Some Customers Would Rather Leave Without Saying Goodbye

Eva Ascarza, Oded Netzer, and Bruce G. S. Hardie

Web Appendix A: Model Estimation

In this appendix we describe the hierarchical Bayesian framework used to estimate the model parameters.

So as to ensure identification of the states (i.e., prevent label switching), we restrict the opening probabilities to be decreasing in the relationship states.[A1] We therefore reparameterize $\boldsymbol{\zeta^o}$ in the following manner,

$$\zeta^o_k = \begin{cases} \zeta'^o_1 & \text{if } k = 1 \\ \zeta'^o_{k-1} - \exp\left(\zeta'^o_k\right) & \text{for } k = 2, \ldots, K \,, \end{cases}$$

and estimate $\boldsymbol{\zeta'^o}$ instead. Note that this restriction does not impose any a priori assumption on the model.

Let $\boldsymbol{\Omega}$ denote the following (population-level) parameters: $\boldsymbol{\zeta'^o}$ (opening), $\boldsymbol{\zeta^c}$ (clicking given open), $\boldsymbol{\zeta^u}$ (unsubscribing given open), $\boldsymbol{\phi}$ (transition probabilities), $\boldsymbol{\rho}$ (initial probabilities), $\boldsymbol{\beta}$ (effect of covariates on the state-dependent probabilities), and $\boldsymbol{\delta}$ (effect of covariates on the transition probabilities). The vector $\boldsymbol{\xi}_i = \{\boldsymbol{\eta}_i, \boldsymbol{\psi}_i\}$ contains all the individual-level unobserved parameters for customer i. Let $\boldsymbol{\xi} = \{\boldsymbol{\xi}_i\}_{i=1,\ldots,I}$ denote all the individual-level parameters (where I is the number of customers in the calibration sample), and $\boldsymbol{\Sigma}_\xi$ denote the variance-covariance matrix for cross-sectional heterogeneity. The full joint posterior distribution can be written as

$$f(\boldsymbol{\Omega}, \boldsymbol{\xi}|\text{data}) \propto \left\{ \prod_{i=1}^{I} \mathscr{L}_i(\boldsymbol{\Omega}, \boldsymbol{\xi}_i \,|\, \text{data}) f(\boldsymbol{\xi}_i \,|\, \boldsymbol{\Sigma}_\xi) \right\} f(\boldsymbol{\Sigma}_\xi) f(\boldsymbol{\Omega}) \,,$$

where $\mathscr{L}_i(\boldsymbol{\Omega}, \boldsymbol{\xi}_i \,|\, \text{data})$ is defined in (11). The term $f(\boldsymbol{\xi}_i \,|\, \boldsymbol{\Sigma}_\xi)$ denotes the prior (or mixing) distribution for $\boldsymbol{\xi}_i$, which is assumed to follow a multivariate normal distribution with mean $\mathbf{0}$ and variance-covariance matrix $\boldsymbol{\Sigma}_\xi$. The terms $f(\boldsymbol{\Omega})$ and $f(\boldsymbol{\Sigma}_\xi)$ denote the (hyper)priors for the population parameters. Uninformative (vague) priors are used for all parameters. We assume $\boldsymbol{\Sigma}_\xi$ has

[A1] By ensuring ordering in one of the state-dependent behaviors (e.g., probability of opening given state membership) we prevent label switching without imposing any restriction on the relationships among behaviors.

