the kth coordinate of the ideal point by consumer i; $IP_\pi^{s,k}$ denotes the kth coordinate of the ideal point for market segment π in the sth context; $IP_{i,z}^{s,k}$ denotes the kth coordinate of the ideal point by consumer i in the sth context $(i \in \Lambda_\pi)$;
w_i denotes the (possibly positive) scale parameter by consumer i; w_i^k denotes the (possibly positive) kth coordinate of scale parameter by consumer i; w_s^k denotes the (possibly positive) kth coordinate of scale parameter in the sth context; w_{is}^k denotes the (possibly positive) kth coordinate of scale parameter by consumer i in the sth context $(i \in \Lambda_\pi)$;
b_i denotes an additive constant by consumer i; b_s denotes an additive constant in the sth context;
$p_{i\pi} = 1$ if consumer i is classified into market segment π, 0 otherwise where $\sum_{\pi=1}^{\Pi} p_{i\pi} = 1$ for non-overlapping segment π or $0 < \sum_{\pi=1}^{\Pi} p_{i\pi} \leq S$ for overlapping segment π;
γ_j^* represents the mean preference for brand j;
B_{ist} is an indicator variable, which takes the value 1 if the context s is associated with consumer i's consumption occasion t;
ϵ_{ijt} is an error term that follows a Type I extreme distribution.

Figure 1-1. SMDU

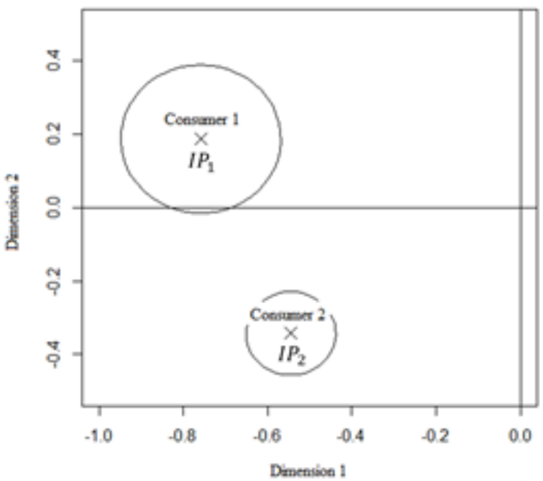


Figure 1-2. WMDU

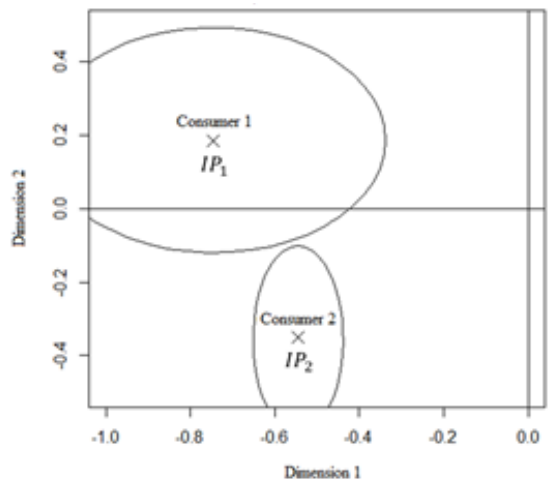


Figure 1-3. Three-way clusterwise

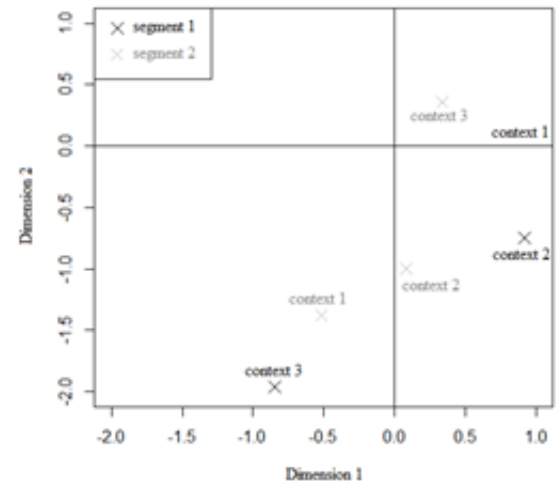


Table 2. Comparison for model applications in consumer research

	Model for context-dependent preference			Data structure			Map
	Context-dependent ideal points IP	Scale parameter w		Stated preference data (ratings)	Transactional (choice) data	Handling data sparseness	Iso-preference contours for repositioning
		Context-dependent	Heterogeneity				
SMDU model	NO	NO	YES (w_i)	YES	NO	NO	NO
WMDU model	NO	NO	YES (w_i^k)	YES	NO	NO	NO
Three-way metric unfolding model	NO	YES (w_s^k)	NO	YES	NO	NO	NO
Three-way clusterwise MDU model	YES ($IP_\pi^{s,k}$)	YES (w_s^k)	NO	YES	NO	NO	NO
CDMDU model (This paper)	YES ($IP_{i,z}^{s,k}$)	YES (w_{is}^k)	YES (w_{is}^k)	NO	YES	YES	YES

Figure 2-1. Illustration of Elrod (1988)

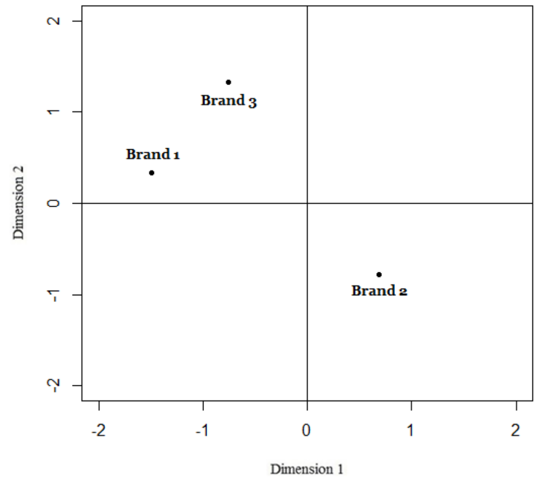


Figure 2-2. Illustration of Kim and Chintagunta (2012)

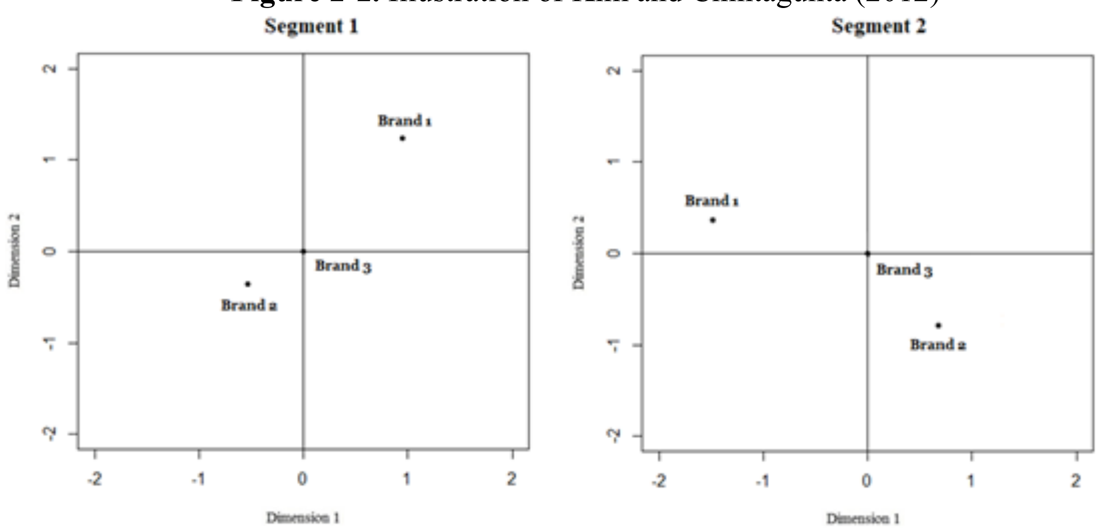


Figure 2-3. Illustration of the CDMDU (this paper)

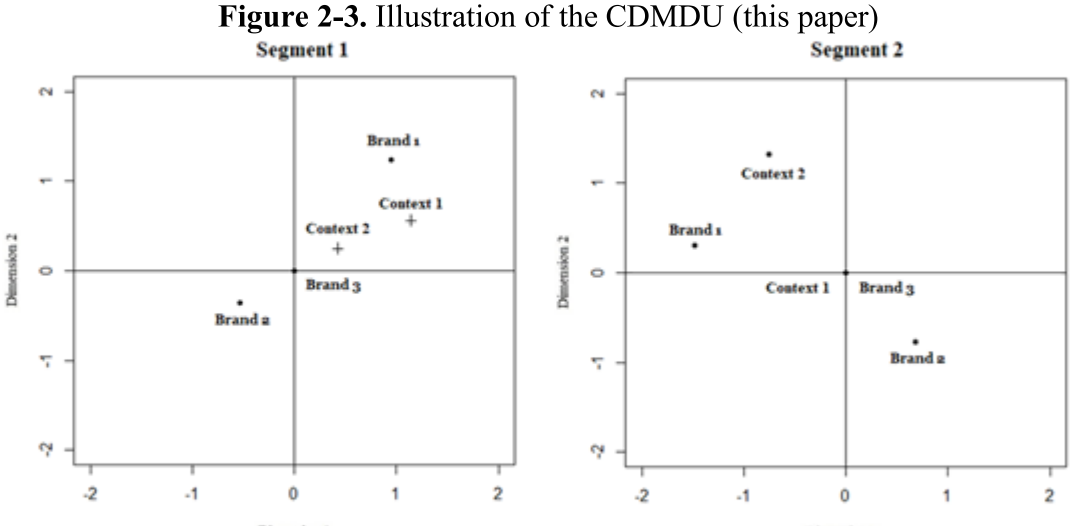


Figure 3. Flow chart for this study

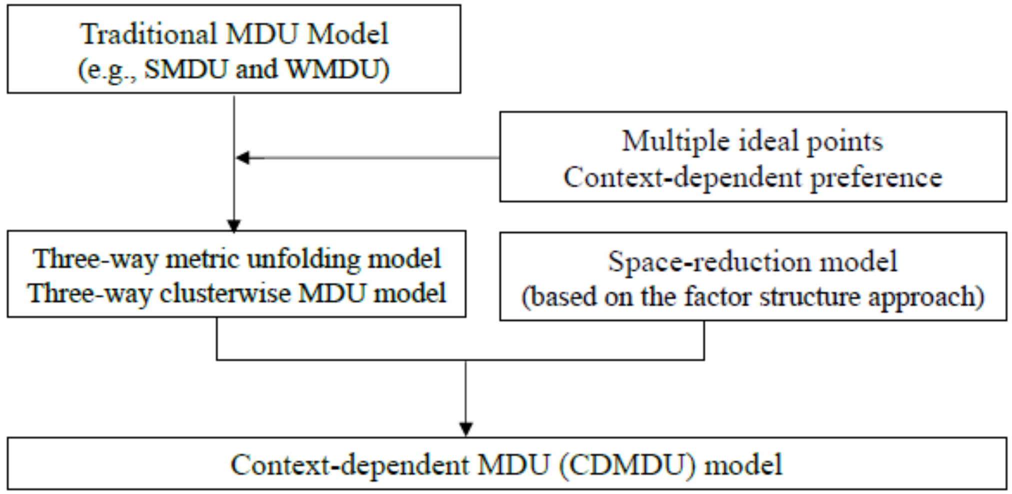


Table 3. Brand choices across contexts

	Context 1 (Relaxing, watching TV at home)		Context 2 (Eating a meal at home)		Context 3 (Working or doing a hobby at home)		Context 4 (Dance club and sports bar)		Context 5 (Local pub)		Total Brand Choice
	Frequency	(%)	Frequency	(%)	Frequency	(%)	Frequency	(%)	Frequency	(%)	Frequency
Bud Light	2,116	63.6%	342	10.3%	311	9.3%	208	6.3%	350	10.5%	3,327
Budweiser	2,297	68.5%	356	10.6%	305	9.1%	179	5.3%	216	6.4%	3,353
Miller Lite	1,241	58.3%	245	11.5%	209	9.8%	151	7.1%	283	13.3%	2,129
Coors Light	1,065	61.3%	183	10.5%	172	9.9%	124	7.1%	194	11.2%	1,738
Corona Extra	393	66.2%	56	9.4%	32	5.4%	57	9.6%	56	9.4%	594
High Life	841	64.0%	246	18.7%	104	7.9%	41	3.1%	82	6.2%	1,314
MGD	628	66.8%	103	11.0%	69	7.3%	49	5.2%	91	9.7%	940
Heineken	310	61.8%	60	12.0%	35	7.0%	33	6.6%	64	12.7%	502
Natural Light	629	68.2%	147	15.9%	112	12.1%	5	0.5%	29	3.1%	922
Busch	558	73.5%	66	8.7%	84	11.1%	11	1.4%	40	5.3%	759
Composite	19	51.4%	4	10.8%	4	10.8%	3	8.1%	7	18.9%	37
Total	10,097	64.7%	1,808	11.6%	1,437	9.20%	861	5.5%	1,412	9.0%	15,615

Table 4. Descriptive statistics on national and local advertising

Brand	National Advertising Expenditures Total U.S. Dollars (000)		Local Advertising Expenditures Total U.S. Dollars (000)	
	Year 2006	Year 2007	Year 2006	Year 2007
Bud Light	103511.5	98697.3	16030.3	21591.5
Budweiser	50198.4	50443.1	11828.3	11701.4
Miller Lite	57043.6	45950.2	8205.7	13763.7
Coors Light	74500.5	66496.1	24476.5	25282.6
Corona Extra	9005.4	13222.3	12529.3	12080.8
High Life	340	13742	2446.5	5384.8
MGD	9664.7	0	19969.4	7119.8
Heineken	7806.8	10594.5	33818.7	39814.2
Natural Light	0	0	85.1	113.6
Busch	2340	165.1	1341.8	1342.1
Composite	350.3	1691	5775.5	6738.7

Table 5. Model fit comparisons for the numbers of segments and dimensions

	-Log-likelihood	The number of parameters	AIC	BIC
$K = 1$ and $\Pi = 2$	13,962.05	57	28,038.11	29,024.89
$K = 2$ and $\Pi = 1$	14,280.59	55	28,671.18	29,623.34
$K = 2$ and $\Pi = 2$	10,949.06	83	22,064.12	23,501.01
$K = 2$ and $\Pi = 3$	9,997.97	111	20,217.93	21,067.75

*The number of observations (N) is 15,615

Table 6. Comparison of model fit and predictive ability

	KC (2012)	CDMDU
Number of consumers in-sample	725	725
Number of choice occasions in-sample	7,641	7,641
Number of parameters	67	83
(in-sample) -Log-likelihood	8,161.89	7,310.35
(in-sample) AIC	16,457.77	14,786.70
(in-sample) BIC	17,521.90	16,104.95
Number of consumers out-of-sample	725	725
Number of choice occasions out-of-sample	7,974	7,974
(Predictive) -Log-likelihood	8,432.109	7,836.92
(Predictive) AIC	16,998.22	15,839.84
(Predictive) BIC	18,068.07	17,165.17

