## Marketing Science

## Dynamic Models Incorporating Individual Heterogeneity: Utility Evolution in Conjoint Analysis

John C. Liechty, Duncan K. H. Fong, Wayne S. DeSarbo,

Please scroll down for article—it is on subsequent pages

# Dynamic Models Incorporating Individual Heterogeneity: Utility Evolution in Conjoint Analysis

## John C. Liechty
Marketing Department of Statistics Department, Smeal College of Business, 701 M BAB, Pennsylvania State University, University Park, Pennsylvania 16802, jcl112@psu.edu

## Duncan K. H. Fong
Marketing Department and Statistics Department, Smeal College of Business, 707 G BAB, Pennsylvania State University, University Park, Pennsylvania 16802, i2v@psu.edu

## Wayne S. DeSarbo
Marketing Department, Smeal College of Business, 701 D BAB, University Park, Pennsylvania 16802, desarbows@aol.com

It has been shown in the behavioral decision making, marketing research, and psychometric literature that the structure underlying preferences can change during the administration of repeated measurements (e.g., conjoint analysis) and data collection because of effects from learning, fatigue, boredom, and so on. In this research note, we propose a new class of hierarchical dynamic Bayesian models for capturing such dynamic effects in conjoint applications, which extend the standard hierarchical Bayesian random effects and existing dynamic Bayesian models by allowing for individual-level heterogeneity around an aggregate dynamic trend. Using simulated conjoint data, we explore the performance of these new dynamic models, incorporating individual-level heterogeneity across a number of possible types of dynamic effects, and demonstrate the derived benefits versus static models. In addition, we introduce the idea of an unbiased dynamic estimate, and demonstrate that using a counterbalanced design is important from an estimation perspective when parameter dynamics are present.

*Key words*: heterogeneity; empirical utility functions; dynamic models; Bayesian analysis; conjoint analysis; unbiased dynamic estimates
*History*: This paper was received November 9, 2001, and was with the authors 21 months for 3 revisions; processed by Michel Wedel.

## 1. Introduction

One of the standard assumptions in conjoint analysis is that a consumer's utilities remain constant during the course of the conjoint study, an assumption that underlies most normative theories of value maximization (e.g., von Neumann and Morgenstern 1947), which assume that individuals have a global set of utilities that is revealed when elicited. There are a number of reasons why this assumption may not hold in practice. Research of human decision making and judgment strongly suggest that preferences are constructed as opposed to being revealed (see Payne et al. 1992, Slovic 1995), implying that individuals may engage in simplification strategies or they may be learning about their preferences during the course of a preference study. Evolution of preferences may also come about because of fatigue and boredom (e.g., Bijmolt and Wedel 1995) or because individuals adjust to the range of possible levels that are revealed during the course of the study (e.g., Hsee et al. 1999).

The manifestation of these phenomenon has been demonstrated in the revealed/stated preference analysis (conjoint measurement) literature (cf. Luce and Tukey 1964), where there is mention of a "burn-in" period during which subjects evolve a systematic approach to providing preference judgments as they examine attributes and attribute levels (Huber et al. 1992). More significantly, Johnson and Orme (1996) have noticed significant shifts in attribute weights during the course of a single survey for such revealed/stated preference studies. DeSarbo et al. (2004) have documented regime shifts in such conjoint preference/utility functions when estimated during the course of the experiment. DeSarbo et al. (2005) have documented such dynamic effects in full-profile conjoint analysis, involving a student apartment application using a counterbalanced design with ordinary least squares and empirical Bayes models.

There is evidence that dynamics because of learning effects may be influenced by the initial profile seen by a participant, e.g., studies on order of entry effects

(cf. Kardes and Kalyanaram 1992) and initial choices (cf. Muthukrishnan and Kardes 2001) suggest that an initial product profile may provide an anchor for subsequent profiles. The standard approach to reduce or eliminate these types of crossover effects is to counterbalance the experimental design (see Keppel 1991). One of the simplest counterbalance arrangements is a cyclic one, where the set of profiles from a design are rotated for different groups of participants (e.g., participants in group 1 rate profiles $1, 2, 3, \ldots, T$, participants in group 2 rate profiles $T, 1, 2, \ldots, T-1$, and so on). In addition to helping reduce order effects, we demonstrate that counterbalancing plays an important role in obtaining estimates that recover the average of underlying aggregate dynamic preference, which may exist independent of the profile order. While design issues are not the main focus of this paper, we remark that Toubia et al. (2003) have recently proposed and tested new adaptive question design and estimation algorithms for partial-profile conjoint analysis.

A number of dynamic models have been proposed in the marketing literature, e.g., Vakratsas et al. (2004). Some studies and the resulting models address aggregate market behavior (see Wildt and Winer 1983, Dekimpe and Hanssens 1995, Bronnenberg et al. 2000), and others look at market behavior by considering individual-level consumer purchases (see Guadagni and Little 1983, Meyer and Sathi 1985, Allenby and Lenk 1994, Erdem 1996, Paap and Franses 2000). Models of aggregate market behavior explicitly consider heterogeneity across time, but do not typically consider individual-level utilities. Individual-level models are designed to uncover persistence in utility functions with respect to brand (Guadagni and Little 1983) and other marketing mix variables, e.g., Paap and Franses (2000) consider a multinomial probit model, which can distinguish between short-term and long-term effects of different marketing mix variables. Such individual-level models use current and lagged marketing mix information and lagged utility and/or purchase information to uncover persistence effects. They are different from the models proposed here, which allow for both individual and time-specific part-worth utilities, and are well suited for investigating potential part-worth evolution during conjoint studies.

