My understanding of BLP

①

1. Product is bundle of attrib
2. Problem of unobservable attrib (called demand shock)
3. Instrument needed for indogeneity
4. ξ_{it} structural error of unobserved but can be estimated
5. data type: (1) product choice (2) product charact. (Dummy: e.g. cereal) (3) demographic dist.
6. all pain due to (1) unobserved prod. attrib (endogeneity) (2) Curse of dimensionality (dependance)
7. modeling dynamic & waiting ⇒ by discount
8. modeling strategic behavior ⇒ by Dynamic Prog.
9. modeling uncertainty game (type) ⇒ by { (1) signaling (2) Bayesian learning
10. modeling search game ⇒ sequential search method
11. modeling decisions & demand (struct) ⇒ logit
12. modeling optimization ⇒ FOC. Game theory
    (1) info about other player
    (2) marginal valuation & dist
    (3) the sequence of optimizing
    (4) act of single agent
13. modeling cooperative game ⇒ Core & deviation
14. modeling uncertainty in action ⇒ (1) distribution (2) indifference & deviation
15. modeling behavioral: (1) trust (2) fairness (3) prospect theory
16. modeling match game ⇒ (1) Best match & deviation (2) guess equil.
17. Bays estimator is more efficient { (1) you need to run multiple times (2) limit parameter space)
18. when you expect something about future and you don't know it today ⇒ you form Expectation (dist. of outcome as weight)
19. Research in empirical:
    (1) Theory, game theory, or theoretical framework (literature)
    (2) Hypothesis & writing
    (3) Data cleaning skills
    (4) Data gathering skills
    (5) Estimation & optimization skills (math, computer, Bayesian, Frequentist GMM) ⇒ Data structure & algorithms
    (6) writing
    (7) Smart targeting & positioning

②

Nevo Code | BLP ①

$ns = 20$; (simulated indiv)
$nmkt = 94$; (# markets)
$nbrn = 24$; (# brands per market)
$n\_inst = 20$; (# inst per Price)

kron : Kronecker tensor product of X, Y

$$A \otimes B = \begin{bmatrix} a_{11}B & \cdots & a_{1n}B \\ \vdots & \ddots & \vdots \\ a_{m1}B & \cdots & a_{mn}B \end{bmatrix}$$

IV, S_{jt}, X_1, X_2 : $\overbrace{[\quad]}^{n\_inst + nbrn = 44}$

nmkt x nbrn → 24 x 94

cdid: $\begin{bmatrix} 1 \\ \vdots \\ 1 \\ 2 \\ \vdots \\ 94 \\ 94 \end{bmatrix}$ } nbrn : 24 $[1 : nmkt]^{94} \otimes \begin{bmatrix} 1 \\ \vdots \end{bmatrix}$ } nbrn = 24

94 = nmkt

cdindex $\begin{bmatrix} 24 \\ 48 \\ \vdots \\ 24 \times 94 \end{bmatrix}$ → nbrn

2256 = 24x94 → [nbrn : nbrn : $\overbrace{nbrn}^{24}$ x $\overbrace{nmkt}^{94}$] last obs for each brand

$$\theta_w = \begin{array}{c|cccc|l} & \text{income} & \text{income}^2 & \text{age} & \text{child} & \\ \sigma_1^2 & \square & 0 & \square & 0 & \text{Constant} \\ \sigma_2^2 & \square & \square & 0 & \square & \text{Price} \\ \sigma_3^2 & \square & 0 & 0 & 0 & \text{Sugar} \\ \sigma_4^2 & \square & 0 & 0 & 0 & \text{mushy} \end{array}$$

σ_{1-4} = variance of Constant, price, sugar, mushy

$invA = inv([IV' \times IV]) = \Phi^{-1}$

- cumsum(A): cumulative sum of matrix (vertical)
- sum1 = temp(cdindex,:); will take only the last element per mkt
- diff(A) : Calculates difference b/w adjacent cell in MATLAB
- outsh : outside goods share
- Error for simple pois: $y = \log(S_{jt}) - \log(\text{outshare})$
- Error for simple pois & β_{pois} $X \times 1$
- mean old : mvalold by running regression Pois
- V : 80 iid. r.n. for 94 obs (mkt)
- vfull = v(cdid,:) Copies each element of v, based on index of cdid
  in this step both $\underset{\sim}{D}$ & $\underset{w}{V}$ Copied for all brands in each mkt
- simplex algorithm for search Linear programming

→ gmmobj optimization
  var-cov for std err

Anonymous function

- sparse : creates sparse matrix
  Full : Converts sparse matrix to full matrix

relation b/w Hessian matrix and Jacobian : $H(f)(x) = J(\nabla f)(x)$

$$\boxed{y = f(x + \Delta x) \approx f(x) + J(x)\Delta x + \tfrac{1}{2}\Delta x^T H(x) \Delta x}$$

Second partial derivative

↳ Jacobian (generalization of gradient of scalar matrix) of multiple var)

Jacobian: amount of stretching, rotating, transformation (how image of neighborhood transforms)

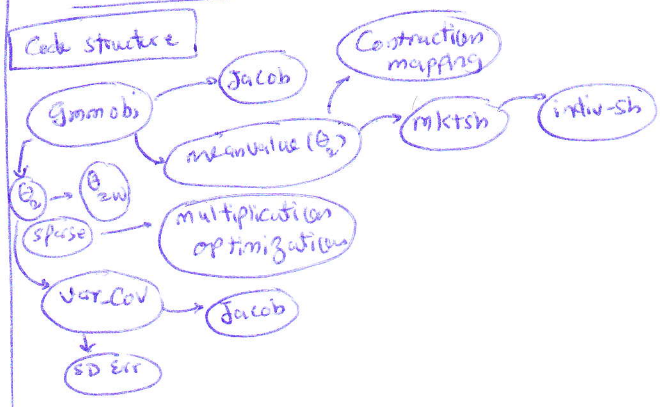
$$J_F(x,y) = \begin{bmatrix} \frac{\partial F_1}{\partial x} & \frac{\partial F_1}{\partial y} \\ \frac{\partial F_2}{\partial x} & \frac{\partial F_2}{\partial y} \end{bmatrix} \quad \text{e.g. } F(x,y) = \begin{bmatrix} x^2 y \\ 5x + \sin(y) \end{bmatrix}$$

$$H(f) = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \vdots & & & \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \frac{\partial^2 f}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{bmatrix} \Rightarrow \text{symmetric matrix}$$

②

- to optimize the code you need to use indexing rather than for loop: usefull functions are: [vectorizing and indexing]
  1. reshape(m, i, j)
  2. sum(m, 2) for rowsum
  3. index Conversion for JxT $[\;]^H$ to TxH $[\;]^J$ to change one vector result of reshape to another
  4. bsxfun(@rdivide, num, den) to divide matrix to vector
  5. eliminate determinants and use chols. for Jacobian

Code structure



Jiang Rossi Manchanda 2009

(5)

Bayesian BLP

(1) orthogonality: it find Z_t: Gmm crit func.: $E[Z_t' \eta_t]$

(2) when price endogeneity instrument is required, otherwise not

(3) moment conditions to identify many elements of Σ

(4) Bays estimator better than Gmm estimator when full specification of likelihood

(5) Gmm performs well under misspecification error without too much efficiency loss under a correctly specified model

(6) Bays estimator performs pretty well in mispecified cases e.g. Hetero, AR(1), Asym Beta, sym Beta

(7) Gmm estimator tends to attenuate the correl. in the rand coeff dist. ⇒ most case tend to bias downward cross price elast. of demand.

(8) on large draws e.g. 200 rather than 50 Gmm improves, but still Bays dominates

(9) to find dist. using Gmm: (1) drive asympt. std err. (2) use delta method or param bootstrap to approx samp. dist.

(10) asymptotic std. error of Gmm are about one half of the size of the actual sampling std. (over optimistic views of percision of estimation): understatement of sampling error

(11) Σ: constraint with positive definiteness

(12) Coverage of Gmm: 63% Bays: 95% for 95% conf interval
Bays estimate & posterior std

(13) Instrumental variable:
- extend by Gibbs sampler for linear struct eq model
- $X_{jt} = \{P_{jt}, W_{jt}\}$
price (log price) ; observed attrib
Correl. with demand shock η_{jt}
- $P_{jt} = Z_{jt}\delta + \xi_{jt}$ (26)

$$\begin{pmatrix}\xi_{jt}\\ \eta_{jt}\end{pmatrix} \sim N\left(\begin{pmatrix}0\\0\end{pmatrix}, \Omega = \begin{bmatrix}\Omega_{11} & \Omega_{12}\\ \Omega_{12} & \Omega_{22}\end{bmatrix}\right)$$

⇒ change of variable

$$\pi(P_t, S_t \mid \bar\theta, r, \delta, \Omega) = \pi(\xi_t, \eta_t \mid \bar\theta, r, \delta, \Omega)\, J_{(\xi_t, \eta_t \to P_t, S_t)}$$
$$= \pi(\xi_t, \eta_t \mid \bar\theta, r, \delta, \Omega)\,(J(P_t, S_t \to \xi_t, \eta_t))^{-1}$$

No direct func rel. demand shock (26)

