

패널자료를 이용한 자영업자의 소득 축소보고율 분석: engel곡선이동과 실질소득 추정*

김 봉 근** · 정 철*** · 박 명 호****

논문 초록 본 연구는 engel곡선 추정방법론을 한국노동패널 데이터에 적용하여 우리나라 자영업자의 소득 축소보고 경향을 새롭게 추정하였다. 방법론상으로 본 연구는 관련 기존 연구에서 존재된 자영업자의 소득의 변동성에 관한 정보와 축소보고경향의 정보를 패널자료를 이용해 분리할 수 있음을 보인다. 또한, 기존 연구에서 필요한 가정을 완화할 수 있음을 보이고, 구간추정에 그치는 기존의 실증분석에서 나아가 보다 정확한 하나의 추정치를 제시한다. 본 연구의 실증분석 결과는 자영업자가 근로소득자에 비해 33% 정도의 소득을 축소보고하는 경향이 있음을 보인다.

핵심 주제어: engel곡선, AIDS 수요추정식, 자영업자 소득축소보고

경제학문헌목록 주제분류: C40, H26, J24

투고 일자: 2008. 4. 22. 심사 및 수정 일자: 2008. 5. 21. 게재 확정 일자: 2008. 6. 25.

* 이 논문은 성균관대학교의 2007학년도 성균학술연구비에 의하여 연구되었음.

** 제1저자, 성균관대학교 경제학부, e-mail: bgkim07@skku.edu

*** 제2저자, 대외경제정책연구원 연구위원, e-mail: cchung@kiep.go.kr

**** 제3저자, 한국조세연구원 전문연구위원, e-mail: ecpmh@kipf.re.kr

I. 연구의 배경 및 선행문헌

본 연구는 앵겔곡선 추정방법론을 한국노동패널 데이터(Korean Labor and Income Panel Study: KLIPS)에 적용하여 우리나라 자영업자의 소득 축소보고 경향을 추정한다. 본 연구에서는 기존 연구 방법론의 문제점을 패널자료를 사용해 보완하고, 보다 정확한 자영업자의 소득 축소보고율의 추정치를 제공한다. 기존의 연구들이 주로 횡단면자료에 기초하여 이루어진데 반해, 본 연구는 패널자료를 활용함으로써 소득축소보고경향의 원인에 대한 구체적인 설명을 제시할 수 있다는 장점이 있다. 또한 실증분석에 있어서 AIDS(Almost Ideal Demand System) 수요추정식이라는 이론적 근거를 바탕으로 도출한 추정방정식을 이용함으로써 소비와 소득 간의 축약방정식에 의존하는 기존 연구들과 차별화된다.

본 연구와 관련된 선행문헌 가운데 앵겔곡선의 이동을 통한 실질소득의 추정 및 소비자물가지수 편익(CPI Bias) 추정에 대한 직관적인 해석은 본 연구의 자영업자의 실질소득 추정과 그 맥을 같이하므로, CPI Bias 관련 연구들을 먼저 간략히 소개하고 이어서 본 연구의 주제인 자영업자(또는 사업소득자, 비근로소득자)들이 근로소득자에 비해 소득축소보고하는 경향과 관련된 기존의 선행연구들을 살펴본다. 여기서 후자는 소득과 소비의 규칙적인 상관관계를 이용하여 자영업자들의 실제 소득을 미시적으로 추정함으로써 거시경제 변수인 지하경제의 규모를 측정하는 연구들로도 분류된다.

Hamilton (2001a)은 소비자물가지수의 상향편의 정도를 측정하는 새로운 방법으로 앵겔(Engel) 계수의 개념을 이용할 것을 제시한다. 그는 가계의 식료품비 비중의 변동과 실질소득의 변동을 비교함으로써 소비자물가지수의 편익을 추정한다. 여기서 실질소득의 변동은 소비자물가지수로 조정한 명목소득의 변동이며 가계의 식료품비 비중은 실질생계표준(real living standards)을 반영한다. 역사적으로 안정적인 것으로 입증된 식료품비와 실질소득간의 관계를 기준으로 삼아 소비자물가지수의 편익을 추정하는 것으로, 시간의 흐름에 따라 식료품비 비중이 감소할 경우 이는 그에 상응하는 실질소득의 증가가 있었음을 암시하는 것이다. 그런데, 만일 실제 데이터상에서 식료품비는 감소하는데 실질소득은 변동이 없는 것으로 관찰되는 경우(예를 들어, 명목소득의 변동이 동일한 정도의 변화율을 가진 소비자물가지수로 조정된 경우), 이는 소비자물가지수가 물가상승을 과도하게 추정한 것으로 해석할 수