We consider extensions of two different classes of dynamic models: (1) models that are a function of time and (2) dynamic linear models. These models are extended to allow for part-worth utilities that display heterogeneity across *both* time and individuals. The basic hierarchical Bayesian models assume that heterogeneity exists across individuals, but not across time (e.g., Lenk et al. 1996). Alternatively, dynamic models such as Bayesian dynamic linear

models assume that heterogeneity exists across time but not across individuals (e.g., West and Harrison 1997). The purpose of this research note is threefold: (1) to introduce two new classes of Bayesian dynamic models that accommodate heterogeneity both across time and across individuals; (2) to investigate the diagnostic power of these models for different potential sources of evolution that may occur during a conjoint study; and (3) to demonstrate that counterbalancing the design used in a conjoint study is essential if utility dynamics exist, which are independent of order in which the profiles are shown to the participants.

We proceed with our description of the models and our investigation into their performance in the following manner. We introduce and discuss our newly proposed models in §2. In §3, we report the results of a simulation conjoint study, where we investigate the performance of these models using synthetic data sets that assume different types of dynamics; we also discuss the importance of using a counterbalanced design when dynamics are present. In §4, we conclude and discuss future research in this area.

## 2. Dynamic Bayesian Models Incorporating Individual Heterogeneity

We propose two different classes of dynamic models that can be used to model rating-based conjoint data; both classes contain the standard hierarchical Bayesian random effects model as a special case (e.g., Lenk et al. 1996). The hierarchical Bayesian random effects model assumes that the likelihood function for a rating $y_{it}$, given by the $i$th individual for the $t$th product profile $X_{it}$, $i = 1, \ldots, n$ and $t = 1, \ldots, T$, is as follows:

$$y_{it} = (\bar{\beta} + \beta_i)' X_{it} + \varepsilon_{it}, \tag{1}$$

where $\varepsilon_{it} =_d N(0, \sigma_i^2)$, $N$ represents the normal distribution and $=_d$ means equal in distribution. The prior distributions for the individual part-worth deviations are given by $\beta_i =_d N(0, \Lambda)$ and we assume conjugate priors for $\Lambda^{-1} =_d Wishart(n_\Lambda, P_\Lambda)$, where $n_\Lambda$ and $P_\Lambda$ are specified, $\bar{\beta} =_d N(\bar{\bar{\beta}}, vI)$, where $I$ is the identity matrix, $\bar{\bar{\beta}}$ and $v$ are specified, and $\sigma_i^{-2} =_d Gamma(Shape, Scale)$, where $Shape$ and $Scale$ are specified. Both classes of models that we introduce build on the standard random effects model by allowing the aggregate part-worth to vary with respect to time or

$$y_{it} = (\bar{\beta}_t + \beta_i)' X_{it} + \varepsilon_{it}, \tag{2}$$

where we assume the same priors as in the random effects model with the exception of the prior for the aggregate part-worth $\bar{\beta}_t$. The first class of models (which we denote as *dynamic functional form* models)

assumes that the aggregate part-worths follow a functional form with respect to time. The second class of models (which we denote as *dynamic linear random effects models*) extends the dynamic linear model framework of West and Harrison (1997) by allowing an aggregate part-worth vector to follow a system equation as in a dynamic linear model.

The *dynamic functional form* models assume that the aggregate part-worth vector follows a quadratic function of time or

$$\bar{\beta}_t = \bar{\beta}_0 + \bar{\beta}_1 t + \bar{\beta}_2 t^2, \qquad (3)$$

where $\bar{\beta}_j = N(\bar{\bar{\beta}}_j, v_j I)$, $j = 0, 1, 2$; we have assumed that $\bar{\bar{\beta}}_j = 0$ (this choice is insignificant if $v_j$ is large). Clearly when $v_1 = v_2 = 0$, the quadratic model reduces to the standard random effects model. To avoid multicollinearity problems, we rescale and center time ($t' = t/T - 1/2$) when applying this quadratic model. Further, if we assume that $\Lambda = 0$, then individual heterogeneity is removed and one is left with an aggregate, quadratic model with respect to time. Regardless of whether individual-level heterogeneity is included, the quadratic model contains the linear model as a special case (by setting $v_2 = 0$) and can act as an approximation of a more complicated functional forms (e.g., an exponential model). Given this framework, there is obviously a broad range of functional form models that could be considered.

For the *dynamic linear random effects* model, we assume that the aggregate part-worth dynamics are given by the system equation

$$\bar{\beta}_t = G\bar{\beta}_{t-1} + w_t, \qquad (4)$$

where $w_t =_d N(0, \Sigma)$, $\Sigma^{-1} =_d Wishart(n_\Sigma, P_\Sigma)$, $n_\Sigma$, and $P_\Sigma$ are specified, and $\bar{\beta}_1 =_d N(A_1, B_1)$, we assume that $A_1 = 0$.

Our major departure from the approach of West and Harrison (1997) is regarding the treatment of the matrix $G$. Because of difficulties in estimation, they assume that $G$ is known. We assume that $G$ is a random diagonal matrix, where the diagonal elements (indexed by $j$) are normally distributed around 1 or

$$G_{jj} =_d N(1, \tau^2), \qquad (5)$$

where $\tau^2$ is specified. With respect to interpretation, if the posterior estimate of $G_{jj}$ is greater than 1, this implies that, on average, the $j$th mean part-worth is becoming more important across time; and, if it is less than 1 but still positive, then, on average, the mean part-worth becomes less important across time.

In general, as the matrix $G$ is diagonal, Equation (4) implies that $\bar{\beta}_t$ follows a vector autoregressive process of order 1. When $\tau^2 = 0$, $G$ becomes an identity matrix and $\bar{\beta}_t$ follows a random walk process. If $G$ is

an identity matrix and the variance-covariance matrix of $w_t$ is zero, then the model will simplify to the standard random effects model. Finally, if we assume that $\Lambda = 0$, then individual heterogeneity is removed, and one is left with the standard dynamic linear model.

