WEB APPENDIX WA

MCMC ESTIMATION OF PURCHASE TIMING AND QUANTITY

Interpurchase Time

Prior Distributions: The prior distributions for the model parameters are as follows:

Parameter	Prior Distribution
λ_i	Inverse generalized gamma (ν, θ, γ) ν = Shape parameter θ = Scale parameter
ν	Uniform (0,100)
θ	Inverse generalized gamma (a_0, b_0, γ) $a_0 = b_0 = 10$
$\delta_1, \ldots \delta_4$	gamma (5,5)
α	Uniform (0,100)

MCMC estimation

MCMC inference of the model defined by (2) through (4) proceeds by recursively generating draws from each of the densities in the following algorithm.

Generate λ_i (one at a time for each customer):

$$\pi\left(\lambda_i \middle| \{t_{ij}, j = 1, \ldots n_i\}, n_i, \alpha, \nu, \gamma\right) = IGG\left(n_i\alpha + \nu, \left[\sum t_{ij}^{\gamma} \Big/ \delta_1^{t^*_{ij-1}} \ldots \delta_4^{Q^*_{ij-2}} + \theta^{-\gamma}\right]^{-1/\gamma}, \gamma\right)$$

where t_{ij} is the j^{th} interpurchase time for customer i and n_i is the number of observations for customer i.

Generate δ_1 : $\pi\left(\delta_1 \mid \lambda_i, \alpha, \delta_2, \delta_3, \delta_4, t^*_{ij-1}, \ldots Q^*_{ij-2}\right) \propto \dfrac{e^{-\left(t_{ij} * \delta_1^{t^*_{ij-1}} \Big/ \lambda_i \delta_2^{t^*_{ij-2}} \delta_3^{Q^*_{ij-1}} \delta_4^{Q^*_{ij-2}}\right)}}{\left(\delta_1^{t^*_{ij-1}}\right)^{\alpha\gamma}}$

The posterior distributions for δ_2, δ_3, and δ_4 are similar to δ_1 and we generate each parameter sequentially.

Generate α: $\pi\left(\alpha \middle| \{\lambda_i\}, \{t_{ij}\}, \gamma\right) \propto \prod_{i=1}^{N} \prod_{j=1}^{n_i} \dfrac{\gamma}{\Gamma(\alpha)\lambda_i^{\alpha\gamma}} t_{ij}^{\alpha\gamma - 1} e^{-(t_{ij}/\lambda_i)^{\gamma}}$

Generate ν: $\pi\left(\upsilon \middle| \{\lambda_i\}, \theta, \gamma\right) \propto \prod_{i=1}^{N} \dfrac{\gamma}{\Gamma(\nu)\theta^{\upsilon\gamma}} \lambda_i^{-\upsilon\gamma - 1} e^{-(1/\theta\lambda_i)^{\gamma}}$

Generate θ: $\pi\left(\theta \middle| \{\lambda_i\}, \upsilon, \gamma\right) = IGG\left(N\upsilon + a_0, \left[\sum \lambda_i^{-\gamma} + b_0^{-\gamma}\right]^{-1/\gamma}, \gamma\right)$

Similar to Allenby et al.(1999), values of γ were obtained via a grid search procedure that minimized the marginal density of the data. The reported standard deviations are estimated conditional on γ and thus are smaller than what would be obtained in an unconditional analysis.

Purchase Quantity

Prior Distributions: The prior distributions for the model parameters are as follows:

Parameter	Prior Distribution
$\{\delta^*\} = \{\delta_1, \delta_2, \delta_3, \delta_4, \delta_5, \delta_6\}$ [1]	$Normal(u_0, v_0)$ where, $u_0 = 0$, $v_0 = 100*I$
$\delta_{i,o}$	$N(0, \tau^2)$ where, τ^2 = Inverse gamma ($r_{\tau 0}, s_{\tau 0}$) $r_{\tau 0} = 5$, $s_{\tau 0} = 5$
σ^2	Inverse gamma (r_0, s_0) where, $r_0 = 5, s_0 = 5$

MCMC estimation

MCMC inference of the model defined by (5) proceeds by recursively generating draws from each of the densities in the following algorithm.

