International Journal of Research in Marketing xxx (2018) xxx–xxx

Contents lists available at ScienceDirect

IJRM
International Journal of Research in Marketing

journal homepage: www.elsevier.com/locate/ijresmar

Full Length Article

A random coefficients mixture hidden Markov model for marketing research

Eelco Kappe [a,*], Ashley Stadler Blank [b], Wayne S. DeSarbo [a]

[a] *Marketing at the Smeal College of Business, Pennsylvania State University, University Park, PA 16802, United States of America*
[b] *Marketing at the Opus College of Business, University of St. Thomas, St. Paul, MN 55105, United States of America*

ARTICLE INFO

Article history:
First received on May 8, 2017 and was under review for 6 1/2 months
Available online xxxx

Senior Editor: Bruce G.S. Hardie

Keywords:
Hidden Markov model
Time-varying effects
Unobserved heterogeneity
Attendance demand model
Major League Baseball

ABSTRACT

The hidden Markov model (HMM) provides a framework to model the time-varying effects of marketing mix variables. When employed in a panel data context, it is important to properly account for unobserved heterogeneity across individuals. We propose a new random coefficients mixture HMM (RCMHMM) that allows for flexible patterns of unobserved heterogeneity in both the state-dependent and transition parameters. The RCMHMM nests all HMMs found in the marketing literature. Results of two simulation studies demonstrate that 1) averaging across a large number of different data generating processes, the RCMHMM outperforms all its nested versions using both in-sample and out-of-sample performance and 2) the RCMHMM is more robust than its nested versions when underlying model assumptions are violated. In addition, we apply the RCMHMM to an empirical application where we examine the effectiveness of in-game promotions in increasing the short-term demand for Major League Baseball (MLB) attendance. We find that the effectiveness of four promotional categories varies over the course of the season and across teams and that the RCMHMM performs best.

1. Introduction

The hidden Markov model (HMM) has made significant inroads into marketing over the past three decades (e.g., Montgomery, Li, Srinivasan, & Liechty, 2004; Netzer, Lattin, & Srinivasan, 2008; Poulsen, 1990). The HMM enables researchers to model the time-varying effects of marketing mix variables via the formulation of unobserved (or hidden) states. The HMM allows for different parameters for the group of observations that belong to each hidden state (state-dependent equation) as well as switching between hidden states over time (transition equation). Recent work has increased the value of the HMM in marketing by incorporating unobserved heterogeneity across individuals (e.g., Montoya, Netzer, & Jedidi, 2010; Schweidel & Knox, 2013), which may be present in both the state-dependent and transition parameters. Unobserved heterogeneity exists in many marketing applications (e.g., panel data) and failure to account for it leads to biased parameters and inaccurate managerial insights (see Netzer, Ebbes, & Bijmolt, 2016 for a discussion of the consequences of ignoring unobserved heterogeneity in the HMM).

While past research incorporates some patterns of unobserved heterogeneity into the HMM (e.g., random coefficients or latent classes), this body of work largely ignores the comparative performance of various alternative forms of unobserved heterogeneity in the state-dependent and transition parameters. An improper representation of the heterogeneity structure can lead to biased inferences (Hsiao, 2003)—especially in dynamic models, such as HMMs, where disentangling unobserved heterogeneity and

* Corresponding author.
E-mail addresses: erk11@psu.edu, (E. Kappe), stad8452@stthomas.edu, (A. Stadler Blank), wsd6@psu.edu (W.S. DeSarbo).

https://doi.org/10.1016/j.ijresmar.2018.07.002

Please cite this article as: Kappe, E., et al., A random coefficients mixture hidden Markov model for marketing research, *International Journal of Research in Marketing* (2018), https://doi.org/10.1016/j.ijresmar.2018.07.002

state dependence can be challenging (Heckman, 1991; Hyslop, 1999). If an HMM fails to properly account for unobserved heterogeneity, the effects are picked up by the dynamic elements of the model, which leads to incorrect conclusions about the time-varying effects.

We contribute to the marketing and HMM literatures by proposing a new random coefficients mixture HMM (RCMHMM) that allows for flexible patterns of unobserved heterogeneity in *both* the state-dependent and transition parameters. To achieve this flexibility, we draw on the random coefficients mixture (RCM) model (see Allenby, Arora, & Ginter, 1998; Lenk & DeSarbo, 2000). The RCM model combines the flexibility of semi-parametric latent class models and allows for parametric variation in the parameters within latent classes through the formulation of random coefficients. Incorporating an RCM representation of unobserved heterogeneity in both the state-dependent and transition parameters within an HMM framework makes it possible to nest all HMMs in the marketing literature and compare model performance of the RCMHMM to its 16 nested versions (combining homogenous, random coefficients, latent classes, and individual-specific independent state-dependent and transition parameters). Thus, the RCMHMM enables researchers to identify and select the optimal heterogeneity structure using a single model structure, reducing parameter bias.

We compare the performance of the RCMHMM to its 16 nested versions in two simulation studies. First, we estimate all 17 models on 17 corresponding datasets. Averaging across datasets, the RCMHMM outperforms its nested model versions and often results in a more parsimonious model (requiring fewer latent classes and hidden states to explain the data). The HMM with random coefficients for both the state-dependent and transition parameters (i.e., the state-of-the-art HMM in marketing) performs second-best. Comparing these two models, the RCMHMM performs better when the unobserved heterogeneity structure is multimodal (i.e., the data includes latent classes or an RCM distribution). When the unobserved heterogeneity structure is unimodal, both models have similar performance (as the RCMHMM perfectly nests the HMM with random coefficients for both the state-dependent and transition parameters), and there is no need to formulate the more complex RCMHMM. In general, the additional flexibility of the RCMHMM can be somewhat harmful when there is little unobserved heterogeneity in the data. Second, we examine what happens with different data dimensions and when the underlying model assumptions are violated. RCMHMM performance improves with more time periods and is more robust to different data generating processes; specifically, when an omitted variable is present, or when the error distribution is misspecified. Importantly, these two simulation studies show that using an incorrect form of unobserved heterogeneity biases parameter estimates. The RCMHMM enables researchers to minimize such bias by nesting 16 HMMs and selecting the model that most accurately reflects the unobserved heterogeneity structure in the data.

We also use the RCMHMM and its 16 nested versions to examine the effectiveness of in-game promotions in increasing the short-term demand for Major League Baseball (MLB) attendance. This empirical application is an ideal context, as past attendance research provides preliminary evidence that promotional effectiveness varies over time (Boyd & Krehbiel, 1999; Lemke, Leonard, & Tlhokwane, 2010) (facilitating hidden states) and across teams (Lemke et al., 2010; Marcum & Greenstein, 1985) (facilitating latent classes). We find that promotional effectiveness varies over the course of the season and across teams and that the RCMHMM performs best.

We organize the remainder of the article as follows. First, we review past research on incorporating unobserved heterogeneity into HMMs. Next, we present our RCMHMM along with the estimation procedure and predictive validation techniques to determine the best model. Then, we test the RCMHMM in two simulation studies and an empirical application. Finally, we highlight practical and methodological implications and discuss directions for future research.

2. Unobserved heterogeneity in HMMs

2.1. Modeling unobserved heterogeneity

Properly incorporating unobserved heterogeneity into a marketing model satisfies both practical and methodological concerns. From a practical standpoint, modeling heterogeneity is important to managers, so they can make inferences for individual consumers, stores, demographic market areas—or, as in our empirical application, teams (Allenby & Rossi, 1998). From a methodological standpoint, ignoring unobserved heterogeneity across individuals or incorporating an improper representation of the heterogeneity structure can lead to biased inferences (Dubé, Hitsch, & Rossi, 2010; Hsiao, 2003; Lenk & DeSarbo, 2000). This problem is even more severe in dynamic models, such as HMMs (Heckman, 1991; Hyslop, 1999). Deciding which heterogeneity structure to use can be difficult, as the true nature of the unobserved heterogeneity is usually unknown a priori and the selection of a discrete or continuous distribution is largely an empirical issue (Andrews, Ansari, & Currim, 2002; Michalek, Ebbes, Adigüzel, Feinberg, & Papalambros, 2011; Otter, Tüchler, & Frühwirth-Schnatter, 2004). Typically, a discrete distribution leads to a latent class model and a continuous distribution leads to a random coefficients model (Wedel et al., 1999). A third and more flexible representation of unobserved heterogeneity is the RCM model (Allenby et al., 1998; Lenk & DeSarbo, 2000), which combines the flexibility of the semi-parametric latent class model and allows for parametric variation within the latent classes through the formulation of random coefficients.[1] The RCM model can capture skewed and multimodal distributions of unobserved heterogeneity and can accommodate both discrete and continuous forms of unobserved heterogeneity in the data (Lenk & DeSarbo, 2000).

[1] This model is also known as the Bayesian mixture approach and the mixture-of-normals model.

Please cite this article as: Kappe, E., et al., A random coefficients mixture hidden Markov model for marketing research, *International Journal of Research in Marketing* (2018), https://doi.org/10.1016/j.ijresmar.2018.07.002

Past marketing research shows that the RCM model outperforms its nested random coefficients and latent class models (Chandukala, Long-Tolbert, & Allenby, 2011; Mehta, Chen, & Narasimhan, 2010; Michalek et al., 2011; Otter et al., 2004).

Note that other flexible specifications exist to capture unobserved heterogeneity. The simplest way to account for unobserved heterogeneity is to include fixed or random effects to control for individual differences; however, these models assume that the effects of independent variables are similar across individuals. Although not yet employed in HMMs, more sophisticated approaches exist. For example, researchers can use Dirichlet process priors (e.g., Ansari & Mela, 2003; Voleti & Ghosh, 2013) to model the distribution of the population parameters in a semiparametric way (instead of assuming a parametric normal distribution, which is common in modeling random coefficients). Another flexible specification is a model where each coefficient has its own finite mixture (Ebbes, Liechty, & Grewal, 2015). Both approaches are promising and can be incorporated within an HMM. However, we focus on the RCM specification because it provides the additional flexibility needed to nest all unobserved heterogeneity structures found in prior HMMs in the marketing literature—enabling researchers to directly compare model performance against all existing HMMs.

2.2. HMMs

An HMM consists of three parts: the state-dependent equation, the transition probability equation, and the initial state probabilities (Hamilton, 1989; Zucchini, MacDonald, & Langrock, 2016). We begin by specifying the linear state-dependent equation:

$$Y_{it} \mid (S_{it} = s) = \beta_{is}' X_{it} + \varepsilon_{it}, \tag{1}$$

where Y_{it} is the dependent variable for individual i at time t, conditional on individual i belonging to hidden state $s \in \{1,\ldots,S\}$ at time t. β_{is} is a Kx1 individual- and state-specific vector of parameters measuring the impact of the Kx1 vector of explanatory variables X_{it}. Finally, ε_{it} represents the error term.