Jacobian

$$J(P_t, S_t \to \xi_t, \eta_t) = \begin{Vmatrix}\nabla_{\xi_t} P_t & \nabla_{\eta_t} P_t\\ \nabla_{\xi_t} S_t & \nabla_{\eta_t} S_t\end{Vmatrix} = \begin{Vmatrix} I & 0\\ \nabla_{\xi_t} S_t & \nabla_{\eta_t} S_t\end{Vmatrix}$$
$$= \| \nabla_{\eta_t} S_t \| = J_{(S_t \to \eta_t)}$$ same jac without instrum

Likelihood

$$L(\bar\theta, r, \delta, \Omega) = \prod_{t=1}^{T} (J(S_t, P_t, W_t, r))^{-1} \prod_{j=1}^{J} \phi \left(\begin{bmatrix}\xi_{jt} = P_{jt} - Z_{jt}\delta\\ \eta_{jt} = h^{-1}(S_t \mid P_t, W_t, \bar\theta, r)\end{bmatrix} \Bigg| \Omega\right)$$

Given δ, r ⇒ recover η_{jt} (demand supply with instr)

(14) advantage of Bays in instrument: uncertainty in estimate of r

(6)

(15) $P_{jt} = Z_{jt}\delta + \xi_{jt}$
$\mu_{jt} = [w_{jt}, P_{jt}]\bar\theta + \eta_{jt}$ $\begin{pmatrix}\xi_{jt}\\ \eta_{jt}\end{pmatrix} = N\left(\begin{pmatrix}0\\0\end{pmatrix}, \Omega\right)$

same prior: $\bar\theta, r$

std prior: $\delta \sim MVN(\bar\delta, V_\delta)$
$\Omega \sim IW(\nu_0, V_\Omega)$

(16) $\bar\theta, \delta, \Omega \mid r, \{S_t, P_t, W_t, Z_t\}_{t=1}^T, \bar\theta_0, V_{\bar\theta}, \bar\delta, V_\delta, \nu_0, V_\Omega$
$r \mid \bar\theta, \delta, \Omega, \{S_t, P_t, W_t, Z_t\}_{t=1}^T, \sigma^2_{r_{jj}}, \sigma^2_{r\text{-}off}$

(17) problem of high variance & param out of range in baysian is out of bound NaN (not a number), so you need to restrict to range (e.g. diag var due to e^x) ⇒ tunning

(18) when long run time run & find new prior and then run with new prior

Rasmusen 2007

⑦

⑧

Summary

① Instrument: to correct for endog. of prices.

② Product charact: Going from 50 prod. to 6 charact. reduces # of param estimate

③ Cons. and prod. char. interact: marginal utilites allow depend on Cons. charact. $\Rightarrow$ allows substitution pattern sensible

④ Struct estimation: Start with model in which indiv. maximize their payoff (choice of action) & disturbance (unobserved) show up in reg

⑤ Contraction mapping: estimate params that are averaged across cons. & otherwise difficult optm. problem

⑥ seperating linear & nonlinear estimation problems.
   ① one part that uses search algorithm (non linearity)
   ② analytical formula to estimate linear params

⑦ GMM: used to estim. other params

⑧ advantage: requires weaker assumptions than max likelihood

GMM std err   Jiang - Rossi, manchanda 2009

① Condition $E[Z_t' \eta_t] = 0$ $Z_t : [J \times \overset{\text{\# moments}}{M}]$ $\eta_t = [J \times 1]$

② $\hat{m}_T(\bar{\theta}, \Sigma) = \frac{\sum_{t=1}^{T} Z_t'(\hat{\mu}_t(\Sigma) - X_t \bar{\theta})}{T}$

GMM obj: $g(\bar{\theta}, \Sigma) = \hat{m}_T(\bar{\theta}, \Sigma)' A^{-1} \hat{m}_T(\bar{\theta}, \Sigma)$

$A_{[M \times M]}$ = Consist est. of $E[(Z_t' \eta_t)(Z_t' \eta_t)']$

FOC. obj w.r.t $\bar{\theta}$: $\hat{\bar{\theta}}(\Sigma) = (X'ZA^{-1}Z'X)^{-1} X'ZA^{-1}Z'\hat{\mu}(\Sigma)$

$X: [JT \times K]$ $Z: [JT \times M]$

③ Two step of GMM: (S1) $A = \frac{\sum_{t=1}^{T} Z_t' Z_t}{T} \Rightarrow$ minimize the GMM obj $\Rightarrow$ obtain resid. $\hat{\eta}_{jt}$

(S2) Construct new $A = \frac{\sum_{t=1}^{T} Z_t' \hat{\eta}_t^{(1)} \hat{\eta}_t^{(1)\prime} Z_t}{T^2}$

min. GMM obj.   after conv. restart GMM est. to ensure convergence. use final conv. result $\Rightarrow \hat{\Sigma}_{GMM}, \hat{\bar{\theta}}_{GMM}, \hat{\eta}_{jt}$

④ GMM std error: asymmptotic:

$\Psi = (\bar{\theta}, r) \Rightarrow Var(\hat{\Psi}_{GMM}) = \frac{1}{T}(D'\hat{V}^{-1}D)^{-1} \Big|_{\Psi = \hat{\Psi}_{GMM}}$

where $D = \left[\frac{\partial m_T}{\partial \bar{\theta}} \quad \frac{\partial m_T}{\partial r}\right]$

$\frac{\partial m_T}{\partial \bar{\theta}} = -\frac{1}{T}\sum Z_t' X_t$ $\frac{\partial m_T}{\partial r} = \frac{1}{T}\sum_t Z_t'\left(\frac{\partial \hat{\mu}_t}{\partial r}\right)$ mean utility numerical computing

∂

Rasmusen 2007

$u_{ijt} = \alpha_i(y_i - p_{jt}) + x_{jt}\beta_i + \xi_{jt} + \epsilon_{ijt} = \alpha_i y_i - (\alpha + \Pi_\alpha D_i + \Sigma_\alpha v_{i\alpha}) p_{jt}$
$+ x_{jt}(\beta + \Pi_\beta D_i + \Sigma_\beta v_{i\beta})\beta_3 + \xi_{jt} + \epsilon_{ijt} =$
$\alpha_i y_i + (-\alpha p_{jt} + x_{jt}\beta + \xi_{jt}) - (\Pi_\alpha D_i + \Sigma_\alpha v_{i\alpha}) p_{jt} +$
$x_{jt}(\Pi_\beta D_i + \Sigma_\beta v_{i\beta}) + \epsilon_{it} = \alpha_i + \delta_{jt} + \mu_{ijt} + \epsilon_{ijt}$ $\quad j=1,\dots,50 \quad t=1,\dots,20$

$\delta_{jt} = -\alpha p_{jt} + x_{jt}\beta + \xi_{jt}$ $\quad \mu_{ijt} = (-p_{jt}, x_{jt})(\Pi D_i + \Sigma v_i) + \epsilon_{jt}$
$\left[\binom{\Pi_\alpha}{\Pi_\beta} D_i + \binom{\Sigma_\alpha}{\Sigma_\beta}(v_{i\alpha}, v_{i\beta})\right]$

unobserved prod charact. (Instrument helps)
→ mean utility; component of utility same across all Cons.

μ_{ijt}: heteroskedastic disturbance
ϵ_{ijt}: homoskedastic disturbance (iid)

$$s_{ijt} = \frac{e^{\delta_{jt}+\mu_{ijt}}}{1+\sum_{k=1}^{50} e^{\delta_{kt}+\mu_{ikt}}}$$

(23) we estimate dist. of Cons. charact. D: $\hat P_D^*(D)$

$$s_{jt} = \int_v\int_D s_{ijt}\, d\hat P_D^*(D)\, dP_v^*(v) = \int_v\int_D \left[\frac{e^{\delta_{jt}+\mu_{ijt}}}{1+\sum_{k=1}^{50} e^{\delta_{kt}+\mu_{ikt}}}\right] d\ldots$$

integration on consumers

(24) price elasticity:

$$\eta_{jkt} \equiv \frac{\partial s_{jt}}{\partial p_{kt}}\frac{p_{kt}}{s_{jt}} = \begin{cases} -\frac{p_{jt}}{s_{jt}}\int_v\int_D \alpha_i s_{ijt}(1-s_{ijt})\, d\hat P_D^*(D)\, dP_v^*(v) & \text{if } j=k \\ \frac{p_{kt}}{s_{jt}}\int_v\int_D \alpha_i s_{ijt}s_{ikt}\, d\hat P_D^*(D)\cdot dP_v^*(v) \end{cases}$$

(25) simulation for integration
real world knowledge of dist. of Cons. type i in town t & prod charact. j ⊕ how diff. Cons. type value diff prod. charact.
⇒ begin init. set of param ⇒ calc. share/town ⇒ match obsrv mkt share?
⇒ pick new set.