Note: 2 dimensions and 2 segments

Table 7. Parameter estimates and standard errors

	Segment 1 (59.31%)				**Segment 2** (40.69%)			
	Dimension 1		*Dimension 2*		*Dimension 1*		*Dimension 2*	
	estimates	*s.e.*	*estimates*	*s.e.*	*estimates*	*s.e.*	*estimates*	*s.e.*
Brand location (z_{ij}^{k})								
Bud Light	-2.3475	0.1111	5.6546	0.1075	-4.6515	0.1065	3.5353	0.1093
Budweiser	0.1351	0.1109	3.5190	0.0851	-1.6744	0.1218	3.8313	0.1559
Miller Lite	-2.1427	0.1237	-0.0356	0.1326	-0.8977	0.0802	-0.4896	0.1416
Coors Light	-5.3511	0.1149	2.5593	0.0897	-4.0102	0.1269	4.1237	0.1430
Corona Extra	-2.1694	0.0993	-0.3636	0.0669	0.2884	0.1138	2.7120	0.1144
High Life	-4.1239	0.3378	5.4119	0.3198	-2.2965	0.0791	5.7402	0.1199
MGD	0.9065	0.2693	4.4755	0.1753	0.2467	0.0791	3.7023	0.1004
Heineken	-3.8314	0.1076	-0.2479	0.0829	-0.9661	0.1002	3.5268	0.1288
Natural Light	2.7241	0.1787	3.7498	0.1418	-4.8814	0.1335	-0.9379	0.1237
Busch	0.4183	0.1339	**0.0000**	**(fixed)**	1.6888	0.1480	**0.0000**	**(fixed)**
Composite	**0.0000**	**(fixed)**	**0.0000**	**(fixed)**	**0.0000**	**(fixed)**	**0.0000**	**(fixed)**
Context-dependent ideal point $(IP_{i,z}^{s,k})$								
Context 1 (Relaxing)	0.5890	0.0893	-2.9611	0.3748	0.4955	0.0867	-2.5885	0.3578
Context 2 (With a meal)	0.4767	0.0926	-2.8029	0.4145	0.5155	0.0966	-2.6840	0.3879
Context 3 (Working at home)	0.4883	0.1051	-3.6502	0.7786	0.2824	0.1193	-3.0644	0.7046
Context 4 (Dancing)	-0.0347	0.0852	-1.7162	0.2653	0.3203	0.0986	-1.2281	0.2338
Context 5 (Local pub)	**0.0000**	**(fixed)**	**0.0000**	**(fixed)**	**0.0000**	**(fixed)**	**0.0000**	**(fixed)**

*-LL = 10,949.06, AIC = 22,064.12 and BIC = 23,501.01 where the number of observations is 15,615 and the number of parameters is 83.

	Dimension 1 *Mean estimate* w_s^1	*s.e.*	*Dimension 2* *Mean estimate* w_s^2	*s.e.*	*Dimension 1* *Square root of variance estimate*†	*s.e.*	*Dimension 2* *Square root of variance estimate*	*s.e.*
Context 1 (Relaxing)	0.8911	0.0352	0.3599	0.0739	0.9442	0.0348	0.5876	0.0254
Context 2 (With a meal)	0.8271	0.0395	0.3790	0.0821	0.8557	0.0410	0.6275	0.0376
Context 3 (Working at home)	1.0201	0.0487	0.3237	0.1347	1.0133	0.0575	0.6091	0.0451
Context 4 (Dancing)	0.9307	0.0445	0.4517	0.0681	0.7952	0.0563	0.4129	0.0320
Context 5 (Local pub)	**1.0000**	**(fixed)**	**1.0000**	**(fixed)**	**1.0000**	**(fixed)**	**1.0000**	**(fixed)**

$$^{**}w_s^k \equiv E\left(w_{is}^k\right) = E\left(e^{ln\left(w_{is}^k\right)}\right) = E\left(e^{\tau_{is}^k}\right) = e^{\tau_s^k + \left(\frac{(\sigma_s^k)^2}{2}\right)}$$

† The square root of variance estimate (i.e., $\sqrt{\sum_{k=1}^{K} -\left(z_{i1}^{k} - IP_{i,z}^{s,k}\right)^4 \times \left(e^{2(\tau_s^K + (\sigma_s^k)^2)} - e^{2\tau_s^K + (\sigma_s^k)^2}\right)}$

Figure 4-1. CDMDU-based joint space map for Segment 1 (59.31% of consumers)

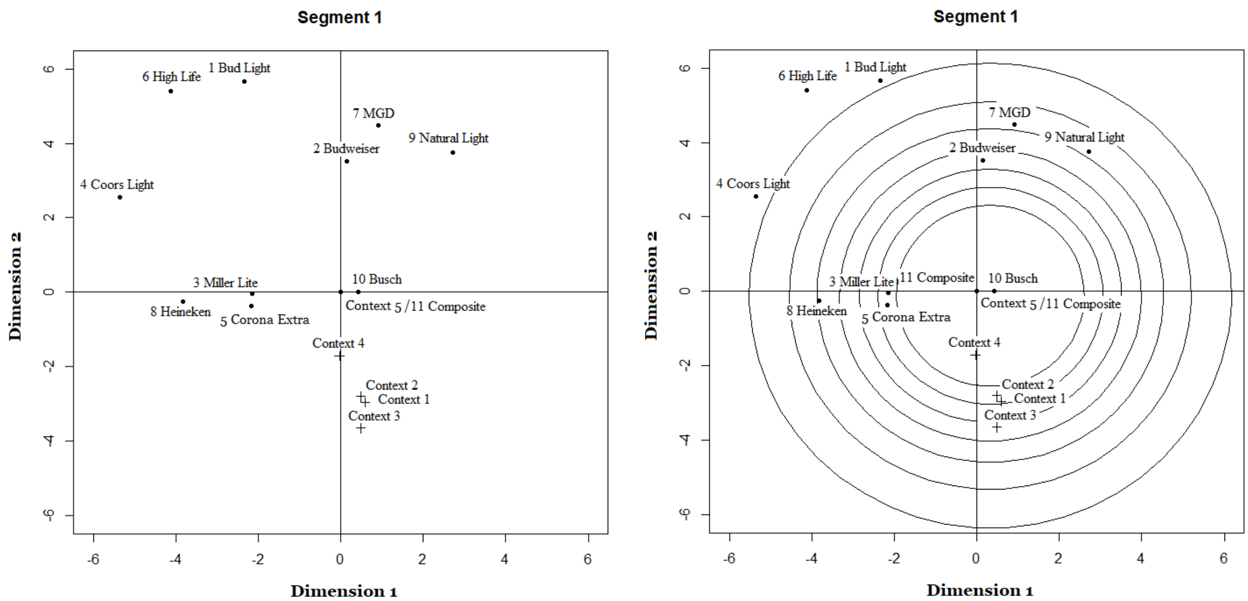


Figure 4-2. CDMDU-based joint space map for Segment 2 (40.69% of consumers)

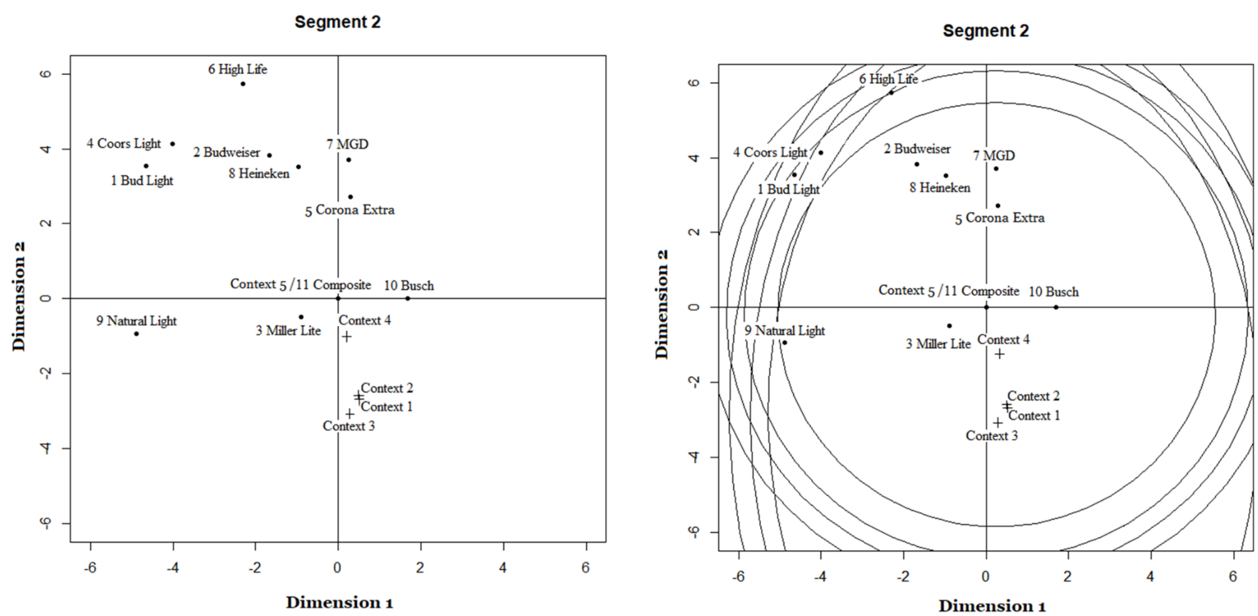


Table 8-1. Context-dependent dispreference for Segment 1

	Brand mean preference		Context-dependent dispreference									
			Context 1 (relaxing)		Context 2 (with a meal)		Context 3 (working at home)		Context 4 (dancing)		Context 5 (local pub)	
	estimates	*s.e.*	*mean*	*s.d.*	*mean*	*s.d.*	*mean*	*s.d.*	*mean*	*s.d.*	*mean*	*s.d.*
1 Bud Light	29.4462	0.8908	-34.3123	0.1654	-33.8089	0.1512	-36.0358	0.0083	-29.5589	0.0567	-37.3198	0.1025
2 Budweiser	16.3280	0.5765	-15.2651	0.0656	-15.1379	0.0036	-16.757	0.1069	-12.4531	0.0484	-12.4672	0.0121
3 Miller Lite	6.3729	0.3332	-9.8598	0.0815	-8.5278	0.1870	-11.1603	0.0482	-5.2619	0.0456	-4.511	0.0353
4 Coors Light	26.9042	0.7856	-42.3956	0.1569	-39.0749	0.0367	-47.1986	0.1364	-34.5955	0.0334	-35.3412	0.0733
5 Corona Extra	8.0570	0.3601	-9.1143	0.0740	-8.0459	0.1063	-10.7375	0.0246	-5.1205	0.0289	-4.7645	0.2265
6 High Life	27.8906	0.8211	-45.1678	0.1887	-42.9885	0.1474	-48.3053	0.0677	-38.4349	0.0380	-46.1065	0.0193
7 MGD	13.7654	0.4555	-19.8369	0.1919	-20.1687	0.0195	-21.6717	0.2592	-18.1583	0.1314	-21.0437	0.0636
8 Heineken	12.8958	0.4770	-20.1056	0.0568	-17.8422	0.0573	-22.721	0.1363	-14.3219	0.0389	-14.6843	0.0278
9 Natural Light	17.5565	0.7841	-20.3543	0.0129	-20.4564	0.0146	-22.9443	0.0164	-20.5809	0.1764	-21.4688	0.0763
10 Busch	3.0769	0.1957	-3.0707	0.0057	-3.0724	0.2129	-4.3021	0.0345	-1.406	0.1905	-0.1806	0.0811
11 Composite	**0.0000**	**(fixed)**	-3.3326	0.0110	-3.1036	0.0924	-4.4476	0.1593	-1.301	0.0045	**0.0000**	**(fixed)**

Note: We compute the means and standard deviations of context-dependent dispreferences for contexts 1 to 5 by bootstrapping. The 10,000 bootstrap statistics (each sample size: 100) are based on the parameter estimates in Table 7.

Table 8-2. Context-dependent dispreference for Segment 2

	Brand mean preference		Context-dependent dispreference									
			Context 1 (relaxing)		Context 2 (with a meal)		Context 3 (working at home)		Context 4 (dancing)		Context 5 (local pub)	
	estimates	*s.e.*	*mean*	*s.d.*	*mean*	*s.d.*	*mean*	*s.d.*	*mean*	*s.d.*	*mean*	*s.d.*
1 Bud Light	29.4462	0.8908	-36.9980	0.1074	-36.5513	0.1896	-38.9937	0.0603	-33.2945	0.0391	-34.1764	0.0416
2 Budweiser	16.3280	0.5765	-19.0686	0.0388	-19.9880	0.0653	-19.1878	0.1125	-15.3436	0.0772	-17.5333	0.0508
3 Miller Lite	6.3729	0.3332	-3.1718	0.1435	-3.5478	0.0710	-3.5734	0.0065	-1.8030	0.1759	-0.9886	0.0570
4 Coors Light	26.9042	0.7856	-34.5107	0.2033	-34.4541	0.0501	-35.6773	0.1532	-30.3945	0.0025	-33.0273	0.0593
5 Corona Extra	8.0570	0.3601	-10.1938	0.0434	-10.8995	0.1773	-10.8042	0.0018	-6.9286	0.0853	-7.4176	0.0205
6 High Life	27.8906	0.8211	-31.8175	0.0962	-33.4947	0.0604	-31.9567	0.0753	-28.4636	0.1556	-38.3692	0.1454
7 MGD	13.7654	0.4555	-14.3441	0.0453	-15.2991	0.2165	-14.7004	0.1246	-10.9267	0.0595	-13.7674	0.0005
8 Heineken	12.8958	0.4770	-15.4451	0.0813	-16.5956	0.1619	-15.6658	0.0110	-11.7094	0.0441	-13.2366	0.1351
9 Natural Light	17.5565	0.7841	-26.7804	0.0357	-25.3616	0.1147	-28.7162	0.0517	-25.2566	0.0362	-24.4727	0.2351
10 Busch	3.0769	0.1957	-3.7896	0.1091	-3.6402	0.2285	-5.1464	0.0886	-2.4132	0.0111	-2.4710	0.3810
11 Composite	**0.0000**	**(fixed)**	-2.6304	0.1612	-2.9498	0.0628	-3.1215	0.0348	-0.7768	0.1194	**0.0000**	**(fixed)**

Table 9. Parameter estimates and standard errors for marketing variables

	Local Ads	Network TV Ads
Mean	0.2198 (0.0669)	0.1758 (0.0798)

Table 10. The own- and cross-elasticities when brands move toward context 3 (Segment 2)