The states are unobserved and follow a first-order Markov process. The transition probability, $p_{itss'}$, represents the probability of switching from state s to s' for individual i at time t. We define the transition probability matrix as:

$$P_{i,t-1\rightarrow t} = \begin{bmatrix} p_{it11} & \cdots & p_{it1S} \\ \vdots & \ddots & \vdots \\ p_{itS1} & \cdots & p_{itSS} \end{bmatrix}, \tag{2}$$

with $0 \le p_{itss'} \le 1$ and $\sum_{l=1}^{S} p_{itsl} = 1, \forall s$. We specify the transition probabilities as a multinomial logit model, where the transition parameters are a function of time-varying explanatory variables and individuals can transition freely between hidden states:

$$p_{itss'} = \frac{\exp(\gamma_{iss'}' W_{it})}{1 + \sum_{l=2}^{S} \exp(\gamma_{isl}' W_{it})}, \quad \text{for } s = 1\ldots S, \tag{3}$$

where W_{it} represents a vector of transition variables, including a constant. The parameter vector $\gamma_{iss'}$ measures the impact of these transition variables for a move from state s to state s' for individual i. Finally, we define the initial state probabilities, π_s, where $0 \le \pi_s \le 1$ and $\sum_{s=1}^{S} \pi_s = 1$.

Researchers can incorporate unobserved heterogeneity in the HMM in both the state-dependent and transition parameters. At one extreme, we can assume parameters are homogenous across individuals; at the other, we can estimate the model separately for each individual, leading to individual-specific independent parameters. Between these two extremes are more parsimonious alternatives, including random coefficients, latent classes, and RCM.

1. *Homogenous parameters*: This model assumes all parameters are similar across individuals. This means that $\beta_{is} = \beta_s$ in Eq. (1) and that $\gamma_{iss'} = \gamma_{ss'}$ in Eq. (3).
2. *Random coefficients*: This model assumes that parameters differ across individuals and come from an underlying continuous distribution. The individual-specific parameters typically follow a multivariate normal distribution, which means that $\beta_{is} \sim MVN(\overline{\beta_s}, \Sigma_{\beta s})$ in Eq. (1) and that $\gamma_{iss'} \sim MVN(\overline{\gamma_{ss'}}, \Sigma_{\gamma ss'})$ in Eq. (3). When the constant is the only individual-specific parameter, this model results in a random-effects model.
3. *Latent classes*: This model assumes that parameters in the HMM differ across individuals, each of whom belongs to one of $m = 1\ldots M$ latent classes. This means that $\beta_{is} = \beta_{ms}$ in Eq. (1) and that $\gamma_{iss'} = \gamma_{mss'}$ in Eq. (3).
4. *Individual-specific independent parameters*: This model assumes that parameters are estimated separately for each individual and that no relationship exists between parameters across individuals. This results in the model outlined in Eqs. (1–3). When the constant is the only unique parameter across individuals, it is equivalent to including fixed effects.
5. *RCM*: This model assumes that parameters differ across individuals, each of whom belongs to one of $m = 1\ldots M$ latent classes. The parameters within each latent class come from an underlying continuous distribution, which results in random coefficients within each latent class. This means that $\beta_{is} = \beta_{ims}$ in Eq. (1) and that $\gamma_{iss'} = \gamma_{imss'}$ in Eq. (3). Assuming these parameters come from a multivariate normal distribution, we obtain:

Please cite this article as: Kappe, E., et al., A random coefficients mixture hidden Markov model for marketing research, *International Journal of Research in Marketing* (2018), https://doi.org/10.1016/j.ijresmar.2018.07.002

$$\beta_{ims} \sim MVN\left(\overline{\beta_{ms}}, \Sigma_{\beta ms}\right) \tag{4}$$

in Eq. (1) and

$$\gamma_{imss'} \sim MVN\left(\overline{\gamma_{mss'}}, \Sigma_{\gamma mss'}\right) \tag{5}$$

in Eq. (3) for each latent class $m = 1 \ldots M$ and hidden state $s = 1 \ldots S$. We refer to this model as the RCMHMM. This model nests all other unobserved heterogeneity structures discussed above. If $M = 1$ and $\Sigma_{\beta ms} = \Sigma_{\gamma mss'} = 0$, we get homogenous parameters. If $M = 1$, we get random coefficients. If $\Sigma_{\beta ms} = \Sigma_{\gamma mss'} = 0$, we get latent classes. Finally, if each individual belongs to its own "latent" class and $\Sigma_{\beta ms} = \Sigma_{\gamma mss'} = 0$, we get individual-specific independent parameters.

2.3. Classification of marketing literature

We performed an extensive search for HMMs in the marketing literature and classified the unobserved heterogeneity structures of each HMM. We began with a systematic search for all articles published in the *International Journal of Research in Marketing*, *Journal of Marketing*, *Journal of Marketing Research*, *Marketing Science*, *Management Science* (marketing-related studies only), and *Quantitative Marketing and Economics* that estimate an HMM.[2] We used the following search terms: "*hidden Markov model*," "*latent Markov*," and "*Markov switching*." Next, we checked references in these articles to identify any additional articles. We identified 38 articles and classified each article based on the unobserved heterogeneity in the state-dependent and transition parameters (see Table 1).[3] In total, six of the 16 nested versions exist in the marketing literature.

The earliest HMMs in marketing contain homogenous state-dependent and transition parameters (e.g., Poulsen, 1990). Montgomery et al. (2004) add random coefficients to the state-dependent parameters reflecting a trend in marketing research to account for individual-level differences. Later, Netzer et al. (2008) use random coefficients to model the transition parameters, which allows for time-varying transition probabilities. Montoya et al. (2010) include random coefficients for both the state-dependent and transition parameters. Schweidel and Knox (2013) include latent classes for both the state-dependent and transition parameters. Lee et al. (2002) use individual-specific independent state-dependent and transition parameters. To our knowledge, no marketing publications exemplify the 10 remaining combinations in Table 1.

To summarize, HMMs in past marketing research only incorporate a single unobserved heterogeneity structure in the state-dependent and/or transition parameters and emphasize homogenous parameters and random coefficients. A recent exception estimates an HMM with homogenous state-dependent and transition parameters and tests whether incorporating latent classes improves model performance (Zhang et al., 2016). We extend this literature with the RCMHMM, which nests all prior HMMs in marketing and 10 additional HMMs.

2.4. Estimation

We develop a Bayesian estimation framework to estimate the RCMHMM and its 16 nested versions. We use data augmentation to sample latent classes and hidden states (following Chib, 1996). We use random-walk Metropolis-Hastings steps to sample the state-dependent and transition parameters. We follow Rosenthal (2011) to adapt the step size of the proposal density to ensure that the acceptance rates stay between 0.30 and 0.35. We use Gibbs steps to sample the remaining parameters. See Web Appendix A for the specification of priors and a technical description of the full estimation procedure.

The parameters of the RCMHMM are not uniquely identified, as the posterior is insensitive to permutations of the hidden states and latent classes. To address this issue, we order the constants for the transition parameters in each state by size to prevent label switching between hidden states (cf., Kumar et al., 2011) (see Step 4 in Web Appendix A for details). We use an unrestricted sampler to sample the latent classes and inspect the MCMC output post hoc to determine an appropriate identifiability constraint (Frühwirth-Schnatter, 2001) that uniquely determines the latent classes. We use the first state of the transition matrix as the reference state and assume that the parameters governing the transition to the first state equal zero.

2.5. Model comparison

We use a predictive information criterion to compare the in-sample performance of the RCMHMM and its nested models. Specifically, we use the widely applicable information criterion (WAIC) (Watanabe, 2010), which selects the best model from a larger set based on the Kullback-Leibler information measure. The WAIC computes model fit based on the log pointwise predictive density adjusted by a penalty for the number of parameters due to overfitting. The WAIC is appropriate to compare model performance for several reasons. First, the WAIC is asymptotically similar to Bayesian leave-one-out cross-validation—the correlation between the WAIC and leave-one-out cross-validation was 0.996 in a simulation experiment using a finite sample of 200

[2] We excluded articles that do not empirically estimate an HMM or estimate a restricted HMM. For example, Platzer and Reutterer (2016) estimate a buy-till-you-die (BTYD) model, which is a restricted HMM with an absorbing state (however, BTYD models do not use unobserved heterogeneity structures outside of those already listed in Table 1).

[3] The only article we could not classify was Erdem, Imai, and Keane (2003), where the authors propose a restricted Markov switching process for households' usage rate of frequently purchased consumer goods. The transition parameters are homogenous and the state-dependent parameters follow an RCM distribution.

Table 1
Classification of unobserved heterogeneity structures in HMMs in marketing.

State-dependent parameters	Transition parameters			
	Homogenous parameters	Random coefficients	Latent classes	Individual-specific independent parameters
Homogenous parameters	Böckenholt and Dillon (1997) Brangule-Vlagsma, Pieters, and Wedel (2002) Du and Kamakura (2006) Ebbes, Grewal, and DeSarbo (2010) Huang, Singh, and Ghose (2015) Poulsen (1990) Ramaswamy (1997) Schweidel, Bradlow, and Fader (2011) Shi, Wedel, and Pieters (2013) Zhang and Feng (2011) Zhang, Watson IV, Palmatier, and Dant (2016)	Netzer et al. (2008) Shi and Zhang (2014)	–	–
Random coefficients	Kumar, Sriram, Luo, and Chintagunta (2011) Lemmens, Croux, and Stremersch (2012) Montgomery et al. (2004) Moon, Kamakura, and Ledolter (2007) Shachat and Wei (2012) van der Lans, Pieters, and Wedel (2008) Zhang, Kumar, and Cosguner (2017)	Ansari, Montoya, and Netzer (2012) Ascarza and Hardie (2013) Holtrop, Wieringa, Gijsenberg, and Verhoef (2017) Hui (2017) Li, Sun, and Montgomery (2011) Luo and Kumar (2013) Ma, Sun, and Kekre (2015) Mehta, Ni, Srinivasan, and Sun (2017) Montoya et al. (2010) Park and Gupta (2011) Schwartz, Bradlow, and Fader (2014) Schweidel, Park, and Jamal (2014) Stüttgen, Boatwright, and Monroe (2012) Yang, Zhao, and Dhar (2010) Zhang, Netzer, and Ansari (2014)	–	–
Latent classes	–	–	Schweidel and Knox (2013) Zhang et al. (2016)	–
Individual-specific independent parameters	–	–	–	Lee, Sudhir, and Steckel (2002)

– Indicates that the model corresponding to that cell is new to the marketing literature.

observations (Watanabe, 2010). Given that our sample sizes are much larger, we can rely on the asymptotics and use the WAIC as a single criterion that takes both in-sample and out-of-sample fit into account. Second, the WAIC is a fully Bayesian information criterion which takes all MCMC draws as well as any prior information into account (Gelman et al., 2013). Third, the WAIC is a robust criterion for singular models (e.g., HMMs, mixture models, and the RCMHMM). Alternative information criteria (e.g., AIC, BIC, DIC) are not robust (and therefore, may be unreliable) for model selection when dealing with singular models (Watanabe, 2010). That being said, we also use the DIC (the most common metric to evaluate HMM performance in the marketing literature) and the root mean squared prediction error (RMSPE) (which tests out-of-sample performance) to test the robustness of our results.

