Bayesian analysis for microeconometric models of discrete choice variables

Shinya Sugawara ee077007@mail.ecc.e.u-tokyo.ac.jp

Graduate school of Economics, University of Tokyo
and DC2 Research Fellowship of Japan Society of Promotion of Science

Dec 2011
Chapter 1: Overview

Chapter 2: Duopoly in the Japanese Airline Market: Bayesian Estimation for the Entry Game

Chapter 3: Separate inference for moral hazard and selection problems in insurance markets: With application to US dental insurance market

Chapter 4: Effects of an intervention for an inefficient Nash equilibrium: Analysis of Japanese private nursing home market via a nonparametric Bayesian approach
Active study of econometrics for discrete variables
- Nobel prize for Daniel McFadden
- Surveys: Maddala(1983), Train(2003) and many others
- New economic theory $\Rightarrow$ Request for new econometric technique
Multiple discrete variables with mutual dependencies

Dependent variables: $y_1, \ldots, y_J$: Some of them are discrete:

$$y_j = f_j(y_1, \ldots, y_{j-1}, y_{j+1}, \ldots, y_J).$$

Need one more step to determine a model:

- **Simultaneous** specification
- **Conditional** specification

Which one? $\Rightarrow$ Depends on situation.
Model: Simultaneous specification

\[ y_j = f_j(y_1, \ldots, y_{j-1}, y_{j+1}, \ldots, y_J). \]

- For all \( j = 1, 2, \ldots, J \) simultaneously.
- Need additional assumption to guarantee the simultaneity.
- In econometrics, it is often market equilibrium.
- \( \Rightarrow \) Active econometric literature of simultaneous equation models.
Model: Conditional specification

\[ y_j = f_j(y_1, \ldots, y_{j-1}, y_{j+1}, \ldots, y_J). \]

- The conditional model of \( y_j | y_1, \ldots, y_{j-1}, y_{j+1}, \ldots, y_J \), separately.
- Need the joint distribution to have a well-defined statistical model.
- Its recoverability, called compatibility, places serious restrictions for feasible models.
Complicate econometric models for discrete variables often have difficulty in evaluating the likelihood function.

⇒ **Bayesian statistics**
- Gelfand and Smith(1990): *Markov chain Monte Carlo* (MCMC)
- Albert and Chib(1993): Binary choice
- McCulloch and Rossi(1994): Multinomial probit
- Surveys: Koop(2003), Koop, Poirier and Tobias(2007)
Chapter 2

Duopoly in the Japanese Airline Market: Bayesian Estimation for the Entry Game (Japanese Economic Review, forthcoming)

1. Introduction
2. Econometrics
3. Empirical study
History of Japanese airlines

- Infant industry protection by aviation constitution in 1972:
  - JAL: International and domestic main routes
  - ANA: Domestic main and local routes
  - JAS (former Toa): Domestic local routes (Merged to JAL in 2001)
- Japanese open sky deregulation from 1985: Only on a halfway?
  - Duopoly of JAL and ANA: More than 90% of share
  - No severe price competition

$\Rightarrow$ Main question: Is the Japanese airline market competitive?
Part II: Econometrics

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>( x'<em>1m \beta_1 + \Delta_1 + u</em>{1m} ), ( x'<em>2m \beta_2 + \Delta_2 + u</em>{2m} )</td>
<td>( x'<em>1m \beta_1 + u</em>{1m} ), 0</td>
</tr>
<tr>
<td>1</td>
<td>0, ( x'<em>2m \beta_2 + u</em>{2m} )</td>
<td>0,0</td>
</tr>
</tbody>
</table>

- **Players:** \( i = 1, 2 \): JAL and ANA
- **Markets:** \( m = 1, 2, \ldots, M \): Route between two airports (ex. Kumamoto - Haneda)
- **Strategies:** \( y_{im} = 0, 1 \):
  - 1: Entrance of \( i \)th player to \( m \)th market
  - 0: No entrance
- **Payoff functions:**
  - For no entrant, 0
  - For entrant, \( x'_im \beta_i + u_{im} \) for sure, plus \( \Delta_i \) if counterpart is present.
Complete information: Everything is known to both players

Payoff components

- $x$: Observed for econometricians, explanatory variables
- $u$: Unobserved for econometricians, error term
- $\theta = (\beta, \Delta)$: Unobserved for econometricians, parameters to be estimated
- $\Delta_i$: Measurement of strategic interaction: Main target
  - $\Delta_i < 0$: Presence of counterpart is damaging: Substitutive competition
  - $\Delta_i > 0$: Presence of counterpart is beneficial
  - $\Delta_1 < 0, \Delta_2 < 0$: Strategic substitution: US airline market. Competitive
Pure Nash equilibrium and simultaneous specification

Pure Nash equilibrium
- Pair of \((y_1m, y_2m)\), from which nobody deviates by himself.
- Best response

\[
\begin{align*}
y_{1m} &= I[x_1m\beta_1 + \Delta_1 y_{2m} + u_{1m} \geq 0] \\
y_{2m} &= I[x_2m\beta_2 + \Delta_2 y_{1m} + u_{2m} \geq 0]
\end{align*}
\]

Nash: Two best response equations hold simultaneously
⇒ Simultaneous specification is required as a statistical model.

i.e.

