

한국의 세대내 소득이동성 측정⁽¹⁾

오 주 현 · 김 봉 근

본 논문은 한국의 세대내 소득이동성(earnings mobility)을 추정한다. 패널자료의 문제점인 비무작위적 탈락(non-random attrition)과 측정오차(measurement error)를 보정하기 위해 가패널 방법론(pseudo-panel approach)을 사용하고, 통상패널자료의 추정편의를 빈곤함정(poverty trap)에 적용해 논의한다.

주제어: 소득이동성, 표본이탈, 가패널, 빈곤함정

1. 서론

소득이동성은 연속하는 두 기간으로부터의 상대적인 소득 순위에 대한 변화율로 정의하고, 낮은 소득이동성은 사회·경제적 상하이동의 감소를 의미한다. 소득이동성은 횡단면적인 소득불평등과 더불어 한 사회의 소득분포 건전도에 대한 지표로 사용한다. 예를 들어, 특정시점에서 관찰되는 빈곤보다 낮은 소득이동성과 결합된 빈곤은 장기빈곤 가능성을 추가적으로 시사한다.

경제위기 이후 소득분배 격차가 확대되고 있는 것으로 나타나고 있다. 경제위기 때 따른 고용변화를 살펴본 김봉근·윤상철(2010)은 외환위기 이후 강조되어온 취약계층의 고용안정성악화가 글로벌 경제위기 이후 더욱 확대될 수 있다는 시사점을 제시하고 있다. 구체적으로 남성에게 비해 여성이, 고학력 계층에 비해 저학력 계층의 고용불안정성 등이 상대적으로 커지고 있으며, 대기업중심의 수출산업보다 취약계층의 비중이 상대적으로 큰 서비스산업의 고용안정성에 더 큰 악영향을 유발했음을 보고하고 있다.

취약계층의 고용안정성악화 및 소득불평등의 심화는 낮은 소득이동성 및 이동성의 감소와 연계되어 빈곤의 고착화, 그리고 세대간 세습으로 이어질 수 있다. 소득이동성에 대한 연구가 중요한 이유 중 또 다른 하나는 저소득층의 소득이동성, 특히 빈곤으로부터 빠져나오지 못하는 빈곤함정의 존재 여부이다. Antman and Mckenzie(2007)는 저소득층의 상방 소득이동성이 매우 낮을 경우, 저소득층은 빈곤함정에 빠질 위험이 있기 때문에 이에 대한

(1) 본 논문은 오주현(2009)을 확대 발전시킨 연구임.

정책대응이 필요하다고 주장하였다. 장기간의 시간에 걸쳐 빈곤선 아래 머무르는 장기빈곤 또는 빈곤의 고착화 가능성을 낮추기 위해 소득이동성의 연구는 매우 중요하다.

소득이동성연구는 크게 세대내와 세대간 소득이동성 연구로 나눌 수 있다. 먼저, 세대간 소득탄력성은 한 세대에서 그다음 세대로의 상대소득계층의 연계성을 나타내는 장기지표로, 횡단면적인 소득 불평등도와 더불어 한 사회의 장기소득분포의 건전성 여부를 직접적으로 나타내는 중요한 지표이다. 세대간 소득이동성이 낮다 또는 세대간 소득상관계수가 크다는 것은 예를 들어 빈곤이 상당한 정도로 세습되는 것으로 한 개인의 소득분포상의 낮은 위치가 부모세대의 낮은 상대소득수준에 의해 사전적으로 상당 부분 결정된다는 것이다. 계층의 세습 및 이로 인한 저소득층의 비관주의는 인적자본투자를 저해하고, 계층간 위화감의 고조 및 이로 인한 범죄율을 높이는 등 전반적인 사회의 갈등수준을 높일 것이다. 세대간 낮은 소득이동성은 낮은 세대내 소득이동성, 특히 빈곤선 아래에 위치한 저소득층은 상방 소득이동성의 부재 (빈곤함정의 존재) 여부와도 밀접하게 연계된다. 계층간 소득이동성의 단절은 저소득층의 낮은 교육투자 및 계층내 결혼 등의 요소가 결합될 경우 세대간 소득이동성을 낮추거나, 낮은 세대간 소득이동성 수준을 고착화시킬 것이다. 세대간 소득이동성은 지표의 중요도에도 불구하고, 부모와 자녀세대의 소득을 직접적으로 연계할 실증자료의 특성상 패널자료의 장기간 축적을 필요로 하고, 패널자료의 형성이 뒤쳐진 우리나라에서는 관련 지표의 제한적인 추정만 이루어져 왔다. 부모와 자녀 간의 교육수준의 세대간 연계성에 관한 안종범·전승훈(2008) 연구가 있고, 김민성·김봉근·하태욱(2009)은 한국의 세대간 소득탄력성을 측정해 미국보다 상대적으로 높은 세대간 소득이동성추정치를 제시하였으나, 상대적으로 짧게 축적된 패널기간을 사용함으로 인해 부모세대의 연령이 상대적으로 높고, 적은 표본가구수와 외환위기에 직접적으로 영향을 받은 1998년 근방의 부모세대 소득분포를 사용했다는 한계면에서 후속보완연구가 필요한 실정이다.

세대내 소득이동성과 관련하여서는 Fields and Ok(1999)이 미국의 소득이동성 연구들을 개괄하여, 미국의 세대내 소득이동성의 크기가 그리 크지 않고, 추세는 1970~1980년대에 일정하게 유지되고 있음을 보고하였다. 성명재·강신욱·이철인(2008)은 한국노동패널(KLIPS)로 우리나라의 세대내 소득이동성을 측정하여, 저소득층과 고소득층의 소득이동성이 중위소득층에 비해 상대적으로 큼을 보였다.

