

최근 패널자료 연구의 동향

한치록 (고려대학교 경제학과)

2015. 10. 30.

1-17차년도 한국노동패널 학술대회
서울대학교 호암교수회관 컨벤션센터

Contents

1	정태적 선형 패널 모형	1
1.1	정태적 선형 패널 모형	1
1.2	추정	4
1.3	Random effects와 fixed effects	9
2	동태적 선형 패널 모형	10
2.1	동태적 선형 패널 모형	10
2.2	동태적 선형 패널 모형의 추정	11
2.3	동태적 모형의 설정과 관련된 검정	12
3	횡단면 상관	13
3.1	Spatial approach	14
3.2	Factor approach	14
3.3	횡단면 상관의 검정	15
4	비균형 패널	16
4.1	비균형 패널과 표본이탈	16
4.2	표본이탈로 인한 편향의 교정	17
5	비선형 모형	18
5.1	패널 비선형 모형의 문제점	18
5.2	고정효과 모형	18
5.3	Pooled estimation	19
5.4	임의효과 모형	19

문: 이 글의 형식은?

답: 문답식

1 정태적 선형 패널 모형

1.1 정태적 선형 패널 모형

문: 선형 모형이란?

답: 모수에 대하여 선형인 모형으로 예를 들어

$$\log(y) = \beta_0 + \beta_1 \log(x_1) + \beta_2 x_2 + \beta_3 x_2^2 + u$$

문: 패널 자료란 무엇인가?

답: 복수의 개체들을 복수의 시간에 걸쳐 추적조사하여 얻은 자료. 예를 들어 한국노동패널.

문: 왜 패널 자료를 사용하는가?

답: (i) 관측치 수가 늘어나서 더 효율적인 추정이 가능함 (ii) 관측 안되는 개체별 이질성을 통제할 수 있음 (iii) 개체들의 동태적 성질을 연구할 수 있음 (iv) 횡단면과 시계열이 결합되어 있으므로 횡단면과 시계열 자료로써 할 수 있는 것보다 더 많은 일을 할 수 있음

문: 패널 모형이란 무엇인가?

답: 패널 자료에 내재하는 관계들을 설명하기 위한 모형

문: 선형 패널 모형의 예를 들어 보라.

답: $\log(y_{it}) = \alpha + \beta_1 \log(x_{1it}) + \beta_2 x_{2it} + \beta_3 x_{2it}^2 + \gamma z_i + u_i + \theta_t + \varepsilon_{it}$

문: 첨자의 뜻은? 굳이 첨자를 붙여야 하는가?

답: it 는 개체별 시간별로 상이할 때, i 는 개체별로 다르지만 시간에 걸쳐 동일할 때, t 는 모든 개체에서 동일한데 시간에 따라 다를 때 사용한다. 패널 자료의 경우에는 시간에 따라 변하는지, 개체에 걸쳐 동일한지 등이 중요하므로 꼭 첨자를 올바르게 붙여서 모형을 기술하는 것이 좋겠다.

문: 정태적 선형 패널 모형을 일반적인 형태로 쓰면?

답: $y_{it} = \alpha + x_{it}\beta + v_{it}$, $v_{it} = u_i + \varepsilon_{it}$, $i = 1, \dots, n$, $t = 1, \dots, T$. 여기서 u_i 를 개별효과 (individual effects), ε_{it} 를 고유오차 (idiosyncratic error)라 한다.

문: 이 모형을 왜 “정태적” 모형이라 하는가?

답: y_{it-1} 이 우변에 없어서. 또한 이 모형에서는 t 기의 정보집합이 x_{it} 이며, t 기 이전 또는 t 기까지의 관측정보가 아니다. 만일 t 기까지의 정보라고 하면 x_{i1}, \dots, x_{it} , y_{i1}, \dots, y_{it-1} 이 정보집합에 포함된다.

문: 앞에서는 θ_t 라고 하여 시간 효과가 있었는데 어디로 갔는지?

답: x_{it} 안에 시간더미로 포함되어 있다. Micro panel data의 경우 보통 횡단면 크기 (n)는 크고 시간 길이 (T)는 작으므로 그래도 괜찮다. 하지만 T 가 크면 좀더 주의를 기울여서 명시적으로 표시하는 것이 좋다. 또, 어떤 모형에서는 시간 더미가 아예 없기도 하다.

문: 시기별 더미변수들을 포함시켰더니 거시변수들(GDP 등)이 모두 소거되었다. 어떻게 된 일인가?

답: 거시변수들은 모든 개체 (i)에 대하여 동일한 값을 가지므로 그 효과가 시기별 효과에 의하여 완전히 흡수되어 버린다. 기술적으로, i 에 걸쳐서 동일한 변수들과

시기별 더미변수들은 완전한 공선성을 갖는다.