For the sake of completeness, the full conditional densities for the two proposed dynamic models are given in Appendix A. All of these models that we present have standard full conditional densities and are estimated using a Markov chain Monte Carlo (MCMC) algorithm based on the Gibbs sampler (see Gilks et al. 1996 for a general discussion of MCMC methods).

## 3. The Conjoint Simulation Study

We assessed the performance of the proposed models on synthetic data that were generated from three different potential cognitive processing dynamics: fatigue, learning, and rule simplification. For each of the scenarios, we generated synthetic parameters and data for 108 individuals, where each individual rated 27 profiles (the first 24 profiles were used for calibration and the last 3 were used for holdout validation). To mimic a conjoint setting, we generated a common set of profiles using a $2^{10}$ fractional factorial design, where each profile consisted of 10 dummy variables and an intercept. In each scenario, every individual had a true part-worth, $\beta_i$, which was perturbed during the course of the study to generate individual and time-specific part-worth values. For the fatigue scenario, we assume that the individual-level part-worths used to generate ratings were centered around an individual's true part-worth but had a variance, which was initially small and then increased toward the end of the sequence of profiles. This dynamic mirrors the type of behavior that one would expect from a burn-in scenario, where the variance is large for the initial profiles and then decreases across time. For the learning scenario, the individual-level part-worth used to generate the ratings converges at roughly an exponential rate to the true part-worth and then is centered around the individual's true part-worth with a relatively small variance. The rule simplification scenario reflects a strategy where respondents suddenly determine that an attribute is no longer relevant to their utility function; the part-worth used to generate the ratings is initially centered around the individual's true part-worth with a small variance, and then at a random time the part-worth is centered around zero and has a larger variance that quickly decreases. We used a random scheme to determine whether a part-worth was no longer relevant and, hence, set to zero at some point during the conjoint study, where the probability of being set to zero depended on how close the part-worth was to zero.

This type of sudden change dynamic could also be related to a structural change in the part-worth where an individual suddenly realizes that the range for some of the attribute levels is different than they initially anticipated, resulting in a sudden change in some of the part-worth values. A detailed description of the algorithms used to generate these synthetic data is given in a separate technical appendix (see http://mktsci.pubs.informs.org).

For each scenario, we generated two sets of synthetic data using the same parameters but different counterbalancing approaches. The first data set did not have any counterbalancing (all of the individuals "saw" the profiles in the same order) and the second data set used a cyclic counterbalancing approach (individuals were divided into 27 groups and each group "saw" the profiles in the same order, but with a different initial profile). To clarify, for the cyclic counterbalancing, group 1 "rated" the profiles $1, 2, 3, \dots, 27$, group two "rated" the profiles $2, 3, \dots, 27, 1$, and so on. Synthetic ratings were obtained by adding a normal error to the utility resulting from the part-worths and profile.

### 3.1. Unbiased Dynamic Estimates and Counterbalancing

To illustrate the impact of counterbalancing from the perspective of estimation, we briefly need to initially discuss maximum likelihood estimates (MLE). This discussion allows us to highlight the origin of problems that can arise when using static Bayesian models with vague prior specifications and dynamic Bayesian models to analyze data from an experimental design that is not counterbalanced. Typically, a good experimental design and estimation scheme result in an unbiased estimate, under the assumption that the model parameters are static or constant across time. If this assumption is in doubt, then it is reasonable to add a criterion for a good design by requiring that the unbiased estimate based on that design is an *unbiased dynamic estimate* or an estimate that, on average, equals the average of the underlying dynamic utility. To illustrate, assume that $y_{it} = \bar{\beta}'_t X_{it} + \varepsilon_{it}$ and $\varepsilon_{it} =_d N(0, \sigma^2)$. If $\bar{\beta}_t = \bar{\beta}$ for all $t$, then the standard aggregate MLE is an unbiased estimate or $E[\hat{\bar{\beta}}_{MLE}] = \bar{\beta}$. However, if $\bar{\beta}_t$ are different for each time $t$, then there is no guarantee that the standard aggregate MLE is an unbiased dynamic estimate or that

$$E\left[\hat{\bar{\beta}}_{MLE}\right] = \frac{1}{T}\sum_t \bar{\beta}_t.$$

Figure 1 illustrates the difference between the MLE for noncounterbalanced and counterbalanced data, by graphically comparing the MLE estimates to the average—at each point in time—of the actual parameter values that were used to generate the synthetic

data; as the learning and rule simplification tend to have similar results, we restricted our attention in this illustration to the fatigue and learning scenarios.

Clearly, as highlighted by the learning scenario estimates, if dynamics are present and counterbalancing is not used, static estimates like the aggregate MLE can be entirely different from the average dynamic. This occurs because the noncounterbalanced design systematically excludes part of the dynamic or weights parts of the dynamic inappropriately. Introducing a simple cyclic counterbalancing approach corrects this by ensuring that every part of the dynamic is evenly represented in the data (see Appendix B for a formal proof that counterbalanced designs will result in unbiased dynamic estimates).

### 3.2. Conjoint Simulation Results

Analyzing the simulation data using the static Bayesian random effects model and the dynamic Bayesian models introduced in §2 yields several interesting results. First, based on our model choice criteria, we find that for these synthetic data sets, accounting for individual-level heterogeneity is relatively more important than accounting for dynamics. Second, we find that when we use noncounterbalanced data, estimates from the static random effects model behave in a fashion similar to the aggregate MLE and they are not representative of the average dynamic; this problem does not tend to occur with counterbalanced data.