\[
\begin{align*}
\{z_m = 1\} &\iff \{u_{1m} < -x_1m\beta_1, u_{2m} < -x_2m\beta_2\}, \\
\{z_m = 2\} &\iff \{u_{1m} \geq -x_1m\beta_1, u_{2m} < -x_2m\beta_2 - \Delta_2\}, \\
\{z_m = 3\} &\iff \{u_{1m} < -x_1m\beta_1 - \Delta_1, u_{2m} \geq -x_2m\beta_2\}, \\
\{z_m = 4\} &\iff \{u_{1m} \geq -x_1m\beta_1 - \Delta_1, u_{2m} \geq -x_2m\beta_2 - \Delta_2\}.
\end{align*}
\]

where \(z_m = 1, 2, 3, 4 \iff (y_{1m}, y_{2m}) = (0, 0), (1, 0), (0, 1), (1, 1)\)
Model A: Non-unique DGP in Region 5: Multiple Nash
Model B: No reasonable DGP in Region 5: No pure Nash
Anyway, $\sum_l Pr(z_m = l | x, \theta) \neq 1$: Incoherency problem (Gourieroux et al, 1980)
Tamer(2003): Selection rule $p_{ml}$: Unobserved proportion of Region 5 where $z_m = l$ realizes.

Choice probability given selection rule:

$$Pr[z_m = l|\mathbf{x}_m, \theta, p_{ml}] = P_l(\mathbf{x}_m, \theta) + p_{ml} P_5(\mathbf{x}_m, \theta), \ l = 1, 2, 3, 4,$$

where $P_l(\mathbf{x}, \theta)$: Area of Region $l$.

$p_{ml}$ is $m$-specific nuisance parameter $\Rightarrow$ No consistent point estimator for $\theta$ 
$\Rightarrow$ Ciliberto and Tamer(2009): Set estimation
Main technique: Estimate sample-specific selection rule $p_m$ using information of $m$th sample

- Bayesian estimator is well-defined for small sample
- Small sample size $\Rightarrow$ Posterior should be similar to prior, but estimation is standard.

Advantage: Flexible inference

- Availability of model selection technique
  - Detection for type of strategic interaction.
  - Adaptation of more DGPs, like mixed-strategy Nash (Abbreviated today)

Prediction:

- Prediction probabilities for players' strategies at new route given $x_{new}$:

  $$Pr(z_{new} = l|\theta, x_{new}, p_{new}) = P_l(\theta, x_{new}) + p_{l,new}P_5(\theta, x_{new}).$$

- Upper and lower bounds:

  $$P_l(\theta, x_{new}) \leq Pr(z_{new} = l|\theta, x_{new}, p_{new}) \leq P_l(\theta, x_{new}) + P_5(\theta, x_{new}).$$
Hierarchical Bayesian modeling

- Add latent dummy variables correspond to \( p_m \)
  - Case A: \( \lambda_m \sim \text{Bernoulli}(p_m) \)
  - Case B: \( \lambda_m = (\lambda_{1m}, \lambda_{2m}, \lambda_{3m}, \lambda_{4m}) \sim \text{Multinomial}(1, p_m) \)

Prior distributions

- Beta (Model A) or Dirichlet (Model B) for \( p \): Specifically, uniform.
- Normal for \( \beta \)
- Truncated normal for \( \Delta \)

Likelihood function

- Distributional assumption: \( u_m \sim \mathcal{N}(0, 1) \)
- Any distributional forms can work, but we need some, to have a closed form of \( P \)

Markov Chain Monte Carlo algorithm

1. Generate \( p|\beta, \Delta, \lambda, z \): Gibbs sampler.
2. Generate \( \lambda|\beta, \Delta, p, z \): Gibbs sampler.
3. Generate \( \beta|\Delta, p, \lambda, z \): Metropolis-Hastings algorithm.
4. Generate \( \Delta|\beta, p, \lambda, z \): Metropolis-Hastings algorithm.
Part III: Empirical Analysis

Data:

- $y$: Entrance
  - 2007 February to March, taken from JR Time Table
  - 39 airport excluding isolated islands. 741 markets as their combinations
- $x$: Explanatory variables for each route
  - Distance, Population, Number of flights of the player from the two airports, Shinkansen bullet train dummy
## Model selection: Result

<table>
<thead>
<tr>
<th>Model</th>
<th>With mixed Nash</th>
<th>Only pure Nash</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\Delta_1 &lt; 0$</td>
<td>$\Delta_1 &lt; 0$</td>
</tr>
<tr>
<td></td>
<td>$\Delta_1 &gt; 0$</td>
<td>$\Delta_1 &gt; 0$</td>
</tr>
<tr>
<td></td>
<td>$\Delta_2 &lt; 0$</td>
<td>$\Delta_2 &lt; 0$</td>
</tr>
<tr>
<td></td>
<td>$\Delta_2 &gt; 0$</td>
<td>$\Delta_2 &gt; 0$</td>
</tr>
<tr>
<td>Likelihood</td>
<td>-530.29</td>
<td>-432.41</td>
</tr>
<tr>
<td></td>
<td>-533.36</td>
<td>-415.65</td>
</tr>
<tr>
<td></td>
<td>-416.02</td>
<td>-410.68</td>
</tr>
<tr>
<td></td>
<td>-418.11</td>
<td>-428.65</td>
</tr>
<tr>
<td>Prior</td>
<td>-28.68</td>
<td>-28.68</td>
</tr>
<tr>
<td></td>
<td>-28.75</td>
<td>-28.60</td>
</tr>
<tr>
<td></td>
<td>-28.59</td>
<td>-31.94</td>
</tr>
<tr>
<td></td>
<td>-28.60</td>
<td>-31.94</td>
</tr>
<tr>
<td>Posterior</td>
<td>29.19</td>
<td>29.19</td>
</tr>
<tr>
<td></td>
<td>20.75</td>
<td>24.88</td>
</tr>
<tr>
<td></td>
<td>24.25</td>
<td>12.16</td>
</tr>
<tr>
<td></td>
<td>25.05</td>
<td>12.16</td>
</tr>
<tr>
<td>ML</td>
<td>-588.16</td>
<td>-490.28</td>
</tr>
<tr>
<td>(S.E.)</td>
<td>(0.06)</td>
<td>(0.05)</td>
</tr>
<tr>
<td></td>
<td>(1.00)</td>
<td>(0.10)</td>
</tr>
<tr>
<td></td>
<td>(0.15)</td>
<td>(0.24)</td>
</tr>
<tr>
<td></td>
<td>(0.10)</td>
<td>(0.36)</td>
</tr>
</tbody>
</table>

