강 석 훈 교수지도

석사학위 청구논문

패널 데이터를 이용한 이진변수 모형 : 서베이

2007

성신여자대학교 대학원

경제학과

김 주 봉

패널 데이타를 이용한 이진변수 모형 : 서베이

Binary Choice Model using Panel Data : A survey

강석훈 교수지도

이 논문을 석사학위논문으로 제출함

2006년 11월

성신여자대학교 대학원

경제학과

김주봉

인 준 서

김주봉의 석사학위 논문으로 인준함.

심사위원 _____________________인

심사위원 _____________________인

심사위원 _____________________인

성신여자대학교 대학원

논 문 개 요

본 논문은 종속 변수가 이진변수인 panel data 모형분석 시 발생하는 몇 가지 사안들을 검토하여, 실증 연구자가 적절한 모형을 비교, 검토할 때 도움이 될 수 있도록 하였다. 우선, 이산형 종속변수 모형의 기반이 되는 이진변수 모형과 패널 데이타의 개념을 검토하였다. 패널 데이타 이진변수 모형으로서 정태적 패널 데이타와 동태적 패널 데이타로 나누어 살펴보았으며 정태적 패널 모형을 고려하는데 있어서는 전통적인 패널 데이타 모형으로서 고정효과 모형과 임의효과 모형의 추정 방법 및 추정 시 발생할 수 있는 제 문제점을 살펴보았다. 추정 방법으로는 오차항의 분포를 가정하는 모수적인 방법 뿐 만 아니라 오차항의 분포를 가정하지 않는 준모수적 방법도 고려하였다.

다음으로는 동태적 데이타 모형의 제 문제점들을 살펴보았다. 그리고 횡단면 자료 시 발생될 수 있는 이분산성과 시계열 자료에서 발생될 수 있는 계열 상관의 문제점을 고려하는 모형으로 확장하였으며, 마지막으로 이러한 이론적 측면을 반영한 실증 분석 사례와 현재까지 진행 중인 이진변수 패널 모형의 연구 현황을 간략하게 정리하였다.

목 차

1. 개관

패널 데이타는 횡단면 데이터와 시계열 데이터를 결합한 자료로서 실증 경제학에서 매우 중요한 역할을 한다. 패널 데이타를 통해 횡단면 데이타로 답해질 수 없는 미시 경제적 동태(dynamic) 행위에 관한 질문에 답할 수 있으며, 서로 다른 시점, 동일한 개체들에 대한 반복적인 관측치를 포함하기 때문이다. 최근 패널 데이터 분야의 계량경제학 발전은 패널 데이타가 적용 가능한 광범위한 분야의 경제 모형의 확장을 이끌었고, 정성적인 분석을 가능하게 하는 이산형 변수를 사용한 분석 역시 널리 사용된다. 본 논문은 종속 변수가 이진변수인 패널 데이타 모형분석 시 발생하는 몇 가지 사안들을 검토하여, 실증 연구자가 적절한 모형을 비교, 검토할 때 도움이 될 수 있도록 하는 것이 목적이다.

논문의 구성은 다음과 같다. 우선, 이산형 종속변수 모형의 기반이 되는 이진변수 모형의 정의를 검토하고, 패널 데이타의 개념을 재검토 한다. 이를 바탕으로 패널 데이타 이진변수 모형의 기초개념을 설명한다. 정태적 패널 데이터 모형을 고려하는데 있어서는 전통적인 패널 데이터 모형으로서 고정효과 모형과 임의효과 모형의 추정 방법을 살펴보고, 추정 시 발생할 수 있는 incidental parameter problem등의 제 문제점을 알아볼 것이다. 또한 선형 모형 추정 시에는 발생하지 않으나 비선형 모형에서 발생하는 문제점들을 고찰해 볼 것이다. 추정 방법으로는 오차항의 분포를 가정하는 모수적인 방법 뿐 만 아니라 오차항의 분포를 가정하지 않는 경우의 방법도 제시한다. 다음으로는 동태적 패널 데이터 모형의 일반적인 모형을 소개하며, 동태모형에서 발생할 수 있는 초기조건(initial condition)의 문제와 상태의존성

(state dependence)의 제 문제점들을 알아볼 것이다. 그리고 나서 횡단면 자료 시 발생될 수 있는 이분산성과 시계열 자료에서 발생될 수 있는 계열 상관의 문제점을 고려하는 모형으로 확장하고자 한다. 마지막으로 이러한 이론적 측면을 반영한 실증 분석 사례와 현재까지 진행 중인 이진변수 패널 모형의 연구 현황을 간략하게 정리하면서 마무리 짓고자 한다.

2. 재검토

2.1 이진변수모형

2.1.1. 추정

이진변수 모형은 종속변수 y 가 두 가지 값, 즉, 사건이 발생한 경우에 1, 그렇지 않은 경우 0인 값을 갖는 모형을 말한다. 예를 들어 주어진 해의 내구재 구매 여부, 노동시장 참가 여부, 대학 진학 결정 여부 등의 결정요인 분석 시 사용된다. 이산형 결과인 y는 잠재적 연속확률 변수의 관측되는 부분으로 나타날 수 있다. 연속적인 잠재 확률변수 $y*$ 가 설명변수 x 의 선형 함수라고 가정하자.

(2.1) $y* = \beta' x + \upsilon$ (오차항 υ는 평균이 0이며, x에 독립)

(2.2) $$y = \begin{cases} 1 & \text{if } \; y* > 0 \\ 0 & \text{if } \; y* \leq 0 \end{cases}$$

사건 발생확률 y의 기대값은

$$(2.3)\quad \begin{aligned} E(y|x) &= 1 \bullet pr(\upsilon > -\beta' x) + 0 \bullet pr(\upsilon \le -\beta' x) \\ &= pr(\upsilon > -\beta' x) \\ &= pr(y=1|x) \end{aligned}$$

가장 흔히 사용되는 모수적 이진변수모형은 다음과 같이 요약된다.

· 선형확률모형 (Linear Probability Model , LPM)

$$(2.4)\quad F(w) = w$$

· 프로빗 모형 (Probit Model)

$$(2.5)\quad F(w) = \int_{-\infty}^{w} \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} du = \Phi(w)$$

· 로짓 모형 (Logit Model)

$$(2.6)\quad F(w) = \frac{e^w}{1+e^w}$$

선형확률모형은 최소자승법(LS) 혹은 가중최소자승법(WLS)으로 추정될 수 있다. 그러나 LPM은 $\beta' x$의 확률이 0과 1사이에 놓여 지지 않는 경우 설명되지 않는 단점을 갖고 있다. 반면 로짓, 프로빗 모형에서는 $\beta' x$가 0과 1 사이에 놓여지게 된다. 로짓, 프로빗모형은 비선형 최소자승 추정법으로도 추정가능하다. 그러나 동 추정량은 일치성은 갖으나 효율적이지 않아 실제적으로 별로 사용되지 않는다. 반면 MLE는 일치추정량이며, 대표본에서 정규분포를 따르고 효율적이다. 즉, $Var(\hat{\beta}_{NLLS}) \ge Var(\hat{\beta}_{ML})$. 이제 N개의 확률표본 (y_i, x_i)이 있는 경우, 이를 MLE로 설명하는 경우를 고려해보자. 먼저 우도함수는 다음과 같다.

(2.7) $$L=\prod_{i=1}^{n} F(\beta' x_i)^{y}[1-F(\beta' x_i)]^{1-y_i}$$

우도함수의 로그값을 β에 대해 미분하여 얻어진 1차 미분값과 2차 미분값은 다음과 같다.

(2.8) $$\frac{\partial \log L}{\partial \beta}=\sum_{i=1}^{n}\frac{[y_i-F(\beta' x_i)]}{F(\beta' x_i)[1-F(\beta' x_i)]}F'(\beta' x_i)x_i$$

(2.9) $$\frac{\partial^2 \log L}{\partial \beta \partial \beta'}=-\sum_{i=1}^{n}\frac{y_i}{F^2(\beta' x_i)}+\frac{1-y_i}{[1-F(\beta' x_i)]^2}[F'(\beta' x_i)^2]$$

$$+\sum_{i=1}^{n}\frac{y_i-F(\beta' x_i)}{F(\beta' x_i)[1-F(\beta' x_i)]}F''(\beta' x_i)\, x_i x_{i'}$$

($F'(\beta' x_i)$와 $F''(\beta' x_i)$는 $F(\beta' x_i)$의 1계 미분, 2계미분을 의미함).

우도함수 (2.7)이 오목하다면, β의 최우추정량 (MLE)를 찾기 위해 다음의 최적화(optimization)방법이 사용될 수 있다[1)]

Newton-Raphson Method

(2.10) $$\hat{\beta}^{(j)}=\hat{\beta}^{(j-1)}-\left(\frac{\partial^2 \log L}{\partial\beta\partial\beta'}\right)^{-1}_{\beta=\hat{\beta}^{(j-1)}}\left(\frac{\partial \log L}{\partial\beta}\right)_{\beta=\hat{\beta}^{(j-1)}}$$

Method of Scoring,

(2.11) $$\hat{\beta}^{(j)}=\hat{\beta}^{(j-1)}-\left[E\frac{\partial^2 \log L}{\partial\beta\partial\beta'}\right]^{-1}_{\beta=\hat{\beta}^{(j-1)}}\left(\frac{\partial \log L}{\partial\beta}\right)_{\beta=\hat{\beta}^{(j-1)}}$$

또한 특정 값 x에 대한 y의 반복적인 관측치가 있는 경우, β를 추정하기 위

1) $\hat{\beta}^{(0)}$는 임의의 초기값, $\hat{\beta}^{(j)}$는 j번째 값을 의미함.

해 가중최소자승법(weighted least squares)을 적용할 수 있으며, 그 결과 추정량을 minimum-chi-square estimator라 불린다. 이는 MLE와 같이 점근적 효율성을 지니며, 계산은 MLE보다 단순하다. 더구나, 표본이 유한한 경우 minimum-chi-square estimator는 MLE보다 더 작은 평균자승오차를 가질 것이다.[2)]

그러나 이러한 통계적 매력에도 불구하고 minimum-chi-square estimator는 설명변수 벡터 값들에 대한 반복적인 관측치를 필요로 하므로 MLE보다 덜 유용하다. 더 나아가 x가 주어진 경우 $y=1$의 확률이 0 혹은 1이면, minimum-chi-square estimator는 정의되지 않는다.[3)] 그러나 MLE는 여전히 적용되므로 MLE방법에 집중한다.

2.1.2. 한계효과

이진변수모형의 한계효과는 모형에 따라 다르며, 로짓 모형의 경우 다음과 같다. (2.12) $E[y|\overline{x}] = \exp(\beta'\overline{x})/[1+\exp(\beta'\overline{x})] = \Lambda(\beta'\overline{x})$

$$\delta = \partial E[y|\overline{x}]/\partial\overline{x} = \Lambda(\beta'\overline{x})[1-\Lambda(\beta'\overline{x})]\beta$$

$$G = \Lambda(\beta'\overline{x})[1-\Lambda(\beta'\overline{x})]\{I+[1-2\Lambda(\beta'\overline{x})\beta\overline{x}'\}$$

즉, $F(\alpha+\beta_1 x_1 \ldots)/\partial x = prob*(1-prob)*\beta$

한편, 프로빗 모형의 경우는 다음과 같이 계산된다.

(2.13) $E[y|\overline{x}] = \Phi(\beta'\overline{x})$

$$\delta = \partial E[y|\overline{x}]/\partial\overline{x} = [\phi(\beta'\overline{x})]\beta$$

$$G = [\phi(\beta'\overline{x})]\{I-(\beta'\overline{x})\beta\overline{x}'\}$$

---

2) Amemiya(1974,1976,1980b); Berson(1944,1955,1957,1980); Fergson(1958); Neyman(1949)).

3) Amemiya(1981).

즉, $F(\alpha+\beta_1 x_1 \ldots)/\partial x = normal\ density * \beta$

2.1.3. 모형설정검정 (Specification Test)

일반적인 선형회귀 모형 $y = x_1\beta_1 + x_2\beta_2 + \varepsilon$에서 x_2를 누락시키고 최소자승추정량 b_1을 계산하면 다음과 같이 된다. (2.14) $E[b_1] = \beta_1 + [x_1'x_1]^{-1}x_1'x_2\beta_2$. x_1과 x_2가 직교되지 않거나 β_2=0 이 아니면, b_1은 편향을 가진다. 만일 이분산성을 무시할 경우 최소자승추정량은 비록 불편성과 일치성을 만족하지만, 효율적이지 못하며, 일반적인 표본 공분산 행렬의 추정치는 적절하지 못하다. Yatchew and Griliches(1984)는 프로빗, 로짓모형에서도 이와 같은 주제를 검토하였다. 그들은 이진변수 모형에서 다음과 같은 사실을 발견했다. (1) x_1과 x_2를 포함하는 모형에서 x_2가 누락되면, $plim\hat{\beta}_1 = c_1\beta_1 + c_2\beta_2$이다. ($c_1$과 c_2 는 미지의 모수를 자세히 기록해 놓은 것임). 이 경우, 누락변수가 모형 내 포함된 변수와 상관관계가 없더라도 모형에 포함된 변수의 계수는 일치추정량이 되지 못한다. (2) 회귀식에서 오차항이 이분산성을 띤다면, MLE는 일치성을 만족하지 못하고 공분산 행렬은 적절하지 못하다.

프로빗모형은 미시 경제학 데이타에서 자주 사용되고, 종종 이분산성을 띠므로 두 번째 결과가 특히 문제가 된다. 가설 검증법으로는 왈드검정(wald test), 우도비 검정(LR), 라그랑지승수 검정(LM)가 사용될 수 있는데, LM은 제약된 모형의 추정치로 수행되므로, 계산상의 편이함이 있다. 특히 이분산성을 검증하는 데는 더욱 그렇다.

1) 누락변수

$$(2.16)\qquad \begin{aligned} H_0 &: y^* = \beta'_1 x_1 + \qquad + \epsilon \\ H_1 &: y^* = \beta'_1 x_1 + \beta'_2 x_2 + \epsilon \end{aligned}$$

$\beta_2=0$인 가설을 검증하며, LM검정을 행할 수 있다. 교란항(disturbance)에 이분산성이 있으면 MLE는 일치추정량이 되지 못하며 공분산 행렬도 적절하지 못하다.

2) 이분산성

이분산 검정법의 일반적인 접근은 어떠한 형태의 이분산이 교란항(disturbance)에 제시된다면 잔차(residual)에서 식별가능하다는 아이디어에 기초하며[4] 이러한 검정은 다양한 형태의 이분산을 검정해 주므로 강건(robust)하다. 모수적 모형의 4가지 검증은 다음과 같다.

2-1) 왈드 검정 (Wald Test)

이 검증법은 제약되지 않은 b_1의 추정치가 0에 가까우면 귀무가설을 채택한다는 것이 아이디어이다.

$$(2.17)\qquad \begin{aligned} \xi_w &= \beta_1'[Var(\beta_1)]^{-1}\beta_1 \\ &= \beta_1'[\widehat{I_{11}} - \hat{I}_{12}\hat{I}_{22}^{-1}\hat{I}_{21}]\beta_1 \end{aligned}$$

($\hat{I}$는 information matrix의 일치추정량)

---

4) White test, the Goldfeld-Quandt test, Breusch-Pagan LM test는 최소자승 잔차의 행위에 기초한 이분산 검정법임.

귀무가설이 얻어지면, 자유도 r의 이 통계치는 점근적으로 χ^2 분포를 따른다. $\chi^2_{95\%}(r)$는 95% 임계치이고, 5% 유의수준의 귀무가설을 왈드검정에 적용하는 경우, $\xi_w < \chi^2{}_{95\%}(r)$이면 귀무가설을 채택하고 $\xi_w > \chi^2{}_{95\%}(r)$이면 귀무가설을 기각한다.

2-2) 우도비 검정 (Likelihood Ratio Test)

이 검증법은 두 MLE(제약된 우도 함수와 제약되지 않은 우도함수를 극대화 함으로써 일반화됨)의 비교에 기반을 둔다. 통계량은 다음과 같고, 이는 점근적으로 $\chi^2(r)$를 따른다. (2.18) $\xi_{LR} = -2(\ln L_c - \ln L_u)$.
만약 $\xi_{LR} < \chi^2{}_{95\%}(r)$이면 귀무가설을 채택하고, $\xi_{LR} > \chi^2{}_{95\%}(r)$이면 귀무가설을 기각한다. 이러한 과정은 점근적으로 wald test와 동일하다.

2-3) LM 검정

귀무가설이 맞을 경우, 두 추정량 β와 β^c는 서로 매우 비슷하다.
(b^c는 제약된 식의 추정량, b^u는 제약되지 않은 식의 추정량)

(2.19) $$\frac{\partial \log[L(y:b^u)]}{\partial \beta} = \begin{bmatrix} \dfrac{\partial \log(L)}{\partial \beta_1}(y:b^u) \\ \dfrac{\partial \log(L)}{\partial \beta_2}(y:b^u) \end{bmatrix} = 0$$

$$\frac{\partial \log[L(y:0:b_2^c)]}{\partial \beta} = \begin{bmatrix} \frac{\partial \log(L)}{\partial \beta_1}(y:0:b_2^c) \\ \frac{\partial \log(L)}{\partial \beta_2}(y:0:b_2^c) \end{bmatrix} = \begin{bmatrix} \frac{\partial \log(L)}{\partial \beta_1}(y:0:b_2^c) \\ 0 \end{bmatrix}$$

통계량은 $\xi_{LM} = \frac{\partial \log[L(y:b^u)]}{\partial \beta_1}\left(\hat{I}_{11} - \hat{I}_{12}\hat{I}_{22}^{-1}\hat{I}_{21}\right)^{-1}\frac{\partial \log(L(y;b^c)}{\partial \beta_1}$ 로 정의되며, 이는 점근적으로 왈드 통계량 ξ_w, 우도비 통계량 ξ_{LR}과 같아진다. 검증은 $\xi_{LM} < \chi^2{}_{95\%}(r)$이면 귀무가설을 채택하고, $\xi_{LM} > \chi^2{}_{95\%}(r)$이면 귀무가설을 기각한다. 실제, LM 검증이 가장 일반적이다.

2-4) 하우스만 검정 (Hausman Specification Test)

장애모수(nuisance parameter) b의 두 추정치와 관련되며, 통계량의 형태는 (2.20) $\xi_H = (\beta-\beta^c)'[\widehat{Var}(\beta)-\widehat{Var}(\beta^c)]^{-1}(\beta-\beta^c)$와 같다.

($[...]^{-1}$은 일반화된 역행렬임.)
귀무가설 하[5] ξ_H가 점근적으로 r^* 자유도[6]를 지닌 χ^2로 분포됨을 보일 수 있고, 분산은 귀무가설 하에서 평가된다. Rank r^*는 항상 제약 조건 r의 수보다 작거나 같으며, $r^*=r$ 인 특수한 경우 통계량 ξ_H는 ξ_w, ξ_{LR}, ξ_{LM}과 같다. Hausman 절차의 장점은 그것이 $\tilde{\beta}$와 β^c 두 추정량의 경우로 쉽게 확정될 수 있다는 것이고, 전자는 귀무가설 하 일치추정량이며, 후자는 귀무가설 하에서 일치추정량, 효율적 추정량이다. 검증은 $\epsilon_H < \chi^2{}_{95\%}(r)$이면 귀무가설

5) 즉, ($[\widehat{Var}(\beta)-\widehat{Var}(\beta^c)]^{-1}$는 $\widehat{Var}(\beta)-\widehat{Var}(\beta^c) = Var(\beta-\beta^c)$의 일반화된 역행렬로 수렴할 때

6) $[rank[var(\beta)-var(\beta^c)]]$

을 채택하고, $\epsilon_H > \chi^2{}_{95\%}(r)$이면 귀무가설 을 기각한다.