(26) non structural way: $s_{jt} = x_{jt}\beta - \alpha p_{jt} + \xi_{jt}$ — product dummies to account for fixed effect
(simple interaction is not specifics of rand coeff but struct.)
→ or: $s_{jt} = x_{jt}\beta - \alpha p_{jt} + d_t\theta_1 + \omega_t e_{\omega t} + \xi_{jt}$
mean val. 4 Cons. charact. [6 prod. charact] $[1\times 24]$ ⇒ interaction 6×4

Problem: ① not result of consistent theoretical model
First advantage: ② not account for relationships b/w the mkt shares of diff prod. e.g. sum of mkt share not more than one (logit does not allow this internal contradiction)

(27) Second advantage: simultaneity problem of est. demand & supply func
Mk share ← Price ← disturbance ⇒ Instrument ⇓ use GMM

(28) m: # instrum (m = 1, …, M)

$E(Z_m\, \omega(\theta^*)) = 0 \quad \forall m$

GMM estimator: $\hat\theta = \arg\min_\theta \omega(\theta)' Z \Phi^{-1} Z' \omega(\theta)$

Φ is consist estimator of $E(Z'\epsilon\epsilon' Z)$

Z = { sing var inst; multi var inst; square & interaction of single vars

(29) min data
① obsrv. unit (town, month, town/month …)
② products
③ prod. charact.
④ Cons. charact.
optional ⑤ common to all Cons: (1) adv (2) Ret weather (3) time trend (4) Dummy var each prod

(30) nonlinear part of estimation: (Π, Σ) for rand coeff indiv. dev. from avg coeff.
X_1: Expla var. common to all Cons. e.g. price adv., not prod dummy
X_2: directly obs explan. var. (nonlinear): price, com var, but "
→ not cont. Cons. charact. (come in aggreg across Cons.)

(31) choose instr. for price, and other common var. (data collection)

(32) Estimation:
(-1) select arbitrary val. for δ ① (→ arbitrary) and (Π, Σ) (for step 1)
② and for (α, β) for step 3 to start
δ: vector of mean util. from each products
(Π, Σ): matrix of params showing how observed & unobsrvd cons. charact & prod. charact interact
(α, β): avg param across Cons
(0) draw rand val. for (v_i, D_i) $i = 1, \dots, n_s$ from dist $P_v^*(v)$, $\hat P_D^*(D)$ for sample size n_s (bigger more accurate estimate)
(1) using starting value:

$$s_{jt} = \left(\frac{1}{n_s}\right)\sum_{i=1}^{n_s} s_{ijt} = \left(\frac{1}{n_s}\right)\sum_{i=1}^{n_s}\left[\frac{e^{\delta_{jt}+\sum_{k=1}^{6} x_{jt}^k(\sigma_k v_i^k + \pi_{k1}D_{i1}+\cdots+\pi_{k4}D_{i4})}}{1+\sum_{m=1}^{50} e^{\delta_{mt}+\cdots}}\right]$$

⇒ output: predicted mkt share given (Π, Σ)
(2) use contraction mapping: find δ by iterating

$$\delta_{\cdot t}^{h+1} = \delta_{\cdot t}^{h} + \ln(S_{\cdot t}) - \ln(s_{\cdot t})$$

observed mkt share; predicted mrkt share step 1 (updated in iteration)
iterate until observed & predicted mkt share are close enough: $\ln(S_{\cdot t}) - \ln(s_{\cdot t}) \simeq 0$
⇒ output δ
(3) figure out values of moment expr. using strt val. $(\alpha, \beta)(0)$ & $\delta(2)$
(3a) $\omega_{jt} = \delta_{jt} - (-\alpha p_{jt} + x_{jt}\beta)$ calc. ω_{jt}
(3b) calc. val. moment expr.: $\omega' Z \Phi^{-1} Z' \omega$ $\quad \Phi^{-1} = (E(Z'\omega\omega' Z))^{-1}$
Consist. estimator of $\Phi^{-1} = (Z'Z)^{-1}$
(4) compute better estimate of all params
(α, β): common param $\quad (\Pi, \Sigma)$: indiv param
Φ: weighting matrix
(4a) find estimate of param common to Cons:
$(\hat\alpha, \hat\beta) = (X'Z\Phi^{-1}Z'X)^{-1}X'Z\Phi^{-1}Z'\delta$

(33) make everything linear & set up minimization algorithm to minimize it
⇒ separating the param's that can be linearly est. from param req search algorithm
⇒ searching takes longer than matrix mult. also less reliable in finding true min & convergence

(4b) $\hat\omega_{jt} = \delta_{jt} - (-\alpha p_{jt} + x_{jt}\beta) = \tilde\xi_{jt}$
then $(\hat\alpha, \hat\beta) = (X'Z\Phi^{-1}Z'X)^{-1}X'Z\Phi^{-1}Z'\delta$
(4c) $\Phi = Z'\hat\omega\hat\omega' Z$
(4d) use search algorithm to find new value (Π, Σ) & keep iterating to step 1 & find param that min $\omega' Z \Phi^{-1} Z' \omega$ moment expression until it is close to zero

GMM

Rasmusen 2007

①

① Combines: ① overidentified inst. Var ② GLS ③ non linear spec.

② $y = x_1\beta_1 + x_2\beta_2 + \varepsilon$ x exogen

two moment Cond. $M_1: E(x_1'\varepsilon)=0$ or $E(M_1)=0$

$M_2: E(x_2'\varepsilon)=0$ or $E(M_2)=0$

T obs $\Rightarrow x_1: [T\times 1]$ $M_1 = x_1'\varepsilon$ $[1\times 1]$

sum of sq of moments: $\min (M_1\ M_2)'(M_1\ M_2) = [1\times 1]$ $\nearrow [2\times 1]$

M_1: rand var $E(M_1)=0$ $Var(M_1)\neq 0$

another way is: find $\hat\beta$ minimizes $M'M$ where $M = x'\varepsilon$

$\Rightarrow$ find $\hat\beta = \underset{\beta}{\mathrm{argmin}}\ \varepsilon' X X' \varepsilon$

$f(\hat\beta) = \hat\varepsilon' X X' \hat\varepsilon = (y - X\hat\beta)' X X'(y - X\hat\beta) = y'XX'y - \hat\beta' X'XX'y - y'XX'X\hat\beta$

$+ \hat\beta' X'XX'X\hat\beta \Rightarrow f'(\hat\beta) = -X'XX'y - y'XX'X + 2\hat\beta X'XX'X$

$= -2X'XX'y + 2\hat\beta' X'XX'X = 2X'X(-X'y + \hat\beta' X'X) = 0$

$\Rightarrow \hat\beta = (X'X)^{-1}X'y$ $E(\varepsilon | x_1, x_2) = 0$ independant

Stronger than uncorrelation

③ independance is information $\Rightarrow$ helps efficient estimation

$\equiv$ i.e. info of variance does not depend on x

$\to$ gives lots of potential moment e.g.:

$E((x_1^2)'\varepsilon) = E(M_3) = 0$ $E((x_2 * x_1)'\varepsilon) = E(M_4) = 0$

④ Some Cond. more reliable than others $\Rightarrow$ weight according to inv. Var.

$\equiv$ inv Var.Cov matrix, since Correlation Call $(\Phi(M))$

⑤ $\Rightarrow$ method of moment:

$$\boxed{(M_1\ M_2\ M_3\ M_4)'(\Phi(M)^{-1})(M_1\ M_2\ M_3\ M_4)}$$

finite sample $\Rightarrow$ less moment better

⑥ higher moment e.g. $E((x_2^3 * x_1^5)'\varepsilon) = E(M_5) = 0$

does not add lot of info, Could have big Var

⑦ GMM like GLS: $\Phi(M)$ estimate of Var-Cov matrix of ε

⑧ need Correction for hetroskedasticity $\Rightarrow$ BLP

need Correction for hetroskedasticity $\Rightarrow$ BLP need ... as a form of instrumental variables est. true strength when used as a form of instrumental variables est.

⑨ if $E(x_2\varepsilon)\neq 0 \Rightarrow$ lost moment M_2, M_4 $\Rightarrow$ need instrument (new basic moment Condition)

$\Rightarrow E(x_1\varepsilon)=0$ $E((x_1^2)'\varepsilon)=0$ $E(z_1\varepsilon)=0$ $E(z_2\varepsilon)=0$

$E((z_1 * x_1)'\varepsilon)=0$ $E((z_1 * z_2)'\varepsilon)=0$

abreviate as: $E(Z'\varepsilon) = E(M) = 0$ $Z = \begin{pmatrix} z_1 \\ z_1 \\ z_2 \\ z_1 x_1 \\ z_1 * z_2 \end{pmatrix}$

⑩ $\Phi(M) = Var(M) = Var(Z'\varepsilon) = E(Z'\varepsilon\varepsilon' Z) - E(Z'\varepsilon)E(\varepsilon' Z)$

$= E(Z'(I\sigma^2)Z) - 0^2 = \sigma^2 Z'Z$ distr. indep. of eachother & Z $\equiv$ no hetrosc.

⑪ GMM: uses var cov matrix (diff weights on 6 moment Cond.)

solves problem of: $f(\hat\beta_{2SLS}) = \hat\varepsilon_{2SLS}' Z(\sigma^2 Z'Z)^{-1} Z'\hat\varepsilon_{2SLS}$

$= (y - X\hat\beta_{2SLS})' Z(\sigma^2 Z'Z)^{-1} Z'(y - X\hat\beta_{2SLS})$

$\Rightarrow f'(\hat\beta_{2SLS}) = -X'Z(\sigma^2 Z'Z)^{-1}Z'(y - X\hat\beta_{2SLS}) = 0$

$\Rightarrow$ solves: $\hat\beta_{2SLS} = [X'Z(Z'Z)^{-1}Z'X]^{-1} X'Z(Z'Z)^{-1}Z'y$ Such as GLS

Both GMM & 2SLS same when homosced.

