Bayesian Estimation in IO

Paul Schrimpf

UBC
Economics 565

April 1, 2020

Section 1

Jiang, Manchanda, and Rossi (2009)

Jiang, Manchanda, and Rossi (2009)

- Bayesian BLP
- BLP uses moment conditions
  - Consumer i utility from buying product j in market t

$$u_{ijt} = \overbrace{x_{jt}}^{1\times K} \underbrace{\overbrace{\theta_{it}}^{K\times 1}}_{=\bar{\theta}+\nu_{it}} + \overbrace{\xi_{jt}}^{1\times 1} + \epsilon_{ijt}$$

  - Common demand shock ξ_{jt} endogenous, have instruments w

$$\mathrm{E}[\xi_{jt}|w_{jt}] = 0$$

- Bayesian needs likelihood, so assume

$$x_{jt} = w_{jt}\delta + u_{jt}$$

and

$$\begin{pmatrix} u_{jt} \\ \xi_{jt} \end{pmatrix} \sim N(0, \Omega).$$

Model & likelihood 1

- Utility:

$$u_{ijt} = \overbrace{x_{jt}}^{1\times K} \overbrace{\underbrace{\theta_{it}}_{=\bar{\theta}+\nu_{it}}}^{K\times 1} + \overbrace{\xi_{jt}}^{1\times 1} + \epsilon_{ijt}$$

- First stage

$$x_{jt} = w_{jt}\delta + u_{jt}$$

- Distributional assumptions:
    - $\epsilon_{ijt} \sim$ type I extreme value
    - $\nu_{it} \sim N(0, \Sigma)$, i.i.d across i, t
    - $\begin{pmatrix} u_{jt} \\ \xi_{jt} \end{pmatrix} \sim N(0, \Omega)$, i.i.d across j, t

Model & likelihood 2

- Share equation:

$$s_{jt} = \int \frac{\exp(x_{jt}(\bar{\theta} + \nu) + \xi_{jt})}{1 + \sum_{k=1}^{j} \exp(x_{kt}(\bar{\theta} + \nu) + \xi_{kt})} dF_{\nu}(\nu; \Sigma)$$
$$= h(\xi_t | x_t, \bar{\theta}, \Sigma)$$

- Likelihood:

$$\pi(s_t, x_t | w_t, \bar{\theta}, \Sigma, \delta, \Omega) = \phi\left(\begin{pmatrix} h^{-1}(s_t | x_t, \bar{\theta}, \Sigma) \\ x_t - w_t \delta \end{pmatrix} | \Omega \right) (J_{s_t \to \xi_t})^{-1}$$

where $J_{s_t \to \xi_t}$ = determinate of $ds_t / d\xi_t$

- $J_{s_t \to \xi_t}$ given shares is function of only Σ

Prior

- $\bar{\theta} \sim N(\theta_0, V_\theta)$
- $\delta \sim N(\delta_0, V_\delta)$
- $\Omega \sim$ inverse Wishart(ν_0, V_Ω)
- $\Sigma = U'U$, $U = \begin{pmatrix} e^{r_{11}} & r_{12} & \cdots & r_{1K} & \\ 0 & e^{r_{22}} & r_{23} & \ddots & \vdots \\ \vdots & & \ddots & & \vdots \\ 0 & \cdots & \cdots & e^{r_{kk}} & \end{pmatrix}$
    - $r_{jj} \sim N(0, \sigma^2_{r_{jj}})$
    - $r_{jk} \sim N(0, \sigma_{r^2_{off}})$ for $j < k$

MCMC

- Combination of Gibbs and random walk Metropolis Hastings
- Gibbs sampler for $\bar{\theta}, \delta, \Omega | r, s, x, w$, priors
- Metropolis for $\Sigma = r | \bar{\theta}, \delta, \Omega, r, s, x, w$
  - Candidate density: $r^{new} = r^{old} + N(0, \sigma^2 D_r)$

Advantages

- No maximization (but problem of chain convergence instead)
- Simulations show lower MSE than GMM (even in simulations with ξ not normally distributed
- Inference natural by-product of MCMC
    - No extra work needed to compute standard errors
    - Inference on functions of parameters straightforward, no need for delta-method
        - Sample from posterior of $f(\theta)$ by drawing θ from posterior and calculating $f(\theta)$
        - e.g. elasticities, any counterfactuals, etc
- In simulations GMM asymptotic confidence intervals too small, MCMC gets closet to correct coverage of credible regions