Before discussing these results in more details, we describe the study and model choice tools that were used to discriminate between models. We analyzed both the noncounterbalanced and the counterbalanced data sets for all of the scenarios, using a range of Bayesian models: assuming that there was no heterogeneity across time or individuals (the aggregate static model), assuming that there was heterogeneity across individuals but not across time (the static random effects model), assuming that there was heterogeneity across time but not across individuals (aggregate quadratic and aggregate dynamic linear models), and assuming that there was heterogeneity across both individuals and time (random effects quadratic and dynamic linear models). We considered a range of prior values and found that the dynamic models performed best when data dependent priors were used for the dynamic portion of the model, and vague or noninformative priors were used for the remaining model parameters (see Appendix A for exact prior values).

We used log marginal probability and holdout prediction root mean square error (RMSE) to determine model performance, where log marginal probabilities were calculated using the estimation method

**Figure 1    MLE Parameter Estimates and Average Parameter Values**
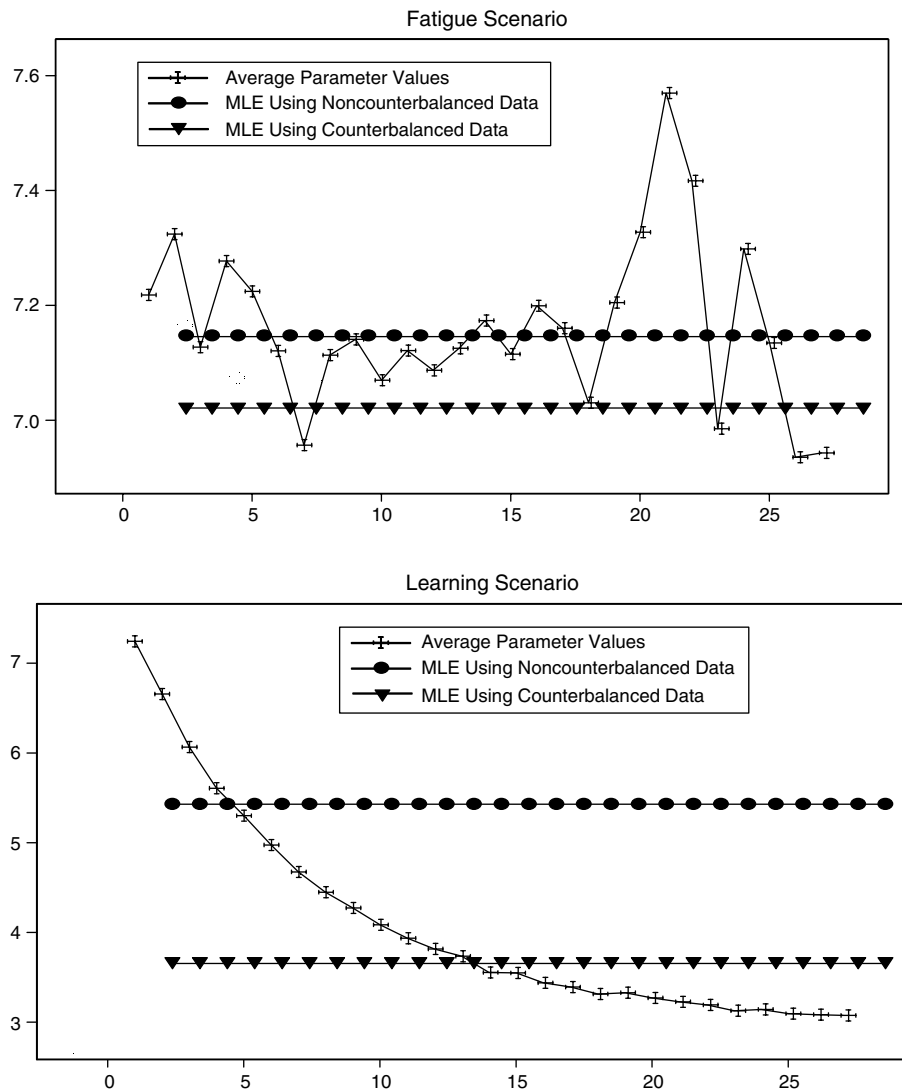


*Note.* The graphs above report estimates and actual values for variable 4, using the noncounterbalanced and counterbalanced data sets from the fatigue and learning scenarios.

provided by Newton and Raftery (1994). We calculated the RMSE for the holdout portion of the data set (the last three observations for each individual) for each sweep of the MCMC sampler resulting in a posterior mean and standard deviation of the holdout RMSE. For the static models, we used the individual-level parameter value at each sweep of the MCMC sampler to forecast out of sample observations; for the dynamic models, we generated parameter values for the holdout portion of the data set based on the current set of parameters and the model dynamics at each sweep of the MCMC sampler to forecast out of sample data. To assess parameter recovery, we compared the average of the actual parameter value at each point in time with the posterior mean estimate of the average parameter value at each point in time, and then calculated the RMSE between these actual and estimated parameter values.

For all of the synthetic data sets, including heterogeneity across individuals appears to contribute the most to performance improvements in terms of the model choice criteria (see Table 1). In terms of holdout predictions, the quadratic random effects model is comparable to the static random effects model for all of the scenarios and clearly outperforms the latter for the learning scenario. When participants learn during the course of the study and a counterbalancing design is employed, the quadratic and dynamic linear random effects models are comparable in terms of holdout predictions and they are much better than the other models.

