

패널 데이터 모형: 기초와 응용

류근관

서울대학교 경제학과 교수

ezstat.co.kr, ryu@snu.ac.kr

발표순서

- 왜 패널데이터를 사용하는가
- 기본적인 패널데이터 모형
- 패널데이터 모형의 확장
- 패널데이터 모형의 응용 사례
- 맺음말

왜 패널데이터를 사용하는가

- **Lexis Diagram**

(그림 여기에)

$$y_{it} = x_{it} * a + b * \text{Cohort} + c * (\text{Years in US}) + \text{error}$$

횡단면 자료를 이용해서는 b와 c 식별 불가함.

이는 횡단면 자료의 경우 $C + (\text{Years in US}) = t$ 이기 때문임

왜 패널데이터를 사용하는가

- 보다 나은 통제가 가능함

예시: 군납유 유류입찰 담합사건

국내 정유5개사가 1998, 1999, 2000 등
3개 연도에 걸쳐 군납 유류 구매 입찰시
담합한 것으로 판정됨. 담합으로 인한
손해액을 산정하는 문제가 발생함.

패널자료 없는 경우의 단순비교

	군납유	비군납유
담합시기	A	B
비담합시기	C	D

- Yardstick Method : A-B (횡단면자료)
- Before & After Method : A-C (시계열자료)

패널자료 이용한 이중비교

	군납유	비군납유
담합시기	A	B
비담합시기	C	D

- D in D Method : $(A-B)-(C-D)$ 또는 $(A-C)-(B-D)$
- $A-B$ =담합+유종별 특수성, $B-D$ =담합+시기별 특수성
- $C-D$ =유종별 특수성, $B-D$ =시기별 특수성

단순비교치와 이중비교치의 비교 (그래프)

(그림 여기에)

기본적인 패널데이터 모형

- **Fixed effect as a special case of random effect**
- **Time constant variables**
- **Endogenous variables**

기본적인 패널데이터 모형

(panel_ohp.pdf 내용 여기에 링크할 것)

- **GLS=OLS to “fractional mean-deviation”**

패널 데이터 모형의 확장

- **Lagged dependent variables**

(Arrelano and Bond 모형 소개)

패널 데이터 모형의 응용 사례

- **Penn World Table** 이용한 **New Growth Theory**

$$y_{it} = x_{it} * a + b * (\text{country dummies}) + c * (\text{time dummies}) + d * (\text{human capital, institution, etc})$$

패널 데이터 모형의 응용 사례

- **Financial Development and Economic Growth**

횡단면 자료: $y_i = x_i * a + c * (\text{financial developm't})$

문제: x_i 로 국가간 차이를 모두 통제하기는 어려움.

Rajan & Zingales 가 다음을 제안:

(국가, 산업) 패널 자료: $y_{is} = x_{is} * a + b * (\text{country dummy}) + c * (\text{industry dummy})$

$+ d * (\text{country dummy}) * (\text{industry dummy})$

패널데이터 모형의 응용 사례

- 기업의 부도예측 및 신용등급 예측모형

(기업의 패널자료 이용한 패널데이터 모형)

Data : 여러 기업의 9개 년에 걸친 패널 자료 이용

부도예측모형

1. Panel Binary Probit Model

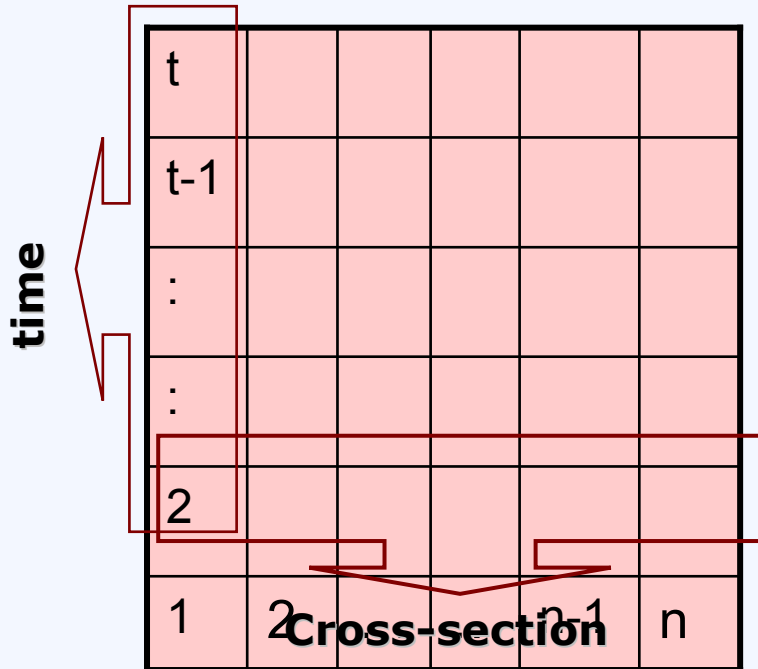
- 기업 수: **165** 기업
- 기간 수: 최장 **9**년에 걸친 연간 자료