Simulation results

- $J = 3$, $T = 300$, $K = 4$ (brand effects and price)
- Setup 1:
  - no endogeneity $x = w$
  - Distributions for ξ
    - Correctly specified: $\xi \sim N(0,1)$
    - Heteroskedasticity: $\xi \sim N(0, \exp(-.5413 + x_{jt}^p))$
    - AR(1) $\xi_{jt} = \rho \xi_{jt-1} + N(0, v)$
    - Different distributoin: $\xi \sim$ Beta with parameters such that either symmetric or assymetric
- Setup 2:
  - One component of x endogenous, $w =$ exogenous x's and one instrument

Table 1
MSE and bias for estimates of τ^2 and $\bar{\theta}$.

		MSE		Bias	
		Bayes	GMM	Bayes	GMM
τ^2	*i.i.d. N*	0.02	0.09	−0.13	−0.03
	Hetero	0.009	0.134	−0.021	−0.011
	AR(1)	0.049	0.227	−0.172	−0.058
	Asym Beta	0.002	0.007	−0.044	−0.002
	Sym Beta	0.001	0.006	−0.03	−0.01
$\bar{\theta}_1$	*i.i.d. N*	0.11	0.54	0.22	0.13
	Hetero	0.53	0.43	−0.37	0.18
	AR(1)	0.22	0.55	0.24	0.16
	Asym Beta	0.12	0.5	0.23	0.02
	Sym Beta	0.17	0.29	0.31	0.33
$\bar{\theta}_2$	*i.i.d. N*	0.26	0.54	0.25	0.29
	Hetero	0.87	1.04	−0.52	0.25
	AR(1)	0.39	1.7	0.22	0.33
	Asym Beta	0.29	2.04	0.45	−0.14
	Sym Beta	0.25	1.52	0.33	0.10
$\bar{\theta}_3$	*i.i.d. N*	0.25	8.51	0.27	−1.51
	Hetero	2.00	12.08	−0.93	−2.02
	AR(1)	0.84	10.92	0.14	−1.46
	Asym Beta	0.41	9.39	0.50	−1.11
	Sym Beta	0.38	5.01	0.32	−1.04
$\bar{\theta}_{\text{price}}$	*i.i.d. N*	0.41	1.71	0.28	0.47
	Hetero	0.85	2.16	0.62	0.67
	AR(1)	0.59	2.39	−0.10	0.33
	Asym Beta	0.51	2.27	0.6	0.37
	Sym Beta	0.34	2.48	0.23	0.29

Table 2
MSE and bias for diagonal Σ elements.

		MSE		Bias	
		Bayes	GMM	Bayes	GMM
Σ_{11}	*i.i.d. N*	1.94	14.89	−1.04	0.13
	Hetero	11.07	25.81	1.85	0.23
	AR(1)	3.91	35.43	−0.38	0.32
	Asym Beta	2.17	66.28	−1.22	1.20
	Sym Beta	2.14	8.49	−1.19	−1.05
Σ_{22}	*i.i.d. N*	2.63	9.52	−0.70	−0.46
	Hetero	15.73	26.65	2.78	−0.33
	AR(1)	5.3	181.16	0.13	0.83
	Asym Beta	4.00	87.09	−1.71	1.96
	Sym Beta	2.34	38.38	−0.92	0.46
Σ_{33}	*i.i.d. N*	1.95	498.86	−0.62	10.68
	Hetero	40.2	566.35	3.6	12.35
	AR(1)	8.08	1927.91	0.50	15.95
	Asym Beta	3.47	601.03	−1.33	8.83
	Sym Beta	3.04	163.88	−0.42	6.04
Σ_{44}	*i.i.d. N*	2.23	21.73	0.45	2.19
	Hetero	5.12	23.11	1.16	2.33
	AR(1)	5.41	24.05	1.44	2.40
	Asym Beta	0.71	64.72	−0.16	2.73
	Sym Beta	2.42	21.58	0.50	1.91

Table 7
MSE and bias for the IV sampling experiment.