(2.21) $\xi_H = (\tilde{\beta}-\beta^c)'[\widehat{Var}(\tilde{\beta})-\widehat{Var}(\beta^c)]^{-1}(\tilde{\beta}-\beta^c)$

($r^* = rank[\widehat{Var}(\tilde{\beta})-\widehat{Var}(\beta^c)] = rank[Var(\tilde{\beta}-\beta^c)]$임, $\tilde{\beta}$는 제약되지 않은 식의 값) 이것이 일치추정량이 되는지를 검증하는데 귀무가설이 틀리다면 반드시 $(\tilde{\beta}-\beta^c)$가 0이 아니다.

2.1.4. 적합도 측정 (Measuring Goodness of Fit)

이진변수모형에서는 적합도를 측정하는 척도로 더 이상 R^2가 적절하지 않으며, 대안적인 적합도 측정 척도는 다음과 같다.

(1) Fraction Correctly Predicted

Y_i=1 이고 $\hat{P}_i = \Phi(\hat{\beta}_0 + \hat{\beta}_1 X_i) > 0.5$ 혹은 Y_i=0 이고 $\hat{P}_i < 0.5$ 인 경우 Y_i는 “correttly predicted"되었다고 한다. 그 반대의 경우는 Y_i가 “incorrectly predicted"되었다고 한다. 여기서 Fraction correctly predicted 는 correctly predicred 개수를 N으로 나누어 정의한다. 이 척도는 이해가 쉽다는 장점이 있지만, 예측의 질을 반영하지 못한다는 단점이 있다. 예를 들어 만일 Y_i=1 이라면 $\hat{p}_i$=0.51 이든 $\hat{p}_i$=0.9이든 둘 다 제대로 예측되었다고 간주하여 “질적” 측면은 반영하지 못한다.

(2) Likelihood Ratio Index

(2.22) $l = \sum_{i=1}^{n} [y_i \log \hat{F}_i + (1-y_i)\log(1-\widehat{F_i})]$

로그우도함수에서 $\hat{F}_i$는 $F(x_i'\hat{\beta}_{ML})$를 의미하고, 모수의 수가 서로 다른 모형들을 비교하는데 적절하다. McFadden[1974]은 R^2를 만들기 위한 l 의 표준화(normalization)를 제시하였는데, 즉, 로그우도함수의 극대값 lnL이 보고되고, 상수항을 제외한 모형의 모든 계수가 0이라는 로그우도함수가 보고되면, 전통적인 R^2와 유사한 likelihood ratio index를 구할 수 있다는 것이다. 이는 0과 1사이에 경계되며, 모형이 확장되는 경우 증가하는 값을 가진다.

(2.23) McFadden's $R^2 = 1 - \frac{l(\hat{\beta}_{ML})}{l_0}$ => LRI = $1 - \frac{\ln L}{\ln L_0}$

(l_0는 상수항을 제외한 모든 계수가 0이라는 제약식 l 의 극대값)

(3) pseudo- R^2

이 방법은 우도함수를 사용하여 모형의 적합도를 측정하는데, 왜냐하면 MLE는 우도함수를 극대화시키기 때문에 프로빗이나 로짓 회귀에 다른 설명변수를 추가하는 것은 마치 OLS에 의한 선형회기 시 하나의 설명변수 추가가 잔차제곱합을 줄여주는 것처럼, 최대우도값을 증가시켜준다. 이는 모든 설명변수를 포함시킨 극대화된 우도함수값과 아무것도 없는 우도함수 값과의 비교에 의해 추정될 수 있다. 따라서 pseudo- R^2는 다음과 같다.

(2.24) pseudo- R^2 = $1 - \frac{\ln(f_{\max\ probit})}{\ln(f_{\max\ bernoulli})}$

($f_{\max\ probit}$은 $x's$를 포함하는 프로빗 우도함수를 극대화 시킨 값이며, $f_{\max\ Bernoulli}$는 $x's$를 제외하는 프로빗모형으로 Bernoulli likelihood 를 극대

화 시킨 값을 의미함).

2.2 패널 데이타

2.2.1 패널자료의 개념[7)]

경제분석 및 정책결정에 가장 많이 사용하는 데이터로 시계열 자료(time-series 데이타)와 횡단면자료(cross-section)의 두 가지가 있다. 시계열 자료는 어떤 변수 값이 시간의 흐름에 따라 어떻게 변화하는지를 보여주는 자료로 거시경제의 분석에 주로 사용되며, 시행된 정책의 효과를 측정하거나 추세자료를 바탕으로 미래의 변화를 예측하는데 유용하게 활용할 수 있다. 횡단면 자료는 어떤 일정한 시점에서 각각의 관찰 표본들이 어떤 값을 지니는가를 보여주는 자료로 미시적 정책분석에 주로 사용되며, 일정 시점의 특정 사건의 상태와 영향을 세부적으로 관찰 할 수 있는 자료이다.

최근 선진국에서는 시계열 자료의 성격과 횡단면자료의 성격을 동시에 지니고 있는 패널 자료가 많은 관심을 불러일으키고 있다. 패널 자료는 1) 동일한 개인 및 가구들을 표본으로 하며, 2) 이들을 지속적으로 추적조사하고, 3) 일정기간마다 4) 동일한 질문을 하여 얻어지는 자료를 말한다. 이처럼 통상의 시계열 자료나 횡단면 자료와 구별하는 차원에서 패널자료는 흔히 종단면(longitudinal)라고 하며, 이 두 용어는 동의어로 흔히 사용된다. 즉, 패널자료는 특정 현상에 대해 입체적이고 일관성 있는 미시적 · 동태적 정보라는 점에서 통상의 자료와 구별되고 있다.

---

7) “빈곤 및 공공부조 패널 데이터 구축 방안“, 한국보건사회 연구원, 2004.

이러한 패널조사에서 가장 중요한 사항은 모든 조사와 마찬가지로 패널로 선택된 응답자 집단이 연구하고자 하는 모집단을 대표할 수 있어야 한다는 것이다. 패널조사는 본래 시장조사에서 소비자의 소비행동의 변화 과정을 분석하기 위하여 고안되었다. 최근에는 여론의 형성과정과 변화과정의 연구, 직업이동의 궤적을 밝혀내기 위한 연구(노동패널), 빈곤과 탈빈곤 과정에 대한 연구(빈곤 패널) 및 퇴직 · 고령화 과정 연구 등에도 널리 활용되고 있다.

2.2.2 패널 데이타의 구성

횡단면 데이타는 k+1(k개의 설명변수와 상수항)개 변수와 N개 관측치로 구성되며, 각 행은 개체(관측치), 각 열은 변수로 $N\times(k+1)$이 된다.

(2.25)
$$\begin{pmatrix} y_1 & x_{10} & x_{11} & x_{12} & \cdots & x_{1k} \\ y_2 & x_{20} & x_{21} & x_{22} & \cdots & x_{2k} \\ \\ . \\ . \\ . \\ y_n & x_{n0} & x_{n1} & x_{n2} & \cdots & x_{nk} \end{pmatrix}$$

반면, 패널 데이타는 각 개체에 대해 여러 개의 관측치를 제공한다.

(2.26)
$$\begin{pmatrix} y_{11} & x_{111} & x_{121} & \cdots & x_{1k1} \\ y_{12} & x_{112} & x_{122} & \cdots & x_{1k2} \\ \\ y_{1T} & x_{11T} & x_{12T} & \cdots & x_{1kT} \\ y_{21} & x_{211} & x_{221} & \cdots & x_{2k1} \\ . \\ . \\ . \\ y_{nT} & x_{n1T} & x_{n2T} & \cdots & x_{nkT} \end{pmatrix}$$

보다 쉬운 설명을 위해 2기 동안, 각 기간에 n개의 표본이 수집된다고 가정하고, 1987년 미국 46개 도시에 대한 데이타를 이용하여, 1000명당 범죄율과 실업율이 다음과 같다고 하자.

(2.27) $\widehat{crime\ rate} = 128.38 - 4.16\,unemp$

이 모형은 실업과 범죄율을 설명하기에 너무 단순하며, 실업과 상관관계가 있는 누락변수는 추정 계수에 편향(bias)을 발생시킬 것이다. 보다 나은 모형은 범죄자의 연령 분포, 이력, 성별 분포, 법률 집행, 교육수준, 범죄에 대한 개별특성 등을 통제하는 것이지만 개별 변수, 이력 변수와 같은 몇몇 변수는 통제하기 힘들고, 다른 변수들은 데이타가 부족할 것이다. 이러한 관점에서 패널 테크닉은 해답을 제시한다.

한 개체에 대한 모형으로 (2.28) $y_{it}=\beta_1+x_{it}\beta_2+u_{it}$을 고려하고, u_{it}을 시간에 따라 변하지만 개체에는 불변인 λ_t , 개체 간에는 변하지만 시간에 따라 불변인 α_i , 개체와 시간에 따라 변하는 v_{it}의 세 부분으로 나누어 고려하자. 이제 모형은 (2.29) $y_{it} = \beta_1 + x_{it}\beta_2 + \lambda_t + \alpha_i + v_{it}$이 되며, 이는 고정효과 모형, error components 모형, unobserved effect 모형으로 다양하게 불린다. 여기서 $E[\alpha_i|x_{it}]=0$이면 모수 추정치는 일치성을 만족하지만, $E[\alpha_i|x_{it}]\neq 0$이면 일치성을 만족하지 못할 것이다. 이는 헤테로 편향(heterogeneity bias)라 불리며 누락변수의 한 형태이다. 선형 모형에서는 1차 차분을 통해 (first-differencing) α_i를 제거할 수 있지만 로짓, 프로빗 모형을 포함하는 비선형 모형에서는 그렇지 않다. 본 논문에서는 비선형 패널데이타 모형 중에서도 이진변수모형에 대한 고정효과모형과 임의효과모형을 논의하고자 한다.

2.2.3 패널 데이타의 특성

(1) 패널 데이타의 장점

패널 자료의 장점은 첫째, 동태적(dynamic) 성격의 양질의 정보를 제공해 준다는 것이다. 특정 시점에서의 관심이 있는 현상의 양태를 보여주는 횡단면 자료로는 파악이 불가능한 특정 시점간의 전환과정을 보여준다. 둘째, 특정 사건에 대한 심층적인 연구를 가능하게 한다. 패널자료는 특정기간 효과, 특정연령 효과, 그리고 동시태생 효과 등에 대한 분석을 가능하게 한다. 셋째, 정책의 합리적인 수립과 평가를 가능하게 한다. 패널조사는 동일한 개인에 대해 주기적이고 장기간에 걸친 추적 조사형식으로 이루어짐에 따라 과거의 사회적 환경이 현재 어떠한 영향을 미치고 있는 가를 심층적으로 전망하고 분석할 수 있다.

(2) 패널 데이타의 한계

패널자료의 단점으로는 첫째, 조사에 비용이 많이들며 둘째, 표본 이탈로 인한 대표성의 상실과 그에 따른 데이터의 질에 대한 것이며 셋째, 자료의 활용이 다소 복잡할 수 있다는 것이다. 먼저 패널자료는 자료의 목적과 활용 의도에 따라 조사기간을 정하지만 동일인을 일정 기간동안 추적조사 하여야 하기 때문에 조사비용이 동일한 규모의 횡단면 자료 조사비용보다 많이 든다. 패널조사가 다른 조사보다 노력과 비용이 많이 드는 주된 이유는 표본 추출, 설문지의 구성 등과 같은 조사 설계가 결정되면 패널조사가 진행되는 한 이를 변경하기 어렵기 때문이다.

3. 정태적 패널 이진변수모형

패널 데이타를 위한 이진변수 모형은 오랫동안 주요 관심사였다[8] . 이제

(3.1) $y_{it}=0$ if $y_{it}^{*}\le 0$ 만을 관찰하는 경우, 잠재 변수 모형 (3.2)
$y_{it}=1$ $otherwise$

$y_{it}^{*}=x_{it}'\beta+\alpha_i+v_{it}$의 틀에서 이진변수모형을 고려해 보고자 한다. v_{it}의 분포에 관한 가정[9]이 F를 결정할 때 $\Pr ob(y_{it}=1|x_{it},\alpha_i)=F(x_{it}'\beta+\alpha_i)$을 모형화한다. T=2 인 균형패널을 고려하면, 1기와 2기 개체 i 에 대한 관측은 다음과 같이 나타내어진다.

(3.3) $y_{i1}^{*}=x_{i1}'\beta+\alpha_i+v_{i1}$

$y_{i2}^{*}=x_{i2}'\beta+\alpha_i+v_{i2}$

이 프로빗모형은 선형의 경우와 같이 고정효과 접근법, 즉, 고정효과를 제거하기 위한 평균추세제거 혹은 차분방법으로 추정될 수 있을까? 이 두 방정식을 차분한 값 $\triangle y_{i2}^{*}=\triangle x_{i2}'\beta+\triangle v_{i2}$ 는 고정효과를 제거할 수 없을 것이다. 또한, (3.4) $\Pr ob(y_{i2}=1|x_{i2},\alpha_i)-\Pr ob(y_{i1}=1|x_{i1},\alpha_i)=F(x_{i2}'\beta+\alpha_i)-F(x_{i1}'\beta+\alpha_i)$인 경우도 고정효과를 제거하지 못할 것이며. 이는 어떠한 함수형태의 비선형 모형에라도 적용될 것이다. 그럼, α_i를 n개의 추정해야할 모수로서 보는 것은 어떠할까? i 번째 개체의 관측치를 다음과 같이 고려해보면 4개 가능성이 존재한다.

(3.5) $y_{i1}^{*}=x_{i1}'\beta+\alpha_i+v_{i1}$

$y_{i2}^{*}=x_{i2}'\beta+\alpha_i+v_{i2}$

8) Arellano and Honore(1999)

9) 표준 정규분포: pribit모형, 로지스틱 분포: 로짓모형

Time	1	2
y_{it}	0	0
y_{it}	0	1
y_{it}	1	0
y_{it}	1	1

그러므로 i 번째 관측치의 우도함수값은 다음과 같으며,

$$(3.6)\quad f(y_i \mid x_i{}', \alpha_i) = (1-F(x_{i1}{}'\beta+\alpha_i))(1-F(x_{i2}{}'\beta+\alpha_i))^{(1-y_{i1})(1-y_{i2})}$$
$$+(1-F(x_{i1}{}'\beta+\alpha_i))F(x_{i2}{}'\beta+\alpha_i)^{(1-y_{i1})(y_{i2})}$$
$$+F(x_{i1}{}'\beta+\alpha_i))(1-F(x_{i2}{}'\beta+\alpha_i))^{(y_{i1})(1-y_{i2})}$$
$$+F(x_{i1}{}'\beta+\alpha_i)(F(x_{i2}{}'\beta+\alpha_i))^{(y_{i1})(y_{i2})}$$

일반적인 우도함수는(3.7) $L = \prod_i f(y_i | x_i{}', \alpha_i)$ 와 (3.8) $\ln L = \sum_{i=1}^{n}\sum_{t=1}^{T_i} \ln(f(y_{it} | x_{it}{}', \alpha_i))$

로 나타낼 수 있다. 이제 $\sum_{i=1}^{n} T_i$개의 관측치와 $n+k$개의 모수를 가질 것이며, 모수의 수가 관측치 수 보다 더 많다. 따라서 통계적인 문제가 발생한다. 그렇다면 선형 확률 모형에 없던 incidental parameter problem이 여기서는 왜 발생했는가? 선형 회귀모형의 추정이 여기서와 같이 원 데이터가 아닌 그룹 평균값으로 부터의 편차에 기반을 두었다는 점을 상기해 보자. $f(y_{it} | x_i)$는 α_i의 함수이지만, $f(y_{it} | x_i, \overline{y_i})$는 α_i의 함수가 아니라는 사실을 β 추정 시에 사용했다. $\overline{y_i}$는 α_i의 최소 충분통계량이며, 이는 앞으로 검토해 볼 로짓 모형에서는 활용 가능하지만, 프로빗 모형에는 그렇지 않다.

대부분의 논문은 오차항의 분포가 공변인(covariates)에 독립이라고 가정하는 임의효과모형을 사용한다. 더구나, 대부분은 분포 함수를 모수적으로(parametrically)정의하며, 비모수적 방법[10]은 거의 없다. 개체효과(individual

10) Chen,Heckman and Vytlacil (1999).

effect)의 분포함수가 공변인(covariates)에 의존할 수 있다는 고정효과 모형은 서로 다른 4가지 접근법으로 구분할 수 있다 : maximum score(Manski,1987), maximum rank correlation(Lee,1999, Abrevaya,2000), "pseudo-regression (Honore and Lewbel,2000), 그리고 조건부로짓 같은 likelihood method .

Manski(1987)의 maximum score estimation은 $\sqrt{n}$ 일치성을 지니며 추정치의 점근적 분포가 정규분포를 따르지 않으나 Horowitz(1992)의 smooth maximum score는 $\sqrt{n}$ 일치성을 지니며 점근적으로 정규분포를 따른다. 조건부 우도함수 접근법은 직교하는 모수(Lancaster,2001)를 이용하는 Cox and Reid(1987)방법과 Bayesian method에 기반을 둔다. 조건부 우도함수 접근의 단순성으로, 임의효과접근을 고려하기 이전에 그것을 먼저 고려하고자 한다.

(3.9) $y_{it}{}^{*} = x_{it}\beta + v_{it} + \alpha_i \quad (i = 1,2,...,n\ ,\ \ t = 1,2,..., T_i)$

$y_{it} = 1 \quad \text{if} \quad y_{it}{}^{*} > 0\ ,\ \text{and}\ 0\ \ \textit{otherwise}$

(α_i는 관측되지 않는 개체 특성 이질성임)

α_i와 x_{it}간의 관계로서 "고정효과"와 "임의효과"를 구분한다. α_i가 x_{it}와 상관관계가 없다는 가정($f(\alpha_i | x_{it})$가 x_{it}에 의존하지 않는다)은 임의효과 모형을 가능하게 하는데, 이것은 heterogeneity의 분포에 제약(즉, α_i와 x_{it}의 상관관계가 없음)을 부과함을 숙지하자. 분포가 제약되지 않는다면(즉, α_i와 x_{it}의 상관관계 있음), 이를 고정효과 모형이라 한다. 이들의 구분은 개체효과 자체의 본질적인 특성과 관련되지 않는다. 임의효과 모형의 추정은 heterogeneity에 대한 강한 가정, 즉, 개체 - 특성 효과와 설명변수가 직교

(orthogonal)함을 가정하는 반면, 고정효과 모형은 MLE를 불일치하게 하는 incidental parameter problem에 직면한다는 어려움이 있다.

패널 데이타 이진변수모형의 일반적인 형태는 다음과 같다.

(3.10) $U_{it} = \alpha + \beta' x_{it} + \epsilon_{it} +$ 개체 i의 특성효과

더미 변수를 이용하는 고정효과 모형은 다음과 같으며,

(3.11) $U_{it} = \alpha_i + \beta' x_{it} + \epsilon_{it}$, $y_{it} = 1[U_{it} > 0]$.