신용등급판정모형

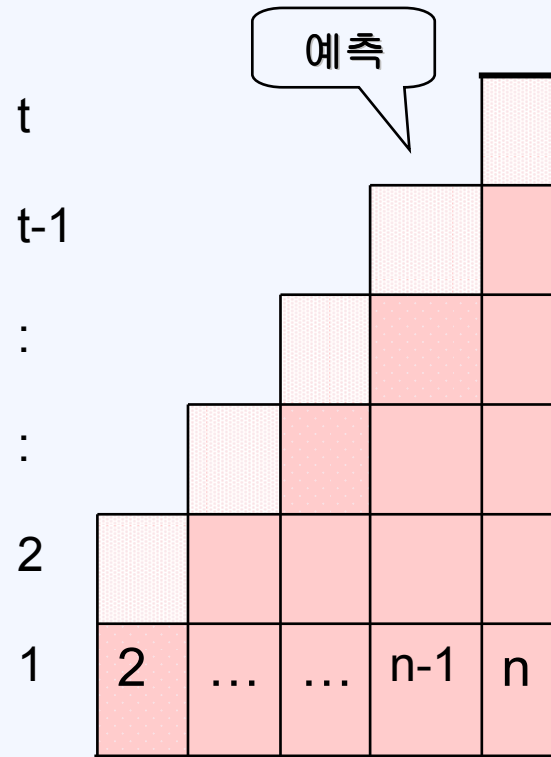
2. Ordered Probit Model

- 기업 수: 453개
- 신용등급 AAA, AA, A, BBB, BB, B (+/- 등의 notch고려 안 함)

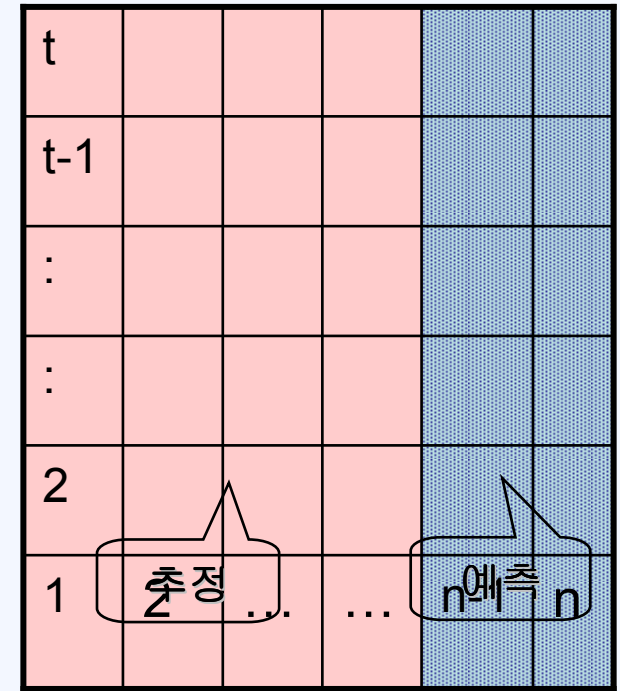
■ In-Sample Test



■ Out-of-Sample Test timewise



■ Out-of-Sample Test cross-section



부도예측모형 - 1. Panel Binary Probit Model

모형의 구조

$y_{it}^* = x_{it}\beta + u_{it}$ Bayesian MCMC기법을 이용하여 Gibbs Run마다 β , $P(y_{it}=1) = \Phi(x_{it}\beta)$ 계산한 뒤, 그 분포 이용

Panel binary probit model (한 기업 내 시기별 부실의 연관성 감안)

$$y_{it}^* = x_{it}\beta + u_{it}$$

$$(A) \quad u_{it} = \rho u_{i,t-1} + \varepsilon_{it} \quad |\rho| < 1 \quad \varepsilon_{it} \sim iid N(0,1)$$

: AR(1) model

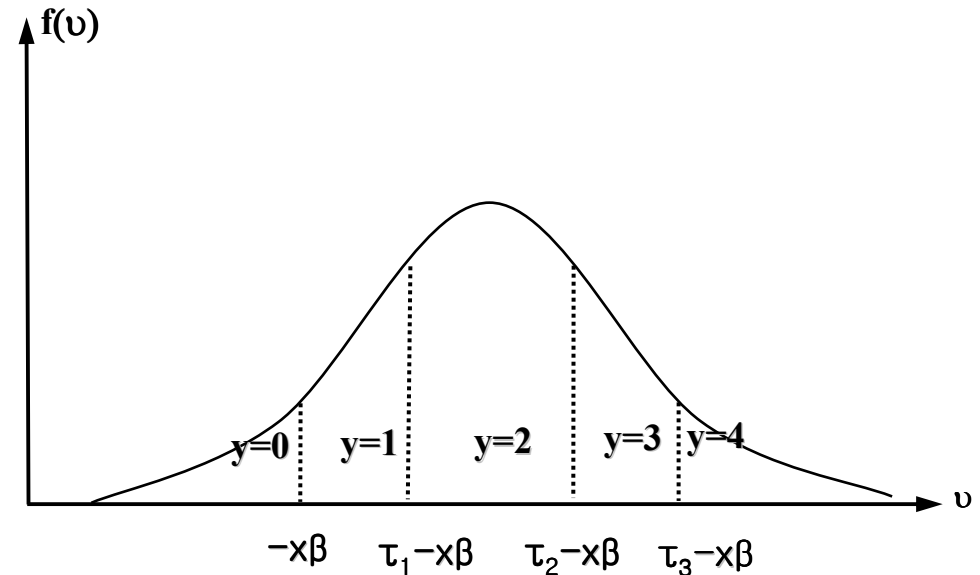
$$(B) \quad u_{it} = \sigma_{\alpha} \alpha_i + \varepsilon_{it} \quad \text{: Random effect model}$$