Marginal likelihood of Chib and Jeliazkov (2001)’s method
Selected Model: Only with pure Nash equilibrium, $\Delta_{ANA} < 0$ and $\Delta_{JAL} > 0$

- Entry of JAL reduces payoff of ANA, while the entry of ANA increases payoff of JAL.
- ANA is leader and JAL is follower: Initial advantage of ANA in the domestic market allows construction of more beneficial networks even after deregulation.
- $\Rightarrow$ Remaining influence of regulation scheme:
- Similar to asymmetric competition of Kmart and Wal-mart in Jia(2008, ECA)
Prediction: Shizuoka airport

- Established in 2009
- Negative forecast
  - Close to Tokyo and Nagoya
  - Located on route of Shinkansen bullet train
- Planned and operated routes
  - ANA: Shin-Chitose and Naha
  - JAL: Shin-Chitose and Fukuoka
  - Requested in 2007: Narita, Komatsu, Matsuyama and Kagoshima
  - Local company Fuji Dream Airline operates to Shin-Chitose, Fukuoka, Kumamoto and Kagoshima
    - Route to Komatsu and Matsumoto were once operated but have been abandoned since March 2011.
    - Route to Kumamoto is under consideration to be abandoned.
No route is attractive, except to Shin-Chitose.
Conclusion

- We established a Bayesian estimation method for entry game estimation
- Our advantage is the flexibility: We presented three inferential techniques:
  - Model selection for form of strategic interaction
  - Extension for mixed Nash model
  - Prediction analysis
- Our empirical result shows
  - Japanese open-sky deregulation has not yet vanished the influence of the former regulation scheme
  - Newly established Shizuoka airport will suffer from the lack of demand.
Separate inference for moral hazard and selection problems in insurance markets: With application to US dental insurance market

1. Introduction and literature review
2. Our econometric models
3. Empirical application
### Insurance under complete information

- **Complete information**: An insurer exactly knows risk of accident and potential loss of each consumer.
- For insurer, realized loss converges to expected loss, as \( \# \) consumers \( \to \infty \):
  - Law of large numbers
- \( \Rightarrow \) Insurance price \( = \) Expected loss.
- \( \Rightarrow \) No relationship between risk and purchase decision of each consumer.

### Insurance under asymmetric information (AI)

- **Adverse selection (AS)**: Risky consumers have larger demand for insurance, but insurers cannot observe consumers’ risk.
- **Moral hazard (MH)**: Insured consumer reduces efforts to avoid an accident. This change is not predicted by insurers.
- \( \Rightarrow \) \( \exists \) Relationship between risk and purchase.
Bivariate probit model

Chiappori and Salanie (2000, JPE)

\[ y_{i1}^{*} = x_{i1}' \beta_1 + \epsilon_{i1}, \]
\[ y_{i2}^{*} = x_{i2}' \beta_2 + \epsilon_{i2}, \]
\[ y_{ij} = I[y_{ij}^{*} \geq 0], \]

where

- \( i = 1, 2, ..., N \): Individuals
- \( j = 1, 2 \).
  - \( y_{i1} \): Insurance purchase dummy (Observable)
  - \( y_{i2} \): Accident occurrence dummy (Observable)

Testable implication of AI: Positive correlation \( \text{Corr}(y_{i1}^{*}, y_{i2}^{*} | \boldsymbol{x}_i) > 0 \)

- MH: Insurance purchase \( \Rightarrow \) Accident risk \( \uparrow \)
- AS: High risk \( \Rightarrow \) High insurance demand
Problem of bivariate probit

### Empirical studies
- Flexible reduced form without detailed model for consumer’s behavior
- ⇒ Many empirical applications
- BUT a positive correlation is not often found

### Advantageous selection
- Counter theory by de Meza and Webb (2001, RANDJ): **Advantageous selection** (ATS)
  - Less risky people purchase more insurance, because of risk-averse preference
- A rationale of no-positive correlation: MH(Positive) and ATS(Negative) cannot be captured by one parameter
- Is it testable? ⇒ Need to identify MH and AS / ATS separately: This is beyond the bivariate probit model
- ⇒ This study
Part II: Our econometric models

Basic setup

\[ y_{i1}^* = x_{i1}' \beta_1 + \alpha_1 y_{i2} + u_{i1}, \]
\[ y_{i2}^* = x_{i2}' \beta_2 + \alpha_2 y_{i1} + u_{i2}, \]
\[ y_{ij} = I[y_{ij}^* \geq 0], \]

Say \( \theta = \{ \beta_1', \beta_2', \alpha_1, \alpha_2 \}' \)