누락된 heterogeneity를 이용하는 임의효과는

(3.12) $U_{it} = \alpha + \beta' x_{it} + (\epsilon_{it} + v_i)$, $y_{it} = 1[U_{it} > 0]$가 된다.

앞에서 언급한 바와 같이 더 이상 선형이 아니므로 차분에 의한 개체효과특성을 제거할 수 없으며, 직접적인 추정법이 필요하다. 이제 고정효과가 존재하는 이진변수모형의 추정과 문제점 및 해결 방안에 대해 살펴보고자 한다.

3.1 고정효과

개체-특성효과 α_i를 고정된 것으로 가정하면, (3.13) $P(y_{it} = 1 | x_{it}, \alpha_i) = F(\beta' x_{it} + \alpha_i)$ 인 모형을 추정하는데 있어 β와 α_i 모두 미지의 모수이다. T가 무한대로 가면, MLE는 일치성을 만족하지만 패널 데이타에서 T는 보통 작다. 고정효과의 문제점은 통계적인 문제이며 T가 고정되는 경우 뿐만 아니라 너무 작은 경우, 상수항의 추정치는 일치성을 만족하지 않는다. 즉, 추정량 β가 α의 함수이고, β의 MLE가 더 이상 일치성을 만족하지 않음을 의미하는데 이를 incidental parameter problem이라 한다[11]. 뿐만 아니라 소규모편향(small-T)도 존재하게 되는데 이 편향이 얼마나 심각한지에 관한 문헌

11) Neyman and soctt(1948), Lancaster (2000).

은 Hsiao (1986), Heckman Macurdy(1980)를 참조할 수 있다.

1차 차분과 같은 선형화(linear transformation)를 통해 개체효과 α_i가 제거될 수 있었던 선형 회귀분석과는 달리, 비선형 모형에서는 incidental parameter를 제거하는 간단한 형태(simple transform)가 존재하지 않는다. binary-choice 모형에서 α_i와 β의 MLE가 서로 독립적이지 않고 T가 고정된 경우, α_i의 불일치성은 β의 MLE로 전이된다. 때문에 N이 무한대로 갈지라도 β의 MLE는 일치성을 지니지 못한다. Neyman and scott(1948)은 incidental parameter α_i가 존재하는 경우, 일치 추정량 β 를 제시한다. 그들의 아이디어는 β가 맞는 값일 경우 incidental parameter α_i에 독립이며 N이 무한대로 감에 따라 $\psi_N(y_1,\ldots,y_N|\beta)$가 0으로 수렴하는 특성을 지닌 K함수 $\psi_N(y_1,\ldots,y_N|\beta)$ $j=1,\ldots,k$를 찾는 것이다. 그리고 적절한 regularity 조건 하에서 $\psi_N(y_1,\ldots,y_N|\hat{\beta})=0$를 풂으로써 유도된 추정량 $\hat{\beta}$은 일치추정량이 된다.

예를 들어 고정효과 로짓모형의 $\hat{\beta}^*=\left(\frac{1}{2}\right)\hat{\beta}$가 그러한 추정치이다. 선형확률모형(LPM)의 경우, 시간에 대해 1차 차분을 취하거나 개체 평균과 관련한 차분은 개체-특성효과를 제거하여, 차분된 방정식의 LS 회귀분석의 경우 N이 무한대로 갈 때, 모수에 대한 일치추정량을 제공한다. 그러나 일반적인 비선형 모형에서 ψ의 간단한 함수를 항상 찾기 쉬운 것은 아니다. 그러나 incidental parameter α_i의 최소 충분 통계량 τ_i가 존재하고, 그것이 β에 의존하지 않는다면 조건부 분포 (3.14) $f^*(y_i|\beta,\tau_i)=\dfrac{f(y_i|\beta,\alpha_i)}{g(\tau_i|\beta,\alpha_i)}$ 는 더 이상 α_i에 의존하지 않는다는 것이다. Anderson(1970,1973)은 $\tau_1,\ldots,\tau_N$이 주어진 경우 $y_1,\ldots,y_N$의 조건부 분포를 극대화 하면[12] 1계 조건 $\psi_{Nj}(y_1,\ldots,y_N|\hat{\beta},\tau_1,\tau_2,\ldots,\tau_N)=0$을 얻을 수 있음을 보였으며, 이는 모수 β의 일치 추정량을 제시한다.

그러면 이제 일반적인 MLE의 불일치성을 검토한 후, 일치성을 만족하는 대안으로서 조건부 우도함수 방법을 살펴보겠다.

3.1.1. 고정효과 로짓 모형

(1) 최우추정량

경제학 응용분야에서는 회귀변수(regressor : 예를 들면 의사결정변수의 경우)와 개체의 이질성(individual heterogeneity : 예를 들면 기호)간의 상관관계가 존재한다. 이러한 관계로 고정효과접근을 채택하여 $\{\alpha_i\}_{i=1,2,\ldots,n}$를 추정할 모수로서 취급할 수 있다. 모수를 추정하기 위해 $\{y_{it}, x_{it}\}$ 를 사용할 수 있으며, ϵ 이 회귀계수(regressor)와 개체효과(individual effect)에 따라 알려진 분포(normal, logistic)를 따른다고 가정하면, 이 모형의 추정법으로 최우추정법(maximum likelihood)을 행할 수 있다.

먼저, 로짓 모형 (3.15) $\mathrm{Prob}(y_{it}=1|x_{it},\alpha_i)=\dfrac{\exp(\beta' x_{it}+\alpha_i)}{1+\exp(\beta' x_{it}+\alpha_i)}$을 고려함으로써 MLE β 의 불일치성을 증명할 것이다. log 우도함수는 다음과 같으며,

(3.16) $$\log L = -\sum_{i=1}^{N}\sum_{t=1}^{T}\log[1+\exp(\beta' x_{it}+\alpha_i)]+\sum_{i=1}^{N}\sum_{t=1}^{T}y_{it}(\beta' x_{it}+\alpha_i)$$

설명의 편의를 위해 T=2이고, x_{i1}=0 그리고 x_{i2}=1을 갖는 설명변수를 고려하자. 이 경우에 1차미분 방정식은,

$$\frac{\partial \log L}{\partial \beta}=\sum_{i=1}^{N}\sum_{T=1}^{2}[-\frac{e^{\beta x_{it}+\alpha_i}}{1+e^{\beta x_{it}+\alpha_i}}+y_{it}]x_{it}$$ 이다.

---

12) $\prod_{i=1}^{N} f^*(y_i|\beta,\tau_i)$ 를 의미함.

(3.17) $$= \sum_{i=1}^{N} [- \frac{e^{\beta x + \alpha_i}}{1 + e^{\beta x_{it} + \alpha_i}} + y_{i2}] = 0$$

(3.18) $$\frac{\partial \log L}{\partial \alpha_i} = \sum_{t=1}^{2} [- \frac{e^{\beta x_{it} + \alpha_i}}{1 + e^{\beta x_{it} + \alpha_i}} + y_{it}] = 0$$

(3.17)을 풀면 다음과 같다.

(3.19) $$\hat{\alpha}_i = \begin{array}{ll} \infty & \text{if} \quad y_{i1} + y_{i2} = 2 \\ -\infty & \text{if} \quad y_{i1} + y_{i2} = 0 \\ -\frac{\beta}{2} & \text{if} \quad y_{i1} + y_{i2} = 1 \end{array}$$

(3.19)를 (3.17)에 넣고, $y_{i1} + y_{i2} = 1$의 개체수를 n_1, $y_{i1} + y_{i2} = 2$의 개체수를 n_2라 놓으면 다음을 얻을 수 있다.

(3.20) $$\sum_{i=1}^{N} \frac{e^{\beta x_{it} + \alpha_i}}{1 + e^{\beta x_{it} + \alpha_i}} = n_1 \frac{e^{\frac{\beta}{2}}}{1 + e^{\frac{\beta}{2}}} + n_2 = \sum_{i=2}^{N} y_{i2}$$

(3.21) $$\hat{\beta} = 2\log(\sum_{i=1}^{N} y_{i2} - n_2) - \log(n_1 + n_2 - \sum_{i=1}^{N} y_{i2})$$

대수의 법칙에 의해 (Rao (1973, chapter2))

(3.22) $$\underset{n \to \infty}{plim} \frac{1}{N} (\sum_{i=1}^{N} y_{i2} - n_2) = \frac{1}{N} \sum_{i=1}^{N} prob(y_{i1} = 0, y_{i2} = 1 | \beta, \alpha_i)$$

$$= \frac{1}{N} \sum_{i=1}^{N} \frac{e^{\beta + \alpha_i}}{(1 + e^{\alpha_i})(1 + e^{\beta + \alpha_i})}$$

$$(3.23) \qquad \underset{n\to\infty}{plim}\frac{1}{N}\left(\sum_{i=1}^{N}n_1+n_2-\sum_{i=1}^{N}y_{i2}\right)=\frac{1}{N}\sum_{i=1}^{N}prob(y_{i1}=1,y_{i2}=0|\beta,\alpha_i)$$

$$=\frac{1}{N}\sum_{i=1}^{N}\frac{e^{\alpha_i}}{(1+e^{\alpha_i})(1+e^{\beta+\alpha_i})}$$

$\widehat{\alpha_i}=-\frac{\beta}{2}$를 (3.21)과 (3.22)에 대입하면 $\underset{n\to\infty}{plim}\,\hat{\beta}=2\beta$ 이고 일치성을 만족하지 않는다. 즉, 고정효과 로짓 모형의 MLE β 는 일치추정량이 아니다.

(2) 조건부 최우추정량 (conditional maximum likelihood estimator)

Anderson(1973)과 Chamberlain(1980)은 로짓 모형을 이용하여 일치추정량을 위한 조건부 우도 접근법을 제시하는데, 먼저 incidental parameter α_i의 충분통계량에 의존하는 우도함수를 고려한다. 로짓 모형에서 α_i의 충분통계량은 $\sum_{t=1}^{T}y_{it}$이며, 조건부 우도함수가 통상조건을 만족하면 β의 조건부 최우추정량은 일치성을 만족한다. T=2인 경우로 설명하면 로짓모형의 경우,

$$(3.24) \qquad \mathrm{Prob}(y_{it}=1)=\frac{\exp(\beta' x_{it}+\alpha_i)}{1+\exp(\beta' x_{it}+\alpha_i)}\ ,\ \mathrm{Prob}(y_{it}=0)=\frac{1}{1+\exp(\beta' x_{it}+\alpha_i)}$$

$$(3.25) \qquad \mathrm{Prob}(Y_{i1}=y_{i1},Y_{i2}=y_{i2})=\frac{\exp(\sum_{t=1}^{T_i}y_{it}\,\beta' x_{it})}{\Sigma_{\Sigma_t d_t=S_i}\exp(\sum_{t=1}^{T_i}d_{it}\beta' x_{it})}$$

(3.26) $\quad \Pr(0,1) = \dfrac{1}{1+\exp(\beta' x_{i1}+\alpha_i)} \cdot \dfrac{\exp(\beta' x_{i2}+\alpha_i)}{1+\exp(\beta' x_{i2}+\alpha_i)}$

(3.27) $\quad \Pr(1,0) = \dfrac{\exp(\beta' x_{i1}+\alpha_i)}{1+\exp(\beta' x_{i1}+\alpha_i)} \cdot \dfrac{1}{1+\exp(\beta' x_{i2}+\alpha_i)}$

(1.0), (0,1)은 상호 배타적이므로

(3.27) $\quad \Pr[(1,0)|(1,0) \text{ or } (0,1)] = \dfrac{prob(1,0)}{prob(1,0)+prob(0,1)} = \dfrac{\exp[\beta'(x_{i1}-x_{i2})]}{D}$

(3.28) $\quad \Pr[(0,1)|(1,0) \text{ or } (0,1)] = \dfrac{1}{D}$

($D = 1+\exp[\beta'(x_{i1}-x_{i2})]$임.)

이와 같이, α_i가 제거되면 일반적인 로짓 모형이 되고, 개체 특성효과(α_i)의 분포를 정의하지 않아도 모수 β를 추정할 수 있다. 그러나 조건부 최우추정량의 문제는 모수 β에 의존하는 충분통계량을 항상 찾을 수 있는 것은 아니라는 것이다. 또한 충분통계량에 의존하는 접근법의 한계는 (x_i, α_i)에 대한 y_i의 모수적 모형을 필요로 한다는 것이다.13) 이 모형에 이질성이 존재하는지 아닌지를 검증하는데 흥미로운 사실들이 있다. 동질성($\alpha_i=\alpha$)이 존재하는 경우, 특별한 문제 없이 로짓모형과 같은 보통의 모형으로 추정될 수 있다. 그러나, 우도비 검정(LR)를 사용하는 가설을 검증할 수는 없다. 왜냐하면 두 개의 우도함수는 비교할 수 없기 때문이다14) 그러나, Hausman(1978)검정은 여기에 사용하기 적합하다.15)

---

13) 조건부 로짓 접근에서 $\sum_{t=1}^{2} y_{it} = 0$ 혹은 2인 그룹은 확률이 1이 되므로 우도함수에 반영하지 않는다 ; $\Pr[(0,0)|(1,0) \text{ or } (0,1)] = 1$, $\Pr[(1,1)|(1,0) \text{ or } (0,1)] = 1$ 임.).

14) 조건부 우도함수는 제약된 데이타 set에 기초하기 때문임.

15) 앞의 [2.1.3] 부분을 참고하라.

동질성의 귀무가설 하에서, chamberlain의 조건부 최우추정량(CMLE)과 일반적인 최우추정량(MLE) 모두 일치성을 만족하지만 효율적이지 않다. 대립가설하에 최우추정량은 불일치한 반면, chamberlain의 추정량은 일치추정량이고 효율적이다. Chamberlain(1980)은 조건부 최우추정법의 경우 로그선형모형 뿐 만 아니라 다항 로짓모형까지도 적용할 수 있음을 보였다.

3.1.2. 고정효과 프로빗 모형

small-T 고정효과 프로빗모형에서는 고정효과가 없어지지 않으므로 조건부 MLE를 사용할 경우, 로짓모형에서와 같이 계산이 단순하지 않다. 이는 N개의 고정효과가 추정과정의 일부로서 추정되어야 한다는 것을 의미하고, 또한 고정효과의 추정량이 소규모 표본(small T)에 대해 일치추정량이 아니므로, 고정효과 프로빗모형은 β의 일치추정량을 제시하지 못한다. 그러므로 패널 데이타 이진변수 모형에 고정효과를 적용하는 경우에는 로짓모형을 선택하여야 한다. 그러나 임의효과의 경우에는 프로빗모형이 로짓모형보다 계산상 다루기 용이하다. 이제 임의효과 프로빗 모형을 살펴보겠다.

3.2. 임의효과 프로빗 모형

3.2.1. 회귀변수(x_{it})와 개체특성(α_i) 간의 상관관계가 없는 경우의 MLE추정.

먼저, 회귀변수 x와 개체특성변수 α_i와의 상관관계가 없는 경우를 살펴보기 위해 임의효과모형은 $\epsilon_{it} = v_{it} + \alpha_i$로 정의하며, 다음과 같다고 하자 ($v_{it}$

와 α_i는 다음과 같이 독립 확률변수이며 X는 표본에 있는 모든 외생변수임).

(3.30)

$$E[v_{it}|X]=0;\ Cov[v_{it},v_{is}|X]=Var[v_{it}|X]=1 \quad \text{if} \ \ i=j \ \text{ and } \ t=s;\ 0 \ \ otherwise$$
$$E[\alpha_i|X]=0;\ Cov[\alpha_i,\alpha_j|X]=Var[\alpha_i|X]=\sigma_\alpha^2 \quad \text{if} \ \ i=j \ ;\ 0 \ \ otherwise$$
$$Cov[v_{it},\alpha_j|X]=0 \ \ (all \ \ i,t,j)$$

그러면

(3.31)
$$E[\epsilon_{it}|X]=0 \qquad\qquad (\sigma_\alpha^2=\rho/(1-\rho)\text{임})$$
$$Var[\epsilon_{it}|X]=\sigma_v^2+\sigma_\alpha^2=1+\sigma_\alpha^2$$
$$Corr[\epsilon_{it},\epsilon_{is}|X]=\rho=\frac{\sigma_\alpha^2}{1+\sigma_\alpha^2}$$

횡단면 데이타의 경우 관측치와 관련된 확률은 아래와 같다.

(3.32)
$$P(y_i|x_i)=\int_{L_i}^{U_i} f(\epsilon_i)d\epsilon_i, \qquad (L_i,\ U_i)=(-\infty,\ -x_i'\beta) \ \text{if} \ y_i=0$$
$$(-x_i'\beta,\ +\infty) \ \text{if} \ y_i=1$$

이것은 정규분포에 대해서는 $\Phi[(2y_i-1)x_i'\beta]$, 로짓 모형에 대해서는 $\Lambda[(2y_i-1)x_i'\beta]$로 단순화 할 수 있다.

제약되지 않은 분산행렬을 지닌 일반적인 경우,

(3.33) $L_i=P(y_{i1},\dots,y_{iT_i}|X)=\int_{L_{iT_i}}^{U_{iT_i}}\cdots\int_{L_{iT_i}}^{U_{iT_i}} f(\epsilon_{i1},\epsilon_{i2},\dots,\epsilon_{iT_i})d\epsilon_{i1},d\epsilon_{i2},\dots,d\epsilon_{iT_i}$ 와 같다.

결합 분포의 적분은 대부분의 경우 활용할 수 없으나 임의효과 모형의 특성은 단순화를 가능하게 한다. $(\epsilon_{i1},\dots,\epsilon_{iT_i},\alpha_i)$의 결합확률로부터 α_i를 적분함으

로써 v_{it}의 결합분포를 얻을 수 있다.

(3.34) $f(\epsilon_{i1},\ldots,\epsilon_{iT_i},\alpha_i)=f(\epsilon_{i1},\ldots,\epsilon_{iT_i}|\alpha_i)f(\alpha_i)$

그리하여 (3.35) $f(\epsilon_{i1},\epsilon_{i2},\ldots,\epsilon_{iT_i})=\int_{-\infty}^{+\infty} f(\epsilon_{i1},\epsilon_{i2}\ldots,\epsilon_{iT_i}|\alpha_i)f(\alpha_i)d\alpha_i$ 가 된다. 이 형태의 장점은 α_i에 의존하는 ϵ_{it} 값들이 독립이라는 것이며, 그리하여 다음과 같이 나타낼 수 있다.

(3.36) $f(\epsilon_{i1},\epsilon_{i2},\ldots,\epsilon_{iT_i})=\int_{-\infty}^{+\infty}\prod_{t=1}^{T_i} f(\epsilon|\alpha_i)f(\alpha_i)d\alpha_i$

이 식을 (3.33)에 삽입하면 아래와 같은 식이 된다.

(3.37) $L_i=P(y_{i1},\ldots,y_{iT_i}|X)=\int_{L_{iT_l}}^{U_{iT_i}}\cdots\int_{L_{iT_1}}^{U_{iT_1}}\int_{-\infty}^{\infty}\prod_{t=1}^{T_i} f(\epsilon|\alpha_i)f(\alpha_i)d\alpha_i,d\epsilon_{i1},d\epsilon_{i2}\ldots,d\epsilon_{iT_i}$

사실 이 식이 훨씬 단순하게 보이지는 않을 지도 모른다. 위 식의 적분 범위는 독립이므로 순서를 바꾸면,

(3.38) $L_i=P(y_{i1},\ldots,y_{iT_i}|X)=\int_{-\infty}^{\infty}\left[\int_{L_{iT_l}}^{U_{iT_i}}\cdots\int_{L_{iT_1}}^{U_{iT_1}}\prod_{t=1}^{T_i} f(\epsilon|\alpha_i)d\epsilon_{i1},d\epsilon_{i2}\ldots,d\epsilon_{iT_i}\right]f(\alpha)d\alpha$

공통의 α_i에 의존하는 ϵ은 독립이므로, 대괄호 안에 있는 항들은 개체 확률들로 다음과 같이 쓸 수 있다.

