Dynamic choice models with panel data

Michel Bierlaire

Transport and Mobility Laboratory
School of Architecture, Civil and Environmental Engineering
Ecole Polytechnique Fédérale de Lausanne

January 3, 2017
Outline

1. Introduction
2. Static model
3. Static model with panel effect
4. Dynamic model
5. Dynamic model with panel effect
6. Application
7. Summary
Panel data

- Type of data used so far: **cross-sectional**.
- Cross-sectional: observation of individuals at the same point in time.
- Time series: sequence of observations.
- **Panel data** is a combination of comparable time series.
Introduction

Panel data
Data collected over multiple time periods for the same sample of individuals.

Multidimensional

<table>
<thead>
<tr>
<th>Individual</th>
<th>Day</th>
<th>Price of stock 1</th>
<th>Price of stock 2</th>
<th>Purchase</th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td>t</td>
<td>$x_{1nt}$</td>
<td>$x_{2nt}$</td>
<td>$i_{int}$</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>12.3</td>
<td>15.6</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>12.1</td>
<td>18.6</td>
<td>2</td>
</tr>
<tr>
<td>1</td>
<td>3</td>
<td>11.0</td>
<td>25.3</td>
<td>2</td>
</tr>
<tr>
<td>1</td>
<td>4</td>
<td>9.2</td>
<td>25.1</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>12.3</td>
<td>15.6</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>12.1</td>
<td>18.6</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>11.0</td>
<td>25.3</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>4</td>
<td>9.2</td>
<td>25.1</td>
<td>1</td>
</tr>
</tbody>
</table>
Introduction

Examples of discrete panel data

- People are interviewed monthly and asked if they are working or unemployed.
- Firms are tracked yearly to determine if they have been acquired or merged.
- Consumers are interviewed yearly and asked if they have acquired a new cell phone.
- Individual’s health records are reviewed annually to determine onset of new health problems.
Model: single time period
Outline

1. Introduction
2. Static model
3. Static model with panel effect
4. Dynamic model
5. Dynamic model with panel effect
6. Application
7. Summary
Static model

\[ x_t = x_{t-1} + \varepsilon_t \]

\[ u_{t-1} = u_t - \varepsilon_{t-1} \]

\[ i_t = i_{t-1} + u_{t-1} \]
Static model

Utility

\[ U_{int} = V_{int} + \varepsilon_{int}, \ i \in \mathcal{C}_{nt}. \]

Logit

\[ P(i_{nt}) = \frac{e^{V_{int}}}{\sum_{j \in \mathcal{C}_{nt}} e^{V_{jnt}}} \]

Estimation: contribution of individual \( n \) to the log likelihood

\[ P(i_{n1}, i_{n2}, \ldots, i_{nT}) = P(i_{n1})P(i_{n2}) \cdots P(i_{nT}) = \prod_{t=1}^{T} P(i_{nt}) \]

\[ \ln P(i_{n1}, i_{n2}, \ldots, i_{nT}) = \ln P(i_{n1}) + \ln P(i_{n2}) + \cdots + \ln P(i_{nT}) = \sum_{t=1}^{T} \ln P(i_{nt}) \]
Static model: comments

- Views observations collected through time as supplementary cross sectional observations.
- Standard software for cross section discrete choice modeling may be used directly.
- Simple, but there are two important limitations:

  **Serial correlation**
  - unobserved factors persist over time,
  - in particular, all factors related to individual $n$,
  - $\varepsilon_{in(t-1)}$ cannot be assumed independent from $\varepsilon_{int}$.

  **Dynamics**
  - Choice in one period may depend on choices made in the past.
  - e.g. learning effect, habits.
Outline

1. Introduction
2. Static model
3. Static model with panel effect
4. Dynamic model
5. Dynamic model with panel effect
6. Application
7. Summary
Dealing with serial correlation

\[ x_{t-1} \]
\[ \varepsilon_{t-1} \]
\[ U_{t-1} \]
\[ i_{t-1} \]
\[ x_t \]
\[ \varepsilon_t \]
\[ U_t \]
\[ i_t \]
Panel effect

Relax the assumption that $\varepsilon_{int}$ are independent across $t$.