	MSE		Bias	
	Bayes	GMM	Bayes	GMM
$\bar{\theta}_1$	0.50	9.89	0.49	−0.93
$\bar{\theta}_2$	0.44	13.46	0.51	−1.28
$\bar{\theta}_3$	0.41	34.11	0.41	−2.16
$\bar{\theta}_{\text{price}}$	0.28	10	0.33	−0.02
Σ_{11}	3.82	315.49	−1.59	6.86
Σ_{22}	3.11	383.2	−1.51	8.74
Σ_{33}	3.68	6301.31	−1.30	19.09
Σ_{44}	0.75	104.68	−0.06	4.02
Σ_{12}	2.33	117.63	−1.24	1.59
Σ_{13}	1.64	82.45	−1.00	1.20
Σ_{23}	1.92	139.48	0.78	2.65
Σ_{14}	0.36	38.25	−0.25	−1.42
Σ_{24}	0.56	24.03	−0.32	−1.05
Σ_{34}	0.20	24.87	0.10	−1.89
δ_1	0.002	0.002	0.003	0.001
δ_2	0.002	0.002	0.002	0.001
δ_3	0.002	0.002	−0.002	−0.003
δ_4	0.004	0.004	−0.005	−0.002
Ω_{11}	0.0002	0.0002	0.0012	−0.0003
Ω_{12}	0.002	0.003	−0.02	−0.01
Ω_{22}	0.03	1.28	−0.16	0.39
Corr_{Ω}	0.003	0.008	−0.002	−0.062

Coverage of confidence intervals

- In setup 1 with correctly specified distribution, GMM 95% confidence intervals have 63% coverage
- In setup 1 with correctly specified distribution, Bayesian 95% credible intervals have 81% coverage

Distributional assumption about ξ

- Why does distributional assumption on ξ not seem to matter?
- Recall that Bayesian OLS with normal distribution $\rightarrow$ frequentist OLS
  - Ignoring priors, gradient of posterior $=$ moment conditions
- Same reasoning implies IV with normal distributions $\rightarrow$ frequentist IV
  - LIML & FIML consistent because gradient of likelihood $=$ moment conditions
- Conjecture: misspecified shape of distribution is fine, but misspecifying heteroskedasticity or dependence could lead to consistent estimates, but inconsistent inference

Allowing heteroskedasticity & dependence

1. Specify distribution of ξ more flexibly
   - E.g. Dirichlet process see Conley et al. (2008)
2. Chernozhukov and Hong (2003): quasi-Bayesian estimation
   - Moments $\mathrm{E}[m_i(\theta)] = 0$, let $g_n(\theta) = \frac{1}{n}\sum_i m_i(\theta)$, $W_n(\theta) =$ consistent estimate of $\lim_{n\to\infty} \mathrm{Var}(\sqrt{n}g_n(\theta))$
   - Quasi-posterior: $\propto \exp(-n/2 g_n(\theta)' W_n(\theta) g_n(\theta))$
   - Bayesian estimation and inference using quasi-posterior is consistent

Zhang (2015)

- Compares Quasi-Bayesian and GMM estimators for BLP demand model
- Uses density tempered sequential Monte Carlo
    - Importance sampling: want sample from π to compute : $\int_\Theta g(\theta)\pi(\theta)d\theta$
        - Draw $\theta_i \sim p(\theta)$
        - $\int_\Theta g(\theta)\pi(\theta)d\theta \approx \frac{1}{S}\sum_{i=1}^S g(\theta_i)\frac{\pi(\theta_i)}{p(\theta_i)}$

      Many draws needed for accuracy if p far from π
    - Density tempered sequential Monte Carlo updates p to get closer to π
- Simulation results a bit incomplete: only shows estimates from a single simulated dataset

Applied papers

- Cohen (2013): vertical supplier relationship's effects on milk prices
- Musalem, Bradlow, and Raju (2008): coupon targeting
- Duan and Mela (2009): pricing and location choice in spatial demand model
- Musalem et al. (2010): effect of out-of-stock

Section 2

Imai, Jain, and Ching (2009)

Imai, Jain, and Ching (2009) – "Bayesian estimation of dynamic discrete choice models"

- Recall likelihood for dynamic discrete choice:

$$\sum_{t=1}^{T} \sum_{i=1}^{N} \log \Lambda \left(a_{it} | v_i^{\mathbf{P}}(\cdot, x_t; \theta) \right)$$

  where $\mathbf{P} = \quad (v^{\mathbf{P}}(\theta))$
- Na ive Metropolis-Hastings:
  - Draw candidate θ
  - Solve for value function
  - Accept or reject with some probability
- Typically infeasible because solving for value function takes too long
- Idea of this paper: combine MCMC iterations with value function iterations
- Each Metropolis step, do one Bellman iteration to update value function