The dynamic linear random effects model gives the lowest log marginal probability for every scenario studied in this paper, but it does not always gives the best holdout predictions. This may be indicative of unreliable parameter estimates from the dynamic

**Table 1    Predictive Performance of Dynamic and Static Models for Three Synthetic Data Sets: Fatigue, Learning, and Simplification**

| | | Aggregate models | | | Random effects models | | |
|---|---|---|---|---|---|---|---|
| Data set | Model choice criteria | Static | Quadratic | Dynamic linear | Static | Quadratic | Dynamic linear |
| Fatigue | Log marginal probability | −7,249.27 | −7,245.71 | −7,233.61 | −6,698.29 | −6,686.5 | −6,668.27 |
| | Holdout RMSE | $6.00_{(0.01)}$ | $6.03_{(0.05)}$ | $6.35_{(0.32)}$ | $5.76_{(0.08)}$ | $5.79_{(0.09)}$ | $6.66_{(0.64)}$ |
| Learning | Log marginal probability | −7,605.12 | −7,028.12 | −6,999.77 | −7,122.61 | −5,452.91 | −5,410.44 |
| | Holdout RMSE | $9.70_{(0.14)}$ | $5.78_{(0.31)}$ | $5.46_{(0.74)}$ | $9.84_{(0.13)}$ | $4.26_{(0.30)}$ | $8.07_{(0.89)}$ |
| Simplification | Log marginal probability | −7,636.63 | −7,620.37 | −7,607.27 | −6,503.89 | −6,422.53 | −6,404.18 |
| | Holdout RMSE | $5.66_{(0.03)}$ | $5.72_{(0.10)}$ | $5.92_{(0.28)}$ | $4.10_{(0.10)}$ | $4.20_{(0.15)}$ | $4.50_{(0.38)}$ |

*Noncounterbalanced data sets (profiles were not rotated)* — above

| | | Aggregate models | | | Random effects models | | |
|---|---|---|---|---|---|---|---|
| Data set | Model choice criteria | Static | Quadratic | Dynamic linear | Static | Quadratic | Dynamic linear |
| Fatigue | Log marginal probability | −7,257.68 | −7,257.37 | −7,112.53 | −6,697.5 | −6,697.43 | −6,482.92 |
| | Holdout RMSE | $6.15_{(0.02)}$ | $6.20_{(0.04)}$ | $6.29_{(0.06)}$ | $6.13_{(0.10)}$ | $6.18_{(0.10)}$ | $6.19_{(0.11)}$ |
| Learning | Log marginal probability | −8,021.65 | −7,106.91 | −6,902.83 | −7,193.42 | −5,436.4 | −5,107.87 |
| | Holdout RMSE | $6.68_{(0.08)}$ | $4.64_{(0.07)}$ | $4.76_{(0.10)}$ | $6.42_{(0.14)}$ | $2.96_{(0.10)}$ | $2.99_{(0.16)}$ |
| Simplification | Log marginal probability | −7,757.54 | −7,720.44 | −7,606.42 | −6,551.26 | −6,437.19 | −6,220.5 |
| | Holdout RMSE | $5.48_{(0.03)}$ | $5.35_{(0.04)}$ | $5.37_{(0.02)}$ | $3.81_{(0.11)}$ | $3.68_{(0.11)}$ | $3.90_{(0.19)}$ |

*Counterbalanced data sets (profiles were rotated)* — above

*Note.* Criteria: log marginal probability and RMSE of holdout in back data posterior mean (std).

linear random effects model and it suggests that the predictive RMSE may be a more dependable model choice diagnostic.

As demonstrated by the parameter RMSE given in Table 2, using counterbalanced data typically results in comparable or better recovery of the true aggregate parameter dynamics when compared with using noncounterbalanced data.[1]

In general, we found that when analyzing noncounterbalanced data, the quadratic random effects model gives estimates that are somewhat stable and that are usually, but not always, close to the average dynamic.[2] The dynamic linear random effects model does not perform well with noncounterbalanced data; the parameter estimates are very unstable and the MCMC algorithm takes much longer to converge. Convergence diagnostics indicated that the static and dynamic models considered here converged relatively quickly, within one or two thousand iterations, for all data sets, with the exception of the dynamic linear random effects model with noncounterbalanced data,

[1] There are a few instances that this is not the case—for example involving the fatigue data and the static and quadratic random effects models—that is because of random variations in the simulated data.

[2] As pointed out by a reviewer, the performance of these models in terms of parameter recovery depends on the number of ratings in the data set; the larger the data set, the better the parameter recovery. However, it is important to realize that based on our investigation, similar size counterbalanced data sets tend to result in much better parameter recovery than noncounterbalanced data sets.

which appears to have not converged after more than 8,000 iterations of the MCMC algorithm. In contrast, the simpler quadratic random effects model appears to be able to use its strong functional form assumptions to bridge the gap in missing information, resulting in better estimates for noncounterbalanced data.

## 4.   Discussion

To investigate preference changes during the course of a conjoint study, it is essential to have statistical models that capture dynamics and accommodate individual heterogeneity. The dynamic Bayesian models introduced in this paper accommodate heterogeneity *both* across time and across individuals. These models offer researchers tools that can be used in assessing the quality of the inference that can be derived from an analysis of their conjoint data.

In addition, we find theoretical and empirical support for using an experimental design that is counterbalanced. The consumer behavior literature supports using counterbalancing to ensure that crossover effects do not create or drive preference dynamics. We demonstrate that, even if preference dynamics are independent of the order of the profiles, it is essential to use a counterbalanced design. If a counterbalanced design is not used and dynamics exist, then static models can fail to capture the average of the dynamics, and dynamic models can give relatively poor estimates, however, these problems are eliminated when a counterbalanced design is used.

