Modeling Dynamic Effects in Repeated-Measures Experiments Involving Preference/Choice: An Illustration Involving Stated Preference Analysis

Wayne S. DeSarbo, Pennsylvania State University
Donald R. Lehmann, Columbia University
Frances Galliano Hollman, Pennsylvania State University

Preference structures that underlie survey or experimental responses may systematically vary during the administration of such measurement. Maturation, learning, fatigue, and response strategy shifts may all affect the sequential elicitation of respondent preferences at different points in the survey or experiment. The consequence of this phenomenon is that responses and effects can vary systematically within the data set. To capture these structural changes, the authors present a maximum likelihood–based change-point multiple regression methodology that explicitly detects discrete structural changes at various points in time/sequence in regression coefficients by simultaneously estimating the number of change points, their location and duration in the sequence of data points, and the respective regression coefficients for each subset of the data defined by the change points. An application involving a stated preference or conjoint analyses study of student apartment choices illustrates that the structure of preferences changes significantly over the sequence of profile responses. *Index terms: preference/choice experiments, behavioral decision making, maximum likelihood estimation, models of structural change, conjoint analysis, consumer psychology.*

Introduction

Much of the existent behavioral literature in behavioral decision making has been involved with the measurement of consumers' preference or utility functions for a specified stimulus set (e.g., a product or service). Normative theories of value maximization (e.g., von Neumann & Morgenstern, 1947) posit that each respondent possesses stable preferences for all possible stimulus options—or what Moore (1999) calls an "internal global preference set." However, as illustrated by Tversky and Simonson (1993), without a global preference set, one quickly faces serious problems applying these value maximization principles. And there is a plethora of research that questions the existence of such an internal global preference set. For example, Moore argues that cognitive storage limitations would make such an elaborate set of established preferences virtually impossible.

Applied Psychological Measurement, Vol. 28 No. 3, May 2004, 186–209

DOI: 10.1177/0146621604264150

Authors such as Herr (1989) and Simonson and Tversky (1992) argue that preferences become more firmly established with experience—a dynamic process. This view is consistent with the vast empirical evidence that preferences are not simply revealed but are actually constructed in the process of their elicitation (Coupey, Irwin, & Payne, 1998; Fischhoff, Slovic, & Lichtenstein, 1980; March, 1978; Payne, Bettman, & Johnson, 1992). This is *not* stating that respondents have no preferences before they are asked about them. It does, however, suggest that respondents have strategies they use for assembling their preferences and that these strategies are *unlikely* to produce preferences that are invariant over time.

Other authors such as Bettman (1979), Bettman and Zins (1979a, 1979b), and Bettman and Park (1980) also support this view in their constructive theories of choice. These authors state that respondents (here consumers) may not have complete rules or heuristics stored in memory that are used in the process of preference and choice formation. Rather, respondents/consumers may have only fragments or elements of heuristics in memory that are then put together during the actual choice process to "construct" a heuristic. Such elements may involve the beliefs about alternatives, evaluations, simple rules of thumb, rules for integrating data, and so on. As Bettman and Park note, the specific elements used for a particular choice and the sequence in which they will be used will be a function of such factors as the following: what external present information is available, the format in which that information is presented, the degree to which various pieces of information "stand out," intermediate processing results, and so on. (See also Payne, Bettman, & Schkade, 1999, who also posit that preferences are constructed rather than recalled.)

Such dynamic processes underlying respondent preference/utility have also been examined from an information integration perspective. Consistent with information integration theory (Anderson, 1981), Johar, Jedidi, and Jacoby (1997) conceptualize stimulus evaluation formation process as a function of prior evaluation (experience) and new information. With this view, respondents are likely to anchor their stimulus evaluations on prior evaluations and adjust these evaluations based on newly acquired information (cf. Einhorn & Hogarth, 1985; Lopes, 1982). When new information is acquired, evaluations may stay the same, become more favorable, or become less favorable. Lopes (1982, p. 2) describes such a serial adjustment process as one in which information is scanned, items are selected for processing, scale values are assessed, and adjustments are made to an interim quantity that summarizes the results from already processed information. Other dynamic adjustments in utility functions based on newly acquired information have also been studied by Gilbert, Krull, and Pelham (1988) and Gilbert (1989). Consequently, as stated explicitly in Kardes and Kalyanaram (1992), respondent preferences are likely to evolve over time through an anchoring and adjustment process (Kahneman, Slovic, & Tversky, 1982; Kahneman & Snell, 1990).

Meyer (1987) examined this dynamic process in the context of how multiattribute judgments are made. He stated that the algebraic cognitive rules underlying such preference elicitation evolve over time/task through learning. At early stages of learning, knowledge about phenomena (e.g., stimulus attributes) is thought likely to consist primarily of recollections of individual experiences (cf. Hayes-Roth, 1977; Taylor & Crocker, 1981). As the pool of experiences grows, these experiences become better organized into associative networks or structures that define generalized knowledge (e.g., Anderson, 1983; Shank & Abelson, 1977). Thus, at early stages, respondents use heuristics that require only episodic knowledge about a class of stimuli. As experience increases, however, a wider range of judgment strategies becomes feasible as the decision maker has both a wider range of experiences to draw on and generalized knowledge about the antecedents of value. Meyer experimentally demonstrated how judgmental processes change over time during the course of learning in a context where respondents have high levels of access to the outcomes of previous judgments.

Although much of this literature focuses on structural changes over time (e.g., prior experience), such changes can occur within the context of an experiment or survey itself. It has been well documented that the reactive nature of the measurement itself can affect such preferences/choices vis-à-vis fatigue, order effects, learning, response shift patterns, context effects, maturation, and so forth (cf. Anderson, 1981; Feldman & Lynch, 1988; Jain & Pinson, 1976; Johnson, Lehmann, & Horne, 1990; Tversky & Simonson, 1993). Thus, the structure underlying preferences in the beginning stages of a repeated-measures or sequential measurements task may often be quite different from the preference structure that has evolved at the end of the same measurements task.

These phenomena have been recently demonstrated in the stated preference or conjoint analysis literature (Huber, Wittink, Johnson, & Miller, 1992), in which there is mention of a "burn-in" period during which respondents evolve a systematic approach to providing preference judgment as they examine attributes and attribute levels. More significantly, Johnson and Orme (1996) have noticed significant changes in attribute weights over the course of a single conjoint survey. (For other literature involving shifts in [time-varying] parameters with respect to time-series data, see also Hanssens, Parsons, & Schultz, 2001; Leeflang, Wittink, Wedel, & Naert, 2000; Wildt & Winer, 1983.)

This article presents a method for determining if and where, during the sequence of responses/observations, significant shifts occur in preference structures in a multiple regression context. The authors describe a maximum likelihood–based methodology that segments observations in a contiguous manner such that sets of optimal response models can be estimated. Section 2 reviews the extensive psychometrics, statistics, econometrics, and engineering literature on such change-point problems. The authors present a modified estimation procedure that is computationally simple and quick, generalizes to multiple change points, can accommodate large data sets and a variety of model selection criteria/heuristics, and is guaranteed to locate globally optimum solutions. In Section 3, an application to a stated preference or conjoint analysis problem, which deals with student apartments in which such structural changes are observed and estimated, is described. Finally, further directions for future research are discussed in Section 4.

Change-Point Multiple Regression

Multiple regression model instability may often be due to a discrete "switch" or change in the regression equation from one subsample period (or regime) to another. A variety of approaches currently exist for attacking such problems. As an approximation, curvilinear or nonlinear models can be fit in such cases in which time/period is an independent variable whose effects vary nonlinearly. Here, one has to specify a priori the exact specification of such nonlinear functions, and that specification places constraints as to the nature of how response can change (e.g., monotonicity). Indeed, in most social science applications with little a priori theory, it is difficult to prescribe a particular functional form to fit. In addition, issues with using such an approximation exist in multiple regression contexts in which other independent variables may also exhibit differential effects over time, especially in a noncontinuous manner.

Chow (1960) proposed an F test (see also Greene, 2003, pp. 130-132) to test for structural changes in multiple regression specifications for the special case in which the dates and duration of these separate subsamples are *known*. In a similar fashion, Neter, Kutner, Nachtsheim, and Wasserman (1996) and von Eye and Schuster (1998) formulate "piecewise regression" (either continuous or discontinuous) methods for estimating separate linear or nonlinear regression functions for subsections of the sample in which the cutoffs for defining the separate functions are *known* and given a priori (typically as ranges of some independent variable). However, in most social science research applications involving repeated measures or sequential measurement, there is no a priori knowledge of whether or when such change points occur. Quandt (1958, 1960), Farley and Hinich (1970), and Kim and Siegmund (1989) propose models that permit at most one switch in the data series with an unknown change point. Quandt (1972); Goldfield and Quandt (1973); Brown, Durbin, and Evans (1975); Ploberger, Kramer, and Kontrus (1989); Kim and Maddala (1991); and

Hamilton (1989, 1990) considered regression models that allow for more than one change point. DeSarbo and Cron (1988) have developed a latent structure multiple regression methodology for segmenting observations into multiple segments based on information indices. However, this procedure does not enforce contiguity constraints on the observations that are segmented (i.e., the segments are not likely to contain consecutive observations). Andrews (1993), Wecker (1979), Sclove (1983), Neftci (1984), and Hamilton (1989) have also proposed similar methods and tests.

As an alternative approach, one can manipulate the Nonlinear Models Module of SYSTAT Software V10.2 to estimate regimes in some multiple regression contexts. Here, Gauss Newton, Quasi-Newton, and/or Simplex algorithms are used to estimate a user-specified nonlinear response model incorporating regime changes in either a least squares or maximum likelihood estimation (MLE) context. However, this approach cannot be gainfully applied to situations involving the presence of both cross-sectional (respondents) and longitudinal (sequence) data. One could average the responses across respondents and apply this approach, but that would destroy the sequence information. Or, if one were to concatenate the responses and independent variables, there would be no guarantee that the estimated change point(s) would occur between different respondents. In addition, the optimization procedure used in SYSTAT assumes a continuous range for the change point(s). (In fact, in the Example 10 provided in the SYSTAT manual, the estimated change point for a two-regime model was *not* an integer value.) This would not be a feasible solution to problems in which the cutoff points represent discrete time or sequence observations in which such structural changes occur. Finally, there is no adjustment for the number of parameters estimated in terms of the criterion to be optimized. How is the "optimal" number of regimes then to be determined when SYSTAT's ordinary least squares (OLS)/MLE loss function is monotonically decreasing/increasing with the number of regimes/parameters estimated? Kim and Nelson (1999), Chen and Gupta (2000), and Gustafsson (2000) have written excellent books summarizing the various classical and Bayesian statistical approaches to change point or switching regression from the statistics and engineering perspective.