	Bud Light	Budweiser	Miller Lite	Coors Light	Corona Extra	High Life	MGD	Heineken	Natural Light	Busch	Composite
Bud Light	0.00043	-0.17432	-0.01249	-0.03668	-0.26098	-0.06584	-0.03293	-0.14798	-0.01905	-0.11035	-0.28948
Budweiser	-0.08341	0.08181	-0.01526	-0.03895	-0.14300	-0.07274	-0.00816	-0.15515	-0.10346	-0.24873	-0.29051
Miller Lite	-0.08427	-0.18852	0.00068	-0.04000	-0.28385	-0.07081	-0.05063	-0.16598	-0.02467	-0.12796	-0.33950
Coors Light	-0.10539	-0.18422	-0.01553	0.42225	-0.26209	-0.15956	-0.03412	-0.15231	-0.02415	-0.11198	-0.29015
Corona Extra	-0.08266	-0.17538	-0.01372	-0.03678	0.09853	-0.14181	-0.03715	-0.14991	-0.01907	-0.11073	-0.29512
High Life	-0.11144	-0.30135	-0.03166	-0.07489	-0.29062	0.41519	-0.08732	-0.24296	-0.02052	-0.12026	-0.29623
MGD	-0.08321	-0.19923	-0.02819	-0.03865	-0.32770	-0.15754	0.05365	-0.19442	-0.01918	-0.11571	-0.32858
Heineken	-0.08283	-0.08208	-0.01443	-0.03729	-0.26437	-0.06876	-0.03804	0.06306	-0.01910	-0.11117	-0.29063
Natural Light	-0.10159	-0.17613	-0.02263	-0.00646	-0.26132	-0.06665	-0.03310	-0.14881	0.13781	-0.11446	-0.29091
Busch	-0.08265	-0.07811	-0.01302	-0.03673	-0.26115	-0.14135	-0.03307	-0.14819	-0.03624	0.19409	-0.28972
Composite	-0.08263	-0.07792	-0.01249	-0.03668	-0.26098	-0.06584	-0.03293	-0.14798	-0.01905	-0.11035	0.10750

Table 11. Competitive clout and vulnerability of brands near context 3 (Segment 2)

	Bud Light	Miller Lite	Corona Extra	MGD	Heineken	Busch
$Competitive\ Clout_{s=1,j}^{\pi=2}$	0.22365	0.29218	0.18282	0.34063	0.18845	0.21093
$Vulnerability_{s=1,j'}^{\pi=2}$	0.08218	0.00366	0.70415	0.01860	0.28202	0.18070

Figure 5. Optimal direction for Budweiser (Brand 2) in Segment 2

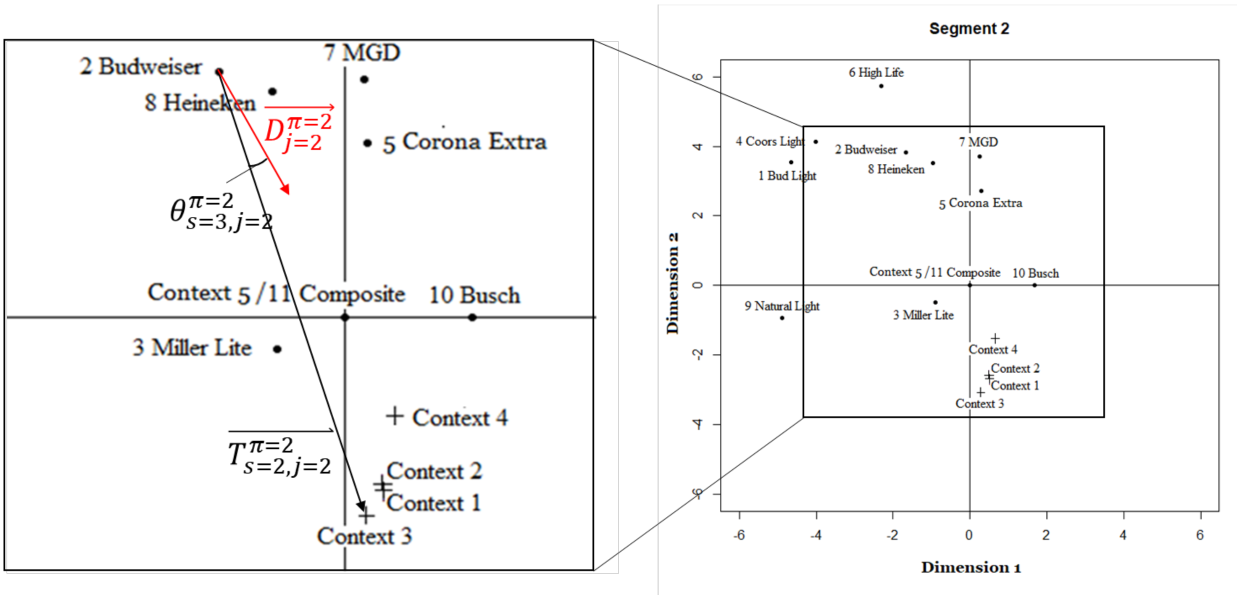
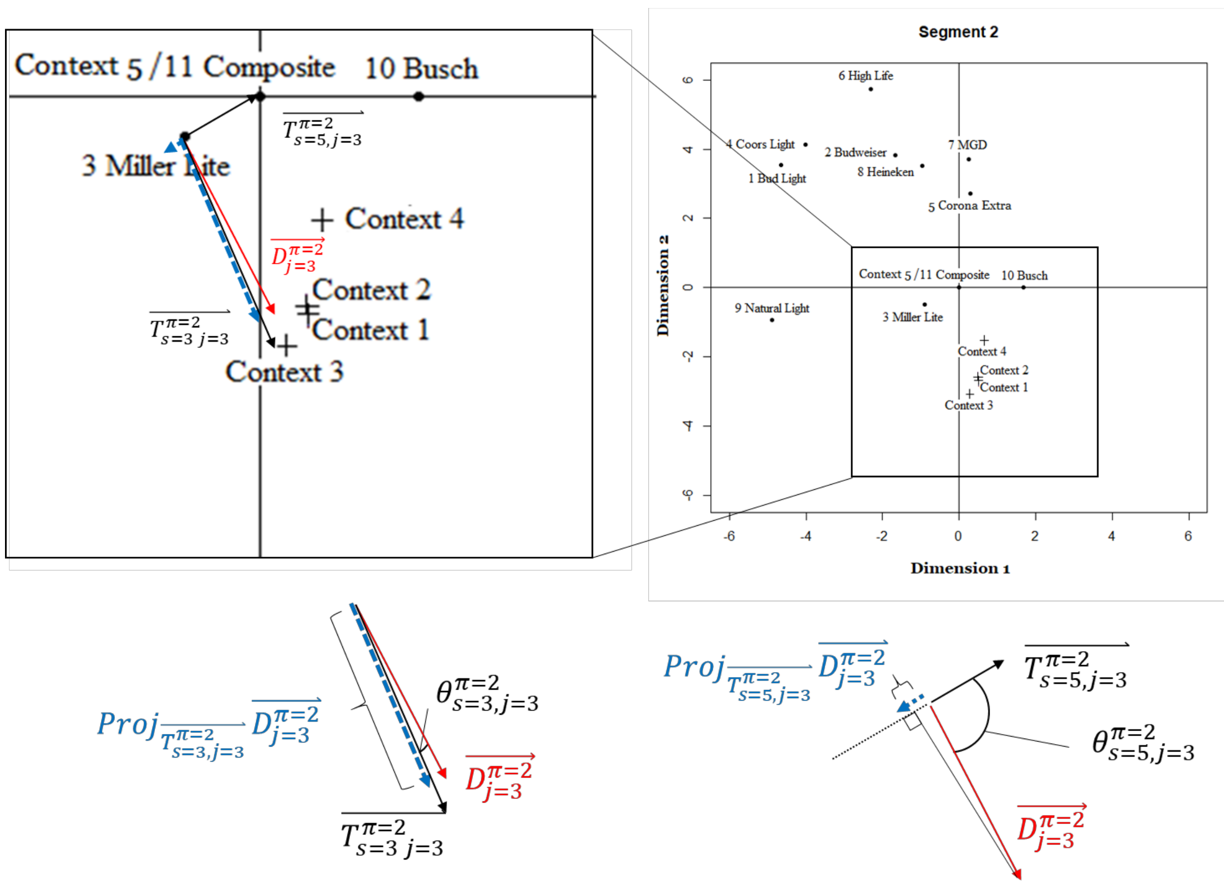


Figure 6. Orthogonal projection and the cosines of the angles between $\overrightarrow{T_{s,J}^{\pi}}$ and $\overrightarrow{D_J^{\pi}}$



Orthogonal projection of Miller Lite for context 3

Orthogonal projection of Miller Lite for context 5

Figure 7. Optimal directions for Miller Lite (Brand 3) in Segment 2

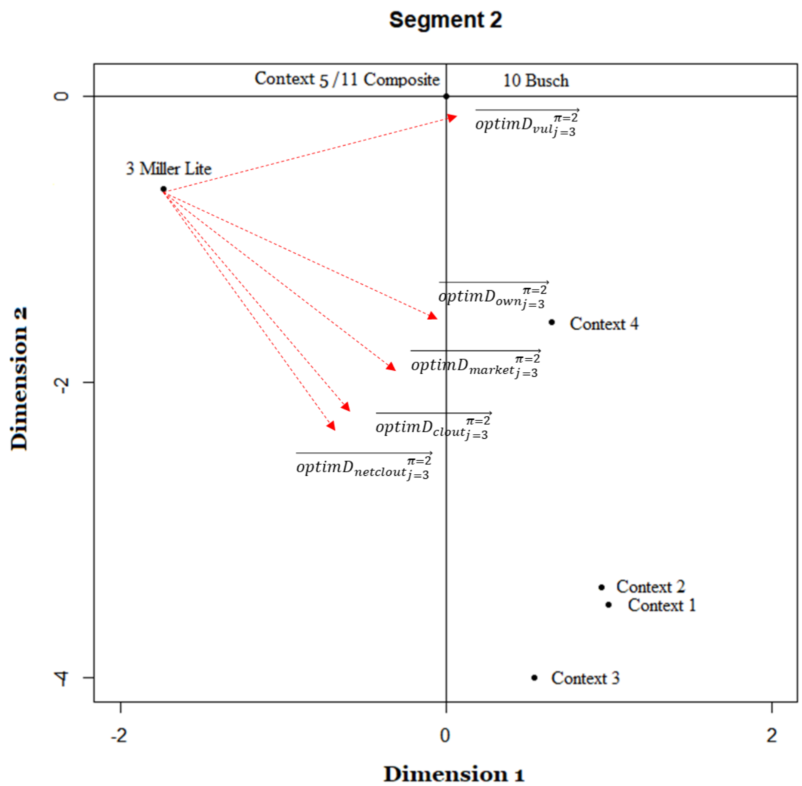


Table 12. Optimal directions and changes in market share:
Miller Lite's (Brand 3) repositioning for Segment 2

		Context 1 (relaxing)	Context 2 (with a meal)	Context 3 (working at home)	Context 4 (dancing)	Context 5 (local pub)
Current Market Shares of Miller Lite		0.33809	0.33027	0.33237	0.32986	0.34063
f_s^{π}		0.63242	0.13152	0.08981	0.05202	0.09423
Alternative Objective Functions						
The Largest Market	$optim\theta^{\pi=2}_{s,j=3}$ degree (°)	0.00000°	0.79361°	8.95201°	25.19539°	85.03251°
	$cos(optim\theta^{\pi=2}_{market,s,j=3})$	1.00000	0.99990	0.98782	0.90486	0.08659
	$ChangeMarket^{\pi=2}_{market,s,j=3}$	6.37.E-05	2.92.E-05	9.75.E-06	4.13.E-06	5.55.E-07
$\sum_{s=1}^{S} f_s^{\pi=2} \times (OwnElasticity^{\pi=2}_{s,j=2})$	$optim\theta^{\pi=2}_{s,j=3}$ degree (°)	17.52604°	18.31965°	26.47805°	7.66935°	67.50647°
	$cos(optim\theta^{\pi=2}_{own,s,j=3})$	0.95358	0.94932	0.89511	0.99105	0.38258
	$ChangeMarket^{\pi=2}_{own,s,j=3}$	3.53.E-04	2.85.E-04	2.88.E-04	4.65.E-05	1.42.E-04
$\sum_{s=1}^{S} f_s^{\pi=2} \times (Competitive\ Clout^{\pi=2}_{s,j'=2})$	$optim\theta^{\pi=2}_{s,j=3}$ degree (°)	5.12302°	4.32942°	3.82898°	30.31841°	90.15554°
	$cos(optim\theta^{\pi=2}_{clout,s,j=3})$	0.99601	0.99715	0.99777	0.86323	-0.00271
	$ChangeMarket^{\pi=2}_{clout,s,j=3}$	2.91.E-04	3.73.E-05	6.70.E-04	1.22.E-05	-2.36.E-08
$\sum_{s=1}^{S} f_s^{\pi=2} \times (-Vulnerability^{\pi=2}_{s,j'=2})$	$optim\theta^{\pi=2}_{s,j=3}$ degree (°)	79.08530°	79.87891°	88.03731°	53.88991°	5.94721°
	$cos(optim\theta^{\pi=2}_{vul,s,j=3})$	0.18935	0.17573	0.03425	0.58934	0.99462
	$ChangeMarket^{\pi=2}_{vul,s,j=3}$	1.21.E-05	5.13.E-06	3.38.E-07	2.69.E-06	6.38.E-06
$\sum_{s=1}^{S} f_s^{\pi=2} \times (Competitive\ Clout^{\pi=2}_{s,j'=2} - Vulnerability^{\pi=2}_{s,j'=2})$	$optim\theta^{\pi=2}_{s,j=3}$ degree (°)	6.13374°	5.34013°	2.81827°	31.32913°	91.16625°
	$cos(optim\theta^{\pi=2}_{netclout,s,j=3})$	0.99428	0.99566	0.99879	0.85419	-0.02035
	$ChangeMarket^{\pi=2}_{netclout,s,j=3}$	2.28.E-04	8.14.E-06	6.61.E-04	8.22.E-06	-4.66.E-08

Table 13. Characteristics of 10 brands

	Price	Calories	Alcohol (ABV)	Color (SRM)	Bitterness (IBU)
1 Bud Light	0.92	1.45	4.20	3	10
2 Budweiser	0.92	0.96	5.00	2	11
3 Miller Lite	0.92	1.02	4.20	3	12
4 Coors Light	0.92	1.27	4.20	3	12
5 Corona Extra	1.52	1.43	4.50	4	10
6 High Life	1.44	1.52	5.00	4	10
7 MGD	1.30	1.50	4.66	4	20
8 Heineken	0.94	0.95	5.00	3	23
9 Natural Light	0.82	1.33	4.75	3	5
10 Busch	0.85	1.27	5.07	3	10

Table 14. Regression results on dimension 2 for Segment 2

Dimension 2	Coeff.	Std. Error	t value
(Intercept)	-6.465	3.968	-1.629
Price	9.920**	3.791	2.617
Calories	9.454	4.788	1.974
Color	-5.378*	2.205	-2.439
Bitterness	0.307*	0.131	2.335

*p<0.10, **p<0.05, and ***p<0.01

Figure 8. Images and product descriptions shown to MTurk workers

Table 15. Comparison of model fit and predictive ability

	at-home/out-of-home	CDMDU
Number of consumers in-sample	725	725
Number of choice occasions in-sample	7,641	7,641
Number of parameters	59	83
(in-sample) -Log-likelihood	10,054.50	7,310.35
(in-sample) AIC	20,227.00	14,786.70
(in-sample) BIC	21,164.07	16,104.95
Number of consumers out-of-sample	725	725
Number of choice occasions out-of-sample	7,974	7,974
(Predictive) -Log-likelihood	11,983.07	7,836.92
(Predictive) AIC	24,084.14	15,839.84
(Predictive) BIC	25,026.25	17,165.17

Note: 2 dimensions and 2 segments

WEB APPENDIX A: Mathematical comparison of models from existing literature with proposed model

Here we describe the evolution of the literature in more mathematical terms. U_{ijt} denotes the indirect utility for individual i for brand j on consumption occasion t, and is usually decomposed as follows: $U_{ijt} = \alpha_{ij} + X_{ijt}\beta + \varepsilon_{ijt}$, where the first term, α_{ij} on the right-hand side, represents the preference for the brand; the second term represents the impact (β) of marketing variables (X_{ijt}); and the third term, ε_{ijt}, represents the component of utility unobserved by the researcher but known to the consumer. Our methodological focus is on the preferences term α_{ij} and the associated parameter explosion (a) when we allow these preferences to vary across consumers following a multivariate normal distribution, $\alpha_i \sim MVN(\alpha, \Sigma)$; and (b) in situations with a large number of brands and where consumers make choices in multiple contexts. Using a factor structure approach achieves parsimony by imposing an a priori factor structure on the covariance matrix of consumer brand preferences (e.g., Chintagunta 1994; Elrod 1988; Elrod and Keane 1995) (see Table A-1).