Assumption about the source of the correlation
- individual related unobserved factors,
- persistent over time.

The model

$$\varepsilon_{int} = \alpha_{in} + \varepsilon'_{int}$$

It is also known as
- agent effect,
- unobserved heterogeneity.
Assuming that $\varepsilon'_{int}$ are independent across $t$, we can apply the static model.

Two versions of the model:
- with fixed effect: $\alpha_{in}$ are unknown parameters to be estimated,
- with random effect: $\alpha_{in}$ are distributed.
Static model with fixed effect

Utility

\[ U_{int} = V_{int} + \alpha_{in} + \varepsilon'_{int}, \ i \in C_{nt}. \]

Logit

\[ P(i_{nt}) = \frac{e^{V_{int} + \alpha_{in}}}{\sum_{j \in C_{nt}} e^{V_{jnt} + \alpha_{jn}}}. \]

Estimation: contribution of individual \( n \) to the log likelihood

\[ P(i_{n1}, i_{n2}, \ldots, i_{nT}) = P(i_{n1})P(i_{n2})\cdots P(i_{nT}) = \prod_{t=1}^{T} P(i_{nt}) \]

\[ \ln P(i_{n1}, i_{n2}, \ldots, i_{nT}) = \ln P(i_{n1}) + \ln P(i_{n2}) + \cdots + \ln P(i_{nT}) = \sum_{t=1}^{T} \ln P(i_{nt}) \]
Comments

- \( \alpha_{in} \) capture permanent taste heterogeneity.
- For each \( n \), one \( \alpha_{in} \) must be normalized to 0.
- The \( \alpha \)'s are estimated consistently only if \( T \to \infty \).
- This has an effect on the other parameters that will be inconsistently estimated.
- In practice,
  - \( T \) is usually too short,
  - the number of \( \alpha \) parameters is usually too high,

for the model to be consistently estimated and practical.
Denote $\alpha_n$ the vector gathering all parameters $\alpha_{in}$.

Assumption: $\alpha_n$ is distributed with density $f(\alpha_n)$.

For instance:

$$\alpha_n \sim N(0, \Sigma).$$

We have a mixture of static models.

Given $\alpha_n$, the model is static, as $\varepsilon_{int}'$ are assumed independent across $t$. 
Utility

\[ U_{int} = V_{int} + \alpha_{in} + \varepsilon'_{int}, \; i \in C_{nt}. \]

Conditional choice probability

\[ P(i_{nt}|\alpha_n) = \frac{e^{V_{int} + \alpha_{in}}}{\sum_{j \in C_{nt}} e^{V_{jnt} + \alpha_{jn}}}. \]
Static model with random effect

Contribution of individual \( n \) to the log likelihood, given \( \alpha_n \)

\[
P(i_{n1}, i_{n2}, \ldots, i_{nT}|\alpha_n) = \prod_{t=1}^{T} P(i_{nt}|\alpha_n).
\]

Unconditional choice probability

\[
P(i_{n1}, i_{n2}, \ldots, i_{nT}) = \int_{\alpha} \prod_{t=1}^{T} P(i_{nt}|\alpha)f(\alpha)d\alpha.
\]
Static model with random effect

Estimation

- Mixture model.
- Requires simulation for large choice sets.
- Generate draws $\alpha^1, \ldots, \alpha^R$ from $f(\alpha)$.
- Approximate

$$P(i_{n1}, i_{n2}, \ldots, i_{nT}) = \int_{\alpha} \prod_{t=1}^{T} P(i_{nt}|\alpha) f(\alpha) d\alpha \approx \frac{1}{R} \sum_{r=1}^{R} \prod_{t=1}^{T} P(i_{nt}|\alpha^r)$$

- The product of probabilities can generate very small numbers.

$$\sum_{r=1}^{R} \prod_{t=1}^{T} P(i_{nt}|\alpha^r) = \sum_{r=1}^{R} \exp \left( \sum_{t=1}^{T} \ln P(i_{nt}|\alpha^r) \right).$$
Comments

- Parameters to be estimated: $\beta$’s and $\sigma$’s
- Maximum likelihood estimation leads to consistent and efficient estimators.
- Ignoring the correlation (i.e. assuming that $\alpha_n$ is not present) leads to consistent but not efficient estimators (not the true likelihood function).
- Accounting for serial correlation generates the true likelihood function and, therefore, the estimates are consistent and efficient.
Choice in one period may depend on choices made in the past
e.g. learning effect, habits.
Simplifying assumption:
- the utility of an alternative at time $t$
- is influenced by the choice made at time $t - 1$ only.

