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

Introduction

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

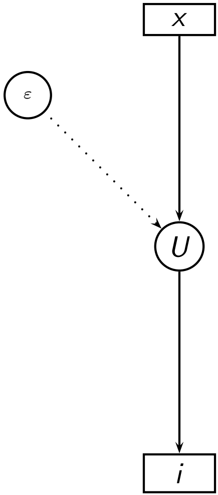
Individual n	Day t	Price of stock 1 x_{1nt}	Price of stock 2 x_{2nt}	Purchase i_{nt}
1	1	12.3	15.6	1
1	2	12.1	18.6	2
1	3	11.0	25.3	2
1	4	9.2	25.1	0
2	1	12.3	15.6	2
2	2	12.1	18.6	0
2	3	11.0	25.3	0
2	4	9.2	25.1	1

Introduction

Examples of discrete panel data

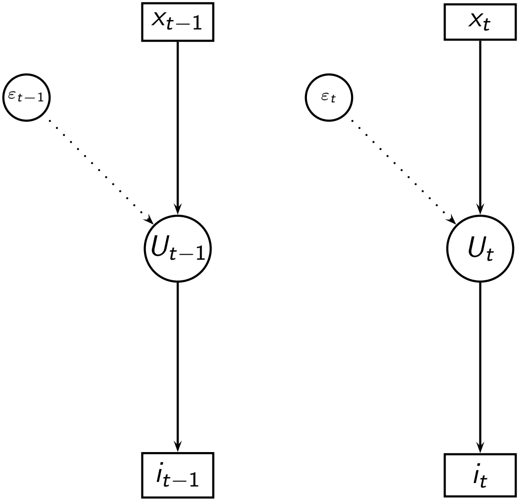
- People are interviewed monthly and asked if they are working or unemployed.
- Firms are tracked yearly to determine if they have been acquired or merged.
- Consumers are interviewed yearly and asked if they have acquired a new cell phone.
- Individual's health records are reviewed annually to determine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

Model: single time period



Outline

Static model



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

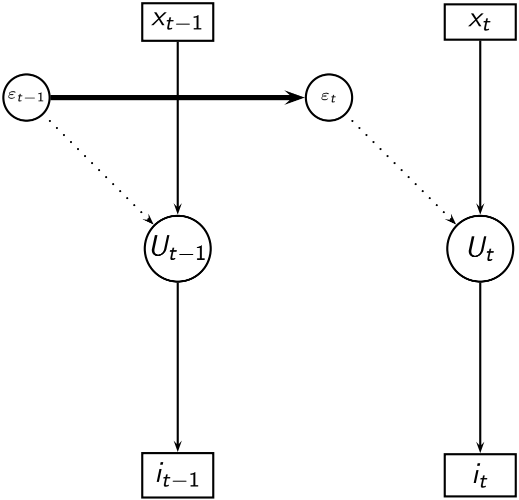
- unobserved factors persist over time,
- in particular, all factors related to individual n,
- $\varepsilon_{in(t-1)}$ cannot be assumed independent from ε_{int}.

Dynamics

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

Outline

Dealing with serial correlation



Panel effect

Relax the assumption that ε_{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.

Panel effect

- Assuming that ε'_{int} are independent across t,
- we can apply the static model.
- Two versions of the model:
  - with fixed effect: α_{in} are unknown parameters to be estimated,
  - with random effect: α_{in} are distributed.

Static model with fixed effect

Utility

$$U_{int} = V_{int} + \alpha_{in} + \varepsilon'_{int}, \ i \in \mathcal{C}_{nt}.$$

Logit

$$P(i_{nt}) = \frac{e^{V_{int}+\alpha_{in}}}{\sum_{j \in \mathcal{C}_{nt}} e^{V_{jnt}+\alpha_{jn}}}$$

Estimation: contribution of individual n to the log likelihood

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

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

Static model with fixed effect

Comments

- α_{in} capture permanent taste heterogeneity.
- For each n, one α_{in} must be normalized to 0.
- The α'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 α parameters is usually too high,

  for the model to be consistently estimated and practical.

Static model with random effect

- Denote α_n the vector gathering all parameters α_{in}.
- Assumption: α_n is distributed with density $f(\alpha_n)$.
- For instance:

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

- We have a mixture of static models.
- Given α_n, the model is static, as ε'_{int} are assumed independent across t.

Static model with random effect

Utility

$$U_{int} = V_{int} + \alpha_{in} + \varepsilon'_{int}, \;\; i \in \mathcal{C}_{nt}.$$

Conditional choice probability

$$P(i_{nt}|\alpha_n) = \frac{e^{V_{int}+\alpha_{in}}}{\sum_{j \in \mathcal{C}_{nt}} e^{V_{jnt}+\alpha_{jn}}}$$

Static model with random effect

Contribution of individual n to the log likelihood, given α_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.