Let α_i denote the $J \times 1$ vector of (unobserved) brand preference for the J brands for household i. Elrod (1988) decomposed these preferences into a household-specific importance weight (w_i) and household-invariant brand location matrix (A). Thus, the vector $\alpha_i = Aw_i$ where A is a $J \times K$ matrix of positions of the J brands on a K-dimensional map, and w_i is an $K \times 1$ vector of household i–specific importance weights for K dimensions with $w_i \sim MVN(w, I_K)$. w denotes the unknown K-element vector, and I_K denotes the $K \times K$ identity covariance matrix. When $K << J$, the number of parameters under the factor structure will be much lower than under the full covariance model. Similarly, Chintagunta (1994) used a similar factor structure to that described above but with w_i following a discrete distribution with Π supports or market segments. Note that to ensure identification, we impose the restriction of a common matrix of brand locations (A) across consumers with consumer-specific importance weights (e.g., Chintagunta 1994; Elrod 1988; Elrod and Keane 1995).

Table A-1. Literature review on the factor structure

		Contribution
Elrod 1988	$\alpha_i = Aw_i$ $w_i \sim MVN(w, I_K)$	This model considers pooling heterogeneous consumers only for a parametric estimation. But there is no further discussion on the outcome related to consumers' heterogeneity.
Chintagunta 1994	$\alpha_i = Aw_{i\pi}$ $w_{i\pi} = w_\pi$ $i \in \Lambda_\pi, \pi = 1, 2, \cdots, \Pi$	This model uses a discrete distribution to account for heterogeneity in intrinsic preferences and imposes a factor structure to recover brand locations in multi-attribute space.
Kim and Chintagunta 2012	$\alpha_{is} = A_{i\pi} w_{is}$ $w_{is} \sim MVN(w_s, \sigma_s^2 I_K)$ $A_{i\pi} = A_\pi$ $i \in \Lambda_\pi, \pi = 1, 2, \cdots, \Pi$	This model suggests a hybrid discrete-continuous heterogeneity distribution to account for preference heterogeneity across brands/contexts.
The CDMDU model (this paper)	$dispreference\ _{ijs} = -\left(z_{i\pi j} - IP_{i\pi,z}^s\right)^2 w_{is}$ $w_{is} \sim MVN(w_s\ , \sigma_s^2 I_K)$ $z_{i\pi j} = z_{\pi j},\ \ IP_{i\pi,z}^s = IP_{\pi,z}^s$ $i \in \Lambda_\pi, \pi = 1, 2, \cdots, \Pi$	The CDMDU model can generate both consumers' context-specific ideal points and brand locations on the same map.

However, most of the model specifications developed in this literature typically involve a single consumption context. Acknowledging the importance of accounting for consumer preference variation across consumption contexts, Kim and Chintagunta (2012) proposed a model of hybrid discrete-continuous heterogeneity that parsimoniously accounts for preference heterogeneity across brands, consumers, and consumption contexts.

More specifically, the distribution of consumer brand preferences for the J brands in a given context s can be decomposed into two components (Kim and Chintagunta 2012): context-specific and context-invariant components, such that $\alpha_{is} = A_{i\pi} w_{is}$, where $A_{i\pi}$ is a $J \times K$ matrix of positions of the J brands on the K-dimensional map that vary across individuals i following a discrete distribution with Π supports or market segments, and w_{is} is an $K \times 1$ vector of consumer i–specific importance weights for K dimensions corresponding to the sth consumption context with $w_{is} \sim MVN(w_s, \sigma_s^2 I_K)$. This approach captures heterogeneity in preferences across scenarios by allowing the brand map itself to have a discrete heterogeneity distribution across consumers. Using data on consumption occasion–specific context information over time, Kim and Chintagunta (2012) selected a "base" consumption context and then decomposed the deviations in consumer preferences from that base context into a low-dimensional brand map where the dimensions' weights vary by context and the brand locations are fixed across contexts but vary across consumer segments.

WEB APPENDIX B. Logarithmic transformation

If w_{is}^{k} is a random variable whose logarithm is normally distributed, then w_{is}^{k} has a log-normal distribution. The probability density function of w_{is}^{k} can be obtained by a straightforward transformation of the normal probability density function, yielding

$$f\left(w_{is}^{k} | \tau_{s}^{k}, (\sigma_{s}^{k})^2\right) = \frac{1}{\sqrt{2\pi}\sigma_{s}^{k}} \frac{1}{w_{is}^{k}} e^{-\left(ln\left(w_{is}^{k}\right)-\tau_{s}^{k}\right)^2 \Big/ (2(\sigma_{s}^{k})^2)} \qquad (B-1)$$
$$0 < w_{is}^{k} < \infty, -\infty < \tau_{s}^{k} < \infty, \sigma_{s}^{k} > 0 .$$

The mean and standard deviation of w_{is}^{k} can be calculated as follows:

$$E\left(w_{is}^{k}\right) = E\left(e^{ln\left(w_{is}^{k}\right)}\right) = E\left(e^{\tau_{is}^{k}}\right) = e^{\tau_{s}^{k}+\left(\frac{(\sigma_{s}^{k})^2}{2}\right)} \qquad (B-2)$$
$$\text{where} \quad \tau_{is}^{k} \equiv ln\left(w_{is}^{k}\right) \sim N(\tau_{s}^{k}, (\sigma_{s}^{k})^2)$$
$$sd\left(w_{is}^{k}\right) = \sqrt{var\left(w_{is}^{k}\right)}$$
$$\text{where} \quad var\left(w_{is}^{k}\right) = e^{2(\tau_{s}^{k}+(\sigma_{s}^{k})^2)} - e^{2\tau_{s}^{k}+(\sigma_{s}^{k})^2} .$$

Thus, when we assume $dispreference_{ijs} = -\left(z_{i\pi j} - IP_{i\pi,z}^{s}\right)^2 w_{is}$, the mean and standard deviation of dispreference for a brand j within a segment π $(i \in \Lambda_{\pi})$ given context s can be written as follows:

$$E\left(dispreference_{ijs}\right) = E\left(\sum_{k=1}^{K} -\left(z_{i\pi j}^{k} - IP_{i\pi,z}^{s,k}\right)^2 w_{is}^{k}\right) \qquad (\mathrm{B}-3)$$
$$= \sum_{k=1}^{K} -\left(z_{i\pi j}^{k} - IP_{i\pi,z}^{s,k}\right)^2 E(w_{is}^{k})$$
$$= \sum_{k=1}^{K} -\left(z_{i\pi j}^{k} - IP_{i\pi,z}^{s,k}\right)^2 \times e^{\tau_{s}^{k}+\left(\frac{(\sigma_{s}^{k})^2}{2}\right)}$$
$$sd\left(dispreference_{\ ijs}\right) = \sqrt{var\left(\sum_{k=1}^{K} -\left(z_{i\pi j}^{k} - IP_{i\pi,z}^{s,k}\right)^2 w_{is}^{k}\right)}$$

$$= \sqrt{\sum_{k=1}^{K} \left(z_{i\pi j}^{k} - IP_{i\pi,z}^{s,k}\right)^{4} \times var\left(w_{is}^{k}\right)}$$

$$= \sqrt{\sum_{k=1}^{K} \left(z_{i\pi j}^{k} - IP_{i\pi,z}^{s,k}\right)^{4} \times \left(e^{2(\tau_s^k + (\sigma_s^k)^2)} - e^{2\tau_s^k + (\sigma_s^k)^2}\right)}$$

where $\tau_{is} = (\tau_{is}^1, \tau_{is}^2, \dots, \tau_{is}^K)$ is assumed to have a multivariate normal distribution across individuals with the context-specific mean ($\tau_s = (\tau_s^1, \tau_s^2, \dots, \tau_s^K)$) and a covariance matrix

$$\begin{bmatrix} (\sigma_s^1)^2 & 0 & 0 & 0 \\ 0 & (\sigma_s^2)^2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & (\sigma_s^K)^2 \end{bmatrix}.$$

The matrix above is the K-dimensional identity matrix scaled by a context-specific variance term $(\sigma_s^k)^2$.

WEB APPENDIX C: Model identification

We lay out (a) the number of parameters that one would have to estimate using a traditional random coefficients approach for each context and (b) the specific restrictions and number of parameters that our proposed approach would entail. Below, we focus mainly on the distribution of brand preferences, since it is this heterogeneity across brands, consumers, and contexts that results in an explosion of parameters.

(a) The traditional random coefficients (RC) model

Consider a dataset with J brands and S contexts where we treat the choices of consumers across the S contexts as being independent. Using a multivariate normal distribution for preference in each context, we would need to estimate $(J-1)$ mean parameters (since one of the brands is normalized to have a preference of 0) and $J \times (J-1)/2$ parameters for the covariance matrix (say Σ). So, the total number of parameters across the S contexts is $S \times (J-1) \times (J+2)/2$ parameters. In other words, with 10 brands and five contexts, we would need to estimate $5 \times 9 \times \frac{12}{2} = 270$ parameters for the distribution of preferences. As is evident, this is a computationally challenging task.

(b) The proposed CDMDU model

First, we consider the case where we have data on a single context. One way in which previous researchers have proposed a more parsimonious version of the random coefficients model is to project the covariance matrix for the traditional RC model above onto lower-dimensional space using a factor structure. In other words, S can be written as $\Sigma = Aww'A'$, where A is a $(J-1) \times K$ matrix of brand "locations" on a space with a lower dimension than J (i.e., $\mathrm{K} \ll J$) and w is a vector of weights that each consumer has for the K dimensions. Heterogeneity in ω is then accomplished by assuming that it has a K-dimensional multivariate normal distribution with mean zero (to ensure identification) and identity matrix of dimension K for the covariance matrix. Note that the covariance matrix in this case has $(J-1)\times K$ parameters. Such a factor model would be very similar to our model for context 5. As shown in Equation 6 of the paper, the utility for context 5 is written as:

$$U_{ijt}=\gamma_j^{*}-\left(z_{ij}^{1}\right)^2 w_{is}^{1}-\left(z_{ij}^{2}\right)^2 w_{is}^{2}-\cdots-\left(z_{ij}^{K}\right)^2 w_{is}^{K}+\epsilon_{ijt} \qquad (C-1)$$

where z_{ij}^{k} is akin to the location of brand *j* along dimension *k* (for consumer *i*) and w_{is}^{k} is the associated weight corresponding to the above factor model and *s* = 5 (for context 5). Identification in this model can be achieved by requiring the number of estimated parameters to be fewer than those corresponding to the RC model. In other words, as long as $(J-1)\times K<(J-1)\times J/2$ or $\mathrm{K}<J/2$.

With *S* contexts, one can imagine creating a factor model for each of the contexts. This would entail having to estimate $S\times(J-1)\times K$ parameters. However, we formulate the model in an even more parsimonious fashion, as follows:

Context 1: $U_{ijt}=\gamma_j^{*}-\left(z_{ij}^{1}-IP_{i,z}^{s=1,1}\right)^2 w_{is}^{1}-\left(z_{ij}^{2}-IP_{i,z}^{s=1,2}\right)^2 w_{is}^{2}-\cdots-\left(z_{ij}^{K}-IP_{i,z}^{s=1,K}\right)^2 w_{is}^{K}+\epsilon_{ijt}$

Context 2: $U_{ijt}=\gamma_j^{*}-\left(z_{ij}^{1}-IP_{i,z}^{s=2,1}\right)^2 w_{is}^{1}-\left(z_{ij}^{2}-IP_{i,z}^{s=2,2}\right)^2 w_{is}^{2}-\cdots-\left(z_{ij}^{K}-IP_{i,z}^{s=2,K}\right)^2 w_{is}^{K}+\epsilon_{ijt}$

Context 3: $U_{ijt}=\gamma_j^{*}-\left(z_{ij}^{1}-IP_{i,z}^{s=3,1}\right)^2 w_{is}^{1}-\left(z_{ij}^{2}-IP_{i,z}^{s=3,2}\right)^2 w_{is}^{2}-\cdots-\left(z_{ij}^{K}-IP_{i,z}^{s=3,K}\right)^2 w_{is}^{K}+\epsilon_{ijt}$

Context 4: $U_{ijt}=\gamma_j^{*}-\left(z_{ij}^{1}-IP_{i,z}^{s=4,1}\right)^2 w_{is}^{1}-\left(z_{ij}^{2}-IP_{i,z}^{s=4,2}\right)^2 w_{is}^{2}-\cdots-\left(z_{ij}^{K}-IP_{i,z}^{s=4,K}\right)^2 w_{is}^{K}+\epsilon_{ijt}$

Context 5: $U_{ijt}=\gamma_j^{*}-\left(z_{ij}^{1}\right)^2 w_{is}^{1}-\left(z_{ij}^{2}\right)^2 w_{is}^{2}-\cdots-\left(z_{ij}^{K}\right)^2 w_{is}^{K}+\epsilon_{ijt} \qquad (C-2)$

In other words, we hold the locations of the brands to be common across all five contexts. However, we introduce an ideal point for each context and dimension. For example, $IP_{i,z}^{s=1,1}$ denotes the ideal point for context *s* = 1 along dimension 1 so the number of parameters scales only as the number of contexts and dimensions and not as the number of contexts and brands, as would have been the case with a natural extension of the factor model to contexts (to recall, the number of dimensions is smaller than number of brands). In this way, we estimate

$(J-1) \times K$ location factors and $(S-1) \times K$ ideal point parameters. At the same time, we are able to relax some of the restrictions on the parameters of the w_{is}^{k} distributions (their means and variances). Finally, since we are still a long way from exhausting the $S \times (J-1) \times J/2$ parameters for the covariance matrices, we allow the location parameters, z_{ij}^{k}, to themselves vary across consumers following a discrete distribution with a small, finite number of supports. By keeping the total number of parameters estimated for the preference parameters to be under $S \times (J-1) \times (J+2)/2$, we ensure estimability of the model parameters.

To understand how the parameters are recoverable, consider the case where we have only one context. In that case, our model would reduce to the one corresponding to the case in which only context 5 is present like Equation C-1. Before examining the recoverability of the parameters, we note that accounting for heterogeneity using a standard random coefficients model with J brands and one context would entail estimating $(J-1)$ means and $J \times (J-1)/2$ covariance parameters. By contrast, for a two-dimensional map, our model above has $(J-1)$ parameters (for γ_j^*), $(J-1) \times 2$ parameters (for z_{ij}^{k}), and two variance parameters for w_{is}^{k}. Recall from the paper that for this context, the means and variances of the weighting parameters, w_{is}^{k}, are set to 1. Thus, the parameters of our proposed model with one context would be identified as long as $J \times (J-1)/2 > 2 \times (J-1)$. Such a model is very similar to models in previous studies by Elrod (1988), Chintagunta (1994), etc. Those studies also established the estimability of the model parameters. As we add more contexts, we note that in principle we can estimate an additional $J-1$ means and $J \times (J-1)/2$ covariance parameters for each context, but as is clear from the specification, we are adding far fewer than those parameters to our proposed specification.