대부분의 기존 연구들은 패널자료를 이용해서 세대내 소득이동성을 측정하게 되는데 패널자료로 측정된 소득이동성은 측정오차(measurement error)와 비무작위적 탈락(non-random attrition)으로 인하여 추정편의(bias)를 갖게 된다. Antman and Mckenzie(2007)는 측정오차는 희석편의(attenuation bias)를 야기 시키고 이동성의 정도를 과대추정

(overestimate)시킬 수 있음을 지적하였다. 패널의 표본이탈(attrition)이 높은 소득변화율을 보이는 코호트에서 주로 발생한다면 추정된 소득이동성은 과소추정편의를 갖게 된다. 기존의 여러 연구는 측정오차문제를 극복하기 위해서 도구변수 추정방법(instrument variable correction)을 이용하였다. 하지만 측정오차가 고전적인 측정오차모형에서 벗어날 때, 도구변수보정방법은 유효하지 못하게 된다. 다음 장에서 이 문제를 보다 자세히 논의한다. 만약, 비고전적인 측정오차(measurement error)와 비무작위적 이탈(non-random attrition)과 같은 문제들이 저소득층 소득이동성의 정확한 측정을 저해한다면, 실증연구결과에 근거한 저소득층에 대한 빈곤정책은 소기의 목적을 달성하기 어려울 것이다. 본 연구에서는 통상적인 소득이동성 연구에서 제기되는 패널자료의 한계를 논의하고, 적절한 보정방법을 사용해 한국의 세대내 소득이동성의 크기를 정확하게 추정하고자 한다. 구체적으로 본 연구에서는 유효한 보정방법으로 가(假)패널(pseudo panel)데이터조사 방식[Antman and McKenzie(2007)]을 이용하여 소득이동성의 측정의 정확도를 높이고 기존 패널자료의 결과물과 비교 분석한다.

가패널 방법론은 기존의 패널자료의 특성처럼 동일한 개인을 시간에 따라서 변하는 양상 및 행태를 관찰할 수 있도록 가상으로 구축하는 패널자료이다. Deaton(1985)은 반복된 횡단면 자료에서 코호트(cohort)패널로 구성된 가패널은 두 가지의 측정오차관련문제를 해결할 수 있다고 제시하였다. 첫 번째로는 각 기간 내에 코호트평균을 택한 가패널로 개인 수준(individual level)의 측정오차를 제거하고 두 번째로 가패널 자료를 이용할 때 원 패널 자료의 동일한 가구주의 자기상관(serial correlation)을 가진 측정오차문제를 해결할 수 있다고 하였다. 패널자료로는 한국노동연구원에서 조사되는 한국노동패널(KLIPS)과 가패널 자료로는 통계청에서 매월 조사되는 가계조사를 이용한다. 2장에서는 패널자료를 이용해서 소득이동성을 측정하는 기존방법론과 도구변수법들의 추정편의에 대해서 설명하고, 가패널을 사용하는 보정방법을 설명한다. 3장에서는 조사 자료와 통상패널과 가패널을 통한 소득이동성추정결과를 논의하며, 이어서 비무작위적 표본이탈현상과 빈곤함정의 존재 여부를 논의하며, 끝으로 4장에서는 실증분석의 결과를 요약하고, 후속연구방향을 논의한다.

2. 통상패널의 소득이동성추정 편의와 보정방법

흔히, 소득이동성은 동일가구의 $t-1$ 기 소득을 t 기 소득에 회귀분석을 통해서 얻게 되는 기울기 계수 값이라고 정의된다[Jarvis and Jenkins(1998)]. 우선 패널조사를 이용해서 소득이동성을 측정해보도록 하면, 기간 t 에 개인 i 의 로그(log) 소득에 대한 소득이동성 측정에

대한 회귀방정식은 식 (2.1)을 통해서 측정될 수 있다.

$$(2.1) \quad Y_{i,t}^* = \alpha + \beta Y_{i,t-1}^* + u_{i,t}$$

식 (2.1)에서 측정된 계수 값 β 는 소득이동성을 의미하며, $Y_{i,t}^*$ 와 $Y_{i,t-1}^*$ 은 각각 t 기와 $t-1$ 기에 관측된 가구주의 로그(log) 소득을 의미한다. 위 식 (2.1)에서 계수 값 β 가 1보다 큰 값을 갖는다면 소득이 동태적으로 발산(divergence)하고, 1보다 작은 값을 갖게 되면 수렴(convergence)하게 된다. 또한, 음의 값은 이동성의 정도의 순위가 ‘역전’(reversal)을 의미하는데 이것은 평균소득 위에 존재하는 개인이 소득의 하락을 경험하는 것을 의미한다.

하지만 기존의 패널조사에서는 잘 알려져 있듯이 자영업자와 고소득층 같은 경우에는 소득을 과소보고하려는 경향이 있기 때문에 자료에서 조사된 소득은 식 (2.2)와 같이 실현된 가구주의 소득과 오차로 구성되어 있을 것이라고 생각할 수 있다.

$$(2.2) \quad Y_{i,t} = Y_{i,t}^* + \varepsilon_{i,t}$$

Bound and Krueger(1991)는 측정오차는 연속하는 기간에 정(positive)의 자기상관(autocorrelation)관계가 있고, 실제 소득과는 부(negative)의 상관관계가 있음을 보였다. 이러한 측정오차는 소득이동성 추정에 편의를 유발한다. 즉, 측정오차를 가진 추정식은

$$(2.3) \quad Y_{i,t} = \alpha + \beta Y_{i,t-1} + \eta_{i,t}, \quad \eta_{i,t} = u_{i,t} + \varepsilon_{i,t} - \beta \varepsilon_{i,t-1}$$

식 (2.3)의 β 의 회귀분석추정치 극한값은

$$(2.4) \quad \begin{aligned} \text{plim } \hat{\beta}_{OLS} &= \beta + \theta_{OLS}, \\ \theta_{OLS} &= [\text{Cov}(u_{i,t}, Y_{i,t-1}) + \text{Cov}(\varepsilon_{i,t}, \varepsilon_{i,t-1}) + \text{Cov}(\varepsilon_{i,t}, Y_{i,t-1}^*) \\ &\quad - \beta \text{Var}(\varepsilon_{i,t-1}) - \beta \text{Cov}(Y_{i,t-1}^*, \varepsilon_{i,t-1})] / \text{Var}(Y_{i,t-1}) \end{aligned}$$

Antman and Mckenzie(2007)는 통상최소자승추정량(Ordinary Least Squares) 편의 가능성을 다음과 같이 설명하였다.

- $\text{Cov}(u_{i,t}, Y_{i,t-1})$ 는 t 기의 소득 충격과 $t-1$ 기의 측정된 소득의 상관정도를 나타낸다. $u_{i,t}$ 에 개인 고정효과(Individual fixed effect)로 인해서 정(positive)의 값을 갖게 될 것이고 $u_{i,t}$ 가 자기상관(autocorrelation)관계가 있으면 0의 값을 갖지 않게 된다.