- Separate parametrization
  - \( \alpha_2 y_{i1} \): Effect of insurance purchase on accident risk: MH (\( \alpha_2 > 0 \))
  - \( \alpha_1 y_{i2} \): Effect of accident occurrence on insurance demand: AS (\( \alpha_1 > 0 \)) or ATS (\( \alpha_1 < 0 \))

- Mutual dependencies \( \Rightarrow \) Conditional or simultaneous?
Conditional specification

\[ y_{ij} = I[\mathbf{x}'_{ij}\beta_j + \alpha_j y_{ik} + u_{ij} \geq 0], k \neq j \]

- Above equations specify conditional distributions \( Y_1|Y_2 \) and \( Y_2|Y_1 \)
- Compatibility condition
  - Besag(1974)'s theorem \( \Rightarrow \) Compatibility condition is
    \[ Pr(y_j|y_k) = \frac{y_j\{\exp[\mathbf{x}'\beta_j + \gamma y_k]\}}{1 + \exp[\mathbf{x}'\beta_j + \gamma y_k]}, \quad j = 1, 2, \ k \neq j \]
  - Common \( \gamma \) for \( j = 1, 2 \)
  - \( \Rightarrow \) No separate identification for MH and AS / ATS
Simultaneous specification

\[ y_{i1} = I[x'_{i1}\beta_1 + \alpha_1 y_{i2} + u_{i1} \geq 0], \]
\[ y_{i2} = I[x'_{i2}\beta_2 + \alpha_2 y_{i1} + u_{i2} \geq 0]. \]

Justification of simultaneity: \textbf{Equilibrium} of optimization of consumers

1. Before actual purchase, consumers consider whether to purchase insurance, taking account of their future accident occurrence\(\text{(No uncertainty)}\)
2. Re-evaluate accidental risk based on the purchase decision
3. \ldots
4. Consumers’ behaviors are decided at steady state, where \(y_s\) before and after these updates equate

\(\Rightarrow\) Same econometric model as entry games
Intuition for specifications

One observed pair to identify two parameters
Model selection

We use model selection to detect true information structure among:

- **Model PP**: $\alpha_1 > 0, \alpha_2 > 0$: Adverse selection and moral hazard
- **Model NP**: $\alpha_1 < 0, \alpha_2 > 0$: Advantageous selection and moral hazard
- **Model PN**: $\alpha_1 > 0, \alpha_2 < 0$: Adverse selection and opposite of moral hazard
- **Model NN**: $\alpha_1 < 0, \alpha_2 < 0$: Advantageous selection and opposite of moral hazard
- **Model CI**: $\alpha_1 = 0, \alpha_2 = 0$: Complete information
- **Model MH**: $\alpha_1 = 0$: Only moral hazard
- **Model AS**: $\alpha_2 = 0$: Only adverse or advantageous selection
Part III: Empirical Analysis

- **Target:** Dental care insurance in US
  - Previous studies detected moral hazard, but selection problems have not been analyzed.

- **Data:** 2004-2005 Medical Expenditure Panel Survey

- **Insurance Purchase** $y_1$
  - Stable coverage: About 50%

- **Accident** $y_2$
  - (at least a) Visit to a doctor: About 50%
  - Moral hazard: Insured consumers are more likely to visit doctors

- **Explanatory variables** $x$
  - Demographic, Health status, Geographic, Firm size (Only for insurance purchase)

- **Sample size**
  - $N = 5090$: Privately-employed workers who are 25 to 65 years old
  - Also analyze 10% highest and lowest income groups
## Bivariate probit result

| Model          | Mean of Corr($y_{1i}^*, y_{2i}^* | x$) | S.D.   | 95% interval          |
|----------------|----------------------------------|-------|-----------------------|
| Whole sample   | 0.129 (0.020)                    | [ 0.089, 0.168] |
| High-income    | 0.054 (0.063)                    | [-0.072, 0.178]  |
| Low-income     | 0.139 (0.062)                    | [ 0.017, 0.259]  |

**Table:** Bivariate probit model: Posterior means, standard deviations and 95% credible intervals of the correlation Corr($y_{1i}^*, y_{2i}^* | x$)

The bivariate probit finds:

- Whole sample and low-income group: Positive correlation (Moral hazard, adverse selection)
- High income: No correlation: Complete information or MH and ATS combination
- A typical interpretation: Rich people are so risk-averse that they exhibit ATS. It does not hold for poor people

But our methodology says...
Model selection

<table>
<thead>
<tr>
<th>Models</th>
<th>Whole Sample</th>
<th>High-income</th>
<th>Low-income</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bivariate probit model</td>
<td>12335</td>
<td>1170.2</td>
<td>1113.8</td>
</tr>
<tr>
<td>Simultaneous models with selection rules</td>
<td>12406</td>
<td>1173.5</td>
<td>1127.6</td>
</tr>
<tr>
<td>NN</td>
<td>12289</td>
<td>1158.3</td>
<td>1100.4</td>
</tr>
<tr>
<td>NP</td>
<td>12319</td>
<td>1175.2</td>
<td>1111.6</td>
</tr>
<tr>
<td>PN</td>
<td>12299</td>
<td>1169.3</td>
<td>1109.2</td>
</tr>
<tr>
<td>Simultaneous models with coherency restriction</td>
<td>12403</td>
<td>1170.4</td>
<td>1123.2</td>
</tr>
<tr>
<td>CI</td>
<td>12307</td>
<td>1169.7</td>
<td>1110.6</td>
</tr>
<tr>
<td>AS</td>
<td>12327</td>
<td>1170.8</td>
<td>1107.7</td>
</tr>
</tbody>
</table>