WEB APPENDIX D. Parameter estimates and standard errors (without imposing a non-negative restriction on w_{is})

	Segment 1 (35.35%)				**Segment 2** (64.65%)			
	Dimension 1		*Dimension 2*		*Dimension 1*		*Dimension 2*	
	estimates	*s.e.*	*estimates*	*s.e.*	*estimates*	*s.e.*	*estimates*	*s.e.*
Brand location (z_{ij}^{k})								
Bud Light	1.7686	0.0351	-2.0359	0.0514	2.3756	0.0514	-0.5115	0.0639
Budweiser	2.3360	0.0372	-2.3482	0.0472	-3.5978	0.0491	-0.4817	0.0673
Miller Lite	1.1320	0.0321	-1.4114	0.0475	0.0219	0.0380	1.8766	0.0690
Coors Light	2.0553	0.0476	-2.2015	0.0677	-1.8844	0.0393	3.5587	0.0889
Corona Extra	1.5198	0.0408	-2.6282	0.0562	-4.9397	0.0895	3.3605	0.1090
High Life	1.1646	0.0364	0.2102	0.0577	0.8698	0.0822	1.9614	0.0746
MGD	-0.0647	0.0652	-1.7466	0.0540	-1.3230	0.0594	2.0708	0.0820
Heineken	1.2993	0.0434	-3.4559	0.0689	-2.3429	0.0770	1.7964	0.1152
Natural Light	3.1736	0.0513	0.6735	0.0468	-2.7737	0.0561	-3.0821	0.2169
Busch	3.7551	0.0525	**0.0000**	**(fixed)**	-2.9550	0.0905	**0.0000**	**(fixed)**
Composite	**0.0000**	**(fixed)**	**0.0000**	**(fixed)**	**0.0000**	**(fixed)**	**0.0000**	**(fixed)**
Context-dependent ideal point ($IP_{i,z}^{s,k}$)								
Context 1	-0.0896	0.0279	2.1946	0.1020	0.0551	0.0235	-2.1220	0.0741
Context 2	-0.1672	0.0464	2.0514	0.0872	0.0337	0.0334	-1.7280	0.0951
Context 3	-0.1350	0.0431	2.5837	0.1756	0.0378	0.0318	-2.1247	0.0951
Context 4	-0.0962	0.0466	1.9580	0.1506	0.1808	0.0339	-1.4000	0.0971
Context 5	**0.0000**	**(fixed)**	**0.0000**	**(fixed)**	**0.0000**	**(fixed)**	**0.0000**	**(fixed)**

* -LL = 10,480.35, AIC = 21,126.71 and BIC = 22,563.60 where the number of observations is 15,615 and the number of parameters is 83.

	Dimension 1		*Dimension 2*		*Dimension 1*		*Dimension 2*	
	Mean estimate w_s^1	*s.e.*	*Mean estimate* w_s^2	*s.e.*	*Square root of variance estimate*	*s.e.*	*Square root of variance estimate*	*s.e.*
Context 1 (Relaxing)	0.9631	0.0233	0.2847	0.0099	1.3540	0.0394	0.4799	0.0174
Context 2 (With a meal)	0.9978	0.0310	0.3414	0.0152	1.4052	0.0514	0.5342	0.0252
Context 3 (Working at home)	1.0141	0.0308	0.2871	0.0156	1.3914	0.0522	0.4445	0.0272
Context 4 (Dancing)	0.9215	0.0300	0.2148	0.0145	1.3093	0.0493	0.3507	0.0208
Context 5 (Local pub)	**1.0000**	**(fixed)**	**1.0000**	**(fixed)**	**1.0000**	**(fixed)**	**1.0000**	**(fixed)**

** We estimated the model leaving w_{is} to be unconstrained (i.e., without imposing a non-negative restriction on w_{is}). This allows us to assess whether the data reveal households for whom $w_{is} < 0$. w_{is} is the $K \times 1$ vector of consumer i-specific importance weights for K dimensions corresponding to the sth consumption context with $w_{is}^{k} \sim N(w_s^{k}, (\sigma_s^{k})^2)$.

WEB APPENDIX E: Interpreting iso-preference contours in a joint space map

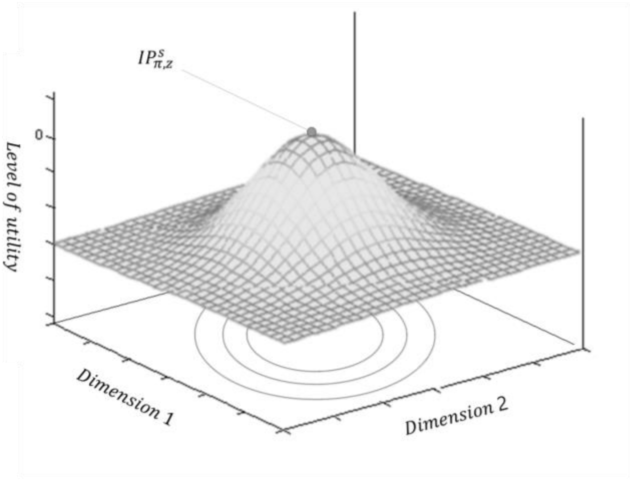
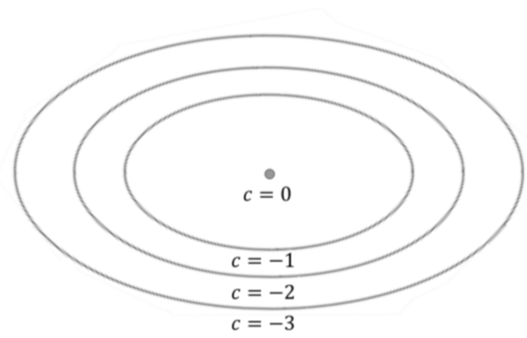
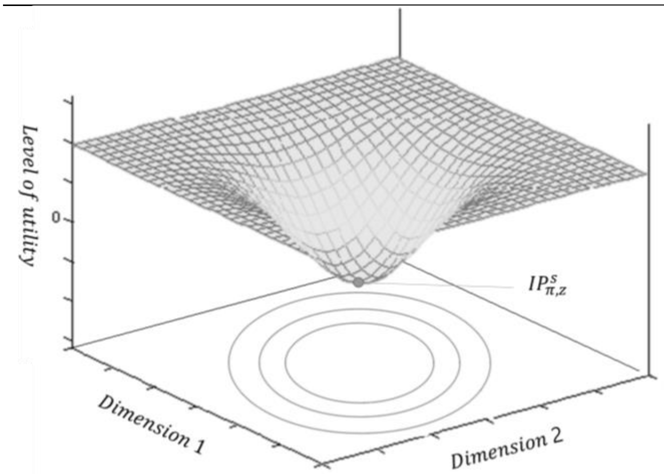
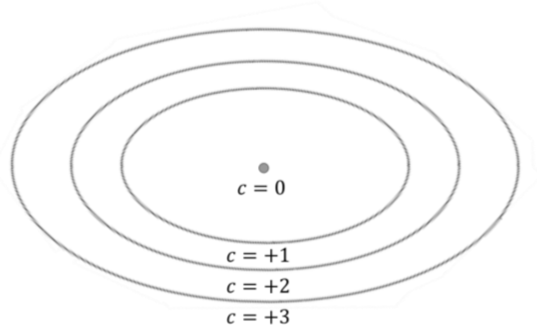
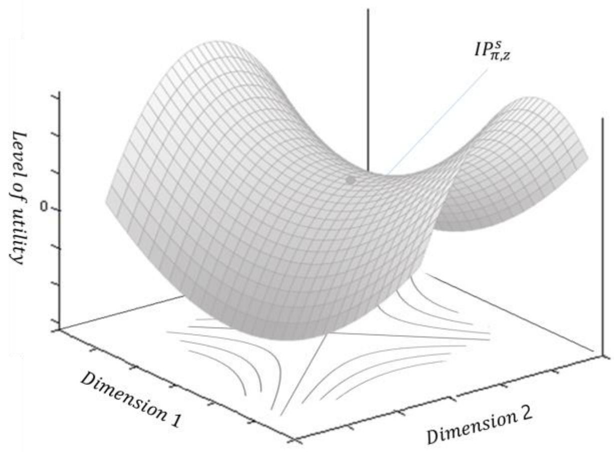
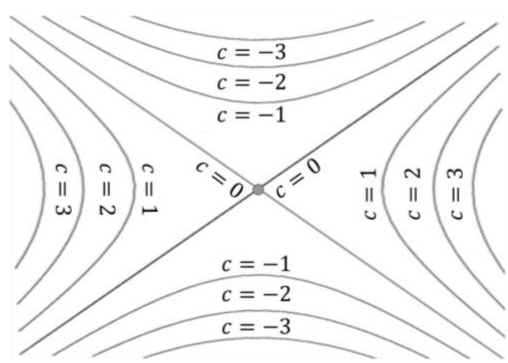
Our empirical results in Section 4 focused on ideal point models with increasing dispreference as we moved away from the ideal point. In this appendix, we consider three situations under which researchers may need to create iso-preference contours and illustrate their construction when the number of dimensions or attributes K is 2. First, note that we can rewrite the dispreference term in Equation 2 as a quadratic form as follows: $\left(\boldsymbol{X} - IP_{\pi,z}^{s}\right)^{T} W_s\left(\boldsymbol{X} - IP_{\pi,z}^{s}\right)$ in K dimensions for segment π. We can define iso-preference contour $\mathcal{E}_\mathrm{c}$ with size (radius) c as the set of all points $\boldsymbol{X}$, with $-\left(\boldsymbol{X} - IP_{\pi,z}^{s}\right)^{T} W_s\left(\boldsymbol{X} - IP_{\pi,z}^{s}\right)$ equal to c for context s, which has context-dependent ideal point $IP_{\pi,z}^{s}$ as a center (or critical point) as follows:

$$\mathcal{E}_\mathrm{c}\left(IP_{\pi,z}^{s}, W_s\right) \coloneqq \left\{\boldsymbol{X}: -\left(\boldsymbol{X} - IP_{\pi,z}^{s}\right)^{T} W_s\left(\boldsymbol{X} - IP_{\pi,z}^{s}\right) = c\right\}, \qquad \text{(E-1)}$$

where W_s is the $K \times K$ diagonal matrix with the mean of w_{is}^{k} (w_s^k) as a diagonal entry for each context s. Next we describe the three situations of interest.

1) <u>When w_s^k is positive for all of k</u>: In this case, the utility surface for the dispreference has negative values and its iso-preference contours are denoted by elliptic paraboloids (see the first row in Figure E-1). The quadratic form of the dispreference term shows a local maximum at context-dependent ideal point $IP_{\pi,z}^{s}$ when $c = 0$.

Figure E-1. Utility surface for the dispreference and iso-preference contours

Utility surface for the dispreference	Iso-preference contours (When $K = 2$)
	$w_s^{k=1} > 0, w_s^{k=2} > 0$ 
	$w_s^{k=1} < 0, w_s^{k=2} < 0$ (anti-ideal point) 
	$w_s^{k=1} < 0, w_s^{k=2} > 0$ (saddle point) 

2) When w_s^k is negative for all of k: As noted above, for the ideal point interpretation of the model, w_{is} should be non-negative for all i and s. However, the CDMDU model can accommodate negative ideal points[1] as well. Note that the negative ideal point makes sense when considering how choices today can influence choices tomorrow owing to attribute satiation (Erdem 1996). The idea here is the following: Suppose I purchase brand A, which is close to the ideal point, this period. When it comes to making a purchase next period, if I feel satiated with the attributes of brand A (e.g., Lattin and McAlister 1985; McAlister 1982), then I am more likely to purchase a product farther away from the ideal point the next period. In that sense, the ideal point is not an attractor but a repeller. However, that model is quite different from the one we are looking at here. If, for a specific context in two dimensions, both weights are negative, then the location of that context-dependent ideal point becomes an anti-ideal point and refers to the point in the joint space of least preference. There, preference increases as a brand moves farther away from the ideal point in either direction. As we can see in the second row in Figure E-1 (where w_s^k is negative for all of k), the utility surface has positive values and the iso-preference contours are also denoted by elliptic paraboloids. The quadratic form has a local *minimum* at anti-ideal point $IP_{\pi,z}^s$ when $c = 0$. Then the question arises as to how a brand should move to maximize consumer utility. In this case, the level of utility can go to infinity without a boundary condition on dimensions 1 and 2 (i.e., brand locations). In other words, the optimizing process with the anti-ideal points is not possible without defining the boundary (i.e., the range of each dimension) (DeSarbo and Rao 1986), which is another constraint researchers have to impose.

3) When the signs of the weights w_s^k are mixed (one is positive and the other is negative): In this case, the location of the ideal point is called a *saddle point*, where preferences increase in

[1] We estimated the model leaving w_{is} to be unconstrained (i.e., without imposing a non-negative restriction on w_{is}). See the result table in Web Appendix D.

one direction as a brand location goes farther away from the ideal point, even as preferences increase in the other dimension as it gets closer to the ideal point. In this case, as with the anti-ideal points, optimizing is not possible without a constraint on the boundary. For the same reason, many previous MDU studies have constrained the weights in the ideal point model to be positive as in case 1 above (Davison 1976; DeSarbo and Kim 2013; DeSarbo and Rao 1986; Srinivasan and Shocker 1973).