- $Cov(\varepsilon_{i,t}, \varepsilon_{i,t-1})$ 는 연속하는 두 기간의 측정오차와의 상관관계를 나타낸다. 측정오차가 자기상관(autocorrelation)관계가 있으면 역시 0의 값을 갖지 않게 된다.
- $Cov(Y_{i,t-1}^*, \varepsilon_{i,t-1})$ 는 측정오차와 실질소득 간의 상관관계를 나타내는데 Bound and Krueger(1991)는 부의 값을 갖게 될 것이라고 하였다. 측정오차가 정(positive)의 자기상관관계가 있으면 $t-1$ 기의 실현된 소득과 t 기의 측정오차 상관관계는 결국 부(negative)의 값을 가진다.
- $Var(\varepsilon_{i,t-1})$ 는 측정오차의 분산(variance)을 나타낸다. 고정효과가 존재하지 않으면 측정오차는 결국 다음 식의 결과를 얻게 되어 희석편의(attenuation bias)를 갖게 된다.

$$(2.5) \quad plim \hat{\beta}_{OLS} = \beta \left[1 - \frac{Var(\varepsilon_{i,t-1})}{Var(Y_{i,t-1})} \right]$$

측정오차의 존재로 인해서 소득이동성(earnings mobility)의 측정의 문제가 발생할 때 독립변수(independent variable)와 상관관계(correlation)가 있고 오차항(error term)과는 상관관계가 없는 도구변수를 채택하여 도구변수 추정을 이용하여 이러한 문제를 해결할 수 있다. 그러나 Antman and McKenzie(2007)에서 보듯이 도구변수추정은 위의 추정편의항목들을 모두 해결하지 못한다. 즉, 소득이동성을 측정하는데 사용되는 도구변수로 교육수준, 지출, 자산 관련변수 등을 사용할 수 있다. $Z_{i,t-1}$ 을 도구변수로 하는 회귀방정식(regression)은 다음과 같다.

$$(2.6) \quad Y_{i,t-1}^* = \phi + \gamma Z_{i,t-1} + v_{i,t-1}$$

$\gamma \neq 0$ 이라고 가정하고, 식 (2.2)에서처럼 소득은 오차항(error term)을 갖고 있으므로 식 (2.6)은 다음과 같이 나타낼 수 있다.

$$(2.7) \quad Y_{i,t-1} = \phi + \gamma Z_{i,t-1} + v_{i,t-1} + \varepsilon_{i,t-1}$$

식 (2.7)을 회귀분석을 통해 얻은 $\hat{Y}_{i,t-1}$ 을 식 (2.4)에 대입하여 식 (2.3)의 회귀식을 통해 $\hat{\beta}$ 을 얻을 수 있다. 이 과정은 2SLS(two stage least square) 방법으로 도구변수추정과 β 추정식은 동일하지만, 정확한 표준오차(standard deviation)를 도출하는 도구변수추정량은 다음과 같다.

$$\begin{aligned}
 \hat{\beta}_{IV} &= \frac{\sum_{i=1}^N Y_{i,t} z_{i,t-1}}{\sum_{i=1}^N Y_{i,t-1} z_{i,t-1}} \\
 (2.8) \quad &= \beta + \frac{\sum_{i=1}^N (u_{i,t} + \varepsilon_{i,t} - \beta \varepsilon_{i,t-1}) z_{i,t-1}}{\sum_{i=1}^N Y_{i,t-1} z_{i,t-1}}
 \end{aligned}$$

식 (2.8)의 확률 극한(probability limit)을 구하려면 도구변수의 시계열(time series)적인 구조가 필요하다. 그러므로 다음과 같이 도구변수에 대해 가정하면

$$(2.9) \quad Z_{i,t} = \mu + \rho Z_{i,t-1} + \omega_{i,t}$$

식 (2.9)에서 ρ 는 도구변수의 시계열적인 상관관계 여부를 나타내는 계수이다. 그러면 식 (2.9)를 통해서 도구변수가 시계열적으로 자기상관관계가 없는 불변한 변수를 사용해야 한다는 것을 추측해볼 수 있다.

$$(2.10) \quad \text{plim } \hat{\beta}_{IV} = \beta + \frac{\gamma(\rho - \beta)\text{Var}(Z_{i,t-1}) + E(\varepsilon_{i,t} Z_{i,t-1}) - \beta E(\varepsilon_{i,t-1} Z_{i,t-1}) + \lambda}{\gamma\text{Var}(Z_{i,t-1}) + E(Z_{i,t-1} \varepsilon_{i,t-1})}$$

$$(2.11) \quad \lambda = \gamma E(\omega_{i,t} Z_{i,t-1}) + E(v_{i,t} Z_{i,t-1})$$

Antman and McKenzie(2007)는 식 (2.10)은 다음과 같은 조건이 성립할 때만 편의가 없음을 설명하였다.

- 도구변수 $Z_{i,t-1}$ 은 $\varepsilon_{i,t}$ 과 $\varepsilon_{i,t-1}$ 상관관계가 없다. 이러한 조건은 교육수준과 같은 도구변수로 측정했을 때는 성립하지만, 소득의 도구변수로 지출관련 변수를 사용했을 때는 성립하지 않는다.
- $\lambda = 0$ 이라는 것은 도구변수인 $Z_{i,t-1}$ 가 오차항인 $\omega_{i,t}$ 와 $v_{i,t}$ 상관관계가 없다. 이 조건은 개인고정효과를 포함하는 $Y_{i,t}^*$ 도구변수와 상관관계가 있거나 도구변수 자체에 개인고정효과를 포함한다면 성립하지 않는다.
- 도구변수의 자기상관관계(autocorrelation) 정도는 소득의 자기상관관계 정도와 완전하게 대등된다. 즉 $\beta = \rho$ 라는 것을 의미한다.

위의 조건 중에 첫 번째, 두 번째 조건을 만족하면

$$(2.12) \quad \text{plim } \hat{\beta}_{IV} = \rho$$

도구변수추정이 자기상관계수에 수렴(converge)할 것이라는 것을 의미하고 교육수준과 같이 시간에 대해 불변한(time invariant) 변수를 도구변수로 사용하면, $\hat{\beta}_{IV}$ 가 1의 값으로 수렴함을 의미한다. 종합하면 일반적으로 도구변수추정법은 소득이동성 추정에서 편의를 피할 수 없게 된다. 대안으로 본 연구에서는 가패널 방법을 소득이동성 측정에 적용해보고 기존의 패널자료의 측정치와 비교해 보고자 한다.