있다. 다시 말해서 소비자물가지수의 상향편의를 바로잡지 않아서 발생하는 물가상승률 과대계상이 실질소득 측정에 오류를 가져온 결과로 볼 수 있다는 것이다.¹⁾ 소비자 물가지수의 측정오차를 앵겔곡선의 이동으로 추정하는 이 방법론은 식료품과 비식료품에 대한 선호가 시간에 따라 변하지 않아 그 비중이 일정하다는 가정이 수많은 실증연구들에서 검증되었다는 점에서, 그리고 필수재인 식료품의 소득탄력성이 1보다 크게 작아, 우하향하는 앵겔곡선의 좌측 이동을 파악하기 쉽다는 점에서 다른 방법론에 비해 선호된다.

동일한 방법론을 두 집단간 앵겔곡선의 이동에 적용한 다른 논문에서 Hamilton (2001b)은 1974~1991년 동안 소비에서 차지하는 식료품비 비중이 흑인의 경우 백인에 비해 더 빠르게 하락했음을 관찰하고 그 원인으로 측정되지 않은 소득 상승분이 흑인의 경우에 상대적으로 높았음을 지목한다. Hamilton은 미국의 대표적 패널자료인 PSID (Panel Study of Income Dynamics) 데이터를 이용하여 표본기간 중 매년 앵겔곡선의 이동을 관찰함으로써 소비자물가지수의 상향편의 정도가 인종 간에 다르게 나타남을 보이고 그 차이를 앵겔곡선의 이동으로 추정한다. 동 연구는 다음의 두 가지 가정에 기초한다. 첫째, 앵겔계수로 대변되는 실질소득과 식료품비지출의 일정한 상관관계가 두 인종집단간에도 차이가 없을 것이라는 가정과 둘째, 주거지역이 분리된 두 인종집단이 접하는 물가수준에 차이가 존재한다는 가정이다. 본 연구에서도 분석대상이 근로소득자와 자영업자, 두 집단으로 구분되고 두 직업군간에 앵겔곡선의 이동이 있다는 점은 Hamilton (2001b)의 경우와 유사하나, 본 연구는 두 직업군간의 앵겔곡선 이동이 자영업자의 실질소득 축소보고에 기인할 것이라는 가설을 바탕으로 하고 있다. 한편, 직업군간 소득변동성의 차이와 이로 인한 식료품지출행태의 차이(앵겔곡선의 직업군간 차이)가 집단간 소비선호의 이질성 (preference heterogeneity)에 따른 결과로 해석될 수 있는 가능성도 동시에 고려해야 한다.

두 집단간 소득과 소비의 정규적 관계를 이용하여 지하경제의 규모를 측정한 연구로는 Pissarides and Weber (1989)와 유일호 (1995)가 있다. Pissarides and Weber (1989)는 근로소득자에 비해 자영업자가 소득을 축소 보고하는 경향이 있으

1) 동일한 방법론으로 Beatty and Larsen (2005)과 Gibson et al. (2007)은 각각 캐나다와 러시아의 소비자물가지수 편의를 추정하였다. 정철, 김봉근, 박명호 (2007)에서는 한국 노동패널 자료(KLIPS)를 사용하여, 한국의 CPI Bias를 추정하였다.

나 동일 소득수준의 근로소득자나 자영업자의 소비지출은 유사하다는 가정하에, 자영업자가 보고한 소득 대신 소득과 소비의 관계에서 추정한 소득을 이용하여 영국의 지하경제 규모를 측정한다. 1982년 영국의 가계소비조사 데이터를 이용한 횡단면 분석 결과, 자영업자의 실제소득은 신고금액의 1.55(추정구간 중간값)배에 달한다. 유일호(1995)는 Pissarides and Weber(1989)의 방법론인, 소비지출을 이용한 소득추계방식을 한국 데이터에 적용하여 사업소득의 탈루규모에서 유사한 결과(1994년 1.52배)를 얻는다. 이러한 소득추계방식은, 보다 신빙성 있는 소비지출 자료를 이용하여 축소보고의 가능성이 있는 소득을 실제치에 근접하게 추정할 수 있다는 점에서, 횡단면 데이터를 사용하는 기존의 연구들에서 널리 채택되어 왔다. 이철인(1998)은 기존의 횡단면분석연구들이 근로소득자와 자영업자의 소비행태에 차이가 없다는 가정하에 근로소득자의 소비지출행태방정식을 이용하여 자영업자들의 소득을 추정하는데 대부분의 연구들이 두 집단간 소비행태에 차이가 없다는 가정에 대한 신중한 검정 없이 이루어져 왔음을 지적한다. 이를 검정하기 위해 이철인(1998)은 동태모형을 제시하고 한국의 대우 패널 데이터를 이용하여 분석한 결과, 실제로 두 집단 간 소비행태에 차이가 없음을 보고한다.