an inverse-Wishart prior with degrees of freedom df $= K(K-1)+7$ and scale matrix $\boldsymbol{R}$, with $\text{diag}(\boldsymbol{R}) = \mathbb{1}_{K(K-1)+2}/(3 \times \text{df})$, where $\mathbb{1}_A$ denotes a $1 \times A$ vector of ones. We assume that $\boldsymbol{\Omega}$ has a multivariate normal prior with mean $\boldsymbol{\mu}_\Omega$ and variance-covariance matrix $\boldsymbol{\Sigma}_\Omega$. The values of $\boldsymbol{\mu}_\Omega$ and $\boldsymbol{\Sigma}_\Omega$ were chosen to ensure uninformative priors in the transformed space. The mean is specified as $\boldsymbol{\mu}_\Omega = \left[0 \times \mathbb{1}_K\, , \, 0.1 \times \mathbb{1}_K\, , \, -3 \times \mathbb{1}_K\, , \, 0.5 \times \mathbb{1}_{K(K-1)}\, , \, -10 \times \mathbb{1}_{K-1}\, , \, -1 \times \mathbb{1}_{C^\delta}\, , \, 0.1 \times \mathbb{1}_{C^o+C^c+C^u}\right]$, where C^δ is the number of covariates included in the transition probability equation multiplied by $K(K-1)$, and C^u, C^o, and C^c are the number of covariates incorporated in the equations for each of the three observed behaviors (excluding the intercept), multiplied by K. The variance-covariance matrix is specified as $\text{diag}(\boldsymbol{\Sigma}_\Omega) = \left[1.2 \times \mathbb{1}_K\, , \, 0.3 \times \mathbb{1}_K\, , \, 0.5 \times \mathbb{1}_K\, , \, 0.01 \times \mathbb{1}_{K(K-1)}\, , \, 0.1 \times \mathbb{1}_{K-1}\, , \, 0.1 \times \mathbb{1}_{C^\delta}\, , \, 0.3 \times \mathbb{1}_{C^o+C^c+C^u}\right]$.

Since there are no closed-form expressions for the posterior distributions of $\boldsymbol{\xi}$ and $\boldsymbol{\Omega}$, we use a Gaussian random-walk Metropolis-Hasting algorithm to draw from these distributions. We draw recursively from the following posterior distributions:

- [Metropolis-Hastings]
$$f(\boldsymbol{\Omega} \,|\, \boldsymbol{\mu}_\Omega, \boldsymbol{\Sigma}_\Omega, \boldsymbol{\xi}, \text{data}) \propto \exp\left(.5(\boldsymbol{\Omega} - \boldsymbol{\mu}_\Omega)' \boldsymbol{\Sigma}_\Omega^{-1} (\boldsymbol{\Omega} - \boldsymbol{\mu}_\Omega)\right) \prod_{i=1}^{I} \mathscr{L}_i(\boldsymbol{\Omega}, \boldsymbol{\xi}_i \,|\, \text{data}).$$

- [Metropolis-Hastings]
$$f(\boldsymbol{\xi}_i \,|\, \boldsymbol{\Sigma}_\xi, \boldsymbol{\Omega}, \text{data}) \propto \exp\left(.5 \boldsymbol{\xi}_i' \boldsymbol{\Sigma}_\xi^{-1} \boldsymbol{\xi}_i\right) \mathscr{L}_i(\boldsymbol{\Omega}, \boldsymbol{\xi}_i \,|\, \text{data}), \; \forall\, i.$$

- [Gibbs]
$$f(\Sigma_\xi \,|\, \boldsymbol{\xi}, \boldsymbol{R}, \text{df}) \sim \text{inv-Wishart}\left(\sum_{i=1}^{I} \boldsymbol{\xi}_i' \boldsymbol{\xi}_i + \text{df}\boldsymbol{R}^{-1}, \text{df} + I \right).$$

For the Metropolis-Hastings steps, we follow the procedure proposed by Atchadé (2006) and adapt the tuning parameters in each iteration to get an acceptance rate of approximately 20%. In the empirical analyses reported in the paper, we ran the simulation for 500,000 iterations. The first 400,000 iterations were used as a "burn-in" period, and the last 100,000 iterations were used to estimate the conditional posterior distributions.

References

Atchadé YF (2006) An adaptive version for the Metropolis adjusted Langevin algorithm with a truncated drift. *Methodol. Comput. Appl. Probab.* 8(2):235–254.

Web Appendix B: Additional Results: Empirical Application 1

B1 Additional results: Three-state specification

For ease of exposition, the main document shows the impact of changes in the covariates on the transition probabilities. Table B1 reports the posterior estimates of the coefficients for the covariates included in the state transition process of the proposed model with three states.

