

Quantifying the Benefits of Individual-Level Targeting in the Presence of Firm Strategic Behavior

Xiaojing Dong, Puneet Manchanda and Pradeep K. Chintagunta

Web Appendix

SOLVING THE SOC INEQUALITY IN EQUATION (9)

In equation (9), the first two items, $markup_b$ and $\exp(u_{pbt})$, are both positive, therefore it requires that

$$\left(\frac{2\beta_{pb,b}}{(dtl_{pbt} + 1)^3} + \left(\frac{\beta_{pb,b}}{(dtl_{pbt} + 1)^2} \right)^2 \right) < 0. \quad (W1)$$

Given that $dtl_{pbt} \geq 0$, it can be simplified as $2\beta_{pb,b} + \frac{\beta_{pb,b}^2}{(dtl_{pbt} + 1)} < 0$

Based on previous literature, detailing has positive effect on prescription, we expect $\beta_{pb,b} < 0$. In this case, solving this inequality, we can get

$$dtl_{pbt} > \frac{-\beta_{pb,b}}{2} - 1. \quad (W2)$$

As mentioned earlier, W2 is satisfied for over 98% of the observations.

In order to specify an MCMC sampler that embeds this constraint, we need to place constraints on *each* $\beta_{pb,b}$ as the constraint is a function of the realized data (detailing) for each physician. In a Bayesian hierarchical setting, this can be done either via a transformation of variables that results in a constrained likelihood or via the prior. In the first case, the transform has to be applied for each of the 1200 (300 physicians time 4 brand coefficients for each brand for each physician) parameters while in the second case, the prior distributions have to be individual specific. Both cases are computationally infeasible. Even in a classical setting, these constraints imply an optimization procedure with 1200 constraints. As can be expected, this optimization procedure will be poorly behaved.

GIBBS SAMPLER OVERVIEW:

In the following equations, we use * to represent all other parameters

1. Define $\beta_p = \{\beta_{pb,0}, \beta_{pb,b}, \beta_{pb,b'}, \beta_{pb,l}\}$ for $\forall b, b' \neq b$, which is the vector of all the parameters for individual i ; and $\beta_{p1} = \{\beta_{pb,b}, \beta_{pb,b'}, \beta_{pb,l}\}$ for $\forall b, b' \neq b$, which is the sub-vector of β_p , without the intercepts. Draw β_p for each physician, all brands.

$$\begin{aligned}
 & \left[\beta_p \mid * \right] \propto \\
 & \prod_{t,b} \text{Poisson}(\lambda_{pbt}) \quad \text{Likelihood from the prescription model} \\
 & \times \prod_{t,b} \left[\frac{\partial r}{\partial dtl_{pbt}} \left(\text{markup}_{pb} \times \exp(u_{pbt}) \times \frac{-\beta_{pb,b}}{(dtl_{pbt} + 1)^2} - (\alpha_{b0} + X_p \alpha_b + s_b) \sim N(0, \Sigma_\eta) \right) \right] \\
 & \quad \text{Likelihood from the detailing model} \\
 & \times \left[\beta_{p1} \sim N(\bar{\beta}, \Sigma_\beta) \right] \times \prod_b \left[\beta_{pb,0} \sim N(Z\theta_b, \Sigma_v) \right] \quad \text{Prior}
 \end{aligned}$$

Where r is the function defined as the left hand side of the FOC in equation (8), that is

$$r = \text{markup}_{pb} \times \exp(u_{pbt}) \times \frac{-\beta_{pb,b}}{(dtl_{pbt} + 1)^2}$$

2. Define $\xi_{pt} = \{\xi_{pbt}\}$ for $\forall b$, draw ξ_{pt} , a B (number of brands) dimensional vector for the prescription model errors for each physician-quarter observation.