②

⑫ when heterosk. diff. (GMM & 2SLS) since GMM will use $(\Phi(M))^{-1} \neq (Z'Z)^{-1}$

⑬ 2SLS $\neq$ IV: $\hat\beta_{IV} = (X'Z)^{-1}Z'y$

Same if # param to estimate = # instrument

not when: $X' \times Z = $ not sq. $\Rightarrow$ not invertable

$[J\times T]\ [T\times K]\quad [J\times K]$

⑭ 2SLS projects X onto Z with proj matrix $Z(Z'Z)^{-1}Z'$

to gen sq matrix that can be invertible

GMM similar but with $\Phi(M)$ instead of $Z(Z'Z)^{-1}Z'$

⑮ nonlinear in GMM:

$y = x_1^{\beta_1} + \beta_1\beta_2 x_2 + \varepsilon$ moment Conditions:

$E(x_1'\varepsilon) = M_1 = 0$ $E(x_2'\varepsilon) = M_2 = 0$ $E(x_1 * x_2'\varepsilon) = M_3 = 0$

$\Rightarrow$ problem $(y - x_1^{\beta_1} + \beta_1 * \beta_2 * x_2)' M(\Phi(M))^{-1} M'(y - x_1^{\beta_1} + \beta_1\beta_2 x_2)$

have to estimate $\Phi(M)$ in search

NEVO 2000

(1)

(1) random coefficient model $\equiv$ mixed logit

(2) logit $\Rightarrow$ projection into space of characteristics

problem: substitution driven by mkt share & not by how similar products are

explicitly model heterogeneity in population (& params of fnct)

problem: treat regressor exogenously determined

(3) BLP advantage: (1) mkt level price & quant

(2) deals with endog. of price

(3) demand elasticity that more realistic (e.g. cross price elasticity)

(4) prices could be correlated

(5) product defined by set of charac. (producer & cons. observe all of charac.; the researcher observes only some of prod charac.)

(6) producer knows charac & take into account when setting prices

$\Rightarrow$ prices are endogeneous $\quad T$: market $\quad I_t$: cons. $\quad$ [obs. aggreg quant, avg P, avg characn.]

$\rightarrow$ indiv taste $[1 \times K]$

(7) $u_{ijt} = \alpha_i(y_i - p_{jt}) + x_{jt}\beta_i + \xi_{jt} + \varepsilon_{ijt}$ (1) $\quad$ Consumer i

marginal util inc. $\quad$ income of cons $\quad$ unobsrved prod charac (vertical prod. diff.)

(8) different cons. could face different price observed $\Rightarrow$ use instrument

(9) Cons. obsrvd charach. D_i $\quad$ ν_i: cons. unobserved charc

(know dist: Demogr.)

Consw: mean & std. $\rightarrow [(K+1)\times(K+1)]$

$$\begin{pmatrix}\alpha_i\\ \beta_i\end{pmatrix} = \begin{pmatrix}\alpha\\ \beta\end{pmatrix} + \Pi D_i + \Sigma \nu_i \quad (2) \qquad \nu_i \sim P_\nu^*(\nu) \qquad D_i \sim \hat{P}_D^*(D)$$

$[d\times 1]$ $\quad [(K+1)\times d]$ $\quad$ parametric dist MVN $\quad$ parametric or nonparam dist.

(10) to allow homogeneous price increase (relative to other sector) of all products change quantities purchase $\Rightarrow$ outside goods

$u_{i0t} = \alpha_i y_i + \xi_{0t} + \pi_0 D_i + \sigma_0 \nu_{i0t} + \varepsilon_{i0t}$ $\quad$ (to identity set to 0, normalize all to 0)

$\theta = (\theta_1, \theta_2)$ all params $\quad \theta_2 = (\Pi, \Sigma)$ nonliner params

$\theta_1 = (\alpha, \beta)$ liner params

(11) Combining (1) and (2): $u_{ijt} = \alpha_i y_i + \delta_{jt}(x_{jt}, p_{jt}, \xi_{jt}; \theta_1)$ (3)

$+ \mu_{ijt}(x_{jt}, p_{jt}, \nu_i, D_i; \theta_2) + \varepsilon_{ijt}$

Common to all cons. $\Rightarrow$ mean utility $\leftarrow$ $\delta_{jt} = x_{jt}\beta - \alpha p_{jt} + \xi_{jt}$ $\quad$ mean zero heteroskedastic deviation from the mean utility $\Rightarrow$ effect of rand coeff

$\mu_{ijt} = [-p_{jt}, x_{jt}](\Pi D_i + \Sigma \nu_i)$

$[1 \times (K+1)]$

vector of demographics and prod-spcfc shocks: $(D_i, \nu_i, \varepsilon_{i0t}, \ldots, \varepsilon_{iJt})$

(12)

(13) set of indiv. attrib: $A_{jt}(x_{\cdot t}, p_{\cdot t}, \delta_{\cdot t}; \theta_2) = \{(D_i, \nu_i, \varepsilon_{i0t}, \ldots, \varepsilon_{iJt}) \mid u_{ijt} \geq u_{ilt} \ \forall l = 0, 1, \ldots, J\}$

$x_{\cdot t} = (x_{2t}, \ldots, x_{Jt})'$ $\quad \delta_{\cdot t} = (\delta_{2t}, \ldots, \delta_{Jt})'$

$p_{\cdot t} = (p_{2t}, \ldots, p_{Jt})'$ $\quad$ pop dist func

(14) market share: $s_{jt}(x_{\cdot t}, p_{\cdot t}, \delta_{\cdot t}; \theta_2) = \int_{A_{jt}} dP^*(D, \nu, \varepsilon) = \int_{A_{jt}} dP^*(\varepsilon \mid D, \nu)$

$dP^*(\nu \mid D)\, dP_D^*(D) = \int_{A_{jt}} dP_\varepsilon^*(\varepsilon)\, dP_\nu^*(\nu)\, d\hat{P}_D^*(D)$ (4)

(15) BLP: account for correl. b/w prices & unobserved product charact.

(16) price elasticities: $\eta_{jkt} = \frac{\partial s_{jt}}{\partial p_{kt}} \frac{p_{kt}}{s_{jt}} = \begin{cases} -\alpha p_{jt}(1 - s_{jt}) & \text{if } j = k \\ \alpha p_{kt} s_{kt} & \text{otherwise} \end{cases}$

(17) problems of nested logit: (1) priori division of prod. into groups not reasonable

(2) assumption of iid shocks not reasonable

(3) segmentation does not fully account for substitution pattern

(4) does not help with problem of own price elasticity

(2)

(18) price elasticities based on (1), (2):

$$\eta_{jkt} = \frac{\partial s_{jt}}{\partial p_{kt}} \cdot \frac{p_{kt}}{s_{jt}} = \begin{cases} -\frac{p_{jt}}{s_{jt}} \int \alpha_i s_{ijt}(1 - s_{ijt})\, d\hat{P}_D^*(D)\, dP_\nu^*(\nu) & \text{if } j = k \\ \frac{p_{kt}}{s_{jt}} \int \alpha_i s_{ijt} s_{ikt}\, d\hat{P}_D^*(D) \cdot dP_\nu^*(\nu) \end{cases}$$

$$s_{ijt} = \frac{\exp(\delta_{jt} + \mu_{ijt})}{1 + \sum_{k=1}^{K} \exp(\delta_{kt} + \mu_{ikt})} \quad \text{prob. of indiv } i \text{ purchase } j$$

$\Rightarrow$ flexible & not constraint by priori segmentation of market

(19) $\mu_{ijt} + \varepsilon_{ijt}$ is no longer indep. of product & cons. charac.

$\Rightarrow$ cons. more likely to switch to brand with similar charact.

(20) need parametric assump. of the functional form of the dist. or by using additional data src's (e.g. sampling CPS)

(21) data: (1) mkt shares (2) prices (3) brand charact.

useful data: (1) dist. demog $\hat{P}_D^*$ (2) marketing mix [promot. activity, adv. expend, avail coupons]

(22) define quant var based on need/context (e.g. # auto sold, cereal serving)

(23) estimation: totals size of mkt − inside goods = outside goods

e.g. cereal per capita/day

office employees for pc. (large enough mkt to allow non zero share of outside goods)

(24) must check sensitivity of result to mkt size

or define constant proportional to population

(25) product charact: physical prod charac & mkt segmentation info

(1) manufac desc. of prod (2) trade press (3) researcher prior

(26) product dummy variable could be included

(27) for demographic dist. assumption: Current Population Survey (CPS) for demographic info

(28) product as bundle of attribute

(29) random coefficient allows when increase price A, rather than equal increase in mkt share B, C, the one similar e.g. B has more increase

(30) GMM estimator obj func (instrument) search over all possible param values to find values minimize objctv func.