문: y_{it} 를 설명하는 데에는 x_{it} 뿐 아니라 x_{it-1} 과 y_{it-1} 을 사용하는 것은 자연스러워 보인다. 그럼에도 이런 시차(lagged) 변수들을 우변에 포함시키지 않는다는 것은 잘못 아닌가?

답: 모형이 그렇게 생겼고, 그래서 “정태적” 모형이라 한다.

문: 어떤 분석을 할 수 있는가?

답: 순전한 횡단면 관계를 볼 수도 있고, 모든 관측치들을 통합(pool)하여 분석할 수도 있으며, 동일한 개체 내에서 일어나는 변동만을 이용하여 분석을 할 수도 있다.

문: 정태적 패널 모형에서 일반적으로 가정되는 것 중 주의할 점이 있는가?

답: 과거 현재 미래를 막론하고 idiosyncratic error (ε_{it})에 대하여 설명변수(x_{it})가 외생적(strictly exogenous)이라는 것이 가정된다. 즉, 모든 s 와 t 에 대하여 $E(x_{is}\varepsilon_{it}) = 0$ 라고 가정된다. u_i 가 통제될 때 x_{it} 는 y_{it} 에 영향을 미칠 수 있지만, 그 외의 y_{it} 의 변화(ε_{it} 의 변화)는 x_{is} 에 평균적으로 어떠한 영향도 미치지 않는다. 예를 들어, 지난 기에 y 에 가해진 충격은 이번 기의 설명변수 값에 평균적으로 영향을 미치지 않아야 한다.

문: 이러한 설명변수 외생성의 가정은 타당한가?

답: 경우에 따라 다르다. 여타 유형의 외생성/내생성에 대해서는 동태적 패널 모형을 보라.

문: 정태적 패널 모형을 $y_{it} = \alpha + X_{it}'\beta + u_i + \varepsilon_{it}$ 라고 쓴다면 β 계수가 모든 i 에서 동일하다는 뜻인데 이것은 너무 강한 가정이지 않는가?

답: 그렇다고 볼 수도 있다. β 대신에 β_i 라고 하여 개체별로 효과가 다르게 하거나 β_t 라고 하여 시기별로 효과가 다르게 할 수도 있다.

문: β_t 라고 하면 어떤 방식으로 처리하는가?

답: 각 시간별 더미와 설명변수들의 상호작용을 포함시킴으로 처리한다. 노동패널 처럼 N 이 크고 T 가 작은(크지 않는) 경우에는 그렇게 하여도 모수의 개수가 그리 크지 않으므로 상관 없다.

문: β_i 라고 하면?

답: 만일 각 i 의 계수인 β_i 자체에 관심이 있다면 곤란할 수 있다. 각 i 별로 관측치 수가 T 개밖에 없으므로 추정량이 정확할 수 없다. 이런 경우에는 β_i 들의 평균 $E(\beta_i)$ 라는 단일 모수를 추정하기도 한다. Keyword: “Mean group”

1.2 추정

문: 정태적 패널 모형의 추정방법들을 크게 구분하면?

답: (1) Between-group 추정, (2) POLS와 RE 추정, (3) Within-group 추정. (1)은 횡단면 관계를 분석하는 것이고, (2)는 모든 자료를 통합하여 추정하는 것이며, (3)은 각 개체별로 존재하는 내재적 관계들을 분석하는 것.

1.2.1 횡단면 관계의 분석

문: 패널 자료가 있지만 횡단면 관계를 분석하고 싶다. 어떻게 하면 되겠는가?

답: 각 t 별로 OLS하여 분석하거나, 각 개체별로 변수들의 평균을 구하여 이 평균들을 이용한 OLS를 하는 방법이 있다. 다른 방법도 가능하다. 다만 각 t 에 대하여 분석할 경우 각각의 모수에 대하여 T 개의 추정값이 얻어지므로 무엇을 사용할지 잘 생각해 보아야 한다. 이것은 “어떤 모수 추정값을 보여줄 것인가” 하는 문제이다.

문: 앞 질문에서 “횡단면 관계”란 무엇을 뜻하는가?

답: “ x 값이 1 단위 높은 개체의 y 값이 평균 얼마나 높은가?”라는 질문에 대한 답변.

문: “횡단면 관계”가 아니면 다른 무엇이 있는가?

답: 동일 개체의 시계열 관계가 있다. “동일 개체에서 x 값이 1 단위 증가할 때 y 값은 평균 얼마나 증가하는가?”라는 질문에 대한 답변.

문: Between-group (BE) 추정이란?

답: 각 개체별로 변수들의 (t 에 걸친) 평균값을 계산한 후 이 평균값들을 이용하여 OLS 회귀를 하는 것. 즉, \bar{y}_i 를 \bar{x}_i 에 OLS 회귀하는 것. Stata에서는 “xtreg y x1 x2, be”라고 하면 됨. 이것이 패널 자료에 내재하는 횡단면 관계를 추정한 것.