$$\begin{aligned}
 & \left[\xi_{pt} \mid * \right] \propto \\
 & \prod_b \text{Poisson}(\lambda_{pbt}) \quad \text{Likelihood from the prescription model} \\
 & \times \prod_b \left[\frac{\partial r}{\partial dtl_{pbt}} \left(\text{markup}_{pb} \times \exp(u_{pbt}) \times \frac{-\beta_{pb,b}}{(dtl_{pbt} + 1)^2} - (\alpha_{b0} + X_p \alpha_b + s_b) \sim N(0, \Sigma_\eta) \right) \right] \\
 & \quad \text{Likelihood from the detailing model} \\
 & \times \left[\xi_{pt} \sim N(0, \Sigma_\xi) \right] \quad \text{Prior}
 \end{aligned}$$

Note that the full conditional posterior distributions for the parameters β_p and the random shock vectors in the prescription model ξ_{pt} are quite similar in their likelihood functions, in that both have the likelihood of Poisson distribution and the likelihood based on the derived distribution of detailing from the FOC. The differences are the data incorporated in the likelihood functions. The likelihood for β_p consists of all the observations for the same physician; and the likelihood for ξ_{pt} contains only one observation for the physician-quarter data.

3. Draw $\bar{\beta}_b, \Sigma_\beta$ with normal and Wishart conjugate prior.

The conditional posterior distribution of $\bar{\beta}_b$ is $N(\mu_\beta, \Delta_\beta)$, where

$$\Delta_\beta = \left(P \times \Sigma_\beta^{-1} + \Delta_0^{-1} \right)^{-1}$$

$$\mu_\beta = \Delta_\beta \times \left(\Sigma_\beta^{-1} \times \sum_p \beta_{p1} + \Delta_0^{-1} \times \mu_0 \right)$$

Where P is the number of physicians in the analysis. $\Delta_0^{-1} = 1000I$, $\mu_0 = \{0\}$.

The conditional posterior distribution of $\Sigma_\beta \propto$ Inverted Wishart $\left(\sum_p (\beta_{p1} - \bar{\beta})(\beta_{p1} - \bar{\beta})' + V_0, P + n_0 \right)$

Where $n_0 = 22$ and $V_0 = n_0 I$

Above three conditional posterior distributions are related directly with the estimates in the prescription model, and the followings are for the estimates in the detailing model.

4. Draw α_{b0} , α_b and s_b

Conditional on all the other parameters, including the physician level response parameters β_p , random shocks in the prescription model ξ_{pt} , we can compute the total marginal cost mc_{pb} using the FOC in equation (8). Based on the equation $mc_{pbt} = \alpha_{b0} + X_p \alpha_b + s_b + \eta_{pbt}$, we can obtain the parameters using multivariate normal regression, as it is assumed that $\eta_{pbt} \sim N(0, \Sigma_\eta)$. For more details of normal regression, please refer to the book by Rossi et al. (2006).

5. Draw the covariance matrices for both prescription model errors Σ_ξ and detailing model errors Σ_η at the same time, as well as the correlations between the two sets of errors.

Compute the detailing model errors by substituting the estimated parameters into the equation $\eta_{pbt} = mc_{pbt} - (\alpha_{b0} + X_p \alpha_b + s_b)$. Put these computed errors from detailing model η_{pbt} right to those from the prescription model (drawn in the second step above) ξ_{pbt} together, which forms a bigger matrix with dimension $N \times 2B$, where N is the number of observations in the data and B is the number of brands. Using this matrix as data, we can draw $\Sigma_{\xi, \eta}$ with dimension $2B \times 2B$, the covariance matrix for all the errors together, using conjugate inverted Wishart prior. From this matrix $\Sigma_{\xi, \eta}$, the first B rows and B columns on the diagonal is the covariance matrix Σ_{ξ} , the last B rows and B columns on the diagonal is the covariance matrix Σ_{η} . The rest entries in $\Sigma_{\xi, \eta}$ are the covariance between the two sets of errors.