It leads to a dynamic Markov model.
Dynamic Markov model

\[ x_t - 1 \varepsilon_t - 1 U_t - 1 i_{t-1} \]

\[ x_t \varepsilon_t U_t i_t \]
Dynamic Markov model

The model

\[ U_{int} = V_{int} + \gamma y_{in(t-1)} + \varepsilon_{int}, \ i \in C_{nt}. \]

\[ y_{in(t-1)} = \begin{cases} 
1 & \text{if alternative } i \text{ was chosen by } n \text{ at time } t - 1 \\
0 & \text{otherwise.} 
\end{cases} \]

Captures serial dependence on past realized state

- Example - utility of bus today depends on whether consumer took bus yesterday (habit).
- Fails if utility of bus today depends on permanent individual taste for bus (tastes) and whether consumer took bus yesterday. No serial correlation.

Estimation: same as for the static model except that observation \( t = 0 \) is lost
Outline

1 Introduction
2 Static model
3 Static model with panel effect
4 Dynamic model
5 Dynamic model with panel effect
6 Application
7 Summary
Dynamic model with panel effect

Dynamic Markov model with serial correlation

\[ x_t - 1 \]

\[ \varepsilon_t \]

\[ U_t - 1 \]

\[ \beta_t \]

\[ i_t - 1 \]

\[ i_t \]
Dynamic Markov model

Extension: combine Markov with panel effect

\[ U_{int} = V_{int} + \alpha_{in} + \gamma y_{in(t-1)} + \varepsilon'_{int}, \ i \in C_{nt}. \]

Dynamic Markov model with fixed effect

- Similar to the static model with FE.
- Similar limitations.

Dynamic Markov model with random effect

- Difficulties depending on how the Markov chain starts.
- If the first choice \( i_0 \) is truly exogenous \( \rightarrow \) similar to the static model with RE.
Dynamic Markov model

What if \( i_{n0} \) is not exogenous (i.e. stochastic)?

\[
U_{in1} = V_{in1} + \alpha_{in} + \gamma y_{in0} + \varepsilon_{in1}', \; i \in C_{n1}.
\]

- The first choice \( i_{n0} \) is dependent on the agent’s effect \( \alpha_{in} \).
- So, the explanatory variable \( y_{in0} \) is correlated with \( \alpha_{in} \).
- This is called endogeneity.
- Solution: use the Wooldridge approach.
Conditional on $y_{i0}$, we have a dynamic Markov model with RE as before

$$U_{int} = V_{int} + \alpha_{in} + \gamma y_{in(t-1)} + \varepsilon'_{int}, \ i \in Cnt.$$  

Contribution of individual $n$ to the log likelihood, given $i_{n0}$ and $\alpha_n$

$$P(i_{n1}, i_{n2}, \ldots, i_{nT}|i_{n0}, \alpha_n) = \prod_{t=1}^{T} P(i_{nt}|i_{n0}, \alpha_n).$$

We integrate out $\alpha_n$

$$P(i_{n1}, i_{n2}, \ldots, i_{nT}|i_{n0}) = \int P(i_{nt}|i_{n0}, \alpha) f(\alpha|i_{n0}) d\alpha.$$
Dynamic Markov model with RE - Wooldridge

The main difference between static model with RE and dynamic model with RE is the term

$$f(\alpha | i_{n0})$$

It captures the distribution of the panel effects, knowing the first choice.