한편, Lyssiotou, Pashardes, and Stengos(2004)는 Pissarides and Weber(1989)의 소비지출을 이용한 소득추계방식에 잠재적 편이가 존재할 수 있음을 지적하면서 이를 보완하기 위한 일환으로 완전수요체계(complete demand system) 접근방식을 제안한다. 이러한 잠재적 편이의 이유로, 고정소득은 생필품 등에 대한 지출에 주로 사용하고 사치품 구입은 상대적으로 변동성이 큰 사업소득에 의존하는 소비행태의 이질성(preference heterogeneity)을 소득의 축소보고라고 잘못 구별할 수 있다는 것이다. 또한 자동차, 컴퓨터 등 사업비용으로 처리할 수 있는 재화들에 대한 소비지출의 경우 근로소득보다는 사업소득과 더 밀접한 연관이 있을 수 있기 때문이다.²⁾ 물론, Pissarides and Weber(1989)에서도 자영업자의 소득변동성이 더 크다는 것을 고려하고 있지만, 소비선호의 이질성보다는 항상소득

2) 이러한 혼동에 따른 편이제거 외에도 Lyssiotou, Pashardes, and Stengos(2004)가 제안하는 완전수요체계 접근방식의 장점은 가구들을 소득원천에 따라 다른 집단으로 분류할 필요가 없다는 것이다. 가구소득의 원천이 혼재하는 경우, 근로소득자와 자영업자에 대한 현행 분류기준이 '가구 소득비중의 얼마' 하는 식의 자의적 기준을 따라 결정되는 등 명쾌하지 않은 점을 감안할 때 소득원천에 따른 가구분류가 문제점이 있을 수 있다는 것이다. 그러나, 자영업자의 분류기준에 따른 실증결과면에서의 그 영향력은 미미한 것으로 나타났다.

으로부터의 실현된 소득의 오차에 초점을 맞추고 있다는 측면에서 그 시각이 상이하다.

Pissarides and Weber(1989)의 정교한 방법론과 Lyssiotou, Pashardes, and Stengos(2004)의 보완에도 불구하고, 기존 방법론들은 다음의 몇가지 문제점을 여전히 내포하고 있다. 첫째는 기존의 방법론과 데이터로는 소득의 변동성과 소득축소보고경향에 대한 정보의 혼재를 분리할 수 없어 축소보고 규모의 구간추정만이 가능하다는 것이다. 둘째, 이와 같은 구간추정도 자영업자의 소득의 축소보고율이 일정하다는, 다소 비현실적인 가정하에서 가능하다는 것이다. 현실적인 상황에서는 자영업자가 탈루를 목적으로 소득을 축소해서 보고한다면, 소득의 변동성에 따라 축소보고율을 조정할 것이라는 것이다. 즉, 기존연구들에서 가정하는 것처럼 축소보고율을 유지하는 것이 아니라 축소된 보고소득의 액수를 일정하게 유지한다는 것이다. 예를 들어, 평소 실제소득 1억원 중 8천만원의 소득을 보고한 자영업자가 경기변동 등에 따라 특정연도에 실현된 소득이 1억 2천만원이라면, 축소보고율을 20 퍼센트로 유지하는 대신 축소율을 높여 평시 보고된 소득 8천만원 정도를 유지하여 보고한다는 것이다. 이 경우, 기존 연구에서 추정된 구간은 그 크기가 더 커질 것이며, 이에 대한 논의는 다음 절에서 자세히 다루기로 한다. 셋째, 만일 직업별 소득변동성의 크기가 경기변동에 따라 다르다면, 특정연도의 횡단면 자료를 사용하는 기존 연구들에서 추정된 구간은 그 신뢰도가 감소될 수 있다.