(3.39) $L_i=P(y_{i1},\ldots,y_{iT_i}|X)=\int_{-\infty}^{\infty}\left[\prod_{t=1}^{T_i}\int_{L_{it}}^{U_{it}} f(\epsilon_{it}|\alpha_i)\right]f(\alpha_i)d\alpha_i$

이제 개체들의 분포를 고려해 보면 α_i에 의존하는 개체들의 분포를 이제 $x_{it}{}'\beta+\alpha_i$에서 계산할 수 있고, 임의효과 이진변수의 일반적인 모형을 생성한다. 모든 항들을 정리하면 다음과 같이 변형할 수 있다.

(3.40) $L_i=P(y_{i1},\ldots,y_{iT_i}|X)=\int_{-\infty}^{\infty}\left[\prod_{t=1}^{T_i}\mathrm{Prob}(Y=y_{it}\,|\,x_{it}{}'\beta+\alpha_i)\right]f(\alpha_i)d\alpha_i$

대괄호 안의 확률은 프로빗, 로짓, weibull과 같이 현재까지 고려한 모형들

이 될 것이며, 밖의 적분을 어떻게 할 것인가의 문제가 존재하게 된다. 우선, α_i가 정규 분포된다고 가정하는 Butler and Moffit's 방법이 간단하므로 먼저 고려하고 그리고 나서 다른 가능성들을 고려할 것이다. 프로빗모형의 경우 괄호 안의 개체 확률은 $\Phi[q_{it}(x_{it}'\beta+\alpha_i)]$가 될 것이며16), 로짓 모형의 경우 $\Phi[\bullet]$는 로지스틱 확률 $\Lambda[\bullet]$로 대체된다. 전체함수를 α_i , $g(\alpha_i)$의 함수로 고려하면 적분은 (3.41) $L_i=\int_{-\infty}^{\infty}\frac{1}{\sigma_\alpha\sqrt{2\pi}}e^{-\frac{\alpha_i^2}{2\sigma_\alpha^2}}g(\alpha_i)\alpha_i$가 되며.17) 변수를 조정하면, (3.42) $L_i=\frac{1}{\sqrt{\pi}}\int_{-\infty}^{\infty}e^{-\gamma_i^2}g(\theta\gamma_i)d\gamma_i$가 된다. 프로빗 모형에 적용하면 (3.43) $L_i=\frac{1}{\sqrt{\pi}}\int_{-\infty}^{\infty}e^{-\gamma_i^2}[\prod_{t=1}^{T_i}\Phi(q_{it}(x_{it}'\beta+\theta\gamma_i)]d\gamma_i$ 와 같아진다. 이 모든 과정은 오직 하나의 적분이 될 수 있으며, 안쪽의 적분은 표준 정규분포 혹은 로지스틱 혹은 extreme value 분포를 지니므로 구하기 쉽다. 이 함수는 계산을 위해 Gauss-Hermite quadrature로 수정가능하며 모든 과정을 정리하면 다음과 같은 로그 우도함수로 근사할 수 있다.18)

(3.44) $$\ln L_H=\sum_{i=1}^{n}\left\{\ln[\frac{1}{\sqrt{\pi}}\sum_{h=1}^{H}\sum_{t=1}^{T_i}w_h\Phi(q_{it}(x'\beta+\theta z_h))]\right\}$$

이 함수를 극대화해도 복잡한 문제는 여전히 남지만, 적분을 1차로 줄이는 변환(transformation)을 통해 해결이 가능하다. Butler and Moffit 방법은 large T_i가 있는 경우 조차도 일반적인 계산절차를 통해 효율적으로 추정할 수 있다는 장점이 있지만, α_i에 대한 정규성을 가정한다는 문제를 지니어

16) $\Phi[\bullet]$는 표준정규분포의 CDF이고, $q_{it}=2y_{it}-1$임.

17) $\gamma_i=\alpha_i/(\sigma_\alpha\sqrt{2})$로 두면 $\alpha_i=(\sigma_\alpha\sqrt{2})\gamma_i=\theta\gamma_i$이고 $du_i=\theta d\gamma_i$임.

18) H는 구적분(quadrature)를 위한 점(points)의 수이고 w_h와 z_h는 구적분(quadrature)를 위한 가중치와 결절점(nodes)임.

다른 분포들의 경우 적분과정을 근사시키는 방법 등에 어려움이 있다. 이에 대한 대안적인 방법은 maximum simulated likelihood(MSL)을 활용하는 방안이 있다.[19)]

위 식(3.40)에서 유도한 변형된 우도함수는 기댓값이며,

$$(3.45)\quad L_i = \int_{-\infty}^{\infty}\left[\prod_{t=1}^{T_i}\mathrm{Prob}(Y_{it} = y_{it}\,|x_{it}'\beta+\alpha_i)\right]f(\alpha_i)d\alpha_i$$
$$= E_{\alpha_i}\left[\prod_{t=1}^{T_i}\mathrm{Prob}(Y_{it} = y_{it}\,|x_{it}'\beta+\alpha_i)\right]$$

이 기댓값은 구적분(quadrature)보다는 simulation으로 근사될 수 있는데(θ는 α_i 분포의 scale parameter라고 두자), 우도 함수는 다음과 같다.[20)]

$$(3.46)\quad L_i = E_{\alpha_i}\left[\prod_{t=1}^{T_i}F(y_{it},\, x_{it}'\beta+\theta\alpha_i)\right] = E_{\alpha_i}[h(\alpha_i)]$$

기대값이 유한하면 대수법칙이 적용되고 이는 관측치 $\alpha_{i1},\cdots,\alpha_{iR}$의 표본이 $plim\frac{1}{R}\sum_{\gamma=1}^{R}h(\alpha_{ir}) = Eu[h(\alpha_i)]$가 됨을 의미한다. simulated log likelihood 함수는 (3.47) $\ln L_{simulated} = \sum_{i=1}^{n}\ln\left\{\frac{1}{R}\sum_{r=1}^{R}\left[\prod_{t=1}^{T_i}F[q(x_{it}'\beta+\sigma_\alpha\alpha_{ir})]\right]\right\}$가 되고, 이 함수는 β와 σ_α와 관련하여 극대화 되며, quadrature에서와 마찬가지로 프로빗, 로짓 모형 등에서 사용될 수 있으나 로짓 모형에 선호된다. 회귀계수 x_i와 개체효과 α_i와의 상관관계가 없다고 가정하는 경우, 임의효과 프로빗모형의 또 다른 추정방법으로는 GMM estimation이 있으며, 이에 관해서는 Avery, Hansen and Hotz(1983), Bertschek and Lechner(1998), Inkmann(2000)이 검토하였다.

---

19) Green(2003), Econometric Analysis, pp.693.
20) 함수는 연속 미분 가능함.

3.2.2. 회귀변수와 개체특성 간의 상관관계가 없는 경우의 GMM 추정[21]

한계분포(marginal distribution)는 $E[y_{it} - \Phi(x_{it}'\beta)|x_i] = 0$로 강한 외생성을 만족하며, 직교성을 가정하면 다음과 같다.

(3.48) $$E\begin{bmatrix} y_{i1} - \Phi(x_{i1}'\beta)|x_{i1} \\ y_2 - \Phi(x_{i2}'\beta)\,|x_{i2} \\ \dots \\ y_{i5} - \Phi(x_{i5}'\beta)|x_{i5} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \dots \\ 0 \end{bmatrix}$$

GMM 추정을 위한 과정은 다음과 같다.

1) Step 1

β를 추정하기 위해 프로빗을 사용하는 경우, 먼저 가중행렬(weighting matrix)을 계산하며, 이 때 MLE가 사용될 수도 있다.

(3.49) $$W = \frac{1}{N}\frac{1}{N}\sum_{i=1}^{N}\begin{bmatrix} y_{i1} - \Phi(x_{i1}'\hat{\beta})x_{i1} \\ y_{i2} - \Phi(x_{i2}'\hat{\beta})x_{i2} \\ \dots \\ y_{i5} - \Phi(x_{i5}'\hat{\beta})x_{i5} \end{bmatrix}\begin{bmatrix} y_{i1} - \Phi(x_{i1}'\hat{\beta})x_{i1} \\ y_{i1} - \Phi(x_{i2}'\hat{\beta})x_{i2} \\ \dots \\ y_{i1} - \Phi(x_{i5}'\hat{\beta})x_{i5} \end{bmatrix}$$

2) Step 2

GMM criterion (3.50) $q = \overline{g}(\beta)'\, W^{-1}\,\overline{g}(\beta)$를 극소화시킴으로써 GMM추정량을 얻을 수 있다.

(3.51) $$\overline{g}(\beta) = \frac{1}{N}\sum_{i=1}^{N}\begin{bmatrix} y_{i1} - \Phi(x_{i1}'\hat{\beta})x_{i1} \\ y_{i2} - \Phi(x_{i2}'\hat{\beta})x_{i2} \\ \dots \\ y_{i5} - \Phi(x_{i5}'\hat{\beta})x_{i5} \end{bmatrix}$$

---

21) Avery- Hansen-Hotz (1983).

이러한 GMM 추정량의 점근적 특성은 $\hat{\theta}_{GMM} \rightarrow^{P} \theta$ 이고 $\hat{\theta}_{GMM} \rightarrow^{a}$ $N[\theta, V_{GMM}]$ 이며, 즉, 일치추정량이고, 점근적으로 정규분포를 따른다. 반면, 직교성의 가정이 성립되지 않으면, 즉, $E[y_{it} - \Phi(x_{it}'\beta)|x_i] \neq 0$인 경우 도구변수 Z를 사용하여, GMM criterion을 극소화시키면 minimum distance estimator $\hat{\beta}$을 얻을 수 있고, 이를 비선형 도구변수 추정량이라 한다.

이제까지 살펴본 바와 같이 α_i와 x_{it}가 독립인 경우, 개체특성효과 α_i가 random임에도 fixed된 것으로 고려한다면, β 추정시 효율성이 감소하고, 이전에 논의된 바와 같이 고정효과 estimator는 일치성을 만족하지 못할 것이다. 다시 말하자면 α_i가 x_{it}와 독립이고 univariate distribution G로 부터 유한히 많은 모수 δ에 의해 지수화 되는 random sample 인 경우, log 우도함수는 다음과 같다.[22]

$$(3.52) \qquad \log L = \sum_{i=1}^{N} \log \int \prod_{t=1}^{T} F(\beta' x_{it} + \alpha_i)^{y_{it}} [1 - F(\beta' x_{it} + \alpha_i)]^{1-y_{it}} dG(\alpha_i|\delta)$$

그러므로 약한 정규성 조건하에서 위의 식을 극대화 하면 N이 무한대로 감에 따라 β와 δ의 일치추정량을 생성할 수 있을 것이다. 위 식의 추정절차(즉, α_i와 x_{it}가 독립)는 앞으로 살펴볼 모형(즉, α_i와 x_{it}와의 상관관계 존재)의 추정절차와 필수적인 특성들이 동일하므로 좀 더 일반적인 모형인 (3.54)를 추정하는 절차만을 논의할 것이다. 그러나 α_i와 x_{it}와의 상관관계가 있다면 위 식을 극대화 하는 값은 누락변수 편향을 제거하지 못할 것이다. 그러므로 α_i와 x_{it}의 의존성을 인정한다면 x_{it}가 주어진 경우 α_i값인 $G(\alpha_i|x_{it})$ 분포를 정의해야 하고, marginal log 우도함수를 고려해야 한다.

---

22) $F(\cdot)$는 x_i와 α_i에 따른 오차항의 분포임.

3.3.3. 회귀변수와 개체특성 간의 상관관계가 있는 경우, Chamberlain 모형.

(3.53) $\log L = \sum_{i=1}^{N} \log \int \prod_{t=1}^{T} F(\beta' x_{it} + \alpha)^{y_{it}} [1 - F(\beta' x_{it} + \alpha)]^{1 - y_{it}} dG(\alpha | x)$

chamberlain(1980,1984)의 $\alpha_i = \sum_{t=1}^{T} a'_t x_{it} + \eta_i = a' x_i + \eta_i$ 가정을 이용하여 이제 우리는 회귀함수 $E(\alpha_i | x_i)$가 실제 선형이며 η_i가 x_i에 대해 독립이고, 특정확률 분포를 가진다고 가정할 수 있다. 이들 가정이 주어지면 임의효과 specification의 log 우도함수는 다음과 같다.[23)]

(3.54) $\log L = \sum_{i=1}^{N} \log \int \prod_{t=1}^{T} F(\beta' x_{it} + \alpha' x_i + \eta)^{y_{it}} [1 - F(\beta' x_{it} + \alpha' x_i + \eta)]^{1 - y_{it}} dG^*(\eta)$

예를 들어, F는 표준 정규 분포이고, 평균이 0 분산이 σ_η^2인 정규분포를 따르는 확률변수의 분포함수 G^*를 선택하면, multivariate 프로빗모형을 제시할 수 있다.

(3.55) $y_{it} = 1 \quad \text{if} \quad \beta' x_{it} + a' x_i + \eta_i + v_{it} > 0$

($v_i + e\eta_i$는 평균이 0, 분산 행렬이 $I_T + \sigma_\eta^2 ee'$인 독립, 정규분포 됨.)

(3.52)와 (3.54)의 차이는 오직 incidental parameter problem α_i와 x_i 간의 의존성을 다루는 $a' x_i$ 항을 포함하는데 있다. 때문에 (3.52)와 (3.54)의 추정과 관련된 필수적인 특성들은 동일하다. T차의 합과 관련된 (3.54)의 극대화는 계산이 복잡하므로 univariate integration의 MLE 추정을 단순히 하는 대안은 α_i에 의존하는 오차항 $\epsilon_{it} = \alpha_i + v_{it}$가 평균 0, 분산 1을 갖는 정규분포로 독립 분포되어 있음을 숙지하는 것이다(확률 분포는 $\phi(\epsilon_{it} | \alpha_i)$로 표시됨). 그렇다면 다음과 같은 식으로 나타낼 수 있다.

---

23) G^*는 η에 대한 univariate 분포함수임.

$$(3.56)\quad \Pr(y_{i1},...,y_{iT}) = \int_{C_{i1}}^{b_{i1}} \cdots \int_{C_{i1}}^{b_{iT}} \prod_{t=1}^{T} \phi(\epsilon_{it}|\alpha_i) G(\alpha_i|x_i) d\alpha_i d\epsilon_i \ldots d\epsilon_{iT}$$
$$= \int_{-\infty}^{\infty} G(\alpha_i|x_i) \prod_{t=1}^{T} [\Phi(b_{it}|\alpha_i) - \Phi(c_{it}|\alpha_i)] d\alpha_i.$$

$\Phi(\bullet|\alpha_i)$는 $\phi(\bullet|\alpha_i)$의 누적분포함수(cdf)이며, $G(\alpha_i|x_i)$는 x_i가 주어진 경우 α_i의 확률 분포함수이다.[24)]

$G(\alpha_i|x_i)$가 분산 σ_α^2으로 정규 분포한다고 가정하면 (3.50)은 T차의 적분을 정규 분포와 정확한 근사를 가능하게 하는 정규분포의 누적분포함수(normal c.d.f)와의 교호항으로 이루어진 하나의 적분과정으로 변환시킬 수 있다. 앞에서 살펴본 바와 같이 Butler and Moffit(1982)은 계산의 효율성을 위해 Gaussian quadrature를 사용할 것을 제안했다. (3.52),(3.54)를 극대화 하는 것이 일치성을 지니며 효율적인 β 추정량을 제시함에도 계산상 여전히 복잡하다. 그러나 v_{it}와 η_i (혹은 α_i)가 모두 정규분포 된다고 가정하는 경우 수리적적분(numerical integration)을 피할 수 있는 단순한 접근은 Heckman(1981a) y_{it}의 분포가 x_i에 의존하며, α_i에 한계적인 프로빗 형태를 띤다는 사실을 이용하는 것이다.

$$(3.57)\quad prob(y_{it}=1) = \Phi[(1+\sigma_\eta^2)^{-\frac{1}{2}}(\beta' x_{it} + a' x_i)]$$

ML에 의한 cross-sectional univariate 프로빗 추정법은 $\hat{\Pi}_t \quad (t=1,2,...,T)$ 를 제시하며 n이 무한대로 감에 따라 (3.58) $\Pi = (1+\sigma_\eta^2)^{-\frac{1}{2}}(I_T \otimes \beta' + ea')$로 수렴할 것이다. 때문에 $(1+\sigma_\eta^2)^{-\frac{1}{2}}\beta$와 $(1+\sigma_\eta^2)^{-\frac{1}{2}}a$의 일치추정량이 유도될 수 있다. 추정된 이들 값을 (3.54)에 대체하여, $(1+\sigma_\eta^2)^{-\frac{1}{2}}\beta$와 $(1+\sigma_\eta^2)^{-\frac{1}{2}}a$ 에 의존

24) $y_{it}=1$이면 $c_{it}=-\beta' x_{it}$, $b_{it}=\infty$, $y_{it}=0$이면 $c_{it}=-\infty$, $b_{it}=-\beta' x_{it}$

하는 σ_η^2와 관련한 함수를 극대화 시킬 수 있다.

또한 수리적 적분과정(numerical integration)을 피할 수 있는 보다 효율적인 추정량으로 minimum-distance estimator를 사용하는 방법이 Chamberlain(1984)에 의해 제안되었으며 Chmberlain(1984)는 (3.56) $[\hat{\Pi}-f(\theta)]'\hat{\Omega}^{-1}[\hat{\Pi}-f(\theta)]'$를 최소화하는 $\hat{\theta}$를 선택할 것을 제안한다. [25)]

3.3. 준모수적 접근법

앞에서 살펴본 binary choice 모형 추정의 모수적 접근법은 두 가지 관점에서 약점을 갖는다. (1) X에 의존하는 $(\varepsilon_{it}, \alpha_i)$의 확률이 사전에 알려지거나 X와 α_i에 의존하는 ϵ_{it}의 확률이 사전에 알려진다는 것. (2) α_i가 고정된 경우, 로짓과 선형확률모형을 제외하면 incidental parameter를 제거할 수 있는 간단한 형태(simple transformation)가 존재하지 않는다는 것이다. 준모수적 접근은 ε의 분포를 가정하지 않을 뿐 아니라, α_i가 fixed이든 random 이든 β의 일치추정을 행한다. 그럼 다음의 준모수적 접근법을 고려하자(t와 s의 2기(期,) ϵ_{is}는 α_i, x_{it}, x_{is}에 동일분포 됨).

$$(3.57)\quad y_{it} = \begin{cases} 1 & \text{if} \quad x_{it}\beta + \alpha_i + \epsilon_{it} \geq 0 \\ 0 & otherwise \end{cases}$$

Manski(1987)는 이 모형에서 다음과 같은 사실을 관찰할 수 있다.

---

25) Hsiao (2003).
$\Pi = Vec(\Pi') = f(\theta)$, $\theta' = (\beta', a', \sigma_\eta^2)$로 (3.52)에 제약을 부과하여, 선형의 경우와 같이 minimum-distance estimator를 사용하는 것이다.