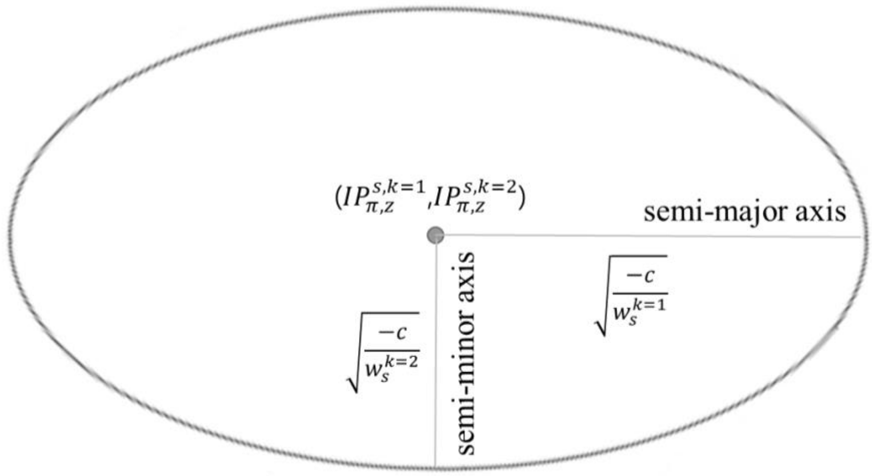
WEB APPENDIX F. Iso-preference contours $\mathcal{E}_c$ in a joint space map

With a non-negative constraint on w_s^k (i.e., w_s^k is positive for $k = 1, 2$), we can express the iso-preference contour ($\mathcal{E}_c$, $c \leq 0$) in quadratic form as an ellipsoid. In other words, the set of all points $\boldsymbol{X} = (X_1, X_2)$ with $-\left(\boldsymbol{X} - IP_{\pi,z}^s\right)^T W_s\left(\boldsymbol{X} - IP_{\pi,z}^s\right) = c$ in context s for Segment π) is given by the following equation:

$$\left(X_1 - IP_{\pi,z}^{s,k=1}\right)^2 w_s^{k=1} + \left(X_2 - IP_{\pi,z}^{s,k=2}\right)^2 w_s^{k=2} = -c \qquad (F-1)$$

$$\frac{\left(X_1 - IP_{\pi,z}^{s,k=1}\right)^2}{\left(\sqrt{\frac{-c}{w_s^{k=1}}}\right)^2} + \frac{\left(X_2 - IP_{\pi,z}^{s,k=2}\right)^2}{\left(\sqrt{\frac{-c}{w_s^{k=2}}}\right)^2} = 1.$$

Figure F-1. Iso-preference contour ($\mathcal{E}_c$) with a non-negative constraint on w_s^k

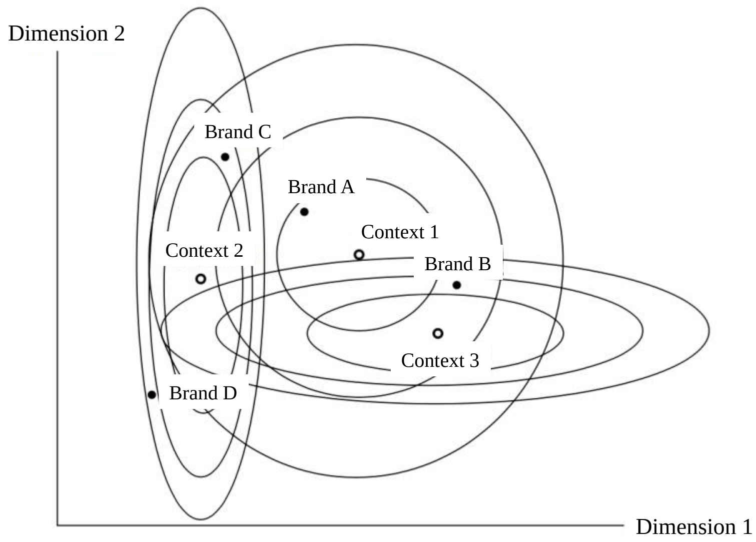


With the origin of the coordinate system at $(IP_{\pi,z}^{s,k=1}, IP_{\pi,z}^{s,k=2})$, if $w_s^{k=2} > w_s^{k=1}$, an ellipse curve in Figure F-1 is represented by $\sqrt{\frac{-c}{w_s^{k=1}}}$ and $\sqrt{\frac{-c}{w_s^{k=2}}}$, which are known as the semi-major axis and the semi-minor axis, respectively.

Hence, depending on the ratio of sensitivity (e.g.,$w_s^{k=1}$: $w_s^{k=2}$), the iso-preference contours in two dimensions can take the shape of either a vertical ellipse (i.e., the horizontal dimension is highly weighted) or a horizontal ellipse (i.e., the vertical dimension is highly weighted). In

Figure F-2, we provide a hypothetical example with three contexts and four brands in two dimensions.

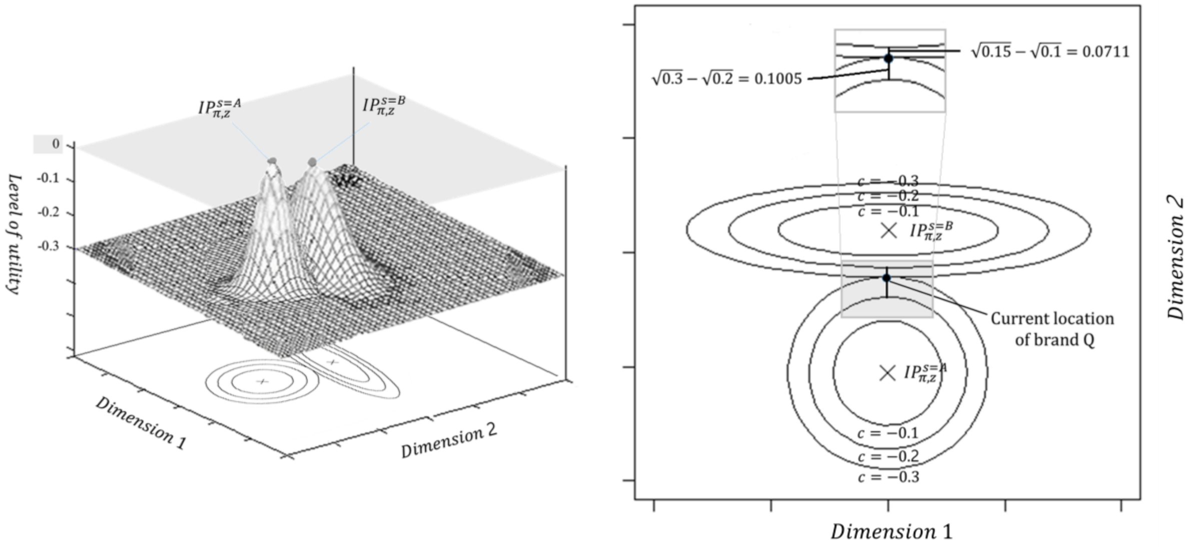
Figure F-2. Joint space map with different weights



In this figure, Context 1 equally weights the two dimensions, Context 2 weights Dimension 1 more than Dimension 2 (i.e., vertical ellipse), and Context 3 weights Dimension 2 more than Dimension 1 (i.e., horizontal ellipse).

To illustrate how the repositioning is associated with the context-specific ideal points and iso-preference contours, Figure F-3 presents two different iso-preference contours for the two context-dependent ideal points. Note that the level of consumer utility is 0 at both context-dependent ideal points $(IP_{\pi,z}^{s,k=1}, IP_{\pi,z}^{s,k=2})$, as seen in Figure F-3 (left side). For context A ($s = A$), the iso-preference contour is a circle around the consumer's ideal points $(IP_{\pi,z}^{s=A})$ with the same weights for both dimensions (i.e., $w_{s=A} = (1,1)$). The iso-preference contour for context B (s $= B$) is a horizontal ellipse with the dimension 2 being highly weighted (e.g., $w_{s=B} = (0.5, 2)$).

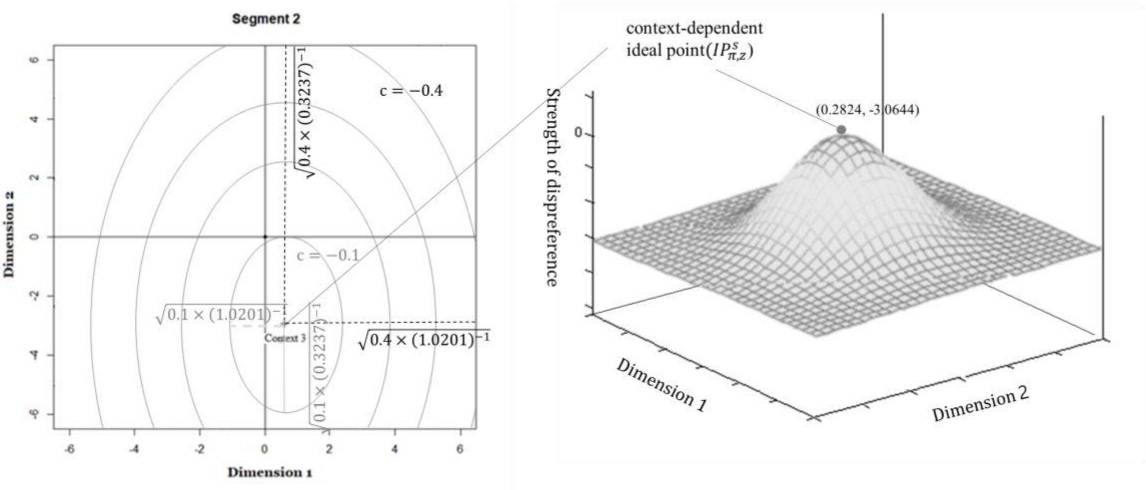
Figure F-3. Iso-preference contours for context A and context B (example)



$$\left(X_1 - IP_{\pi,z}^{s=A,k=1}\right)^2 w_{s=A}^{k=1} + \left(X_2 - IP_{\pi,z}^{s=A,k=2}\right)^2 w_{s=B}^{k=2} = -c$$
$$\left(X_1 - IP_{\pi,z}^{s=B,k=1}\right)^2 w_{s=B}^{k=1} + \left(X_2 - IP_{\pi,z}^{s=B,k=2}\right)^2 w_{s=B}^{k=2} = -\mathrm{c}$$
where $w_{s=A} = (1,1)\ and\ w_{s=B} = (0.5,2)$.

Though the utility difference between the two hypothetical iso-preference contours in Figure F-3 (right side) is 0.1 (= –0.1 – (–0.2)) in both contexts A and B, the Euclidean distance between the two iso-preference contours is much smaller in the ellipse for context B (= $\sqrt{0.15}-\sqrt{0.1}=0.0711$) than in the ellipse for context A (= $\sqrt{0.3}-\sqrt{0.2}=0.1005$). Now we can discuss a potential repositioning strategy of brand Q whose mean preference is assumed to be 0 (i.e., $\gamma_{\text{brand Q}}^{*}$ = 0). At its current location (i.e., black dot), the level of consumer utility for brand Q is the same for both contexts A and B, at –0.3, with the two iso-preference contours tangential to each other. If brand Q slightly changes its location toward $IP_{\pi,z}^{s=B}$ by 0.0711 along dimension 2, then the utility gain for context B is greater than the utility loss for context A. This example has implications because a firm's repositioning direction should be set to improve its overall profit by pursuing additional utility gain while minimizing the utility loss generated by such a movement.

Figure F-4. Brand *j*'s iso-preference contour for context 3 in Segment 2



Using the parameter estimates from our sample data, we can also illustrate the iso-preference contours with $IP_{\pi,z}^{s,k} \pm (\sqrt{-\mathrm{c} \times (w_s^k)^{-\mathbf{1}}})$ for context s in K dimensions. In Figure F-4, for example, we provide the iso-preference contours for context 3 in Segment 2. In Figure F-4 (left side), we overlay the iso-preference contours centered at the location of the ideal point for context 3 (0.2824, –3.0644), highlighting the cases where $c = -0.1$ and $\mathrm{c} = -0.4$ in a two-dimensional space.

WEB APPENDIX G. Elasticities

The own- (or cross-) elasticities along dimension 1 within a segment π given context s, denoted by $\eta^{\pi}_{s,F_1,j'j'}$ (or $\eta^{\pi}_{s,F_1,jj'}$), can be computed as follows:

$$\begin{aligned}
\eta^{\pi}_{s,F_1,j'j'} &= \frac{1}{TC_{\pi s}}\sum_{i\in\Lambda_\pi}\left\{\frac{1}{TS_{is}}\sum_{t=1}^{T_i}\left(\frac{\partial P_{ij't}}{\partial z^1_{ij'}}\right)\left(\frac{z^1_{ij'}}{P_{ij't}}\right)\times B_{is't}\{s'=s\}\right\}\\
&= \frac{1}{TC_{\pi s}}\sum_{i\in\Lambda_\pi}\left\{\frac{1}{TS_{is}}\sum_{t=1}^{T_i}\left(\frac{\partial V_{ij't}}{\partial z^1_{ij'}}\right)P_{ij't}\left(1-P_{ij't}\right)\times\left(\frac{z^1_{ij'}}{P_{ij't}}\right)\times B_{is't}\{s'=s\}\right\}\\
&= \frac{1}{TC_{\pi s}}\sum_{i\in\Lambda_\pi}\left\{\frac{1}{TS_{is}}\sum_{t=1}^{T_i}2w^1_{is}z^1_{ij'}\left(z^1_{ij'}-IP^{s,1}_{i,z}\right)\left(1-P_{ij't}\right)\times B_{is't}\{s'=s\}\right\}\\
&= 2z^1_{ij'}\left(z^1_{ij'}-IP^{s,1}_{\pi,z}\right)\times\frac{1}{TC_{\pi s}}\sum_{i\in\Lambda_\pi}\left\{\frac{1}{TS_{is}}\sum_{t=1}^{T_i}\left(w^1_{is}\left(1-P_{ij't}\right)\times B_{is't}\{s'=s\}\right)\right\} \qquad (G-1)\\
\eta^{\pi}_{s,F_1,jj'} &= \frac{1}{TC_{\pi s}}\sum_{i\in\Lambda_\pi}\left\{\frac{1}{TS_{is}}\sum_{t=1}^{T_i}\left(\frac{\partial P_{ijt}}{\partial z^1_{ij'}}\right)\left(\frac{z^1_{ij'}}{P_{ijt}}\right)\times B_{is't}\{s'=s\}\right\}\\
&= \frac{1}{TC_{\pi s}}\sum_{i\in\Lambda_\pi}\left\{\frac{1}{TS_{is}}\sum_{t=1}^{T_i}-\left(\frac{\partial V_{ij't}}{\partial z^1_{ij'}}\right)z^1_{ij'}P_{ijt}\times B_{is't}\{s'=s\}\right\}\\
&= -2z^1_{ij'}\left(z^1_{ij'}-IP^{s,1}_{\pi,z}\right)\times\frac{1}{TC_{\pi s}}\sum_{i\in\Lambda_\pi}\left\{\frac{1}{TS_{is}}\sum_{t=1}^{T_i}\left(w^1_{is}P_{ij't}\times B_{is't}\{s'=s\}\right)\right\}\\
&\qquad j'\neq j, i\in\Lambda_\pi
\end{aligned}$$

where T_i denotes the total number of consumption occasions for consumer *i*; TS_{is} denotes the number of consumption occasions associated with context s for consumer *i* $\left(\text{i.e.}, TS_{is}=\sum_{t=1}^{T_i}B_{is't}\{s'=s\}\right)$; $TC_{\pi s}$ denotes the total number of consumers who have at least one consumption occasion associated with context s within segment π $\left(\text{i.e.}, TC_{\pi s}=\sum_i B_{is't}\{i\in\Lambda_\pi\}\times B_{is't}\{s'=s\}\right)$; $B_{is't}$ is an indicator variable, which takes the value 1 if the context s' is associated with consumer i's consumption occasion t. Also note that

$$\begin{aligned}
\frac{\partial P_{ij't}}{\partial z^1_{ij'}} &= \frac{\partial\left[e^{V_{ij't}}\big/\sum_k e^{V_{ikt}}\right]}{\partial z^1_{ij'}} = \left(\frac{e^{V_{ij't}}}{\sum_k e^{V_{ikt}}}\right)\frac{\partial V_{ij't}}{\partial z^1_{ij'}} - \left[e^{V_{ij't}}\big/\left(\sum_k e^{V_{ikt}}\right)^2\right]\left(e^{V_{ij't}}\right)\frac{\partial V_{ij't}}{\partial z^1_{ij'}} = \frac{\partial V_{ij't}}{\partial z^1_{ij'}}P_{ij't}\left(1-P_{ij't}\right)\\
&= 2w^1_{is}\left(z^1_{ij'}-IP^{s,1}_{i,z}\right)P_{ij't}\left(1-P_{ij't}\right)
\end{aligned}$$