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

Static model with random effect

Comments

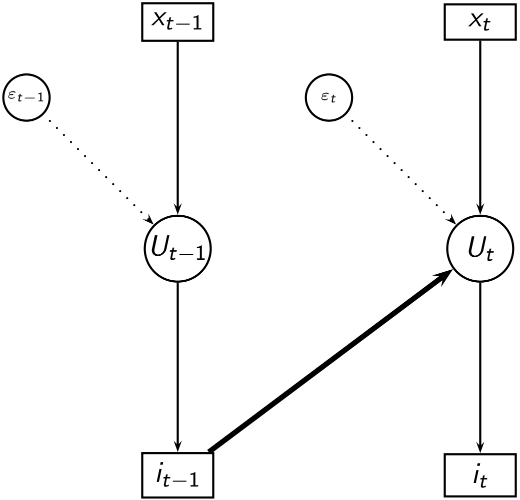
- Parameters to be estimated: β's and σ's
- Maximum likelihood estimation leads to consistent and efficient estimators.
- Ignoring the correlation (i.e. assuming that α_n is not present) leads to consistent but not efficient estimators (not the true likelihood function).
- Accounting for serial correlation generates the true likelihood function and, therefore, the estimates are consistent and efficient.

Outline

Dynamics

- 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



Dynamic Markov model

The model

$$U_{int} = V_{int} + \gamma y_{in(t-1)} + \varepsilon_{int}, \ i \in \mathcal{C}_{nt}.$$

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

Captures serial dependence on past realized state

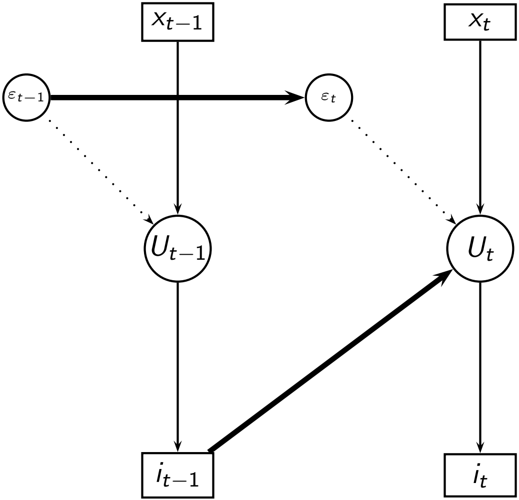
- Example - utility of bus today depends on whether consumer took bus yesterday (habit).
- Fails if utility of bus today depends on permanent individual taste for bus (tastes) and whether consumer took bus yesterday. No serial correlation.

Estimation: same as for the static model

except that observation $t = 0$ is lost

Outline

Dynamic Markov model with serial correlation



Dynamic Markov model

Extension: combine Markov with panel effect

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

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)?

$$U_{in1} = V_{in1} + \alpha_{in} + \gamma y_{in0} + \varepsilon'_{in1}, \; i \in \mathcal{C}_{n1}.$$

- The first choice i_{n0} is dependent on the agent's effect α_{in}.
- So, the explanatory variable y_{in0} is correlated with α_{in}.
- This is called endogeneity.
- Solution: use the Wooldridge approach.

Dynamic Markov model with RE - Wooldridge

Conditional on y_{in0}, we have a dynamic Markov model with RE as before

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

Contribution of individual n to the log likelihood, given i_{n0} and α_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 α_n

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

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,

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

  - a, b and c are vectors and Σ_α 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

Application

Cherchi and Ortuzar (2002) Mixed RP/SP models incorporating interaction effects, Transportation 29(4), pp. 371--395.

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.

Application

Estimation results

	Logit		with panel effect	
Variable	Estimate	*t*-test	Estimate	*t*-test
Cte. train	-0.727	-3.130	-0.745	-3.047
Cte. car	-2.683	-6.378	-2.770	-5.775
Travel time (min)	-0.061	-4.120	-0.067	-3.722
Travel cost/wage rate (euros)	-1.895	-3.198	-2.364	-4.454
Waiting time (min)	-0.252	-6.247	-0.270	-6.705
Comfort low	-1.990	-7.328	-2.075	-6.219
Comfort avg.	-1.107	-6.330	-1.187	-5.546
Transfers	-0.286	-1.378	-0.316	-1.000
Panel effect std. dev.			0.840	6.348
Log likelihood	-511.039		-502.959	
ρ^2	0.116		0.130	

Application

Average value of time by purpose (euros/min)

		Logit	with panel effect
Work	321 obs.	0.20	0.17
Study	285 obs.	0.05	0.04
Personal business	164 obs.	0.13	0.11
Leisure	64 obs.	0.16	0.14

Application

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

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.