Given the types of social science research applications typically dealt with for this class of problems (large samples, repeated measures, sequential trials, long time series), computational feasibility is an important issue. In addition, the ability to perform "parameter segmentation" while optimizing a relevant objective function is appealing. Finally, having the ability to solve for a globally optimal solution would be a major plus. As in Yao (1988) and Chen and Gupta (2000), the authors start with a maximum likelihood framework. Later, they demonstrate that this is equivalent to selecting solutions for a specified number of regimes that minimizes the total error sum of squares in the prediction.

Let

$t = 1, \ldots, T$ observations/time periods;

$k = 1, \ldots, K$ independent variables;

$r = 1, \ldots, R$ (unknown) regime or change points.

The standard multiple regression model examined is

$$y_t = \mathbf{X}_\mathbf{t}' \boldsymbol{\beta} + \varepsilon_t, t = 1, \ldots, T, \tag{1}$$

where

$$\mathbf{X}_\mathbf{t}' = (1, X_{1t}, X_{2t}, \ldots, X_{Kt}), \tag{2}$$

and

$$\boldsymbol{\beta}' = (\beta_0, \beta_1, \ldots, \beta_K) \tag{3}$$

is a $(K+1)$ unknown regression parameter vector; ε_t is a random error distributed as $N(0, \sigma^2)$, with σ^2 unknown; and ε_t is uncorrelated from observation to observation (extensions of the methodology to deal with such restrictions are discussed later). Thus, $y_t \sim N(\mathbf{X}_t' \beta, \sigma^2)$; that is, the dependent variable is conditionally normally distributed.

The authors examine single-stimulus judgment situations in which there is interest in testing whether there is a change in the regression model at some unknown location S, where $(K+2) \leq S \leq (T-K-2)$ due to the number of parameters to be estimated per regime. For now, the simple case of a single change point ($R = 1$) whose location (S) is unknown will be examined. The generalization to multiple change points unfolds easily from this development. In essence, the authors (cf. Chen and Gupta) are interested in testing the null hypothesis:

$$H_0 : \boldsymbol{\mu}_{\mathbf{y}_t} = \mathbf{X}_\mathbf{t}' \boldsymbol{\beta} \quad \text{for} \quad t = 1, \ldots, T \tag{4}$$

versus the alternative hypothesis:

$$H_1 : \begin{bmatrix} \boldsymbol{\mu}_{\mathbf{y_t}} = \mathbf{X}_\mathbf{t}' \boldsymbol{\beta}_1 \\ \boldsymbol{\mu}_{\mathbf{y_t}} = \mathbf{X}_\mathbf{t}' \boldsymbol{\beta}_2 \end{bmatrix} \begin{matrix} t = 1, \ldots, S \\ t = S+1, \ldots, T \end{matrix} \tag{5}$$

where S, $S \in (K+2, \ldots, T-K-2)$, is the unknown location of the change point, and $\boldsymbol{\beta}$, $\boldsymbol{\beta}_1$, $\boldsymbol{\beta}_2$ are the unknown regression parameters.

Now, define

$$\mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_T \end{pmatrix}, \quad \mathbf{X} = \begin{pmatrix} 1 & x_{11} & \cdots & x_{k1} \\ 1 & x_{12} & \cdots & x_{k2} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & x_{1T} & \cdots & x_{KT} \end{pmatrix} \equiv \begin{pmatrix} \mathbf{x}_1' \\ \mathbf{x}_2' \\ \vdots \\ \mathbf{x}_T' \end{pmatrix}, \quad \text{and} \quad \boldsymbol{\beta} = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_K \end{pmatrix};$$

then the null hypothesis H_0 corresponds to the following model:

$$\boldsymbol{\mu}_\mathbf{y} = \mathbf{X}\boldsymbol{\beta},$$

where

$$\boldsymbol{\mu}_\mathbf{y} = \begin{pmatrix} \mu_{y_1} \\ \mu_{y_2} \\ \vdots \\ \mu_{y_T} \end{pmatrix}.$$

Here, the likelihood function under H_0 in matrix notation is:

$$\begin{aligned} L_0(\boldsymbol{\beta}, \sigma^2) &= f(y_1, y_2, \ldots, y_T; \boldsymbol{\beta}, \sigma^2) \\ &= (2\pi)^{-T/2}(\sigma^2)^{-T/2} \exp\{-(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})'(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})/2\sigma^2\}, \end{aligned} \tag{6}$$

and the well-known MLEs of $\boldsymbol{\beta}$ and σ^2 are

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}, \tag{7}$$

$$\hat{\sigma}^2 = \frac{1}{T}(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}})'(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}). \tag{8}$$

Then, the maximum likelihood function under H_0 is

$$L_0(\hat{\boldsymbol{\beta}}, \hat{\sigma}^2) = (2\pi)^{-T/2}\left[\frac{1}{T}(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}})'(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}})\right]^{-T/2} e^{-T/2}. \tag{9}$$

Burnham and Anderson (2002) provide an excellent review and comparison of various model selection criteria for such maximum likelihood problems. In particular, a number of information-based heuristics are examined and related to Kullback-Leibler information. A general form for such criteria for a given model or parameterization m can be expressed as follows:

$$\text{GIC}_m = c_1(-2lnL_m) + c_2(P_m) + c_3(C_m), \tag{10}$$

where c_1, c_2, c_3 are specified scaling/weighting constants; L_m is the likelihood function calculated at its maximum value under model m; P_m is the number of free parameters associated with the estimation of model m; and C_m is the complexity of model m. The complexity of a particular model is often indicated by the variances and covariances of the estimated parameters vis-à-vis the determinant of the estimated information matrix. A variety of such information-based heuristics, including Akaike's information criterion (AIC), the corrected or modified Akaike's information criterion (AIC_c), Takeuchi's information criterion (TIC), the quasi-likelihood AIC (QAIC), the Bayesian or Schwartz information criterion (BIC), the consistent AIC (CAIC), ICOMP, MDL, and so forth can all be expressed as special cases of GIC above. Given the work by Bozdogan (1987, 1994) showing the superior performance of CAIC, that heuristic for model selection was chosen in this context, although the combinatorial optimization framework to be presented can be trivially extended for any such GIC-based measure.

Under the null hypothesis for this regression problem, one can define the consistent AIC, denoted by CAIC(T), as

$$\begin{aligned}\text{CAIC(T)} &= -2\log L_0(\hat{\boldsymbol{\beta}}, \hat{\sigma}^2) + (K+2)(\log(T)+1)\\ &= T\ \log[(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}})'(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}})] + T(\log 2\pi + 1) + (K+2-T)\,(\log(T)+1).\end{aligned} \tag{11}$$

Now define

$$\mathbf{y}_1 = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_S \end{pmatrix}, \mathbf{y}_2 = \begin{pmatrix} y_{S+1} \\ y_{S+2} \\ \vdots \\ y_T \end{pmatrix},$$

$$\mathbf{X}_1 = \begin{pmatrix} 1 & x_{11} & \cdots & x_{K1} \\ 1 & x_{12} & \cdots & x_{K2} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & x_{1S} & \cdots & x_{KS} \end{pmatrix} \equiv \begin{pmatrix} \mathbf{x}_1' \\ \mathbf{x}_2' \\ \vdots \\ \mathbf{x}_\mathbf{S}' \end{pmatrix},$$

$$\mathbf{X}_2 = \begin{pmatrix} 1 & x_{1(S+1)} & \cdots & x_{K(S+1)} \\ 1 & x_{1(S+2)} & \cdots & x_{K(S+2)} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & x_{1T} & \cdots & x_{KT} \end{pmatrix} \equiv \begin{pmatrix} \mathbf{x}_{\mathbf{S+1}}' \\ \mathbf{x}_{\mathbf{S+2}}' \\ \vdots \\ \mathbf{x}_\mathbf{T}' \end{pmatrix},$$

$$\beta_1 = \begin{pmatrix} \beta_0^1 \\ \beta_1^1 \\ \vdots \\ \beta_K^1 \end{pmatrix}, \boldsymbol{\beta}_2 = \begin{pmatrix} \beta_0^2 \\ \beta_1^2 \\ \vdots \\ \beta_K^2 \end{pmatrix}$$

for $S \in (K+2, \ldots, T-K-2)$. Then the alternative hypothesis, H_1, corresponds to the following models:

$$\boldsymbol{\mu}_{\mathbf{y}_1} = \mathbf{X}_1\boldsymbol{\beta}_1,\ \ \boldsymbol{\mu}_{\mathbf{y}_2} = \mathbf{X}_2\boldsymbol{\beta}_2, \sigma_1, \sigma_2 \tag{12}$$

where

$$\boldsymbol{\mu}_{\mathbf{y}_1} = \begin{pmatrix} \mu_{y_1} \\ \mu_{y_2} \\ \vdots \\ \mu_{y_S} \end{pmatrix}, \text{ and } \boldsymbol{\mu}_{\mathbf{y}_2} = \begin{pmatrix} \mu_{y_{S+1}} \\ \mu_{y_{S+2}} \\ \vdots \\ \mu_{y_T} \end{pmatrix} \text{ for } S \in (K+2, \ldots, T-K-2). \tag{13}$$