3. Simulation studies: design

We design two simulation studies to answer the following questions: 1) Under which conditions does the RCMHMM beat its nested model versions? 2) Under which conditions is the additional flexibility of the RCMHMM harmful? 3) Under which conditions does the RCMHMM perform relatively better or worse? and 4) How robust is the RCMHMM when underlying model assumptions are violated?

We answer the first two questions in our first simulation study by generating 17 datasets that correspond to the RCMHMM and its 16 nested versions for 50 individuals and 50 time periods. We estimate each of the 17 models on the 17 datasets and

Please cite this article as: Kappe, E., et al., A random coefficients mixture hidden Markov model for marketing research, *International Journal of Research in Marketing* (2018), https://doi.org/10.1016/j.ijresmar.2018.07.002

Table 2
Model selection for the 17 datasets in simulation study 1.

DGP	Estimated model																
	Hom, Hom	Hom, RC	Hom, LC	Hom, Het	RC, Hom	RC, RC	RC, LC	RC, Het	LC, Hom	LC, RC	LC, LC	LC, Het	Het, Hom	Het, RC	Het, LC	Het, Het	RCM[b]
Hom, Hom	3	1[a]	3[a]	8	11	9	11	13	3[a]	1[a]	3[a]	7	14	16	14	17	9
Hom, RC	14	3	7	4	16	6	12	8	10	1[a]	13	2[a]	17	9	15	11	5
Hom, LC	13	3	1	5	16	7	6	10	13	2	13	4	17	11	9	12	7
Hom, Het	14	1[a]	3[a]	4	16	6	8	10	13	1[a]	14	4[a]	17	9	11	12	6
RC, Hom	14	16	14	17	1	3[a]	1[a]	8	13	10	12	11	5	7	5	9	3[a]
RC, RC	17	12	16	13	8	1	6	3[a]	14	10	15	11	9	5	7	4	1[a]
RC, LC	16	14	13	15	9	3	1	5	16	8	12	10	11	7	2	6	3
RC, Het	17	12	13	14	9	1[a]	6	4	16	8	15	11	10	7	3[a]	5	1[a]
LC, Hom	16	15	14	17	7	5	7	12	1	2[a]	3	4	9	11	9	13	5
LC, RC	17	15	14	16	12	7	10	5	9	1	8	2[a]	13	4	11	6	3
LC, LC	15	14	12	17	13	8	5	10	11	2	1	3	16	7	6	9	4
LC, Het	17	14	15	16	12	3	6	8	11	1[a]	10	2	13	5	7	9	3
Het, Hom	13	17	13	16	6[a]	1[a]	3[a]	8	13	12	11	10	4	7	4[a]	9	1[a]
Het, RC	15	13	14	11	8	3[a]	5	2[a]	15	10	17	11	9	1	6	7	3[a]
Het, LC	16	12	14	11	9	3	1[a]	7	15	12	16	10	8	5	2	6	3
Het, Het	15	11	14	12	8	1[a]	4[a]	7[a]	15	10	17	12	9	3[a]	5[a]	6	1[a]
RCM	17	14	16	15	8	2	7	3	11	9	12	10	13	5	6	4	1
Average	14.65	11.00	11.53	12.41	9.94	4.06	5.82	7.24	11.71	5.88	11.29	7.29	11.41	7.00	7.18	8.53	3.47

Notes. Hom = homogenous parameters, RC = random coefficients, LC = latent classes, Het = individual-specific independent, and RCM = random coefficients mixture. The first abbreviation refers to the state-dependent parameters and the second refers to the transition parameters (e.g., "Hom, RC" has homogenous state-dependent parameters and random coefficients for the transition parameters.) We rank the 17 estimated models from 1 = the best to 17 = the worst using the WAIC. When one model is perfectly nested within another (e.g., the RCMHMM with one latent class reduces to the RC, RC model), the model fit (and thus, rank) are the same.

[a] Indicates that the fit of the estimated model is not significantly different from that of the true model.

[b] For the estimated RCM models, two hidden states and one latent class fit the data best for all datasets, except when the DGP is Hom, RC; LC, RC; LC, LC; or RCM. In these four cases, two hidden states and two latent classes fit the data best.

determine the optimal number of latent classes (where applicable) and hidden states for each model and compare model performance. See Web Appendix B for details on the choice of priors and parameter values.

We answer the third and fourth question in our second simulation study by modifying the data dimensions from the first study and violating two underlying model assumptions. Specifically, we manipulate five factors (cf. Andrews et al., 2002; Ebbes et al., 2015; Vriens, Wedel, & Wilms, 1996). The first factor is the data generating process (DGP). To keep the simulation study manageable, we generate data using: homogeneous, random coefficients, latent class, individual-specific independent, and RCM state-dependent and transition parameters (i.e., the four most representative HMMs in marketing—see the diagonal in Table 1—plus the RCMHMM). The second factor is the number of individuals in the data: 25 (low) vs. 100 individuals (high). The third factor is the number of time periods: 25 (low) vs. 100 time periods (high). The second and third factors allow us to assess the RCMHMM's relative performance as data dimensions change. The fourth factor is the presence of an omitted variable: absent vs. present. Past research shows that HMMs can account for an omitted variable in the hidden states (Moon et al., 2007), so we test whether the RCMHMM and its 16 nested versions can detect the effects of an omitted variable. The fifth factor is the error term distribution: normal vs. gamma. The fourth and fifth factors violate two of the underlying assumptions of the RCMHMM and allow us to test relative performance when the model is misspecified.[4]

Combining all five factors leads to a 5 (DGPs) × 2 (individuals) × 2 (time periods) × 2 (omitted variable) × 2 (error term) design resulting in 80 conditions (i.e., 80 different datasets). As estimating all five models (corresponding to the five DGPs) on each condition is computationally expensive, we reduce the number of models we estimate by generating a fractional factorial design (cf., Bodapati & Gupta, 2004; Carmone Jr, Ali, & Maxwell, 1999). This allows us to estimate the main effects for all factor levels applied to each of the five models. This process results in 16 conditions. Applying each of the five models to these 16 conditions results in the estimation of 80 different models, which we replicate four times each (for a total of 320 models).

4. Simulation studies: results

The results for both simulation studies are based on 100,000 iterations of the MCMC sampler—half of which we discarded for burn-in. We retained every 25th draw (due to storage space constraints) and found no significant autocorrelation in the draws after thinning. The Gelman and Rubin (1992) test statistic for convergence was below the 1.1 threshold for over 90% of the parameters in each HMM based on four independent chains with dispersed starting values, indicating convergence.

[4] We do not manipulate factors related to the amount of heterogeneity within segments or the separation of the mixture components (e.g., Ansari & Mela, 2003; Ebbes et al., 2015), as these factors only apply to the models with within-segment heterogeneity or multiple segments, respectively.

Please cite this article as: Kappe, E., et al., A random coefficients mixture hidden Markov model for marketing research, *International Journal of Research in Marketing* (2018), https://doi.org/10.1016/j.ijresmar.2018.07.002

Table 3
Model selection for the RCMHMM dataset in simulation study 1.

State-dependent parameters	Transition parameters	WAIC
Homogenous	Homogenous ($s = 8$)	9813.08 (17)
	Random coefficients ($s = 5$)	9622.01 (14)
	Latent classes ($m = 7, s = 2$)	9680.12 (16)
	Individual-specific independent ($s = 8$)	9624.34 (15)
Random coefficients	Homogenous ($s = 4$)	8552.55 (8)
	Random coefficients ($s = 2$)	8206.39 (2)
	Latent classes ($m = 4, s = 3$)	8392.63 (7)
	Individual-specific independent ($s = 2$)	8217.91 (3)
Latent classes	Homogenous ($m = 4, s = 4$)	8949.88 (11)
	Random coefficients ($m = 6, s = 2$)	8691.87 (9)
	Latent classes ($m = 4, s = 2$)	9028.49 (12)
	Individual-specific independent ($m = 10, s = 2$)	8692.12 (10)
Individual-specific independent	Homogenous ($s = 2$)	9053.38 (13)
	Random coefficients ($s = 2$)	8222.75 (5)
	Latent classes ($m = 6, s = 2$)	8295.95 (6)
	Individual-specific independent ($s = 2$)	8217.97 (4)
Random coefficients mixture	Random coefficients mixture ($m = 2, s = 2$)	8171.70 (1)

Note. We provide the optimal number of latent classes (m) and hidden states (s) and rank each model from 1 = the best to 17 = the worst (in parentheses).

4.1. Simulation study 1

We rank the 17 estimated models on each of the 17 simulated datasets (DGPs) in Table 2 using the WAIC (from 1 = best to 17 = worst). Tables WA-B2 and WA-B3 in Web Appendix B show the corresponding results using the DIC and RMSPE. Results based on the RMSPE are very similar, and while results based on the DIC lead to similar conclusions, the DIC is less able to identify the true model. Returning to Table 2, in cases where one model is perfectly nested within another (e.g., the RCMHMM with one latent class reduces to the RC, RC model), the model fit (and thus, rank) are the same. For nine of the 17 cases, the best model is the true model (i.e., there is a "1" in the Table 2). For the remaining eight cases, we calculate Akaike weights (Burnham & Anderson, 2004) to determine whether the difference between WAIC values for the best and true model is significant. If the difference is >5.89, we conclude that the difference is significant at the 95% level. For six of the remaining cases, the difference between the best and true models is not significant. For example, when the true model has homogenous state-dependent and transition parameters (Hom, Hom row in Table 2), the model with homogenous state-dependent parameters and random coefficients for the transition parameters (Hom, RC column) performs better. However, the WAIC for this model is not significantly different from the true model. Upon inspection, the population means of the transition parameters do not significantly differ from the true model (and the population variance is near zero). For two of these eight cases, the WAIC for the true model was significantly worse than the WAIC for the best model. In both cases, the true model includes individual-specific independent state-dependent parameters (Het, Hom and Het, Het rows) and the best model includes random coefficients or RCM for both the state-dependent and transition parameters (RC, RC and RCM columns). This is not entirely unexpected as the random-coefficients distribution is much more parsimonious than the individual-specific independent parameters and this result depends on how the individual-specific independent parameters are generated.

Next, we examine the RCMHMM simulated dataset (DGP) more closely. We provide the optimal number of latent classes and hidden states as well as the corresponding WAIC values and rank for all 17 models (from 1 = best to 17 = worst) in Table 3 (the RCM row in Table 2 is the same as the last column in Table 3). The RCMHMM fits the data best and recovers the true parameters, latent classes, and hidden states. The 16 nested models only recover the state-independent parameters correctly. All other parameters were biased.[5] Web Appendix B discusses the bias of the marketing parameters in these models in more detail. We also note that while the RCMHMM is the most complex model, it leads to one of the most parsimonious solutions. The 16 nested models include up to 10 latent classes and eight hidden states to compensate for the misspecification of unobserved heterogeneity.