(31) straight forward to solve: $\min_\theta \| s(x, p, \delta(x, p, \xi; \theta_1); \theta_2) - S$

problem of nonlinearity $\Rightarrow$ BLP $\Rightarrow$ make linear

(32) unobserved: $(D_i, \nu_i, \varepsilon_i), \xi_i$ $\quad m = 1, \ldots, M$

(33) $Z = [z_1, z_2, \ldots, z_m]$ instruments $\quad E[Z_m \omega(\theta^*)] = 0$

func model param $\quad$ true value

$\Rightarrow$ GMM Estimate: $\hat{\theta} = \arg\min_\theta \omega(\theta)' Z \Phi^{-1} Z' \omega(\theta)$

gives less weight to $\Leftarrow$ consistent estimates of $E[Z'\omega]$ [measure how close to zero]

Reason: like weighted least square, higher variance lower weight, so all are equally minimized.

simple MM does: $\arg\min_{\mu, \gamma_0, \gamma_1, \gamma_2} \frac{\sum}{T}(\xi_t - \mu)^2 + ((\xi_t - \mu)^2 - \gamma_0)^2 + \ldots$

(34) not diff observed & predicted err, but structural err ξ_{jt} $\Rightarrow$ instrum

$\Rightarrow$ explit func of param of model & data (enters mean utility $\delta(\cdot)$)

(35) need mean utility as linear func. of var & params of model:

$s(\delta_{\cdot t} \mid \theta_2) = S_{\cdot t}$ $\quad t = 1, \ldots, T$

observed mkt share

to solve: $s_{jt}(p_{\cdot t}, x_{\cdot t}, \delta_{\cdot t}, P_{ns} \mid \theta_2) = \frac{1}{ns} \sum_{i=1}^{ns} s_{jti} =$

$$\frac{1}{ns} \sum_{i=1}^{ns} \frac{\exp(\delta_{jt} + \sum_{k=1}^{K} x_{jt}^k (\sigma_k \nu_i^k + \pi_{k1} D_{i1} + \cdots + \pi_{kd} D_{id}))}{1 + \sum_{m=1}^{J} \text{same} \ldots \equiv [\delta_{mt} + \cdots]}$$

$(\nu_i^1, \ldots, \nu_i^K) \sim$ draws of $P_\nu^*(\nu)$ $\quad x_{jt}^k \sim$ var with rand slope coeff

$(D_{i1}, \ldots, D_{id}) \sim$ draws of $\hat{P}_D^*(D)$

NEVO 2000

(36) Second: using mkt share computation: invert the system of equations: analytically:

$\delta_{jt} = \ln S_{jt} - \ln S_{ot}$ → mkt share of outside goods

$\delta^{h+1}_{\cdot t} = \delta^{h}_{\cdot t} + \ln S_{\cdot t} - \ln S(P_{\cdot t}, x_{\cdot t}, \delta^{h}_{\cdot t}, P_{ns} | \theta_2)$

(37) Once inversion done analytically or numerically

Error term is: $\omega_{jt} = \delta_{jt}(S_{\cdot t} | \theta_2) - (x_{jt}\beta + \alpha P_{jt}) \equiv \xi_{jt}$

θ_1: linear θ_2: nonlinear — observed mkt share — unobserved product charact.

(38) logit & nested logit, with appropriate weight matrix ⇒ 2SLS

(39) Full rand Coeff model: Compt. mkt shares, and inversion to get $\delta_{jt}(\cdot)$: numerically & nonlinear search to find $E(Z_m \omega(\theta)) = 0$

$\hat{\theta} = \arg\min_{\theta} \omega(\theta)' Z \Phi^{-1} Z' \omega(\theta)$

(40) Key stem is $E(Z_m \omega(\theta^*)) = 0$ identifying assumption

(41) price = f(Cost, markup) markup = g(unobsrved Prod. charac.)

(42) demand side instrument: shift Cost & uncorrelated with demand shock

(43) location of product is exogenous, or determined before customer revelation of evaluation of unobserved. Prod. charac.

(44) ① sum of the values of the same charact. of other prod. offered by the firm ② sums of the values of the same charact. of prod. offered by other firms

(45) observed charact. may not be uncorrelated with unobserved (e.g. time to change observed charach. is short)

(46) Control for brand specific intercept, demographics, City specific valuations of prod (indep. across cities, but Cor. within cities over time) ⇒ Price in other cities valid instrument

indep violated when: local ads & promo. Coordinated national awareness of fiber

(47) Brand fix effect should be included: Problem No dimensionality

(48) $d = (d_1, \dots, d_J)'$ $[J\times1]$ dummy brand $X(J\times K)(K<J)$ prod charac.

$\xi = (\xi_1, \dots, \xi_J)'$ $[J\times1]$ unobsrvd prod. qual. $\hat{\xi} = \hat{d} - X\hat{\beta}$

$d = X\beta + \xi$ $\hat{\beta} = (X' V_d^{-1} X)^{-1} X' V_d^{-1} \hat{d}$ $E[\xi | x] = 0$

(49) statistical significance by: (1) use more data

(2) increasing the # of simulation (ns)

(3) improving simulation method

(4) adding supply side

(50) Exogeneous: instrument is crucial

Rasmusen 2007

① t: towns $q_t(P_t) = \alpha - \beta P_t + \varepsilon_t$ supply estimate & not demand

② add income: $q_t(P_t) = \alpha - \beta P_t + \gamma y_t + \varepsilon_t$

still not able to untangle supply shifts from demand shifts

need observed var that shifts supply

③ individual $q_{it}(P_t) = \alpha - \beta P_t + \gamma y_{it} + \varepsilon_{it}$

④ weight each Cons. by likely freq of his type of pset. in pop.

⑤ indogeneity of price, unobservable, noise

⑥ instrumental var. that Correlated with Price but not anything else in demand Equations

⑦ root of problem: unobservable demand disturbance

(8) ① production start b/c somebody demands it ② entrepreneur discovers supply ③ other Entrepreneur ⇒ industry compete

④ demand depends: Price all firm indst ⇒ oligopoly

⑤ demand depends: Prod charact ⇒ monoplstc Compt. Election theory

⑥ demand depends on Cons. info ⇒ search theory, advertising, advers select, moral hazard

(Start from estim. demand)

(9) primitive: utility func, prod func (Starting Point)

players, actions, info., payoffs, equil strtgy

(10) e.g. 20 towns of Cereal mkt data, 50 brands

6 charct. of each Cereal brand, 4 Cons. Charact. dist in each town

(11) Curse of dim.: 50 products × 50 other Prices = 2500 param when only 1000 data points.

demand (interaction)

(12) Solution: Cons. utility func. of prod. charact.

indirect utility — Probability of whether to buy rather than how much

Cross elast ⇒ by Combining charact. effect param

(13) $u_{ijt} = \alpha_i(y_i - P_{jt}) + x_{jt}\beta_i + \xi_{jt} + \epsilon_{ijt}$ $i = 1, \dots, 400$; $j = 1, \dots, 50$; $t = 1, \dots, 20$

6dim vect. of observed charct. of product

disturbance scalar summarizing unobserved charct product

(14) not buy any product at all: $u_{i0t} \equiv \alpha_i y_i + \varepsilon_{i0t}$

(normalize or assume zero price, zero charact.)

(13) Consumer's problem:

$\max_{q_0, q_1, \dots, q_{50}} \{q_0 + u_1(q_1) + u_2(q_2) + \dots + u_{50}(q_{50})\}$

s.t. $\sum_{j=0}^{50} \{q_j P_{jt}\} \le y$ budg. Const.

$\sum_{j=0}^{50} \{q_j\} = 1$ only one unit.

(14) Cons util depend on: ① prod charact. x_{jt} ② product, same for all Cons ξ_{jt} prod (fixed effect) ③ unobsrvd on Cons, prod, town (ε_{ijt})

(15) Cons. Charct. affect β_i (marg. util. of prod. charact.)

(16) $\varepsilon_{it} \sim f(x) = e^{-x} e^{-e^{-x}}$ $F(x) = e^{-e^{-x}}$ thicker tail  ← left skew

(17) identical customers: $\beta_i = \beta$, $\alpha_i = \alpha$

(18) identical ⇒ $S_{jt} = \dfrac{e^{x_{jt}\beta - \alpha P_{jt} + \xi_{jt}}}{1 + \sum_{k=1}^{50} e^{x_{kt}\beta - \alpha P_{kt} + \xi_{kt}}}$ ← normalized zero outside good

Prob prod. 1 higher util:

$Prob(\alpha(y - P_{1t}) + x_{1t}\beta + \xi_{1t} + \varepsilon_{1t} > \alpha(y - P_{2t}) + x_{2t}\beta + \xi_{2t} + \varepsilon_{2t}) \times$

$Prob(\alpha(y - P_{1t}) + x_{1t}\beta + \xi_{1t} + \varepsilon_{1t} > \alpha(y - P_{2t}) + x_{3t}\beta + \xi_{3t} + \varepsilon_{3t}) \times$

$\dots \times Prob(\alpha(y - P_{1t}) + x_{1t}\beta + \xi_{1t} + \varepsilon_{1t} > \alpha y + \varepsilon_{0t})$

(19) ① Problem of low share ⇒ $\alpha(1 - S_{jt}) = -\alpha P_{jt}$ = own elastic ⇒ lower price ⇒ demand is less elastic ⇒ Can have ↑ markup when low marg. Cost

② But wrong, markup is higher for luxury in reality

red bus/blue bus & bicycle ⇒ ↑red Price ⇒ switch more to blue bus ⇒ due to cross $\alpha P_{kt} S_{kt}$

(20) nested logic problem: how to define categories

(21) random Coeff ≡ indiv Coeff.

(22) $\begin{pmatrix}\alpha_i \\ \beta_i\end{pmatrix} = \begin{pmatrix}\alpha \\ \beta\end{pmatrix} + \Pi D_i + \Sigma \nu_i = \begin{pmatrix}\alpha \\ \beta\end{pmatrix} + \begin{pmatrix}\Pi_\alpha \\ \Pi_\beta\end{pmatrix} D_i + \begin{pmatrix}\Sigma_\alpha \\ \Sigma_\beta\end{pmatrix} [\nu_{i\alpha}, \nu_{i\beta}]$

Cons i obsrv charact ← $[4\times1]$; $[7\times1]$ Cons i unobs charact; $[7\times4] = \Pi$ → $6\times\beta_i$, $1\times\alpha_i$; Σ = $[7\times7]$

$D \sim P^*_D(D)$ $\nu \sim P^*_{\nu}(\nu) \equiv MVN$

BLP

- ① Logic?
- ② Gauss Code? → Crack
- ③ Convertable Bayesian?