(3.58) $median(y_{it}-y_{is}|\alpha_i, x_{it}, x_{is}, y_{it} \neq y_{is}) = sgn((x_{it}-x_{is})\beta)$

중요한 것은 우변이 α_i에 의존하지 않고, β에 의존한다는 사실이다.

결국 (3.59) $\sum_{i=1}^{n} |(y_{it}-y_{is}) - sgn((x_{it}-x_{is})b)|$를 최소화 함으로써 β를 추정할 수 있고, 이는 (3.60) $\sum_{i=1}^{n} sgn(y_{it}-y_{is}) \cdot sgn((x_{it}-x_{is})b)$를 극대화함과 같다. 결과 추정량은 일치추정량이지만 $\sqrt{n}$ 점근적 분포를 따르지는 않는다.

3.3.1. Maximum score estimator (MSE)

Manski(1975,1985,1987)는 (3.61)과 같은 표본평균함수를 극대화하는 maximum score estimator를 제안한다.

(3.61) $H_N(b) = \frac{1}{N}\sum_{i=1}^{N}\sum_{t=2}^{T} sgn(\triangle x_{it}'b)\triangle y_{it}.$

(표준화 조건 $b'b=1$하에서 $\triangle x_{it} = x_{it} - x_{i,t-1}$, $\triangle y_{it} = y_{it} - y_{i,t-1}$이고,

(3.62) $w > 0$ 이면 $sgn(w)=1$ 로 정의함.)
$w = 0$ 이면 $sgn(w)=0$
$w < 0$ 이면 $sgn(w)=-1$

이는 일반적인 조건 하에서 (3.63) $H(b) = E[sgn(\triangle x_{it}'b)\triangle y_{it}]$로 항상 수렴하기 때문이다.[26] 이를 자세히 알아보기 위해 binary choice모형에서는 (3.64)

$$y_{it} = \begin{cases} 1 & if \;\; y^* > 0 \\ 0 & if \;\; y^* \leq 0 \end{cases}$$

의 형태로 쓰여질 수 있음을 상기하자.

($y_{it^*} = \beta' x_{it} + \epsilon_{it}$ 이며, $\epsilon_{it} = \alpha_i + v_{it}$ 임)

---

26) $H(b)$는 $b=\beta$에서 극대화되며, $\beta^* = \frac{\beta}{\|\beta\|}$ 임. $\|\beta\|$ 는 euclidean norm $\sum_{k=1}^{K}\beta_k^2$ 임.

v_{it}가 독립적으로 동일분포 되어있고 (즉, i.i.d), x_i와 α_i에 독립이라는 가정 하에서 다음을 얻을 수 있다.

$$\begin{aligned}(3.65)\quad & x_{it}'\beta > x_{i,t-1}'\beta \ \Leftrightarrow E(y_{it}|x_{it}) > E(y_{i,t-1}|x_{i,t-1}) \\ & x_{it}'\beta = x_{i,t-1}'\beta \ \Leftrightarrow E(y_{it}|x_{it}) = E(y_{i,t-1}|x_{i,t-1}) \\ & x_{it}'\beta < x_{i,t-1}'\beta \ \Leftrightarrow E(y_{it}|x_{it}) < E(y_{i,t-1}|x_{i,t-1})\end{aligned}$$

이 식을 1차 차분 형태로 다시 쓰면,

$$\begin{aligned}(3.66)\quad & \triangle x_{it}'\beta > 0 \ \Leftrightarrow E(y_{it} - y_{i,t-1}|\triangle x_{it}) > 0 \\ & \triangle x_{it}'\beta = 0 \ \Leftrightarrow E(y_{it} - y_{i,t-1}|\triangle x_{it}) = 0 \\ & \triangle x_{it}'\beta < 0 \ \Leftrightarrow E(y_{it} - y_{i,t-1}|\triangle x_{it}) < 0\end{aligned}$$

c>0 인 경우, $\overline{\beta}=\beta_c$를 계속 유지하며, 때문에 nomalized vector $\beta^* = \frac{\beta}{\|\beta\|}$ 만을 고려하겠다. 그러면 $b \neq \beta^*$와 같은 어떤 b에 대해서 ($b'b=1$만족) 다음과 같다.

$$\begin{aligned}(3.67)\quad H(\beta^*) - H(b) &= E\{[sgn(\triangle x_{it}'\beta^*) - sgn(\triangle x_{it}'b)](y_{it} - y_{i,t-1})\} \\ &= 2\int_{w_b} sgn(\triangle x_{it}'\beta^*) E[y_t - y_{t-1}|\triangle x_{it}] dF_{\triangle X}\end{aligned}$$

($W_b = [\triangle x : sgn(\triangle x_{it}'\beta^*) \neq sgn(\triangle x_{it}'b)]$이고, $F_{\triangle x}$는 $\triangle x$의 분포를 표시함)

(3.66)으로 인해 (3.67)은 모든 $\triangle x$에 대하여 다음의 내용을 함축한다.

$$(3.68)\quad sgn(\triangle x'\beta^*) E[y_t - y_{t-1}|\triangle x] = |E[y_t - y_{t-1}|\triangle x]|$$

그러므로 $x's$ 에 대한 가정 하에 다음과 같다.

$$(3.69)\quad H(\beta^*) - H(b) = 2\int_{w_b} E[y_t - y_{t-1}|\triangle x] dF_{\triangle X} > 0$$

Manski(1985,1987)는 일반적인 조건 하에서 (3.61)을 극대화하는 추정량이 β^*의 일치추정량을 생성함을 보인다. 준모수적 접근은 잠재변수표현의 선형 구조를 사용하며, 개체특성효과 α_i는 차분함으로써 다시 제거될 수 있으므

로 α_i는 더 이상 β에 영향을 미치지 않는다. Manski의 maximum score estimator는 $N \rightarrow \infty$로 감에 따라 일치성을 만족하나 이는 모수적 접근의 $N^{\frac{1}{2}}$보다 훨씬 느린 $N^{\frac{1}{3}}$ 속도로 수렴한다.

4. 동태적 패널 이진변수모형

4.1. 일반 모형

이제까지는 사건의 확률이 과거 사건의 발생, 혹은 비발생에 독립인 정태적 모형을 논의했다. 그러나 다양한 문헌들로 부터 과거에 사건을 경험한 사람은 그렇지 않은 사람보다 미래에 그 사건을 경험할 가능성이 높다고 고려된다.[27] 달리 말하면, 개개인이 미래에 그 사건을 경험할 조건부 확률은 과거 경험의 함수라는 것이다. 이산적인 변수들 간의 이시(異時)적인 관계를 분석하기 위해 Heckman(1978,1981b)은 잠재적-연속적 변수의 관점에서 일반적인 틀을 제시했다. 그는 연속 확률 변수 y^*_{it}를 x_{it}와 과거 사건 발생의 함수로 두었다.

$$(4.1) \qquad y^*_{it} = \beta' x_{it} + \sum_{l=1}^{t-1} \gamma_l y_{i,t-1} + \phi \sum_{s=1}^{t-1} \prod_{l=1}^{s} y_{i,t-l} + \epsilon_{it}$$

$$(4.2) \qquad \begin{aligned} y_{it} &= 1 \qquad if \quad y_{it^*} > 0 \\ &\ 0 \qquad if \quad y_{it^*} \le 0 \end{aligned}$$

오차항 ϵ_{it}는 x_{it}에 독립이며, i 에 독립 분포되고 일반적인 이시적 분산 행

---

27) Bates and Neyman(1951).
브랜드 로열티 : Chintagunta Kyriazidou and Perktold(2001),
노동시장 참가 : Heckman and Wills(1977), Hyslop(1999),
실업 : Layton(1978).

렬 $Ev_i v_i' = \Omega$ 를 가지는 것으로 가정한다. 계수 γ_l은 y^*_{it} 값에 대한 l 기 이전의 사건 경험 효과를 나타내고, 계수 ϕ는 y^*_{it} 값에 대한 사건의 누적된 효과를 나타낸다. 이전에 언급한 바와 같이 시간에 대한 개체의 반복적인 관측은 사건을 경험하는 개체들의 성향이 다르다는 모형을 설정하게 해주며 그러한 heterogeneity는 다음과 같이 분해할 수 있다.

(4.3) $\epsilon_{it} = \alpha_i + v_{it}$ $(i = 1,\ldots,N \quad T = 1,\ldots,T)$

v_{it}는 임의의 계열상관을 지니며, i에 독립 분포된다. 또한 α_i는 개체효과로서 고정된 상수 혹은 확률변수로 고려된다. fixed-T인 동태 모형의 경우 고정효과모형은 일치성을 만족하지 못한다(T=2인 경우).

(4.1)~(4.3)은 관측된 설명변수 x를 통제한 후 세 가지 지속성(persistence)를 따른다. 지속성의 원인으로는 오차항 v_{it}의 계열상관의 결과, 혹은 관측되지 않는 이질성 α_i의 결과, 혹은 $\gamma y_{i,t-l}$, $\phi \sum_{s=1}^{t-l} \prod_{l=1}^{s} y_{i,t-l}$ 항의 상태의존성(state dependence)의 결과가 될 수 있다. 개체가 그 상태로 유지되는 조건부 확률이 과거 경험의 함수인 경우 두 가지 새로운 이슈가 발생할 수 있는데 하나는 초기 관측치(initial observations)를 다루는 법이고, 다른 하나는 관측되지 않는 개체효과의 대변수(proxy)로서 과거 값에 있는 가성상태의존을 진성상태의존과 어떻게 구분하느냐 하는 것이다.

첫 번째 이슈는 주어진 모형에 대한 일치추정량을 유도하는 데 있어서 중요하며, 두 번째 이슈는 관측된 사건의 상태의존성이 실제 사건의 경험이 개인의 행동을 바꾸었다는 사실로부터, 혹은 시간과 상관관계가 있는 관측되지 않는 부분으로부터, 아니면 이 둘의 결합으로 발생될 수 있기때문에 중요하다.

4.2 초기조건 (Initial conditions)

초기조건의 문제와 해결책에 대한 논의를 위해서는 외생변수가 존재하지 않으며, 관측된 데이터는 1차 Markov 과정으로 생성됨을 가정한다.[28)]

즉, (4.4) $y_{it^*} = \beta_0 + \gamma y_{i,t-1} + v_{it}$

$$y_{it} = \begin{cases} 1 & if \ \ y_{it}^* > 0 \\ 0 & if \ \ y_{it}^* \le 0 \end{cases}$$

사회과학의 많은 연구에서는 전형적으로 초기조건에 대한 두 가지 가정이 행해진다.

(1) 초기조건은 외생적인 것으로 가정하거나

(2) process는 균형을 이룬다고 가정한다.

y_{i0}가 개체 i에 대해 고정된 상수라는 가정 하에서 α_i가 주어진 경우 $y_i' = (y_{i1}, \ldots, y_{iT})$의 결합 확률은 다음과 같다.

(4.5) $$\prod_{t=1}^{T} F(y_{it}|y_{i,t-1}, \alpha_i) = \prod_{t=1}^{T} \Phi\{(\beta_0 + \gamma y_{i,t-1} + \alpha_i)(2y_{it} - 1)\}$$

process가 균형을 이룬다는 가정 하 주어진 α_i에 대해 y_{i0}=1일 한계확률은 (Karlin and Taylor(1975)) 다음과 같고,

(4.6) $$P_i = \frac{\Phi(\beta_0 + \alpha_i)}{1 - \Phi(\beta_0 + \gamma + \alpha_i) + \Phi(\beta_0 + \alpha_i)}$$

y_{i0}=0일 한계확률은 1- P_i가 된다.

α_i가 주어진 경우 $(y_{i1}, \ldots, y_{iT})$는 다음과 같다.

---

28) u_{it}의 평균은 0, 분산 σ_u^2은 1로 표준화되어 정규 분포됨을 가정한다 .

(4.7) $\prod_{t=1}^{T}\Phi\{(\beta_0+\gamma y_{i,t-1}+\alpha_i)(2y_{it}-1)\}p_i^{y_{i0}}(1-p_i)^{1-y_{y0}}$

α_i가 $G(\alpha)$의 분포로 임의적이면, 첫 번째 가정 하 임의효과 모형의 우도함수는 다음과 같으며, (4.8) $L=\prod_{i=1}^{N}\int\prod_{t=1}^{T}\Phi\{(\beta_0+\gamma y_{i,t-1}+\alpha_i)(2y_{it}-1)\}dG(\alpha)$

두 번째 가정 하 우도함수는 다음과 같다.

(4.9) $L=\prod_{i=1}^{N}\int\prod_{t=1}^{T}\Phi\{(\beta_0+\gamma y_{i,t-1}+\alpha_i)(2y_{it}-1)\}p_i^{y_{i0}}(1-p_i)^{1-y_{y0}}dG(\alpha)$

α_i가 random으로 고려되면 β_0, γ, σ_α^2의 MLE는 N이 무한대로 감에 따라 혹은 N, T가 무한대로 감에 따라 일치성을 만족한다. α_i가 고정효과로 고려되면, β_0, γ, α_i의 MLE는 T가 무한대로 감에 따라 일치성을 만족하지만, T가 유한한 경우 MLE는 편향을 지니며, Monte carlo 실험 결과 정태적인 경우와는 반대로 이 편향은 유의한 결과를 갖는다. [29]

그러나 초기조건이 고정된 상수라는 가정은 교란항이 계열 간에 독립이고 새로운 process가 표본 초기에 우연히 관측되는 경우에만 정의될 수 있으며 process가 균형을 이룬다는 가정 또한 응용 시에 많은 문제를 발생시킨다. process가 표집 되기 이전에 수행된다면 혹은 모형의 교란항이 개체-특성 임의효과 존재 시 계열 간에 종속되면 초기조건은 외생적이지 않다. 분석자가 처음부터 그 process에 접하지 않는다고 가정하면 개체 i의 초기상태 y_{i0}는 고정된 것으로 가정될 수 없으며 초기상태는 패널 표본을 만드는 과정에서 결정된다. 고정효과 모형의 표본 우도함수는 다음과 같으며

(4.10) $L=\prod_{i=1}^{N}\prod_{t=1}^{T}\Phi\{(\beta_0+\gamma y_{i,t-1}+\alpha_i)(2y_{it}-1)\}f(y_{i0}\mid\alpha_i)$.

임의효과 모형의 표본 우도함수는 (4.11)과 같이 나타낼 수 있다.

29) Heckman (1981b).

(4.11) $L=\prod_{i=1}^{N}\int_{-\infty}^{\infty}\prod_{t=1}^{T}\Phi\{(\beta_0+\gamma y_{i,t-1}+\alpha_i)(2y_{it}-1)\}f(y_{i0}\mid\alpha_i)dG(\alpha).$

($f(y_{i0}\mid\alpha)$는 α_i가 주어진 경우 y_{i0}의 한계확률임)

그러므로 T가 매우 크지 않으면 (4.5) 혹은 (4.8)을 극대화하는 값은 일치추정량을 제공하지 않는다. y_{i0}는 관측되지 않는 과거 값의 함수이고, $f(y_{i0}\mid\alpha)$는 유도하기 어렵기 때문에 (4.10) 혹은 (4.11)을 극대화 하는 것 또한 상당히 복잡하다. 그리하여 Heckman(1981b)는 동태 이산형변수모형의 초기조건을 다음과 같이 근사할 것을 제안했다.

1. 표본의 초기 값 y_{i0}의 확률을 프로빗 모형으로 근사한다.

(4.12) $y_{i0}^{*}=Q(x_i)+\epsilon_{io}$ 이고 $y_{i0}=\begin{cases}1 & if\ \ y_{i0}^{*}>0\\ 0 & if\ \ y_{i0}^{*}\le 0\end{cases}$ 이다.

2. ϵ_{i0}가 v_{it}와 자유롭게 상관관계를 갖도록 한다.

3. structural system의 모수와 표본 초기 상태에 대한 축약형(reduced)확률을 근사하는 모수 간의 어떠한 제약 없이 ML로 추정한다.

4.3 조건부 접근법 (conditional approach)

T가 고정되고 N이 무한대로 가는 동태 패널에서 개체효과가 고정되면 우도함수 접근은 일치추정량을 제공하지 않는다. 개체 효과가 random이고 x에 독립이면 MLE의 일치성은 the effect와 초기 관측치의 확률 분포 공식에 의존한다. 그리고 준모수적 접근은 동태 모형에서 활용될 수 없다. 왜냐하면 시차종속변수를 설명변수로 고려함에 따라 준모수 모형의 외생성 조건을 위반하기 때문이다.[30] 그러므로 Manski(1985)의 maximum score estimation은 수행할 수 없고 어느 쪽에서도 조건부 접근법은 행할 수 없다. T=2인 경우를 고려해 보면 조건부 접근의 기본적인 아이디어는 $y_{i1} \neq y_{i2}$이고, 두 기 모두의 설명변수에 의존하는 y_{i2}=1 or 0의 확률을 고려하는 것이다. 만약 $\mathrm{Pr}ob(y_{i2}=1)$의 설명변수가 y_{i1}을 포함하면 조건부 확률은 y_{i1}=0 혹은 y_{i1}=1인 확률 모두를 포함한다. y_{i1}=0 or 1임에 따르는 조건부 확률은 1 혹은 0 이 될 것이고 γ와 β에 관한 정보를 제공하지 않을 것이다.

그러나 $T \geq 3$ 이고 x_{it}가 특정 패턴을 따르는 경우, Honore and Kyriazidou는 로짓 모형의 일치추정량을 찾는 조건부 확률접근법을 일반화 할 수 있으며, α_i 혹은 초기조건에 대한 확률 분포 없이 maximum score접근을 일반화 할 수 있음을 보였다. 그러나 추정량의 (진정한 값으로의)수렴속도는 보통의 $\sqrt{n}$ 보다 느리다.

4.3.1. 로짓 모형의 경우

외생변수 x_{it}가 나타나는 잠재 반응 함수 $y_{it}^* = x_{it}\beta + \gamma y_{i,t-1} + \alpha_i + u_{it}$를 고려하여, 다음의 고정효과 로짓 모형을 살펴보자.

(4.13) $P(y_{i0}=1 | x_i, \alpha_i) = p_o(x_i, \alpha_i)$

$$P(y_{it}=1 | x_i, \alpha_i, y_{i0}, \ldots, y_{it-1}) = \frac{\exp(x_{it}\beta + \gamma y_{it-1} + \alpha_i)}{1+\exp(x_{it}\beta + \gamma y_{it-1} + \alpha_i)}$$

[($x_i=(x_{i1}, \ldots, x_{iT})$이며, 4.3.1에서는 모두 T=3인 경우를 고려한다.]

우리는 y_{i0}가 관측되는 것으로 가정하지만, 설명변수들이 초기에 관측되는 것으로 가정할 필요는 없다. (4.13)을 유도하는 임계치모형의 오차항들이 로지스틱분포로 시간에 따라 i.i.d이며, 모든 기간 (x_i, α_i, y_{i0})에 독립이라는 함

30) 조건 위반시 $E(\triangle u_{it} | x_{it}, x_{i,t-1}, y_{i,t-1}, y_{i,t-2}) \neq 0$ 이 됨.

축적 가정을 숙지하는 중요하다. 우리는 조건부 로짓 접근법을 따라 개체효과에 의존하지 않는 확률 셋(set)을 유도하는 것이 목적이다.

Chamberlain(1985)에 따르면, y_{i0}와 y_{i3}가 0과 1 둘 중 하나인 경우 다음과 같은 사건을 고려할 수 있다.