Variable	Posterior mean	95% CPI	
Lag(QualStock)			
From state 1 to 1	−0.842	−2.232	0.567
From state 1 to 2	**1.456**	0.210	2.827
From state 2 to 1	0.975	−0.360	2.264
From state 2 to 2	0.768	−0.849	2.237
From state 3 to 1	**−1.589**	−2.701	−0.387
From state 3 to 2	0.777	−0.815	2.304
Number of periods since last email			
From state 1 to 1	−0.038	−0.710	1.005
From state 1 to 2	−1.130	−2.639	0.638
From state 2 to 1	0.572	−0.276	1.584
From state 2 to 2	−0.437	−1.198	0.294
From state 3 to 1	0.449	−0.073	1.019
From state 3 to 2	0.114	−1.056	1.080

Table B1: Posterior means of the coefficients for the covariates included in the state transition process. Numbers in bold are associated with 95% CPIs not including 0.

B2 Results for the one-, two-, and four-state specifications

In this appendix we present the results for the one-, two-, and four-state specifications of our proposed model. We focus on the two main sets of results: the state-specific probabilities for all three behaviors (open, click, and unsubscribe) and the transition probabilities among the latent states.

B2.1 Model with one state ($K = 1$)

We start by describing the results of the most restricted model, in which only one latent state is allowed. As such, it is a static specification and there are no parameters governing dynamics.

		Posterior mean	95% CPI	
Prob(Open)	State 1	0.286	0.268	0.304
Prob(Click)	State 1	0.059	0.048	0.072
Prob(Unsubscribe)	State 1	0.002	0.001	0.002

Table B2: State-dependent probabilities for the one-state model.

B2.2 Model with two states ($K = 2$)

We now turn to the results of the simplest dynamic specification, one with two hidden states. Combining the results from Tables B3 and B4, and comparing them with those obtained when three hidden states are allowed (Tables 4 and 5 in the main manuscript) we observe that a state of high activity is identified (state 1). This state capture customers with high activity in *all three behaviors*. Unlike the insights derived from the three-state specification, the model with two states fails to separate the very engaged customers from those who open frequently, but rarely click and are more likely to unsubscribe. The model clearly captures the state of the "silently gone" customers, who rarely interact with the service. As we find in the three-state specification, such a state is highly absorbing.

		Posterior mean	95% CPI	
Prob(Open)	State 1	0.600	0.553	0.641
	State 2	0.068	0.053	0.084
Prob(Click)	State 1	0.178	0.135	0.229
	State 2	0.026	0.011	0.043
Prob(Unsubscribe)	State 1	0.003	0.001	0.005
	State 2	0.001	0.000	0.001

Table B3: State-dependent probabilities for the two-state model.

	To state	
From state	1	2
1	0.796	0.204
	[0.546 , 0.987]	[0.013 , 0.454]
2	0.040	0.960
	[0.001 , 0.141]	[0.859 , 0.999]

Table B4: Mean transition probabilities and the 95% heterogeneity interval of individual posterior means for the two-state model.

B2.3 Model with four states ($K = 4$)

We finally consider the results of an HMM with four states and compare them with the more parsimonious specification of an HMM with three states. Comparing Tables B5 and B6 with Tables 4 and 5, we see that allowing additional flexibility in the model (i.e., adding an additional state) sees the "silently gone" state in the three-state model divided into two states, both of which are very sticky and have low probabilities of any kind of activity. Moreover, in contrast to the three-state specification, the four-state model is not as effective at separating the "at risk" and "engaged" states, as suggested by the distributions of clicking and unsubscribing probabilities— they have closer posterior means and wider 95% CPIs than those obtained with the three-state model.

		Posterior mean	95% CPI	
Prob(Open)	State 1	0.656	0.617	0.695
	State 2	0.683	0.637	0.729
	State 3	0.069	0.055	0.084
	State 4	0.080	0.067	0.093
Prob(Click)	State 1	0.115	0.068	0.171
	State 2	0.239	0.140	0.344
	State 3	0.015	0.007	0.031
	State 4	0.001	0.000	0.002
Prob(Unsubscribe)	State 1	0.003	0.002	0.005
	State 2	0.004	0.001	0.007
	State 3	0.001	0.000	0.002
	State 4	0.001	0.000	0.002

Table B5: State-dependent probabilities for the four-state model.

