A Bayesian Approach to Estimation of Dynamic Models with Small and Large Number of Heterogeneous Players and Latent Serially Correlated States

A. Ronald Gallant
Penn State University

Han Hong
Stanford University

Ahmed Khwaja
Yale University

Paper: http://www.aronaldg.org/papers/socc.pdf
Appendix: http://www.aronaldg.org/papers/socc_web.pdf
Slides: http://www.aronaldg.org/papers/soccclr.pdf

History

Application: Gallant, A. Ronald, Han Hong, and Ahmed Khawaja (2016), "The Dynamic Spillovers of Entry: An Application to the Generic Drug Industry," *Management Science*, forthcoming.

Theory: This paper.

Outline

- Overview
  - Econometric Problem
  - Econometric Approach
  - Results
- Examples
- Econometrics
- Simulation

Econometric Problem

- Estimate a dynamic game
  - ▷ with partially observed state
  - ▷ with serially correlated state
  - ▷ with (possibly) endogenous state
  - ▷ with (possibly) complete information
  - ▷ with continuous or discrete choice
  - ▷ with (mixed) continuous or discrete state
- Applications:
  - ▷ Entry and exit from industry, technology adoption, technology upgrades, introduction of new products, discontinuation of old products, relocation decisions, etc.

Econometric Approach

- Bayesian econometrics
  - ▷ accommodates a nondifferentiable, nonlinear likelihood
  - ▷ easy to parallelize
  - ▷ allows the use of prior information
- Develop a general solution algorithm
  - ▷ computes pure strategy subgame perfect Markov equilibria
  - ▷ using a locally linear value function
- Use sequential importance sampling (particle filter)
  - ▷ to integrate unobserved variables out of the likelihood
  - ▷ to estimate ex-post trajectory of unobserved variables

Results

- Method is exact
  - ▷ Stationary distribution of MCMC chain is the posterior.
  - ▷ Because we prove the computed likelihood is unbiased.
  - ▷ Efficient, number of required particles is small.
- Regularity conditions minimal.

Outline

- Overview
- Examples
  - ▷ An entry game, three players
    - ∗ Equilibrium: Markov sub-game perfect
  - ▷ Monopolistic competition, twenty players
    - ∗ Equilibrium: oblivious
- Abstraction
- Econometrics
- Simulation

Table 1. Generic pharmaceuticals, Scott-Morton (1999)

Drug / Active Ingredient	ANDA Date	Dominant Firms (enter = 1, not enter = 0) Mylan	Novopharm	Lemmon	Geneva	Total Entrants	Revenue ($'000s)
Sulindac	03 Apr. 90	1	0	1	1	7	189010
Erythromycin Stearate	15 May 90	0	0	0	0	1	13997
Atenolol	31 May 90	1	0	0	0	4	69802
Nifedipine	04 Jul. 90	0	1	0	0	5	302983
Minocycline Hydrochloride	14 Aug. 90	0	0	0	0	3	55491
Methotrexate Sodium	15 Oct. 90	1	0	0	0	3	24848
Pyridostigmine Bromide	27 Nov. 90	0	0	0	0	1	2113
Estropipate	27 Feb. 91	0	0	0	0	2	6820
Loperamide Hydrochloride	30 Aug. 91	1	1	1	1	5	31713
Phendimetrazine	30 Oct. 91	0	0	0	0	1	1269
Tolmetin Sodium	27 Nov. 91	1	1	1	1	7	59108
Clemastine Fumarate	31 Jan. 92	0	0	1	0	1	9077
Cinoxacin	28 Feb. 92	0	0	0	0	1	6281
Diltiazem Hydrochloride	30 Mar. 92	1	1	0	0	5	439125
Nortriptyline Hydrochloride	30 Mar. 92	1	0	0	1	3	187683
Triamterene	30 Apr. 92	0	0	0	1	2	22092
Piroxicam	29 May 92	1	1	1	0	9	309756
Griseofulvin Ultramicrocrystalline	30 Jun. 92	0	0	0	0	1	11727
Pyrazinamide	30 Jun. 92	0	0	0	0	1	306
Diflunisal	31 Jul. 92	0	0	1	0	2	96488
Carbidopa	28 Aug. 92	0	0	1	0	4	117233
Pindolol	03 Sep. 92	1	1	0	1	7	37648
Ketoprofen	22 Dec. 92	0	0	0	0	2	107047
Gemfibrozil	25 Jan. 93	1	0	1	0	5	330539
Benzonatate	29 Jan. 93	0	0	0	0	1	2597
Methadone Hydrochloride	15 Apr. 93	0	0	0	0	1	1858
Methazolamide	30 Jun. 93	0	0	0	1	3	4792
Alprazolam	19 Oct. 93	1	1	0	0	7	614593
Nadolol	31 Oct. 93	1	0	0	0	2	125379
Levonorgestrel	13 Dec. 93	0	0	0	0	1	47836
Metoprolol Tartrate	21 Dec. 93	1	1	0	1	9	235625
Naproxen	21 Dec. 93	1	1	1	1	8	456191
Naproxen Sodium	21 Dec. 93	1	1	1	1	7	164771
Guanabenz Acetate	28 Feb. 94	0	0	0	0	2	18120
Triazolam	25 Mar. 94	0	0	0	0	2	71282
Glipizide	10 May 94	1	0	0	0	1	189717
Cimetidine	17 May 94	1	1	0	0	3	547218
Flurbiprofen	20 Jun. 94	1	0	0	0	1	155329
Sulfadiazine	29 Jul. 94	0	0	0	0	1	72
Hydroxychloroquine Sulfate	30 Sep. 94	0	0	0	0	1	8492
Mean		0.45	0.28	0.25	0.25	3.3	126901

Entry Game Characteristics

- Costs
  - ▷ Serially correlated.
  - ▷ Partially observed.
- Endogenous state
  - ▷ Entry changes future costs.
    - ∗ Capacity constraint: increased costs.
    - ∗ Learning: decreased costs.
  - ▷ Induces heterogeneity.
- Complete information
  - ▷ Firms know each other's revenue and costs.
- Simultaneous move dynamic game.

An Entry Game I

- There are $i = 1, \ldots, I$, firms that are identical ex ante.
- Firms maximize PDV of profits over $t, \ldots, \infty$
- Each period t a market opens and firms make entry decisions:
  - ▷ If enter $A_{i,t} = 1$, else $A_{i,t} = 0$.
- Number of firms in the market at time t, is $N_t = \sum_{i=1}^{I} A_{i,t}$.

An Entry Game II

- Gross revenue R_t is exogenously determined.
- A firm's payoff is $R_t/N_t - C_{i,t}$ where $C_{i,t}$ is "cost".
- Costs are endogenous to past entry decisions:
  - ▷ $c_{i,t} = c_{i,u,t} + c_{i,k,t}$ (lower case denotes logs)
  - ▷ $c_{i,u,t} = \mu_c + \rho_c \, (c_{i,u,t-1} - \mu_c) + \sigma_c e_{it}$
  - ▷ $c_{i,k,t} = \rho_a \, c_{i,k,t-1} + \kappa_a A_{i,t-1}$
  - ▷ Source of the dynamics
- Coordination game: If multiple equilibria (rare), the lowest cost firms are the entrants.