Finally, we can answer the two questions that motivated this simulation study. First: Under which conditions does the RCMHMM beat its nested model versions? The RCMHMM performs best across all DGPs with an average rank of 3.47 (see Table 2). The second-best model is the HMM with random coefficients for both the state-dependent and transition parameters with an average rank of 4.06. It is encouraging that this model is second, given that it is currently the state-of-the-art HMM in marketing. Comparing these two models, the RCMHMM performs better when the unobserved heterogeneity structure is multimodal (i.e., the data includes latent classes or an RCM distribution). When the unobserved heterogeneity structure is unimodal, both models perform similarly (as the RCMHMM perfectly nests the HMM with random-coefficients state-dependent and transition parameters). Hence, if there is evidence that the unobserved heterogeneity is unimodal, there is no need for the researcher to specify the more complex RCMHMM and a simpler model would suffice.

[5] Because we estimate our models with Bayesian techniques, the term "distortion" is more precise than "bias;" but because the insights also apply to non-Bayesian estimation situations, we continue to use the term "bias."

Please cite this article as: Kappe, E., et al., A random coefficients mixture hidden Markov model for marketing research, *International Journal of Research in Marketing* (2018), https://doi.org/10.1016/j.ijresmar.2018.07.002

Second: Under which conditions is the additional flexibility of the RCMHMM harmful? The RCMHMM performs relatively worse than the true model when there is little unobserved heterogeneity in the data. Specifically, the RCMHMM performs worse than the true model when the DGP includes either homogeneous parameters or latent classes for the state-dependent parameters or latent classes for the transition parameters (see the superscript a in the last column of Table 2). However, parameter estimates based on the RCMHMM only show a slight bias. This is because the RCMHMM assumes some level of heterogeneity due the presence of prior hyperparameters. Thus, we conclude that RCMHMM performance improves as the unobserved heterogeneity structure becomes more complex (either through random coefficients or individual-specific independent parameters). However, when this is not the case, choosing a more diffuse prior mitigates this problem.

4.2. Simulation study 2

We estimate five different models (corresponding to the four diagonal cells in Table 1 and the RCMHMM) on each condition in the fractional factorial design. We select the optimal number of latent classes and hidden states for each model using the WAIC. Because the size of each dataset differs across conditions, we compute the average WAIC per observation for each model. We use the average WAIC per observation for each run i of the simulation study ($\overline{WAIC_i}$) as the dependent variable and the five simulation factors as independent variables to assess how each factor affects model performance. This leads to the following regression model.

$$\overline{WAIC}_i = \mu_1 DGPHom_i + \mu_2 DGPRC_i + \mu_3 DGPLC_i + \mu_4 DGPHet_i + \mu_5 DGPRCM_i + \mu_6 100Ind_i + \mu_7 100Time_i + \mu_8 Omitted_i + \mu_9 Gamma_i + \omega_i.$$

All explanatory variables are dummies. The first five explanatory variables refer to the five DGPs. We omit a constant in this model and rather include all five DGPs to facilitate the comparison of the parameters across the five estimated models. The last four explanatory variables represent, respectively, the datasets with 100 (vs. 25) individuals, 100 (vs. 25) time periods, the presence (vs. absence) of an omitted variable, and an error term with a Gamma (vs. normal) distribution. We run this model separately for each of the five estimated models and the results are in Table 4 (where lower values indicate better performance). For example, the first row of Table 4 shows how each of the five estimated models performs when the DGP has homogenous state-dependent and transition parameters. Specifically, the HMM with homogenous or latent-class parameters (3.18) performs best and the HMM with individual-specific independent parameters (3.27) performs worst. In Table 4, we also indicate whether the parameters for the first four estimated models (in the first four columns) are significantly different from the corresponding parameter for the RCMHMM (in the last column). For the first row, we find that when the DGP has homogenous parameters, the estimated models with homogenous parameters and latent classes fit the data significantly better than the RCMHMM. The model with individual-specific independent parameters fits the data significantly worse than the RCMHMM.

We can use the results in Table 4 to answer the two questions that motivated this simulation study. First: Under which conditions does the RCMHMM perform relatively better or worse? HMMs with random coefficients, individual-specific independent, and RCM state-dependent and transition parameters are relatively robust to the DGP, whereas HMMs with homogenous parameters or latent classes for the state-dependent and transition parameters perform worse. The RCMHMM performs best overall—scoring either first or second (after the true model). Interestingly, the RCMHMM performs better than the HMM with individual-specific independent parameters when the DGP has individual-specific independent parameters. Regarding data dimensions, the number of individuals does not affect model performance (consistent with Ebbes et al., 2015). However, all models except the HMM with random coefficients perform significantly better with more time periods (a negative coefficient indicates an increase in model performance). Improvement is greatest for the RCMHMM (−0.46), indicating the RCMHMM performs better with more time periods.

Table 4
Parameter estimates for simulation study 2.

Manipulated factors	Estimated model				
	Homogenous	Random coefficients	Latent classes	Individual-specific independent	Random-coefficients mixture
Homogenous DGP	3.18[a,b]	3.23[a]	3.18[a,b]	3.27[a,b]	3.23[a]
Random coefficients DGP	3.73[a,b]	3.27[a]	3.64[a,b]	3.27[a]	3.27[a]
Latent classes DGP	3.46[a,b]	3.32[a,b]	3.15[a,b]	3.31[a,b]	3.19[a]
Individual-specific independent DGP	4.73[a,b]	3.34[a,b]	4.79[a,b]	3.28[a,b]	3.24[a]
Random coefficients mixture DGP	3.93[a,b]	3.42[a,b]	3.61[a,b]	3.41[a,b]	3.27[a]
100 individuals	0.03	−0.05	−0.03	0.06	0.01
100 time periods	−0.17[a,b]	−0.13[b]	−0.22[a,b]	−0.32[a,b]	−0.46[a]
Omitted variable present	4.12[a,b]	3.88[a,b]	4.01[a,b]	3.77[a,b]	3.52[a]
Gamma distribution	0.64[a,b]	0.36[a]	0.59[a,b]	0.64[a,b]	0.25[a]

Notes. The dependent variable is the average WAIC per observation. Parameter estimates indicate how each factor affects model performance. Lower values indicate better performance.

[a] Indicates that the parameter is significantly different from zero at $\alpha = 0.05$.

[b] Indicates that the parameter for the corresponding estimated model is significantly different from the corresponding parameter for the random-coefficients mixture HMM at $\alpha = 0.05$.

Please cite this article as: Kappe, E., et al., A random coefficients mixture hidden Markov model for marketing research, *International Journal of Research in Marketing* (2018), https://doi.org/10.1016/j.ijresmar.2018.07.002

Second: How robust is the RCMHMM when the underlying model assumptions are violated? The RCMHMM performs better than other models when an omitted variable is present in the DGP (i.e., it has the smallest coefficient among all estimated models, indicating that model performance suffers least for the RCMHMM. The RCMHMM also performs significantly better than other models, except the HMM with random coefficients, when researchers specify the incorrect error distribution. Hence, we conclude that the RCMHMM is relatively robust when underlying model assumptions are violated.

5. Empirical application: Major League Baseball (MLB)

We focus on the effectiveness of in-game promotions in increasing the short-term demand for MLB attendance. MLB consists of 30 professional baseball teams in the U.S. and Canada, each of which plays a 162-game regular season (81 home games and 81 away games). Methodologically, this is an ideal context for the RCMHMM, as the demand for attendance can vary both over the course of the season (facilitating hidden states) and across teams (facilitating latent classes). Practically, regular-season attendance fell to 74,026,895 in 2013 (MLB, 2013)—just over 70% of total stadium capacity. Multiplying the number of unsold tickets by the average ticket price of $27.48 (Team Marketing Report, 2013), MLB forfeited over $850 million in potential revenue. Thus, the demand for attendance has important implications for MLB.

5.1. The data

We collected data related to the demand for MLB attendance for all 2430 games played during the 2013 regular season.[6] Past attendance research typically uses either raw attendance (Boyd & Krehbiel, 1999, 2003, 2006; DeSarbo, Hwang, Stadler Blank, & Kappe, 2015; Kappe, Stadler Blank, & DeSarbo, 2014; Lemke et al., 2010) or the log of raw attendance (Borland & Lye, 1992; Lemke et al., 2010; Peel & Thomas, 1988; Siegfried & Eisenberg, 1980) as the dependent variable. We use the log of attendance because it "is easy to interpret and explains a greater fraction of the variation in attendance" (Siegfried & Eisenberg, 1980, p. 60), allows us to interpret parameters as elasticities, and allows us to compare estimates across teams, which all benefit the managerial interpretability of results.[7]

Our primary interest is the effectiveness of in-game promotions in increasing short-term demand for attendance. We are specifically interested in promotions because it is one of the few variables MLB teams can control. We recorded the promotions listed on each team's website every day of the 2013 season and classified each promotion into one of four promotional categories (i.e., entertainment, event, giveaway, or price promotions) to conserve degrees of freedom (see Table 5). We observe substantial variation across teams in terms of the number of promotions employed in each category as well as total promotions (see Table 6).

There is some preliminary evidence that promotional effectiveness varies over the course of the season and across teams. First, promotional effectiveness may vary over the course of the season because of changes in fans' preferences to attend games (e.g., due to team performance or distinct milestones within the season—the MLB draft, All-Star break, trade deadline, mathematical elimination from playoff contention, etc.). This argument is akin to differential effectiveness of marketing mix variables over a product's lifecycle (Narayanan, Manchanda, & Chintagunta, 2005; Parsons, 1975; Thietart & Vivas, 1984). Supporting this notion, Lemke et al. (2010) find that giveaways are more effective in April and May (vs. June through August) and fireworks are more effective in June through August (vs. September); McDonald and Rascher (2000) demonstrate that increasing the number of promotions decreases the marginal effect of each promotion over the course of the season; and Boyd and Krehbiel (1999) show that the effectiveness of promotions can differ across multiple seasons. Second, promotional effectiveness may vary across teams because teams operate in different markets with distinct fan bases. Indeed, past research shows that the effectiveness of in-game promotions varies by team (Boyd & Krehbiel, 1999, 2003) and that promotions are more effective for small-market and poorer-performing teams (Lemke et al., 2010; Marcum & Greenstein, 1985). While this work offers initial support that promotional effectiveness varies over the course of the season and across teams, it tests variation via split-sample regressions (i.e., the unobserved heterogeneity structure is imposed a priori). Because an incorrect specification of the unobserved heterogeneity structure biases parameters and managerial insights, we allow these short-term promotional effects to vary over time and across teams in the RCMHMM.

Finally, we gathered (and control for) a number of other variables related to the demand for MLB attendance (DeSarbo et al., 2015; Lemke et al., 2010) (see Table 7).