문: “Between-group”에서 “group”이란?

답: i 가 동일한 관측치들의 집합

문: Stata에서 개체별 평균을 수동으로 구하는 방법은?

답: by id: egen x1bar = mean(x1)

문: BE 추정량이 consistent할 조건은?

답: 설명변수(\bar{x}_i)와 오차항(\bar{v}_i)이 비상관. 만일 개체의 관측 안되는 내재적 속성 (개별효과)과 설명변수가 서로 상관되면 BE 추정량은 inconsistent.

1.2.2 자료의 통합(pooling)

문: 모든 관측치들을 사용하고 싶은데 방법은?

답: 전체 관측치들을 사용하여 OLS를 하거나(pooled OLS, POLS), 오차항의 공분산 구조에 대한 가정을 한 후 이 가정에 기초한 FGLS를 함(Random-effects FGLS, RE).

문: POLS가 consistent하기 위해서는 어떤 조건이 필요한가?

답: x_{it} 와 v_{it} 가 비상관

문: POLS 시 특별히 주의할 점은?

답: v_{it} 는 시간에 걸쳐 상관되기 쉽다. 그러므로 “클러스터” 표준오차를 사용하는 것이 안전하다. Stata에서는 `vce(cluster id)` 옵션 사용. 물론 x_{it} 와 v_{it} 가 비상관 일지 제일 먼저 생각해 보아야 한다.

문: RE FGLS를 하는 경우 어떤 가정을 세우는가?

답: (i) 모든 s 와 t 에서 x_{is} 와 v_{it} 가 비상관, (ii) $v_{it} = u_i + \varepsilon_{it}$ 이고, u_i 와 $(\varepsilon_{i1}, \dots, \varepsilon_{iT})$ 가 서로 비상관이며, $\varepsilon_{i1}, \dots, \varepsilon_{iT}$ 가 서로간에 동분산 및 비상관

문: 가정 (i)이 위배되면 어떻게 되는가?

답: RE 추정량은 inconsistent하므로 RE 추정량은 그 자체로서는 쓸모가 없다. 이 경우 POLS 대신에 RE를 사용하면 더 나을까 하는 생각은 버려야 한다.

문: 가정 (i)은 성립하지만 가정 (ii)가 위배되면 무슨 일이 일어나는가?

답: RE는 진짜 FGLS가 아니므로 효율성을 갖지 않으며, 통상적으로 보고되는 표준오차가 잘못될 수 있다. 표준오차의 문제는 견고한 표준오차를 사용함으로써 해결할 수 있다. Stata에서는 `vce(r)` 또는 `vce(cluster id)` 옵션 사용.

문: 견고한 표준오차를 사용할 때 주의할 점은?

답: 오차가 t 에 걸쳐서는 임의의 이분산과 자기상관을 가질 수 있으나 i 에 걸쳐서는 독립이라는(이분산은 가능) 가정이 성립하여야 한다. 또, 반드시 개체의 수(N)가 커야 한다.

문: 가정 (ii)가 위배되면 RE 추정량보다 더 나은 선형 추정량이 존재할 수 있다는 뜻인데 이 더 나은 추정량을 사용해야 하는 것은 아닌가?

답: 그래도 되지만 RE 추정량이 이미 충분히 좋으므로 그냥 사용해도 괜찮다고 본다. 표준오차만 잘 계산하면 된다.

문: 가정 (ii)가 실제로 위배되는지 검정해 보아야 하는 것은 아닌가?

답: 검정할 필요 없이, 견고한 표준오차(위 참조)를 사용하면 안전하다.

문: 내가 이러저러한 연구주제를 가지고 있다. RE 추정방법을 사용해도 되는가?

답: v_{it} 가 x_{is} 와 상관되는지 확인해 보아야 한다. 특히 v_{is} 에는 개체 i 의 시간불변 속성 (개별효과) u_i 가 포함되어 있을 수 있으므로 이 개별효과가 설명변수들과 상관되어 있는지 점검(생각)해 보아야 한다.

1.2.3 Within-group 추정

문: Within-group이란 무슨 뜻인가?

답: 동일한 i 로 이루어진 관측치들의 집합(“group”) 내(“within”).

문: Within-group (WG) 추정이란 무엇이며, Stata 명령어는?

답: 각 i 별로 within-group deviations ($y_{it} - \bar{y}_i, x_{it} - \bar{x}_i$)를 구한 후 이 within-group deviations을 이용하여 pooled OLS를 하는 것. `xtreg y x1 x2, fe`

문: WG 추정의 핵심은?

답: 각 개체별로 자신의 평균을 차감하므로 개체별 수준의 높고 낮음은 무시됨. 결국 Between-group variability로부터 오는 정보는 전적으로 무시됨. 한 극단에 BE가 있다면 다른 극단에 WG이 있음.

문: WG 추정량은 언제 consistent한가?