Here, the likelihood function is

$$\begin{aligned} L_1(\boldsymbol{\beta}_1,\ \boldsymbol{\beta}_2, \sigma_1^2, \sigma_2^2) &= f\left(y_1, y_2, \ldots, y_T; \boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \sigma_1^2, \sigma_2^2\right) \\ &= (2\pi)^{-T/2}\,(\sigma_1^2)^{-S/2}\exp\{-(\mathbf{y}_1 - \mathbf{X}_1\boldsymbol{\beta}_1)'(\mathbf{y}_1 - \mathbf{X}_1\boldsymbol{\beta}_1)/2\sigma_1^2\} \\ &\quad \cdot (\sigma_2^2)^{-(T-S)/2}\exp\{-(\mathbf{y}_2 - \mathbf{X}_2\boldsymbol{\beta}_2)'(\mathbf{y}_2 - \mathbf{X}_2\boldsymbol{\beta}_2)/2\sigma_2^2\}, \end{aligned} \tag{14}$$

and the MLEs of the parameters are

$$\hat{\boldsymbol{\beta}}_1 = (\mathbf{X}_1'\mathbf{X}_1)^{-1}\mathbf{X}_1'\mathbf{y}_1, \tag{15}$$

$$\hat{\boldsymbol{\beta}}_2 = (\mathbf{X}_2'\mathbf{X}_2)^{-1}\mathbf{X}_2'\mathbf{y}_2, \tag{16}$$

$$\hat{\sigma}^2 = \frac{1}{S}[(\mathbf{y}_1 - \mathbf{X}_1\hat{\boldsymbol{\beta}}_1)'(\mathbf{y}_1 - \mathbf{X}_1\hat{\boldsymbol{\beta}}_1)], \tag{17}$$

$$\hat{\sigma}_2^2 = \frac{1}{T-S}(\mathbf{y}_2 - \mathbf{X}_2\hat{\boldsymbol{\beta}}_2)'(\mathbf{y}_2 - \mathbf{X}_2\hat{\boldsymbol{\beta}}_2). \tag{18}$$

Then, the concentrated maximum likelihood function can be written as (cf. Chen and Gupta, 2000)

$$\begin{aligned} &L_1(\hat{\boldsymbol{\beta}}_1,\ \hat{\boldsymbol{\beta}}_2, \hat{\sigma}_1^2, \hat{\sigma}_2^2) \\ &= (2\pi)^{-T/2}e^{-T/2}\left[\frac{(\mathbf{y}_1 - \mathbf{X}_1\hat{\boldsymbol{\beta}}_1)'(\mathbf{y}_1 - \mathbf{X}_1\hat{\boldsymbol{\beta}}_1)}{S}\right]^{-\frac{S}{2}}\left[\frac{(\mathbf{y}_2 - \mathbf{X}_2\hat{\boldsymbol{\beta}}_2)'(\mathbf{y}_2 - \mathbf{X}_2\hat{\boldsymbol{\beta}}_2)}{T-S}\right]^{-\frac{(T-S)}{2}}. \end{aligned} \tag{19}$$

Note that an option exists in the methodology that permits the estimation of common regime-specific variance terms ($\sigma_1 = \sigma_2$) as well. Therefore, under H_1, the consistent Akaike information criterion, CAIC(S), for $S \in (K+2, \ldots, T-K-2)$, is

$$\begin{aligned} \text{CAIC(S)} &= -2\log L_1(\hat{\boldsymbol{\beta}}_1,\ \hat{\boldsymbol{\beta}}_2, \hat{\sigma}_1^2, \sigma_2^2) + (2K+4)(\log(T)+1) \\ &= T(\log 2\pi + 1) + S\log\left[\frac{(\mathbf{y}_1 - \mathbf{X}_1\hat{\boldsymbol{\beta}}_1)'(\mathbf{y}_1 - \mathbf{X}_1\hat{\boldsymbol{\beta}}_1)}{S}\right] \\ &\quad + (T-S)\log\left[\frac{(\mathbf{y}_2 - \mathbf{X}_2\hat{\boldsymbol{\beta}}_2)'(\mathbf{y}_2 - \mathbf{X}_2\hat{\boldsymbol{\beta}}_2)}{T-S}\right] + (2K+4)(\log(T)+1). \end{aligned} \tag{20}$$

According to the principles of the information criterion in model selection, H_0 will be accepted if CAIC(T) $\leq \min_{K+2 \leq S \leq T-K-2}$ CAIC(S), and H_1 will be accepted if CAIC(T) $>$

$\min_{K+2 \le S \le T-K-2}$ CAIC(S). If H_1 is accepted, the estimated position of the change in the linear model will be $\hat{S}$ such that

$$\mathrm{CAIC}(\hat{S}) = \min_{K+2 \le S \le T-K-2} \mathrm{CAIC}(S). \tag{21}$$

Thus, for a specified number of change points $R^* \geq 1$, one selects the set of regressions that results in minimum CAIC.[1] Conditioned on R^*, CAIC is monotone with the sum of the error sum of squares in both regimes, and an equivalent (and much less computationally burdensome) approach is adopted to select the change points, given $R^* \geq 1$, which renders minimum total error sums of squares. The modification reduces the computational effort immensely, especially when programming in an interpreted matrix-oriented language such as APL, MATLAB, or Gauss. Such minimization is equivalent, given R^*, to selecting a solution that renders the largest confidence value in rejecting H_0 in the traditional Chow (1960) F test for structural change. Then, across these optimal $R^* = 1, 2, 3, \ldots$ solutions, one can use the min CAIC heuristic for determining the optimal number of change points. Given the ease of computation, a complete enumeration of all possible feasible, contiguous solutions, given R^*, is used. As such, a globally optimum solution is guaranteed for every regime analysis at very reasonable computational speeds.[2]

Consumer Psychology Application: Apartment Preference Analysis

Introduction to Stated Preference or Conjoint Analysis

Luce and Tukey (1964) published a seminal paper concerning the conditions under which measurement scales for both dependent and independent variables exist, given only (a) order information on the joint effects of the independent variables and (b) a hypothesized composition rule. This "conjoint measurement" idea further spawned several theoretical extensions and algorithmic contributions (cf. Carroll, 1969; Kruskal, 1965; Young, 1969). One of the areas that grew under this paradigm was called *conjoint analysis*, in which interest was focused primarily on parameter estimation and scaling in the context of the multiattribute preference measurement problem (cf. Green & Rao, 1971). That is, conjoint analysis is a "decompositional" multivariate technique used specifically to understand how respondents develop preferences for a specified stimulus (e.g., product, service, activity, apartments, etc.). According to Hair, Anderson, Tatham, and Black (1995), conjoint analysis is based on the simple premise that respondents evaluate the value or utility of a stimulus (real or hypothetical) by combining the separate amounts of utility provided by each attribute. Conjoint analysis is unique among multivariate methods in that the researcher first constructs a set of hypothetical stimuli by combining the selected levels of each attribute. These hypothetical stimuli are then presented to respondents, who provide *only* their overall evaluations. Respondents need not tell the researcher anything else, such as how important an attribute is to them or how well the stimulus performs on a number of attributes. Because the researcher constructs the hypothetical stimuli via efficient experimental designs, the importance of each attribute and each value of each attribute can be determined from the respondents' overall ratings.

The researcher must describe the stimuli in terms of both its attributes and the relevant values for each attribute. The term *factor* is used when describing a specific attribute or other characteristic of the stimuli. The possible values for each factor are called *levels*. In conjoint terms, a stimulus is described in terms of *its levels on the set of factors* characterizing it. When the researcher selects the factors and the levels to describe a stimulus according to a specific experimental design or plan, the combination is known as a *treatment* or *profile*.

Then, as described in Hair, et al. (1995), by constructing specific combinations (treatments or profiles), the researcher is attempting to understand a respondent's *preference structure*. The preference structure "explains" not only how important each factor is in the overall decision but also how differing levels within a factor influence the formation of an overall preference. This overall preference, which represents the total worth or utility of an object, can be thought of as based on the *part-worths* for each level. The general form of a conjoint model can be shown as follows:

$$\begin{aligned}\text{Total utility for stimulus}_{ij,\ldots,n} &= \text{Part worth of level}_i \text{ for factor}_1\\ &+ \text{Part worth of level}_j \text{ for factor}_2\\ &+ \cdots + \text{Part worth of level}_n \text{ for factor}_m,\end{aligned}$$

where the stimulus has m attributes, each having two or more levels (cf. Hair, et al., 1995). The profile consists of level$_i$ of factor$_1$, level$_j$ of factor$_2$, ..., up to level$_n$ of factor$_m$. Given the potential for the creation of literally millions of potential hypothetical profiles of stimuli, conjoint analysts make extensive use of highly fractionated factorial designs (cf. Addelman, 1962; Green, 1974) to reduce the number of profile descriptions to a small fraction of the total number of combinations while still allowing for main-effects estimation. Most often, metric ratings of overall preference are collected from each respondent for each of the profiles contained in such fractional factorial main-effects designs. Multiple regression is typically used for such conjoint analysis when the fractional design, converted either to dummy variables or effects codings, is used as the set of independent variables, and overall preference is the dependent variable. The estimated regression coefficients are the factor part-worths.

Study Design

The current study involves descriptions of privately offered, unfurnished student apartments located near a large, eastern university. Respondents for the experiment were undergraduate business students (juniors and seniors), most of whom were living in a student apartment or were contemplating renting one during the next school year. The context of this experiment was adapted from Johnson and Meyer (1984), Huber and Hansen (1986), and Green, Helsen, and Shandler (1988), who each performed stated preference (cf. Louviere, Hensler, & Swait, 2000) or conjoint analyses regarding student apartments. Based on this past research and a series of in-depth personal interviews with this particular student population, 10 attributes or factors were derived that these students claimed to be important in making their decisions for apartment selections. Table 1 lists these 10 factors as well as the three levels tested for each factor. As mentioned, most of these factors have appeared in past empirical stated preference/conjoint research on student apartments cited above.