This can be approximated by, for instance,

$$\alpha_n = a + by_{n0} + cx_n + \xi_n, \quad \xi_n \sim N(0, \Sigma_\alpha).$$

- $a$, $b$ and $c$ are vectors and $\Sigma_\alpha$ a matrix of parameters to be estimated.
- $x_n$ capture the entire history ($t = 1, \ldots, T$) for agent $n$.
- This addresses the endogeneity issue.
Outline

1. Introduction
2. Static model
3. Static model with panel effect
4. Dynamic model
5. Dynamic model with panel effect
6. Application
7. Summary
**Application**

**Context**
- Study done in 1998, Sardinia Island, Italy
- Cagliari-Assimini corridor (20km)
- Modal shares: car (75%), bus (20%), train (3%), other (2%)
- RP/SP data.
- Not time series, but panel structure of SP data.
- $t$ is the index of the choice experiment instead of time.
- $t = 0$ corresponds to the RP observation.
- Panel effect is captured.
## Estimation results

<table>
<thead>
<tr>
<th>Variable</th>
<th>Logit Estimate</th>
<th>Logit t-test</th>
<th>with panel effect Estimate</th>
<th>t-test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cte. train</td>
<td>-0.727</td>
<td>-3.130</td>
<td>-0.745</td>
<td>-3.047</td>
</tr>
<tr>
<td>Cte. car</td>
<td>-2.683</td>
<td>-6.378</td>
<td>-2.770</td>
<td>-5.775</td>
</tr>
<tr>
<td>Travel time (min)</td>
<td>-0.061</td>
<td>-4.120</td>
<td>-0.067</td>
<td>-3.722</td>
</tr>
<tr>
<td>Travel cost/wage rate (euros)</td>
<td>-1.895</td>
<td>-3.198</td>
<td>-2.364</td>
<td>-4.454</td>
</tr>
<tr>
<td>Waiting time (min)</td>
<td>-0.252</td>
<td>-6.247</td>
<td>-0.270</td>
<td>-6.705</td>
</tr>
<tr>
<td>Comfort low</td>
<td>-1.990</td>
<td>-7.328</td>
<td>-2.075</td>
<td>-6.219</td>
</tr>
<tr>
<td>Comfort avg.</td>
<td>-1.107</td>
<td>-6.330</td>
<td>-1.187</td>
<td>-5.546</td>
</tr>
<tr>
<td>Transfers</td>
<td>-0.286</td>
<td>-1.378</td>
<td>-0.316</td>
<td>-1.000</td>
</tr>
<tr>
<td>Panel effect std. dev.</td>
<td></td>
<td></td>
<td>0.840</td>
<td>6.348</td>
</tr>
<tr>
<td>Log likelihood</td>
<td>-511.039</td>
<td></td>
<td>-502.959</td>
<td></td>
</tr>
<tr>
<td>$\rho^2$</td>
<td>0.116</td>
<td></td>
<td>0.130</td>
<td></td>
</tr>
</tbody>
</table>
### Average value of time by purpose (euros/min)

<table>
<thead>
<tr>
<th>Purpose</th>
<th>Observations</th>
<th>Logit</th>
<th>with panel effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Work</td>
<td>321</td>
<td>0.20</td>
<td>0.17</td>
</tr>
<tr>
<td>Study</td>
<td>285</td>
<td>0.05</td>
<td>0.04</td>
</tr>
<tr>
<td>Personal business</td>
<td>164</td>
<td>0.13</td>
<td>0.11</td>
</tr>
<tr>
<td>Leisure</td>
<td>64</td>
<td>0.16</td>
<td>0.14</td>
</tr>
</tbody>
</table>
Comments

- Panel effect is significant.
- Significant improvement of the fit.
- With small samples, the gain in efficiency obtained from the panel effect may significantly improve the estimates.
Outline

1. Introduction
2. Static model
3. Static model with panel effect
4. Dynamic model
5. Dynamic model with panel effect
6. Application
7. Summary
Summary

Static model
- Straightforward extension of cross-sectional specification.
- Two main limitations: serial correlation and dynamics.

Panel effect
- Deals with serial correlation.
- Fixed effect:
  - Static model with additional parameters.
  - Not operational in most practical cases.
- Random effect:
  - Modifies the log likelihood function.
  - Must integrate the product of the choice probabilities over time.
Summary

Dynamic model, with a Markov assumption
Static model with an additional variable: the previous choice.

Dynamic model with panel effect
- Both can be combined.
- Must capture the relation between the first choice and the panel effect.

Application
Illustrates the importance of the panel effect.