답: $v_{it} = u_i + \varepsilon_{it}$ 이고 모든 s 와 t 에서 $E(x_{is}\varepsilon_{it}) = 0$, 즉 “ x ”가 strictly exogenous하면

그러함.

문: WG 추정은 within-group deviations를 효율적으로 이용하는가?

답: 만일 $v_{it} = u_i + \varepsilon_{it}$ 이고 ε_{it} 에 이분산과 자기상관이 없으면 그러하다.

문: $v_{it} = u_i + \varepsilon_{it}$ 인데 ε_{it} 에 이분산이나 자기상관이 있으면?

답: WG 추정량보다 더 효율적인 추정량이 있을 수 있으나, WG 추정량도 충분히 효율적이라고 본다. 다만, 표준오차를 구할 때에는 이분산과 임의의 자기상관(t 에 걸친)이 허용되는 견고한 방법(“클러스터”)을 사용한다. Stata에서는 vce(r) 옵션 이용.

문: 이러한 견고한 방법이 잘 작동하기 위해서는 어떠한 조건이 만족되어야 하는가?

답: 오차가 i 에 걸쳐서 독립이고 개체의 수(N)가 커야 한다.

문: ε_{it} 에 이분산만 있고 자기상관이 없으면 표준오차를 어떻게 구할 것인가?

답: N 이 크면 vce(r) 옵션을 사용하는 클러스터 표준오차 사용. N 이 크지 않는 경우, 만일 i 에 걸쳐서만 이분산이 있고 t 에 걸쳐서는 iid이면 통상적인 White 표준오차(Stata에서 areg와 vce(r) 결합)를 사용하면 됨. 만일 자기상관이 없지만 이분산적일 수 있으면 Stock and Watson (2008, *Econometrica* 76, 155–174)의 방법 사용.

문: N 이 작은데 ε_{it} 에 이분산이나 자기상관이 있으면 표준오차를 어떻게 구할 것인가?

답: ε_{it} 에 t 에 걸친 자기상관이 존재하고 N 이 작으면 GLS의 방법을 선택할 수 있음. 오차항에 AR(1) 자기상관이 존재한다는 가정이 맞다면 Bhargava, Franzini, and Narendranathan (1982, *Review of Economic Studies*)의 방법을 사용할 수 있음(Stata의 xtregar). Greenaway-McGrevy, Han and Sul (2014, *Advances in Econometrics*)은 X-differencing (HPS, 2011, 2014)의 방법을 이용하여 BFN (1982)의 방법을 AR(p)

로 확장하였음. N 이 작고 T 가 큰 경우 적절한 가정 하에서 t_{N-1} 분포와 비교하는 방법도 있음(C. Hansen, 2007, *Journal of Econometrics*).

문: 표준오차 계산 시 시계열 상관의 존재가 중요해 보인다. 시계열상관을 검정하는 방법은?

답: (i) Wooldridge의 패널 교과서에 나온 방법을 Drukker (2003, Testing for serial correlation in linear panel-data models, *Stata Journal*)가 Stata 모듈로 만들었음. 이것은 “serial correlation” 검정이 아니라 “serial correlation + homoskedasticity” 검정 (“white noise 검정”)임. White noise 검정에는 다른 것들도 있음. (ii) 순전한 serial correlation 검정의 방법들도 있으나 N 이 클 것을 요구하고 random walk에 가까울 때 검정력이 낮음(차분에 기초하기 때문).

1.3 Random effects와 fixed effects

문: 임의효과(random effects)와 고정효과(fixed effects)란?

답: 오차항이 $v_{it} = u_i + \varepsilon_{it}$ 일 때, 이 중 개체별 효과(u_i)가 어떠한 것인지 지칭하는 용어이다. u_i 가 독립변수들과 상관(i 에 걸친 상관)되어 있으면 이 개체별 효과를 “고정효과”라 한다. 독립변수들과 상관되어 있지 않은 개체별 효과를 “임의효과”라 한다. u_i 를 두고 “고정효과”이거나 “임의효과”라고 하는 것은 적절하지만, 모형이 고정효과라거나 임의효과라고 하는 것은 타당하지 않다. 하지만 u_i 가 임의효과(고정효과)로 간주되는 모형이라는 뜻에서 “임의효과 모형(고정효과 모형)”이라고 하는 것은 적절하다.

문: 고정효과 추정량과 임의효과 추정량이란?

답: 좁은 의미에서 고정효과 추정량은 WG 추정량을 말하고, 임의효과 추정량은 RE FGLS 추정량을 말한다. 넓은 의미에서는, 개별효과를 고정효과로 간주할 때의

정보를 이용하여 추정하는 방법을 고정효과 추정량이라 한다.

문: 개별효과가 임의효과이면 어떤 추정량들이 consistent하고, 개별효과가 고정효과이면 어떤 추정량들이 consistent한가?