A main-effects, fractional factorial design (as used in past research cited above) was implemented involving 27 profiles. Three additional profiles were formulated for predictive validation. Each conjoint profile contained a description of a hypothetical apartment available for rent where specific levels of all 10 factors were displayed according to the experimental design. The respondent was to use a 100-point preference scale to indicate his or her stated interest in actually renting such an apartment for the upcoming academic year. A total of 27 different orderings of the profiles were formulated, with each profile having a different order of presentation (from 1-27). In addition, the presentation of the three profiles was altered for validation to occur either prior to the conjoint task or immediately following it. Thus, $27 \times 2 = 54$ different questionnaires were used in this experiment. A student sample of $n = 162$ students was obtained, allowing

Table 1
Apartment 3^{10} Experimental Design

Factor	Level	Code	Effects	Dummy	
1. Rent	$200	1	−1	1	0
	$350	2	0	0	0
	$500	3	1	0	1
2. Distance	5 min.	1	−1	1	0
	13 min.	2	0	0	0
	21 min.	3	1	0	1
3. Size	1 bedroom	1	−1	1	0
	2 bedroom	2	0	0	0
	3 bedroom	3	1	0	1
4. Amenities	None	1	−1	1	0
	Cable TV or utilities	2	0	0	0
	Cable TV and utilities	3	1	0	1
5. Maintenance	Poor	1	−1	1	0
	Average	2	0	0	0
	Good	3	1	0	1
6. Condition	Poor	1	−1	1	0
	Average	2	0	0	0
	Newly renovated	3	1	0	1
7. Noise	Very quiet	1	−1	1	0
	Average	2	0	0	0
	Very noisy	3	1	0	1
8. Safety	Very unsafe	1	−1	1	0
	Average	2	0	0	0
	Very safe	3	1	0	1
9. Cleanliness	Very dirty	1	−1	1	0
	Average	2	0	0	0
	Very clean	3	1	0	1
10. Privacy	Unannounced inspections	1	−1	1	0
	1-day notice	2	0	0	0
	3- to 5-day notice	3	1	0	1

for three complete replications of these 54 questionnaires in which each student was randomly assigned to a questionnaire containing all $27 + 3$ profiles. Thus, for any/all particular sequence(s), the orthogonality of the fractional factorial design remains in place, ensuring efficient estimation. All questionnaires contained additional (identical) sections measuring memory, stated attribute importance, demographics, willingness to pay, budget, current rent levels, and current apartment descriptions.

Table 2

Aggregate Conjoint Regression Analysis: Effects Coding

Intercept: 45.45

Variable	Coefficient
Rent	−5.50**
Distance	−3.95**
Size	5.11**
Amenities	1.56**
Maintenance	3.33**
Condition	4.06**
Noise	−4.89**
Safety	6.15**
Cleanliness	3.97**
Privacy	0.73

ANOVA Table

Source	Sum of Squares	*df*	Mean Square	*F* Calculated
Regression	524862.60	10	52486.26	114.75
Error	1995648.63	4,363	457.40	
Total	2520511.23	4,373		

Standard error of the estimate = 21.39
$R^2 = .208$
Adjusted $R^2 = .206$

$^{**}p \leq .01$.

Traditional Aggregate-Level Preference Analysis

An analysis over the entire sample was first conducted at an aggregate level, using all 27 responses per respondent. Main-effects models were tested using both continuous effects and dummy variable coding. Based on the fact that the dummy variable coding increased R-square by less than 2% at a cost of nearly double the number of parameters, the simpler but more restrictive continuous-effects coding scheme (which treats the three levels of each factor as an interval scale) was retained because the levels within each factor were a priori constructed to be monotone. Table 2 presents the results of the aggregate regression analysis for these 10 factors. All independent design variables, except privacy, were significant, and all coefficient signs were in the anticipated direction. As shown by the magnitudes of the absolute values of the coefficients, safety, rent, size, noise, condition, and cleanliness are the most important apartment design variables in order of importance. Such trade-off analysis also suggests specific preference orderings. For example, as implied by Table 2, a high-rent, high-safety combination ($-5.50 + 6.15 = 0.65$) appears to be preferred to a low-rent, low-safety one ($-.065$), holding all other factor levels the same. A key issue, therefore, is whether this is a “stable” result or whether such comparisons change over the course of the study (e.g., rent increases in importance over the course of the conjoint experiment). Put differently, if respondents’ burden were restricted here to a few profiles (say, eight), would the resulting interpretation be the same as an analysis based on subsequent profiles?

Table 3

First Versus Last Profile Conjoint Regression Solutions

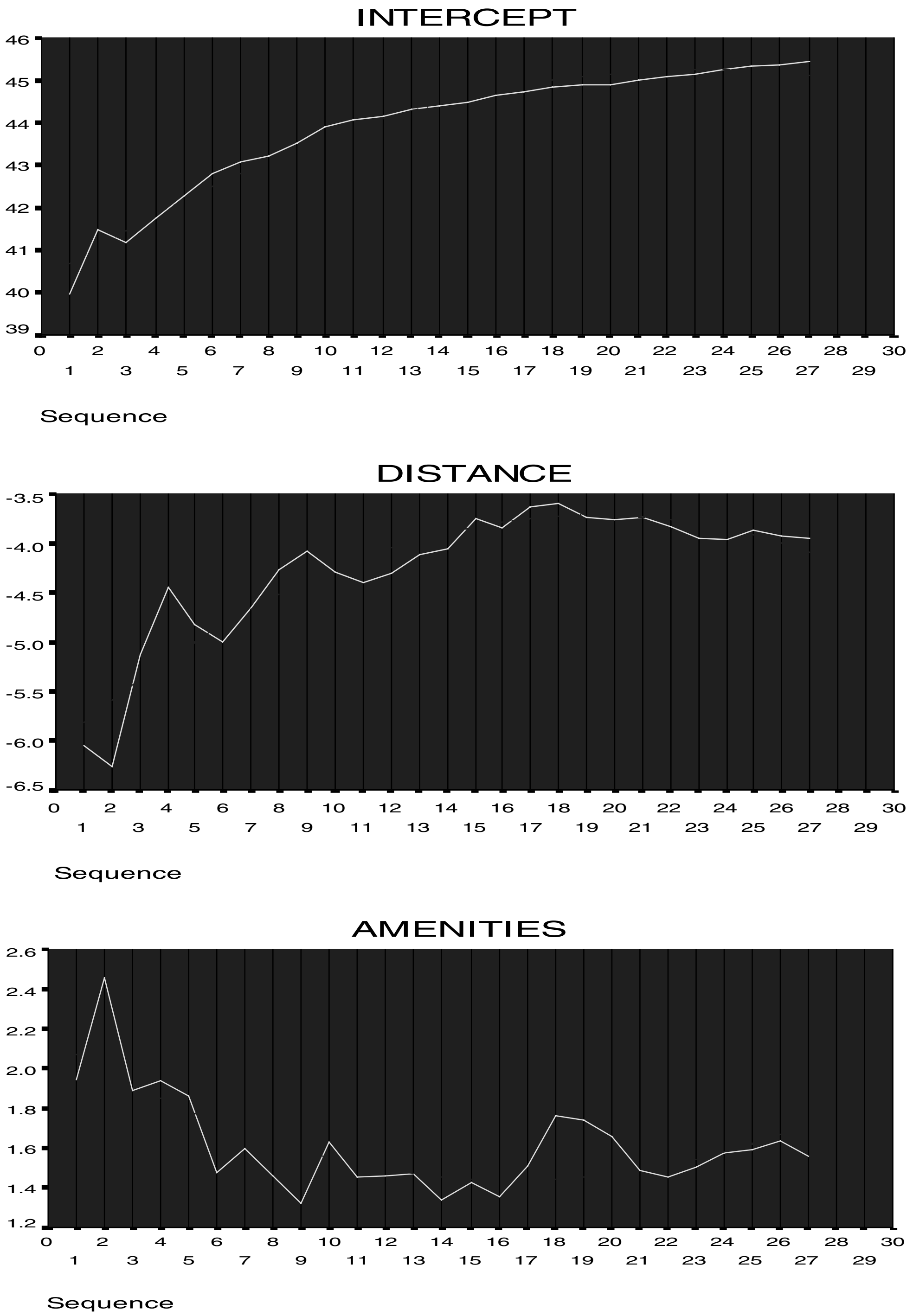
	First	Last	Aggregate
Intercept	39.95**	47.61**	45.45**
Rent	−5.18**	−6.08**	−5.50**
Distance	−6.05**	−4.48*	−3.95**
Size	6.37**	5.63**	5.11**
Amenities	1.94	−0.41	1.56**
Maintenance	3.30	−2.39	3.33**
Condition	2.94	2.00	4.06**
Noise	−6.94**	−3.35	−4.89**
Safety	4.60*	4.65*	6.15**
Cleanliness	7.83**	5.39**	3.96**
Privacy	0.62	−2.05	0.73
Standard error	20.76	21.40	21.39
F	6.47**	3.89**	114.75**
Adjusted R^2	.25	.15	.21

$^{*}p \leq .05$.
$^{**}p \leq .01$.

Note that given the design of this particular experiment, part-worths can be estimated at each of the 27 stages of the experiment (after each profile is completed) separately at the aggregate level. For example, Table 3 displays the estimated part-worths from the responses collected in Step 1 (first profile) and Step 27 (last profile). Some substantial changes are seen in these part-worths compared to each other and from those obtained from the aggregate analysis in Table 2. Part of this variability reflects sampling error, small numbers of observations per analysis, and perhaps the design itself. Figure 1 represents the estimations performed in a cumulative fashion in which the observations from previous steps are included in a particular stage of analysis. Uniformly, an erratic sequence or time series is initially observed, which then starts to converge to some stationary level. It is clear from these two analyses that changes occur in the part-worths (especially early in the task) as the sample progresses through the task. Part of this may be due to fatigue, part may be due to sample participants learning their preference structures as they accumulate experience, part may be due to simplifications in the response patterns to expedite the conjoint task (e.g., focusing on fewer attributes), and part may be a result of more data being cumulated in the analyses. As shown in Table 3, the factor importances change in order when computed from Step 1 versus Step 27. In Step 1, the top five factors in order of importance were as follows: (1) cleanliness, (2) noise, (3) size, (4) distance to class, and (5) rent. In Step 27, that order changes to the following: (1) rent, (2) size, (3) cleanliness, (4) distance to class, and (5) safety. There appears to be a change toward stressing more concrete and quantitative attributes (rent and size). Again, only the first and last profiles are examined here. These orders both differ from that obtained from the total or complete analysis computed over all collected 27 profiles (Table 2), in which the order was the following: (1) safety, (2) rent, (3) size, (4) noise, and (5) condition.