Solution I: Bellman Equation

For each player

$$\begin{aligned} & V_i(C_{it}, C_{-i,t}, R_t) \\ & \quad = A^E_{it}\left(R_t/N^E_t - C_{it}\right) \\ & \qquad + \beta\,\mathcal{E}\Big[V_i(C_{i,t+1}, C_{-i,t+1}, R_{t+1}) \mid A^E_{i,t}, A^E_{-i,t}, C_{i,t}, C_{-i,t}, R_t\Big] \end{aligned}$$

The value function for all players is

$$V(C_t, R_t) = \Big(V_1(C_{1t}, C_{-1t}, R_t), \ldots, V_I(C_{It}, C_{-It}, R_t)\Big)$$

– $V(c_t, r_t)$ is approximated by a local linear function.

– The integral is computed by Gauss-Hermite quadrature.

Solution II: Subgame Perfect Markov Equilibrium

Equilibrium condition (Nash)

$$V_i(A^E_{i,t}, A^E_{-i,t}, C_{i,t}, C_{-i,t}, R_t) \geq V_i(A_{i,t}, A^E_{-i,t}, C_{i,t}, C_{-i,t}, R_t) \quad \forall\, i, t.$$

where

$$\begin{aligned} & V_i(A_{i,t}, A_{-i,t}, C_{i,t}, C_{-i,t}, R_t) \\ & \quad = A_{it}\,(R_t/N_t - C_{it}) \\ & \qquad + \beta\mathcal{E}\Big[V_i(A^E_{i,t+1}, A^E_{-i,t+1}, C_{i,t+1}, C_{-i,t+1}, R_{t+1}) | A_{i,t}, A_{-i,t}, C_{i,t}, C_{-i,t}, R_t\Big] \end{aligned}$$

is the choice-specific payoff function.

Complete information: C_t, R_t known implies A^E_t known whence

$$V_i(A^E_{i,t+1}, A^E_{-i,t+1}, C_{i,t+1}, C_{-i,t+1}, R_{t+1}) = V_i(C_{i,t+1}, C_{-i,t+1}, R_{t+1})$$

Solution III: Local Linear Approximation

- The value function V is approximated as follows:
  - ▷ Define a coarse grid on $s = (c_{u,1}, \ldots, c_{u,I}, r, c_{k,1}, \ldots, c_{k,I})$.Each hypercube of the grid is indexed its centroid K, called its key. The local linear approximation over the Kth hypercube is $V_K(s) = b_K + (B_K)s$.
  - ▷ For a three player game V_K is 3×1, b_K is 3×1 , B_K is 3×7, and s is 7×1.

- The local approximator is determined at key K by (1) solving the game at a set $\{s_j\}$ of states within the Kth hypercube, (2) computing $\{V_j = V(s_j)\}$ using the Bellman equation, and (3) computing the coefficients b_K and B_K by regressing $\{V_j\}$ on $\{s_j\}$. Continue until b_K and B_K stabilize.
  - ▷ Usually only 6 hypercubes are visited.

An Entry Game – Summary

- Log revenue: r_t

- Log costs: $c_{i,t} = c_{i,u,t} + c_{i,k,t} \qquad i = 1, \dots, I$
  - ▷ $c_{i,u,t} = \mu_c + \rho_c \, (c_{i,u,t-1} - \mu_c) + \sigma_c e_{it}$
  - ▷ $c_{i,k,t} = \rho_a \, c_{i,k,t-1} + \kappa_a A_{i,t-1}$

- Parameters: $\theta = (\mu_c, \rho_c, \sigma_c, \mu_r, \sigma_r, \rho_a, \kappa_a, \beta, p_a)$

- Solution: $A_t^E = S(c_{u,t}, c_{k,t}, r_t, \theta)$
  - ▷ A deterministic function.

Outcome Uncertainty

- Error density
  - ▷ $p(A_t \mid A_t^E, \theta) = \prod_{i=1}^{I} (p_a)^{\delta(A_{it}=A_{it}^E)} (1 - p_a)^{1-\delta(A_{it}=A_{it}^E)}$
  - ▷ $A_t^E = S(c_{u,t}, c_{k,t}, r_t, \theta)$
- Equilibrium
  - ▷ Firms take outcome uncertainty into account.
  - ▷ Bellman equations modified to include error density.

Abstraction

The state vector is

$$x_t = (x_{1t}, x_{2t}), \tag{1}$$

where x_{1t} is not observed and x_{2t} is observed. The observation (or measurement) density is

$$p(a_t \,|\, x_t, \theta). \tag{2}$$

The transition density is

$$p(x_t \,|\, a_{t-1}, x_{t-1}, \theta). \tag{3}$$

Its marginal is

$$p(x_{1t}|a_{t-1}, x_{t-1}, \theta). \tag{4}$$

The stationary density is

$$p(x_{1t} \,|\, \theta). \tag{5}$$

Assumptions

- We can draw from $p(x_{1t}\,|\,a_{t-1}, x_{t-1}, \theta)$ and $p(x_{1t}\,|\,\theta)$.
  - ▷ Can draw a sample from $p(x_{1t}\,|\,\theta)$ by simulating the game, and discarding a_t and x_{2t}.
  - ▷ Can draw from $p(x_{1,t}\,|\,a_{t-1}, x_{t-1}, \theta)$ by drawing from $p(x_t\,|\,a_{t-1}, x_{t-1}, \theta)$ and discarding x_{2t}.
- There is an analytic expression or algorithm to compute $p(a_t\,|\,x_t, \theta)$, $p(x_t\,|\,a_{t-1}, x_{t-1}, \theta)$, and $p(x_{1t}|a_{t-1}, x_{t-1}, \theta)$.
- If evaluating or drawing from $p(x_{1t}|a_{t-1}, x_{t-1}, \theta)$ is difficult some other importance sampler can be substituted.

Outline

- Overview
- Example
- Econometrics
  - Overview
  - Eliminating unobservables
  - Theory
- Simulation

Estimation Overview

1. In an MCMC loop, propose a parameter value **and a seed**.

2. Given the parameter value **and the seed**, compute an unbiased estimator of the integrated likelihood.
   - Compute by averaging a likelihood that includes latent variables over particles for those latent variables.

3. Use the estimate of the integrated likelihood to make the accept/reject decision of the MCMC algorithm.

Main point:
Deliberately put Monte Carlo jitter into the particle filter.