본 연구에서는 한국의 패널자료를 이용하여 기존의 연구방법론을 발전시켜 조세회피 등의 목적으로 실행되는 자영업자의 소득축소보고 정도를 추정하려 한다. 먼저 방법론 측면에서 패널자료를 활용해 소득의 변동성과 소득의 축소보고 경향에 대한 정보를 분리할 수 있음을 보인다. 나아가 패널자료를 이용함으로써 기존의 횡단면분석자료를 이용한 연구와는 달리 자영업자의 소득축소보고 정도에 대한 구간추정이 아니라 보다 정확한 하나의 추정치를 제시한다. 또한, 본 연구는 소득축소보고경향의 원인이 국세청의 자영업자 소득에 대한 정보의 불완전성이라는 직업특성에 따라 조세회피 등을 목적으로 하는 경제적 유인에 의한 것인지 또는 비근로 소득 직종을 선택한 개인들의 특성에 의한 것인지도 검증한다. 한편, 기존의 연구들이 소비를 소득에 회귀분석하는 축약추정식을 채택함으로써 수요추정식의 이론적인 정합성의 미비가 우려되는 반면, 본 연구에서는 재화간의 상대가격 통제 등 이론적 정합성을 갖춘 AIDS 수요추정식을 사용한다는 점에서 기존 연구들과 차별된

다.

본 고의 구성은 다음과 같다. 제 II장에서는 식료품 앵겔곡선 추정방정식을 발전시키고, 직종별 앵겔곡선 이동의 추정과정을 설명하며, 제 III장에서는 실증분석의 자료인 노동패널자료를 소개하고, 실증분석에 사용된 변수들의 기술통계량 및 기간 내 변화를 요약한 후, 추정된 자영업자의 소득 축소보고율을 논의한다. 끝으로 IV 장에서는 연구결과를 간략히 요약하고, 결과의 시사점과 추후 연구방향을 제시한다.

II. 식료품 앵겔곡선 추정방정식

본 연구는 Hamilton (2001a) 에서 사용된 Deaton and Muellbauer (1980) 의 AIDS (Almost Ideal Demand System) 수요추정식을 이용한다. 본고에서 사용된 수요추정식은 개별 가구의 식료품비 지출비중과 식료품의 비식료품에 대한 상대가격 및 실질소득과의 관계를 나타내는 식으로 앵겔곡선으로 해석될 수 있다.

$$w_i = \phi + \gamma(\ln P_F - \ln P_N) + \beta \ln y_i^P \mathbf{X}'\theta + \epsilon_i, \tag{1}$$

w_i 는 가구 i 의 지출 중 식료품 비중을 나타내며, P_F 와 P_N 는 식료품, 비식료품 가격수준을 나타낸다. y_i^P 는 가구의 실질항상소득, X 는 가구원수 등 식료품 예산비중에 영향을 미칠 수 있는 개별가구의 특성변수들의 벡터를 나타낸다. ϵ_i 는 확률오차를 나타낸다. 기존 연구는 t 년도의 보고된 소득 y_{it}^* 는 항상소득에 비해 두가지 오차가 있을 수 있으며, 이를 다음의 식으로 나타낸다.³⁾

$$\begin{aligned} y_{it} &= g_{it} y_i^P, \quad y_{it} = k_{it} y_{it}^* \\ \Leftrightarrow \ln y_{it}^* &= \ln g_{it} + \ln y_i^P - \ln k_{it} \end{aligned} \tag{2}$$

y_{it} 는 t 년도의 실현된 소득이다. 보고된 소득과 항상소득의 차이는 항상소득과

3) 사용된 식 (3)-(9)는 Pissarides and Weber (1989)와 Lyssiotou, Pashardes, and Stengos (2004)의 기본모형에 따른다.

실현된 소득의 차이와 실현된 소득과 보고된 소득의 차이로 나눌 수 있다. 이러한 차이를 가져오는 첫번째 비례적인 오차는 특정년도의 소득의 변동성을 나타내는 g_{it} 값이 1보다 크면 그 해의 실현된 소득이 항상소득보다 큰 것을 나타낸다. 소득축소 보고 경향을 나타내는 k_{it} 값은 자영업자의 경우 1보다 큰 값을 가질 것으로 예상되며, 두 변수 g_{it} 와 k_{it} 는 기존 연구들과 동일하게 다음과 같은 log-normal 분포를 따른다고 가정한다.