Thus, even where the signs or relative sizes of the part-worth estimates for significant factors do not change dramatically, there is a noticeable change in the order of importance for the various design factors. Obviously, because these analyses were all done at the aggregate sample level, there is little insight as to the dynamics of learning, fatigue, response patterns, order bias, and so on

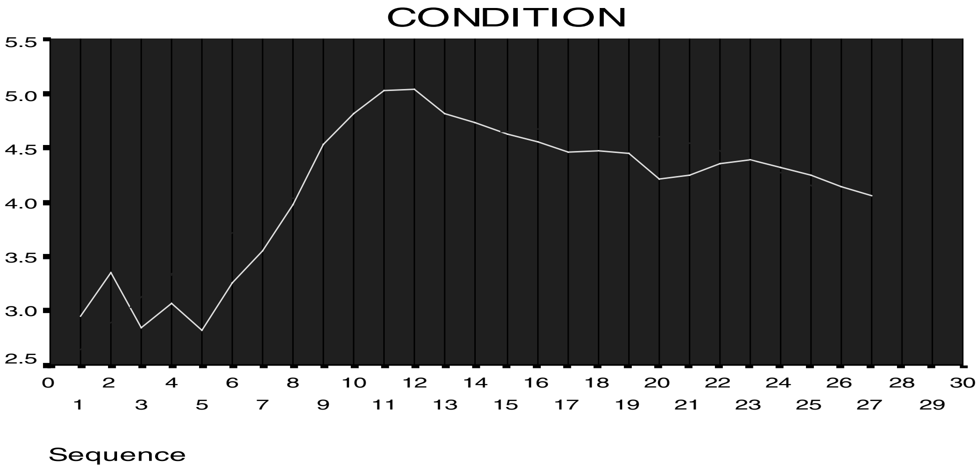
Figure 1
Cumulative Effects Coding Regression Part-Worths

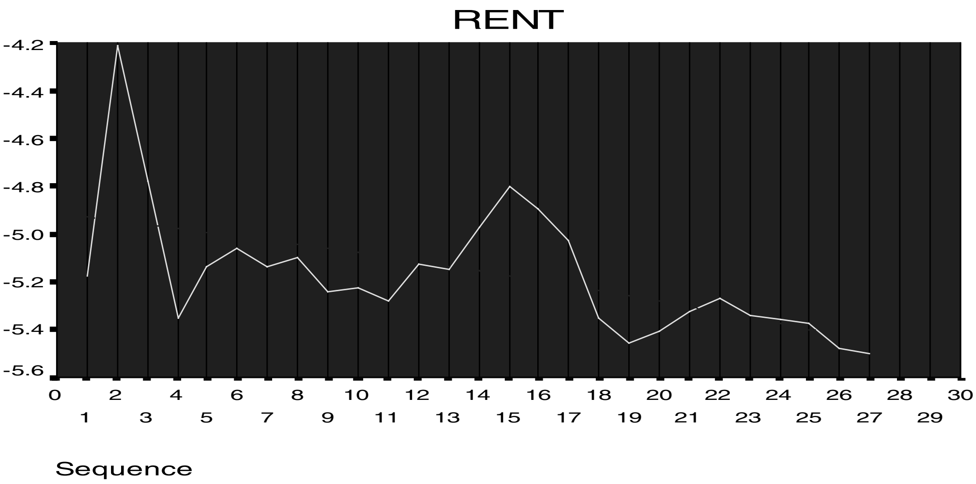


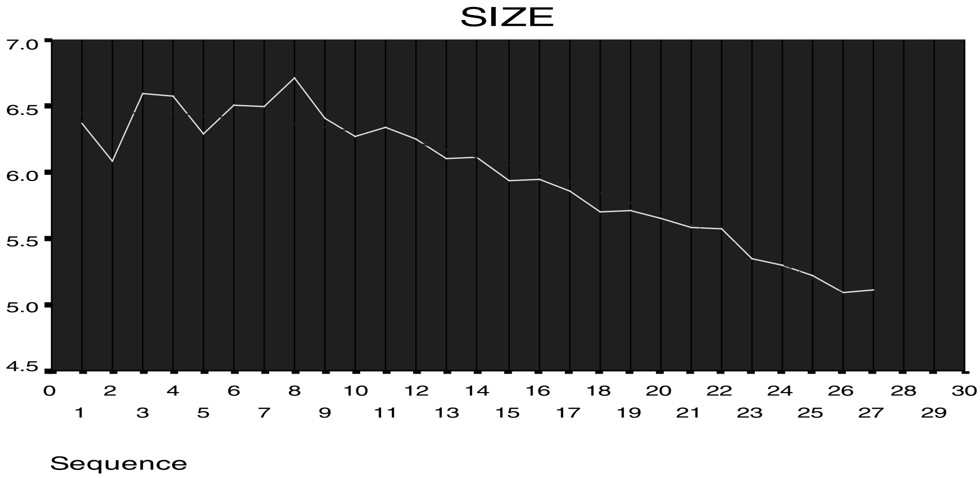
(continued)

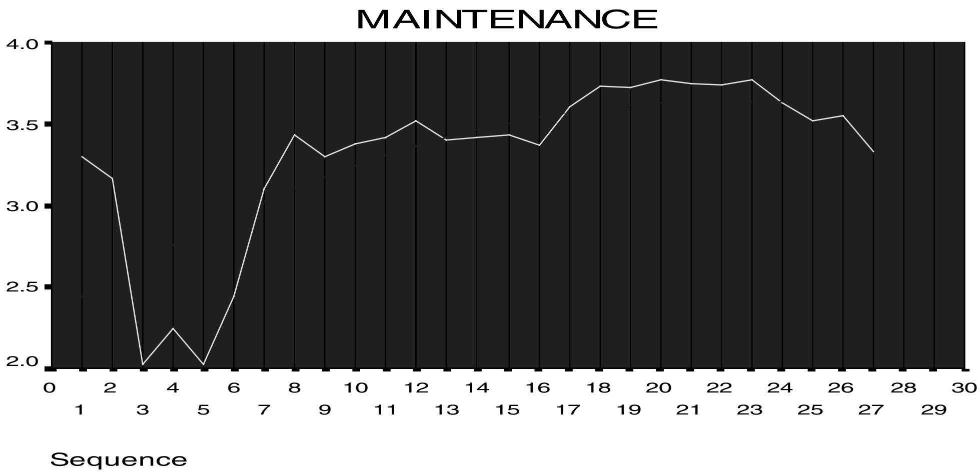
Figure 1

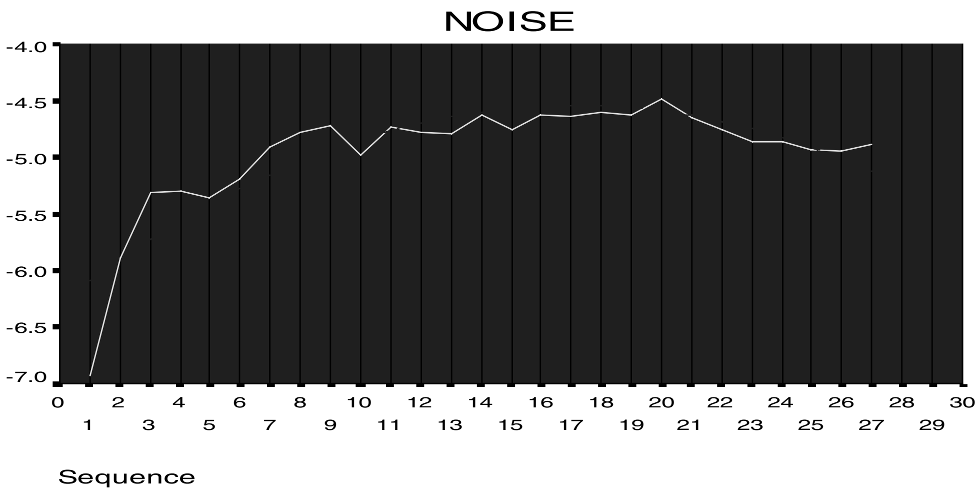
(continued)







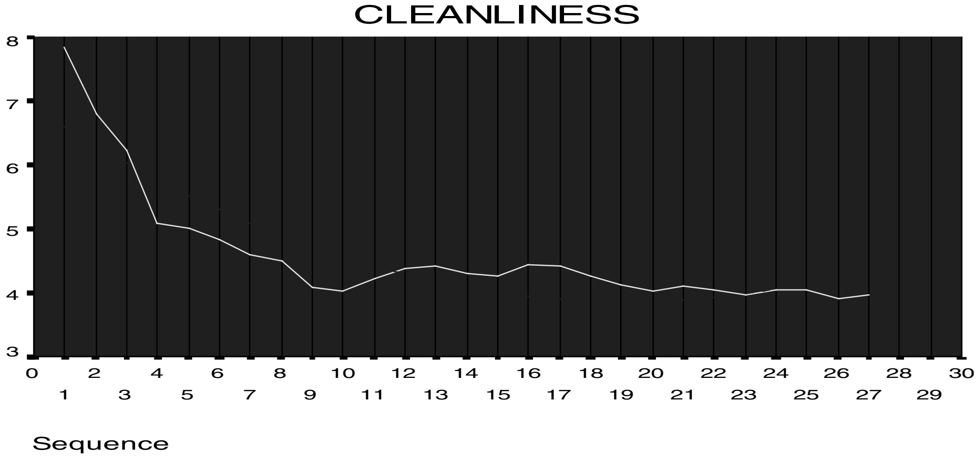


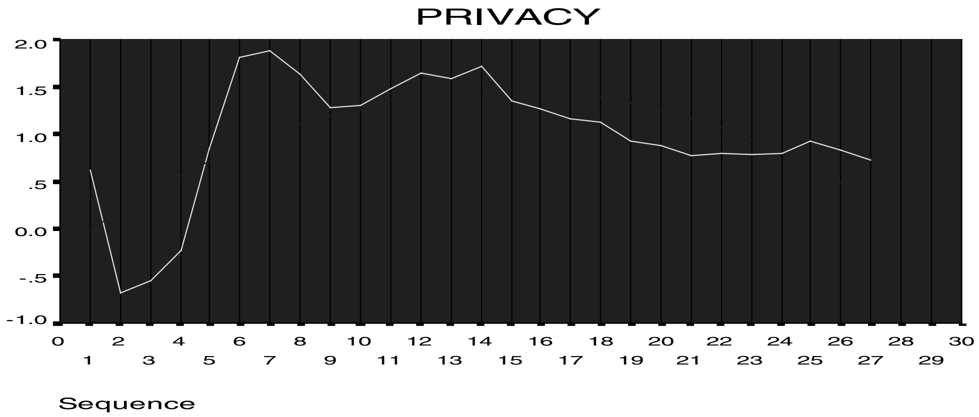


(continued)

Figure 1

(continued)





Note. The vertical axis is identified by the label at the top of each graph.