The Likelihood

- With latent variables

$$L_t(\theta) = \left[\prod_{s=1}^{t} p(a_t \,|\, x_s, \theta)\, p(x_s \,|\, a_{s-1}, x_{s-1}, \theta) \right] p(a_0, x_0 \,|\, \theta)$$

- Without latent variables

$$\mathcal{L}(\theta) = \prod_{t=1}^{T} \int \cdots \int L_t(\theta) \prod_{s=0}^{t} dx_{1,s}$$

- Integrate by averaging sequentially over progressively longer particles. Concatenated draws for fixed k that start at time s and end at time t are denoted

$$\tilde{x}_{1,s:t}^{(k)} = (\tilde{x}_{1,s}^{(k)}, \ldots, \tilde{x}_{1,t}^{(k)});$$

$\tilde{x}_{1,0:t}^{(k)}$ is called a particle.

Particle Filter

1. For $t = 0$

   (a) Start N particles by drawing $\tilde{x}_{1,0}^{(k)}$ from $p(x_{1,0} \,|\, \theta)$ using s as the initial seed and putting $\bar{w}_0^{(k)} = \frac{1}{N}$ for $k = 1, \ldots, N$.

   (b) If $p(a_t, x_{2t} \,|\, x_{1,t-1}, \theta)$ is available, then compute $\widehat{C}_0 = \frac{1}{N} \sum_{k=1}^{N} p(a_0, x_{2,0} \,|\, \tilde{x}_{1,0}^{(k)}, \theta)$ otherwise put $\widehat{C}_0 = 1$.

   (c) Set $x_{1,0:0}^{(k)} = \tilde{x}_{1,0}^{(k)}$.

2. For $t = 1, \ldots, n$

   (a) For each particle, draw $\tilde{x}_{1t}^{(k)}$ from the transition density

   $$p(x_{1t} \mid a_{t-1}, x_{1,t-1}^{(k)}, x_{2,t-1}, \theta).$$

   (b) Compute

   $$\bar{v}_t^{(k)} = \frac{p\left(a_t \mid \tilde{x}_{1,t}^{(k)}, x_{2,t}, \theta\right) p\left(\tilde{x}_{1,t}^{(k)}, x_{2,t} \mid a_{t-1}, x_{1,t-1}^{(k)}, x_{2,t-1}, \theta\right)}{p\left(\tilde{x}_{1,t}^{(k)} \mid a_{t-1}, x_{1,t-1}^{(k)}, x_{2,t-1}, \theta\right)}$$

   $$\hat{C}_t = \frac{1}{N} \sum_{k=1}^{N} \bar{v}_t^{(k)}$$

(c) Set

$$\tilde{x}_{1,0:t}^{(k)} = \left(x_{1,0:t-1}^{(k)}, \tilde{x}_{1,t}^{(k)} \right).$$

(d) Compute the normalized weights

$$\hat{w}_t = \frac{\bar{v}_t^{(k)}}{\sum_{k=1}^{N} \bar{v}_t^{(k)}}$$

(e) For $k = 1, \ldots, N$ draw $x_{1,0:t}^{(k)}$ by sampling with replacement from the set $\{\tilde{x}_{1,0:t}^{(k)}\}$ according to the weights $\{\hat{w}_t^{(k)}\}$.

(f) Note the convention: Particles with unequal weights $\bar{v}_t^{(k)}$ are denoted by $\{\tilde{x}_{0:t}^{(k)}\}$. After resampling the particles have equal weights $\frac{1}{N}$ and are denoted by $\{x_{0:t}^{(k)}\}$.

3. Done

   (a) An unbiased estimate of the likelihood is

$$\ell' = \prod_{t=0}^{T} \widehat{C}_t$$

   and s' is the last seed returned in Step 2e.

Why Does This Work?

- For each particle, draw $\tilde{x}_{1t}^{(k)}$ from the transition density

$$p(x_{1t} \,|\, a_{t-1}, x_{1,t-1}^{(k)}, x_{2,t-1}, \theta).$$

- Compute

$$\begin{aligned}
\bar{v}_t^{(k)} &= \frac{p\left(a_t \,|\, \tilde{x}_{1,t}^{(k)}, x_{2,t}, \theta\right) p\left(\tilde{x}_{1,t}^{(k)}, x_{2,t} \,|\, a_{t-1}, x_{1,t-1}^{(k)}, x_{2,t-1}, \theta\right)}{p\left(\tilde{x}_{1,t}^{(k)} \,|\, a_{t-1}, x_{1,t-1}^{(k)}, x_{2,t-1}, \theta\right)} \\
\widehat{C}_t &= \frac{1}{N} \sum_{k=1}^{N} \bar{v}_t^{(k)}
\end{aligned}$$

- An unbiased estimate of the likelihood is

$$\ell(\theta, s) = \prod_{t=0}^{T} \widehat{C}_t$$

Verification is Remarkably Simple

- Theorem 1 establishs a recursion using Bayes theorem. The idea is straightforward and is expressed as one four line equation. The remainder of the proof is algebra to reduce the basic expression to model primitives.

- Corollary 1 establish unbiasedness via a simple two line telescoping expression.

- Theorem 2 shows that resampling is a mere footnote requiring only three sentences to dismiss.

Verification Requires Some Notation

- In the Bayesian paradigm, θ and $\{a_t, x_t\}_{t=-\infty}^{\infty}$ are defined on a common probability space. Let $\mathcal{F}_t = \sigma\left\{\{a_s, x_{2s}\}_{s=-T_0}^{t}, \theta\right\}$.

- Particle filters are implemented by drawing independent uniform random variables $u_t^{(k)}$ and then evaluating a function of the form $X_{1t}(u)$ and putting $\tilde{x}_{1t}^{(k)} = X_{1t}(u_t^{(k)})$ for $k = 1, \dots, N$.
  - ▷ Let $\tilde{\mathcal{E}}_{1t}$ denote integration with respect to $\left(u_t^{(1)}, \dots, u_t^{(N)}\right)$ with $\tilde{x}_{1t}^{(k)} = X_{1t}(u_t^{(k)})$ substituted into the integrand.
  - ▷ $\tilde{\mathcal{E}}_{1,0:t}$ is defined similarly.

- Unbiasedness is a corollary of the following result.

THEOREM 1 If particles $\tilde{x}_{1,0:t}^{(k)}$ and weights $\tilde{w}_t^{(k)}$, $k = 1, \ldots, N$, satisfy

$$\int g(x_{1,0:t})\, dP(x_{1,0:t}|\mathcal{F}_t) = \tilde{\mathcal{E}}_{1,0:t} \left\{ \mathcal{E} \left[\sum_{k=1}^{N} \tilde{w}_t^{(k)} g(\tilde{x}_{1,0:t}^{(k)}) \mid \mathcal{F}_t \right] \right\} \tag{6}$$

then draws $\tilde{x}_{1,t+1}^{(k)}$ from $p(x_{1,t+1}|\tilde{x}_{1,0:t}^{(k)}, \mathcal{F}_t)$ and weights

$$\tilde{w}_{t+1}^{(k)} = \frac{\bar{v}_{t+1}^{(k)}}{C_{t+1}}\, \tilde{w}_t^{(k)} \tag{7}$$