Model

- State = s (observed) & ϵ (unobserved)
- Parameters θ
- Value function:

$$V(s,\epsilon,\theta) = \max_{a \in A} R(s,a,\epsilon_a,\theta) + \beta \mathrm{E}_{\theta_s}[V(s',\epsilon',\theta)|s,a]$$
$$= \max_{a \in A} v(s,a,\epsilon_a,\theta)$$

- Choice probabilities

$$\mathrm{P}[a = a_{i,t}|s_{i,t},V,\theta] = \mathrm{P}[\epsilon : a_{it} = \arg\max v(s,a,\epsilon_a,\theta)]$$

- State transition pmf $f(s'|s,a;\theta_s)$
- Conditional likelihood:

$$L(Y|\theta) = \prod_{i,t} \mathrm{P}[a = a_{i,t}|s_{i,t},V,\theta]$$

Bayesian DP

- θ_s estimated separately? Paper is unclear, but fine
- Iteration t, draw $\theta^{*t} \sim q(\theta^{t-1}, \cdot)$
- At iteration t, have history of draws of $V^\tau, \epsilon^\tau, \theta^{*\tau}$ for $\tau < t$
- Expected value:

$$\hat{\mathrm{E}}^t[V(s', \epsilon', \theta^*)|s, a] = \sum_{s'} f(s'|s, a, \theta) \begin{pmatrix} \sum_{n=1}^{N(t)} V^{t-n}(s', \epsilon^{t-n}, \theta^{*(t-n)}) \times \\ \times \frac{K_h(\theta^* - \theta^{*(t-n)})}{\sum_{k=1}^{N(t)} K_h(\theta^* - \theta^{*(t-k)})} \end{pmatrix}$$

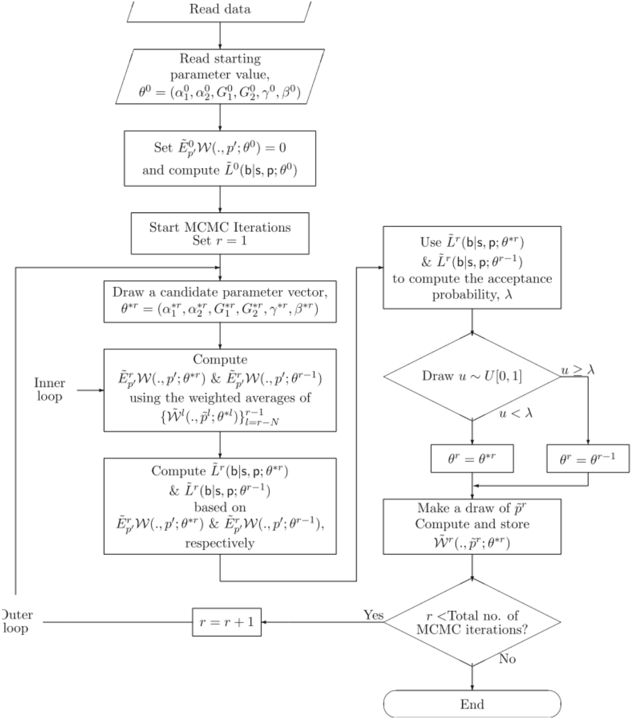
- Choice specific:

$$v(s, a, \epsilon_a, \theta^*) = R(s, a, \epsilon_a, \theta^*) + \beta \hat{\mathrm{E}}^t[V(s', \epsilon', \theta^*)|s, a]$$

- Accept or reject θ^* based on likelihood using $v(s, a, \epsilon_a, \theta^*)$
- Draw $\epsilon^t \sim F(\epsilon; \theta^*_\epsilon)$ calculate and save

$$V^t(s, \epsilon^t, \theta^{*t}) = \max_{a \in A} v(s, a, \epsilon^t_a, \theta^{*t})$$

- Flow chart from Ching et al. (2012) "Practitioner's guide ..."