Draw δ^* from:

$$P\left(\delta^* \mid \delta_{i,0}, y_{ij}, x_{ij}, \sigma^2\right) = Normal\left(\left(\sum_{i=1}^{N}\sum_{j=1}^{n_i}\left(\frac{x_{ij}^T x_{ij}}{\sigma^2}\right)\right)^{-1} * \left(\sum_{i=1}^{N}\sum_{j=1}^{n_i}\frac{x_{ij}^T\left(y_{ij}-\delta_{i,0}\right)}{\sigma^2}\right), \left(\sum_{i=1}^{N}\sum_{j=1}^{n_i}\frac{\left(x_{ij}^T x_{ij}\right)}{\tau^2}\right)^{-1}\right)$$

where, $x_{ij} = \{Q^*_{i,j-2}, Q^*_{i,j-3}, t^*_{i,j-1}, t^*_{i,j-2}, size_i, ind_i\}$, $y_{ij} = (\Delta Q_{i,j})$

For each i, draw $\delta_{i,0}$ from: $P\left(\delta_{i,o} \middle| \delta^*, \sigma^2, \tau^2, y, x\right) = N\left(\left(\frac{n_i}{\sigma^2} + \frac{1}{\tau^2}\right)^{-1} * \sum_{j=1}^{n_i}\left(\frac{Y_{ij} - X_{ij}\delta^*}{\sigma^2}\right), \left(\frac{n_i}{\sigma^2} + \frac{1}{\tau^2}\right)^{-1}\right)$

Draw $\Sigma_{\delta_{i,o}}$ from: $P\left(\tau^2 \middle| \delta_{i,o}, r_{\tau 0}, s_{\tau 0}\right) =$ Inverse gamma $(r_{\tau n}, s_{\tau n})$

$r_{\tau n} = r_{\tau 0} + n_i$

$S_{\tau n} = S_{\tau 0} + \sum_i \delta_{i0}\delta_{i0}^T$

Draw σ^2 from: $P\left(\sigma^2 \middle| y, \delta^*, \delta_{i,o}, x\right) =$ Inverse gamma(r_n, s_n),

$r_n = r_o + N$

$$s_n = s_o + \sum_{i=1}^{N}\sum_{j=i}^{n}\left[\left(y_{ij} - x_{ij}\delta^* + \delta_{i,0}\right)^T\left(y_{ij} - x_{ij}\delta^* + \delta_{i,0}\right)\right]$$

[1] δ_6 represents all the C industry category coefficients, $Ind_i = \{ind_{i1} \ldots ind_{ic}\}$.

Mixture Model

The mixture model defined by (6) - (11) can be estimated through MCMC methods using the standard trick of introducing a latent random variable (s_{ijk}), which is used to index which of the *k* components each of the observations belong. On each iteration of the procedure, only those observations currently assigned to a component are used to arrive at $\{\lambda_{ijk}\}, \alpha_k, \nu_k, \theta_k$, $\delta_k^*, \delta_{i,0,k}$ and σ_k^2. The assignment of observations to components is straightforward if $\{\phi_k\}$, the mass points of the components, are known.

$\Pr(S_{ijk} = k) \propto \phi_k f(t_{ij} \mid \alpha_k, \lambda_{ijk}, \gamma_k)\, p(y_{ij} \mid x_{ij}, \delta^*, \delta_{i0}, \sigma_k^2)$

Where ϕ_k are estimated from (6) if k=1 and (7) if k=2

We impose the ordering $\theta_1 > \theta_2 > \ldots > \theta_k$ to achieve model identification.