가패널 방법을 이용하려면 코호트(cohort)분석을 이용하게 된다. 특정코호트는 어떠한 특성을 기준으로 분류된다. 각 코호트별로 관측된 소득의 평균인 코호트평균을 사용하면 식 (2.3)과 소득이동성 추정식은 동일한 형식이지만 상이한 점은 코호트평균을 사용한 회귀분정식이다. t 기간 코호트 c 에서 n 명의 개인에 대한 회귀분석 식은 다음과 같다.

$$(2.13) \quad \begin{aligned} \overline{Y_{c(t),t}} &= \alpha + \beta \overline{Y_{c(t-1),t-1}} + \overline{u_{c(t),t}} + \overline{\varepsilon_{c(t),t}} - \beta \overline{\varepsilon_{c(t),t-1}} + \lambda_{c(t),t}, \\ \lambda_{c(t),t} &= \beta (\overline{Y_{c(t),t-1}} - \overline{Y_{c(t-1),t-1}}) \end{aligned}$$

$\overline{Y_{c(t),t}}$ 와 $\overline{Y_{c(t-1),t-1}}$ 는 각각 t 기와 $t-1$ 기에 코호트 c 에서 관측된 로그소득의 평균을 나타내고 있다. 각 기간에서 서로 다른 개인을 관찰했으므로 어떠한 측정오차의 함은 확률극한에서 서로 상쇄된다. Antman and Mckenzie(2007)는 각 코호트 개인의 수가 무한히 커진다고 가정하게 되면 $\lambda_{c(t),t}$ 는 0으로 수렴함을 보였다. 하지만 이러한 가정은 동일한 $t-1$ 기의 관측된 개인은 t 기와 $t+1$ 기에도 조사되어야 한다는 전제가 필요로 하므로 너무 고령자이거나 또 아직 가구형태가 안정되지 않은 너무 젊은 층의 가구주를 선택하게 되면 위의 가정을 유지하는 것이 쉽지 않을 것이다. 그러므로 본 연구에서도 이러한 점을 감안하여서 가구주의 연령을 1998년 기준으로 20세~65세로 제한하였다. $\overline{\varepsilon_{c(t),t}}$ 는 t 기에서 코호트 c 개인소득의 평균 측정오차(mean measurement error)이다. $n_c \rightarrow \infty$ 이면, 다음과 같이 측정오차에 대한 가정이 성립함을 알 수 있다.

$$(2.14) \quad \text{plim } \overline{\varepsilon_{c(t),t}} = \text{plim } \frac{1}{n_c} \sum_{i=1}^{n_c} \varepsilon_{i(t),t} = 0$$

3. 실증분석

3.1. 자료

식 (2.1)을 측정하기 위해서, 한국노동패널(KLIPS)을 사용하는데 KLIPS는 한국노동연구

원에서 1998년부터 진행된 한국을 대표하는 종단면조사(longitudinal household)로 소득 활동 및 소비에 대한 다양한 정보를 포함하고 있다. KLIPS의 소득 자료는 금융소득, 부동산 소득, 사회보험소득, 이전소득, 기타소득으로 구성되어 있다. 소득항목에 대해서 살펴보면 금융소득 같은 경우는 은행이나 금융기관 같은 곳에서 연계 되는 이자소득 또 주식이나 채권을 통해 얻는 배당금 등을 포함하고 있고 이전소득 항목은 정부보조금 및 사회단체 보조금, 기타보조금에 관련된 항목이 조사되고 있다. KLIPS자료의 경우 1차년도의 소득은 세부변수가 없고 조사시점의 문제로 소득 자료는 한계성이 있다고 알려져 있다. 본 연구에서는 소득이동성의 정도와 그 추세를 알아보기 위해 1998년부터 2007년까지의 자료를 사용하게 되는데 KLIPS자료에서 1차년도의 조사 자료의 한계성이 존재하지만 1차년도의 소득 자료도 포함하였다.

또한, KLIPS자료의 가구주 소득은 근로소득, 금융소득, 부동산소득, 사회보험소득, 이전소득을 모두 합한 변수를 이용하였고 기타소득⁽²⁾은 제외했고 KLIPS자료에서 소득을 구성할 때 당해 연도에 분가한 가구의 가구소득은 지난해 1년간의 소득이 아니라 분가한 이후의 소득이기 때문에 낮게 추정된 소득이므로 분가가구는 제외했다.

가패널 방법에 사용되는 가계조사는 1942년부터 실시된 법정 국가기본 통계자료로서 1963년부터는 매년 통계청에서 조사되고 있으며, 2003년부터는 조사대상을 읍면지역 비농가까지 확대되면서 가계조사로 불린다. 가계조사는 전국에 거주하는 가구의 수입과 지출을 조사하여 가구의 생활 실태와 변동사항을 파악하는 것이 목적이다. 가계조사의 조사방법은 매월 조사되고 농가, 어가, 1인 가구, 가계수지 파악이 곤란한 가구, 외국인 가구를 제외한 전국에 거주하는 가구를 대상으로 하고 있다. 하지만 2006년부터는 조사대상을 1인 가구까지 확대되었다. 가계조사의 소득부분은 크게 정상소득과 비정상소득으로 구분된다. 정상소득은 가계지출에 가장 큰 영향을 주는 정기적이고 반복적으로 재현 가능성이 있는 소득으로서 근로소득, 사업소득, 재산소득 및 이전소득으로 구성되어 있고 비정상소득은 일시적이며 비정상적으로 발생한 소득을 의미한다. 가계지출은 가계를 영위하기 위한 지출로 소비지출과 비소비지출로 구성되어 있으며 소비지출은 일상생활을 영위하기 위하여 필요한 상품 및 서비스의 구입을 위해 지출된 현금지출과 카드 유가 증권 등 일체의 지출을 말하며 비소비지출은 조세, 공적연금, 사회보험 등 가구의 자유의지 없는 지출 및 소비지출에 포함시킬 수 없는 이전지출을 의미한다.

소득이동성을 측정하기 위해서는 가구주의 소득에 대해서 정의하여 회귀분석에 사용되

(2) 기타소득은 세부 소득항목 중 따로 구별하기 힘든 것을 묶어놓은 소득범주인데 퇴직금, 보험으로 인해 받은 돈, 증여 및 상속 등이 속한다.