From state	To state 1	2	3	4
1	0.568	0.205	0.041	0.187
	[0.128 , 0.992]	[0.000 , 0.581]	[0.000 , 0.080]	[0.007 , 0.303]
2	0.186	0.370	0.265	0.179
	[0.102 , 0.304]	[0.225 , 0.588]	[0.057 , 0.453]	[0.087 , 0.229]
3	0.005	0.061	0.903	0.031
	[0.000 , 0.013]	[0.001 , 0.206]	[0.715 , 0.996]	[0.003 , 0.067]
4	0.039	0.006	0.013	0.941
	[0.022 , 0.060]	[0.004 , 0.009]	[0.008 , 0.02]	[0.917 , 0.964]

Table B6: Mean transition probabilities and the 95% heterogeneity interval of individual posterior means for the four-state model.

B3 Benchmark model details and results

In Section 4.6 we compare the out-of-sample predictions generated by our model with benchmark models in which each of the three behaviors — opening, clicking, and unsubscribing — are modeled as a function of lagged covariates and/or past behavior (recency and frequency). We model each of the three behaviors — unsubscribing, opening, and clicking — using binary logit models. More formally,

$$\bar{o}_{it} = P(Y^o_{it} = 1) = \frac{e^{\mathbf{z}^o_{it}\boldsymbol{\lambda}^o_i}}{1 + e^{\mathbf{z}^o_{it}\boldsymbol{\lambda}^o_i}} \tag{B1}$$

$$\bar{c}_{it} = P(Y^c_{it} = 1) = \begin{cases} \dfrac{e^{\mathbf{z}^c_{it}\boldsymbol{\lambda}^c_i}}{1 + e^{\mathbf{z}^c_{it}\boldsymbol{\lambda}^c_i}} & \text{if } y^o_{it} = 1 \\ 0 & \text{if } y^o_{it} = 0\,, \end{cases} \tag{B2}$$

$$\bar{u}_{it} = P(Y^u_{it} = 1) = \begin{cases} \dfrac{e^{\mathbf{z}^u_{it}\boldsymbol{\lambda}^u_i}}{1 + e^{\mathbf{z}^u_{it}\boldsymbol{\lambda}^u_i}} & \text{if } y^o_{it} = 1 \\ 0 & \text{if } y^o_{it} = 0\,, \end{cases} \tag{B3}$$

where $\mathbf{z}^o_{it}$, $\mathbf{z}^c_{it}$, and $\mathbf{z}^u_{it}$ represent the covariates affecting each of the behaviors. Using the same logic as in (8), the probability that customer i has observed behavior $\boldsymbol{y}_{it} = [\, y^o_{it}, y^c_{it}, y^u_{it} \,]$ in period t is

$$\begin{aligned} P(\boldsymbol{Y}_{it} = \boldsymbol{y}_{it}) = \mathbb{1}(y^o_{it} = 1)\bar{o}_{it}\Big\{ &\big[\mathbb{1}(y^c_{it} = 1)\bar{c}_{it} + \mathbb{1}(y^c_{it} = 0)(1 - \bar{c}_{it})\big] \\ &\times \big[\mathbb{1}(y^u_{it} = 1)\bar{u}_{it} + \mathbb{1}(y^u_{it} = 0)(1 - \bar{u}_{it})\big]\Big\} \\ &+ \mathbb{1}(y^o_{it} = 0)(1 - \bar{o}_{it})\,. \end{aligned} \tag{B4}$$

It follows that customer i's likelihood function is

$$\mathscr{L}_i(\boldsymbol{\lambda_i} \,|\, \text{data}) = \prod_{t=1}^{T_i} P(\boldsymbol{Y}_{it} = \boldsymbol{y}_{it}) \,, \tag{B5}$$

where $\boldsymbol{\lambda_i}$ contains $\boldsymbol{\lambda}_i^u$, $\boldsymbol{\lambda}_i^o$, and $\boldsymbol{\lambda}_i^c$ (i.e., the parameters in (B1)–(B3)). As described in Section 4.6, we consider three specifications of this model: lagged covariates, RF (no covariates) and RF (covariates). Tables B7–B9 report the parameter estimates for these models.

Our fourth benchmark model, HMM (static) is a version of our HMM in which transitions between states are not allowed (i.e., $\boldsymbol{Q}_{it}$ is the identity matrix). The associated parameter estimates are reported in Table B10.