답: 개별효과가 임의효과이면 위에서 이야기한 모든 추정량들이 consistent하다. 개별효과가 고정효과이면 위의 WG 추정량, 1계차분(FD) 추정량 [Stata 명령어는 `reg d.(y x1 x2), vce(c1 id)`] 등이 consistent하다.

문: 개별효과가 고정효과인지 임의효과인지 검정하는 방법은?

답: 널리 사용되는 방법은 WG 추정값과 RE FGLS 추정값을 비교하는 Hausman의 방법이다(`hausman fe re, sigma`). 이 방법은 RE FGLS가 실제 FGLS이기 위해 필요한 가정이 충족되어야만 타당한 검정법이다. ε_{it} 가 이분산적이거나 자기상관되어 있어도 타당한 검정법은 우변에 \bar{X}_i 를 포함시키고 POLS나 RE FGLS를 하여 \bar{X}_i 에 대하여 robust한 검정을 하는 것이다(Wooldridge 패널 교과서).

문: FE vs RE 검정을 꼭 해야 하는가?

답: 표본크기가 작을 때에는 효율적 추정이 중요하였고, RE FGLS가 특히 효율적이므로 RE 추정값의 타당성을 확인하기 위해서는 이 검정이 중요하다. 반면 노동패널의 경우처럼 표본크기(N)가 매우 크다면 고정효과 추정을 하여도 충분히 보인다.

2 동태적 선형 패널 모형

2.1 동태적 선형 패널 모형

문: 흔히 사용되는 동태적 선형 패널 모형은?

답: $y_{it} = \alpha + \delta y_{it-1} + X'_{it}\beta + u_i + \varepsilon_{it}$

문: 정태적 모형과는 달리 동태적 모형의 경우에는 임의효과와 고정효과에 관한 논의가 별로 없다. 왜 그런가?

답: 동태성 자체가 개별 개체 내의 관계를 의미하므로 cross sectional variation을 이용하는 “임의효과” 접근법은 별로 사용되지 않음

문: 동태적 모형에서 설명변수의 외생성이나 내생성은 어떤 종류가 있는가?

답: (strictly) exogenous, predetermined, (contemporaneously) endogenous 세 종류.
Exogenous: $E(X_{is}\varepsilon_{it}) = 0$ for all s, t ; predetermined: $E(X_{is}\varepsilon_{it}) = 0$ for all $s \leq t$; endogenous: $E(X_{is}\varepsilon_{it}) = 0$ for all $s < t$

문: X_{it} 가 strictly exogenous하지 않을 수 있다면 모형이 어떻게 식별되는가?

답: 오차항 ε_{it} 에 특별한 가정이 추가된다. 이 가정은 ε_{it} 가 t 에 걸쳐 uncorrelated라는 것.

2.2 동태적 선형 패널 모형의 추정

문: “차분GMM”과 “시스템GMM”의 차이는?

답: 차분GMM은 $\Delta\varepsilon_{it}$ 가 $X_{1,i1}, \dots, X_{1,iT}, X_{2,i1}, \dots, X_{2,it-1}, X_{3,i1}, \dots, X_{3,it-2}, y_{i1}, \dots, y_{it-2}$ 와 uncorrelated라는 조건을 (특정한 방식으로) 이용하는 추정방법이고, 시스템GMM은 여기에 Δy_{it-1} 이 $u_i + \varepsilon_{it}$ 와 uncorrelated라는 조건을 추가로 이용하는 방법임.

문: 두 추정량은 어떤 조건들을 필요로 하는가?

답: 차분GMM은 ε_{it} 가 t 에 걸쳐 uncorrelated라는 표준적 가정에 기초하고, 시스템GMM은 이에 추가하여 y_{it} 가 mean-stationary하다는 조건을 필요로 한다. 효율성은

시스템GMM이 훨씬 좋지만 mean-stationarity라는 조건은 항상 만족되는 것이 아니므로 주의해야 한다.

문: Stata 명령어는?

답: xtabond와 xtdpdsys. 오차항이 i 와 t 에 걸쳐 이분산적일 수 있으므로 vce(r) 옵션을 사용하여 클러스터 표준오차를 구하도록 한다.

문: Stata 명령어에서 “one step”이란 무엇을 의미하는가?

답: ε_{it} 가 IID라는 가정 하에서는 1단계 추정으로도 efficient한 추정량(xtabond의 경우에만 해당, xtdpdsys의 경우에는 suboptimal)을 구할 수 있다. “One step”이란 이 추정량을 의미한다.

2.3 동태적 모형의 설정과 관련된 검정

문: 위의 추정법들은 GMM으로서 특정한 적률조건들(moment conditions)이 충족된다는 가정을 바탕으로 한다. 이 적률조건들이 충족되는지 어떻게 검정할 수 있는가?