Draws of β_i are obtained from a Metropolis-Hastings algorithm with a random walk chain. That is, let $\beta_i^{(p)}$ denote the previous draw for β_i. The next draw $\beta_i^{(n)}$ is then given by

$$\beta_i^{(n)} = \beta_i^{(p)} + \Delta\beta$$

where, $\Delta\beta$ is a draw from the normal density (0, .05). The choice of parameters for this density yields an acceptance rate of about 50%. The acceptance probability is given by

$$\Pr(accept) = \min\left[\frac{\exp\left[-\frac{1}{2}\left(\beta_i^{(n)} - \bar{\beta}\right)' V^{-1}\left(\beta_i^{(n)} - \bar{\beta}\right)\right]\prod_j \phi_{ijk}(\beta_i^n)}{\exp\left[-\frac{1}{2}\left(\beta_i^{(p)} - \bar{\beta}\right)' V^{-1}\left(\beta_i^{(p)} - \bar{\beta}\right)\right]\prod_j \phi_{ijk}(\beta_i^p)}, 1\right]$$

where $\phi_{ijk}(\beta_i)$ denotes the value obtained by evaluating (6)-(7) at β_i. The remaining conditional distributions are given by

$$\pi\left(\bar{\beta} \mid \{\beta_i\}, V\right) = normal\left(\sum_{i=1}^{N} \beta_i / N, V / N\right)$$

and

$$\pi\left(V \mid \{\beta_i\}, \bar{\beta}\right) = invertedWishart\left(\sum_{i=1}^{N} \left(\beta_i - \bar{\beta}\right)\left(\beta_i - \bar{\beta}\right)' + G, N + g\right)$$

where, G and g are prior parameters that are set to G=15I and g=15.

Prediction of Interpurchase time and Purchase Quantity

For every iteration in the posterior sample, the interpurchase time and purchase quantity for a customer i (for the sake of simplicity we remove the subscript i in the following steps), is measured as follows:

1. Set up X_d as the frequency of rich and standard modes between t_{n-1} and t_n (where n is the total number of purchases for customer i). Set up the covariates – cross buying, upgrading as the cumulative number until t_n. Set up the current effect covariates as the activities between t_{n-1} and t_n.
2. Compute $\hat{\phi}_{1k}$ using eq. (6) – (7).
3. Draw $\hat{t}^*_{n+1,1}$ from generalized gamma $(\alpha_1, \lambda_{1i}, \gamma_1)$.
4. Draw $\hat{t}^*_{n+1,2}$ from generalized gamma $(\alpha_2, \lambda_{2i}, \gamma_2)$.
5. Draw $\Delta\hat{Q}_{n+1,1}$ from Normal $(\mu_{1,n+1}, \sigma_1^2)$ [2].
6. Draw $\Delta\hat{Q}_{n+1,2}$ from Normal $(\mu_{2,n+1}, \sigma_2^2)$.
7. Generate U; a uniform random number between 0 and 1.
8. If $U \le \hat{\phi}_1$, then $\hat{t}^*_{n+1} = \hat{t}^*_{n+1,1}$ and $\Delta\hat{Q}_{n+1} = \Delta\hat{Q}_{n+1,1}$
   else $\hat{t}^*_{n+1} = \hat{t}^*_{n+1,2}$ and $\Delta\hat{Q}_{n+1} = \Delta\hat{Q}_{n+1,2}$.
9. Calculate $\hat{t}_{n+1} = \hat{t}^*_{n+1} + (48 - \sum_{j=1}^{n} t_j)$. We use 48 because the calibration data has 48 months for Cohort 1. It is changed to 36 months for Cohort 2.
10. Calculate $\hat{Q}_{n+1} = \Delta\hat{Q}_{n+1} + Q_n$.
11. Set sumt=$\hat{t}_{n+1}$. Set j=1;
12. Repeat the steps 13-17 until sumt>36 (we use 36 because CLV is measured over 3 years)
13. Increment j by one.
14. Repeat step 1 and update the lagged interpurchase time in X_{COV} as the predicted interpurchase time in the previous step, i.e., $\hat{t}_{n+j-1}$. If j is equal to 2, then set the covariates in the Quantity regression as {Q_n,Q_{n-1}, $\hat{t}_{n+1}$,t_n,size,Ind}. If j is equal to 3, then set the covariates in the Quantity regression as {$\hat{Q}_{n+1}$,Q_n, $\hat{t}_{n+j-1}$,$\hat{t}_{n+j-2}$,size,Ind}. If j is greater than 3, then set the covariates in the Quantity regression as {$\hat{Q}_{n+j-2}$,$\hat{Q}_{n+j-3}$, $\hat{t}_{n+j-1}$,$\hat{t}_{n+j-2}$,size,Ind}.
15. Repeat steps 2-8 to obtain $\hat{t}^*_{n+j}$, and $\hat{Q}_{n+j}$.
16. Calculate $\hat{t}_{n+j} = sumt + \hat{t}^*_{n+j}$.
17. Repeat step 10 and Calculate sumt=sumt+$\hat{t}_{n+j}$.