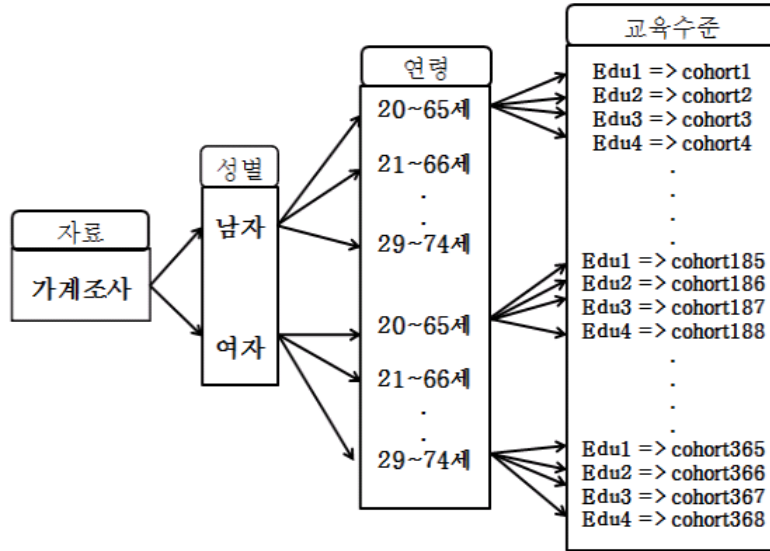
는 소득을 결정해야 한다. 가계조사는 <표 2>에서처럼 소득에 대한 정보가 다양하게 제공 되기 때문에 연구에서 사용될 소득 자료를 선택해야 한다. 소득에 대한 정의는 매우 다양하다. OECD나 세계은행 등과 같은 경우에서 시장소득, 총소득(gross income), 세후소득(after-tax income) 또는 가처분소득(disposable income) 등으로 소득을 구분하여 사용하고 있다. 여러 연구에서는 가처분소득을 이용하는 경향이 있지만 본 연구에서는 가구주의 소득으로 총소득에서 비경상소득을 제외한 경상소득만을 이용하여 분석하였다. KLIPS와 가계조사의 소득변수들은 물가수준을 고려한 실질소득(real income)으로 변환되어 사용한다. 조사된 소득은 명목소득(nominal income)이기 때문에 가구주의 소비자물가지수를 이용해서 실질소득을 구할 수 있다. 실질소득⁽³⁾으로 모두 변환한 후에 1998년~2007년까지의 가구주의 소득변수를 만들 수 있다.

3.2. 가패널구성 및 기초통계량

본 연구에서는 가계조사를 토대로 가패널 방법론을 적용하였다. 가패널 방법론이 반복된 횡단면 조사를 바탕으로 이루어지기 때문이다. 한국노동패널조사에서는 조사차수에 따라서 부여되는 가구의 번호에 따라서 해당되는 가구주의 소득을 이용하여 이동성을 측정했었다면, 가패널 방법에 사용되는 가구주는 특별한 특성을 적용해서 만들어진 가구주의 소득을 이용해서 이동성을 측정할 수 있다. 즉, 가패널 방법론은 동일한 특성을 가진 임의의 가구주를 만들어내기 위해 코호트(cohort)를 분석단위로 선택한다.

<그림 1>은 가패널에서 이용되는 코호트 분석에 대한 그림이다. 그림을 토대로 가패널 방법론에 사용되는 코호트는 성별, 교육정도와 같은 시 불변(time invariant)한 변수로 생성할 수 있다. 가구주를 분류하는 기준인 연령은 가구주가 가구를 형성하는 시기와 노동시장에 참여하고 또 이탈 시기를 고려하여 20세부터 65세로 제한하였다. 우선 가계조사자료에서 제일 먼저 성별을 분류하는데 남자와 여자로 구분하게 된다. 가구주와 배우자의 성별이 남자인 경우와 반대로 가구주와 배우자의 성별이 여자인 경우로 분류할 수 있다. 성별로 분류하게 되면 지금까지 가계조사자료는 두 가지 코호트로 이루어져 있는 것이다. 가계조사에서 성별을 모두 구별한 후에는 가구주를 연령별로 구분할 수 있다. 본 연구에서는 가구주의 연령을 1998년을 기준으로 20세에서 65세까지로 분류하였고 1998년에 20세인 가구주는 1999년에는 21세, 2007년에는 29세가 되고 1998년에 65세인 가구주는 2007년에는 74세가 되도록 연도별로 시간이 지남에 따라서 가구주의 연령도 변하기 때문에 가구주의 연령을 상이하게 분류하였다. 그러면 연도별로 46개의 코호트가 만들어지게 되고 전체 가

(3) 실질소득은 명목소득/(2005년 = 100 기준 소비자물가지수)×100으로 계산되었다.



〈그림 1〉 코호트구성

계조사자료 상에서는 1998년도를 기준으로 한다면 92개의 코호트가 생성된 것으로 계산될 수 있다.

마지막으로 가구주의 교육정도를 분류할 수 있는데 교육정도는 고졸, 전문대졸, 4년제 대학졸업, 대학원 이상과 같이 총 4가지로 분류하였다. 결국, 가계조사에서 사용된 코호트는 성별, 연령, 교육정도로 세분화되어서 한국노동패널조사의 가구주와는 다르게 고유한 특성을 갖는 가구주를 만들어 낼 수 있고 각 코호트에 있는 개인들은 서로 공통적인 요인들을 포함하고 있기 때문에 서로 비슷한 정도의 소득을 갖고 있을 거라고 추측한다.

가패널 자료에서 연도별로 만들어질 수 있는 코호트의 개수는 $\text{성별} \times \text{연령} \times \text{교육정도} = 2 \times 46 \times 4 = 368$ 과 같이 총 368개의 코호트가 될 수 있다. Deaton(1985)은 각 코호트마다 적어도 100개 이상의 개인이 포함되어 있어야 가패널 계수의 유의성을 보장해주기 때문에 본 연구에서 사용된 코호트에도 적어도 100개 이상의 개인을 포함해서 분석하였다.