**Table 2**    **Parameter Recovery of Dynamic and Static Models for Three Synthetic Data Sets: Fatigue, Learning, and Simplification**

| Data set | Type of data set | Aggregate models | | | Random effects models | | |
|---|---|---|---|---|---|---|---|
| | | Static | Quadratic | Dynamic linear | Static | Quadratic | Dynamic linear |
| Fatigue | Noncounterbalanced | 1.578 | 1.581 | 1.914 | 1.433 | 1.440 | 2.174 |
| | Counterbalanced | 1.577 | 1.582 | 1.767 | 1.469 | 1.473 | 1.639 |
| Learning | Noncounterbalanced | 2.752 | 1.811 | 1.528 | 2.848 | 2.022 | 3.420 |
| | Counterbalanced | 1.859 | 1.451 | 1.546 | 2.364 | 1.853 | 1.850 |
| Simplification | Noncounterbalanced | 1.990 | 1.935 | 2.383 | 1.539 | 1.454 | 2.049 |
| | Counterbalanced | 1.951 | 1.899 | 2.071 | 1.476 | 1.398 | 1.554 |

*Note.* Parameter RMSE based on actual average parameter values and posterior mean estimates.

From our simple conjoint simulation study, we find that for noncounterbalanced data, the static Bayesian random effects model estimates behave in a fashion similar to the aggregate MLE and they are not representative of the average dynamic; this problem does not occur with counterbalanced data. We also find that the dynamic models, particularly the quadratic models, result in estimates that are much closer to the average dynamic when noncounterbalanced data are used, but that these estimates do not always represent the dynamics properly; again, these problems do not occur when counterbalanced data are used. Based on our model choice criteria, we find that accounting for individual-level heterogeneity is relatively more important than accounting for the dynamics employed in our modest simulation. Finally, we find that predictive RMSE is a more reliable model choice diagnostic when compared with the log marginal probability.

There are, of course, a number of limitations with respect to this simulation study. First, the empirical results are based on our synthetic conjoint data sets and may change for real conjoint data sets. Second, the models that we introduce do not accommodate potential evolution in the variance of the error term, which could be another important source of dynamics. In addition, they do not offer a way for investigating whether the change in one part-worth impacts the dynamics of other part-worths. Obviously, these dynamic models could be extended to accommodate these types of complex dynamics. For example, the dynamic linear random effects model could be modified so that the updating matrix $G$ in the system equation is a full matrix, which would allow for changes in one part-worth to impact the dynamics of other part-worths; these types of extensions are the subject of future research. Finally, several applications with real conjoint data sets are necessary to examine in detail the exact nature of dynamics in actual conjoint settings. In this same line of thought, extensions of these models to choice-based conjoint analysis, adaptive conjoint analysis, hybrid conjoint analysis, and so on would also prove beneficial.

## Appendix A. Full Conditional Densities and Prior Values

**Full Conditional Densities**
Some of the full conditional densities for both the quadratic and dynamic linear models are identical given $\bar{\beta}_t$. We present these full conditional densities first, and then the unique full conditional densities for the quadratic model, followed by the unique full conditional densities for the dynamic linear model. The common full conditional densities are given by

$$\beta_i \mid \sigma_i^2, \Lambda, \bar{\beta}_t, D$$

$$=_d N\left(\left[\frac{1}{\sigma_i^2}\sum_t x_{it}x_{it}' + \Lambda^{-1}\right]^{-1}\left[\frac{1}{\sigma_i^2}\sum_t x_{it}(y_{it} - \bar{\beta}_t'x_{it})\right],\right.$$

$$\left.\left[\frac{1}{\sigma_i^2}\sum_t x_{it}x_{it}' + \Lambda^{-1}\right]^{-1}\right), \tag{A.1}$$

$$\sigma_i^{-2} \mid \beta_i, \Lambda, \bar{\beta}_t, D$$

$$=_d Gamma\left(\frac{T}{2} + Shape, \frac{\sum_t (y_{it} - (\bar{\beta}_t + \beta_i)'x_{it})^2}{2} + Scale\right), \tag{A.2}$$

and

$$\Lambda^{-1} \mid \beta_i =_d Wishart\left(n_\Lambda + N, \left[P_\Lambda + \sum_i(\beta_i)(\beta_i)'\right]\right), \tag{A.3}$$

where $D$ represents all of the data.

The unique full conditional densities for the quadratic model are given by

$$\bar{\beta}_0 \mid \sigma_i^2, \bar{\beta}_1, \bar{\beta}_2, \beta_i, D$$

$$=_d N\left(\left[\sum_{it}\frac{1}{\sigma_i^2}(x_{it}x_{it}') + \frac{1}{v_0}I\right]^{-1}\right.$$

$$\cdot\left[\sum_{it}\frac{1}{\sigma_i^2}x_{it}(y_{it} - (\bar{\beta}_1 t + \bar{\beta}_2 t^2 + \beta_i)'x_{it})\right],$$

$$\left.\left[\sum_{it}\frac{1}{\sigma_i^2}(x_{it}x_{it}') + v_0 I\right]^{-1}\right), \tag{A.4}$$

$$\bar{\beta}_1 \mid \sigma_i^2, \bar{\beta}_0, \bar{\beta}_2, \beta_i, D$$

$$=_d N\left(\left[\sum_{it}\frac{1}{\sigma_i^2}t^2(x_{it}x_{it}') + v_1 I\right]^{-1}\right.$$

$$\cdot\left[\sum_{it}\frac{1}{\sigma_i^2}x_{it}t(y_{it} - (\bar{\beta}_0 + \bar{\beta}_2 t^2 + \beta_i)'x_{it})\right],$$

$$\left.\left[\sum_{it}\frac{1}{\sigma_i^2}t^2(x_{it}x_{it}') + \frac{1}{v_1}I\right]^{-1}\right), \tag{A.5}$$

$$\bar{\beta}_2 \mid \sigma_i^2, \bar{\beta}_0, \bar{\beta}_1, \beta_i, D$$

$$=_d N\left(\left[\sum_{it} \frac{1}{\sigma_i^2} t^4 (x_{it} x_{it}') + \frac{1}{v_2} I\right]^{-1}\right.$$

$$\cdot \left[\sum_{it} \frac{1}{\sigma_i^2} x_{it} t^2 (y_{it} - (\bar{\beta}_0 + \bar{\beta}_1 t + \beta_i)' x_{it})\right],$$

$$\left.\left[\sum_{it} \frac{1}{\sigma_i^2} t^4 (x_{it} x_{it}') + \frac{1}{v_2} I\right]^{-1}\right). \quad (A.6)$$