Behavior	Variable	Posterior mean	95% CPI	
Open	Intercept	**−0.895**	−0.921	−0.867
	Sunday	**0.103**	0.037	0.167
	Lag(QualStock)	**−1.437**	−1.626	−1.283
	Number of periods since last email	**0.089**	0.035	0.143
	Lag(Sunday)	0.030	−0.026	0.085
	Lag2(QualStock)	**−0.071**	−0.130	−0.010
	Lag(Number of periods since last email)	**0.072**	0.011	0.131
Click | Open	Intercept	**−0.907**	−1.121	−0.672
	Sunday	**−0.194**	−0.335	−0.042
	log(#deals)	**0.495**	0.331	0.665
	Discount	**−0.056**	−0.084	−0.033
	Time left	**0.002**	0.000	0.005
	Food	**−0.074**	−0.105	−0.046
	Fitness	**−0.039**	−0.062	−0.020
	Source	−0.001	−0.015	0.017
	Order	**−0.030**	−0.047	−0.016
	Lag(QualStock)	**3.052**	1.848	4.215
	Number of periods since last email	0.008	−0.081	0.095
	Lag(Sunday)	**−0.131**	−0.262	−0.003
	Lag(log(#deals))	0.025	−0.084	0.139
	Lag(Avg. Discount)	0.168	−0.635	0.935
	Lag(Avg. Time left)	−0.010	−0.026	0.008
	Lag(%Food)	−0.062	−0.265	0.139
	Lag(%Fitness)	**1.079**	0.738	1.421
	Lag(Avg. Source)	0.098	−0.088	0.301
	Lag2(QualStock)	−0.005	−0.114	0.107
	Lag(Number of periods since last email)	0.006	−0.107	0.115
Unsubscribe | Open	Intercept	**−5.149**	−5.546	−4.773
	Avg. Discount	**1.248**	0.499	1.951
	Avg. Time left	0.005	−0.062	0.067
	%Food	**1.244**	0.702	1.819
	%Fitness	**0.935**	0.318	1.567
	Avg. Source	0.164	−0.384	0.738
	Sunday	**−0.603**	−1.182	−0.003
	log(#deals)	−0.284	−0.654	0.093
	Lag(QualStock)	**1.915**	0.870	2.943
	Number of periods since last email	0.161	−0.171	0.446
	Lag(Avg. Discount)	0.723	−0.034	1.383
	Lag(Avg. Time left)	−0.058	−0.165	0.031
	Lag(%Food)	**−0.661**	−1.190	−0.092
	Lag(%Fitness)	**0.980**	0.147	1.756
	Lag(Avg. Source)	0.005	−0.647	0.707
	Lag(Sunday)	0.151	−0.356	0.650
	Lag(log(#deals))	**0.703**	0.222	1.209
	Lag2(QualStock)	0.089	−0.274	0.541
	Lag(Number of periods since last email)	−0.090	−0.543	0.273

Table B7: Parameter estimates for the lagged covariates model. Numbers in bold are associated with 95% CPIs not including 0.

Behavior	Variable	Posterior mean	95% CPI	
Open	Intercept	$\mathbf{-1.980}$	-2.064	-1.894
	rec^o	$\mathbf{-0.072}$	-0.083	-0.062
	freq^o	**3.512**	3.392	3.633
	rec^c	0.004	-0.001	0.009
	freq^c	$\mathbf{-1.106}$	-1.275	-0.939
Click | Open	Intercept	$\mathbf{-0.426}$	-0.587	-0.266
	rec^o	**0.052**	0.030	0.075
	freq^o	$\mathbf{-1.733}$	-1.944	-1.520
	rec^c	$\mathbf{-0.083}$	-0.096	-0.071
	freq^c	**2.535**	2.269	2.804
Unsubscribe | Open	Intercept	$\mathbf{-4.694}$	-5.270	-4.143
	rec^o	**0.064**	0.003	0.122
	freq^0	-0.636	-1.345	0.064
	rec^c	-0.002	-0.041	0.034
	freq^c	0.539	-0.458	1.512

Table B8: Parameter estimates for the RF (no covariates) model. Numbers in bold are associated with 95% CPIs not including 0.