5.2. Model specification

Note that attendance is censored at stadium capacity when a game is sold out, making it important to account for right-censoring (Lemke et al., 2010) in the state-dependent equation[8]:

$$D^*_{it} \mid (S_{it} = s) = \beta'_{ims} X_{it} + \delta' Z_{it} + \varepsilon_{it}, \tag{6}$$

[6] MLB defines attendance as ticket sales (vs. fans who attend). We classified one CIN home game as an SFG home game because the game was played in San Francisco. Secondary data came from: www.mlb.com, www.baseball-reference.com, www.pro-football-reference.com, www.hockey-reference.com, www.mls.com, www.sportsbookreview.com, www.wunderground.com, www.teammarketing.com, and Badenhausen, Ozanian, and Settimi (2013).

[7] While percent of stadium capacity is another option, variation in stadium size (e.g., OAK = 34,077 vs. LAD = 56,000) may cause this variable to be misleading in that two identical percentages may have different managerial and financial implications.

[8] We also right-censored approximately 5% of observations above stadium capacity (which occurs when teams sell standing room only tickets).

Table 5
MLB promotions by category.

Entertainment	Event	Giveaway	Price
Concert	Appreciation	Accessory	Discounted food and beverage
Festival	Autographs	Automotive	Discounted tickets
Fireworks	Awareness	Bag	Ticket package with food and beverage
	Charity	Baseball	Ticket package with giveaway
	Education	Beverage	
	Faith	Bobblehead	
	History	Card	
	Kids/family	Clothing	
	Military/first responders	Coupon	
	On-field	Figurine	
	Pet	Headwear	
	Run the bases	Jersey	
	Student	Kids	
	Theme	Memorabilia	
		Miscellaneous	
		Rally	
		Schedule	
		Technology	
		T-shirt	
		Wall hanging	

$$Att_{it} = D^*_{it} \text{ if } D^*_{it} < C_i,$$

$$Att_{it} = C_i \text{ if } D^*_{it} \geq C_i, \text{ and}$$

$$\beta_{ims} \sim N\left(\overline{\beta_{ms}}, \Sigma_{\beta ms} I\right). \tag{7}$$

Table 6
Descriptive statistics for MLB attendance and promotions.

Team	Average attendance	Stadium capacity	No. of entertainment promotions	No. of event promotions	No. of giveaway promotions	No. of price promotions	No. of total promotions
Arizona Diamondbacks (ARI)	26,355	48,633	19	37	23	29	108
Atlanta Braves (ATL)	31,465	49,377	27	68	9	38	142
Baltimore Orioles (BAL)	29,106	45,971	4	55	15	18	92
Boston Red Sox (BOS)	34,979	37,495	0	116	0	16	132
Chicago Cubs (CHC)	32,626	41,160	0	26	45	4	75
Chicago White Sox (CHW)	22,186	40,615	20	32	23	1	76
Cincinnati Reds (CIN)	31,151	42,319	18	13	25	33	89
Cleveland Indians (CLE)	19,706	42,241	36	37	18	15	106
Colorado Rockies (COL)	34,492	50,398	6	31	41	37	115
Detroit Tigers (DET)	38,067	41,255	20	38	40	4	102
Houston Astros (HOU)	20,394	40,963	18	18	25	34	95
Kansas City Royals (KCR)	21,614	37,903	26	57	27	37	147
Los Angeles Angels (LAA)	37,278	45,957	20	16	27	117	180
Los Angeles Dodgers (LAD)	46,216	56,000	28	15	41	0	84
Miami Marlins (MIA)	19,584	37,000	26	88	14	98	226
Milwaukee Brewers (MIL)	31,248	41,900	0	19	29	90	138
Minnesota Twins (MIN)	30,588	39,021	12	53	57	44	166
New York Mets (NYM)	26,677	41,480	6	105	37	94	242
New York Yankees (NYY)	40,489	50,281	0	3	43	0	46
Oakland Athletics (OAK)	22,337	34,077	8	25	29	24	86
Philadelphia Phillies (PHI)	37,190	43,500	4	62	32	26	124
Pittsburgh Pirates (PIT)	28,281	38,362	11	16	45	3	75
San Diego Padres (SDP)	26,749	43,633	9	48	19	10	86
Seattle Mariners (SEA)	21,747	47,860	4	28	30	50	112
San Francisco Giants (SFG)	41,603	41,915	2	10	30	2	44
St. Louis Cardinals (STL)	41,602	46,861	26	36	36	11	109
Tampa Bay Rays (TBR)	18,646	36,973	8	5	22	26	61
Texas Rangers (TEX)	38,710	48,114	18	16	24	25	83
Toronto Blue Jays (TOR)	31,316	48,278	9	37	15	8	69
Washington Nationals (WSN)	32,746	41,418	8	38	15	94	155
Mean	30,505	43,365	13	38	28	33	112

Please cite this article as: Kappe, E., et al., A random coefficients mixture hidden Markov model for marketing research, *International Journal of Research in Marketing* (2018), https://doi.org/10.1016/j.ijresmar.2018.07.002

Table 7
MLB attendance drivers.

Attendance driver	Predictor variable(s)
Broadcast and electronic media	Game broadcast on network TV dummy
Game schedule characteristics	Day (vs. evening) game dummy
	Day of the week dummies
	Month dummies
	Holiday dummy
Home team	Home team dummies
Opponent	Opponent dummies
	Distance from home team in thousands of miles
Performance	Home team winning percentage (0 to 1)
	Opponent winning percentage (0 to 1)
	Home team number of games back (vs. division leader)[a]
	Opponent number of games back (vs. division leader)[a]
	Probability that the home team wins (0 to 1)[b]
	Competitiveness of the game (0 to 0.5)[b]
Pricing	Ticket price in USD ($)[c]
Promotions	Number of entertainment promotions
	Number of event promotions
	Number of giveaway promotions
	Number of price promotions
Substitute forms of entertainment	Number of competing MLB, MLS, NBA, NFL, or NHL events
Weather	Temperature in °F
	Precipitation in inches

[a] Calculated by: [(division leader wins − team wins) + (team losses − division leader losses)] / 2.
[b] Calculated using the money line for each game; we followed Lemke et al. (2010, pp. 326–327) for the probability that the home teams wins and subtracted 0.5 from its absolute value to determine competitiveness of the game (the closer this value is to zero, the more competitive the game).
[c] Because not all teams used dynamic pricing in 2013, we recorded fixed/variable pricing for each game at the beginning of the season for a representative section of the ballpark (i.e., upper level, centermost section). Although not including dynamic pricing is a limitation, most tickets subject to dynamic pricing (e.g., 90% in Zhu, 2014) are sold before dynamic pricing kicks in (typically about two weeks before the game) (Xu, Fader, & Veeraraghavan, 2016).

Here, $D^*_{it} \mid (S_{it} = s)$ represents the log of the latent demand for attendance, conditional on team i belonging to hidden state $s \in \{1,\ldots,S\}$ at game t. Att_{it} denotes the observed log of attendance and each team has its own right-censoring limit, C_i, which equals the log of the stadium capacity for team i. β_{ims} is a Kx1 vector of parameters for team i in latent class m and hidden state s measuring the impact of the log-transformed, endogenous explanatory variables X_{it}. δ is an Lx1 vector that captures parameters that are common across teams and hidden states and measures the impact of the log-transformed, exogenous explanatory variables Z_{it}. We assume that the error term, ε_{it}, is i.i.d. normally distributed—but relax this assumption when we account for potential endogeneity.

To select variables for the X and Z matrices, we distinguish between our primary variables of interest and the other attendance drivers listed in Table 7. We include the four promotional categories in X and the remaining attendance drivers along with lagged attendance (to capture any unobserved dynamic effects) and fixed effects for each home team and opponent in Z.[9] In Eq. (3), W_{it} represents a vector of transition variables, including a constant. In Eq. (5), $\gamma_{imss'}$ measures the impact of these transition variables for team i, belonging to latent class m, for a move from state s to state s'. To select variables for the W matrix, we build on past attendance research, which suggests that promotional effectiveness may change over time, with team performance, and corresponding playoff propensity (e.g., Lemke et al., 2010; Marcum & Greenstein, 1985). Thus, we test four sets of transition variables—including time (operationalized as home game number), the number of games back the team is from the division leader, the probability that the team reaches the playoffs, and all three variables combined.

5.3. Estimation

We update our estimation routine to allow for the censoring of the dependent variable and correction for endogeneity (which may arise if MLB teams design their promotional schedule strategically). To correct for endogeneity, we include instrumental variables for the promotional variables (Rossi, 2014) and include a full covariance matrix for the error terms (see Web Appendix C for details). To estimate the model, we discard the first home game for each team due to the presence of lagged attendance in our model and follow the procedure in Section 2.4. First, we estimate the model with the appropriate number of latent classes (where applicable) but without hidden states. For the parameters that do not vary across hidden states, we use the resulting estimates as starting values. For the parameters that vary across hidden states, we use different percentiles of the estimates as starting values to aid convergence (see Web Appendix A for details).

[9] The home team fixed effects capture season ticket holders while the promotional effects capture the impact of in-game promotions on non-season ticket holders.

Please cite this article as: Kappe, E., et al., A random coefficients mixture hidden Markov model for marketing research, *International Journal of Research in Marketing* (2018), https://doi.org/10.1016/j.ijresmar.2018.07.002

Table 8
Model selection for empirical application based on the WAIC.

State-dependent parameters	Transition parameters	WAIC
Homogenous parameters	Homogenous parameters ($s = 4$)	−1547.47 (17)
	Random coefficients ($s = 4$)	−1768.40 (12)
	Latent classes ($m = 6, s = 3$)	−1573.49 (16)
	Individual-specific independent parameters ($s = 4$)	−1601.92 (15)
Random coefficients	Homogenous parameters ($s = 2$)	−1903.78 (5)
	Random coefficients ($s = 5$)	−1951.26 (2)
	Latent classes ($m = 5, s = 2$)	−1911.92 (4)
	Individual-specific independent parameters ($s = 3$)	−1922.16 (3)
Latent classes	Homogenous parameters ($m = 6, s = 2$)	−1848.72 (9)
	Random coefficients ($m = 6, s = 2$)	−1864.52 (6)
	Latent classes ($m = 6, s = 2$)	−1854.22 (8)
	Individual-specific independent parameters (m = 5, s = 2)	−1855.62 (7)
Individual-specific independent parameters	Homogenous parameters ($s = 2$)	−1650.75 (14)
	Random coefficients ($s = 3$)	−1833.48 (10)
	Latent classes ($m = 2, s = 2$)	−1794.13 (11)
	Individual-specific independent parameters ($s = 2$)	−1754.67 (13)
RCM	RCM ($m = 2, s = 2$)	**−1966.48 (1)**

Notes. We provide the optimal number of latent classes (m) and hidden states (s) and rank each model from 1 (the best) to 17 (the worst) (in parentheses).

5.4. Results

Results are based on 100,000 iterations of the MCMC sampler—half of which we discarded for burn-in. We retained every 25th draw and found no significant autocorrelation in the draws after thinning. The Gelman and Rubin (1992) test statistic for convergence was below the 1.1 threshold for over 90% of the parameters in each HMM based on three independent chains with dispersed starting values, indicating convergence.