---

[2] $\mu_{k,n+1} = x'_{n+1}\delta^*_k + \delta_{0,k}$, where k =1 or 2.

WEB APPENDIX WB

MODEL SPECIFICATION AND SELECTION EXERSICE FOR THE BENCHMARK MODEL

Model Specification

The joint model of purchase timing and quantity proposed by Boatwright et al. (2003) consists considering separately the distribution of quantity conditional on purchase timing and the marginal distribution of purchase timing. We use a generalized gamma distribution for the marginal distribution of purchase timing in our context because it provides the best in-sample fit in our initial comparison of the various distributions of interpurchase times. The generalized gamma distribution, as indicated in equation (2) is given by:

(B1) $$\mathrm{t_{ij}} \sim \mathrm{GG}(\alpha, \lambda_{ij}, \gamma) = \frac{\gamma}{\Gamma(\alpha)\lambda_{ij}^{\alpha\gamma}} t_{ij}^{\alpha\gamma-1} e^{-(t_{ij}/\lambda_{ij})^{\gamma}}$$

where, α, and γ establish the shape of the distribution and λ_{ij} is the individual specific purchase rate parameter.

The individual specific purchase rate parameter, λ_{ij}, is related to covariates via a multiplicative model,

(B2) $$\lambda_{ij} = \lambda_i \beta_1^{x_{1ij}} .....\beta_k^{x_{kij}}$$

where, x_{1ij}…x_{kij} are the time varying marketing decision variables and covariates. β_k=1 implies that x_k has no influence on the interpurchase time. For β_k>1, x_k has a positive influence (increase) on the interpurchase time, and for β_k<1, x_k has a negative influence (decrease) on interpurchase time. Heterogeneity in the expected customer interpurchase times is accommodated by allowing:

(B3) $$\lambda_i \sim gamma(\nu, \theta)$$

We specify uniform priors on α, and ν, an inverse gamma prior on θ, and gamma priors on β. As in the proposed model, γ is estimated through a grid search that minimized the marginal density of the data.

In the benchmark model, the quantity purchased is modeled conditional on interpurchase time and is given by:

(B4) $$\Delta Q_{i,j} \sim normal(\mu_{ij}, \sigma^2)$$

where, $\mu_{i,j} = \delta_{i,0} + \delta_{i,1} * Q_{i,j-2} + \delta_{i,2} * Q_{i,j-3} + \delta_{i,3} * \hat{t}_{i,j} + \delta_{i,4} * t_{i,j-1} + \delta_{i,5} * size_i + \sum_{c=1}^{C} \delta_{i,6,c} * Ind_{ic} + e_{ij}$.

$\hat{t}_{i,j}$ is the predicted interpurchase time for customer i in purchase occasion j obtained from (B1), and σ^2 is the variance parameter. The individual specific coefficients $\delta_i = (\delta_{i,0} \ldots \delta_{i,6})$ are specified using a random effects specification,

(B5) $\delta_i \sim normal(\bar{\delta}, \Sigma_\delta)$.

We assume a conjugate normal prior for $\bar{\delta}$, inverted wishart prior for Σ_δ, and a inverse gamma prior for σ^2. The benchmark model was estimated using Winbugs, we used first 10,000 iterations as burn-in and the last 10,000 iterations as the posterior sample. The signs of the parameter estimates of β, and δ were in the expected direction as listed in Table 2[3]. The aggregate log CPO for the Benchmark model is equal to -6,542 which is lesser than the log CPO the Proposed Model (-5,945). The MAD for purchase time from the Benchmark model is equal to 3.0 months, and the MAD for purchase quantity is equal to 7.3, as compared to the MAD of 2.2 months and 6.1 for purchase timing and quantity from the proposed model (Model II).

[3] The parameter estimates from the benchmark model are available from the author(s).