satisfy

$$\begin{aligned} &\int g(x_{1,0:t}, x_{1,t+1})\, dP(x_{1,0:t}, x_{1,t+1}|\mathcal{F}_{t+1}) \\ &\quad = \tilde{\mathcal{E}}_{1,t+1}\tilde{\mathcal{E}}_{1,0:t} \left\{ \mathcal{E} \left[\sum_{k=1}^{N} \tilde{w}_{t+1}^{(k)}\, g(\tilde{x}_{1,0:t}^{(k)}, \tilde{x}_{1,t+1}^{(k)}) \mid \mathcal{F}_{t+1} \right] \right\}, \end{aligned} \tag{8}$$

where

$$\bar{v}_{t+1}^{(k)} = \frac{p\Big(a_{t+1} \mid \tilde{x}_{1,t+1}^{(k)}, x_{2,t+1}, \theta\Big)\, p\Big(\tilde{x}_{1,t+1}^{(k)}, x_{2,t+1} \mid a_t, \tilde{x}_{1,t}^{(k)}, x_{2,t}, \theta\Big)}{p\Big(\tilde{x}_{1,t+1}^{(k)} \mid a_t, \tilde{x}_{1,t}^{(k)}, x_{2,t}, \theta\Big)} \tag{9}$$

and

$$C_{t+1} = p(a_{t+1}, x_{2,t+1}|\mathcal{F}_t). \tag{10}$$

Proof We show the result for the weights

$$\tilde{w}_{t+1}^{(k)} = \frac{p(a_{t+1}, x_{2,t+1}|\tilde{x}_{1,0:t}^{(k)}, \tilde{x}_{1,t+1}^{(k)}, \mathcal{F}_t)}{p(a_{t+1}, x_{2,t+1}|\mathcal{F}_t)} \tilde{w}_t^{(k)}, \tag{11}$$

then show that (11) and (7) are equivalent expressions for $\tilde{w}_{t+1}^{(k)}$.

Bayes theorem states that

$$p(x_{1,0:t}, x_{1,t+1}|a_{t+1}, x_{2,t+1}, \mathcal{F}_t) = \frac{p(a_{t+1}, x_{2,t+1}, x_{1,0:t}, x_{1,t+1}|\mathcal{F}_t)}{p(a_{t+1}, x_{2,t+1}|\mathcal{F}_t)}. \tag{12}$$

Note that

$$p(x_{1,0:t}, x_{1,t+1}|a_{t+1}, x_{2,t+1}, \mathcal{F}_t) = p(x_{1,0:t}, x_{1,t+1}|\mathcal{F}_{t+1}) \tag{13}$$

and that

$$\begin{aligned} & p(a_{t+1}, x_{2,t+1}, x_{1,0:t}, x_{1,t+1}|\mathcal{F}_t) \\ = \;\; & p(a_{t+1}, x_{2,t+1}|x_{1,0:t}, x_{1,t+1}, \mathcal{F}_t) p(x_{1,t+1}|x_{1,0:t}, \mathcal{F}_t) p(x_{1,0:t}|\mathcal{F}_t). \end{aligned} \tag{14}$$

Then

$$
\begin{aligned}
&\int g(x_{1,0:t}, x_{1,t+1})\, dP(x_{1,0:t}, x_{1,t+1}|\mathcal{F}_{t+1}) \\
&\quad = \iint g(x_{1,0:t}, x_{1,t+1}) \frac{p(a_{t+1}, x_{2,t+1}|x_{1,0:t}, x_{1,t+1}, \mathcal{F}_t)}{p(a_{t+1}, x_{2,t+1}|\mathcal{F}_t)} p(x_{1,t+1}|x_{1,0:t}, \mathcal{F}_t) \\
&\qquad \times dx_{1,t+1} dP(x_{1,0:t}|\mathcal{F}_t) \qquad (15) \\
&\quad = \tilde{\mathcal{E}}_{1,0:t} \int \mathcal{E}\left[\sum_{k=1}^{N} g(\tilde{x}^{(k)}_{1,0:t}, x_{1,t+1})\, w^{(k)}_{t+1}\, p(x_{1,t+1}|\tilde{x}^{(k)}_{1,0:t}, \mathcal{F}_t) \mid \mathcal{F}_t\right] dx_{1,t+1} \qquad (16) \\
&\quad = \tilde{\mathcal{E}}_{1,t+1}\tilde{\mathcal{E}}_{1,0:t}\, \mathcal{E}\left[\sum_{k=1}^{N} g(\tilde{x}^{(k)}_{1,0:t}, \tilde{x}^{(k)}_{1,t+1})\, \tilde{w}^{(k)}_{t+1} \mid \mathcal{F}_{t+1}\right] \qquad (17)
\end{aligned}
$$

where (15) is due to (12) after substituting (13) and (14), (16) is due to (6) and (11), and (17) is due to the fact that $\tilde{x}^{(k)}_{1,t+1}$ is a draw from $p(x_{1,t+1}|\tilde{x}^{(k)}_{1,0:t}, \mathcal{F}_t)$. This proves the result for the weights (11).

Showing (11) and (7) are equivalent expressions for $\tilde{w}^{(k)}_{t+1}$ is just algebra.

COROLLARY 1 If one starts the recursion of Theorem 1 with draws from the marginal stationary density (5) and weights $\tilde{w}_0^{(k)} = 1/N$, then

$$\widehat{\ell}' = \left(\sum_{k=1}^{N} \bar{v}_T^{(k)} \frac{\tilde{w}_{T-1}^{(k)}}{\sum_{k=1}^{N} \tilde{w}_{T-1}^{(k)}}\right)\left(\sum_{k=1}^{N} \bar{v}_{T-1}^{(k)} \frac{\tilde{w}_{T-2}^{(k)}}{\sum_{k=1}^{N} \tilde{w}_{T-2}^{(k)}}\right)\cdots\left(\sum_{k=1}^{N} \bar{v}_1^{(k)} \frac{\tilde{w}_0^{(k)}}{\sum_{k=1}^{N} \tilde{w}_0^{(k)}}\right)\left(\sum_{k=1}^{N} \tilde{w}_0^{(k)}\right) \tag{18}$$

is an unbiased estimator of ℓ'.