The unique full conditional densities for the dynamic linear model are given by

$$\bar{\beta}_1 \mid \sigma_i^2, G, \beta_i, \bar{\beta}_2, \Sigma, D$$

$$=_d N\left(\left[\sum_{i=1}^n \frac{1}{\sigma_i^2} x_{i1} x_{i1}' + G'\Sigma^{-1}G + B_1^{-1}\right]^{-1}\right.$$

$$\cdot \left[\sum_{i=1}^n \frac{1}{\sigma_i^2} x_{i1} (y_{i1} - \beta_i' x_{i1}) + G'\Sigma^{-1}\bar{\beta}_2 + B_1^{-1} A_1\right],$$

$$\left.\left[\sum_{i=1}^n \frac{1}{\sigma_i^2} x_{i1} x_{i1}' + G'\Sigma^{-1}G + B_1^{-1}\right]^{-1}\right), \quad (A.7)$$

$$\bar{\beta}_t \mid \sigma_i^2, G, \beta_i, \bar{\beta}_{t-1}, \bar{\beta}_{t+1}, \Sigma, D$$

$$=_d N\left(\left[\sum_{i=1}^n \frac{1}{\sigma_i^2} x_{it} x_{it}' + G'\Sigma^{-1}G + \Sigma^{-1}\right]^{-1}\right.$$

$$\cdot \left[\sum_{i=1}^n \frac{1}{\sigma_i^2} x_{it} (y_{it} - \beta_i' x_{it}) + G'\Sigma^{-1}\bar{\beta}_{t+1} + \Sigma^{-1}G\bar{\beta}_{t-1}\right],$$

$$\left.\left[\sum_{i=1}^n \frac{1}{\sigma_i^2} x_{it} x_{it}' + G'\Sigma^{-1}G + \Sigma^{-1}\right]^{-1}\right), \quad \text{for } T > t \geq 2, \quad (A.8)$$

$$\bar{\beta}_T \mid \sigma_i^2, G, \beta_i, \bar{\beta}_{T-1}, \Sigma, D$$

$$=_d N\left(\left[\sum_{i=1}^n \frac{1}{\sigma_i^2} x_{iT} x_{iT}' + \Sigma^{-1}\right]^{-1}\right.$$

$$\cdot \left[\sum_{i=1}^n \frac{1}{\sigma_i^2} x_{iT} (y_{iT} - \beta_i' x_{iT}) + \Sigma^{-1}G\bar{\beta}_{T-1}\right],$$

$$\left.\left[\sum_{i=1}^n \frac{1}{\sigma_i^2} x_{iT} x_{iT}' + \Sigma^{-1}\right]^{-1}\right), \quad (A.9)$$

$$\Sigma^{-1} \mid G, \bar{\beta}_t, D$$

$$=_d Wishart\left(n_\Sigma + T, \left[P_\Sigma + \sum_{t=2}^T (\bar{\beta}_t - G\bar{\beta}_{t-1})(\bar{\beta}_t - G\bar{\beta}_{t-1})'\right]\right), \quad (A.10)$$

and

$$diag(G) \mid \bar{\beta}_t, \Sigma, D$$

$$=_d N\left(\left[\left(\sum_{t=2}^T (\bar{\beta}_{t-1}\bar{\beta}_{t-1}')\right) \% \Sigma^{-1} + \frac{1}{\tau^2} I_p\right]^{-1}\right.$$

$$\cdot \left[\sum_{t=2}^T \bar{\beta}_{t-1} \% (\Sigma^{-1}\bar{\beta}_t) + \frac{1}{\tau^2} diag(I_p)\right],$$

$$\left.\left[\left(\sum_{t=2}^T (\bar{\beta}_{t-1}\bar{\beta}_{t-1}')\right) \% \Sigma^{-1} + \frac{1}{\tau^2} I_p\right]^{-1}\right), \quad (A.11)$$

where $diag()$ is a vector comprised of the diagonal elements of a matrix, and % is the binary operator for element-by-element matrix multiplication (i.e., $(A \% B)_{ij} = \{A_{ij}B_{ij}\}$).

### Prior Values
Although a range of hyperparameter values were considered, the following values were used in the models reported in this study. Results for the static models were very robust to the prior specifications, while the results for the dynamic models were sensitive to the hyperparameters that contributed to the dynamics. Some of the hyperparameters are common across all models; we will first specify the common hyperparameters and then give the hyperparameters for the dynamic models: $Shape = 1$, $Scale = 0.1$, $n_\Lambda = COV + 1$, $P_\Lambda = 0.1I$, where $COV$ is the number of covariate and $I$ is the identity matrix. For the quadratic model, the variance of the intercept was set to a large value but the variance of the linear and quadratic terms were small: $v_0 = 1,000$, $v_1 = 1$, and $v_2 = 0.1$. For the dynamic linear random effects models, we used $n_\Sigma = 10$, $P_\Sigma = 100$, $B_1 = 0.1I$, and $\tau^2 = 0.1$.

### Appendix B. Proof of Unbiased Dynamic Estimates
DEFINITION. COUNTERBALANCED EXPERIMENTAL DESIGNS. Let $x = (x_1, \ldots, x_T)$ be an experimental design. A counterbalance of this design is achieved by assigning the same number of participants to $T$ groups, where each element of $x$ is applied to at least one of the groups during every trial of the experiment.