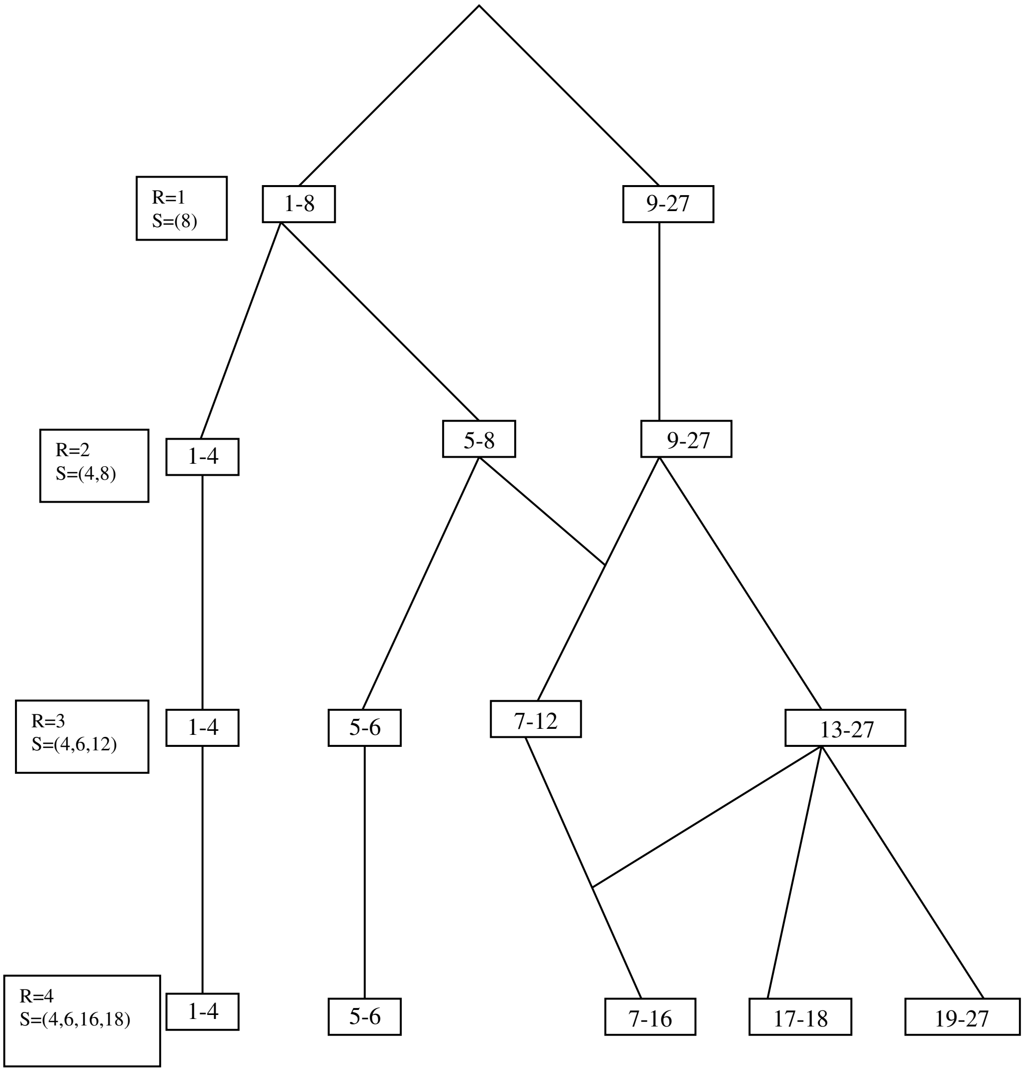
that may be occurring at the individual level. However, the location and duration of any *significant* structural changes are now examined in the aggregate utility function that might occur as these respondents progress through this task.

Change-Point Multiple Regression Analysis

The change-point multiple regression methodology is applied next, as developed in Section 2, to the 27 conjoint profiles' responses. Here, different regimes are allowed for that involve differences in both part-worths and variances. Initially, the optimal splits (i.e., those that provide minimum CAIC given R) of the sequence of profiles for sequential values of R (from 0-4 change points) are tracked. Figure 2 presents a graphical decomposition of this profile segmentation for various optimal splits (recall that $R =$ the number of change points, $R + 1 =$ the number of splits or regimes, and $S =$ the locations of the change points). As depicted in Figure 2, the sequence of splits appears hierarchical or nested until $R + 1 = 5$ splits/regimes for this application. What is

Figure 2

Recursive Regime Splits for Conjoint Application



of particular interest here is the fact that most of the splits or regimes for $R = 1, 2$, and 3 occur early in the sequence of profiles, suggesting that adaptations in the aggregate utility function occur early in the task. This may reflect their adaptation or learning as they familiarize themselves with this full-profile conjoint task. That is, it may take some experience before respondents are able to calibrate the heuristics they use to complete the task.

Using the CAIC to select the most parsimonious solution over all optimal splits presented in Figure 2, the $R + 1 = 2$ split/regime solution provides the minimum CAIC[3] value. The first regime contains the data from Profiles 1 through 8, whereas the second regime contains the remaining sequence of Profiles 9 through 27. Table 4 provides the optimal two-regime solution in comparison to the aggregate regression analysis computed over all 27 profiles. Note that the computed F statistic

Table 4

The Two-Regime Optimal Conjoint Regression Solution

	Regime 1	Regime 2	Aggregate
Intercept	43.20**	46.40**	45.45**
Rent	−5.10**	−5.67**	−5.50**
Distance	−4.27**	−3.81**	−3.95**
Size	6.71**	4.44**	5.11**
Amenities	1.46*	1.60**	1.56**
Maintenance	3.43**	3.29**	3.33**
Condition	3.97**	4.10**	4.06**
Noise	−4.78**	−4.93**	−4.89**
Safety	4.75**	6.73**	6.15**
Cleanliness	4.49**	3.74**	3.96**
Privacy	1.63*	0.34	0.73
Standard error	21.70	21.16	21.39
F	34.35**	82.76**	114.75**
Adjusted R^2	.20	.21	.21

$^{*}p \leq .05$.
$^{**}p \leq .01$.

for the Chow test of structural change ($R = 0$ vs. $R = 1$) is 3.34, which is significant at $p < .01$. (Theoretically, this test is questionable because the $R = 1$ optimal split involved a pretest analysis of the same data.)

A number of observations are appropriate here. First, there is a somewhat higher propensity or preference for apartment profiles in general as respondents progress through the response task. This is reflected in the different intercept values between the two regimes. Second, the relative order of importance of the conjoint factors changes between the two regimes. This was noted earlier in comparing the first versus last profile in Table 2. In Regime 1, size, rent, noise, safety, and cleanliness are the top five (in order) important factors. In Regime 2, the order is safety, rent, noise, size, and condition. This suggests an adaptation or evolution of the utility function as respondents progress through this task, with safety increasing and size decreasing in importance. As such, the implicit preferences and resulting trade-offs among profiles change. For example, for the first regime of eight profiles, a low-rent, low-safety option (+0.35) would be preferred to a high-rent, high-safety one (−0.35). By contrast, for the second regime involving the later observations, the high-rent, high-safety option would be preferred (+1.06), holding all other factor levels constant. Thus, the managerial implications of the conjoint analysis differ depending on which estimated regime is used for decision making. Third, the coefficient estimates from the second regime are much closer to that of the aggregate utility function in sign and magnitude (as the second regime contains more than two thirds of the data points). Finally, there are sharper spikes witnessed in the computed factor importances (higher peaks/larger coefficient magnitudes) for those part-worths estimated in the first regime.

Validation

Recall that each respondent was given three holdout profiles for validation that were not included in the estimation sample. Half of the respondents (**FIRST**) were given these profiles as the first three profiles they received, whereas the other half (**LAST**) received the validation profiles as the

Table 5

The Two-Regime Optimal Conjoint Regression Validation

	Regime 1	Regime 2	Aggregate
FIRST	21.09	21.38	21.28
LAST	17.45	16.79	16.97

Note. Numbers represent average mean square error.

last three conjoint profiles to evaluate. Table 5 presents the average mean square errors (MSE) in predictive validation for each of these two groups of respondents (**FIRST, LAST**) over the three holdout validation profiles. The results are examined in a two-way table in which the mean square errors are calculated for each group using one of three sets of estimated part-worth functions: those computed over Regime 1, Regime 2, and the aggregate sample. One would expect that more accurate predictions would occur in using the part-worth coefficients computed from data closer in sequence to when the validation occurred. There is a slight improvement in MSE for the **FIRST** group when Regime 1's coefficients were used. A more pronounced improvement is seen with respect to the **LAST** group in using the coefficients from Regime 2. The aggregate function is always in the middle in terms of MSE in this predictive validation, as expected. What is most striking concerning the MSEs in Table 5 is the stark contrast in magnitudes of these error rates across the different rows (**FIRST** vs. **LAST**) of the table. That is, it appears that there is smaller prediction error in predictive validation when the validation occurs after the conjoint task. This suggests that respondents may require a "burn-in" period to properly calibrate their utility function, that these first eight observations do not help prediction, and/or that later responses are artificially more consistent in terms of response structure as a result of respondents adopting a simple and consistent process as they progress through the response task.