$$\frac{\partial P_{ijt}}{\partial z_{ij'}^{1}} = \frac{\partial \left[e^{V_{ijt}} / \sum_k e^{V_{ikt}} \right]}{\partial z_{ij'}^{1}} = -\left[e^{V_{ijt}} / (\sum_k e^{V_{ikt}})^2 \right] \left(e^{V_{ij't}} \right) \frac{\partial V_{ij't}}{\partial z_{ij'}^{1}} = -\frac{\partial V_{ij't}}{\partial z_{ij'}^{1}} P_{ijt} P_{ij't}$$

$$= -2 w_{is}^{1} \left(z_{ij'}^{1} - IP_{i,z}^{s,1} \right) \left(P_{ijt} P_{ij't} \right) \qquad (G-2)$$

where

$$V_{ijt} = \gamma_j^{*} - \sum_{s=1}^{S} \sum_{k=1}^{K} \left(z_{ij}^{k} - IP_{i,z}^{s,k} \right)^2 w_{is}^{k} B_{1,ist} + X_{ijt} \beta$$

$$TC_{\pi s} = \sum_{i} B_{is't} \{ i \in \Lambda_\pi \} \times B_{is't} \{ s' = s \}$$

$$TS_{is} = \sum_{t=1}^{T_i} B_{is't} \{ s' = s \}.$$

The (own- and cross-) elasticities with respect to a change along dimension 2 are

$$\eta_{s,F_2,j'j'}^{\pi} = \frac{1}{TC_{\pi s}} \sum_{i \in \Lambda_\pi} \left\{ \frac{1}{TS_{is}} \sum_{t=1}^{T_i} \left(\frac{\partial P_{ij't}}{\partial z_{ij'}^{2}} \right) \left(\frac{z_{ij'}^{2}}{P_{ij't}} \right) \times B_{is't} \{ s' = s \} \right\} \qquad (G-3)$$

$$= \frac{1}{TC_{\pi s}} \sum_{i \in \Lambda_\pi} \left\{ \frac{1}{TS_{is}} \sum_{t=1}^{T_i} \left(\frac{\partial V_{ij't}}{\partial z_{ij'}^{2}} \right) P_{ij't} \left(1 - P_{ij't} \right) \times \left(\frac{z_{ij'}^{2}}{P_{ij't}} \right) \times B_{is't} \{ s' = s \} \right\}$$

$$= 2 z_{\pi j'}^{2} \left(z_{\pi j'}^{2} - IP_{\pi,z}^{s,2} \right) \times \frac{1}{TC_{\pi s}} \sum_{i \in \Lambda_\pi} \left\{ \frac{1}{TS_{is}} \sum_{t=1}^{T_i} \left(w_{is}^{2} \left(1 - P_{ij't} \right) \times B_{is't} \{ s' = s \} \right) \right\}$$

$$\eta_{s,F_2,jj'}^{\pi} = \frac{1}{TC_{\pi s}} \sum_{i \in \Lambda_\pi} \left\{ \frac{1}{TS_{is}} \sum_{t=1}^{T_i} \left(\frac{\partial P_{ijt}}{\partial z_{ij'}^{2}} \right) \left(\frac{z_{ij'}^{2}}{P_{ijt}} \right) \times B_{is't} \{ s' = s \} \right\}$$

$$= \frac{1}{TC_{\pi s}} \sum_{i \in \Lambda_\pi} \left\{ \frac{1}{TS_{is}} \sum_{t=1}^{T_i} - \left(\frac{\partial V_{ij't}}{\partial z_{ij'}^{2}} \right) z_{ij'}^{2} P_{ijt} \times B_{is't} \{ s' = s \} \right\}$$

$$= -2 z_{\pi j'}^{2} \left(z_{\pi j'}^{2} - IP_{\pi,z}^{s,2} \right) \times \frac{1}{TC_{\pi s}} \sum_{i \in \Lambda_\pi} \left\{ \frac{1}{TS_{is}} \sum_{t=1}^{T_i} \left(w_{is}^{2} P_{ij't} \times B_{is't} \{ s' = s \} \right) \right\}.$$

Also, the (own- and cross-) elasticities corresponding to a change in the K-dimensional Euclidean distance between $z_{\pi j}$ and $IP_{\pi,z}^{s}$ are derived using the total derivative as follows:

$$\eta^{\pi}_{s,E,j'j'} = \frac{1}{TC_{\pi s}} \sum_{i \in \Lambda_\pi} \left\{ \frac{1}{TS_{is}} \sum_{t=1}^{T_i} \left(\frac{dP_{ij't}}{dz_{ij'}} \right) \left(\frac{z_{ij'}}{P_{ij't}} \right) \times B_{is't}\{s' = s\} \right\} \qquad (G-4)$$

$$= \frac{1}{TC_{\pi s}} \sum_{i \in \Lambda_\pi} \left\{ \frac{1}{TS_{is}} \sum_{t=1}^{T_i} \left(\sum_{k=1}^{K} \left(\frac{\partial V_{ij't}}{\partial z^{\mathrm{k}}_{ij'}} \right) P_{ij't} \left(1 - P_{ij't}\right) \times \left(\frac{z_{ij'}}{P_{ij't}} \right) \right) \times B_{is't}\{s' = s\} \right\}$$

$$= \frac{1}{TC_{\pi s}} \sum_{i \in \Lambda_\pi} \left\{ \frac{1}{TS_{is}} \sum_{t=1}^{T_i} \left(\sum_{k=1}^{K} 2 w^{k}_{is} z^{k}_{ij'} \left(z^{k}_{ij'} - IP^{s,k}_{i,z} \right) \left(1 - P_{ij't}\right) \right) \times B_{is't}\{s' = s\} \right\}$$

$$= \sum_{k=1}^{K} \left[2 z^{k}_{ij'} \left(z^{k}_{ij'} - IP^{s,k}_{i,z} \right) \times \times \frac{1}{TC_{\pi s}} \sum_{i \in \Lambda_\pi} \left\{ \frac{1}{TS_{is}} \sum_{t=1}^{T_i} \left(w^{k}_{is} \left(1 - P_{ij't}\right) \times B_{is't}\{s' = s\} \right) \right\} \right]$$

$$\eta^{\pi}_{s,E,jj'} = \frac{1}{TC_{\pi s}} \sum_{i \in \Lambda_\pi} \left\{ \frac{1}{TS_{is}} \sum_{t=1}^{T_i} \left(\frac{dP_{ijt}}{dz_{ij'}} \right) \left(\frac{z_{ij'}}{P_{ijt}} \right) \times B_{is't}\{s' = s\} \right\}$$

$$= \frac{1}{TC_{\pi s}} \sum_{i \in \Lambda_\pi} \left\{ \frac{1}{TS_{is}} \sum_{t=1}^{T_i} \left(\sum_{k=1}^{K} - \left(\frac{\partial V_{ij't}}{\partial z^{\mathrm{k}}_{ij'}} \right) z^{k}_{ij'} P_{ijt} \right) \times B_{is't}\{s' = s\} \right\}$$

$$= \sum_{k=1}^{K} \left[-2 z^{k}_{ij'} \left(z^{k}_{ij'} - IP^{s,k}_{\pi,z} \right) \times \frac{1}{TC_{\pi s}} \sum_{i \in \Lambda_\pi} \left\{ \frac{1}{TS_{is}} \sum_{t=1} \left(w^{\mathrm{k}}_{is} P_{ij't} \times B_{is't}\{s' = s\} \right) \right\} \right]$$

where $z_{ij'} = \left[z^{1}_{ij'}, z^{2}_{ij'}, \cdots\ z^{K}_{ij'} \right]$.

WEB APPENDIX H. Attribute selection with multiple regression models

We ran regressions on all possible combinations of potential independent variables. With the four potential independent variables, we have 2^4 distinct combinations to test. We selected the subset of predictors based on R^2 and mean squared error of prediction (MSEP hereafter).[2] Note that R^2 is the fraction of the total sum of squares that is explained by the regression and the RMSE is a measure of the average deviation of the estimates from the observed values (Darlington 1968; Stein 1960). R^2 is conveniently scaled between 0 and 1, whereas RMSE is not scaled to any particular values. This can be good or bad; obviously R^2 can be more easily interpreted, but with RMSE we know explicitly how much our predictions deviate, on average, from the actual values in the dataset. Below are the subset regressions with both large R^2 and small MSEP for Segment 2.

Table H-1. Regressions on dimension 2 for Segment 2

Model index on dimension 2	Predictors	R^2	MSEP	
Model 1-1	Price	0.2653	5.2506	
Model 1-2	Price, Color	0.3704	5.9988	
Model 1-3	Price, Color, Bitterness	0.4868	6.8459	
Model 1-4	Price, Calories, Color, Bitterness	0.7116	5.7701	BEST
Model 1-5	Price, Calories, Alcohol, Color, Bitterness	0.7139	9.5404	

$$Dimension\ 2 = -6.465 + (9.920) \times Price + (9.454) \times Calories + (-5.378) \times Color + (0.307) \times Bitterness$$

Dimension 2	**Coeff.**	**Std. Error**	**t value**
(Intercept)	-6.465	3.968	-1.629
Price	9.920**	3.791	2.617
Calories	9.454	4.788	1.974
Color	-5.378*	2.205	-2.439
Bitterness	0.307*	0.131	2.335

*p<0.10, **p<0.05, and ***p<0.01

[2] The MSEP computes the estimated mean square error of prediction assuming that both independent and dependent variables are multivariate normal (Darlington 1968; Stein 1960). $MSEP = MSE \times \frac{(n+1)(n-2)}{n(n-p-1)}$, where $MSE = \frac{SSE}{(n-p)}$, n is the sample size, and p is the number of predictors including the intercept.

WEB APPENDIX I. Advertising-based respositioning via Amazon Mechanical Turk

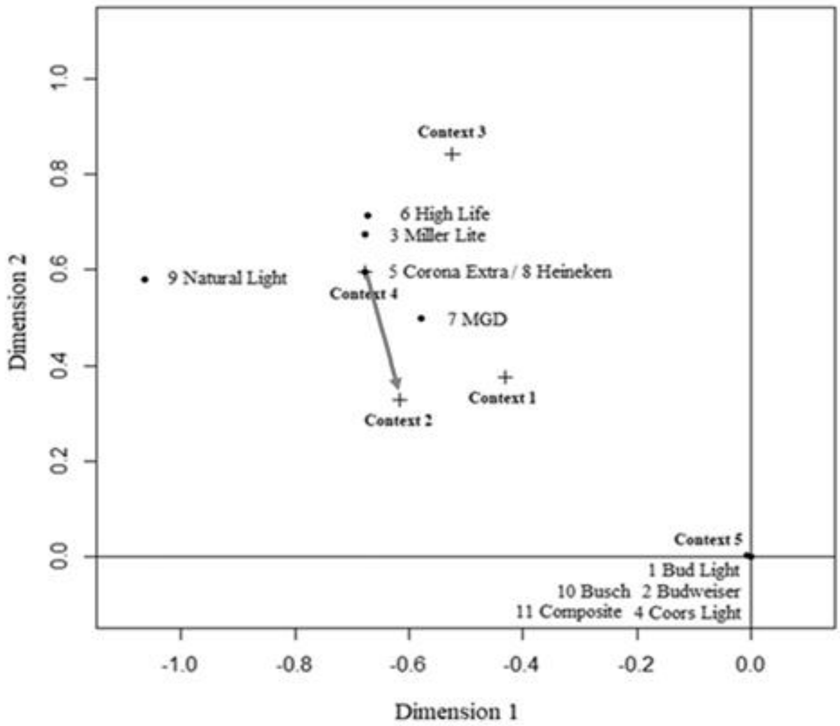
We conduct two small-scale discrete choice experiments via MTurk and verify whether repositioning based on the brand and context locations translates, at least directionally, to the changes in market shares predicted by our model. In total, 84 participants passed attention check questions and successfully completed the survey. These participants provided us with their brand choices in the five consumption contexts (i.e., their context-dependent preferences as seen in Table I-1).

Table I-1. Brand choices across contexts in MTurk experiment 1

	Context 1 (Relaxing)		Context 2 (With a meal)		Context 3 (Working at home)		Context 4 (Dancing)		Context 5 (Local pub)		Total
	Freq.	%	Freq.	%	Freq.	%	Freq.	%	Freq.	%	Freq.
1 Bud Light	14	16.67%	14	16.67%	12	14.29%	11	13.10%	10	11.90%	61
2 Budweiser	8	9.52%	5	5.95%	5	5.95%	7	8.33%	10	11.90%	35
3 Miller Lite	7	8.33%	7	8.33%	6	7.14%	10	11.90%	5	5.95%	35
4 Coors Light	8	9.52%	3	3.57%	6	7.14%	6	7.14%	3	3.57%	26
5 Corona Extra	11	13.10%	2	2.38%	19	22.62%	15	17.86%	10	11.90%	57
6 High Life	1	1.19%	1	1.19%	1	1.19%	3	3.57%	2	2.38%	8
7 MGD	0	0.00%	20	23.81%	0	0.00%	0	0.00%	1	1.19%	21
8 Heineken	10	11.90%	14	16.67%	14	16.67%	17	20.24%	23	27.38%	78
9 Natural Light	1	1.19%	2	2.38%	1	1.19%	1	1.19%	0	0.00%	5
10 Busch	6	7.14%	1	1.19%	2	2.38%	0	0.00%	1	1.19%	10
11 Composite	18	21.43%	17	21.86%	18	21.43%	14	16.67%	19	22.62%	86
	84	100.00%	84	100.00%	84	100.00%	84	100.00%	84	100.00%	420

As the data show evidence of both within and across variation in the participants' choices, we used this panel dataset to run our proposed CDMDU model. Based on the parameter estimates, we obtained both brand locations and context-dependent ideal points (see Figure I-1).

Figure I-1. CDMDU-based joint space map



Since we are interested in a brand's repositioning and the consequent changes in consumers' brand choices, we re-contacted all of the 84 MTurk workers who had completed the survey (Part 1) and gave them the option of participating in a follow-up survey (Part 2). The main purpose of this part of the survey was to check whether advertising could effectively move a firm's position closer to the ideal point. We considered Corona Extra's (Brand 5) repositioning from its current position near context 4 to a position closer to the ideal point for context 2. To make survey participants perceive Corona Extra's repositioning, we showed the participants an advertisement (photo) of Corona Extra that emphasizes consumption context 2, eating a meal at home. Indeed, companies often create TV commercials to establish links between their products and consumption occasions, as seen in Figure I-2 (left[3] and right[4]).

---

[3] Source: http://www.adweek.com/creativity/fans-crammed-bud-house-world-cup-12579/ (needs subscription)
[4] Source: http://adsoftheworld.com/media/print/goldfish(needs subscription) (needs subscription)

Figure I-2. Budweiser's advertising for context 1 (left) and context 5 (right)

Note that there might be a positive effect of advertising for Corona Extra (i.e., market shares of Corona Extra can increase regardless of contexts because of the advertising). Thus, for other brands, we also had the participants exposed to advertisements that emphasize a consumption context closely located to their current brand locations (i.e., without repositioning), as seen in Figure I-3.