〈표 1〉은 KLIPS와 가패널의 기초통계량을 나타내고 있다. 사용된 모든 연도를 모두 표시하지 않고 전체, 첫 번째 연도, 중간, 마지막 연도로 전체적인 흐름을 보여주려 했다. 전체 패널에서의 비교를 하면 KLIPS의 가구주의 소득은 평균은 16.75이고 가패널은 17.27 정도로 더 높은 편이다. 연도별 흐름을 볼 때도 가패널의 소득의 평균이 KLIPS 보다 모두 높은 편으로 나타났다. 이것은 가패널의 특성상 코호트 분석을 이용하였기 때문에 KLIPS에 비해서 표준편차도 대체적으로 적은 것으로 나타났다. 가구주의 연령도 KLIPS의 경우

〈표 1〉 기초통계량

	연도	관측치	변수설명	평균	표준편차	최소값	최대값	
한국 노동 패널	전체	6394	가구주소득	16.75	0.99	10.37	22.09	
			교육연한	10.83	4.43	0	23	
			교육수준 더미	고졸=1 전문대졸=2 4년대졸=3 석사이상=4 박사이상=5				
			가구주연령	49.54	13.76	16	97	
			성별	1.17	0.37	1(83.4)	2(16.58)	
			4747	가구주소득	16.73	1.12	11.20	22.09
	1998		교육연한	10.73	4.43	0	23	
			가구주연령	46.76	13.01	18	94	
			성별	1.14	0.35	1	2	
			4085	가구주소득	16.77	1.01	10.69	20.54
	2002		교육정도	10.63	4.40	0	23	
			가구주연령	49.65	13.55	18	89	
			성별	1.15	0.36	1	2	
			4902	가구주소득	16.93	0.94	11.65	21.60
	2007		교육정도	11.22	4.41	0	23	
			가구주연령	51.24	14.40	19	91	
			성별	1.20	0.40	1	2	
			가 패 널	전체	326	가구주소득	17.27	0.49
	1998	281	가구주소득	17.17	0.43	15.75	18.36	
	2002	287	가구주소득	17.35	0.43	15.98	18.84	
2007	320	가구주소득	17.28	0.54	14.63	18.93		

에는 전체를 기준으로 49.4세이고 가구주의 연령 평균이 제일 높은 연도는 2007년이다.

KLIPS에서 가구주의 성별은 남자는 1, 여자는 2로 표시되는데 가구주 중 남자의 비율은 83.4%로 남자가 가구주의 대부분을 차지한다고 할 수 있다. 교육연한은 최소값이 0인 경

우는 교육을 받지 않은 경우를 의미하고 최대값인 23은 교육수준이 박사 이상을 의미한다. 교육수준더미변수 또한 KLIPS에서 조사된 방식대로 기준을 나눠서 더미변수를 각각 1부터 5까지의 값을 부여했다.

가패널의 관측치는 각 연도의 코호트수를 의미하는데 앞에서 코호트 분석에서도 설명하였듯이 연도별로 생성될 수 있는 코호트수는 총 368개이지만 정확히 368개의 코호트가 만들어질 수는 없다. 예를 들어 1998년 경우에는 코호트가 281개인데 만일 23세 남자 가구주 중 교육수준이 대학원 이상의 조건을 가진 가구주가 아무도 없거나 소득항목을 무응답 하였다면 가구주가 남자, 23세, 대학원 이상의 조건을 만족하는 코호트는 생성될 수가 없다. 그러므로 코호트는 가계조사에서 연도별로 조사된 결과에 따라서 상이한 코호트수를 가져온다.

3.3. 소득이동성 추정

〈표 2〉는 한국노동패널과 가계조사를 이용해서 소득이동성을 측정한 것을 나타낸 표이다. 1998년부터 2007년까지의 한국의 소득이동성의 계수를 비교해보면 전체적으로 가패널 방법론을 이용해서 소득이동성을 얻은 가계조사의 계수 값이 한국노동패널의 값에 비해 1에 가까운 값을 갖는다는 것을 알 수 있다. 즉, 가패널의 추정치들은 높은 소득상관계수를 보여, 낮은 소득이동성을 보이고 있다.

가장 두드러지게 계수차이를 크게 보이는 연도는 1998년과 1999년 사이인데 한국노동패널의 경우에는 0.401 정도로 낮게 측정되었지만 가패널 방법을 이용한 가계조사의 소득이동성은 0.809 정도로 측정되었다. 이것은 외환위기 이후에 패널조사에서 소득에 대한 조사가 저소득층에 대해서는 과다하게 보고가 되었거나 고소득층은 과소보고 되었을 거라는 추측을 해 볼 수 있다. 그리고 이 수치들은 앞에서 언급한 1998년도 KLIPS자료의 잠재적인 문제점을 반영하고 있다.

그 외의 경우에도 1999년에서 2000년 사이는 각각 0.658과 0.87이고, 가장 최근의 경우인 2006년과 2007년에는 0.78과 0.82로 앞서 설명했던 두드러진 차이에 비해서는 완화되었지만, 여전히 가패널을 이용한 소득상관계수가 KLIPS보다는 큰 값을 가진다. 즉, 지난 10년 동안 소득상관계수를 볼 때 KLIPS가 가패널에 비해서 작고, 따라서 상대적으로 높은 이동성을 시사하고 있다.

연도별로 측정한 소득이동성이 아닌 전체 패널자료에서 소득이동성을 측정해도, KLIPS의 소득상관계수는 0.67이고 가패널은 0.819로 크다. 이러한 가패널의 소득상관계수가 큰 값을 갖는 것은 기존의 패널의 소득이동성을 측정 시에 나타나는 측정오차와 비무작위적 이탈(non-random attrition)을 어느 정도 해결해 주고 있음을 함의한다. 그리고 가패널

〈표 2〉 통상패널(KLIPS)과 가패널의 연도별 소득이동성

	한국노동패널(KLIPS)		가패널	
	상수항	가구주의 소득	상수항	가구주의 소득
1998~ 1999	9.885 (0.190)	0.401 (0.011)	3.308 (0.565)	0.808 (0.033)
1999~ 2000	5.726 (0.205)	0.657 (0.012)	2.275 (0.703)	0.870 (0.040)
2000~ 2001	4.733 (0.224)	0.715 (0.013)	4.662 (0.591)	0.732 (0.034)
2001~ 2002	5.730 (0.205)	0.665 (0.012)	3.755 (0.703)	0.786 (0.040)
2002~ 2003	4.891 (0.194)	0.709 (0.011)	4.238 (0.824)	0.752 (0.047)
2003~ 2004	1.995 (0.085)	0.741 (0.011)	2.920 (0.485)	0.831 (0.028)
2004~ 2005	5.676 (0.155)	0.666 (0.009)	1.279 (0.528)	0.925 (0.030)
2005~ 2006	4.049 (0.163)	0.762 (0.009)	2.692 (0.660)	0.840 (0.038)
2006~ 2007	3.733 (0.164)	0.781 (0.009)	3.028 (0.618)	0.826 (0.035)
전체	5.557 (0.063)	0.670 (0.004)	3.145 (0.211)	0.818 (0.012)

은 KLIPS에 비해 표본크기(sample size)가 크기 때문에 소득상관계수 값에 대한 유의성도 크다.