(4.14) $$A=\{y_{i0},\, y_{i1}=0, y_{i2}=1, y_{i3}\}$$
$$B=\{y_{i0},\, y_{i1}=1, y_{i2}=0, y_{i3}\}$$

간단한 계산과정으로 다음과 같다.

(4.15) $$P(A|x_i,\alpha_i)=p_0(x_i,\alpha_i)^{y_{i0}}(1-p_0(x_i,\alpha_i))^{1-y_{i0}} \times \frac{1}{1+\exp(x_{i1}\beta+\gamma y_{i0}+\alpha_i)} \times \frac{\exp(x_{i2}\beta+\alpha_i)}{1+\exp(x_{i2}\beta+\alpha_i)} \times \frac{\exp(y_{i3}x_{i3}\beta+y_{i3}\gamma+y_{i3}\alpha_i)}{1+\exp(x_{i3}\beta+\gamma+\alpha_i)}$$

(4.16) $$P(B|x_i,\alpha_i)=p_0(x_i,\alpha_i)^{y_0}(1-p_0(x_i,\alpha_i))^{1-y_{i0}} \times \frac{\exp(x_{i1}\beta+\gamma y_{i0}+\alpha_i)}{1+\exp(x_{i1}\beta+\gamma y_{i0}+\alpha_i)} \times \frac{1}{1+\exp(x_{i2}\beta+\alpha_i+\gamma)} \times \frac{\exp(y_{i3}x_{i3}\beta+y_{i3}\alpha_i)}{1+\exp(x_{i3}\beta+\alpha_i)}$$

일반적으로 1기와 2기 간의 종속변수 변화 events에 비례하는 확률 $P(A|x_i, \alpha_i, A\cup B)$와 $P(B|x_i, \alpha_i, A\cup B)$는 α_i에 의존할 것이다. 이것은 조건부 우도함수 접근법이 고정효과를 제거하지 못할 것이기 때문이다. 그러나 Honore and Kyriazidou (2000a)는 만약 $x_{i2}=x_{i3}$라면 조건부 확률은 다음과 같이 α에 의존하지 않음을 보인다.

(4.17) $$P(A|x_i,\alpha_i,A\cup B, x_{i2}=x_{i3})=\frac{1}{1+\exp((x_{i1}-x_{i2})\beta+\gamma(y_{i0}-y_{i3}))}$$

$$(4.18)\qquad P(B|x_i,\alpha_i,A\cup B,x_{i2}=x_{i3})=\frac{\exp((x_{i1}-x_{i2})\beta+\gamma(y_{i0}-y_{i3}))}{1+\exp((x_{i1}-x_{i2})\beta+\gamma(y_{i0}-y_{i3}))}$$

이것이 모든 결과들의 핵심이며 Honore and Kyriazidou (2000a)는 다음 식을 극대화 함으로써 β와 γ를 추정할 것을 제안한다.

$$(4.19)\qquad \sum_{i=1}^{n}1\{y_{i1}+y_{i2}=1\}1\{x_{i2}-x_{i3}=0\}\times\ln\left(\frac{\exp((x_{i1}-x_{i2})b+g(y_{i0}-y_{i3}))^{y_{i1}}}{1+\exp((x_{i1}-x_{i2})b+g(y_{i0}-y_{i3}))}\right)$$

모든 설명변수들이 이산형[31]이고 x_{it}가 $P(x_{i2}=x_{i3})>0$를 만족하는 특수한 경우, β와 γ에 관한 추론을 위해 (4.17)와 (4.18)을 사용할 수 있다. 결과 추정량은 표준적인 가정 하에서 일치성과 $\sqrt{n}$ 점근적 분포의 보통의 모든 특성 가진다. 그러나 $x_{i2}=x_{i3}$인 관측치 만에 기초한 추론은 몇몇 경우에서는 합리적이겠지만[32], 경제 응용에서는 그렇지 않은 경우도 많다. 그리하여 위의 목적함수에 있는 지시함수 1{$x_{i2}-x_{i3}$=0}을 $x_{i2}-x_{i3}$의 차이에 역으로 의존하는 가중치(x_{i2}가 x_{i3}에 가까운 관측치에 더 많은 가중치를 주는)로 대체하는 것이 아이디어이다. Honore and Kyriazidou(2000)는 특히, 다음을 극대화함으로써 β와 γ를 추정하는 방법을 제안한다.

$$(4.20)\qquad \sum_{i=1}^{n}1\{y_{i1}+y_{i2}=1\}K\left\{\frac{x_{i2}-x_{i3}}{\sigma_n}\right\}\times\ln\left(\frac{\exp((x_{i1}-x_{i2})b+g(y_{i0}-y_{i3}))^{y_{i1}}}{1+\exp((x_{i1}-x_{i2})b+g(y_{i0}-y_{i3}))}\right)$$

여기서 $K(\cdot)$는 관측치 i에 적절한 가중치를 주는 kernel density function 이며, σ_n은 n 이 무한대로 증가함에 따라 0으로 줄어든다. 점근적 이론에 의하면 $K(\cdot)$는 $K(v)\to 0\ as\ \|v\|\to\infty$와 같은 통상조건을 만족하도록 선택된다. $x_{i2}-x_{i3}$는 0 근방에서 연속적으로 분포되며, $x_{i1}-x_{i2}$는 $x_{i2}-x_{i3}$=0인 사

31) x_{it}가 연속인 경우 $x_{i2}=x_{i3}$는 드물 것이다.

32) 특히, x_i가 연구자의 통제 하에 있는 경우.

건에 비례하는 충분한 분산을 지닌다. (4.20)을 극대화하는 추정량들의 수렴속도는 비록 $n^{-1/2}$보다 느릴지라도 그것이 일치추정량이며 점근적으로 정규분포포를 따르고, x_{it}의 연속적인 공변인(covariate)의 수에 의존한다.

조건부 접근은 로짓모형에서는 잘 수행하지만, 일반적인 비선형 모형에 적용가능해 보이지는 않는다. 그러나 비선형성이 transformation function F를 갖는 하나의 index 형태 $F(\alpha)$로 나타날 수 있다면, Manski(1987)의 (정태적)maximum score estimator는 다음을 고려함으로써 시차 종속변수를 설명변수로 갖는 경우로 일반화 할 수 있다.

4.3.2. 준모수적 접근법의 경우

4.3.2에서는 (4.13)의 임계치모형의 정의에 기반을 둔다. 시간에 따라 변하는 오차의 분포에 대한 로짓 가정을 완화시키기 위해 Manski(1987)의 통찰을 살펴볼 것이다. 4.3.1과 같이 시간이 흐름에 따른 독립성 가정은 유지될 것이다.

(4.21) $P(y_{i0}=1\,|\,x_i,\alpha_i) = p_0(x_i,\alpha_i)$

(4.22) $P(y_{i0}=1\,|\,x_i,\alpha_i,y_{i0},...,y_{it-1}) = F(x_{it}\beta+\gamma y_{it-1}+\alpha_i) \qquad (t=1,...,T)$

전과 같이, 개개인당 4개 관측치가 존재하는 경우(즉, T=3)에 초점을 맞춘다. A와 B는 4.3.1에서와 같이 정의된다.[33]

33) $A=(\,y_{i0}\,,\,y_{i1}=0,\;y_{i2}=1\,,\;y_{i3})$
$B=(\,y_{i0}\,,\,y_{i1}=1,\;y_{i2}=0\,,\;y_{i3})$의 경우.

(4.23)

$$P(A \mid x_i, \alpha_i, x_{i2} = x_{i3})$$
$$= p_0(x_i, \alpha_i)^{y_{i0}} (1 - p_0(x_i, \alpha_i))^{1 - y_{i0}} \times [1 - F(x_{i1}\beta + \gamma y_{i0} + \alpha_i)]$$
$$\times F(x_{i2}\beta + \alpha_i) \times [1 - F(x_{i2}\beta + \gamma + \alpha_i)]^{1 - y_{i3}}$$
$$\times\ F(x_{i2}\beta + \gamma + \alpha_i)^{y_{i3}}$$

$$P(B \mid x_i, \alpha_i, x_{i2} = x_{i3})$$
(4.24)
$$= p_0(x_i, \alpha_i)^{y_{i0}} [1 - p_0(x_i, \alpha_i)]^{1 - y_{i0}} \times F(x_{i1}\beta + \gamma y_{i0} + \alpha_i)$$
$$\times [1 - F(x_{i2}\beta + \gamma + \alpha_i] \times [1 - F(x_{i2}\beta + \alpha_i)]^{1 - y_{i3}}$$
$$\times\ F(x_{i2}\beta + \alpha_i)^{y_{i3}}$$

만약 y_{i3}=0이면 다음과 같고

(4.25)
$$\frac{P(A \mid x_i, \alpha_i, x_{i2} = x_{i3})}{P(B \mid x_i, \alpha_i, x_{i2} = x_{i3})}$$
$$= \frac{1 - F(x_{i1}\beta + \gamma y_{i0} + \alpha_i)}{1 - F(x_{i2}\beta + \alpha_i)} \times \frac{F(x_{i2}\beta + \alpha_i)}{F(x_{i1}\beta + \gamma y_{i0} + \alpha_i)}$$
$$= \frac{1 - F(x_{i1}\beta + \gamma y_{i0} + \alpha_i)}{1 - F(x_{i2}\beta + \gamma y_{i3} + \alpha_i)} \times \frac{F(x_{i2}\beta + \gamma y_{i3} + \alpha_i)}{F(x_{i1}\beta + \gamma y_{i0} + \alpha_i)}$$

y_{i3}=1이면 다음과 같다(y_{i3}=1 이면 $\gamma y_{i3} = \gamma$).

(4.26)
$$\frac{P(A \mid x_i, \alpha_i, x_{i2} = x_{i3})}{P(B \mid x_i, \alpha_i, x_{i2} = x_{i3})}$$
$$= \frac{1 - F(x_{i1}\beta + \gamma y_{i0} + \alpha_i)}{1 - F(x_{i2}\beta + \gamma + \alpha_i)} \times \frac{F(x_{i2}\beta + \gamma + \alpha_i)}{F(x_{i1}\beta + \gamma y_{i0} + \alpha_i)}$$
$$= \frac{1 - F(x_{i1}\beta + \gamma y_{i0} + \alpha_i)}{1 - F(x_{i2}\beta + \gamma y_{i3} + \alpha_i)} \times \frac{F(x_{i2}\beta + \gamma y_{i3} + \alpha_i)}{F(x_{i1}\beta + \gamma y_{i0} + \alpha_i)}$$

어느 경우에서나 F는 다음 사실을 함의하며,

(4.27) $$\frac{P(A)}{P(B)} > 1 \qquad \text{if } x_{i2}'\beta + \gamma y_{i3} > x_{i1}'\beta + \gamma y_{i0}$$

$$\frac{P(A)}{P(B)} < 1 \qquad \text{if } x_{i2}'\beta + \gamma y_{i3} < x_{i1}'\beta + \gamma y_{i0}$$

(4.28)

$$\begin{aligned} & sgn[P(A \mid x_i, \alpha_i, x_{i2} = x_{i3}) - P(B|x_i, \alpha_i, x_{i2} = x_{i3})] \\ & = sgn[(x_{i2} - x_{i1})\gamma + \beta(y_{i3} - y_{i0})]. \end{aligned}$$ 가 된다.

만약 $P(x_{i2} = x_{i3}) > 0$이면 maximum score estimator는 $A \cup B$와 $x_{i2} = x_{i3}$를 만족시키는데 적용될 것이다. β와 γ 는 다음을 극대화시킴으로써 추정될 것이다.

(4.29) $$\sum_{i=1}^{n} 1\,(x_{i2} - x_{i3} = 0)(y_{i2} - y_{i1})\,sgn((x_{i2} - x_{i1})\beta + \gamma(y_{i3} - y_{i0}))$$

$y_{i1} + y_{i2} = 1$을 만족하는 관측치 만이 추정 시에 사용되었으며 로지스틱 경우와 유사하게, $x_{i2} - x_{i3}$는 0 근방에서 연속적으로 분포된다. Honore and Kyriazidou (2000a)는 다음의 스코아 함수를 극대화하는 maximum score estimator를 제안하며, $x_{i2} - x_{i3}$의 분포 $f(x_{i2} - x_{i3})$가 0에서 양수라면(즉, $f(0) > 0$) 그 추정량은 일치성을 만족한다.

(4.30) $$\sum_{i=1}^{n} K\!\left(\frac{x_{i2} - x_{i3}}{\sigma_n}\right)(y_{i2} - y_{i1})\,sgn((x_{i2} - x_{i1})b + g(y_{i3} - y_{i0}))$$

조건부 접근법은 표본 초기 관측치에 대한 가정을 필요로 하지 않고, 관측된 설명변수 혹은 초기조건과 개체 효과와의 통계적 관계를 가정 하지 않는다. 그러나 임계치 모형의 오차항이 시간에 따라 독립이라는 가정과 $x_{is} - x_{it}$ $(t \neq s)$가 0 근방에 분포된다는 가정의 한계를 가지며, 그것은 설명변수로서 시간 더미를 배제한다. 개체효과가 추정될 수 없다는 사실 또한

설명변수의 값으로 나타난 개인의 탄력성을 계산할 수 없으며, 예측할 수 없음을 의미한다.

4.4 MMLE (Modified MLE)

MMLE는 점근적 분산의 증가 없이 MLE의 편향을 $O(T^{-1})$에서 $O(T^{-2})$로 줄이는데 사용되고 직교되는 모수를 필요로 하지는 않는다. 내생변수와 외생변수의 시차를 포함하는 프로빗, 로짓모형의 추정량은 패널데이타에서 편향을 지니는데, MMLE의 차별적인 장점은 그것의 일반성이다. Monte carlo experiments는 고정효과를 지닌 비선형 패널 데이타의 전통적 MLE가 T가 크지 않은 경우, 상당한 편향을 나타냄을 보인다. 선형 모형에서는 T가 고정된 경우 일치추정량이 가능하므로 추정방법과 문제해결이 잘 알려져 있는 반면, 비선형 모형의 일반적인 해법은 존재하지 않으며 $\sqrt{n}$일치성을 만족하지 않는다. 예를 들어 Honore and Kyriazidou(2000)는 (T가 고정된 경우) 제한적 가정을 필요로 하는 연속적인 외생변수를 지닌 동태 이산형 패널의 일치추정량을 제시하지만 그것은 $\sqrt{n}$일치성을 만족하지 않는다.

더불어 Honore and Tamer(2004)는 동태적 이산형 변수 패널의 모수가 간단한 경우에도, T가 고정된 경우에서는 정의되지 않음을 보였다. ML에 의한 고정효과를 지닌 비선형 모형의 추정은 incidental parameter problem이라는 문제를 겪는다. Cox and Reid(1987)는 장애모수에 관한 정보 부족 시 모수 추정의 일반적인 문제를 고려하였는데, 해결책은 장애모수의 영향을 제한하고 다른 모수에 직교하는 장애모수를 재모수화(reparametrization)하는데 기초한다. 그리고 MLE의 편향을 점근적 분산의 증가 없이 줄이는 수정을 했다. 일반적인 틀은 고정효과가 장애모수인 Arellano(2003)의 고정

효과를 지닌 정태 이진변수 패널 데이타에서 사용되며, Arellano(2003)는 기존의 모수 관점에서 수정을 행한다. Arellano(2003)는 그가 유도한 modification이 직교 재모수화 존재의 가정 하, MLE의 편향을 줄임을 보인다. 그러나 문제는 직교 재모수화의 정보가 존재하지 않을지도 모른다는 것이다.

Arellano(2003)에서와 같이 모형의 기존 모수 관점에서 수정한 MMLE는 직교 재모수화에 대한 정보의 존재와 상관없이 편향을 줄이며, 이 결과는 동태 이진변수 모형 뿐만 아니라, 일반적인 패널 데이타 비선형 모형에서도 지속된다. 비록 MMLE는 T가 무한대로 갈 때만 일치추정량을 제시하지만 외생, 내생 변수의 시차변수를 지닌 패널 프로빗모형 추정 시에 유용하다. 이는 T가 그리 크지 않은 경우 유한한 표본에서 편향을 야기하는 MLE의 점근적 편향을 줄이기 때문이다. 이 방법은 고정효과가 존재하는 비선형 모형 추정 시 일반적인 틀을 제시하는데, 즉, 로지스틱 경우로 국한되지 않고 사용할 수 있다(일치성과 점근적 정규성 만족).

5. 확장

5.1 이분산성 존재시, 임의효과 프로빗 모형

이분산이 존재하는 프로빗모형을 일반화하고, $y_j(j=1,...,N)$를 0 혹은 1의 값을 갖는 이진변수로 두자. y_j가 1의 값을 취할 확률은 K개 독립변수 $x_j=(x_{1j},x_{2j},...,x_{kj})$의 선형 조합의 비선형 함수, 즉, $\Pr(y_j=1)=\Phi(x_jb)$로 설정할 수 있다. (Φ는 표준정규분포를 갖는 확률변수의 누적분포함수이며, 0의 평균과 1의 분산을 갖는 정규분포된 확률변수임)

이분산성을 띠는 프로빗모형은 더 이상 1로 고정된 분산을 갖지 않고 독립변수의 함수로써 변하는 정규분포된 누적분포 함수를 일반화함으로써 행할 수 있다. m개 변수 $Z_i = (Z_{1i}, Z_{2i}, \ldots, Z_{mi})$의 배수(multiplicative)함수를 갖는 분산은 Harvey(1976)에 따르면, (5.1) $\sigma_{it}^2 = [\exp(z_{it}\gamma)]^2$가 되므로 모든 독립변수가 1이 될 확률은 다음과 같다. (5.2) $\Pr(y_{it}=1) = \Phi(x_{it}b/\exp(z_{it}\gamma))$.

이진변수 y_{it}가 평균이 $x_{it}b$, 분산이 1인 정규분포된 관측되지 않는 변수w_{it}를 임계치로 함으로써 생성된다 가정하면 다음과 같다.

(5.3) $$y_{it} = \begin{cases} 1 & \text{if} \quad w_{it} > 0 \\ 0 & \text{if} \quad w_{it} \le 0 \end{cases}$$

이 과정을 프로빗 모형으로 나타내면 (5.4) $\Pr(y_{it}=1) = \Pr(w_{it} > 0) = \Phi(x_{it}b)$와 같으며, 이제 w_{it}가 (5.5) $Var(\epsilon_{it}) = \sigma_{it}^2 = [\exp(z_{it}\gamma)]^2$의 분산으로 이분산성을 띤다 가정하자. 이러한 방식으로 프로빗 모형의 이분산성을 완화하면 다음과 같은 배수(multiplicative) 이분산 프로빗 모형을 제시한다.

(5.6) $$\Pr(y_{it}=1) = \Phi\left(\frac{x_{it}b}{\exp(z_{it}\gamma)}\right)$$

이러한 이분산 프로빗 모형의 로그우도함수는 다음과 같고,

(5.7) $$\ln L = \sum_{i \in s}\sum_{t=1} w_{it} \ln \Phi(x_{it}b/\exp(z_{it}\gamma)) + \sum_{i \notin s}\sum_{t=1} w_{it} \ln(1-\Phi(x_{it}b/\exp(z_{it}\gamma)))$$

위의 $\ln L$을 극대화하는 값을 구하는 방식으로 일반화 할 수 있다.[34] 이 경우 한계효과는 매우 복잡해진다.