답: Sargan 검정(estat sargan)이나 Hansen 검정. 이 검정들의 귀무가설은 적률조건들이 모두 충족된다는 것이므로 귀무가설을 기각한다는 것은 모형이 잘못 설정되었음을 뜻한다.

문: 위의 동태적 패널 모형에서는 ε_{it} 가 t 에 걸쳐 uncorrelated일 것이 중요하다. 어떻게 이 “자기 비상관” 조건이 충족되는지 검정할 수 있는가?

답: Arellano and Bond (1991) 검정(estat abond). 이 검정은 차분한 오차항($\Delta\varepsilon_{it}$)에 시계열상관이 없음을 검정하는 것이므로, ε_{it} 에 시계열상관이 없다면 $\Delta\varepsilon_{it}$ 는 1계 시계열상관을 가져야 하고(Order = 1에서는 기각), 2계 이상의 시계열상관은 갖지 말아야 한다(2 이상의 order 기각하지 않음).

문: 우변에 y_{it-2} 나 그 이상의 차수를 포함시켜도 되는가?

답: 그렇다.

문: 차수를 어떻게 결정하는가?

답: 위의 모형설정 검정들이 제대로 되는 가장 작은 차수로 하면 된다. 그보다 더 큰 차수를 이용하여도 괜찮아야 한다.

3 횡단면 상관

문: 횡단면 상관이란 무엇인가?

답: i 에 걸친 상관.

문: 어떤 문제가 있는가?

답: (i) v_{it} 에 i 에 걸친 상관이 있으면 이를 고려한 표준오차를 계산하여야 한다. (ii) 횡단면 상관을 야기하는 요인들을 통제하고 나서 β 를 추정하고자 한다.

문: 추정량(예를 들어 WG 추정량)이 consistent하다면 횡단면 상관을 고려한 표준오차는 어떻게 계산하는가?

답: i 에 걸쳐서 임의의 상관관계를 가정한 후 i 에 걸쳐서 평균을 구한 후 t 에 걸쳐서 stationary하다는 가정 하에서 Newey-West의 방법과 유사하게 처리할 수 있다 (Vogelsang 2012, Driscoll and Kraay 1998). 이 방법이 작동하려면 T 가 커야 한다.

문: 횡단면 상관을 야기하는 요인들을 통제하는 방법은?

답: 두 가지 방법이 사용된다. (i) “Spatial” approach, (ii) “Factor” approach.

3.1 Spatial approach

문: “Spatial” approach에 대하여 좀더 자세히 설명한다면?

답: y_{it} 방정식의 우변에 $\sum_{j \neq i} w_{ij} y_{jt}$ 를 포함시킨다 (spatial lag model). 여기서 w_{ij} 는 모두 알려져 있으며 (“economic distance”, 예를 들어 국경을 마주하는 국가들에 flat한 가중치를 줌), 모든 i 에서 $\sum_{j \neq i} w_{ij} = 1$ 이다. 가장 간편한 추정법은 x_{it} 와 $\sum_{j \neq i} w_{ij} x_{jt}$ 를 도구변수로 사용하여 추정 (예를 들어 2SLS, Kelejian and Prucha 2001) 하는 것.

문: 우변에 $\sum_{j \neq i} w_{ij} y_{jt}$ 를 포함시키지 않고 오차항에 spatial correlation을 갖도록만 하는 모형은 없는가?

답: 있다. 이런 모형을 Spatial error model이라고 한다. 이 모형은 오차항의 횡단면 상관을 모형화한 것으로 인접 개체들을 “통제” 하는 것과는 다른 방법이다. WG 추정량도 consistent하지만 표준오차 계산이 어렵다. Spatial error model이 맞는 경우 그 MLE는 더 효율적이다. Lee and Yu (2010).

3.2 Factor approach

문: “Factor” 모형을 한 마디로 설명하면?

답: 소수의 공통된 요인이 모든 개체들에 영향을 미치도록 하는 모형으로서, 일반적으로 다음과 같음

$$y_{it} = \lambda_i' F_t + X_{it}' \beta + \varepsilon_{it}$$

문: 이 모형은 얼마나 일반적인가?

답: 고정효과 모형, 고정효과 + 기간별 효과 모형, $\alpha_i \lambda_i$ 같은 모형 등을 모두 포함함

문: 이것이 어떻게 하여 횡단면 상관을 설명하는가?

답: X_{it} 의 영향을 제거하더라도 t 기에 y_{it} 와 y_{jt} 의 공분산은 $\lambda_i' E(F_t F_t') \lambda_j$ 로서 0이 아닐 수 있음. y_{it} 에 존재하는 횡단면 상관이 F_t 가 공유되기 때문에 발생한다고 봄

문: Factor 모형은 횡단면 상관을 모형화하는 데에서만 사용되는가?

답: 그렇지 않다. 노동경제학에서 Y_{it} 가 임금이라면 λ_i 는 능력, 동기부여, 근면성 등 관측되지 않는 속성이고 F_t 는 이러한 관측되지 않은 요소들의 가격일 수 있다 (application들에 대해서는 Bai, 2009, *Econometrica*, p. 1233 참조).