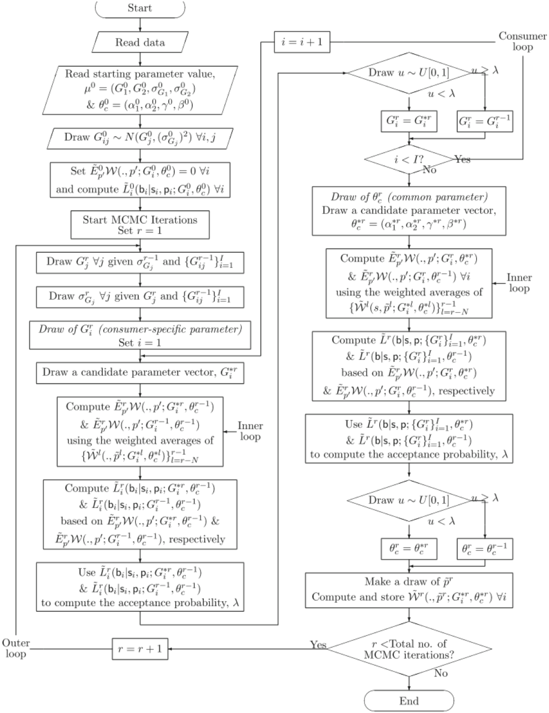

Read data
Read starting parameter value, $\theta^0 = (\alpha_1^0, \alpha_2^0, G_1^0, G_2^0, \gamma^0, \beta^0)$
Set $\tilde{E}^0_{p'}\mathcal{W}(., p'; \theta^0) = 0$ and compute $\tilde{L}^0(\mathsf{b}|\mathsf{s}, \mathsf{p}; \theta^0)$
Start MCMC Iterations Set $r = 1$
Draw a candidate parameter vector, $\theta^{*r} = (\alpha_1^{*r}, \alpha_2^{*r}, G_1^{*r}, G_2^{*r}, \gamma^{*r}, \beta^{*r})$
Inner loop
Compute $\tilde{E}^r_{p'}\mathcal{W}(., p'; \theta^{*r})$ & $\tilde{E}^r_{p'}\mathcal{W}(., p'; \theta^{r-1})$ using the weighted averages of $\{\tilde{\mathcal{W}}^l(., \tilde{p}^l; \theta^{*l})\}_{l=r-N}^{r-1}$
Compute $\tilde{L}^r(\mathsf{b}|\mathsf{s}, \mathsf{p}; \theta^{*r})$ & $\tilde{L}^r(\mathsf{b}|\mathsf{s}, \mathsf{p}; \theta^{r-1})$ based on $\tilde{E}^r_{p'}\mathcal{W}(., p'; \theta^{*r})$ & $\tilde{E}^r_{p'}\mathcal{W}(., p'; \theta^{r-1})$, respectively
Use $\tilde{L}^r(\mathsf{b}|\mathsf{s}, \mathsf{p}; \theta^{*r})$ & $\tilde{L}^r(\mathsf{b}|\mathsf{s}, \mathsf{p}; \theta^{r-1})$ to compute the acceptance probability, λ
Draw $u \sim U[0, 1]$
$u \geq \lambda$
$u < \lambda$
$\theta^r = \theta^{*r}$
$\theta^r = \theta^{r-1}$
Make a draw of $\tilde{p}^r$ Compute and store $\tilde{\mathcal{W}}^r(., \tilde{p}^r; \theta^{*r})$
$r <$Total no. of MCMC iterations?
Yes
No
$r = r + 1$
Outer loop
End

Statistical properties

- Theorem 1: $\hat{\mathrm{E}}^t[V] \xrightarrow{p} \mathrm{E}[V]$ uniformly over s, θ, as $t \to \infty$
- Theorem 2: $\theta^{(t)} \xrightarrow{p} \tilde{\theta}^{(t)}$ where $\tilde{\theta}^{(t)}$ is Markov chain generated by usual Metropolis-Hastingings

Extensions

- Continuous s: draw s^t along with ϵ^t, add importance weights to $\hat{\mathrm{E}}^t$
- Unobserved heterogeneity: Metropolis draws for each i, Gibbs updating for hyperparameters
- Norets (2009): uses nearest-neighbor instead of kernel approximation to V
    - Incorporates serially correlated unobservables
    - Argues more computationally efficient and also applicable to more general model specifications