3.4. 비무작위적 표본이탈

전장에서 논의하였듯이 세대내 소득이동성 측정편의의 주요 원인으로 지목되는 것이 통상패널의 비무작위적 이탈(non-random attrition)이다. 만약, 표본이탈이 소득이나 고용여부와 체계적인 관련성을 가진다면, 추정된 소득이동성은 편의를 가질 것이다. 예를 들어, 표본이탈이 높은 소득변화율을 보이는 코호트에서 주로 발생한다면 추정된 소득이동성은 과소추정편의를 보일 것이다. 본 절에서는 이러한 가설을 직접 검증해본다. KLIPS도 통상적인 패널과 유사하게, 초기에 큰 표본이탈률을 보이고, 이후 안정화되어 매년 1%포인트 이

내에서 표본이탈률을 보이고 있다. 최근 자료인 7차년도부터 11차년도까지의 원표본 유지률은 각각 77.3%, 76.5%, 76.5%, 75.5%, 74.2%로서 매년 1%포인트 이내에서 표본이탈률을 보이고 있다.

표본이탈이 소득과 체계적인 연관을 점검하기 위해, t 기의 응답여부를 전기의 소득수준에 대해 회귀분석하였고, 응답여부 (역으로 표본이탈여부)는 소득과 음(양)의 상관관계를 가지고 있는 것으로 분석된다. 예를 들어, 본 연구에서 사용된 패널자료에 한 기간을 더한 11차년도의 응답여부는 전기(여기서는 7차년도 사용)의 소득수준에 통계적으로 유의하게 음의 상관관계를 갖는 것으로 분석된다. 프로빗분석에서 소득변수의 추정된 기울기는 -0.0000545 ($SE=0.0000144$)로 1% 수준에서 통계적으로 유의한 값을 가진다. 비선형관계를 통제하기 위해 제곱항을 추가하였고, 제곱항은 5% 수준에 유의하지 않은 것으로 나타났다. 추가적으로 소득이동성은 고용여부와 밀접한 관련성이 있을 것이고, 이상호(2005)도 표본이탈이 가구주의 고용(실업)상태와 통계적으로 유의하게 상관성을 가짐을 보고하고 있다.

3.5. 빈곤함정(Poverty trap)

본 절에서는 통상패널(KLIPS)과 가패널의 상이한 소득이동성의 추정치를 빈곤함정(poverty trap) 존재라는 정책적 시사점과 관련하여 논의한다. 낮은 세대내 소득이동성, 특히 저소득층의 낮은 소득이동성, 보다 정확하게 빈곤선 아래에 위치한 저소득층의 낮은 상방 소득이동성은 빈곤상태에서 빠져나오지 못하는 빈곤함정을 시사한다. 빈곤함정의 존재와 관련하여, Jalan and Ravallion(2004)에 따르면 소득이동성에 대한 선형함수를 변형한 삼차비선형 소득함수(cubic non-linear income mapping)를 사용해 빈곤함정의 존재조건을 다음과 같이 정의한다.

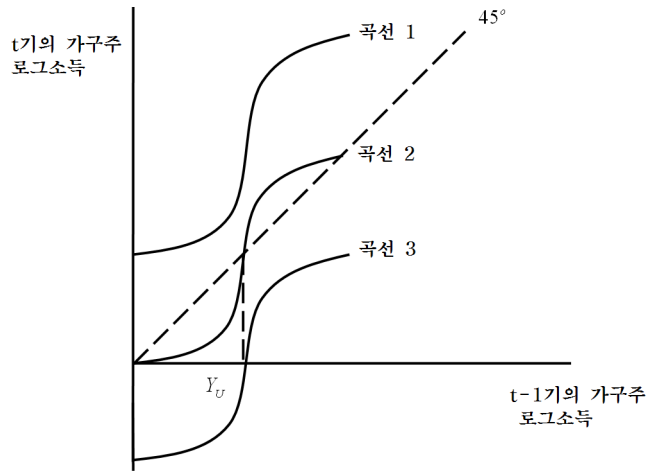
$$(3.1) \quad Y_{i,t}^* = \beta_1 Y_{i,t-1}^* + \beta_2 (Y_{i,t-1}^*)^2 + \beta_3 (Y_{i,t-1}^*)^3 + u_{i,t}$$

의 식을 미분한 형태인 $g(\cdot)$ 로 다음과 같이 정의하게 되면

$$(3.2) \quad g_i'(Y_{i,t-1}^*) Y_{i,t}^* = Y_{i,t-1}^* > 1$$

$$(3.3) \quad \beta_1 + 2\beta_2 Y_{i,t-1}^* + 3\beta_3 Y_{i,t-1}^{*2} > 1$$

식 (3.3)은 삼차비선형 소득함수를 미분한 형태를 나타내고 있으며 도함수 값이 1보다 큰 값을 갖는다면, 위의 차분방정식 모형은 빈곤선 아래에 있는 가구들이 아래쪽으로 발산하



〈그림 2〉 삼차비선형모델에서의 빈곤함정조건

는 경향이 있음을 나타낸다. 즉, 빈곤선 아래에 있는 가구는 상방소득이동성이 제한되어, 빈곤함정에 빠지게 되는 것이다. 위 조건을 통상패널(KLIPS)과 가패널 자료로 비교하면 소득이동성정도를 과소추정하는 KLIPS가 빈곤함정의 존재를 과소보고하려는 경향성이 있을 수 있다.