5.4.1. Model selection

First, we select the transition variable and estimate the RCMHMM. Models that include a constant and the number of games back from the division leader as the transition variable perform best according to the WAIC. Next, we estimate all 16 nested model versions, selecting the optimal number of latent classes and hidden states for each model using the WAIC. The results based on the DIC and RMSPE are in Tables WA-D1 and WA-D2 in Web Appendix D. Note that the RMSPE is now based on the in-sample observations. We report and rank model performance based on the WAIC as well as the optimal number of latent classes and hidden states for each model in Table 8. The WAIC, DIC, and RMSPE all point to our proposed RCMHMM with two latent class and two hidden states as the best model (see Table WA-D3 in Web Appendix D for results of the RCMHMM with a varying number of latent classes and hidden states). We also computed posterior model probabilities based on the WAIC (Burnham & Anderson, 2004) and the probability that the HMM with an RCM specification for both the state-dependent and transition parameters is the best model is >0.999.[10] As in our simulation studies, we see that a suboptimal unobserved heterogeneity structure often leads to more latent classes and hidden states. Thus, while the RCMHMM is more complex, it leads to a more parsimonious solution than its nested versions. Finally, we tested five additional models where we account for unobserved heterogeneity in the state-dependent parameters (i.e., homogeneous, random coefficients, latent classes, individual-specific independent, and RCM) but not in the transition parameters (i.e., there is only one hidden state). The RCMHMM performs better than these five models (see Table WA-D4 in Web Appendix D).

5.4.2. Parameter estimates

In this section, we discuss the results of our "best" model—the RCMHMM with two latent classes and two hidden states. See Web Appendix D for parameters that do not vary over time or across teams and the parameters we used to control for endogeneity.

5.4.2.1. Latent class membership. Before interpreting latent classes, we evaluated label switching but did not find much evidence after burn-in in our main model. To interpret the latent classes (see Table 9), we correlated membership probabilities with various team-specific variables. We interpret the latent classes by the team-specific variables that are correlated with latent class membership ($p < .05$). Teams in the first latent class have lower ticket prices ($r = -0.51$), percent stadium capacity filled ($r = -0.39$), and ticket revenue ($r = -0.46$), which results in more wins per dollar of ticket revenue ($r = 0.48$) and total revenue ($r = 0.40$) as well as more wins per dollar spent on player expenses ($r = 0.38$). As we only have two latent classes, the second latent class is the opposite of the first. Based on these correlations, we argue that latent class 1 is more efficient given financial resources. Thus, we label latent class 1 "overachievers," as performance exceeds expectations based on financial resources, and latent class 2 "underachievers," as performance falls short of expectations based on financial resources.

[10] Holdout sample prediction may help when two or more models have similar performance.

Please cite this article as: Kappe, E., et al., A random coefficients mixture hidden Markov model for marketing research, *International Journal of Research in Marketing* (2018), https://doi.org/10.1016/j.ijresmar.2018.07.002

Table 9
Latent class membership.

Latent class 1 (overachievers)	Latent class 2 (underachievers)
ATL[a]	ARI
BAL	BOS[a]
CIN[a]	CHC
CLE[a]	CHW
DET[a]	COL
HOU	LAD[a]
KCR	MIA
LAA	MIL
NYM	MIN
PIT[a]	NYY
SDP	OAK[a]
SEA	PHI
TBR[a]	SFG
TEX	STL[a]
TOR	WSN

[a] Indicates teams that made the 2013 post-season playoffs.

5.4.2.2. Transition parameters. We report the results for the population mean of the transition parameters $\overline{\gamma_{mss'}}$ in Table 10 (see Table WA-D5 in Web Appendix D for the population variance parameters, $\Sigma_{\gamma mss'}$). Recall that the transition variable is the number of games back from the division leader, which is measured using positive numbers. Larger values indicate poorer performance, as the gap between a team and its division leader expands. Teams are 7.5 games back on average. In other words, the average team lost 7.5 more games than its division leader.

We model the probability that a team transitions from state 1 in game *t-1* to state 2 in game *t*, as well as the probability that the team stays in state 2 to achieve identification. Based on the initial state probabilities (0.67 for state 1 and 0.33 for state 2), more teams begin the season in state 1. For both latent classes, the constant for moving from state 1 to state 2 is negative (−2.21 and −0.83, respectively), indicating that division leaders (at zero games back) in state 1 are more likely to stay in state 1 than to move to state 2. The difference between these two constants is significant ($p = .02$). For both overachievers and underachievers, the probability of moving from state 1 to state 2 increases with the number of games back from the division leader (0.30 and 0.09, respectively). The difference between these parameters is not significant ($p = .12$). So as performance declines (relative to the division leader), teams in state 1 are more likely to transition to state 2. After a team transitions to state 2, the probability of staying there increases with the number of games back from the division leader for both overachievers (0.23) and underachievers (0.39). So, as team performance declines (relative to the division leader), teams are more likely to stay in state 2. Based on these findings, we label state 1 "high-performance" and state 2 "low-performance." While more teams begin the season in the high-performance state, we find that teams spend more time in the low-performance state during our data period (53% of time periods for overachievers and 56% for underachievers, on average), which makes intuitive sense in that over half of MLB teams do not make the playoffs.

5.4.2.3. State-dependent parameters. We summarize the results for $\overline{\beta_{ms}}$, the population mean for the parameters that vary over the course of the season (hidden states) and across teams (latent classes) in Table 11 (see Table WA-D6 in Web Appendix D for the population variance parameters, $\Sigma_{\beta ms}$). We begin with the results for the high-performance state. For overachievers, entertainment and giveaway promotions increase the demand for attendance (i.e., a 1% increase in entertainment and giveaway promotions increases the demand for attendance by 0.31% and 0.28%, respectively). These effects are significantly greater than the effects for event and price promotions. For underachievers, only giveaways are significant, and this effect is significantly greater than the effects for event and price promotions. The results for the low-performance state tell a somewhat different story. For overachievers, while still significant, entertainment and giveaway promotions are less effective, and event promotions become significant. The

Table 10
Transition parameters.

		Latent class 1 (overachievers)			Latent class 2 (underachievers)		
		Median	2.50%	97.50%	Median	2.50%	97.50%
Switch from state 1 to state 2 (high- to low-performance)	Constant	−2.21[a]	−3.25	−1.32	−0.83	−1.58	0.01
	Home team number of games back (vs. division leader)	0.30[a]	0.02	0.66	0.09[a]	0.02	0.16
Stay in state 2 (low-performance)	Constant	−1.46[a]	−2.14	−0.27	−2.26[a]	−3.03	−1.55
	Home team number of games back (vs. division leader)	0.23[a]	0.15	0.30	0.39[a]	0.23	0.73

[a] Indicates that the parameter is significantly different from zero based on the 95% confidence interval.

Please cite this article as: Kappe, E., et al., A random coefficients mixture hidden Markov model for marketing research, *International Journal of Research in Marketing* (2018), https://doi.org/10.1016/j.ijresmar.2018.07.002

Table 11
State-dependent parameters.

	Variable	Latent class 1 (overachievers)			Latent class 2 (underachievers)		
		Median	2.50%	97.50%	Median	2.50%	97.50%
State 1 (high-performance)	No. of entertainment promotions	0.31[a]	0.09	0.57	0.07	−0.02	0.16
	No. of event promotions	0.04	−0.06	0.15	−0.03	−0.11	0.07
	No. of giveaway promotions	0.28[a]	0.10	0.43	0.14[a]	0.09	0.21
	No. of price promotions	0.07	−0.04	0.15	0.05	−0.02	0.12
State 2 (low-performance)	No. of entertainment promotions	0.15[a]	0.09	0.21	0.14[a]	0.10	0.19
	No. of event promotions	0.05[a]	0.01	0.09	0.07[a]	0.01	0.13
	No. of giveaway promotions	0.10[a]	0.05	0.16	0.08[a]	0.02	0.14
	No. of price promotions	0.02	−0.02	0.05	0.04	−0.01	0.10

[a] Indicates that the parameter is significantly different from zero based on the 95% confidence interval.

effect for entertainment promotions is significantly greater than the effect for event and price promotions and the effect of giveaways is significantly greater than the effect for price promotions. For underachievers, entertainment and event promotions become significant and giveaways are less effective.

These results lead to three valuable managerial insights. First, while event and price promotions were the most commonly used, these categories are the least effective. Second, concerning the two latent classes, the average promotional effectiveness for overachievers (0.13) is almost twice that of underachievers (0.07). This makes sense in that overachieving teams need to be more creative/strategic in planning promotions to increase the demand for attendance. Promotions may also be especially effective for overachievers given that the average ticket price ($22.75) is about one-third lower than underachievers ($32.22); thus, promotions for overachievers provide fans with more value or bang for their buck. Finally, concerning hidden states, the average effect across promotional categories for games in the high-performance state (0.12) is somewhat higher than for games in the low-performance state (0.08) (although this pattern only holds for overachievers).

5.4.2.4. Alternative specifications. We also compare the results of the RCMHMM to the next best model in Table 8—an HMM with random coefficients for both the state-dependent and transition parameters (the most popular HMM in marketing). This model has five hidden states and is less parsimonious than the RCMHMM (while models with three, four, and five hidden states were not significantly different, the model with five hidden states had the lowest WAIC). We highlight a few major insights from this model. While the effects of the four promotional categories averaged across hidden states and teams are fairly comparable, the effects of each promotional category at the team level are quite different. To compare estimates at the team level, we extracted the average effectiveness of each of the four promotional categories for all 30 teams (across hidden states). This resulted in a vector of 120 promotional effects for each model. The correlation between the two vectors was only 0.59, supporting our argument that specifying a suboptimal unobserved heterogeneity structure leads to biased results and managerial insights.

5.5. Managerial implications

Using results from the RCMHMM, we can answer three managerial questions related to promotional schedule design (we focus on relatively minor changes to stay within the variation of our data and avoid the Lucas (1976) critique). First, what promotions are most effective? Teams could decrease the frequency of less effective promotions and increase the frequency of more effective promotions. For example, if NYM replaced 10 event promotions with 10 entertainment promotions—irrespective of the hidden state—revenues would increase by 33.6% or $2.27 million. Second, when are promotions most effective? Teams could rearrange promotions based on the recovery of hidden states to ensure that promotions are employed in the period when they are most effective. For example, if WSN (an underachiever that spent 46% of the season in the low-performance state) swapped 10 event promotions with 10 giveaways during high-performance periods, and vice versa during periods of low performance, revenues for those 20 games would increase by 12.6% or $2.15 million. Third, when should teams update their promotional schedule during the season? On average, teams have equal probabilities of being in the high- and low-performance state at around 7.5 games back from the division leader. Thus, teams could update promotional schedules as soon as they go below or above this benchmark. For example, while COL was below this benchmark until home game 56, by home game 57, they were 13 games back and never returned below this benchmark for the remainder of the season. If COL (underachievers) updated the promotional schedule after home game 56, any game with an additional entertainment promotion would have increased revenues by 24% or $200,000.