Proof Set $g(x_{1,0:t}, u) \equiv 1$ in Theorem 1 whence $1 = \tilde{\mathcal{E}}_{1,0:T}\left\{\mathcal{E}\left[\sum_{k=1}^{N} \tilde{w}_t^{(k)} \mid \mathcal{F}_T\right]\right\}$. Write

$$\sum_{k=1}^{N} \tilde{w}_T^{(k)} = \frac{1}{C_T}\left(\frac{\sum_{k=1}^{N} \bar{v}_T^{(k)} \tilde{w}_{T-1}^{(k)}}{\sum_{k=1}^{N} \bar{v}_{T-1}^{(k)} \tilde{w}_{T-2}^{(k)}}\right)\left(\frac{\sum_{k=1}^{N} \bar{v}_{T-1}^{(k)} \tilde{w}_{T-2}^{(k)}}{\sum_{k=1}^{N} \bar{v}_{T-2}^{(k)} \tilde{w}_{T-3}^{(k)}}\right)\cdots\left(\frac{\sum_{k=1}^{N} \bar{v}_1^{(k)} \tilde{w}_0^{(k)}}{\sum_{k=1}^{N} \tilde{w}_0^{(k)}}\right)\left(\sum_{k=1}^{N} \tilde{w}_0^{(k)}\right)$$

$$= \frac{1}{\ell'}\left(\sum_{k=1}^{N} \bar{v}_T^{(k)} \frac{\tilde{w}_{T-1}^{(k)}}{\sum_{k=1}^{N} \tilde{w}_{T-1}^{(k)}}\right)\left(\sum_{k=1}^{N} \bar{v}_{T-1}^{(k)} \frac{\tilde{w}_{T-2}^{(k)}}{\sum_{k=1}^{N} \tilde{w}_{T-2}^{(k)}}\right)\cdots\left(\sum_{k=1}^{N} \bar{v}_1^{(k)} \frac{\tilde{w}_0^{(k)}}{\sum_{k=1}^{N} \tilde{w}_0^{(k)}}\right)\left(\sum_{k=1}^{N} \tilde{w}_0^{(k)}\right)$$

The result follows.

THEOREM 2 Theorem 1 and Corollary 1 remain valid if resampling is applied between recursive steps.

Proof If a set of particles and weights satisfy condition (6) then so will the particles and weights generated from them by resampling. Because a set of particles and weights satisfy condition (6) at the end of an iterate, the set of particles and weights generated from them by resampling will satisfy (6) when used at the beginning of an iterate. The only formal change to the development required is that $\tilde{\mathcal{E}}_{1,0:t}$ becomes expectation both with respect to the uniform draws that advance the filter and to the uniform draws involved in resampling.

REMARK 1 For any resampling scheme that produces equal weights, the conclusion of Corollary 1 becomes

$$\widehat{\ell}' = \left(\frac{1}{N}\sum_{k=1}^{N}\bar{v}_T^{(k)}\right)\left(\frac{1}{N}\sum_{k=1}^{N}\bar{v}_{T-1}^{(k)}\right)\cdots\left(\frac{1}{N}\sum_{k=1}^{N}\bar{v}_1^{(k)}\right)$$

is an unbiased estimator of ℓ'.

Outline

- Overview
- Example
- Econometrics
- Simulation
  - ▷ Small Entry Game Design
  - ▷ Small Entry Game Results
  - ▷ Large Game Design
  - ▷ Large Game Results

Entry Game Design – 1

- Three firms, time increment one year.
  - ▷ β is 20% internal rate of return
  - ▷ μ_c and μ_r imply 30% profit margin, persistent ρ_c
  - ▷ κ_a is a 20% hit to margin with ρ_a at 6 mo. half life.
  - ▷ σ_c and σ_r chosen to prevent monopoly
  - ▷ Outcome uncertainty $1 - p_a$ is 5% (from an application).
- Simulated with

$$
\begin{aligned}
\theta &= (\mu_c, \rho_c, \sigma_c, \mu_r, \sigma_r, \rho_a, \kappa_a, \beta, p_a) \\
&= (9.7, 0.9, 0.1, 10.0, 2.0, 0.5, 0.2, 0.83, 0.95) \\
T_0 &= 160,\ \text{sm} : T = 40,\ \text{md} : T = 120,\ \text{lg} : T = 360
\end{aligned}
$$

Entry Game Design – 2

1. Fit with blind importance sampler, and multinomial resampling.

2. Fit with adaptive importance sampler, and multinomial resampling.

3. Fit with adaptive importance sampler, and systematic resampling.

Results – 1

- A large sample size is better. In Tables 2 through 4 the estimates shown in the columns labeled ”lg” would not give misleading results in an application.

Results – 2

- Constraining β is beneficial: compare Figures 1 and 2. The constraint reduces the bimodality of the marginal posterior distribution of σ_r and pushes all histograms closer to uni-modality.

- Constraining p_a is irrelevant except for a small savings in computational cost: compare columns "β" and "β & p_a" in Tables 2 through 4.

Results – 3

- Improvements to the particle filter are helpful. In particular, an adaptive importance sampler is better than a blind importance sampler; compare Tables 2 and 3 and compare Figures 3 and 4. Systematic resampling is better than multinomial resampling; compare Tables 3 and 4.

Table 2. Blind Sampler, Multinomial Resampling

Parameter		Unconstrained			Constrained					
					β			β & p_a		
	value	sm	md	lg	sm	md	lg	sm	md	lg
μ_c	9.70	10.10	9.72	9.68	9.94	9.67	9.68	9.86	9.72	9.68
		(0.15)	(0.12)	(0.06)	(0.19)	(0.11)	(0.06)	(0.18)	(0.12)	(0.06)
ρ_c	0.90	0.58	0.86	0.92	0.69	0.92	0.91	0.69	0.85	0.91
		(0.25)	(0.09)	(0.03)	(0.26)	(0.05)	(0.03)	(0.25)	(0.11)	(0.03)
σ_c	0.10	0.16	0.09	0.09	0.17	0.08	0.10	0.15	0.09	0.10
		(0.05)	(0.03)	(0.01)	(0.06)	(0.03)	(0.01)	(0.07)	(0.03)	(0.01)
μ_r	10.00	9.87	9.98	9.96	9.88	9.99	9.98	9.84	9.99	9.99
		(0.10)	(0.03)	(0.02)	(0.10)	(0.03)	(0.02)	(0.13)	(0.06)	(0.02)
σ_r	2.00	1.95	1.97	1.98	2.02	2.00	2.02	2.04	2.00	2.03
		(0.09)	(0.05)	(0.01)	(0.08)	(0.02)	(0.02)	(0.10)	(0.03)	(0.01)
ρ_a	0.50	0.76	0.56	0.58	0.59	0.57	0.56	0.76	0.57	0.52
		(0.09)	(0.07)	(0.06)	(0.22)	(0.09)	(0.05)	(0.10)	(0.07)	(0.04)
κ_a	0.20	0.04	0.24	0.19	0.15	0.26	0.20	0.14	0.22	0.22
		(0.05)	(0.05)	(0.02)	(0.07)	(0.05)	(0.03)	(0.06)	(0.06)	(0.03)
β	0.83	0.90	0.95	0.87	0.83	0.83	0.83	0.83	0.83	0.83
		(0.07)	(0.04)	(0.04)						
p_a	0.95	0.97	0.94	0.95	0.96	0.94	0.95	0.95	0.95	0.95
		(0.02)	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)			