Behavior	Variable	Posterior mean	95% CPI	
Open	Intercept	**−1.923**	−2.010	−1.835
	rec^o	**−0.078**	−0.089	−0.068
	freq^o	**3.435**	3.320	3.550
	rec^c	0.002	−0.003	0.007
	freq^c	**−0.909**	−1.085	−0.731
	Sunday	**0.131**	0.056	0.205
Click | Open	Intercept	−0.008	−0.248	0.236
	rec^o	**0.052**	0.030	0.073
	freq^o	**−1.751**	−1.951	−1.558
	rec^c	**−0.082**	−0.094	−0.070
	freq^c	**2.625**	2.324	2.907
	Sunday	**−0.200**	−0.333	−0.062
	log(#deals)	**0.409**	0.262	0.562
	Discount	**−0.019**	−0.036	−0.004
	Time left	**−0.020**	−0.035	−0.005
	Food	0.001	−0.001	0.004
	Fitness	**−0.050**	−0.083	−0.020
	Source	**−0.024**	−0.046	−0.007
	Order	0.000	−0.003	0.002
Unsubscribe | Open	Intercept	**−4.826**	−5.426	−4.212
	rec^o	**0.064**	0.001	0.123
	freq^o	**−0.697**	−1.391	−0.015
	rec^c	−0.001	−0.040	0.035
	freq^c	0.720	−0.340	1.719
	Avg. Discount	0.525	−0.653	1.725
	Avg. Time left	−0.006	−0.078	0.056
	%Food	0.474	−0.131	1.057
	%Fitness	−0.144	−1.205	0.926
	Avg. Source	−0.080	−0.727	0.631
	Sunday	**−0.620**	−1.316	−0.022
	log(#deals)	−0.084	−0.392	0.240

Table B9: Parameter estimates for the RF (covariates) model. Numbers in bold are associated with 95% CPIs not including 0.

	Segment 1	Segment 2	Segment 3
Effect on Opening			
Sunday	**−1.307**	**0.145**	**2.522**
	[−2.449 , −0.195]	[0.046 , 0.252]	[1.403 , 3.520]
Effect on Clicking, given Opening			
Sunday	0.604	**−0.228**	1.270
	[−0.76 , 2.02]	[−0.414 , −0.062]	[−0.429 , 2.901]
log(#deals)	0.100	**0.740**	0.826
	[−1.498 , 1.858]	[0.498 , 0.982]	[−0.777 , 2.414]
Discount	0.112	**0.147**	−0.255
	[−0.16 , 0.506]	[0.041 , 0.264]	[−1.291 , 0.413]
Time left	−0.082	**0.003**	−0.166
	[−0.218 , 0.020]	[0.000 , 0.006]	[−0.561 , 0.010]
Food	0.135	**−0.125**	−0.113
	[−0.102 , 0.506]	[−0.163 , −0.087]	[−0.855 , 0.511]
Fitness	0.022	**0.148**	0.412
	[−0.243 , 0.337]	[0.057 , 0.242]	[−0.216 , 1.660]
Source	0.025	0.003	−0.092
	[−0.223 , 0.303]	[−0.035 , 0.041]	[−0.956 , 0.577]
Order	**−0.160**	**−0.056**	−0.392
	[−0.347 , −0.011]	[−0.086 , −0.028]	[−1.22 , 0.012]
Effect on Unsubscribing, given Opening			
Avg. Discount	0.057	−0.404	−0.012
	[−1.726 , 1.803]	[−1.898 , 1.224]	[−1.903 , 1.975]
Avg. Time left	0.081	−0.023	−0.339
	[−0.173 , 0.366]	[−0.127 , 0.063]	[−1.134 , 0.308]
%Food	−0.334	0.765	−0.108
	[−2.096 , 1.497]	[−0.155 , 1.696]	[−1.605 , 1.449]
%Fitness	1.222	**−2.197**	−1.377
	[−0.562 , 3.14]	[−3.904 , −0.566]	[−3.529 , 0.867]
Avg. Source	−0.702	0.068	−0.365
	[−2.568 , 1.277]	[−1.021 , 1.290]	[−2.141 , 1.566]
Sunday	−0.216	**−1.033**	−0.103
	[−1.931 , 1.48]	[−2.096 , −0.221]	[−1.692 , 1.402]
log(#deals)	−1.273	0.530	0.501
	[−2.379 , 0.017]	[−0.259 , 1.605]	[−1.687 , 2.609]

Table B10: Parameter estimates (posterior means) for the HMM (static) model. Numbers in bold are associated with 95% CPIs (in brackets) not including 0.