예를 들어 $y^* = x_{it}'\beta + \alpha_i + \varepsilon_{it}$ 인 경우, 이분산이 존재하면

(5.8) $$pr[y_{it} = 1 | x_{it}] = pr[\alpha_i + \epsilon_{it} > -x_{it}'\beta]$$

---

34) s는 $y_j \neq 0$인 모든 관측치 j의 set이며, w_j는 최적의 가중치를 의미함.

$$(5.9) \quad pr[y_{it}=1|x_{it}] = F[\frac{x_{it}'\beta}{\sqrt{1+\sigma_\alpha^2}}] = F[x_{it}'\delta]$$ 가 되며, 이분산성을 무시하면 β가 아닌 δ를 추정하여 한계효과는 $\beta f(x_{it}'\beta)$가 아닌 $\delta f(x_{it}'\delta)$가 된다. 이분산성을 검증하는 방법은 full 모형과 상수항만 있는 모형을 통해 왈드검정을 행할 수 있고, 이분산성이 있는 full 모형과 이분산성이 없는 full모형을 검증하는 이분산성 우도비검정도 가능하다.

5.2 편향 조정방법(bias correction)

패널 데이타는 관측되지 않는 특성 중에서 시간에 따라 변하지 않는 개체특성을 통제할 수 있다. 그러나 개체의 이질성이 완전히 제약되지 않을 때, 비선형 모형에서는 incidental parameter problem을 겪는다. 이로 인해 고정효과추정치는 편향을 지닐 가능성이 많다. 즉, 고정효과 MLE 방법은 관측되지 않는 개체특성 효과가 표본 추정치로 전이되므로 비선형 모형에서 모수 추정은 개체특성 효과로부터 분리 될 수 없다. 개체 특성 추정시의 오차는 모수 추정 시 편향을 만들게 되는 것이다.

이에 대한 대안적인 접근[35]으로 1. 추정량의 편향 조정법, 2. moment equation의 편향 조정법 3. concentrated likelihood의 편향 조정법을 고려할 수 있다. 또한 직교성에 기초한 접근(Cox and Reid's Adjusted Profile Likelihood Approach 등)과 automatic methods(Panel Jacknife 등)에 의존하는 접근법도 고려할 수 있다. 더 나아가 각 방법들이 특정모형과 데이타 셋에 얼마나 잘 수행하는지에 관해서는 Carro(2004) and Fernandez-val(2005)의 이진변수 모형의 추정치와 Monte-carlo 결과를 참고하라.

---

35) 구체적인 방법에 대해서는 Manuel Arellano and Jinyoung Hahn(2005), "Understanding Bias on Nonlinear Panel models : Some Recent Developments", CEMFI, UCLA. 참조.

5.3 상태의존성(state dependence)과 계열상관(serial correlation)

이진변수모형에서 시차효과는 w_{it}의 serial correlation, α_i의 heterogeneity, $\rho y_{i,t-1}$항의 true state dependence의 세 가지 요인에 의해 발생될 수 있다. 동태적 모형에서는 계열상관과 상태의존성의 구분이 더욱 중요한데, 상태의존성(state dependence)의 경우는 앞서 기술한 동태모형의 방법과 유사하므로, 간략히 재검토 해볼 것이며, 계열상관(serial correlation)이 존재하는 경우에 대해서는 추가적으로 연구가 필요하다.

- 계열상관(serial correlation) 모형

(5.10) $$y_{it} = \gamma' z_i + \beta' x_{it} + \alpha_i + w_{it}$$

$$w_{it} = \rho w_{i,t-1} + u_{it}$$

- 상태의존성(state dependence) 모형

(5.11) $$y_{it} - \rho y_{i,t-1} = \gamma' z_i + \beta' x_{it} + \alpha_i + u_{it}$$

$$y_{it} = \rho y_{i,t-1} + \gamma' z_i + \beta' x_{it} + \alpha_i + u_{it}$$

5.3.1. 상태의존성(state dependence)이 존재하는 경우

(1) 자기회귀 로짓모형 (Autoregressive Logit Model)

개체효과 α_i가 허용되고, 상태의존성(state dependence)가 존재하면 개체의 과거 상태 $y_{i,t-1}$가 현재 상태 y_{it} 예측에 영향을 끼치는 문제가 발생한다. Chamberlain(1978,1985)은 Cox(1958)의 자기회기 로짓 모형을 추가적으

로 분석하였으며, 그 내용은 다음과 같다.

$$(5.12)\qquad \mathrm{Prob}(y_{it}=1|y_{i,t-1})=\frac{\exp(\alpha_i+\gamma y_{i,t-1})}{1+\exp(\alpha_i+\gamma y_{i,t-1})}$$

α_i와 γ의 결합우도함수(joint 우도함수)를 극대화하면 γ는 (fixed T, $N\rightarrow\infty$ 함에 따라) 일치추정량이 되지 못하며, N이 증가함에 따라 추정해야할 모수의 수가 증가하게 되고, 불일치성은 자기회귀모형의 경우 특히 심할 것이다. 때문에 Chamberlain은 조건부 우도함수를 제안한다. α_i에 대한 충분 통계량은 $\Sigma_t y_{it}$와 y_{iT}이며, y_{i1}에 의존하는 초기조건을 다루면,

$$(5.13)\quad \Pr(y_{i1},y_{i2},...,y_{iT}|y_{i1},\Sigma_t y_{it},y_{iT})=\frac{\exp(\gamma\sum_{t=2}^{T}y_{it},y_{i,t-1})}{\sum_{d\in B_i}\exp(\gamma\sum_{t=2}^{T}d_t,d_{t-1})}$$

$$(B_i=[d=(d_1,d_2,...,d_T)|d_t=1 \text{ or } d_1=y_{i1},\sum_t d_t=\sum_t y_{it},d_T=y_{iT}]\text{임})$$

개체들에 대한 y값들의 합으로 주어진 개체 관측치의 조건부 확률은 모든 개체들 값의 합으로($y_{i1},y_{iT},\Sigma_t y_{it}$와 같은) 나뉘어진다. 이 모형은 조건부 로짓 모형의 형태이고 α_i가 지워지므로 쉽게 추정될 수 있다.[36]

(5.13)의 자기회귀 모형에는 두 가지 한계점이 존재한다. 첫 번째는 조건부 로짓 방법의 사용은 Corcoran and Hill(1985)에서 제시된 바와 같이 표본 관측치의 많은 부분을 버리고, 데이터의 매우 작은 부분을 사용할 것이라는 것이다. 두 번째 문제는 이 모형이 외생변수를 허용하지 않으므로 정책 분석을 위한 사용에 제한적이라는 것이다. 즉, 개체 특성 효과 α_i를 허용한 후 y_t가 y_{t-1}에 의존하는지 아닌지 만을 말해줄 뿐이다. 이러한 문제들은 임의

36) Chamberlain ann Hill(1985), Corcoran(1982), Corcoran and Hill(1985)등에 의해 노동시장 참가 연구에 사용됨.

효과 프로빗모형의 수행으로 해결될 수 있으며, 대안적인 모형은 외생변수를 포함하는 Chamberlain 모형에 의해 다음과 같다.

$$(5.14)\qquad \Pr(y_{it}=1|x_{it},y_{i,t-1})=\frac{\exp(\alpha_i)+\gamma y_{i,t-1}}{\exp(\alpha_i)+\exp(-\beta' x_{it})}$$

여기서 γ=0이면 이는 상태의존성(state dependence)이 없는 고정효과 로짓 모형과 같다. chamberlain은 가설 γ=0를 검증하는 것이 고정효과 모형의 적절성을 판단하게 해준다고 주장한다. 동태 모형에서 시간에 따라 변하는 이질적인 설명된 변수 x_{it}의 영향은 임의효과 프로빗 모형에 의해 해결될 수 있다.

(2) 자기회귀 프로빗모형(Autoregressive Probit Model)

상태의존성(state dependence)이나 종속변수(y)의 시차변수를 포함하는 경우를 고려한 임의효과 프로빗 모형을 일반화 할 수 있다. 추가적인 문제는 초기조건의 문제로 y_{i0}를 고정된 값으로 처리하는 ML 추정법은 문제에 직면하게되므로 y_{i0}의 분포를 정의하는 것이 중요한데, 가장 좋은 것은 모형 자체에서 직접 유도한 것이다. 좀 더 간편한 방법은 y_{i0}를 확률변수(random variable)로 가정하는 것이다. [37]

5.3.2. 자기상관이 존재하는 경우

비선형 이진변수 모형(예를 들면 프로빗 모형) 추정 시 잔차 간에 계열상관이 존재하는 경우 프로빗 모형 표준오차의 추정치는 불일치성을 야기한다.

---

37) $prob(y_{i0}=1)=F(x_{i0}'\delta)$, δ 는 추정되는 미지의 모수

오차 간에 계열상관이 존재하는 경우 자기상관된 오차항들을 통해 모형화 할 수 있으며, 선형의 경우와 유사하게 자기회귀(AR)방법을 사용하여 추정한다.

$y_{it} = X_{it}'\beta + e_{it}$

$e_{it} = \rho_1 e_{it-1} + \dots + \rho_l e_{it-l} + u_{it}$ (u_{it}는 백색잡음 N(0,1)임)

또한 Estrella and Rodrigues(1998)에 따르면 HAC(Heteroskedasticity and autocorrelation-corrected)공분산 행렬뿐만 아니라 통상적인 프로빗 준우도함수(Quasi Maximum Likelihood)의 추정치 β를 추정할 수 있다. ($F(\cdot)$와 $f(\cdot)$를 각각 누적정규분포와 표준정규분포로 놓자)

(5.15) $$\log L = \sum_{t=1}^{T} y_{it} F_t + (1 - y_{it})(1 - F_t)$$

이 우도함수를 극대화하면 QML으로서의 통상적인 프로빗 모형 결과를 얻을 수 있고 β의 1계조건은 $\sum_{t=1}^{T} h_t = 0$ 이다. ($h_t = \dfrac{[y_{it} - F(\beta' X_t)] f(\beta' X_t) X_t}{F(\beta' X_t)[1 - F(\beta' X_t)]}$ 임)

β의 결과추정량 $\hat{\beta}^{QML}$은 표준적인 프로빗 추정치이고 계열상관이 존재하는 경우 일치성을 지닌다.

6. 응용 예

6.1 노 동

6.1.1 "A life cycle 모형 of female labour supply", James J. Heckman , Tomas E. Macurdy, 1980.

기혼여성 노동공급 결정의 life cycle 모형을 실증적으로 제시한 이 논문은 두 가지 측면의 노동공급 활동을- 연(年) 노동시간과 연(年) 노동시장 참가- 이시적 틀에서 분석하였다. 이 논문은 신용제약이 없는 완전한 확실성 하 의사결정에서 “항상”, “일시적” 소득과 임금에 반응하는 노동공급의 의미와 측정을 고려한다. 그리고 항상소득 가설과 일치하는 가설을 제시 한다. 일시적 소득에 반응하는 기혼 여성의 증거는 찾아 볼 수 없었다.

고정효과 Tobit모형을 사용하였으며, 고정효과는 바람직한 경제적 해석과 노동공급 결정하는 누락된 생애 변수를 대체하여 준다. Michigan Panel Survey of Income Dynamics로 부터의 8차년도 데이타를 이용하여 분석하였다. 전술한 결과를 제외하고, 이론 예측과 일치하는 실증 분석 결과들을 발견할 수 있었는데, 노동 공급은 생애 부와 역의 관계에 있으며, 자녀는 생애 노동공급에 영향을 미친다.

6.1.2. “자격취득의 결정요인 및 취업 · 임금효과”, 김안국 · 강순희, 2004.

이 논문에서는 한국노동패널 1~4차년도 자료를 개인별로 묶어 고정효과 로짓모형, 임의효과 모형으로 회귀분석 하였다. 관측되지 않는 개체특성을 통제할 수 있는 Chamberlain(1980)의 고정효과 로짓모형으로 자격 취득의 요인과 자격의 취업효과를 추정하였다. 그리고 임금에 미치는 자격 효과를 보기 위해 임금함수를 회귀분석하였는데, 임의효과 모형으로 관측되지 않는 개인별 특성을 통제하면서 자격의 임금효과를 추정하였다.

추정 결과 고졸 미만 저학력층의 자격 취득 확률이 높게 나타났다. 남자는 실업상태에 있는 경우에 자격 취득이 많았고, 여성은 비경제활동의 기타

상태에 있는 경우 자격취득이 많았다. 취득한 자격의 임금효과는 남녀 모두 유의하지 않은 것으로 나타났다.

6.1.3. "한국여성의 취업과 비취업 선택 : 5년간의 경험", 김종숙, 한국여성개발원(2005).

한국노동연구원의 “노동 패널” 자료를 이용하여 여성 근로자의 취업과 비취업상태 전환을 패널자료로 분석하고 노동시장 유출입을 빈번한게 경험하는 여성 근로자가 취업할 확률은 어떠한 요인에 의하여 결정되는가를 살펴보았다. 추정 모형으로는 Chamberlain(1980)이 사용한 고정효과 로짓모형을 이용하였다.

추정 결과는 개체 효과가 없는 로짓 모형과 개체 특성을 통제한 후 얻을 수 있는 고정효과(고정효과) 로짓 모형의 분석으로 나누어 보였다. 결과 개인의 연령이 증가할 수록 취업할 확률이 높아지고 가구원 수가 많을 수록 취업할 확률이 높아진다. 산업별로는 도소매음식숙박업의 취업할 확률이 타 산업에 비하여 높은 것으로 나타나며, 직종별로는 전문가에 비하여 사무직, 판매직, 단순노무직 근로 모두가 (+)의 유의한 효과를 보였다. 고정효과 모형을 일반적인 모형과 비교하여 보면 산업더미에서 양자 간 결과에 차이가 있음이 나타난다.

6.1.4. “교육훈련의 경제적 효과 - 임금근로자를 중심으로”, 김안국(2004).

본 논문은 교육훈련의 경제적 효과를 분석하기 위해 한국 노동패널의 개

인별 패널 자료를 이용하여 분석하였다. 교육훈련 이수 가능성을 위해 임의효과 프로빗 모형을 이용하여 교육훈련 이수확률을 추정해 본 결과 남자의 경우 여자보다 교육훈련 이수확률이 높았고, 대기업의 경우 교육훈련을 이수 받을 확률이 더 높았다. 임금효과를 임의효과 모형으로 추정한 결과 교육훈련을 이수한 경우 (+)의 유의한 임금효과를 가지며, 그 크기는 근속년수의 두 배 정도가 되는 것으로 나타났다. 이하 내용은 본 논문에서 제시하고 있는 기존연구 검토 중 일부이다.

연구자	년도	자료	계량모형	주요 발견
Lillard and Tam	1992	CPS, NLS, NLSY	프로빗	교육훈련은 실업의 확률을 낮춤. 교육훈련 효과는 양의 임금 증가로 나타남.
Lynch	1992	NLSY	프로빗 고정효과 모형	이전 직장에서의 집체훈련이 임금에 유의한 효과를 보임
Loewenstein and Spletzer	1997	CPS, NLSY	프로빗	교육훈련의 기업 특수성에 따라 임금효과가 차이남.

6.2 금융

6.2.1 “Ownership of Stock and Mutual Fund; A panel data Analysis”, Rob Alessie, Stefan Hochguertel, Arthur van Soest.

대표적인 panel survey를 이용하여 분석한 이 논문은 위험 금융자산으로서 무츄얼 펀드와 개인 주식에 초점을 맞추었다. 일변량 동태 이진변수모형뿐만 아니라 두 유형의 자산 간의 상호영향을 고려하기 위해 다변량 모형도 고려하였다. 모형은 CentER Saving Survey의 1993-1998 웨이브로 부터 추정되었으며, 두 유형의 자산 모두에서 중요한 역할을 하는 관측되지 않는 헤테로와 상태의존성(state dependence)을 발견하였다. 1기의 무츄얼 펀드의 소유와 그 다음기의 개인 주식 소유와 관계는 대체로 양의 상관관계를 보였으며, 이는 관찰되지 않는 헤테로로 설명된다. 두 binary choice 방정식에서 가구 특성 효과 간의 상관관계를 계산하면, 무츄얼 펀드 소유에 대한 과거 주식 소유의 효과는 음의 상관관계임을 알수 있다.

자산 포트폴리오의 동태적 소유 구조를 설명하기 위해 동태적 이진변수 패널 데이타를 사용하였으며, 관찰되지 않는 헤테로와 상태의존성(state dependence) 간의 구별을 가능하게 하는 일변량 임의효과 패널 모형으로 분석하였다. 또한 두 자산 간의 상호 관계를 계산하기 위해 이변량의 경우로 확장을 하였다.

6.2.2 "개인연금 가입 결정 및 가입상태 변화 분석", 전승훈, 임병인, 강성호 (2006).

가계의 저축 목적, 구체적인 저축 수단 등에 대한 조사가 이루어진「한국노동패널 」4-7차년도 개인 및 가구 자료에 binomial 로짓분석을 적용하여, 가구의 개인 연금 가입결정요인 및 가입상태 변화 요인을 분석하였다. 분석 결과 국민연금 가입자 일수록, 저축성 보험 납입액이 많을 수록, 경제

활동상태가 취업인 상태로 지속될수록, 저축목적이 노후 대비일수록, 개인연금의 신규 가입 또는 유지가능성이 높아지나, 가구주 연령이 높을수록 중도탈퇴 또는 가입하지 않을 가능성이 높아진다는 결과를 얻을 수 있었다.

이러한 분석결과는 '자발적인 노후대비의 중요성'에 대한 교육 및 홍보강화, 노후대비를 위한 공적 · 사적 저축에 대한 유인정책, 중장년층의 개인연금 가입에 대한 유인강화와 중장년층의 욕구를 충족시킬 수 있는 연금상품 개발이 요구된다는 시사점을 얻을 수 있다. 이렇듯 패널 데이타분석은 사회, 금융을 포함한 제 분야의 정책적 시사점을 제공하며, 민간과 국가 차원의 정책적인 접근을 가능하게 해준다.

6.2.3 "Towards an Explanation of Household Portfolio Choice Heterogeneity : Nonfinancial Income and Participation Cost Structure", Vissing-Jorgensen(2002).

과거 주식시장 참가와 현재의 주식시장 참가의 유의성이 고정, 가변 거래 비용 때문이라기보다 관측되지 않는 개체 특성 때문인지를 알아보기 위해, Heckman (1981) 동태 프로빗 추정량의 일반화된 버전을 사용한 주식시장 참가의 프로빗 모형을 고려하였다. 이 추정은 전기의 참가 결정 여부, 전기의 주식 금액, 현재의 참가 결정 여부에 관한 방정식을 정의하며, 그 결과 과거 주식 투자 금액은 현재의- 참가 결정에 (-) 효과를 가진다. 이는 이전에 알려진 (+) 효과가 허구적이며, 관측되지 않는 개체 효과 때문임을 나타낸다. 데이타는 Panel Study of Income Dynamics를 이용하였다. (N=1081)

6.2.4. "세대(Cohort)별 자산- 연령 프로파일 분석", 남준우(2005).

본 논문은 아직도 논쟁이 진행 중인 자산- 연령의 프로파일을 추정한다. 이를 위해 한국 노동연구원의 한국노동패널 1999-2004년 자료를 이용하며 우리나라의 경우 노년층에 있어 실제로 자산의 축적이 이루어지고 있는지 아니면 자산의 처분이 이루어지고 있는지를 준모수(semiparametric) 추정방법을 통해 분석하였다. 또한 노년층의 자산 축적/처분율을 준모수 추정방법을 통해 추정하였다. 주어진 자료의 극심한 비대칭성을 고려하여 90% 절삭된(trimmed) 표본을 구하였고 횡단면 분석에서 나타나는 생산성 효과로 인한 편향을 고려하여 세대별 자료를 구축하였다.