Figure I-3. Images of advertisements shown to MTurk workers

Advertisement for Brand 5 (Corona Extra): Repositioning toward context 2 (eating a meal at home)

Advertisement for Brand 8 (Heineken): Staying close to context 4 (dance club or sports bar)

Advertisement for Brand 2 (Budweiser): Staying close to context 5 (local pub)

Advertisement for Brand 11 (Composite): Staying close to context 5 (local pub)

Brand	Brand repositioning	Context associated with the current brand location	Context emphasized in the ad
Budweiser (Brand 2)	NO	Context 5 (local pub)	
Corona Extra (Brand 5)	YES	Context 4 (dancing)	Context 2 (with a meal)
Heineken (Brand 8)	NO	Context 4 (dancing)	
Composite (Brand 11)	NO	Context 5 (local pub)	

In Table I-2, we provide new brand choices associated with each of the five consumption contexts.

Table I-2. Brand choices across contexts in MTurk experiment 2

	Context 1 (Relaxing)		Context 2 (With a meal)		Context 3 (Working at home)		Context 4 (Dancing)		Context 5 (Local pub)		Total
	Freq.	%	Freq.	%	Freq.	%	Freq.	%	Freq.	%	Freq.
1 Bud Light	6	15.79%	0	0.00%	4	10.53%	5	13.16%	11	28.95%	15
2 Budweiser	5	13.16%	7	18.42%	4	10.53%	4	10.53%	4	10.53%	24
3 Miller Lite	1	2.63%	1	2.63%	2	5.26%	2	5.26%	4	10.53%	10
4 Coors Light	2	5.26%	3	7.89%	3	7.89%	2	5.26%	1	2.63%	11
5 Corona Extra	3	7.89%	4	10.53%	4	10.53%	6	15.79%	5	13.16%	22
6 High Life	2	5.26%	1	2.63%	1	2.63%	1	2.63%	2	5.26%	7
7 MGD	0	0.00%	0	0.00%	0	0.00%	2	5.26%	0	0.00%	2
8 Heineken	8	21.05%	9	23.68%	11	28.95%	9	23.68%	11	28.95%	48
9 Natural Light	0	0.00%	0	0.00%	1	2.63%	1	2.63%	0	0.00%	2
10 Busch	1	2.63%	0	0.00%	0	0.00%	0	0.00%	0	0.00%	1
11 Composite	10	26.32%	13	34.21%	8	21.05%	6	15.79%	11	28.95%	48
	38	100.00%	38	100.00%	38	100%	38	100.00%	38	100.00%	190

The results of this exercise showed that advertisement appears to have a positive impact. For example, in the case of Budweiser (Brand 2), market shares increased in most contexts except for context 5. In particular, Heineken's advertising appears to be quite effective, with its market share increasing in all five contexts. Importantly, however, in the presence of other brands' advertising, Corona Extra's shares in contexts 3 and 4 decreased but its share in context 2 (i.e., targeted context) increased from 2.38% to 10.53%.

WEB APPENDIX J: Correspondence analysis

Correspondence analysis (CA hereafter) is a graphic method of exploring the relationship between variables in a contingency table. Scaling the rows and columns of a rectangular data matrix with non-negative entries and displaying each corresponding unit graphically in the same low-dimensional space, CA is also known as dual scaling, optimal scaling, method of reciprocal averages, and canonical analysis of contingency tables in the statistical and psychometric literature (Hoffman and Franke 1986).

Compared to other familiar multivariate approaches to graphical data representation such as principal components analysis and canonical correlation anlaysis, CA has the distinct advantage of producing two dual displays whose row and column geometries have similar interpretations. While we believe CA to be a useful technique that enables researchers to easily detect relationships with a seemingly easy-to-read visualization, CA has several well-known limitations such as loss of information due to aggregation, a possible curse of dimensionality, and an inability to address consumer heterogeneity (i.e., segmentation) as well as the roles of marketing variables (e.g., advertising) for marketing research. Basically, CA is not appropriate for the hypothesis testing required for marketing research (Hoffman and Franke 1986). In what follows, however, our discussion focuses instead on the limitations of its visual interpretation: what comparisons can be made of distances—between row points (i.e., beer brands), column points (i.e., contexts), and/or row and column points (i.e., beer brand-context)—in the geometric representation obtained via CA.

For that, we conducted CA using a contingency table (see Table 3) with the row and column representing beer brand and consumption context, respectively. We first computed the averages for each row and column and then computed each cell's expected value by multiplying the row average for that cell with the column average and dividing it by the overall average. Subtracting the expected values from the original data, we computed the residuals, which show the associations between the row (beer brand) and column (context) labels, as seen in Table J-1.

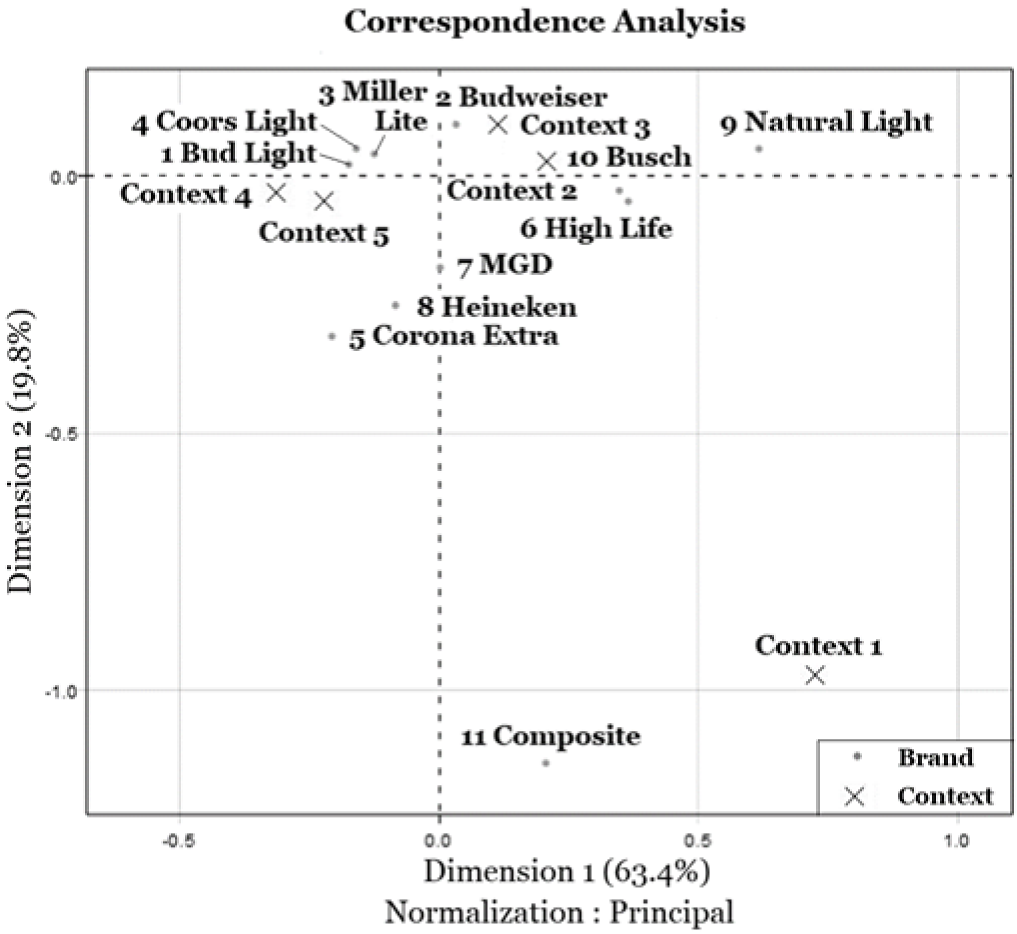
Table J-1. Residuals for brand choices across contexts based on Table 4

	Context 1 (Relaxing)	Context 2 (With a meal)	Context 3 (Working at home)	Context 4 (Dancing)	Context 5 (Local pub)
Bud Light	-34	-42	4	25	50
Budweiser	128	-31	-5	-5	-86
Miller Lite	-136	-1	13	34	91
Coors Light	-60	-18	11	28	37
Corona Extra	8	-13	-23	24	2
High Life	-9	94	-17	-31	-37
MGD	20	-6	-18	-3	6
Heineken	-13	2	-11	6	19
Natural Light	34	41	27	-46	-54
Busch	67	-22	14	-31	-29
Composite	-4	0	1	1	4

Note that big positive numbers mean a strong positive relationship, and the opposite is true for negatives. From the residuals for Miller Lite, we can see that its biggest score is for context 5 and its lowest score is for context 1. Comparing the residuals for Bud Light with those for Coors Light across contexts, all of the residuals are in the same direction. And Budweiser is least like the composite because most residuals across contexts are in opposite directions. Plotting the labels with similar residuals close together, CA produces a visualization as a set of row and column points in two-dimensional space in Figure J-1. We find that the relative positions of beer brands in the visualization are consistent with the similarities of their respective residuals.

In Figure J-1, the points vary much more on the horizontal than on the vertical. As shown in the axes' labels, the horizontal dimension explains 63.4% of the variance in the data whereas the vertical dimension explains only 19.8%, meaning that 83.2% of the variance is explained by these two dimensions. CA places the row labels on the plot such that the closer two rows (i.e., beer brands) are to each other, the more similar their residuals. This also applies to the column (i.e., consumption context) labels.

Figure J-1. CA of beer brand and context (Normalization: Principal)

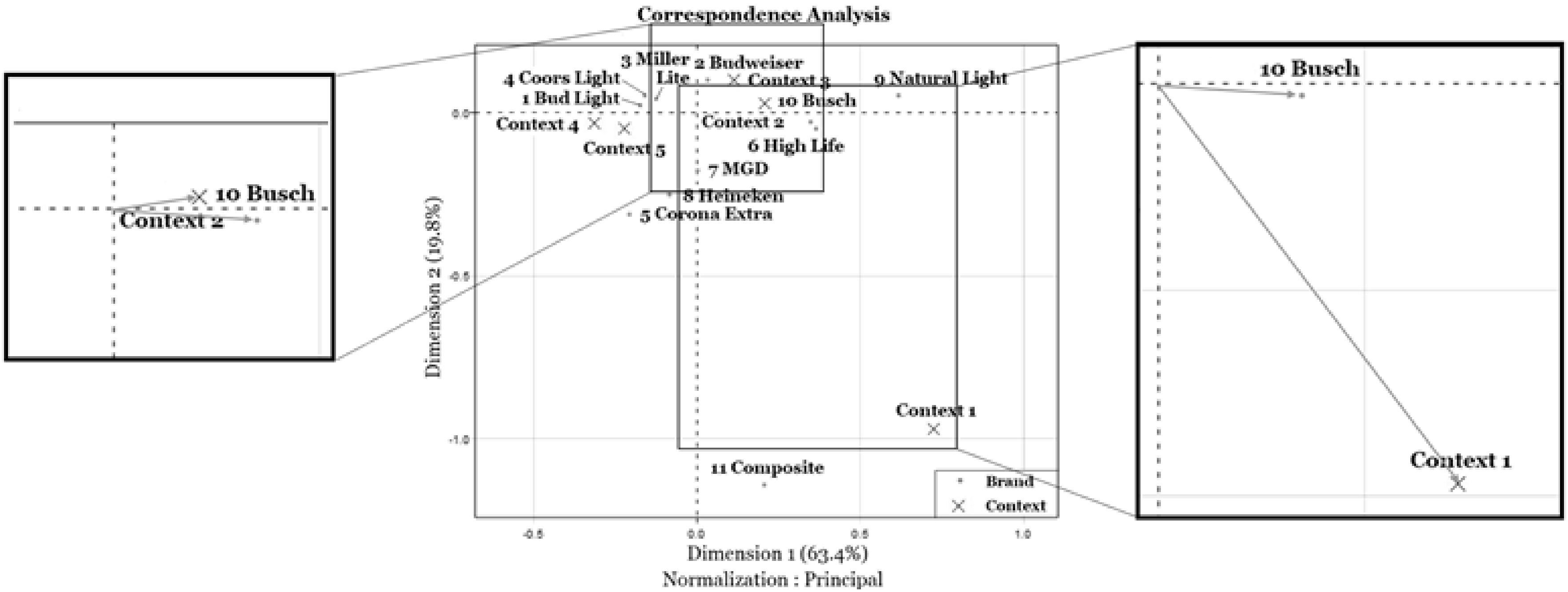


However, we cannot conclude that the greater proximity between a row label and a column label indicates the higher association. Because the measure of association used in CA is the chi-square distance between the response categories (Clausen 1998), an important caveat is that the between-set distances cannot be interpreted for CA. In other words, the joint display of coordinates shows the relationship between a point from one set and all the points of the other set, not between individual points from each set. Although Carroll, Green, and Schaffer (1986) proposed a metric joint space model in which all within-set and between-set distances are comparable, this scaling is not only different from the conventional scaling used in CA but also requires additional technical procedures with computational work. To the best of our knowledge, most of the canned software programs for correspondence analysis (e.g., SPSS) do not provide such a detailed procedure for users' cautious interpretation of results. In fact, we searched several canned programs and found that the latest R package, which considers several procedures for interpretation, nevertheless does not discuss between-set distances.

To better understand this issue, we compare Busch, High Life, and Natural Light in context 2. Busch and High Life are almost identically proximate and close together, although the residuals for High Life–context 2 and Busch–context 2 are 41 and –22, respectively. In addition, whereas it appears that Busch has relatively high preference in context 2 in a map, the residual is negative and the number of consumption situations in Table 3 is only 66. In other words, CA does not show us which rows have the highest numbers, nor which columns have the highest numbers. It instead shows the relativities. If we are interested in how beer brand preferences differ across contexts, we are better off plotting the raw data than using CA. At the same time, the residual for Budweiser–context 1 is the highest at 128, but they are not close together on the map. What is going on here? There is, in fact, no way to position the labels on the CA factor map to sensibly communicate these residuals.

If we want to compare a row label to a column label, we need to check at least three things. First, we have to check the length of the line connecting the row label to the origin. Longer lines indicate that the row label is highly associated with some of the column labels (i.e., it has at least one high residual). Second, we can look at the length of the label connecting the column label to the origin. Longer lines again indicate a high association between the column label and one or more row labels. Third, we have to check the angle connecting a row and column label to the origin.

Figure J-2. Associations between row and column labels (i.e., brands and contexts)



For example, while Busch and context 2 are closely located in a map, the angle formed between the two indicates a positive association (i.e., the angle is less than 90 degree). Owing to the short lines, however, the correct interpretation is that there is either no association or a very weak one. In contrast, Busch and context 1 are not closely located. Since the arrow to context 1 is longer and the angle is less than 90 degrees, however, Busch is more strongly associated with context 1 than with context 2.

However, note that interpreting the angles is strictly valid only when you have either row principal or column principal. The normalization, which is a technical option in CA software, needs to have been set to either principal or row/columns principal. For instance, the aspect ratio of the map needs to be fixed at 1. That is, the horizontal and vertical coordinates of the map need to match each other. Thus, to make inferences about the relationships between the rows and columns, researchers should choose the default principal normalization (Figure J-1 is the CA map based on principal normalization). However, principal normalization is not a default option in all statistical packages. For instance, it is not the default in SPSS, so comparing the distances between row labels in a map created by SPSS with the default settings leads to misinterpretations. In summary, when complex multivariate relationships are examined, correspondence analysis is limited only by the researcher's ingenuity in interpreting the derived spatial map.