ID – Expenditure – income – Household size – Brand – feature

① create share of sales assuming mkt size
↳ creat out share

② Stack brands

① price ② sales ③ feature ④ Display

matrix form

① Brand intercept, price, feature, Display, Fiture & Display

nu data

person1 Brand 1 … 4 [ ]

person2 Brand 1 … 4 [ ]

person3 Brand 1 … Brand 4 [ ]

people N ↑ Brand ↑

ou: ① all sales 115 × 4
② share from sales mkt size 1 mil
③ transpose 4 × 115 same
④ outside share transposed; reshap (ovovovov) → form; four by four outside option the same

role: 1-p to have $\ln \frac{p}{1-p}$

② $b_{ols} = (X'X)^{-1}(X'Y)$ → # Sales / total share; $\ln\left(\frac{p}{1-p}\right)$

nu data

$\hat{e}_{ols} = Y - Xb_{ols}$

$V(\hat{e}_{ols}) = (Y - X\beta_{ols})'(X'X)^{-1}(Y - X'b_{ols})/N$

$(X'X)^{-1}X'\Omega X(X'X)^{-1}$

③ instrument: standard price

→ Fixed effect of Brand

0	$O \times$ inst ~~...~~ is item multipl	f	P	fP

↳ instead, X use put instrument and only effective as fixed effect

$\left\{ \begin{array}{l} y = X\alpha + \\ X = Y\beta + Q\alpha\gamma \end{array}\right.$

$bb = \left(\frac{X'ZZ'X}{\sigma_Z^2}\right) \times \left(\frac{X'ZZ'Y}{\sigma_Z}\right)$ 2SLS

$\hat{e}_{2SLS} \Rightarrow V_{2SLS} \Rightarrow$ Approx optimal weight matrix

observed aggregate share is sales share per store

$$s_{jts} = \frac{e^{\alpha_j + X_{jt}\beta + \xi_{jt}}}{1 + \sum e^{\alpha_k + X_{kt}\beta + \xi_{kt}}}$$

→ randomness from shock; problem of correlation

$\Rightarrow \ln\left(\frac{S_{jt}}{1 - S_{jt}}\right) = \alpha_j + X_{jt}\beta + \xi_{jt}$

price ↗ advertising; instrument does not capture; instrument captures

Z_{it}: suck out part of X that uncorrelated with ξ_{it}

$\int pr_{ijt|\theta_i} f(\theta_i)d\theta_i = S_{it}$ indiv aggreg $\frac{S_{jt}}{S_{0t}} = \frac{\int \dots}{\int \dots}$

$$S_{jt} = \int \frac{e^{\delta_{jt} + \Delta M_{ijt}}}{1 + \sum_{k=1}^{K} e^{\delta_{kt} + \Delta M_{ikt}}} f(\theta_i)\theta_i$$ → seperation of mean & idiosyncratic difference

$\delta_{jt} = \alpha_j + X_{jt}\beta + \xi_{kt}$ $\quad \Delta M_{ikt} = \Delta\alpha_{ik} + X_{jt}\Delta\beta_j$

$\delta_{jt}^{n+1} = \delta_{jt}^{n} + \ln(S_{jt}) - \ln(\hat{S}_{jt}^{n}) \Rightarrow |\delta^{n+1} - \delta^{n}| << \epsilon$

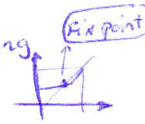
$\delta_{jt}^{1} = \delta_{jt}^{0} + \ln(S_{jt}) - \ln(\hat{S}_{jt}|\delta^{0})$

outer loop: Guess the value of Γ (lower triang matrix)

inner loop

Loop ① Given Γ compute $(J+L)$ δ using Contraction map $\frac{(J+L)(J+L+1)}{2}$

Fix point



② Given δ run 2SLS

③ get ξ_{jt}

outer loop $\xi' Z(Z'Z)^{-1}Z'\xi$

solving correlation with instrument

not optimal weighting matrix (GMM)

Hetrogenity

① create draws for Brand pref & price sensitivity variance

$(w_1, w_2, w_3, w_4, w_p)$ ↳ price

$[114, 1000] \sim$ Normal × 4 then stack; N rows; columns ⇓ # of people

② procedure of optimization:

③ Contraction matrix

Finds best δ that explains

(3-1) add random components of of Brands (aw) and price (awp) to δ (de) & calculate probability

(3-2) check maximum element of $|\delta^{n+1} - \delta^{n}|$

↓ update based on difference of MKT shr

④ 2SLS: matrix, take out instrument

$\beta = \left(\frac{X'ZZ'X}{\sigma_Z}\right)^{-1}\left(\frac{X'ZZ'Y}{\sigma_Z}\right)$ ↳ δ: data(δ)

ZZ; dd

$\boxed{\sigma_Z = Z'Z}$

estimate vs real; based on next draw

$\xi_{it} = \delta - X\beta$

er; dd

⑤ return of procedure: $\xi' Z(Z'Z)^{-1}Z'\xi$

er' × oins × [ ] × oins' er

mmm

Optimization procedure to find variance that best explains data

e.g.

$u_{ijt} = X_{jt}\beta + \beta_{ib}B_{jt}(\text{Branded Product}) + \alpha_i P_{jt} + \xi_{jt} + \epsilon_{ijt}$

↳ $= \sigma_B v_i$ ↳ draw from a standard random distrib.

price coeff $\alpha_i = \alpha + \sigma_I I_i$ (Income)

instrument. ① cost ② price of the same product at other stores in the same period

$Z'Z$ → weighting

Bayesian

① $\hat{\beta} \xrightarrow{d} N(\beta_1, \sigma_1^2) \rightarrow \Gamma\left(\frac{\alpha_0}{2}, \frac{S_0}{2}\right)$ → draw

↳ $N(\beta_0, B_0)$

σ_1 → variance of estimator (stat infer.) and not variance of β mean B_0

iteration

②

$B_1 = \left(\frac{X'X}{\sigma_1^2} + B_0^{-1}\right)^{-1}$

$\beta_1 = B_1 \times \left(\frac{X'Y}{\sigma^2} + B_0^{-1}\beta_0\right) \rightarrow 0$

$\beta_1 = \beta_1 + \text{chol}(B_1) \times \text{randn}$ → used for σ_1^2 draw effect in ete

③

$e = y - \beta_1 X$

$\alpha_1 = \frac{\alpha_0 + N}{2}$

$S_1 = \frac{(S_0 + e'e)}{2}$

$\sigma_1^2 = \text{randGam}\left(\frac{\alpha_1}{2}, \frac{S_1}{2}\right)$ → uses in next draw of β_i

② mixture & Random Coefficient Bayesian models ①

③ + Selection Model

Random coefficien

- Common demand shock
- aggregate random coeff logit
- Good w.r.t GMM (MSE lower)
- random walk metropolis hasting
- Robust to misspecified shock dist
- normal distribution of parameters over individuals
- price elasticity [important] utility weights
- GMM procedure $\Rightarrow$ understate sampling variance
- Data Augmentation $\longrightarrow$ pseudo consumers $\longrightarrow$ add set of param

integrate posterior with augmented param = likelihood without param

latent indirect utility:

$$U_{ijt} = f(X_{jt} \mid \theta^i) + \eta_{jt} + \varepsilon_{ijt} = X_{jt}\theta^i + \eta_{jt} + \varepsilon_{jt} \sim \text{iid Extreme Value } (0,1)$$

θ^i: household preference ≈ sensitivity $\sim N(\bar\theta, \Sigma)$

η_{jt}: ideosyncratic shock, ~~common across consumers~~/ ~~households~~

Consumer, Product, Time

X_{jt} $[1 \times K]$: includes all observed product attrib e.g. Brand intercept & price P_{jt}

η_{jt}: aggregate demand shock; common across cons/ households (or time varying unobserved product attribute) → Chang Tax money Govmt

Predicted share: (2)

$$s_{jt} = \int s_{ijt}\,\phi(\theta^i \mid \bar\theta, \Sigma)\,d\theta^i = \int \frac{e^{X_{jt}\theta^i + \eta_{jt}}}{1 + \sum_{k=1}^{J} e^{X_{kt}\theta^i + \eta_{kt}}}\,\phi(\theta^i \mid \bar\theta, \Sigma)\,d\theta^i$$

$\phi \longrightarrow$ multivariate normal density

mean utility using identity:

$$\theta^i = \bar\theta + v_i \qquad v_i \sim N(0, \Sigma)$$

$$s_{jt} = \int \frac{e^{\mu_{jt} + X_{jt} v_i}}{1 + \sum_{k=1}^{J} e^{\mu_{kt} + X_{kt} v_i}}\,\phi(v_i \mid 0, \Sigma)\,dv$$

Observed product attrib

$\longrightarrow$ observed covariate $(X'_{1t}, \dots, X'_{Jt})' = X_t$

$\mu_{jt} = X_{jt}\bar\theta + \eta_{jt}$ $\longrightarrow$ aggregate demand shock $(\eta_{1t}, \dots, \eta_{Jt})'$

$\longrightarrow$ share $(s_{1t}, \dots, s_{Jt})'$

$\Rightarrow$ density of share as func. of density of aggreg. demand shock: $h(\cdot)$

$$s_{jt} = h(\eta_t \mid X_t, \bar\theta, \Sigma)$$

assumption: Common shocks independently distributed across all products with identical variance: $\eta_{jt} \sim N(0, \tau^2)$ (**)