Web Appendix C: Additional Results: Empirical Application 2

C1 Additional results: Three state specification

In this appendix we report the results from the second empirical application that were not discussed in the main manuscript, namely the initial state probabilities (Table C1) and the posterior estimates of the coefficients for the covariates included in the models of state-dependent behavior (Table C2) for the proposed model with three states.

	Posterior mean	95% CPI	
State 1	0.486	0.428	0.547
State 2	0.291	0.236	0.356
State 3	0.223	0.151	0.295

Table C1: Initial state probabilities.

	State		
	"At Risk"	"Engaged"	"Silently gone"
Effect on Opening			
Ticket	−0.086	**0.852**	**1.261**
	[−0.231 , 0.064]	[0.470 , 1.256]	[0.731 , 1.941]
Donation	−0.070	−0.543	−0.951
	[−0.274 , 0.146]	[−1.295 , 0.101]	[−2.361 , 0.152]
Effect on Clicking, given Opening			
Ticket	0.331	0.357	**−1.623**
	[−0.160 , 0.862]	[−0.883 , 1.575]	[−2.808 , −0.377]
Donation	**−1.794**	**−2.277**	−1.452
	[−3.021 , −0.754]	[−3.604 , −1.004]	[−3.763 , 1.090]
Effect on Unsubscribing, given Opening			
Ticket	−0.052	−1.569	0.483
	[−0.770 , 0.647]	[−3.621 , 0.292]	[−0.923 , 1.900]
Donation	0.724	0.937	0.450
	[−0.004 , 1.478]	[−1.572 , 3.120]	[−2.097 , 2.688]

Table C2: Posterior means of the effect of the covariates on the state-dependent probabilities. Numbers in bold are associated with 95% CPIs (in brackets) not including 0.

C2 Results for the four-state specification

Recall from Section 5 that the models with three and four latent states provided very similar measures of fit. While the model with three states had the best fit in terms of MSE, the model with four states had lower log predictive density and WAIC. In this section we present the results of the model with four states and discuss how they differ from those obtained using the three-state model.

		Posterior mean	95% CPI	
Prob(Open)	State 1	0.637	0.589	0.682
	State 2	0.992	0.981	0.998
	State 3	0.092	0.069	0.118
	State 4	0.013	0.007	0.019
Prob(Click)	State 1	0.017	0.011	0.023
	State 2	0.897	0.814	0.963
	State 3	0.904	0.820	0.973
	State 4	0.218	0.183	0.256
Prob(Unsubscribe)	State 1	0.007	0.007	0.005
	State 2	0.009	0.007	0.001
	State 3	0.000	0.000	0.000
	State 4	0.001	0.001	0.001

Table C3: State-dependent probabilities.

Comparing Table C3 with Table 12, we see that allowing additional flexibility in the model (i.e., adding an additional state) results in the "engaged" state from the three-state model being divided into two states, one state in which customers open almost every e-mail (state 2) and another state in which customers open less frequently (state 3). Moreover, if we label state 1 as "at risk" and state 4 as "silently gone," these two states clearly coincide in terms of customer behavior with their corresponding states in the three-state model. The transition matrix (Table C4) is consistent with this intuition. The transition probabilities related to the "at risk" and "silently gone" states are consistent with those obtained in the three-state model. Regarding the two "engaged" states, the state with very high opening behavior is not very sticky, while the other state is more stable.

From state	To state 1	2	3	4
1	0.965	0.000	0.000	0.035
	[0.867 , 0.998]	[0.000 , 0.000]	[0.000 , 0.001]	[0.002 , 0.132]
2	0.000	0.319	0.170	0.511
	[0.000 , 0.000]	[0.149 , 0.411]	[0.037 , 0.36]	[0.242 , 0.800]
3	0.155	0.034	0.809	0.002
	[0.013 , 0.395]	[0.001 , 0.181]	[0.541 , 0.983]	[0.001 , 0.003]
4	0.000	0.001	0.002	0.996
	[0.000 , 0.000]	[0.000 , 0.003]	[0.001 , 0.005]	[0.993 , 0.998]

Table C4: Mean transition probabilities and the 95% heterogeneity interval of individual posterior means.