Discussion

This article has proposed a method for assessing whether significant changes occur in the parameters of a multiple regression model over a sequence of data points. Such changes might occur within a consumer survey due to order effects, learning, boredom, fatigue, and so on. The authors demonstrate the method with a stated preference or conjoint study that, through randomization, controls for order effects. Even in these data, evidence is found for two distinct sets of parameters over the sequence of data.

Several implications emerge. First, it will often be advisable to check for changes in parameters and, to the extent these are substantial, adjust interpretations and forecasting accordingly. In the specific case of conjoint analysis, deciding which set of parameters to use is less clear. Parameters from the end of a sequence of responses may represent greater thought and expertise with the choice options themselves or merely an adaptation to the survey and represent an efficient method for responding to questions. On the other hand, early responses could represent "true" reactions to low-involvement choices or just random responses as the respondent learns to think about the alternatives. Put differently, it may be that experts and highly involved respondents give their most accurate responses early (before fatigue sets in) and novices and less involved respondents their most useful responses late in the sequence, or at least in the middle after they develop a defined preference structure. In any event, a clear direction for future research is to investigate which responses are most predictive for different categories of respondents, as well as to understand what mechanisms (e.g., boredom, fatigue, learning, etc.) are at work in producing such dynamic effects.

Another direction worth pursuing is to develop a general typology of which attributes become more (or less) important through studies in terms of either specific attributes (e.g., price) or general type (e.g., concrete vs. abstract). More detailed estimation is required on either an individual or market segment level. An aggregate solution as presented here may not be appropriate in the presence of severe heterogeneity or alternative utility structures. Finally, further validation is required in hopes of identifying application guidelines as how to properly weight the derived regimes in obtaining more accurate forecasting models.

On the methodological side, several enhancements to the current procedure are possible. One of the clear limitations of the current work stems from the fact that respondent heterogeneity is not explicitly modeled in the proposed framework. Theoretically, one could argue for individual-level coefficients and/or regimes. Although such detailed analyses in a classical statistical framework are implausible due to overparameterization and limited degrees of freedom, other potential intermediate solutions are plausible concerning this issue of respondent heterogeneity. Given the similarity of the present data collected in such conjoint studies to panel data involving both cross-sectional (here, different respondents) and longitudinal (here, the sequence of profiles completed), one could theoretically fit a random-effects model (or, more generally, a fully random-coefficients model) in which the intercept (and part-worth coefficient) varies across respondents (see Diggle, Heagerty, Liang, & Zeger, 2002; Lindsey, 1999). If respondent-level covariates are available with a theory as to how they affect the conjoint part-worth functions, one might also employ what Greene (2003) calls "hierarchical models" to attempt to explain the specific nature of respondent heterogeneity. Unfortunately, these types of more complex models involve much more difficulty in terms of computational complexity. Although these can all be cast in a maximum likelihood framework à la generalized least squares, the associated computational effort increases dramatically, especially when the elements of the error covariance matrix are unknown. In addition, difficulties involving sufficient data points evolve when the number of regimes is large and the subsequent sequence of observations become sparse per regime.

A related potential extension of the proposed methodology would be to consider general linear mixed models that would remove the assumption of the independent errors and allow for correlated errors in repeated measures in such conjoint analysis applications. Laird, Lange, and Stram (1987) introduce an EM-based algorithm for cases involving serial measurements. Verbeke and Molenberghs (2000) present an excellent review of such mixed linear models for longitudinal data as well as applications using SAS PROC MIXED. Note that no one has yet considered the case of regime switching in such longitudinal data using these mixed linear models. Obviously, one starts to worry about the number of parameters to be estimated when covariance matrices by regime are to be estimated, in addition to supplementary parameters for respondent heterogeneity, as mentioned above. Also, there are computational issues because many of the proposed EM procedures for this type of regime estimation would involve iterative estimation procedures with only linear convergence properties and susceptibility to local optimum solutions.

As fertile ground for future research, one might also consider the application of dynamic linear/nonlinear models for such repeated-measures conjoint analysis applications. Lindsey (1999) discusses the use of Markov autoregressive models, Kalman filtering, and recent Bayesian dynamic models for such "panel"-type data. Such representations would provide a more continuous evolution or path of the part-worth functions over the sequence of profiles as opposed to the discrete dynamic representation provided with the proposed methodology (this discrete representation of such dynamics approaches the continuous or evolutionary approach as the number of regimes increases). Unfortunately, there is usually insufficient theory available to appropriately specify a particular nonlinear structure, especially in the multivariate situation such as this study's application, in which some 10 independent variables are present.

In addition, extensions to other forms of multivariate analyses are possible. For example, such change point models can be extended to structural equation models in which different theoretical structures may hold over different portions of the data. Alternatively, change-point multidimensional scaling models can be constructed where one expects an evolutionary process to occur with respect to evoked perceptions and preferences. In these contexts, comparisons with "hidden Markov" switching models and dynamic linear models can be devised using more complex Bayesian schemes (cf. Carlin, 1992; DeSarbo, Hollman, Fong, & Leichty, 2001; Henderson & Matthews, 1993; Stephens, 1994) and more complicated estimation routines (e.g., hierarchical Bayes methods) that allow for individual-level estimation. More experimental work is needed to identify optimal stopping rules in conjoint analysis, as well as to locate and estimate optimal part-worths. Monte Carlo simulations with synthetically constructed data with known structures are needed to investigate the performance of the proposed methodology as a number of different factors are experimentally varied (e.g., number of regimes, sample size, amounts and types of error, heterogeneity across respondents, etc.). Finally, extending the proposed procedure to choice-based conjoint analysis, hybrid conjoint analysis, and self-explicated preference analysis would be also desirable.

Notes

1. Any form of the GIC class of heuristics in equation (11) can be so optimized with use of this complete enumeration-based procedure.
2. In the proposed methodology, computational speed varies according to the size of R. For the application to be discussed, global optimum estimates were all obtained for one, two, three, four, and five regimes under 20 minutes each.
3. The number of observations here reflects both the number of respondents and the number of conjoint profiles as an adjustment to the CAIC formula developed in Section 2.

References

Addelman, S. (1962). Orthogonal main-effect plans for asymmetrical factorial experiments. *Technometrics*, *4*, 21-46.

Anderson, J. R. (1981). *Foundations of information integration theory*. San Diego: Academic Press.

Anderson, J. R. (1983). *The architecture of cognition*. Cambridge, MA: Harvard University Press.

Andrews, D. W. K. (1993). Tests for parameter instability and structural change with unknown change point. *Econometrica*, *62*, 1383-1414.

Bettman, J. R. (1979). *An information processing theory of consumer choice*. Reading, MA: Addison-Wesley.

Bettman, J. R., & Park, W. C. (1980). Effects of prior knowledge and experience and phase of the choice process on consumer decision processes: A protocol analysis. *Journal of Consumer Research*, *7*, 234-248.

Bettman, J. R., & Zins, M. A. (1979a). Constructive processes in consumer choice. *Journal of Consumer Research*, *4*, 75-85.

Bettman, J. R., & Zins, M. A. (1979b). Information format and choice task effects in decision making. *Journal of Consumer Research*, *6*, 141-153.

Bozdogan, H. (1987). Model selection and Akaike's information criterion (AIC): The general theory and its analytical extensions. *Psychometrika*, *52*, 345-370.

Bozdogan, H. (1994). Mixture model cluster analysis using model selection criteria and a new informational measure of complexity. In H. Bozdogan (Ed.), *Multivariate statistical modeling* (Vol. 2, pp. 69-113). Dordrecht, The Netherlands: Kluwer Academic.

Brown, R. L., Durbin, J., & Evans, J. M. (1975). Techniques for testing the constancy of regression relationships over time. *Journal of the Royal Statistical Society*, *B37*, 149-192.

Burnham, K. P., & Anderson, D. R. (2002). *Model selection and multimodal inference* (2nd ed.). New York: Springer-Verlag.

Carlin, B. P. (1992). Hierarchical Bayesian analysis of change point problems. *Applied Statistics*, *41*, 389-405.

Carroll, J. D. (1969). Categorical conjoint measurement. Unpublished manuscript. Murray Hill, N. J.: Bell Laboratories.

Chen, J., & Gupta, A. K. (2000). *Parametric statistical change point analysis*. Boston: Birkhäuser.

Chow, G. (1960). Tests of the equality between two sets of coefficients in two linear regressions. *Econometrica*, *28*, 561-605.

Coupey, E., Irwin, J. R., & Payne, J. W. (1998). Product category familiarity and preference construction. *Journal of Consumer Research*, *24*, 459-468.

DeSarbo, W. S., & Cron, W. L. (1988). A conditional mixture maximum likelihood methodology for clusterwise linear regression. *Journal of Classification*, *5*, 249-289.

DeSarbo, W. S., Hollman, F., Fong, D., & Leichty, J. (2001). *Evolutionary consumer preference/utility functions: A dynamic perspective*. Working paper, Pennsylvania State University.

Diggle, P. J., Heagerty, P., Liang, K., & Zeger, S. L. (2002). *Analysis of longitudinal data* (2nd ed.). Oxford, UK: Oxford University Press.

Einhorn, H. J., & Hogarth, R. M. (1985). Ambiguity and uncertainty in probabilistic inference. *Psychological Review*, *92*, 433-461.

Farley, J. U., & Hinich, M. J. (1970). A test for a shifting slope coefficient in a linear model. *Journal of the American Statistical Association*, *65*, 1320-1329.

Feldman, J. M., & Lynch, J. G., Jr. (1988). Self-generated validity and other effects of measurement on belief, attitude, intention, and behavior. *Journal of Applied Psychology*, *73*, 421-435.

Fischhoff, B., Slovic, P., & Lichtenstein, S. (1980). Knowing what you want: Measuring labile values. In T. Wallstein (Ed.), *Cognitive processes in choice and decision behavior* (pp. 117-141). Hillsdale, NJ: Lawrence Erlbaum.