Change of variable Theorem: $\pi(s_{1t}, \dots, s_{Jt} \mid X_t, \bar\theta, \Sigma, \tau^2) =$

$$\phi(h^{-1}(s_{1t}, \dots, s_{Jt} \mid X_t, \bar\theta, \Sigma) \mid \tau^2)\, J_{\eta_t \to s_t} = \phi(h^{-1}(s_{1t}, \dots, s_{Jt} \mid X_t, \bar\theta, \Sigma) \mid \tau^2) \cdot (J_{(s_t \to \eta_t)})^{-1}$$

$$\Rightarrow L(\bar\theta, \Sigma, \tau^2) = \prod_{t=1}^{T} \pi(s_t \mid X_t, \bar\theta, \Sigma, \tau^2)$$

Rand Coeff mean; Cover of Random Coeff; Var shock

Q: ① share inversion? ② Jacobian

one fix point

Contraction mapping: $\eta_{jt}^{new} = \eta_{jt}^{old} + \ln(s_{jt}) - \ln(h(\eta_{1t}^{old}, \dots, \eta_{Jt}^{old} \mid \bar\theta, \Sigma))$

$$u_{ijt} = \mu_{jt} + X_{jt} v_i + \varepsilon_{ijt} \qquad \mu_{jt} = X_{jt}\bar\theta + \eta_{jt} \qquad v_i \sim N(0, \Sigma)$$

iterative procedure of BLP to obtain $\mu_t = (\mu_{1t}, \dots, \mu_{Jt})'$

- avg over finite draws from $N(0, \Sigma)$ to evaluate integral (2) $20 \le H \le 50$

$\hookrightarrow$ do in every single eval of objective func

$$\pi(s_t \mid \bar\theta, \Sigma, \tau, X_t) = \phi(h^{-1}(s_t \mid \bar\theta, \Sigma, X_t) \mid \tau)\, J_{(\eta_t \to s_t)}$$

$s_t, \eta_t \sim [J \times 1]$ vectors at time t ②

$$J_{(s_t \to \eta_t)} = \| \nabla_{\eta_t} s_t \| = \left\| \begin{matrix} \frac{\partial s_{1t}}{\partial \eta_{1t}} & \frac{\partial s_{1t}}{\partial \eta_{2t}} & \cdots & \frac{\partial s_{1t}}{\partial \eta_{Jt}} \\ \vdots & & & \\ \frac{\partial s_{Jt}}{\partial \eta_{1t}} & & \cdots & \frac{\partial s_{Jt}}{\partial \eta_{Jt}} \end{matrix} \right\|_{J \times J}$$

matrix elements from (2)

$$\frac{\partial s_{jt}}{\partial \eta_{kt}} = \begin{cases} \int -s_{ijt} s_{ikt}\,\phi(\theta^i \mid \bar\theta, \Sigma)\,d\theta^i & \text{if } k \ne j \\ \int s_{ijt}(1 - s_{ijt})\,\phi(\theta^i \mid \bar\theta, \Sigma)\,d\theta^i & \text{if } k = j \end{cases}$$

Jacobian only function of Σ (not $\bar\theta$ or τ^2)

$$\bar\theta \sim MVN(\bar\theta_0, V_{\bar\theta})$$

$$\tau^2 \sim \nu_0 s_0^2 / \chi^2_{\nu_0}$$

Cover matrix: $\frac{K(K+1)}{2}$ unique elements of cholesky root (K: # random coeff)

$\rightarrow$ positive definiteness: reparametrize log of the diagonal elements of the root $\Sigma = U'U$

$$U = \begin{bmatrix} e^{r_{11}} & r_{12} & \cdots & r_{1K} \\ 0 & e^{r_{22}} & & \vdots \\ \vdots & & \ddots & r_{K-1K} \\ 0 & \cdots & 0 & e^{r_{KK}} \end{bmatrix} \qquad r = \{r_{jk}, j, k = 1, \dots, K,\ j \le K\}$$

Priors of r: $r_{jj} \sim N(0, \sigma^2_{r-jj})$ for $j = 1, \dots, K$

$r_{jk} \sim N(0, \sigma^2_{r-off})$ for $j, k = 1, \dots, K$, $j < K$

Goal: diffuse, symmetric prior

next $\{\bar\theta_0, V_{\bar\theta}, \nu_0, s_0^2\}$

joint posterior: $\pi(\bar\theta, r, \tau^2 \mid \{s_t, X_t\}_{t=1}^{T}) \propto L(\bar\theta, r, \tau^2) \times \pi(\bar\theta, r, \tau^2)$

$$= \prod_{t=1}^{T} \left(J^{-1}(s_t, X_t, r) \prod_{j=1}^{J} \phi\!\left(\frac{h^{-1}(s_t \mid X_t, \bar\theta, r)}{\tau} \right) \right) \times |V_{\bar\theta}|^{-\frac{1}{2}} e^{-\frac{1}{2}(\bar\theta - \bar\theta_0)' V_{\bar\theta}^{-1} (\bar\theta - \bar\theta_0)} \times \prod_{j=1}^{K} e^{\frac{-(r_{jj})^2}{2\sigma^2_{r-jj}}} \times \prod_{j=1}^{K-1} \prod_{k=j+1}^{K} e^{\frac{-(r_{jk})^2}{2\sigma^2_{r-off}}} \times (\tau^2)^{-\frac{\nu_0}{2}+1} e^{\frac{-\nu_0 s_0^2}{2\tau^2}}$$

$\theta^i \rightarrow$ sensitivity

$r_{ij} \rightarrow$ corr of sensitivity e.g. reason leave together

$\sigma_{r-jj} \rightarrow$ var of corr of sensitivity e.g. reason they do not live full time together

$\tau^2 \rightarrow$ variance of shock (e.g. shock could be fusion) or crisis in factory or media scandal

$\mu_{jt} \rightarrow$ shock and mean role of prod attrib

Hybrid algorithm

- two sets of conditional distribution
  ① pure Gibbs sampler, using standard natural conjugate reg
  ② for $\Sigma, r \rightarrow$ metropolis Hasting

$$\bar\theta, \tau^2 \mid r, \{s_t, X_t\}_{t=1}^{T}, \bar\theta_0, V_{\bar\theta}, \nu_0, s_0^2 \quad ①$$

$$r \mid \bar\theta, \tau^2, \{s_t, X_t\}_{t=1}^{T}, \sigma^2_{r-off}, \sigma^2_{r-jj} \quad ②$$

① Compute $\mu_{jt} \mid r$ & $\mu_{jt} = X_{jt}\bar\theta + \eta_{jt}$ $\quad \eta_{jt} \sim N(0, \tau^2)$

② Random walk metropolis chain (RW)

$$r^{new} = r^{old} + MVN(0, \sigma^2 D_r)$$

$\longrightarrow$ Candidate Cover matrix

$\hookrightarrow$ scaling constant

or reject candidates that are not positive definite (but problem in high dimension)

aggregate demand price elasticity

$$\varepsilon_{jk}(\underline{X} \mid \bar\theta, \Sigma, \tau) = \frac{\partial \ln s_j(\underline{X} \mid \bar\theta, \Sigma, \tau)}{\partial \ln P_k} = \frac{P_k}{s_j(\underline{X} \mid \bar\theta, \Sigma, \tau)}$$

$$\cdot \iint_{\theta^i, \eta} \frac{\partial}{\partial P_k} Pr(j \mid \underline{X}, \theta^i, \eta) \times \phi(\theta^i \mid \bar\theta, \Sigma)\, p(\eta \mid \tau)\, d\theta^i\, d\eta$$

where $\frac{\partial}{\partial P_k} Pr(j|\underline{X}, \theta^i, \eta) = \begin{cases} -\theta^i_{Price} Pr(j|\underline{X}, \theta^i, \eta) \cdot Pr(k|\underline{X}, \theta^i, \eta) & j \neq k \\ \theta^i_{Price} Pr(j|\underline{X}, \theta^i, \eta) \cdot (1 - Pr(j|\underline{X}, \theta^i, \eta)) & j = k \end{cases}$ (3)

$p(\eta|\tau) \rightarrow$ distribution of common demand shock

GMM - orthogonality conditions   moment condition

J number of alternatives - matrix Z_t $[J \times M]$ orthogonal to shock η_t : $E[Z_t' \eta_t]$ $[J \times 1]$

- # unique param $\dim(\bar{\theta}) + \dim(r) < M$: # moment conditions $\hookrightarrow$ not necessary

$$-\hat{m}_T(\bar{\theta}, \Sigma) = \frac{\sum^T}{T} Z_t'(\hat{\mu}_t(\Sigma) - X_t\bar{\theta})$$

GMM objective $g(\bar{\theta}, \Sigma) = \hat{m}_T(\bar{\theta}, \Sigma)' A^{-1} \hat{m}_T(\bar{\theta}, \Sigma)$

Consistent estimate of cov. matrix of moments $E[(Z_t' \eta_t)(Z_t' \eta_t)']$ $[M \times M]$ Weighting matrix

$$\frac{\partial \text{GMM obj}}{\partial \bar{\theta}} = 0 \Rightarrow \hat{\bar{\theta}}(\Sigma) = (X' Z A^{-1} Z' X)^{-1} X' Z A^{-1} Z' \hat{\mu}(\Sigma)$$