문: β 의 추정방법은?

답: (i) Pesaran (2006)은 변수들의 각 시점별 횡단면 평균들을 우변에 추가하여 통제하는 방법을 제안하였음. (ii) Bai (2009)는 $\sum_i \sum_t (y_{it} - \lambda_i' F_t - X_{it}' \beta)^2$ 을 최소화하는 LS 방법에 대하여 분석하였음 (LS의 구현을 위해 iterated least squares를 이용하며, 추정시 주요인분석이 사용됨). (iii) Ahn, Lee and Schmidt (2013)는 large N , fixed T 의 경우 일반적인 상황에서 좋은 성질을 갖는 nonlinear GMM의 방법을 제시함.

문: 주의할 점은?

답: (i) Pesaran (2006)의 방법이 작동하기 위해서는 F_t 가 \bar{y}_t 와 \bar{X}_t 의 선형결합으로 표현될 수 있어야 함. (ii) Iterated least squares 추정량은 일반적으로 (이분산, 자기상관 존재 시) 분포가 편향되어 있으며, Bai (2009)가 편향의 교정방법을 제시함. (iii) Ahn, Lee and Schmidt (2013)의 방법은 비선형이라서 계산이 복잡함

3.3 횡단면 상관의 검정

문: 횡단면 상관을 검정할 방법은?

답: N 이 고정되고 T 가 큰 경우 Breusch and Pagan (1980)의 방법을 사용할 수 있음 (Stata의 xttest2 모듈 참조). 이 방법은 $\hat{\rho}_{ij}^2$ 의 합을 이용함. N 이 큰 경우, Pesaran,

Ullan and Yamagata (2008, *Econometrics Journal*)가 $\hat{\rho}_{ij}^2$ 들을 적절히 정규화하여 합하는 방법을 제안하였음. Pesaran (2004, “General diagnostic tests for cross section dependence in panels”)은 $\hat{\rho}_{ij}$ 들을 사용하는 방법을 제안하였음.

4 비균형 패널

문: 비균형 패널이란?

답: 균형 패널이란 자료 내 모든 개체들이 동일한 기간 동안에 빠짐없이 관측되어 있는 패널자료를 말하며, 비균형 패널이란 그렇지 않은 패널자료를 말한다.

4.1 비균형 패널과 표본이탈

문: 비균형 패널이 생성되는 이유는?

답: 표본이탈 (attrition, 마모)과 신규 진입

문: 비균형 패널을 다루는 방법은?

답: 균형화된 부분패널을 이용하거나 전체 비균형패널을 모두 이용하면서 표본이탈을 무시하는 방법과, 전체를 이용하면서 표본이탈을 설명하는 방법이 있다.

문: 표본이탈은 언제 문제가 되는가?

답: 표본이탈이 내생변수에 의존할 경우 문제(편향)가 발생한다. 예를 들어, 종속 변수가 임금인데 임금이 너무 낮거나 너무 높은 경우 표본으로부터 이탈하는 경우. 표본이탈이 외생변수에만 의존할 경우 통상적인 WG 추정량은 여전히 일치성을 갖는다.

4.2 표본이탈로 인한 편향의 교정

문: 표본이탈은 표본선택의 일종으로 간주되곤 한다. 둘을 완전히 동일한 방식으로 다루어도 되는가?

답: 그렇지 않다. 표본선택의 경우에는 표본선택을 결정하는 외생변수가 관측되는 상황을 고려하는 반면, 표본이탈의 경우에는 자명한(deterministic) 경우를 제외하고는 외생변수가 아예 관측되지 않으므로 문제가 간단하지 않다.

문: 표본이탈이 y_{it-1} 에 의하여 발생한다. 표본이탈로 인한 편향을 교정할 수 있는가?

답: 교정할 수 있다. 이는 표본으로부터 이탈하여도 표본이탈을 야기하는 y_{it-1} 과 이를 결정하는 외생변수인 x_{it-1} 이 관측되기 때문이다. 가장 간단한 교정방법은 Heckman의 방법(inverse Mills ratio 이용)을 이용하는 것이다. 자세한 내용은 Wooldridge의 대학원 교과서 참조.

문: 표본이탈이 y_{it} 에 의존한다. 표본이탈로 인한 편향을 교정할 수 있는가?

답: 독립변수들이 이탈 이후에도 여전히 관측되는 특수한 상황(나이처럼 deterministic 한 변수들의 경우)이 아니면 교정이 안된다.

문: 표본이탈률이 낮는데, 그렇다면 편향이 작을 것 아닌가?

답: 그렇지 않을 수도 있다. 특히 모형의 설명력이 낮은 경우 편향이 매우 클 수도 있다.