Table 3. Adaptive Sampler, Multinomial Resampling

Parameter	value	Unconstrained sm	Unconstrained md	Unconstrained lg	Constrained β sm	Constrained β md	Constrained β lg	Constrained β & p_a sm	Constrained β & p_a md	Constrained β & p_a lg
μ_c	9.70	10.00	9.82	9.77	9.93	9.74	9.70	9.85	9.73	9.65
		(0.24)	(0.07)	(0.05)	(0.12)	(0.07)	(0.06)	(0.15)	(0.09)	(0.05)
ρ_c	0.90	0.95	0.85	0.87	0.87	0.92	0.93	0.87	0.92	0.94
		(0.03)	(0.07)	(0.05)	(0.08)	(0.04)	(0.03)	(0.09)	(0.04)	(0.02)
σ_c	0.10	0.14	0.09	0.10	0.12	0.08	0.08	0.12	0.09	0.08
		(0.02)	(0.02)	(0.01)	(0.04)	(0.02)	(0.01)	(0.04)	(0.03)	(0.01)
μ_r	10.00	9.93	10.00	10.01	10.00	9.99	9.97	9.94	9.96	9.96
		(0.06)	(0.02)	(0.01)	(0.05)	(0.02)	(0.02)	(0.07)	(0.03)	(0.03)
σ_r	2.00	1.93	1.98	1.99	2.01	1.98	2.00	2.03	1.97	1.99
		(0.10)	(0.02)	(0.02)	(0.09)	(0.01)	(0.01)	(0.09)	(0.02)	(0.02)
ρ_a	0.50	-0.11	0.51	0.47	0.56	0.59	0.57	0.47	0.51	0.61
		(0.21)	(0.09)	(0.06)	(0.17)	(0.06)	(0.06)	(0.20)	(0.07)	(0.05)
κ_a	0.20	0.19	0.20	0.17	0.17	0.21	0.18	0.24	0.20	0.19
		(0.02)	(0.03)	(0.02)	(0.06)	(0.02)	(0.02)	(0.03)	(0.02)	(0.02)
β	0.83	0.87	0.95	0.92	0.83	0.83	0.83	0.83	0.83	0.83
		(0.10)	(0.03)	(0.04)						
p_a	0.95	0.95	0.94	0.95	0.96	0.95	0.95	0.95	0.95	0.95
		(0.01)	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)			

Table 4. Adaptive Sampler, Systematic Resampling

Parameter		Unconstrained			Constrained					
					β			β & p_a		
	value	sm	md	lg	sm	md	lg	sm	md	lg
μ_c	9.70	9.87	9.82	9.72	9.81	9.78	9.68	9.78	9.76	9.65
		(0.24)	(0.07)	(0.05)	(0.12)	(0.07)	(0.06)	(0.15)	(0.09)	(0.05)
ρ_c	0.90	0.77	0.82	0.91	0.93	0.94	0.94	0.86	0.92	0.94
		(0.03)	(0.07)	(0.05)	(0.08)	(0.04)	(0.03)	(0.09)	(0.04)	(0.02)
σ_c	0.10	0.14	0.10	0.09	0.14	0.08	0.08	0.11	0.08	0.08
		(0.02)	(0.02)	(0.01)	(0.04)	(0.02)	(0.01)	(0.04)	(0.03)	(0.01)
μ_r	10.00	10.05	10.00	9.97	9.95	9.96	9.94	9.78	9.95	9.96
		(0.06)	(0.02)	(0.01)	(0.05)	(0.02)	(0.02)	(0.07)	(0.03)	(0.03)
σ_r	2.00	1.94	1.99	1.99	1.93	1.97	2.01	2.07	1.98	1.97
		(0.10)	(0.02)	(0.02)	(0.09)	(0.01)	(0.01)	(0.09)	(0.02)	(0.02)
ρ_a	0.50	0.61	0.53	0.56	0.41	0.36	0.61	0.71	0.58	0.64
		(0.21)	(0.09)	(0.06)	(0.17)	(0.06)	(0.06)	(0.20)	(0.07)	(0.05)
κ_a	0.20	0.21	0.22	0.18	0.20	0.18	0.18	0.17	0.19	0.18
		(0.02)	(0.03)	(0.02)	(0.06)	(0.02)	(0.02)	(0.03)	(0.02)	(0.02)
β	0.83	0.93	0.96	0.90	0.83	0.83	0.83	0.83	0.83	0.83
		(0.10)	(0.03)	(0.04)						
p_a	0.95	0.96	0.94	0.95	0.95	0.93	0.95	0.95	0.95	0.95
		(0.01)	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)			

Figure 1. Posterior Distributions, Unconstrained, Blind Sampler, Md.

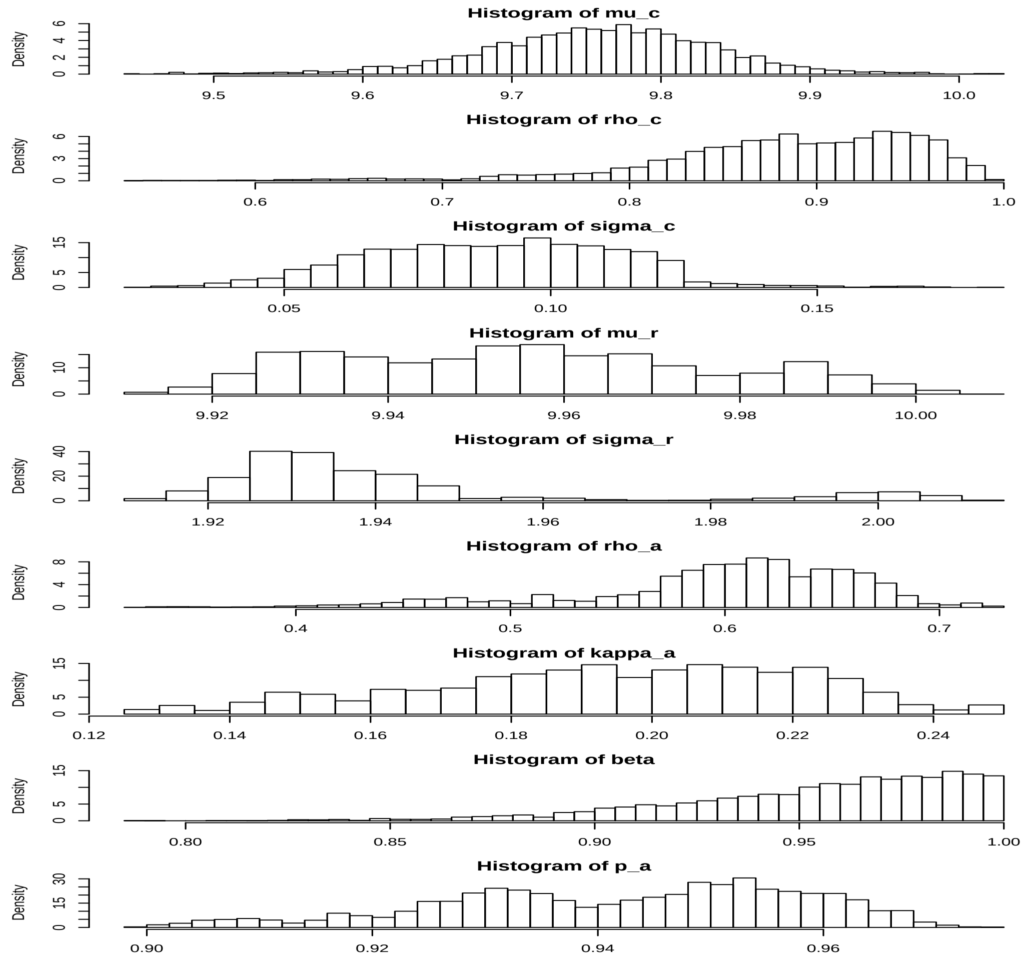


Figure 2. Posterior Distributions, β Constrained, Blind Sampler, Md.