$[JT \times M]$ $[JT \times K]$

Two step GMM: **step 1** $A = \frac{\sum^T}{T} Z_t' Z_t$ minimize GMM obj func

get $\hat{\eta}_{jt}$

step 2 Construct $A = \frac{\sum}{T^2} Z_t' \hat{\eta}_t^{(1)} \hat{\eta}_t^{(1)'} Z_t$

minimize GMM obj func. after convergence run optimization routin with converge to make sure converged $\Rightarrow$ results:

$\hat{\Sigma}_{Gmm}, \hat{\bar{\theta}}_{Gmm}, \hat{\eta}_{jt}$

$\psi = (\bar{\theta}, r) \Rightarrow Var(\hat{\psi}_{Gmm}) = \frac{1}{T}(D' \Omega^{-1} D)^{-1} \Big|_{\psi = \hat{\psi}_{Gmm}}$

matrix of derivatives of the GMM criterion $g(\cdot)$ w.r.t. params $\hookrightarrow$ consistent estimate of the variance $\hat{m}_T$

$$D = \left[\frac{\partial m_T}{\partial \bar{\theta}} \quad \frac{\partial m_T}{\partial r}\right] \qquad \frac{\partial m_T}{\partial \bar{\theta}} = -\frac{\sum_t Z_t' X_t}{T} \qquad \frac{\partial m_T}{\partial r} = \frac{\sum_t}{T} Z_t' \left(\frac{\partial \hat{\mu}_t}{\partial r}\right)$$

μ_t mean utility $\rightarrow$ it's derivative w.r.t. r computed numerically

example $J = 4$ goods: 3 inside, 1 outside

- 3 brand intercept
- log price $\sim$ unif(0,1)
- Neg brand intercept: outside good share is large relative to any given product

$\bar{\theta} = (-2, -3, -4, -5)$ $\hookrightarrow$ log price coeff

$$\Sigma = \begin{bmatrix} 3 & 2 & 1.5 & 1 \\ & 4 & -1 & 1.5 \\ & & 4 & -0.5 \\ & & & 3 \end{bmatrix}$$

$\rightarrow$ relative large neg correl b/w log price coeff & brand intercepts

$T = 300$ (the # of observations) $\hookrightarrow$ informative data regarding Σ

5 experiment cells $\hookrightarrow$ 50 data sets each   observable price (fixed)

$V_{jt} = e^{-.5413 + P_{jt}}$ $\rightarrow$ log price, for method $E(V_{jt}) = 1$

draw $\eta \sim N(0, \tau^2)$ $\tau^2 = 1$

for conditional heteroskedastic case: $\eta_{jt} \sim N(0, \tau^2 V_{jt})$

- assumption of indep aggreg. shocks questionable when short time intervals (eg week or month) $\Rightarrow$ AR(1)

$\Rightarrow \begin{cases} \eta_{jt} = \rho \eta_{j,t-1} + u_{jt} & AR(1) \\ u_{jt} \sim N(0, .4359^2) \end{cases}$ $\hookrightarrow$ to make sure $Var(\eta_{jt}) = 1$

(4)

shock $\sim$ Beta dist (to depart from normal)

asymmetric: $\eta_{jt} \sim 3.3135 \times Beta(2, 5) - .8949$

Symmetric: $\eta_{jt} \sim 1.4142 \times Beta(.5, .5) - .7071$

MCMC

params $\begin{cases} \Sigma: \text{sensitivity variance} \\ \bar{\theta}: \text{mean of sensitivity} \\ \tau^2: \text{variance of shock} \end{cases}$ $\theta^i = \bar{\theta} + v_i$, $v_i \sim N(0, \Sigma)$

Prior: $\bar{\theta} \sim MVN(\bar{\theta}_0, V_{\bar{\theta}})$

$r_{jj} \sim N(0, \sigma_j^2)$ $r_{jk} \sim N(0, \sigma_{off}^2)$ $j \neq k$

$\tau^2 \sim \nu_0 s^2 / \chi^2_{\nu_0}$ $\Sigma = U'U$, $U = \begin{bmatrix} e^{r_{11}} & r_{12} & \cdots & r_{1k} \\ 0 & e^{r_{22}} & & \vdots \\ 0 & \cdots & 0 & e^{r_{kk}} \end{bmatrix}$

Posterior

$$\propto \prod_t \cancel{\bar{\theta}, r, \tau^2}\ \pi(S_t | \bar{\theta}, r, \tau, X_t) \times Prior(\bar{\theta}, r, \tau)$$

likelihood

reparameterization

$$\Sigma = U'U = \begin{bmatrix} e^{2r_{11}} & r_{12}e^{r_{11}} & r_{13}e^{r_{11}} & r_{14}e^{r_{11}} \\ & r_{12}^2 + e^{2r_{22}} & r_{12}r_{13} + r_{23}e^{r_{22}} & r_{12}r_{14} + r_{24}e^{r_{22}} \\ \text{Symmetry} & & r_{13}^2 + r_{23}^2 + e^{2r_{33}} & r_{13}r_{14} + r_{23}r_{24} + r_{34}e^{r_{33}} \\ & & & r_{14}^2 + r_{24}^2 + r_{34}^2 + e^{2r_{44}} \end{bmatrix}$$

Priors $r_{jj} \sim N(0, \sigma_j^2)$

$r_{jk} \sim N(0, \sigma_{off}^2)$ $j \neq k$

$\Rightarrow$ prior variance & mean on diagonals of Σ $(k = 1, 2, 3)$

$Var(\Sigma_{kk}) = 2(k-1)\sigma_{off}^4 + e^{8\sigma_k^2} - e^{4\sigma_k^2}$

$E(\Sigma_{kk}) = (k-1)\sigma_{off}^2 + e^{2\sigma_k^2}$

$Var[\Sigma_{k,k+1}] = (k-1)\sigma_{off}^4 + \sigma_{off}^2 e^{2\sigma_k^2}$

$E(\Sigma_{k,(k+1)}) = 0$

Goal $\rightarrow$ diagonals of Σ have same prior variance

i.e. $Var[\Sigma_{11}] = Var[\Sigma_{22}] = Var[\Sigma_{33}] = Var[\Sigma_{44}]$

MCMC Algorithm

1. $r | \bar{\theta}, \tau^2, \{S_t, X_t\}_{t=1}^T, \sigma^2_{r\text{-diag}}, \sigma^2_{r\text{-off}}$
2. $\bar{\theta}, \tau^2 | r, \{S_t, X_t\}_{t=1}^T, \bar{\theta}_0, V_{\bar{\theta}}, \nu_0, s_0^2$

[1] Random walk metropolis chain for r

$r^{new} = r^{old} + MVN(0, \sigma^2 D_r)$

[2] Gibbs sampler for $\bar{\theta}, \tau^2$

univar Bayes reg $\mu_{jt} = X_{jt}\bar{\theta} + \eta_{jt}$, $\eta_{jt} \sim N(0, \tau^2)$

(normal reg Gibbs sampler)

- Shock of demand to different people & they respond differently to it (different sensitivity)

τ: response variance (or sensitivity variance)

- Hessian is covar matrix of sensitivity

intuition of Bayesian

1. you use same formula, but in each step prob of picking true param increases by correction, & then propagates to whole data to move you closer

① $\mu_{jt} = x_{jt}\bar{\theta} + \eta_{jt}$ Gibbs Sampler

$\eta_{jt} \sim N(0, \tau^2)$

updated τ^2: Taosq
updated $\bar{\theta}$: thetabar1

② $r_{ij} \sim N(0, \sigma_j^2)$ → mean of sensitivity $\mu_{jt} \equiv S_{jt}$

inver fun → to do Contraction mapping

Var-Cov of sensitivity

elements of Jacobian: v

$\int f(s_{ijt})\, d\Phi(\theta_i|\bar{\theta},\Sigma) = \int f\left(\frac{e^{\mu_{jt} + x_{jt} v_i}}{1 + \sum e^{\mu_{kt} + x_{kt} v_i}}\right) d\Phi(v_i|0,\Sigma)$

H: person with different sensitivity (i) # random coeff = K

T: time (t) → change in attribute

J: Brand (j) → utility of which brand

X_{jt} [1×K] vector → person

repitition?

$\bar{\theta} \sim MVN(\theta_0, V_{\bar{\theta}})$

X → [T×J, K]

~~Gen v → H~~ ~~vector person~~

→ prior needs to be revised

iter & saving

to save thetat = thetabar

tbeot
rt

sh → $\hat{s}$
S → real share

① ~~G: aaa?~~ ✓   ③ ~~sigmasqR DIAG?~~ ✓
② ~~Jacobian?~~ ✓   ④ ~~sigmasqR_off?~~ ✓
③ ~~normaldetarm?~~

$\prod_{t=1}^{T} (J^{-1}(s_t, x_t, r) \prod_{j=1}^{J} \frac{1}{\tau} \phi\left(\frac{h(s_t|X_t,\bar{\theta},r)}{\tau}\right)$

① Data? loop ② signal

② Debug?

Caution: without instrument

⓪ ~~make sure functions have enough params~~ ✓

① V_{Σ} ~~variance of effect~~ → random effect which we are searching to find variance of

X → depend
S → indep

~~H×K~~

H?

~~K×J~~   [J×T K] × [K K] [K H]

mu   [J×T]   → mu vertical vect [NT×1]

Sum in item by item

Thetabar   [X] [J T × K]   [K H]

[K × 1]

① Debug Data generating process
② add iteration & saving
③ Debug the estimation