Start
Read data
Read starting parameter value, $\mu^0 = (G_1^0, G_2^0, \sigma_{G_1}^0, \sigma_{G_2}^0)$ & $\theta_c^0 = (\alpha_1^0, \alpha_2^0, \gamma^0, \beta^0)$
Draw $G_{ij}^0 \sim N(G_j^0, (\sigma_{G_j}^0)^2)\ \forall i, j$
Set $\tilde{E}_{p'}^0 \mathcal{W}(., p'; G_i^0, \theta_c^0) = 0\ \forall i$ and compute $\tilde{L}_i^0(\mathsf{b}_i|\mathsf{s}_i, \mathsf{p}_i; G_i^0, \theta_c^0)\ \forall i$
Start MCMC Iterations Set $r = 1$
Draw $G_j^r\ \forall j$ given $\sigma_{G_j}^{r-1}$ and $\{G_{ij}^{r-1}\}_{i=1}^I$
Draw $\sigma_{G_j}^r\ \forall j$ given G_j^r and $\{G_{ij}^{r-1}\}_{i=1}^I$
Draw of G_i^r (consumer-specific parameter) Set $i = 1$
Draw a candidate parameter vector, G_i^{*r}
Compute $\tilde{E}_{p'}^r \mathcal{W}(., p'; G_i^{*r}, \theta_c^{r-1})$ & $\tilde{E}_{p'}^r \mathcal{W}(., p'; G_i^{r-1}, \theta_c^{r-1})$ using the weighted averages of $\{\tilde{\mathcal{W}}^l(., \tilde{p}^l; G_i^{*l}, \theta_c^{*l})\}_{l=r-N}^{r-1}$
Inner loop
Compute $\tilde{L}_i^r(\mathsf{b}_i|\mathsf{s}_i, \mathsf{p}_i; G_i^{*r}, \theta_c^{r-1})$ & $\tilde{L}_i^r(\mathsf{b}_i|\mathsf{s}_i, \mathsf{p}_i; G_i^{r-1}, \theta_c^{r-1})$ based on $\tilde{E}_{p'}^r \mathcal{W}(., p'; G_i^{*r}, \theta_c^{r-1})$ & $\tilde{E}_{p'}^r \mathcal{W}(., p'; G_i^{r-1}, \theta_c^{r-1})$, respectively
Use $\tilde{L}_i^r(\mathsf{b}_i|\mathsf{s}_i, \mathsf{p}_i; G_i^{*r}, \theta_c^{r-1})$ & $\tilde{L}_i^r(\mathsf{b}_i|\mathsf{s}_i, \mathsf{p}_i; G_i^{r-1}, \theta_c^{r-1})$ to compute the acceptance probability, λ
$i = i + 1$
Consumer loop
Draw $u \sim U[0, 1]$
$u \geq \lambda$
$u < \lambda$
$G_i^r = G_i^{*r}$
$G_i^r = G_i^{r-1}$
$i < I?$
Yes
No
Draw of θ_c^r (common parameter) Draw a candidate parameter vector, $\theta_c^{*r} = (\alpha_1^{*r}, \alpha_2^{*r}, \gamma^{*r}, \beta^{*r})$
Compute $\tilde{E}_{p'}^r \mathcal{W}(., p'; G_i^r, \theta_c^{*r})$ & $\tilde{E}_{p'}^r \mathcal{W}(., p'; G_i^r, \theta_c^{r-1})\ \forall i$ using the weighted averages of $\{\tilde{\mathcal{W}}^l(s, \tilde{p}^l; G_i^{*l}, \theta_c^{*l})\}_{l=r-N}^{r-1}$
Inner loop
Compute $\tilde{L}^r(\mathsf{b}|\mathsf{s}, \mathsf{p}; \{G_i^r\}_{i=1}^I, \theta_c^{*r})$ & $\tilde{L}^r(\mathsf{b}|\mathsf{s}, \mathsf{p}; \{G_i^r\}_{i=1}^I, \theta_c^{r-1})$ based on $\tilde{E}_{p'}^r \mathcal{W}(., p'; G_i^r, \theta_c^{*r})$ & $\tilde{E}_{p'}^r \mathcal{W}(., p'; G_i^r, \theta_c^{r-1})$, respectively
Use $\tilde{L}^r(\mathsf{b}|\mathsf{s}, \mathsf{p}; \{G_i^r\}_{i=1}^I, \theta_c^{*r})$ & $\tilde{L}^r(\mathsf{b}|\mathsf{s}, \mathsf{p}; \{G_i^r\}_{i=1}^I, \theta_c^{r-1})$ to compute the acceptance probability, λ
Draw $u \sim U[0, 1]$
$u \geq \lambda$
$u < \lambda$
$\theta_c^r = \theta_c^{*r}$
$\theta_c^r = \theta_c^{r-1}$
Make a draw of $\tilde{p}^r$ Compute and store $\tilde{\mathcal{W}}^r(., \tilde{p}^r; G_i^{*r}, \theta_c^{*r})\ \forall i$
$r <$Total no. of MCMC iterations?
Yes
No
$r = r + 1$
Outer loop
End