세대별 자료의 추정결과 대체로 젊은 세대의 경우 자산은 연령에 따라 꾸준히 증가하나 고령층인 광복이전 세대의 경우에는 은퇴시기를 맞아 자산이 감소하는 전형적인 일생주기가설과 부합하나 일정 연령을 지나면 다시 증가하는 형태를 이룬다. 세대별 비교에 있어서는 젊은 세대는 선배세대에 비해 확실히 생산성이 증가하는 생산성 효과가 작용하고 있었으며, 고령층에 대해 자산의 축적/처분율을 계산한 결과 자산의 처분은 65-74세의 경우 0.6 ~ 0.7% 처분율을 가지며 전반적으로 고령층에서는 3~5% 정도의 자산 축적을 이루고 있었다.

6.3 환경

6.3.1. “에너지절약투입자금에 대한 경제적 성과 분석”, 유동헌(2005), 에너지경제연구원

에너지절약 투자액이 에너지원단위를 개선시키는 효과가 있는지를 실증

분석하기 위해 에너지 절약설비 투자 자료가 이용 가능한 1982년부터 2004까지의 데이터를 사용하였다. 에너지절약 투자 자료는 한국산업은행 자료를 이용하고, 부가가치자료는 한국은행의 업종별 부가가치 2000년 가격 자료를 사용하고 있으며, 업종별 에너지 소비 자료는 에너지경제연구원 자료를 사용하였다.

패널자료에 대한 추정방법으로 고정효과(고정효과) 모형과 랜덤효과(임의효과) 모형을 채택하고, 고정효과모형과 랜덤효과모형 중 어느 모형을 선택하느냐는 Hausman 검증법을 적용하여 결정하였다. 추정결과에 의하면, 에너지절약투자는 에너지원단위를 개선시키는 효과가 있는 것으로 추정되었으나, 에너지절약 절대량 측면에서 그 효과는 그리 크지 않는 것으로 나타났다.

6.4 정 책

6.4.1. “법인세 부담이 기업의 투자활동에 미치는 효과 분석”, 김우철(2006), 한국조세연구원

본 논문은 한국신용평가정보의 개별기업에 관한 미시자료(Kis-Value)를 바탕으로 법인세 부담의 지표 중 하나인 평균유효세율을 측정하고, 이를 투자수요 모형에 결합하여 법인의 세부담이 투자에 미치는 효과를 분석했다. 개별 기업 자료의 속성과 여러 기간에 걸쳐 시행 또는 조정되는 투자의 특성으로 인해, 투자함수는 전형적인 동태 패널모형의 형태로 주어지며, 이의 식별과 추정을 위해서 Arellano & Bond(1991)와 Blundell & Bond(1998)에 의해 개발된 GMM을 이용하였다.

실증분석 결과 대체로 경제이론의 예측과 부합된 결과를 보여주고 있는데, 전기의 투자비중의 계수 추정치는 통계적으로 유의한 (+)의 효과, 매출액 증가율과 현금흐름비중은 투자에 유의미한 (+)의 효과를 미치는 것으로 나타났다. 법인세부담의 기업투자효과는 (-)의 방향으로 나타나났으나, 그 효과의 크기는 통계적으로 유의미하지 않았다. 부분적으로는 통계적으로 유의한 경우가 있었으나, 계수추정치의 값이 매우 낮은 수준이어서 본 연구에서 제시하는 세부담 인하의 투자증대효과는 일반적인 기대보다 훨씬 작았다고 결론짓는다.

7. 요약 및 결론

패널 데이타분야의 계량경제학 발전으로 패널 데이타가 적용 가능한 광범위한 분야의 경제 모형의 확장을 이끌었으며, 정성적인 분석을 가능하게 하는 이산형 변수를 사용한 분석 역시 널리 사용된다. 본 논문은 종속 변수가 이진변수인 패널 데이타 모형분석 시 발생하는 몇 가지 사안들을 검토하여, 실증 연구자가 적절한 모형을 비교, 검토할 때 도움이 될 수 있도록 하였다.

우선, 이산형 종속변수 모형의 기반이 되는 이진변수 모형과 패널 데이타의 개념을 검토하였다. 패널 데이타 이진변수 모형은 정태적 패널과 동태적 패널로 나누어 살펴볼 수 있는데, 먼저 정태적 패널 데이타 모형을 고려하는데 있어서는 전통적인 패널 데이타 모형으로서 고정효과 모형과 임의효과 모형의 추정 방법 및 추정 시 발생할 수 있는 incidental parameter problem등의 제 문제점을 살펴보았다. 추정 방법으로는 오차항의 분포를 가정하는 모수적인 방법 뿐 만 아니라 오차항의 분포를 가정하지 않는 준모수

적 접근법도 고려하였다.

다음으로는 동태적 패널 데이타 모형의 일반적인 모형, 초기조건의 문제와 상태의존성(state dependence)의 제 문제점들을 살펴보았다. 그리고 횡단면 자료 분석 시 발생될 수 있는 이분산성과 시계열 자료에서 발생될 수 있는 계열 상관의 문제점을 고려하는 모형으로 확장하였으며, 마지막으로 이러한 이론적 측면을 반영한 실증 분석 사례와 현재까지 진행 중인 이진변수 패널 모형의 연구 현황을 간략하게 정리하였다.

1. 고정효과모형

개체효과(effect)와 회귀변수(regressor) 간의 상관관계를 허용하는 소규모표본 존재(small T)고정효과 로짓 모형에서 MLE는 추정해야 하는 모수들이 많은 incidental parameter problem(small T bias)을 발생시키며, 그 경우 MLE는 일치추정량이 되지 않는다.

반면, 개체효과 α_i의 충분통계량을 이용하여 로짓 모형에 적용한 조건부 MLE는 일치추정량이 된다. (small T)고정효과 프로빗 모형에서는 고정효과가 없어지지 않으므로, 조건부 MLE를 사용할 경우, 계산이 단순하지 않으며 일치추정량을 제시하지 않는다. 그러므로 이진변수 모형에 고정효과를 적용하는 경우에는 로짓 모형을 선택하는 것이 좋다.

2. 임의효과모형

개체효과(effect)와 회귀변수(regressor) 간의 독립을 가정하는 임의효과모형에서 MLE를 사용할 경우 일치추정량을 얻을 수 있으나, 수리적적분

(numerical integration)문제에 직면한다. 이 경우, Butler and Moffit 방법과 같은 효율적인 계산절차를 통해 추정할 수 있으며, Maximum simulated likelihood(MSL)를 활용하는 방법도 있다. 또 다른 추정법으로는 GMM방법을 이용할 수 있다. 그러나 개체효과(effect)와 회귀변수(regressor) 간의 상관관계가 있는 경우, 이전과 방식을 고려하면 편향을 제거하지 못할 것이다. 이 경우에는 Chamberlain의 정의를 이용하여 일치성을 만족하고, 효율적인 모수추정량을 얻을 수 있다.

3. 준모수적 접근법

이 접근법은 오차항의 분포를 가정하지 않을 뿐만 아니라 개체효과가 고정효과이든 임의효과이든 일치추정량을 제시한다. Manski의 maximum score estimator는 N이 무한대로 감에 따라 일치성을 만족하지만, 이 추정량은 $\sqrt{N}$ 일치성을 만족하지 않는다.

4. 동태적 패널

동태 고정효과모형의 (fixed T, $N\rightarrow\infty$) MLE는 일치추정량을 제시하지 않으며, 임의효과 모형은 개체효과와 초기관측치의 확률분포에 의존하여 일치성이 결정된다. 반면 준모수적(semiparemetric)방법은 동태 패널에서 사용될 수 없다. 그러나 Honore and Kyriazidou는 로짓모형의 조건부 확률 접근법을 일반화하여 $\sqrt{N}$보다는 느리지만, 일치성을 만족하는 추정량을 제시한다. 또한 maximum score접근법을 일반화하여 일치성을 지니지만, 점근적 분포가 정규분포가 아닌 추정량을 제시한다

5. 확장

이분산성 존재 시에는 우선 $Var(\epsilon_i) = [\exp(\gamma' z_i)]^2$를 모수화 한다. 그리고 다음과 같은 이분산 프로빗모형 $\mathrm{Prob}[y_i = 1] = \Phi[\frac{\beta' x}{\exp(\gamma' z)}]$의 로그우도함수를 극대화하는 값을 구하는 방식으로 일반화 한다. 이 경우 partial effect는 매우 복잡해진다. 시차종속변수가 설명변수로서 포함되는 경우는 위의 1과 2의 확장으로 조건부 로짓 모형이 이 경우로 확장될 수 있으나, 조건부 로짓 모형은 제한적인 가정 하에서만 확장될 수 있다. 반면 임의효과 프로빗 모형은 이 경우로 쉽게 확장 될 수 있다. 상태의존성(state dependence)존재 시 자기회귀 로짓 모형과 프로빗 모형을 사용할 수 있다.

< 표 1 > binary choice model using panel data: survey

<table>
<tr><th>분야</th><th>년도</th><th>연구자</th><th>세부내용</th></tr>
<tr><td rowspan="9">QR모형 및 panel 전반</td><td>1981</td><td>Amemiya</td><td>Qualitative response model</td></tr>
<tr><td>1987</td><td>Maddala</td><td>Limited dependent variable model</td></tr>
<tr><td>1991</td><td>Arellano, Stephen Bond</td><td>specificaion test방법</td></tr>
<tr><td>2001</td><td>Yoshitsugu Kitazawa</td><td>패널 전반</td></tr>
<tr><td>1994</td><td>McFadden</td><td>추정법</td></tr>
<tr><td>2002</td><td>Honore</td><td>비선형 패널</td></tr>
<tr><td>2003</td><td>Hsiao</td><td>discrete choice</td></tr>
<tr><td>2003</td><td>Arellano</td><td>ML, MML, GMM</td></tr>
<tr><td>2003;2006</td><td>Green</td><td>비선형 패널 전반</td></tr>
<tr><td rowspan="5">binary choice 패널 데이타 전반</td><td>1944;1955;1957; 1980</td><td>Berson</td><td rowspan="5">estimation 및 specification 전반</td></tr>
<tr><td>1949</td><td>Neyman</td></tr>
<tr><td>1958</td><td>Fergson</td></tr>
<tr><td>1974;1976;1980b</td><td>Amemiya</td></tr>
<tr><td>1984</td><td>Yatchew and Griliches</td></tr>
</table>

			Chamberlain	
		1999	Arellano and Honore	
		2001	Horowitz, Savin	
이분산		1976	Harvey	이분산
binary choice 패널 데이타	fixed effect conditional likelihood	1973	Anderson	conditional likelihood
		1980	Chamberlain	
		1987	Cox and Reid	
		2000	Honore and Kyrizidou	fixed effect logit, conditional likelihood
		2001	Thiery Magnac, INRA	fixed effect conditional logit
		2001	Lancaster	conditional likelihood
		2002	Green	nonlinear fixed effect
		2003	Arellano and Carrasco	GMM
		2006	Stock and Watson	Heteroskedasticity-robust standard error
	incidental parameter problem	1948	Neyman and Scott	-
		2001	Arellano	
	semiparametric	1987	Manski	Maximum score

	estimation			estimation
		1988	P. M. Robinson	semiparanetric 전반
		1992 1997	Horowitz Charlier,Erwin	smooth maximum score
		1994	Powell, J. L	semiparametric
		1997	Chunrong Ai	MLE
		1999	Lee, Myong-jae	root-N consistent estimation
		2002	Honore and Lewbel	-
	nonparametric estimation	1999	Chen, Heckman and Vytlacil	nonparametric
		1997	Laisney and Pohlmeier	Semi-Nonparametric
	random effect	1980;1984	Chamberlain	specification 전반
		1981a	Heckman	numerical integration
		1982	Butler and Moffit	
		1983	Avery-Hansen-Hotz	GMM
		1984	Chamberlain	numerical integration
			Heckman and Wills	random effect probit

		1998	Bertschek and Lechner	GMM
		2000	Inkmann	
	dynamic model	1951	Bates and Neyman	-
		1977	Hecman and Wills	
		1978b	Chamberlain	state dependence
		1978	Layton	Dynamic
		1999	Hyslop	state dependence, serial correlation, Heterogeneity
		2000	Honore and Kyriazidou	lagged dependence variable 포함한 discrete패널
		2001	Chintagunta, Kyriazidou and Perktold	Dynamic
		2001	Hahnm Jinyong	dynamic fixed effect logit
		2006	Jesus. M. Carro	fixed effect, MMLE
	bias correction	2004	Carro	bias correction

		2005	Fernandez-val	bias correction, fixed effect binary choice panel
		2005	Arellano, Jinyoung Hahn	Bias correction
application		1980	Heckman	fixed effect Tobit model
		2001	Rob Alessie, Stefan Hochguertel, Arthur van Soest	dynamic binary choice, dinamic multivariate choice
		2001	Thiery Magnag	binary fixed logit
		2002	Vissing-Jorgensen	dynamic probit
		2004	김안국, 강순희	fixed effect logit, random effect
		2005 2006 2006	남준우 전승훈, 임병인, 강성호 김우철	semi parametric. binomial logit model. Dynamic, GMM.

참고문헌

국내문헌

김안국 · 강순희, (2004) : "자격취득의 결정요인 및 취업 · 임금효과", 노동 경제 논집, 제 27권(1), 한국노동경제학회.

김안국(2004) : “교육훈련의 경제적 효과 - 임금근로자를 중심으로”, 한국노동경제학회.

김우철(2006) : “법인세 부담이 기업의 투자활동에 미치는 효과분석”, 한국조세연구원.

남준우(2005) : “ 세대 (Cohort)별 자산- 연령 프로파일 분석”, 서강대학교.

백화종, 김안나(2004) : “빈곤 및 공공부조 패널 데이터 구축 방안”, 한국보건사회 연구원,

유동헌(2005) : “에너지절약투입자금에 대한 경제적 성과 분석”, 에너지경제연구원.

전승훈, 임병인, 강성호 (2006) : “개인연금 가입 결정 및 가입상태 변화 분석”, 보험개발원.

외국문헌

Andrew Berg and Rebecca N. Coke (2004) : " Autocorrelation-Corrected Standard Errors in Panel Probit: An Application to Currency Crisis Prediction", *IMF Working Paper* WP/04/39.

Amemiya, T. (1981) : "Qualitative Response Model: A survey", *Journal of Economic Literature*,Vol.19, No.4, pp.1483-1536.

Bo E. Honore (2002) : Non-linear Model with panel data, *Cemmap Working paper* CWP13/02.

Bo E. Honore and Ekaterini Kyriazidou(2000) : "panel data Discrete Choice Models with Lagged Dependence Variables", *Econometrica*, Vol.68, No.4, pp.839-874.

Charles F. Manski(1987) : "Semiparametric Analysis of Random Effects Linear Models from Binary Panel Data", *Econometrica*, Vol.55, No.2, pp.357-362.

Cheng Hsiao(2003) : Analysis of panel data , Cambridge, second edition ch.15 logit and probit models.

Christian Gourieroux (2000) : Econometrics of Qualitative Dependent Variable, Cambridge university press. ch.4.

Dean R. Hyslop(1999) : "State Dependence, Serial Correlation and Heterogeneity in Intertemporal Force Participation of Married Woman", *Econometica*, Vol.67, No.6, pp.1255-1294.

Cristian Abmann(2006) : "Determinants and Costs of Current Account Reversals under Heterogeneity and Serial Correlatiom".

Gary Chamberlain(1980) : "Analysis of Covariance with Qualitative panel", *The Review of Economic Studies*, Vol.47, No.1, Econometrics Issue, pp.225-238.

G. S. Maddala(1987) : "Limited Dependence Variable Models Using panel data", *The Journal of Human Resources*, Vol.22, No.3, pp. 307-338.

Ivan Fernandez-Val(2005) : "Estimation of Structure Parameters and Marginal Effects in Binary Choice panel data Models with Fixed Effects", *MIT JOB MARKET PAPER*.

James H. Stock and Mark W, Watson(2003) : Introduction to Econometrics, Addison Wesley, ch.9 Regression with a Binary Dependent Variable.

James J. Heckman and Tomas E. Macurdy (1980) : "A Life Cycle Model of Female Labour Supply", *The Review of Economic Studies*, Vol.47, No.1, Econometrics Issue, pp.47-74.

Jesus M. Carro(2006) : "Estimating dynamic panel data discrete choice models with fixed effect", *Journal of Econometrics*.

Jeffrey M. Wooldridge(2002) : " Econometric Analysis of Cross Section and panel data", The MIT Press.

Joel L. Horiwitz; N.E.Savin(2001) : " Binary Response Models: Logits, Probits and Semiparametrics", *The Journal of Economic Perspectives*, Vol.15, No.4, pp. 43-56.

Kenneth Y. Chay and Dean R. Hyslop(2000) : "Identification and Estimation of Dynamic Binary Response Panel Data Models : Empirical Evidence using Alternative Approachs".

Laszlo Matyas and Patrick Sevestre kluwer academic publishers : The econometics of panel data , a handbook of the theory with application ch.7 discrete data.

Manuel Arellano, Stephen Bond(1991) : "Some Test of Specification for the panel data", *The Review of Economic Studies*, Vol.58, No.2, pp.277-297.

Manuel Arellano and Bo E. Honore(2000) : “panel data Models "Some Recent Developments”, *Working paper* No. 0016.

Manuel Arellano(2003) : “Discrete Choice with panel data”, Investigaciones economicas, Vol. 27, pp.423-458.

Manuel Arellano, Jinyoung Hahn(2005) : “Understanding Bias on Nonlinear Panel Models : Some Recent Developments”, *CEMFI*, UCLA.

Michael Lechner(2003) ; "Topics in Econometrics of Nonlinear Panel data Models", CIDE, Berino.

Rob Alessie, Stefan Hochguertel, Arthur van Soest (2001) : “Ownership of Stock and Mutual Fund; A panel data Analysis”.

Vissing-Jorgensen(2002) : “Towards an Explanation of Household Portfolio Choice Heterogeneity : Nonfinancial Income and Participation Cost Structure"..

William H.Green(2002) : "The Behavior of the Fixed Effects Estimator in Nonlinear Models", Stern School of Business, New york University.

__________________(2003) : Econometric Analysis, New york University, fifth edition.

__________________(2006) : Econometric Analysis of panel data, Stern S School of Business Spring, ch.15,16,17,21.

Yoshitsugu Kitazawa : "Recent Denelopment in panel data Econometrics", Kyushu Sangyo University.

ABSTRACT

Binary Choice Model using Panel Data : A survey

Kim, Joo-Bong
Department of Economics
Graduate School of
Sungshin Women's University

This paper present a survey of the method used in the estimation binary choice model with panel data. It first review some issues in the analysis of panel data when the dependent variables are discrete. This paper reviews the existing approaches to deal with panel data binary choice models with individual effects and consider the estimation of static, dynamic binary choice panel data.

The problems of fixed effects vs. random effects and heteroscedasticity vs. serial correlation are discussed with reference to binary choice model. The paper then discuss these problems with reference to the panel logit, panel probit, and panel semiparametric models. Finally, the paper presents a comparative assesment of these models and illustrate the empirical research.