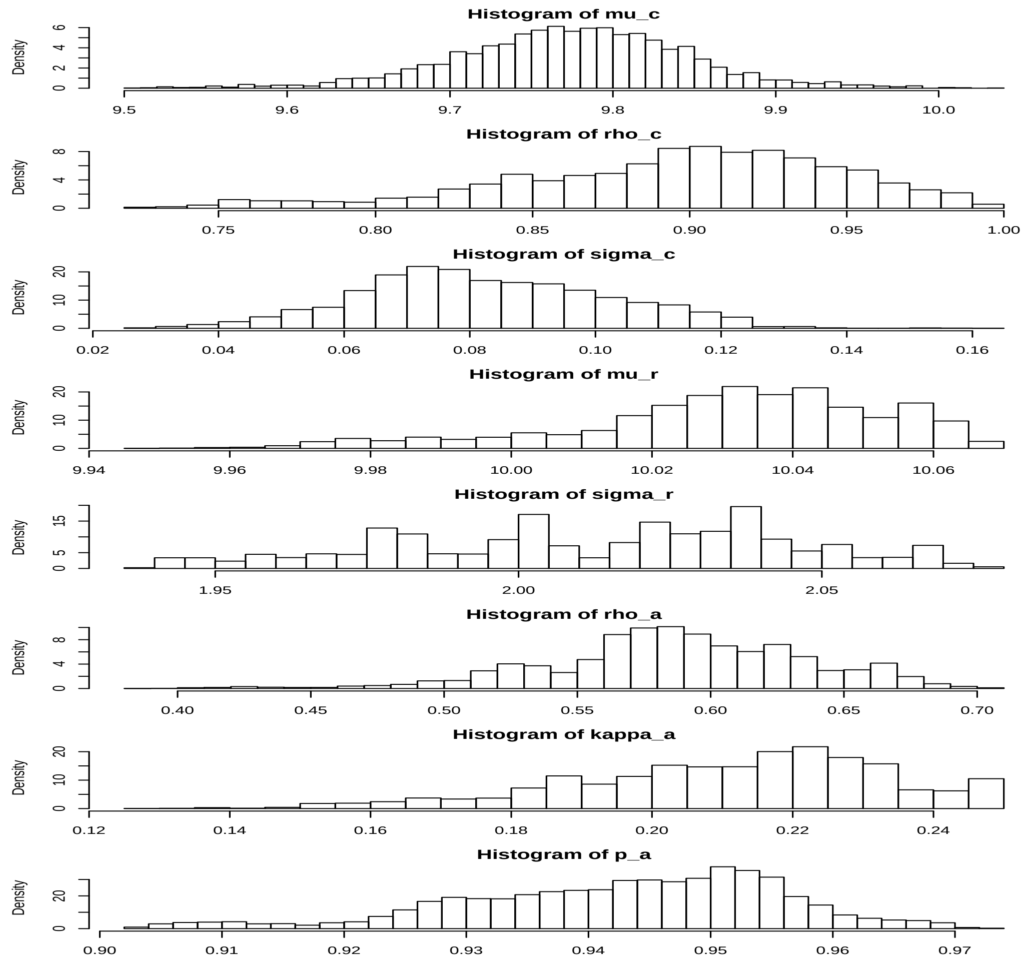
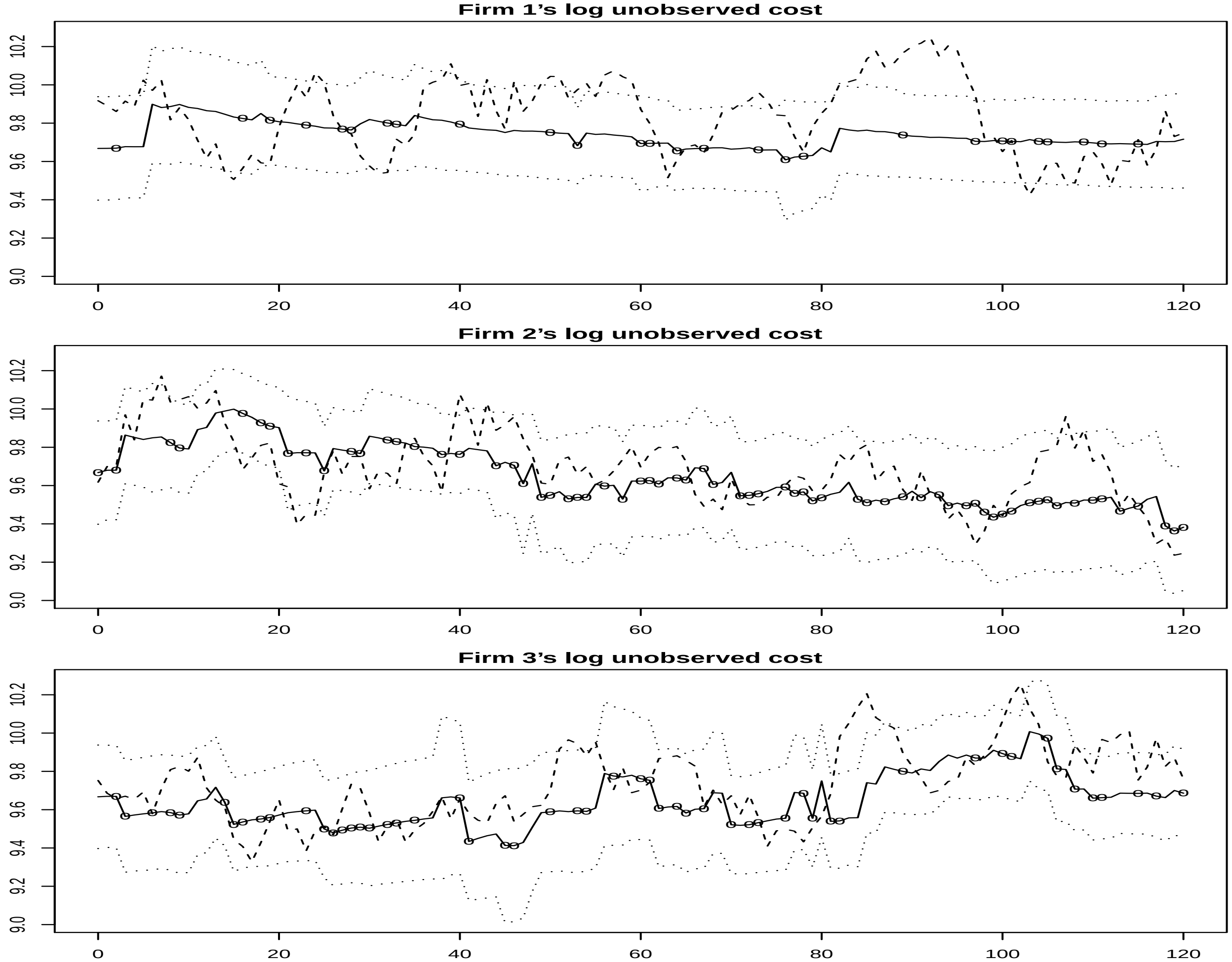
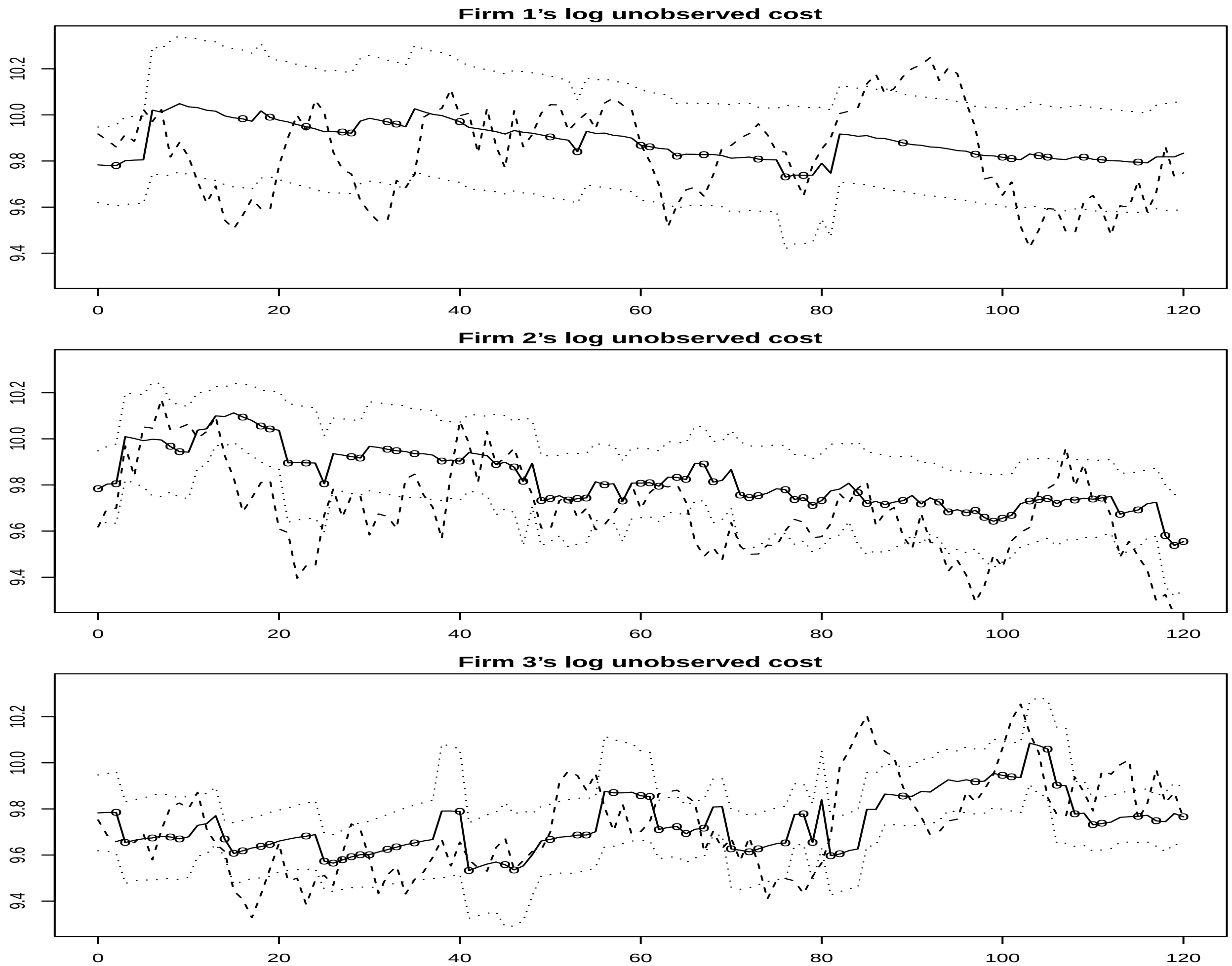