Table: DIC (Spiegelhalter et al, 2002)

- The best model is with moral hazard and advantageous selection for each samples
- Everybody shows ATS and moral hazard.
- ⇒ Bivariate probit results are misleading: The two distinct mechanisms are not captured by only one correlation parameter.
Alternative parametrization

\[ y_{i1}^* = x_{i1}\beta_1 + \alpha_1 y_{i2}^* + u_{i1}. \]
\[ y_{i2}^* = x_{i2}\beta_2 + \alpha_2 y_{i1} + u_{i2}. \]
\[ y_{ij} = I[y_{ij}^* \geq 0]. \]

- Original parametrization: Reverse causality (Accident occurrence affects insurance purchase decision.)
- This parametrization: Accidental risk affects insurance purchase decision
Conclusion

- We have proposed a simple econometric method for insurance market
- We have showed that simultaneity assumption is necessary when we have only cross-section data
- We have found evidence of moral hazard and advantageous selection in US dental insurance market, even for low income households, unlike the previous studies.
Chapter 4

Effects of an intervention for an inefficient Nash equilibrium: Analysis of the Japanese private nursing home market via a nonparametric Bayesian approach

1. Introduction
   - Peculiar price mechanism in Japanese private nursing home market

2. Model

3. Econometrics
   - What happens if the custom is vanished?: Simulation study
   - For simulation, GMM does not work ⇒ Put distributional assumption
   - To minimize cost of assumption, use nonparametric Bayesian method called Polya tree mixture

4. Empirical Analysis
   - Data: Shuukan Asahi Mook
Part I: Market

- Nursing homes: Institutions to provide a permanent residence for elders, not specializing in a medical care
- Coexistence of public and private sector
  - The Long-Term Care Insurance (LTCI) covers both
  - Waiting list of 430,000 elders for public homes
  - Our target: Private nursing homes in which the long-term care is provided as a default option

See also P.51

<table>
<thead>
<tr>
<th>Year</th>
<th>1999</th>
<th>2000 (LTCI)</th>
<th>2001</th>
<th>2002</th>
<th>2003</th>
</tr>
</thead>
<tbody>
<tr>
<td>Private # Homes</td>
<td>298</td>
<td>350</td>
<td>400</td>
<td>508</td>
<td>694</td>
</tr>
<tr>
<td>Capacities</td>
<td>32302</td>
<td>37467</td>
<td>41445</td>
<td>46561</td>
<td>56837</td>
</tr>
<tr>
<td>Public # Homes</td>
<td>4214</td>
<td>4463</td>
<td>4651</td>
<td>4870</td>
<td>5084</td>
</tr>
<tr>
<td>Capacities</td>
<td>283822</td>
<td>298912</td>
<td>314192</td>
<td>330916</td>
<td>346069</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Year</th>
<th>2004</th>
<th>2005</th>
<th>2006</th>
<th>2007</th>
<th>2008</th>
<th>2009</th>
</tr>
</thead>
<tbody>
<tr>
<td>Private # Homes</td>
<td>1045</td>
<td>1406</td>
<td>1968</td>
<td>2671</td>
<td>3400</td>
<td>3565</td>
</tr>
<tr>
<td>Capacities</td>
<td>76128</td>
<td>96412</td>
<td>123155</td>
<td>147981</td>
<td>176935</td>
<td>18324</td>
</tr>
<tr>
<td>Public # Homes</td>
<td>5291</td>
<td>5535</td>
<td>5716</td>
<td>5892</td>
<td>6015</td>
<td>5876</td>
</tr>
<tr>
<td>Capacities</td>
<td>363747</td>
<td>383326</td>
<td>399352</td>
<td>412807</td>
<td>422703</td>
<td>41466</td>
</tr>
</tbody>
</table>
Need to pay both monthly fee and initial payment

- Monthly fee $p$: For daily needs and long-term care costs
- Initial payment $F$: For rents during an expiration period $T$, predetermined by the home

Property of initial payment

- Exit before the expiration $\rightarrow$ rest of rents are paid back
- Stay longer than the expiration $\Rightarrow$ No additional rent

$\Rightarrow$ Homes cover all longevity risks

It has been established during the bubble economy, when the private nursing home is a luxury good and consumers are generous for high initial payments with long expiration periods.

In our data, 35% of homes offer the contract without an initial payment, but 65% still receive the initial payment.

We guess welfare of most consumers would be reduced when the initial payment is vanished.
Part II: Model

- Berry et al. (1995): Simultaneous demand and supply
- $m = 1, 2, \ldots, M$: Markets, each has demand and supply sides
- Demand side: $i_m = 1, 2, \ldots, I_m$ Potential consumers
  - Consumers’ utility maximization (Multinomial choice)
  - Utility function can be estimated without observing individual-level variables, but only with aggregate sales data of homes
- Supply side: $h_m = 1, 2, \ldots, H_m$ homes. $h_m = 0$ is a outside option
  - Homes’ profit maximization through Bertrand competition
  - We extend the model to have two control variables, $p$ and $T$.
  - We assume the monthly rent, say $f$, is an exogenous variable which is determined in a housing market
- $H = \sum H_m$: Sample size. Markets are pooled.
- $I_m$: Unobserved
Private information: \( \tau_{im} \): Lifetime, \( \delta_{im} \): Time-discount factor. Unknown to homes

\[
U_{imh_m} = \sum_{t=1}^{\tau_i} \delta_{im}^{t-1} [\check{x}'_{hm} \tilde{\beta}_d + \xi_{hm} + \tilde{\eta}_{imh_m}] - P(p_{hm}, f_{hm}, T_{hm}, \tau_{im}, \delta_{im}) \alpha,
\]