6. Conclusion

We propose a new random coefficients mixture HMM (RCMHMM) that allows for flexible patterns of unobserved heterogeneity in both the state-dependent and transition parameters in an HMM framework. The RCMHMM nests all previous HMMs found in the marketing literature and enables researchers to test and select the best unobserved heterogeneity structure for their data. Failure to do so can lead to biased parameters and inaccurate managerial insights (as we confirm in simulation study 1). We

Please cite this article as: Kappe, E., et al., A random coefficients mixture hidden Markov model for marketing research, *International Journal of Research in Marketing* (2018), https://doi.org/10.1016/j.ijresmar.2018.07.002

apply the RCMHMM to two simulation studies and averaging across the 17 different data generating processes we test, the RCMHMM outperforms its 16 nested versions. We also apply the RCMHMM to an empirical application and show that the RCMHMM outperforms its 16 nested versions as well as several additional benchmark models.

Our simulation studies answer four questions. First, under which conditions does the RCMHMM beat its nested model versions? The RCMHMM performs better than its nested versions when more complex unobserved heterogeneity structures are present in the data. The HMM with random coefficients for both the state-dependent and transition parameters, the state-of-the-art HMM in marketing, performs second-best. Comparing these two models, the RCMHMM performs better when the unobserved heterogeneity structure is multimodal. When the unobserved heterogeneity structure is unimodal, the performance of both models is similar (as the RCMHMM perfectly nests the HMM with random-coefficient state-dependent and transition parameters). Thus, when there is theoretical or empirical evidence to suggest that the unobserved heterogeneity structure is multimodal, the RCMHMM will reduce bias and improve managerial insights, but when there is evidence that the structure is unimodal, the random-coefficients HMM suffices. Second, under which conditions is the additional flexibility of the RCMHMM harmful? The RCMHMM performs somewhat worse when the true model includes either homogeneous parameters or latent classes for the state-dependent parameters or latent classes for the transition parameters. When this is the case, researchers can employ more diffuse priors for the population variance parameters to improve RCMHMM performance. Third, under which conditions does the RCMHMM perform relatively better or worse? The RCMHMM is relatively robust to the DGP and performs better when more time periods are available in the data. The specific number of time periods necessary to reliably estimate the RCMHMM depends on the empirical application and data at hand (e.g., amount of variation in the data, distribution of the latent classes and hidden states, nature of the dependent variable). Fourth, how robust is the RCMHMM when the underlying assumptions of the RCMHMM are violated? The RCMHMM is more robust than its nested versions when an omitted variable is present in the data or when the error distribution is misspecified.

In our empirical application, we investigate the effectiveness of in-game promotions in increasing the short-term demand for MLB attendance and find that promotional effectiveness varies over the course of the season and across teams. Thus, the RCMHMM performs best. We offer three valuable managerial insights: 1) the average effectiveness of entertainment and giveaway promotions is highest—even though event and price promotions were employed most often, 2) promotional effectiveness is higher for overachieving teams than underachieving teams, and 3) promotional effectiveness changes for both overachievers and underachievers when teams transition from a high-performance to a low-performance state. As a result, MLB teams could benefit from knowing which latent class they belong to and which hidden state they are in.

While the RCMHMM performs best in our simulation studies and empirical application, the optimal unobserved heterogeneity structure for other applications remains an open question. The ability of the RCMHMM to test a number of alternative unobserved heterogeneity structures makes it possible for researchers to select the most appropriate structure across a variety of applications using a single model. We believe that the RCMHMM could be valuable in panel data contexts where the effects for cross-sectional units may vary over time. Examples range from the latent relationship between a firm and consumer (Netzer et al., 2008) to modeling latent attrition (Schweidel & Knox, 2013), missing competitive firm data (Moon et al., 2007), or biases in survey responses (Yang et al., 2010). Finally, as we illustrate in this paper, researchers could use the RCMHMM to identify how marketing mix variables vary over time and across individuals (Montoya et al., 2010).

As with all research, our work has limitations that provide opportunities for future research. First, we focus on the RCMHMM and its 16 nested versions to control for unobserved heterogeneity. However, other promising approaches exist to incorporate unobserved heterogeneity, such as Dirichlet process priors (e.g., Ansari & Mela, 2003; Voleti & Ghosh, 2013) or a model in which each coefficient has its own finite mixture (Ebbes et al., 2015). Future research could combine either of these approaches with the HMM and assess its performance relative to our proposed RCMHMM. Next, while our estimation algorithm accurately recovered all parameters in a reasonable amount of time, the size of the datasets in our simulation studies and empirical application were relatively small. Future research could examine more efficient estimation procedures when applying the RCMHMM to larger datasets (with thousands of cross-sectional and/or time series observations) or when making real-time decisions (e.g., Foti, Xu, Laird, & Fox, 2014). Ultimately, we hope the RCMHMM helps researchers and practitioners make better decisions across a variety of contexts.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ijresmar.2018.07.002.

References

Allenby, G. M., Arora, N., & Ginter, J. L. (1998). On the heterogeneity of demand. *Journal of Marketing Research*, *35*(3), 384–389.

Allenby, G. M., & Rossi, P. E. (1998). Marketing models of consumer heterogeneity. *Journal of Econometrics*, *89*(1–2), 57–78.

Andrews, R. L., Ansari, A., & Currim, I. S. (2002). Hierarchical Bayes versus finite mixture conjoint analysis models: A comparison of fit, prediction, and partworth recovery. *Journal of Marketing Research*, *39*(1), 87–98.

Ansari, A., & Mela, C. F. (2003). E-customization. *Journal of Marketing Research*, *40*(2), 131–145.

Ansari, A., Montoya, R., & Netzer, O. (2012). Dynamic learning in behavioral games: A hidden Markov mixture of experts approach. *Quantitative Marketing and Economics*, *10*(4), 475–503.

Ascarza, E., & Hardie, B. G. S. (2013). A joint model of usage and churn in contractual settings. *Marketing Science*, *32*(4), 570–590.

Badenhausen, K., Ozanian, M., & Settimi, C. (2013). *MLB team values: The business of baseball.* Forbes (March 27) http://www.forbes.com/mlb-valuations/list/.

Please cite this article as: Kappe, E., et al., A random coefficients mixture hidden Markov model for marketing research, *International Journal of Research in Marketing* (2018), https://doi.org/10.1016/j.ijresmar.2018.07.002