Gilbert, D. T. (1989). Thinking lightly about others: Automatic components of the social inference process. In J. S. Uleman & J. A. Bargh (Eds.), *Unintended thought* (pp. 189-211). New York: Guilford.

Gilbert, D. T., Krull, D. S., & Pelham, B. W. (1988). Of thoughts unspoken: Social inference and the self-regulation of behavior. *Journal of Personality and Social Psychology*, *55*, 685-694.

Goldfield, S. M., & Quandt, R. E. (1973). A Markov model for switching regression. *Journal of Econometrics*, *1*, 3-16.

Green, P. E. (1974). On the design of choice experiments involving multifactor alternatives. *Journal of Consumer Research*, *1*, 61-68.

Green, P. E., Helsen, K., & Shandler, B. (1988). Conjoint internal validity under alternative profile presentations. *Journal of Consumer Research*, *15*, 392-397.

Green, P. E., & Rao, V. R. (1971). Conjoint measurement for quantifying judgmental data. *Journal of Marketing Research*, *8*, 355-363.

Greene, W. H. (2003). *Econometric analysis* (5th ed.). Upper Saddle River, NJ: Prentice Hall.

Gustafsson, F. (2000). *Adaptive filtering and change detection*. New York: John Wiley.

Hair, J. E., Anderson, R. E., Tatham, R. L., & Black, W. C. (1995). *Multivariate data analysis* (4th ed.). Englewood Cliffs, NJ: Prentice Hall.

Hamilton, J. D. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica*, *57*(2), 357-384.

Hamilton, J. D. (1990). Analysis of time series subject to changes in regime. *Journal of Econometrics*, *45*, 39-70.

Hanssens, D. M., Parsons, L. J., & Schultz, R. L. (2001). *Market response models: Econometric and time series analysis*. Boston: Kluwer.

Hayes-Roth, B. (1977). Evolution of cognitive structures and processes. *Psychological Review*, *84*, 260-278.

Henderson, R., & Matthews, J. N. S. (1993). An investigation of change points in the annual number of cases of haemolytic uraemic syndrome. *Applied Statistics*, *42*, 461-471.

Herr, P. M. (1989). Priming price: Prior knowledge and context effects. *Journal of Consumer Research*, *16*, 67-75.

Huber, J., & Hansen, D. (1986). Testing the impact of dimensional complexity and affective differences of paired concepts in adaptive conjoint analysis. In M. W. Wallendorf & P. Anderson (Eds.), *Advances in consumer research* (Vol. 14, pp. 159-163). Provo, UT: Association for Consumer Research.

Huber, J., Wittink, D. R., Johnson, R. M., & Miller, R. (1992). Learning effects in preference tasks: Choice-based versus standard conjoint.

In *Sawtooth Software conference proceedings*. Sequim, WA: Sawtooth Software, Inc.

Jain, A. K., & Pinson, C. (1976). The effects of order of presentation of similarity judgments on multidimensional scaling results: An empirical examination. *Journal of Marketing Research*, *13*, 435-439.

Johar, G. V., Jedidi, K., & Jacoby, J. (1997). A varying-parameter averaging model of on-line brand evaluations. *Journal of Consumer Research*, *24*, 232-246.

Johnson, E. J., & Meyer, R. J. (1984). Compensatory choice models of noncompensatory processes: The effect of varying context. *Journal of Consumer Research*, *11*, 528-541.

Johnson, M. D., Lehmann, D. R., & Horne, D. R. (1990). The effects of fatigue on judgments of interproduct similarity. *International Journal of Research in Marketing*, *7*(1), 35-43.

Johnson, R. M., & Orme, B. K. (1996). *How many questions should you ask in choice-based conjoint studies?* Technical report, Sawtooth Software, Inc., Sequim, WA.

Kahneman, D., Slovic, P., & Tversky, A. (Eds.). (1982). *Judgment under uncertainty: Heuristics and biases*. Cambridge, UK: Cambridge University Press.

Kardes, F. R., & Kalyanaram, G. (1992). Order-of-entry effects on consumer memory and judgment: An information integration perspective. *Journal of Marketing Research*, *29*, 343-357.

Kardes, F. R., & Snell, J. (1990). Predicting utility. In R. M. Hogarth (Ed.), *Insights in decision making: A tribute to Hillel J. Einhorn* (pp. 295-310). Chicago: University of Chicago Press.

Kim, C.-J., & Nelson, C. R. (1999). *State-space models with regime switching*. Cambridge, MA: MIT Press.

Kim, H. J., & Siegmund, D. (1989). The likelihood ratio test for a change point in simple linear regression. *Biometrika*, *76*, 409-423.

Kim, I.-M., & Maddala, G. S. (1991). *Multiple structural breaks and unit roots in the nominal and real exchange rates*. Unpublished manuscript, University of Florida, Department of Economics.

Kruskal, J. B. (1965). Analysis of factorial experiments by estimating monotone transformations of the data. *Journal of the Royal Statistical Society, Series B*, *27*, 251-263.

Laird, N., Lange, N., & Stram, D. (1987). Maximum likelihood computations with repeated measures: Application of the EM algorithm. *Journal of the American Statistical Society*, *82*, 97-105.

Leeflang, P. S. H., Wittink, D. R., Wedel, M., & Naert, P. (2000). *Building models for marketing decisions*. Boston: Kluwer.

Lindsey, J. K. (1999). *Models for repeated measurements* (2nd ed.). Oxford, UK: Oxford University Press.

Lopes, L. L. (1982). *Toward a procedural theory of judgment*. Madison: University of Wisconsin, Department of Psychology, Wisconsin Information Processing Program.

Louviere, J. J., Hensher, D. A., & Swait, J. D. (2000). *Stated choice models: Analysis and application*. Cambridge, UK: Cambridge University Press.

Luce, R. D., & Tukey, J. W. (1964). Simultaneous conjoint measurement: A new type of fundamental measurement. *Journal of Mathematical Psychology*, *1*, 1-127.

March, J. G. (1978). Bounded rationality, ambiguity, and the engineering of choice. *Bell Journal of Economics*, *9*, 587-608.

Meyer, R. J. (1987). The learning of multiattribute judgment policies. *Journal of Consumer Research*, *14*, 155-173.

Moore, D. A. (1999). Order effects in preference judgments: Evidence for context dependence in the generation of preferences. *Organizational Behavior and Human Decision Processes*, *78*, 146-165.

Neftci, S. N. (1984). Are economic time series asymmetric over the business cycle? *Journal of Political Economy*, *92*(2), 306-328.

Neter, J., Kutner, M. H., Nachtsheim, C. J., & Wasserman, W. (1996). *Applied linear statistical models* (4th ed.). New York: McGraw-Hill.

Payne, J. W., Bettman, J. R., & Johnson, E. J. (1992). Behavioral decision research: A constructive processing perspective. *Annual Review of Psychology*, *43*, 87-131.

Payne, J. W., Bettman, J. R., & Schkade, D. (1999). Measuring constructed preferences: Towards a building code. *Journal of Risk and Uncertainty*, *19*(1-3), 243-270.

Ploberger, W., Kramer, W., & Kontrus, K. (1989). A new test for structural stability in the linear regression model. *Journal of Econometrics*, *40*, 307-318.

Quandt, R. E. (1958). The estimation of the parameters of a linear regression system obeying two separate regimes. *Journal of the American Statistical Association*, *53*, 873-880.

Quandt, R. E. (1960). Tests of the hypothesis that a linear regression system obeys two separate regimes. *Journal of the American Statistical Association*, *55*, 324-330.

Quandt, R. E. (1972). A new approach to estimating switching regressions. *Journal of the American Statistical Association*, *67*, 306-310.

Sclove, S. L. (1983). Time-series segmentation: A model and a method. *Information Sciences*, *29*, 7-25.

Shank, R., & Abelson, R. (1977). *Scripts, plans, goals, and understanding: An inquiry into human knowledge structures*. Hillsdale, NJ: Lawrence Erlbaum.

Simonson, I., & Tversky, A. (1992). Choice in context: Tradeoff contrast and extremeness aversion. *Journal of Marketing Research*, *29*, 281-295.

Stephens, D. A. (1994). Bayesian retrospective multiple change point identification. *Applied Statistics*, *44*, 159-178.

Taylor, S. E., & Crocker, J. (1981). Schematic bases of social information processing. In E. T. Higgins, P. Herman, & M. Zanna (Eds.), *Social cognition: The Ontario Symposium* (Vol. 1). Hillsdale, NJ: Lawrence Erlbaum.

Tversky, A., & Simonson, I. (1993). Context dependent preferences. *Management Science*, *39*(10), 1179-1189.

Verbeke, G., & Molenberghs, G. (2000). *Linear mixed models for longitudinal data*. New York: Springer-Verlag.

von Eye, A., & Schuster, C. (1998). *Regression analysis for social science*. New York: Academic Press.

von Neumann, J., & Morgenstern, O. (1947). *Theory of games and economic behavior*. Princeton, NJ: Princeton University Press.

Wecker, W. E. (1979). Predicting the turning points of a time series. *Journal of Business*, *52*(1), 35-50.

Wildt, A. R., & Winer, R. S. (1983). Modeling and estimation in changing market environments. *Journal of Business*, *56*(3), 365-388.

Yao, Y. C. (1988). Estimating the number of change points by Schwartz's criterion. *Statistics and Probability Letters*, *6*, 181-187.

Young, F. W. (1969). *Polynomial conjoint analysis of similarities: Definitions for a special algorithm* (Research Paper No. 76). Chapel Hill: University of North Carolina, Psychometric Laboratory.

Acknowledgments

The authors wish to acknowledge and thank the editor and three anonymous reviewers whose excellent comments helped improve the quality of our manuscript.

Author's Address

Address correspondence to Dr. Wayne S. DeSarbo, Marketing Department, Smeal College of Business, Pennsylvania State University, University Park, PA 16802; phone: (814) 865-8150; fax: (814) 865-3015; e-mail: desarbows@aol.com.