\( P: \) the present value of the payment stream:

\[
P(p_{hm}, f_{hm}, T_{hm}, \tau_{im}, \delta_{im}) = \sum_{t=1}^{\tau_{im}} \delta_{im}^{t-1} p_{hm} + F_{hm} - I[T_{hm} \geq \tau_{im}] \delta_{im}^{\tau_{im}} (T_{hm} - \tau_{im}) f_{hm}.
\]

The payment stream consists of:

1. Present value of monthly fees
2. Initial payment
3. Returns of initial payment

We assume no voluntary exit.
After several manipulation, we have that consumer chooses the $h_m$th home when $V_{i_m h_m} = \max_{k_m} V_{i_m k_m}$ where

$$V_{i_m h_m} = x'_{h_m} \beta_d - p_{h_m} \alpha + \xi_{h_m} + \eta_{i_m h_m}.$$

Suppose $\eta_{i_m h_m}$ follows i.i.d. type I extreme value distribution and integrate it out, we have

$$s_{h_m} = \begin{cases} \exp[x'_{h_m} \beta_d - p_{h_m} \alpha + \xi_{h_m}] / (1 + \sum_{k_m=1}^{H_m} \exp[x'_{k_m} \beta_d - p_{k_m} \alpha + \xi_{k_m}]) & \text{for } h_m = 1, \ldots, H_m \\ 1 / (1 + \sum_{k_m=1}^{H_m} \exp[x'_{k_m} \beta_d - p_{k_m} \alpha + \xi_{k_m}]) & \text{for } h_m = 0 \end{cases}.$$

$s_{h_m}$ is a share of the $h_m$th home, defined as the number of residents in $h_m$ over the potential consumers $I_m$ in the $m$th market.

Say

$$q_{h_m} \defn \ln s_{h_m} = x'_{h_m} \beta_d - p_{h_m} \alpha + \xi_{h_m} + \ln \left(1 - \sum_{k_m=1}^{H_m} s_{k_m}\right), \ h_m = 1, 2, \ldots, H_m.$$
Supply side

- Home’s one-shot expected (wrt. \( \tau \)) profit at steady state:

\[
\Pi_{hm} = I_m s_{hm} \left[ p_{hm} + f_{hm} - mc_{hm} - f_{hm} Pr^s(T_{hm} < \tau | f_{hm}) \right].
\]

\( mc \): a marginal cost:

\[
\log(mc_{hm}) = w_{hm} \beta_s + \omega_{hm}
\]

and \( Pr^s(T_{hm} < \tau | f_{hm}) \) is home’s subjective probability, for which we define

\[
\Gamma(T_{hm}, f_{hm}; \gamma) \stackrel{def}{=} 1 - Pr^s(T_{hm} < \tau | f_{hm}) \text{ s.t.}
\]

\[
\Gamma(T_{hm}, f_{hm}; \gamma) \stackrel{def}{=} \frac{\exp(\gamma_0 + \gamma T_1 T_{hm} + \gamma T_2 T_{hm}^2 + \gamma f_1 f_{hm} + \gamma f_2 f_{hm}^2 + \gamma TF T_{hm} f_{hm})}{1 + \exp(\gamma_0 + \gamma T_1 T_{hm} + \gamma T_2 T_{hm}^2 + \gamma f_1 f_{hm} + \gamma f_2 f_{hm}^2 + \gamma TF}.
\]

- First order condition for \( p \) yields:

\[
\log \left\{ p_{hm} + \frac{1}{\alpha(s_{hm} - 1)} + f_{hm} \Gamma(T_{hm}, f_{hm}; \gamma) \right\} = w_{hm} \beta_s + \omega_{hm}.
\]
Part III Econometrics

\[ q_{hm} = \bm{x}'_{hm} \bm{\beta}_d - p_{hm} \alpha + \xi_{hm} + \ln \left( 1 - \sum_{k_m=1}^{H_m} \exp(q_{km}) \right) \]

\[ \bm{w}_{hm} \bm{\beta}_s + \omega_{hm} = \log \left\{ p_{hm} + \frac{1}{\alpha [\exp(q_{hm}) - 1]} + f_{hm} \Gamma(T_{hm}, f_{hm}; \gamma) \right\} \]

- Conventional method: GMM under \( E[Z\xi] = E[Z\omega] = 0 \), with appropriate instrument \( Z \)
- It is difficult to conduct a simulation
  - Exogenous intervention to vanish initial payment: \( f_{hm}^{new} = 0, T_{hm}^{new} = 0 \)
  - Target: Comparison of total payments before and after intervention
  - For prediction of \( p_{hm}^{new} \), a candidate is the predictive mean
  \[ E[p_{hm}^{new} | f_{hm}^{new} = 0, T_{hm}^{new} = 0, \text{Data}] \]
  - Note that the demand and supply side equations must occur at the same market \( m \) as equilibrium: They are structural equations for a simultaneous equation system for \( (q_1, \ldots, q_{H_m}, p_1, \ldots, p_{H_m}) \).
  - Since the reduced home for this system is difficult to obtain, the moment condition is not enough to obtain the predictive mean.
Instead, we put distributional assumptions on the error terms \((\xi, \omega)\).