Böckenholt, U., & Dillon, W. R. (1997). Some new methods for an old problem: Modeling preference changes and competitive market structures in pretest market data. *Journal of Marketing Research*, *34*(1), 130–142.
Bodapati, A., & Gupta, S. (2004). A direct approach to predicting discretized response in target marketing. *Journal of Marketing Research*, *41*(1), 73–85.
Borland, J., & Lye, J. (1992). Attendance at Australian rules football: A panel study. *Applied Economics*, *24*(9), 1053–1058.
Boyd, T. C., & Krehbiel, T. C. (1999). The effect of promotion timing on Major League Baseball attendance. *Sport Marketing Quarterly*, *8*(4), 23–34.
Boyd, T. C., & Krehbiel, T. C. (2003). Promotion timing in Major League Baseball and the stacking effects of factors that increase game attractiveness. *Sport Marketing Quarterly*, *12*(3), 173–183.
Boyd, T. C., & Krehbiel, T. C. (2006). An analysis of the effects of specific promotion types on attendance at Major League Baseball games. *American Journal of Business*, *21*(2), 21–32.
Brangule-Vlagsma, K., Pieters, R. G. M., & Wedel, M. (2002). The dynamics of value segments: Modeling framework and empirical illustration. *International Journal of Research in Marketing*, *19*(3), 267–285.
Burnham, K. P., & Anderson, D. R. (2004). Multimodel inference: Understanding AIC and BIC in model selection. *Sociological Methods & Research*, *33*(2), 261–304.
Carmone, F. J., Jr., Ali, K., & Maxwell, S. (1999). HINoV: A new model to improve market segment definition by identifying noisy variables. *Journal of Marketing Research*, *36*(4), 501–509.
Chandukala, S. R., Long-Tolbert, S., & Allenby, G. M. (2011). A threshold model for respondent heterogeneity. *Marketing Letters*, *22*(2), 133–146.
Chib, S. (1996). Calculating posterior distributions and modal estimates in Markov mixture models. *Journal of Econometrics*, *75*(1), 79–97.
van der Lans, R., Pieters, R., & Wedel, M. (2008). Competitive brand salience. *Marketing Science*, *27*(5), 922–931.
DeSarbo, W. S., Hwang, H., Stadler Blank, A., & Kappe, E. (2015). Constrained stochastic extended redundancy analysis. *Psychometrika*, *80*(2), 516–534.
Du, R. Y., & Kamakura, W. A. (2006). Household life cycles and lifestyles in the United States. *Journal of Marketing Research*, *43*(1), 121–132.
Dubé, J. P., Hitsch, G. J., & Rossi, P. E. (2010). State dependence and alternative explanations for consumer inertia. *The RAND Journal of Economics*, *41*(3), 417–445.
Ebbes, P., Grewal, R., & DeSarbo, W. S. (2010). Modeling strategic group dynamics: A hidden Markov approach. *Quantitative Marketing and Economics*, *8*(2), 241–274.
Ebbes, P., Liechty, J., & Grewal, R. (2015). Attribute-level heterogeneity. *Management Science*, *61*(4), 885–897.
Erdem, T., Imai, S., & Keane, M. P. (2003). Brand and quantity choice dynamics under price uncertainty. *Quantitative Marketing and Economics*, *1*(1), 5–64.
Foti, N. J., Xu, J., Laird, D., & Fox, E. B. (2014). Stochastic variational inference for hidden Markov models. *Advances in Neural Information Processing Systems*, *27*, 1–9.
Frühwirth-Schnatter, S. (2001). Markov chain Monte Carlo estimation of classical and dynamic switching and mixture models. *Journal of the American Statistical Association*, *96*(453), 194–209.
Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., & Rubin, D. B. (2013). *Bayesian data analysis* (3rd ed.). Boca Raton, FL: CRC Press.
Gelman, A., & Rubin, D. B. (1992). Inference from iterative simulation using multiple sequences. *Statistical Science*, *7*(4), 457–472.
Hamilton, J. D. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica*, *57*(2), 357–384.
Heckman, J. J. (1991). Identifying the hand of past: Distinguishing state dependence from heterogeneity. *American Economic Review*, *81*(2), 75–79.
Holtrop, N., Wieringa, J. E., Gijsenberg, M. J., & Verhoef, P. C. (2017). No future without the past? Predicting churn in the face of customer privacy. *International Journal of Research in Marketing*, *34*(1), 154–172.
Hsiao, C. (2003). *Analysis of panel data* (2nd ed.). New York, NY: Cambridge University Press.
Huang, Y., Singh, P. V., & Ghose, A. (2015). A structural model of employee behavioral dynamics in enterprise social media. *Management Science*, *61*(12), 2825–2844.
Hui, S. K. (2017). Understanding repeat playing behavior in casual games using a Bayesian data augmentation approach. *Quantitative Marketing and Economics*, *15*(1), 29–55.
Hyslop, D. R. (1999). State dependence, serial correlation and heterogeneity in intertemporal labor force participation of married women. *Econometrica*, *67*(6), 1225–1294.
Kappe, E., Stadler Blank, A., & DeSarbo, W. S. (2014). A general multiple distributed lag framework for estimating the dynamic effects of promotions. *Management Science*, *60*(6), 1489–1510.
Kumar, V., Sriram, S., Luo, A., & Chintagunta, P. K. (2011). Assessing the effect of marketing investments in a business marketing context. *Marketing Science*, *30*(5), 924–940.
Lee, J. K. H., Sudhir, K., & Steckel, J. H. (2002). A multiple ideal point model: Capturing multiple preference effects from within an ideal point framework. *Journal of Marketing Research*, *39*(1), 73–86.
Lemke, R. J., Leonard, M., & Tlhokwane, K. (2010). Estimating attendance at Major League Baseball games for the 2007 season. *Journal of Sports Economy*, *11*(3), 316–348.
Lemmens, A., Croux, C., & Stremersch, S. (2012). Dynamics in the international market segmentation of new product growth. *International Journal of Research in Marketing*, *29*(1), 81–92.
Lenk, P. J., & DeSarbo, W. S. (2000). Bayesian inference for finite mixtures of generalized linear models with random effects. *Psychometrika*, *65*(1), 93–119.
Li, S., Sun, B., & Montgomery, A. L. (2011). Cross-selling the right product to the right customer at the right time. *Journal of Marketing Research*, *48*(4), 683–700.
Lucas, R. E., Jr. (1976). Econometric policy evaluation: A critique. In K. Brunner, & A. H. Meltzer (Eds.), *Carnegie-Rochester conf. ser. public policy. 1*. (pp. 19–46).
Luo, A., & Kumar, V. (2013). Recovering hidden buyer-seller relationship states to measure the return on marketing investment in business-to-business markets. *Journal of Marketing Research*, *50*(1), 143–160.
Ma, L., Sun, B., & Kekre, S. (2015). The squeaky wheel gets the grease – An empirical analysis of customer voice and firm intervention on Twitter. *Marketing Science*, *34*(5), 627–645.
Marcum, J. P., & Greenstein, T. N. (1985). Factors affecting attendance of Major League Baseball: II. A within-season analysis. *Sociology of Sport Journal*, *2*(4), 314–322.
McDonald, M., & Rascher, D. (2000). Does bat day make cents? The effect of promotions on the demand for major league baseball. *Journal of Sport Management*, *14*(1), 8–27.
Mehta, N., Chen, X., & Narasimhan, O. (2010). Examining demand elasticities in Hanemann's framework: A theoretical and empirical analysis. *Marketing Science*, *29*(3), 422–437.
Mehta, N., Ni, J., Srinivasan, K., & Sun, B. (2017). A dynamic model of health insurance choices and healthcare consumption decisions. *Marketing Science*, *36*(3), 338–360.
Michalek, J. J., Ebbes, P., Adigüzel, F., Feinberg, F. M., & Papalambros, P. Y. (2011). Enhancing marketing with engineering: Optimal product line design for heterogeneous markets. *International Journal of Research in Marketing*, *28*(1), 1–12.
MLB (2013). MLB finishes 2013 with sixth best attendance total ever. MLB.com (October 1) http://m.mlb.com/news/article/62282120/mlb-finishes-2013-with-sixth-best-attendance-total-ever/.
Montgomery, A. L., Li, S., Srinivasan, K., & Liechty, J. C. (2004). Modeling online browsing and path analysis using clickstream data. *Marketing Science*, *23*(4), 579–595.
Montoya, R., Netzer, O., & Jedidi, K. (2010). Dynamic allocation of pharmaceutical detailing and sampling for long-term profitability. *Marketing Science*, *29*(5), 909–924.
Moon, S., Kamakura, W. A., & Ledolter, J. (2007). Estimating promotion response when competitive promotions are unobservable. *Journal of Marketing Research*, *44*(3), 503–515.
Narayanan, S., Manchanda, P., & Chintagunta, P. K. (2005). Temporal differences in the role of marketing communication in new product categories. *Journal of Marketing Research*, *42*(3), 278–290.
Netzer, O., Ebbes, P., & Bijmolt, T. (2016). In L. PSH, J. E. Wieringa, B. THA, & K. H. Pauwels (Eds.), *Hidden Markov Models in marketing, forthcoming as chapter XX in advanced methods in modeling markets*. Springer International Series in Quantitative Marketing.
Netzer, O., Lattin, J. M., & Srinivasan, V. (2008). A hidden Markov model of customer relationship dynamics. *Marketing Science*, *27*(2), 185–204.
Otter, T., Tüchler, R., & Frühwirth-Schnatter, S. (2004). Capturing consumer heterogeneity in metric conjoint analysis using Bayesian mixture models. *International Journal of Research in Marketing*, *21*(3), 285–297.
Park, S., & Gupta, S. (2011). A regime-switching model of cyclical category buying. *Marketing Science*, *30*(3), 469–480.
Parsons, L. J. (1975). The product life cycle and time-varying advertising elasticities. *Journal of Marketing Research*, *12*(4), 476–480.

Please cite this article as: Kappe, E., et al., A random coefficients mixture hidden Markov model for marketing research, *International Journal of Research in Marketing* (2018), https://doi.org/10.1016/j.ijresmar.2018.07.002

Peel, D., & Thomas, D. (1988). Outcome uncertainty and the demand for football: An analysis of match attendances in the English football league. *Scott Journal of Political Economy*, *35*(3), 242–249.
Platzer, M., & Reutterer, T. (2016). Ticking away the moments: Timing regularity helps to better predict customer activity. *Marketing Science*, *35*(5), 779–799.
Poulsen, C. S. (1990). Mixed Markov and latent Markov modelling applied to brand choice behaviour. *International Journal of Research in Marketing*, *7*(1), 5–19.
Ramaswamy, V. (1997). Evolutionary preference segmentation with panel survey data: An application to new products. *International Journal of Research in Marketing*, *14*(1), 57–80.
Rosenthal, J. S. (2011). Optimal proposal distributions and adaptive MCMC. In S. Brooks, A. Gelman, G. L. Jones, & X. L. Meng (Eds.), *Handbook of Markov chain Monte Carlo* (pp. 93–111). Boca Raton, FL: Chapman & Hall/CRC.
Rossi, P. E. (2014). Even the rich can make themselves poor: A critical examination of IV methods in marketing applications. *Marketing Science*, *33*(5), 655–672.
Schwartz, E. M., Bradlow, E. T., & Fader, P. S. (2014). Model selection using database characteristics: Developing a classification tree for longitudinal incidence data. *Marketing Science*, *33*(2), 188–205.
Schweidel, D. A., Bradlow, E. T., & Fader, P. S. (2011). Portfolio dynamics for customers of a multiservice provider. *Management Science*, *57*(3), 471–486.
Schweidel, D. A., & Knox, G. (2013). Incorporating direct marketing activity into latent attrition models. *Marketing Science*, *32*(3), 471–487.
Schweidel, D. A., Park, Y. H., & Jamal, Z. (2014). A multiactivity latent attrition model for customer base analysis. *Marketing Science*, *33*(2), 273–286.
Shachat, J., & Wei, L. (2012). Procuring commodities: First-price sealed-bid or English auctions? *Marketing Science*, *31*(2), 317–333.
Shi, S. W., Wedel, M., & Pieters, F. G. M. (2013). Information acquisition during online decision making: A model-based exploration using eye-tracking data. *Management Science*, *59*(5), 1009–1026.
Shi, S. W., & Zhang, J. (2014). Usage experience with decision aids and evolution of online purchase behavior. *Marketing Science*, *33*(6), 871–882.
Siegfried, J. J., & Eisenberg, J. D. (1980). The demand for minor league baseball. *Atlantic Economic Journal*, *8*(2), 59–69.
Stüttgen, P., Boatwright, P., & Monroe, R. T. (2012). A satisficing choice model. *Marketing Science*, *31*(6), 878–899.
Team Marketing Report (2013). 2013 MLB fan cost index. Accessed May 10, 2015 https://www.teammarketing.com/btSubscriptions/fancostcontroller/index.
Thietart, R. A., & Vivas, R. (1984). An empirical investigation of success strategies for businesses along the product life cycle. *Management Science*, *30*(12), 1405–1423.
Voleti, S., & Ghosh, P. (2013). A robust approach to measure latent, time-varying equity in hierarchical branding structures. *Quantitative Marketing and Economics*, *11*(3), 289–319.
Vriens, M., Wedel, M., & Wilms, T. (1996). Metric conjoint segmentation methods: A Monte Carlo comparison. *Journal of Marketing Research*, *33*(1), 73–85.
Watanabe, S. (2010). Asymptotic equivalence of Bayes cross validation and widely applicable information criterion in singular learning theory. *Journal of Machine Learning Research*, *11*(December), 3571–3594.
Wedel, M., Kamakura, W., Arora, N., Bemmaor, A., Chiang, J., Elrod, T., ... Poulsen, C. S. (1999). Discrete and continuous representations of unobserved heterogeneity in choice modeling. *Marketing Letters*, *10*(3), 219–232.
Xu, J. J., Fader, P., & Veeraraghavan, S. (2016). *The revenue impact of dynamic pricing policies in Major League Baseball ticket sales.* Working paper Philadelphia: University of Pennsylvania.
Yang, S., Zhao, Y., & Dhar, R. (2010). Modeling the underreporting bias in panel survey data. *Marketing Science*, *29*(3), 525–539.
Zhang, J. Z., Netzer, O., & Ansari, A. (2014). Dynamic targeted pricing in B2B relationships. *Marketing Science*, *33*(3), 317–337.
Zhang, J. Z., Watson, G. F., IV, Palmatier, R. W., & Dant, R. P. (2016). Dynamic relationship marketing. *Journal of Marketing*, *80*(5), 53–75.
Zhang, X., & Feng, J. (2011). Cyclical bid adjustments in search-engine advertising. *Management Science*, *57*(9), 1703–1719.
Zhang, X., Kumar, V., & Cosguner, K. (2017). Dynamically managing a profitable email marketing program. *Journal of Marketing Research*, *54*(6), 851–866.
Zhu, J. D. (2014). *Effect of resale on optimal ticket pricing: Evidence from Major League Baseball tickets.* Working paper Texas A&M University, College Station.
Zucchini, W., MacDonald, I. L., & Langrock, R. (2016). *Hidden Markov models for time series: An introduction using R* (2nd ed.). Boca Raton, FL: CRC Press.

Please cite this article as: Kappe, E., et al., A random coefficients mixture hidden Markov model for marketing research, *International Journal of Research in Marketing* (2018), https://doi.org/10.1016/j.ijresmar.2018.07.002