우선 빈곤선을 정하는 기준은 상대적인 빈곤과 절대적인 빈곤으로 나눌 때, 본 연구에서는 상대적인 빈곤선인 중위소득 50%를 사용하였다. 중위소득이란 총가구 중 소득순위가 정확히 중간인 가구를 기준으로 삼는 것을 말한다. 〈표 3〉은 식 (3.2)의 회귀분석 결과를 토대로 식 (3.3)에 추정된 β 값들을 대입하고 빈곤선을 Y_u 로 정의해서 계산된 값을 보여주고 있다. 삼차비선형 소득함수는 삼차곡선으로 표현할 수 있는데 곡선의 위치에 따라서 빈곤함정여부를 파악할 수 있다. Antman and McKenzie(2006)에서 〈그림 2〉에서 곡선 1은 45도선 위에 있으므로 빈곤함정이 존재하지 않고 곡선 2는 빈곤선 Y_u 이하에 빈곤함정이 존재한다고 하였다. 곡선 3은 모든 가구주가 빈곤함정에 빠져 있음을 알 수 있다. 본 연구에서의 결과는 기초통계량을 근거로 모든 가구주의 소득은 정(positive)의 값을 갖고 있으므로 곡선 2에 해당된다. 곡선 2를 기준으로 Y_u 이하에서 식 (3.3)처럼 기울기가 1을 기준으로 어떤 값을 갖는지 판단해야 한다. 기울기가 1보다 큰 값을 갖는다는 것은 $t-1$ 기에 빈곤함정에 빠졌다면 t 기에도 빈곤함정에 빠질 위험이 크다는 것을 의미한다.

〈표 3〉에서의 결과에서 보듯이 KLIPS에 비해 가계조사 결과 값이 1을 넘거나 더 큰 값을 갖는다는 것을 확인할 수 있다. 이것은 KLIPS가 빈곤함정의 존재를 과소보고하고 있다는 것을 함의한다.

〈표 3〉 통상패널(KLIPS)과 가패널의 빈곤함정 존재여부

자료	전 체		1998~1999		1999~2000		2005~2006	
	KLIPS	가패널	KLIPS	가패널	KLIPS	가패널	KLIPS	가패널
$g'_i(Y_{i,t-1}) Y_{i,t}^* = Y_{i,t-1}^*$	0.849	0.865	0.781	0.844	0.810	1.014	0.836	0.931

위의 추정결과를 통해 빈곤함정에 빠진 가구주에 대해서 가패널 자료에서 보다 패널자료인 KLIPS에서 과소보고 되는 것을 발견하였고 이는 저소득 빈곤층에 대해 정확하게 파악하지 못하게 되는 결과를 가져오게 된다. 그러면 빈곤정책방향에 잘못된 정보를 주게 되어 실질적인 빈곤층 외의 계층에서 정책 수혜를 받을 수 있는 문제가 우려된다.

4. 결 론

통상적인 패널자료(KLIPS)와 도시가계조사로 만든 가패널을 비교해 본 결과 통상적인 패널은 소득이동성을 과다추정하는 것으로 나타났고, 특별히 저소득층에 대한 빈곤함정의 존재에 대해서는 과소추정하고 있는 것으로 나타난다. 가패널을 통한 패널추정치의 편의를 보정할 때, 저소득층에 대한 빈곤함정의 존재가 확인되었고, 장기빈곤에 대한 정책대응이 필요함을 시사하고 있다. 나아가, 보다 구체적인 빈곤탈출정책의 필요성이 제기된다. 나아가 본 연구의 후속연구로 이동성의 추세 및 소득이동성 크기를 결정짓는 주요 정책 및 설명변수의 인과관계를 파악하는 연구가 요구되며, 나아가 소득이동성 향상을 위한 정책변수의 적절한 조합가능성에 대한 연구가 필요한 것으로 사료된다. 또한, 거주지역과 소득이동성을 연계해 계층별 지역간 분리현상에 대한 연구도 요구되며, 장기 주제로는 세대내 소득이동성의 추세에 관한 연구도 가능할 것이다.

성균관대학교 경제학부
 110-745 서울특별시 종로구 명륜동 3가 53
 전화: (02)760-0941
 팩스: (02)760-0946
 E-mail: zooty2640@naver.com

교신저자, 서울대학교 경제학부 부교수
 151-746 서울특별시 관악구 관악로 599

전화: (02)880-6240

팩스: (02)886-4231

E-mail: bgkim07@snu.ac.kr

참 고 문 헌

- 김민성 · 김봉근 · 하태욱(2009): “한국의 세대간 소득탄력성,” 『국제경제연구』, **15, 2**, 87-102.
- 김봉근 · 윤상철(2010): “글로벌 경제위기가 한국의 경제구조에 미친 단기영향과 경기회복 메커니즘 분석,” 『경제연구』, **28, 3**, 1-26.
- 성명재 · 강신욱 · 이철인(2008): “소득이동성과 빈곤특성 분석,” 『저소득층 소득보전정책의 개선방안 연구』, 83-119.
- 안중범 · 전승훈(2008): “교육 및 소득수준의 세대간 이전,” 『재정학연구』, **1, 1**, 119-142.
- 오주현(2009): “한국의 소득이동성과 빈곤함정의 존재여부,” 석사학위 논문, 성균관대학교 대학원.
- 이상호(2005): “한국노동패널(KLIPS)의 표본이탈분석: 가구소득을 중심으로,” 『노동리뷰』, **1, 11**, 66-80.
- Antman, Francisca, and David J. McKenzie(2006): “Poverty Traps and Nonlinear Income Dynamics Heterogeneity,” *Journal of Development Studies*, Forthcoming.
- _____ (2007): “Earnings Mobility and Measurement Error: A Pseudo-Panel Approach,” *Econometric Development and Cultural Change*, **56, 1**, 125-161.
- Bound, John, and Alan Krueger(1991): “The Extent of Measurement Error in Longitudinal Earnings Data: Do Two Wrongs Make a Right?” *Journal of Labor Economics*, **9**, 1-24.
- Deaton, Angus(1985): “Panel Data from Time Series of Cross-sections,” *Journal of Econometrics*, **30**, 109-126.
- Fields, Gary, and Efe A. Ok(1999): “The Measurement of Income Mobility: An Introduction to the Literature,” *Handbook of Inequality Measurement*.
- Jalan, J., and M. Ravallion(2004): “Household Income Dynamics in Rural China,” in S. Dercon(ed.), *Insurance Against Poverty*, Oxford, Oxford University, 108-124.
- Jarvis, Sarah, and Stephen P. Jenkins(1998): “How Much Income Mobility is There in Britain,” *The Economic Journal*, **108**, 428-443.

Abstract

Earnings Mobility in Korea

Ju Hyun Oh · Bonggeun Kim

We measure the degrees of earnings mobility in Korea, which is defined as the rate of changes of relative income ranking over two terms, with a conventional panel subject to measurement error and non-random attrition and a pseudo panel. The different degrees of earnings mobility from two panels were applied to the empirical detection of poverty traps.

Keywords: Earnings mobility, Pseudo panel, Poverty trap