MCMC-like numerical integration:

Joint predictive density: using variable transformation

\[
\pi(p_{new}^m, q_{new}^m | f_{new}^m = 0, T_{new}^m = 0, \text{Data}) = \pi_{\omega,\xi} \left[ \omega_{hm}(p_{new}^m, q_{new}^m ; f_{new}^m = 0, T_{new}^m = 0) \right],
\]

where \(J_m\): Jacobian for \((\omega_m, \xi_m)\) to \((p_m, q_m)\) and

\[
\xi_{hm}(p_m, q_m; f_{hm}, T_{hm}) = q_{hm} - \ln \left[ 1 - \sum_{k_m=1}^{H_m} \exp(q_{km}) \right] - \tilde{x}_{hm}' \tilde{\beta}_d + T_{hm} f_{hm} \alpha_F + p_{hm} \alpha,
\]

\[
\omega_{hm}(p_m, q_m; f_{hm}, T_{hm}) = \ln \left[ p_{hm} + \frac{1}{\alpha[\exp(q_{hm}) - 1]} + f_{hm} \Gamma(T_{hm}, f_{hm}; \gamma) \right] - w_{hm}' \beta_s.
\]

We get conditional predictive distributions:

\[
q_{hm}^{new} | p_{new}^m, q_{(-hm)}^{new}, f_{hm}^{new} = 0, T_{hm}^{new} = 0, \text{Data} \quad \text{and} \quad p_{hm}^{new} | q_{new}^m, p_{(-hm)}^{new}, f_{hm}^{new} = 0, T_{hm}^{new} = 0, \text{Data}
\]

Iterated draws from these conditionals \(\Rightarrow\) marginal samples
To minimize cost of distributional assumptions, we adopt nonparametric Bayesian method.

Nonparametric Bayes: The likelihood for an arbitrary distribution: Ex. Dirichlet process

We assume $\omega$ and $\xi$ follow nonparametric Bayesian models, independently. Now we concentrate on $\omega$.

To identify $\omega$ and the coefficient for the constant term separately, we put additional constraints on the distribution of $\omega$.

Here we adopt median assumption: $Med(\omega) = 0$.

For such a median constraint, the Polya tree mixture is suitable.
Polya tree

- Original form of Polya tree mixture
- Let \( \epsilon_j = (\epsilon_1, ..., \epsilon_j) \in E^j = \{0, 1\}^j \), \( E^* = \bigcup_{j \geq 1} E^j \) be a set of all sequence of zeros.
- Suppose that support for \( \omega_{hm} \) is \( \Omega \).
- \( \Upsilon \): Iterative partitions of \( \Omega \):

\[
\Upsilon_1 = \{B_0, B_1\}, \ \Omega = B_0 \cup B_1,
\]

and for \( j \geq 2 \), given the level \( j \) partition \( \Upsilon_j = \{B_{\epsilon_j}\} \), the level \( j + 1 \) partition \( \Upsilon_{j+1} = \{B_{\epsilon_j0}, B_{\epsilon_j1}\} \), where \( B_{\epsilon_j0} \cup B_{\epsilon_j1} = B_{\epsilon_j} \) for each \( \epsilon_j \).
- \( A = \{\alpha_{\epsilon} : \epsilon \in E^*\} \): Nonnegative hyperparameters.

A probability measure \( \mu \) follows a Polya tree with respect to \( \Upsilon = \{\Upsilon_j : j \geq 1\} \) and \( A \) and is written as \( \mu \sim PT(A, \Gamma) \) if and only if

\[
\mu(B_{\epsilon_0}|B_{\epsilon}) \sim Beta(\alpha_{\epsilon_0}, \alpha_{\epsilon_1}),
\]

where \( Beta \) denotes the Beta distribution.
Basic properties

Conjugacy

\[ \mu | \omega_1, \ldots, \omega_h \sim PT(A^*_h, \Upsilon), \]

where \( A^*_h = \{ \alpha^*_{h, \varepsilon} : \varepsilon \in E^* \} \), \( \alpha^*_{h, \varepsilon} = \alpha_\varepsilon + n_\varepsilon \), and \( n_\varepsilon = \sum_{i=1}^{h} I[\omega_i \in B_\varepsilon] \).

Predictive distribution:

\[
Pr[\omega_{h+1} \in B_{\varepsilon_1, \ldots, \varepsilon_j} | \omega_1, \ldots, \omega_h] = \prod_{k=1}^{j} E[\mu(B_{\varepsilon_1 \ldots \varepsilon_k} | B_{\varepsilon_1 \ldots \varepsilon_{k-1}})]
\]
\[
= \prod_{k=1}^{j} \frac{\alpha_{\varepsilon_1 \ldots \varepsilon_k} + n_{\varepsilon_1 \ldots \varepsilon_k}}{\alpha_{\varepsilon_1 \ldots \varepsilon_k 0} + \alpha_{\varepsilon_1 \ldots \varepsilon_k 1} + n_{\varepsilon_1 \ldots \varepsilon_{k-1}}},
\]

where we define \( n_{\varepsilon_0} = h \).