신용등급 판정모델의 구조

Panel Ordered Probit Model

$$y_i^* = x_i \tau \beta + u_i$$

$$\left\{ \begin{array}{ll} y_i^* = 0 & \text{if } y_i^* \leq 0 \\ y_i^* = 1 & \text{if } 0 \leq y_i^* \leq \tau_1 \\ y_i^* = 2 & \text{if } \tau_1 \leq y_i^* \leq \tau_2 \\ \vdots & \vdots \\ y_i^* = J & \text{if } \tau_{J-1} \leq y_i^* \end{array} \right.$$



Panel Ordered Probit Model

$$(A) \quad u_{it} = \rho u_{i,t-1} + \varepsilon_{it} \quad : \text{AR(1) model}$$

$$(B) \quad u_{it} = \sigma_\alpha \alpha_i + \varepsilon_{it} \quad : \text{random effect}$$

추론방식

- **Use Bayesian Approach**
 - see Ryu (mimeo), “Panel Probit Models.pdf”
 - data augmentation (y^*)
 - Gibbs sampling
 - numerical MCMC

패널 데이터 모형의 응용 사례

- 기업의 소유구조와 기업성과
- **Demsetz and Lehn (1985), Morck, Shleifer, and Vishny (1988), and McConnell and Servaes (1990)**
- **Lee and Ryu (mimeo) suggests panel data model**

패널 데이터 모형의 응용 사례

- **Lee and Ryu: re-estimate the relation between management ownership and firm's value after controlling for the history of management ownership as well as inter-firm differences using panel data.**

패널 데이터 모형의 응용 사례

- **Lee and Ryu also consider the possibility that the current ownership structure is jointly determined with the firm value.**
- **They find that history of the management ownership, not its current level, matters in determining the firm value, which is consistent with information asymmetry arguments.**

패널 데이터 모형의 응용 사례

- **Natural Experiments**와 패널 데이터

이중차분법 이용한 두 가지 응용 사례

패널 데이터 모형의 응용 사례

Jeffrey Milyo and Joel Waldfogel (1999), “The Effect of Price Advertising on Prices: Evidence in the Wake of 44 Liquormart,” American Economic Review, 89, 1081-1096

- **The US supreme court decision (44 Liquormart decision), eliminating Rhode Island’s ban on liquor price advertising, made Rhode Island the subject of a natural experiment for measuring the effect of advertising on prices.**
- **Using Massachusetts prices as controls, they find that advertising stores substantially cut prices of the products that they advertise. (Prices of other products do not change.)**

패널 데이터 모형의 응용 사례

Bruce Meyer, W. Viscusi and D. Durbin (1995), “Workers’ Compensation and injury Duration: Evidence from a Natural Experiment,” American Economic Review, 85, 322-340

- This paper examines **the effect of workers’ compensation on time out of work**. It compares two groups of individuals injured **before and after** increases in the maximum weekly benefit amount.
- The increase in maximum benefit raised the benefit amount for high-earnings individuals, but not for low-earnings individuals.
- Time out of work increased for those eligible for the higher benefits and remained unchanged for those whose benefits were constant.

국내외 패널자료 예시

- **PSID, NLS**
- **Daewoo** 패널 (사장됨)
- 한국노동패널
- 한국직업능력개발원 패널

국내외 패널자료 예시

- **Scanner data**
- 카드가맹점 별 카드 매출액 자료
- 카드 소유자별 카드 매출액 실적 자료
- 통신회사 (전화, 인터넷) 보유한 개인별 사용자료
- 전기 등 각종 사용자료 자료
- 학생들의 성적자료
- 기업 회계자료
- 은행의 기업/가계별 대출 및 사후 관리 자료
- 결혼정보회사의 **matching history** 자료

맞음말

- **“All Models are Wrong. Some are Useful.”
(G. Box)**
- **“Panel Data is like Red Wine: The Longer, the Better”**
- **Some Panel Programs written in SAS and Gauss available upon request**