LEMMA. *Let* $(\theta_1, \ldots, \theta_T)$ *represent dynamics from a statistical model. Let* $x$ *be an experimental design with* $T$ *trials and* $\hat{\theta}$ *be an unbiased estimate of* $\theta$ *(* $E[\hat{\theta}] = \theta$ *) when* $\theta_t = \theta$ *for all* $t$. *If* $x$ *is a counterbalanced experimental design, then* $\hat{\theta}$ *will be an unbiased dynamic estimate.*

PROOF. If $E[\hat{\theta}] = \theta$ when $\theta_t = \theta$ for all $t$, then $E[\hat{\theta}] = \sum_{t=1}^T w_t \theta_t$, where $w_t$ may depend on the design and $\sum_{t=1}^T w_t = 1$. If $x$ is counterbalanced, then

$$E[\hat{\theta}] = \frac{1}{T} \sum_{j=1}^T E[\hat{\theta}_{Gj}] = \frac{1}{T} \sum_{t=1}^T \theta_t (w_1 + \cdots + w_T) = \frac{1}{T} \sum_{t=1}^T \theta_t. \quad \square$$

### References

Allenby, G., P. Lenk. 1994. Modeling household purchase behavior with logistic normal regression. *J. Amer. Statist. Association* **89** 1218–1231.

Bijmolt, T., M. Wedel. 1995. The effects of alternative methods of collecting similarity data for multidimensional scaling. *Internat. J. Res. Marketing* **12** 363–371.

Bronnenberg, B., V. Mahajan, W. Vanhonacker. 2000. The emergence of new repeat-purchase categories: The interplay of market share and retailer distribution. *J. Marketing Res.* **37** 16–21.

Dekimpe, M., D. Hanssens. 1995. The persistence of marketing effects on sales. *Marketing Sci.* **14** 1–21.

DeSarbo, W., D. Lehmann, F. Hollman. 2004. Modeling dynamic effects in repeated measures experiments involving preference/choice: An illustration involving stated preference analysis. *Appl. Psych. Measurement* **28** 186–209.

DeSarbo, W., D. K. H. Fong, J. Liechty, J. C. Copland. 2005. Evolutionary preference/utility functions: A dynamic perspective. *Psychometrika*. Forthcoming.

Erdem, T. 1996. A dynamic analysis of market structure based on panel data. *Marketing Sci.* **14** 359–378.

Gilks, W., S. Richardson, D. Spiegelhalter. 1996. *Markov Chain Monte Carlo in Practice*. Chapman & Hall, New York.

Guadagni, P., J. Little. 1983. A logit model of brand choice calibrated on scanner data. *Marketing Sci.* **2** 203–238.

Hsee, C., G. Loewenstein, S. Blount, M. Bazerman. 1999. Preference reversals between joint and separate evaluations of options: A review and theoretical analysis. *Psych. Bull.* **125**(5) 576–590.

Huber, J., D. Wittink, R. Johnson, R. Miller. 1992. Learning effects in preference tasks: Choice-based versus standard conjoint. *Sawtooth Software Conference Proceedings*. Sawtooth Software, Inc., Sequim, WA.

Johnson, R., B. Orme. 1996. How many questions should you ask in choice-based conjoint studies. Technical report, Sawtooth Software, Inc., Sequim, WA.

Kardes, F., G. Kalyanaram. 1992. Order-of-entry effects on consumer memory of judgment: An information integration perspective. *J. Marketing Res.* **29** 343–357.

Keppel, G. 1991. *Design and Analysis: A Researcher's Handbook*, ed. 3. Prentice-Hall, Englewood Cliffs, NJ.

Lenk, P., W. DeSarbo, P. Green, M. Young. 1996. Hierarchical Bayes conjoint analysis: Recovery of part-worth heterogeneity from incomplete designs in conjoint analysis. *Marketing Sci.* **15** 173–191.

Luce, D., J. Tukey. 1964. Simultaneous conjoint measurement: A new type of fundamental measurement. *J. Math. Psych.* **1** 1–27.

Meyer, R., A. Sathi. 1985. A multiattribute model of consumer choice during product learning. *Marketing Sci.* **4** 41–61.

Muthukrishnan, A. V., F. Kardes. 2001. Persistent preferences for product attributes: The effects of the initial choice context and uninformative experience. *J. Consumer Res.* **28** 89–104.

Newton, M., A. Raftery. 1994. Approximate Bayesian inference by the weighted likelihood bootstrap (with discussion). *J. Roy. Statist. Soc. Ser. B* **56** 3–48.

Paap, R., P. Franses. 2000. A dynamic multinomial probit model for brand choice with different long-run and short-run effects of marketing-mix variables. *J. Appl. Econometrics* **15** 717–744.

Payne, J., J. Bettman, E. Johnson. 1992. Behavioral decision research: A constructive processing perspective. *Ann. Rev. Psych.* **43** 87–131.

Slovic, P. 1995. The construction of preference. *Amer. Psych.* **50** 364–371.

Toubia, O., D. I. Simester, J. R. Hauser, E. Dahan. 2003. Fast polyhedral adaptive conjoint estimation. *Marketing Sci.* **22**(3) 273–303.

Vakratsas, D., F. M. Feinberg, F. M. Bass, G. Kalyanaram. 2004. The shape of advertising response functions revisited: A model of dynamic probabilistic thresholds. *Marketing Sci.* **23**(1) 109–119.

von Neumann, J., O. Morgenstern. 1947. *Theory of Games and Economic Behavior*. Princeton University Press, Princeton, NJ.

West, M., J. Harrison. 1997. *Bayesian Forecasting and Dynamic Models*, ed. 2. Springer-Verlag, New York.

Wildt, A., R. Winer. 1983. Modeling and estimation in changing market environments. *J. Bus.* **56** 365–388.