Figure 3. Posterior Cost Estimates, β Constrained, Blind Sampler, Md.



Circles indicate entry. Dashed line is true unobserved cost. The solid line is the average of β constrained estimates over all MCMC repetitions, with a stride of 25. The dotted line is $\pm$ 1.96 standard deviations about solid line. The sum of the norms of the difference between the solid and dashed lines is 0.186146.

Figure 4. Posterior Cost Estimates, β Constrained, Adaptive Sampler, Md.



Circles indicate entry. Dashed line is true unobserved cost. The solid line is the average of β constrained estimates over all MCMC repetitions, with a stride of 25. The dotted line is $\pm$ 1.96 standard deviations about solid line. The sum of the norms of the difference between the solid and dashed lines is 0.169411.

Large Number of Players Design – 1

- Oblivious equilibrium: Weintraub, Benkard, and Roy (2008)
  - ▷ Logit utility $u_{ijt} = \theta_1 \ln\left(\frac{x_{it}}{\psi} + 1\right) + \theta_2 \ln\left(Y - p_{it}\right) + v_{ijt}$,
  - ▷ Investment strategy $\iota_{it} = \iota(x_{it}, s_{-i,t})$ that increases quality one level with probability $\frac{a\iota}{1+a\iota}$
  - ▷ Quality depreciates by one level with probability δ.
  - ▷ x is product quality, Y income, p price, $s = x$ state, ijt indexes firm, consumer, time.
  - ▷ Many other details.
- We estimate utility and transition dynamics $\theta = (\theta_1, \theta_2, \psi, a, \delta)$.
- All else the same as in the Matlab code on the authors' website.

Large Number of Players Design – 2

- Unique equilibrium p_{it}^* yielding a multinomial for number of customers attracted by firm and a transition matrix for the state.

- Customers are the observable, the state is the unobservable.

  ▷ 50 customers

  ▷ 20 firms, hence 20 dimensional state

  ▷ 5 time periods

- Prior has positive support conditions, otherwise uninformative.

Table 5. Large Game, Blind Sampler, Stratified Resampling

		Posterior	
Parameter	Value	Mean	Std. Dev.
θ_1	1.00000	0.97581	0.04799
θ_2	0.50000	0.53576	0.07317
ψ	1.00000	1.01426	0.07070
a	3.00000	2.96310	0.06846
δ	0.70000	0.64416	0.05814

The data were generated according to the oblivious equilibrium model with parameters for the consumer's utility function and firm's transition function set as shown in the column labeled "Value" and all others set to the author's calibrated values. The number of firms is 20 and the number of consumers is 50. $T = 5$. The prior is uninformative except for a support conditon that all values be positive. The number of MCMC repetitions is 109,000 and the number of particles per repetition is 32696.