Applied papers

- Ishihara and Ching (2012): dynamic demand
- Toubia and Stephen (2013): contributing to Twitter

Chernozhukov, Victor and Han Hong. 2003. “An {MCMC} approach to classical estimation.” *Journal of Econometrics* 115 (2):293 – 346. URL http://www.sciencedirect.com/science/article/pii/S0304407603001003.

Ching, AndrewT., Susumu Imai, Masakazu Ishihara, and Neelam Jain. 2012. “A practitioner’s guide to Bayesian estimation of discrete choice dynamic programming models.” *Quantitative Marketing and Economics* 10 (2):151–196. URL http://dx.doi.org/10.1007/s11129-012-9119-6.

Cohen, Michael A. 2013. “A Study of Vertical Integration and Vertical Divestiture: The Case of Store Brand Milk Sourcing in Boston.” *Journal of Economics & Management Strategy* 22 (1):101–124. URL http://dx.doi.org/10.1111/jems.12002.

Conley, Timothy G., Christian B. Hansen, Robert E. McCulloch, and Peter E. Rossi. 2008. "A semi-parametric Bayesian approach to the instrumental variable problem." *Journal of Econometrics* 144 (1):276 – 305. URL http://www.sciencedirect.com/science/article/pii/S0304407608000171.

Duan, Jason A. and Carl F. Mela. 2009. "The Role of Spatial Demand on Outlet Location and Pricing." *Journal of Marketing Research* 46 (2):260–278. URL http://www.jstor.org/stable/20618888.

Imai, Susumu, Neelam Jain, and Andrew Ching. 2009. "Bayesian estimation of dynamic discrete choice models." *Econometrica* 77 (6):1865–1899. URL http://onlinelibrary.wiley.com/doi/10.3982/ECTA5658/abstract.

Ishihara, Masakazu and Andrew Ching. 2012. “Dynamic demand for new and used durable goods without physical depreciation: The case of japanese video games.” *Rotman School of Management Working Paper* (2189871). URL http://papers.ssrn.com/sol3/Papers.cfm?abstract_id=2189871.

Jiang, R., P. Manchanda, and P.E. Rossi. 2009. “Bayesian analysis of random coefficient logit models using aggregate data.” *Journal of Econometrics* 149 (2):136–148. URL http://www.sciencedirect.com/science/article/pii/S0304407608002297.

Musalem, Andrés, Eric T. Bradlow, and Jagmohan S. Raju. 2008. “Who’s Got the Coupon? Estimating Consumer Preferences and Coupon Usage from Aggregate Information.” *Journal of Marketing Research* 45 (6):715–730. URL http://www.jstor.org/stable/20618859.

Musalem, Andrés, Marcelo Olivares, Eric T. Bradlow, Christian Terwiesch, and Daniel Corsten. 2010. "Structural Estimation of the Effect of Out-of-Stocks." *Management Science* 56 (7):1180–1197. URL http://dx.doi.org/10.1287/mnsc.1100.1170.

Norets, Andriy. 2009. "Inference in Dynamic Discrete Choice Models With Serially orrelated Unobserved State Variables." *Econometrica* 77 (5):1665–1682. URL http://dx.doi.org/10.3982/ECTA7292.

Toubia, Olivier and Andrew T. Stephen. 2013. "Intrinsic vs. Image-Related Utility in Social Media: Why Do People Contribute Content to Twitter?" *Marketing Science* 32 (3):368–392. URL http://dx.doi.org/10.1287/mksc.2013.0773.

Zhang, Xing. 2015. “A Sequential Monte Carlo Approach to the Random-Coefficient Logit Demand Models.” Tech. rep. URL https://dornsife-usc-edu.ezproxy.library.ubc.ca/assets/sites/524/docs/job_market_cand_2015_2016/JMP_XingZhang_v1.pdf.