Intuition:

- Collection of binary histograms: Separate a bin \( B_\varepsilon \) to two bins \( B_{\varepsilon 0} \) and \( B_{\varepsilon 1} \).
- Large \( \alpha \): Probability of \( B_{\varepsilon 0} \) is \( \alpha_{\varepsilon 0} / [\alpha_{\varepsilon 0} + \alpha_{\varepsilon 1}] \), regardless of data
- Small \( \alpha \): Only histogram
Statistician needs to set

- **Truncated point** $J$:
  - Essentially, infinite-level partitions are required
  - In practice, truncate it.
- **Partition** $\Upsilon$
  - using a base probability measure $G$ as:
    $$B_{\epsilon_1, \ldots, \epsilon_j} = \left[ G^{-1}\left(\frac{m}{2^j}\right), G^{-1}\left(\frac{m + 1}{2^j}\right) \right],$$
    where $m$ is a decimal number representation for $\epsilon_1, \ldots, \epsilon_j$.
- **Hyperparameters** $A$
  - $\alpha_{\epsilon_1, \ldots, \epsilon_j} = cj^2$ with a constant $c$
  - $c \to \infty$: Regardless of data, binary partitions place probabilities $1/2$
  - $c \to 0$: Only histogram
  - $j$: Deeper levels histograms have fewer weights
Problem of Polya tree: If there are many observations near borders of bins, unstable

Polya tree mixture (Hanson and Johnson, 2002): \( G \) is assumed to depend on a parameter \( \tau_\omega \) which is also a parameter to be estimated

Border is smoothed

Likelihood: Constructed via predictive densities

\[
\pi_\omega(\omega_1, \ldots, \omega_H|\tau_\omega) = f(\omega_1|\tau_\omega)f(\omega_2|\omega_1, \tau_\omega)f(\omega_3|\omega_1, \omega_2, \tau_\omega)\ldots f(\omega_H|\omega_1, \ldots, \omega_{H-1}, \tau_\omega)
\]

where

\[
f(\omega_1|\tau_\omega) = g_{\tau_\omega}(w_1),
\]

\[
f(w_{h+1}|w_1, \ldots, w_h, \tau_\omega) = \left[ \prod_{j=1}^{J} \frac{cj^2 + n_{\epsilon(j, \tau_\omega, w_{h+1})(w_h)}}{2cj^2 + n_{\epsilon(j-1, \tau_\omega, w_{h+1})(w_h)}} \right] 2^{J-1} g_{\tau_\omega}(w_{h+1}),
\]

and

\[
\epsilon(j, \tau_\omega, w) = \{ \epsilon_1, \ldots, \epsilon_j : w \in B_{\epsilon_1\ldots\epsilon_j}^{\tau_\omega} \},
\]

\[
n_{\epsilon(j, \tau_\omega, w_{h+1})(w_h)} = \sum_{i=1}^{h} I[w_i \in B_{\epsilon(j, \tau_\omega, w_{h+1})}^{\tau_\omega}],
\]

If \( G(0) = 1/2 \), \( Pr(w_h \leq 0|\tau_\omega) = 0.5 \) holds for any \( h \)
Part IV: Empirical analysis

**Data**

- **Source:** Shuukan Asahi (2011)
- **Market:** Prefectures, except those only with 0 or 1 home
- **Sample size:** \( H = 1265 \)
- **\( s_h \):** Share: \# residents in \( h \)th home over \# potential consumers in \( m \)th market
  - Potential consumer: number of the elders with the age 65 years old or more and with the eligibility level of Care Required 1 or more: Typically required for institutional care costs to be covered by the LTCI
- **Multiple answers for** \( p, F, T \)
  - Several homes answer maximum and minimum values. We use the average.
  - We assume variations \( p \) and \( F \) are from qualities of care and facility
  - There are only 39 homes which have variations in \( T \): No separating equilibria using product differentiation
Estimation result

- Random walk MH (See paper for simulation data analysis)
- Primitives: $c = 10$, $J = 5$

Coefficient estimates: Table: P.77

- $\beta_d$: Utility components:
  - $\#$ residents per worker: $-$: consumers prefer homes with an enough capacity of care workers
  - Occupancy rate: $+$: good public image is an important element to attract consumers
  - Chain dummy: $+$: Active advertisement positively stimulates the demand
  - Years from opening: ambiguous: old facilities is not preferable, but credibility from a long history is attractive

- $\beta_s$: Marginal cost component
  - $\#$ residents per worker: $-$: Natural
  - Chain dummy: ambiguous: Scale economy and high advertisement cost are canceled out
  - Years from opening: $-$: Accumulation of know-hows decreases the cost
  - Local average rent: $-$: Natural
  - Local average wage: ambiguous: medical and welfare sector is too large to capture the care worker
For MCMC-like procedure, large $H_m$ is burdensome: We concentrate on Shizuoka ($H_m = 32$)

Total payment before and after intervention for a match of an individual with $\tau_i$ lifetime and a home with $T_h$ expiration

- After: $p^{new}_h \tau_i$
- Before:
  - $T_h = 0$ (without initial payment): $p_h \tau_i$: Comparison: $p_h$ vs $p^{new}_h$
  - $\tau_i \leq T_h$ (low longevity risk): $(p_h + f_h)\tau_i$: Comparison: $p_h + f_h$ vs $p^{new}_h$
  - $\tau_i > T_h$ (high longevity risk): $p_h \tau_i + f_h T_h$: Need to specify $\tau_i$

if the intervention is realized, consumers with low longevity risks can reduce the total payment for the homes which currently offer a contract with an initial payment
- 5, 10, 20 and 30 years of remaining lifetimes
- The total payments after the intervention ≤ before the intervention only in the case where a consumer with 30 years of remaining lifetime chooses a specific home
- ⇒ Extremely risky consumers prefer a contract with an initial payment
We analyzed Japanese private nursing home market using an extended model of Berry et al (1995)

To evaluate effects of a governmental intervention, we employed a nonparametric Bayesian analysis

Prediction results showed that the intervention increases the welfare of many consumers, except those who have extremely high longevity risks, through a reduction of their total payment.