5 비선형 모형

5.1 패널 비선형 모형의 문제점

문: 비선형 모형은 왜 특별히 취급하는가?

답: 선형모형의 경우에는 차분이나 WG 변환으로써 u_i 를 제거할 수 있으나, 비선형 모형의 경우 고정효과(u_i)를 처리할 수 없기 때문.

문: 예를 들면?

답: 이항반응 모형 $y_{it} = I(\alpha + X'_{it}\beta + u_i + \varepsilon_{it} > 0)$ 의 경우 차분이나 다른 변환을 해도 u_i 가 소거되지 않는다.

문: u_i 가 없는 모형은?

답: 아무런 문제가 없다. 하지만 패널분석의 의의도 크지 않다.

5.2 고정효과 모형

문: 고정효과는 어떻게 처리하는가?

답: 특별한 몇 가지 경우는 적당한 변환을 통하여 u_i 를 모형으로부터 제거할 수 있음. 다른 경우에는 Chamberlain의 방법을 사용하나(예를 들어 $u_i = \bar{X}'_i\delta + a_i$ 라고 모형화한 후 a_i 가 임의효과라고 가정) 이는 실상 “고정효과” 모형이 아니다.

문: 고정효과 추정이 가능한 비선형 모형으로는 어떤 것이 있는가?

답: • 이항반응 모형: 정태적 모형에서 Fixed effects logit (probit은 안됨) 동태적 로짓 모형에서 Honoré and Kyriazidou (2000)의 추정방법. 이상은 모수적 방법. 준

모수적 방법으로는 Manski (1987)의 maximum score estimation, Horowitz (1992)의 smoothed MSE, Lee (1999) 등 • Truncated and censored regressions: Pairwise trimmed least-squares, LAD (Honoré 1992) • Count data: Poisson CMLE • Nonnegative data: Gamma CMLE. 등등

5.3 Pooled estimation

문: Pooled estimation의 예를 들면?

답: $y_{it} = I(\alpha + X_{it}'\beta + v_{it} > 0)$ 이고 $v_{it} \sim N(0,1)$ 이라고 한다. 만일 X_{it} 와 v_{it} 가 서로 독립이라면 (u_{it} 가 임의효과) 이렇게 가정해도 괜찮아 보인다.

문: 위의 예에서 그냥 pool하여 추정해도 괜찮은가?

답: “Pooled” 방법은 X_{it} 와 v_{it} 가 서로 독립이고 $v_{it} \sim N(0,1)$ 이라는 가정만을 이용한다. y_{it} 에 내재하는 시계열 상관은 전적으로 무시하는데 그래도 괜찮다.

문: 이 pooled estimation은 무엇을 하지 않는가?

답: (i) u_i 가 X_{it} 와 상관될 수 있는 가능성을 허용하지 않는다. (단, 나중에 보듯이 Chamberlain의 방법을 사용하여 이를 고려하기도 한다.) (ii) 또한 v_{it} 에 존재하는 시계열상관을 전혀 고려하지 않는다. 첫째 점이 고정효과 추정과의 차이점이고 둘째 점이 임의효과 추정과의 차이점이다.

5.4 임의효과 모형

문: 임의효과는 어떻게 처리하는가?

답: 임의효과 u_i 가 ε_{it} 와 독립이고 (설명변수들과도 독립이며) 특정한 분포를 갖는다고

가정하고 나서 MLE를 함. 예를 들어 $y_{it} = I(X'_{it}\beta + u_i + \varepsilon_{it} > 0)$ 이라면 $u_i \sim N(0, \sigma_u^2)$ 이라고 가정하고 나서 MLE.

문: 응용연구에서는 어떤 방법을 사용하는 것이 좋겠는가?

답: \bar{X}_i 를 우변에 포함시키고 나서 (또는, X_{i1}, \dots, X_{iT} 를 모두 설명변수로 추가) pooled 추정이나 RE 추정 (Chamberlain의 방법). 동태적 모형도 추정할 수 있음 (Wooldridge 대학원 교과서 참조, Example 15.6: Dynamic Women's LFP Equation).

문: 구체적으로?

답: y_{it} 를 $x_{1,it}$ 와 $x_{2,it}$ 에 대하여 회귀하는 이항반응모형이라면 $\bar{x}_{1,i}$ 와 $\bar{x}_{2,i}$ 를 우변에 포함시켜 RE 추정을 한다.

문: Stata 코드는?

답: 다음과 같다.

```
xtset id year
by id: egen x1bar = mean(x1)
by id: egen x2bar = mean(x2)
xtprobit y x1 x2 x1bar x2bar, re
```

문: 동태적 모형은 어떻게 처리하는가?

답: Wooldridge (2010, 대학원 교과서)는 이상의 Chamberlain의 방법에 추가하여 우변에 y_{i0} 까지 포함시켜 RE 추정을 하는 것을 제안한다(예: Wooldridge, 2010, Example 15.6).