$$\begin{aligned}\ln k_{it} &= \mu_k + v_{it} \\ \ln g_{it} &= \mu_g + u_{it}\end{aligned}\quad (3)$$

(2) 식에서 t 년도의 실현된 소득의 로그값인 $\ln y_{it}$ 는 참값인 $\ln y_i^P$ 에 확률오차를 더한 고전적인 측정오차(임시소득)를 포함한 값으로 해석될 수 있으며, 보고된 소득은 그 값에 비고전적인 측정오차의 일종인 Mean-Reverting 측정오차를 더한 것으로 해석될 수 있다. (2) 식과 (3) 식을 (1) 식에 대입하면

$$w_i = \phi + \gamma(\ln P_F - \ln P_N) + \beta \ln y_i^* + \beta(\mu_k - \mu_p) + \beta(v_i - u_i) + \mathbf{X}'\theta + \epsilon_i. \quad (4)$$

자영업자의 축소보고를 추정하기 위해 자영업자의 더미변수를 사용한 t 년도 자료의 엔겔곡선 회귀방정식은 다음과 같다.

$$w_i = \phi + \gamma(\ln P_{Ft} - \ln P_{Nt}) + \beta \ln y_{it}^* + \delta D_{it} + \mathbf{X}'\theta + \epsilon_i \quad (5)$$

D_{it} 는 자영업자가구 더미변수이다. 자영업자의 더미변수의 계수는 다음과 같은 정보를 나타낸다.

$$\begin{aligned}\delta &= \beta[(\mu_{kSE} - \mu_{kEE}) - (\mu_{gSE} - \mu_{gEE})] \\ &= \beta[(\mu_{kSE} + \frac{1}{2}(\sigma_{uSE}^2 - \sigma_{uEE}^2))]\end{aligned}\quad (6)$$

하첨자 SE 와 EE 는 각각 자영업자와 근로소득자를 나타낸다. 본 연구에서 추정하고자 하는 자영업자 가구의 평균적인 축소보고율인 \bar{k} 는 log-normal 분포의 특성과 식 (6)을 이용해 다음과 같이 나타낸다.

$$\ln \bar{k} = \mu_{kSE} + \frac{1}{2} \sigma_{vSE}^2 = \frac{\delta}{\beta} + \frac{1}{2} [\sigma_{vSE}^2 - (\sigma_{uSE}^2 - \sigma_{uEE}^2)] \quad (7)$$

식 (7)의 두번째 항목에 대한 추가적인 정보로 다음의 보고된 소득을 활용한 항상 소득의 추정방정식을 활용한다.

$$\ln y_{it}^* = Z' \pi + \zeta_{it} \quad (8)$$

여기서 Z 는 교육이나 연령 등 항상소득을 나타내는 변수를 의미한다. 소득의 오차항(ζ_{it})에는 소득의 변동성과 축소보고 경향을 모두 포함하며, 잔차항의 분산정보는 위의 식 (7)의 정보와 다음과 같이 연계된다.

$$\sigma_{\zeta SE}^2 - \sigma_{\zeta EE}^2 = \sigma_{vSE}^2 + (\sigma_{uSE}^2 - \sigma_{uEE}^2) - 2cov(uv)_{SE} \quad (9)$$

(9) 식을 가지고 각각 하한추정치($\sigma_{vSE}^2 = 0$ 경우)와 상한추정치($\sigma_{uSE}^2 = \sigma_{uEE}^2$ 경우)를 가정하여 자영업자 가구의 평균적인 축소보고율인 \bar{k} 의 구간을 다음과 같이 설정할 수 있다.

$$\begin{aligned} \ln \bar{k} \in & \left[\frac{\delta}{\beta} - \frac{1}{2} (\sigma_{\zeta SE}^2 - \sigma_{\zeta EE}^2) - cov(uv)_{SE}, \right. \\ & \left. \frac{\delta}{\beta} + \frac{1}{2} (\sigma_{\zeta SE}^2 - \sigma_{\zeta EE}^2) + cov(uv)_{SE} \right] \end{aligned} \quad (10)$$

그러나, 계산된 구간은 관측되지 않는 정보($cov(uv)_{SE}$)를 포함하므로, 축소보고율은 소득의 변동성과 관계가 없다는, 즉 축소보고율이 소득변동성 정도와 독립적이라는 가정을 도입해 기존 연구들은 최종적으로 관측 가능한 정보들로 이루어진

구간을 제시하였다.

$$\ln \bar{k} \in \left[\frac{\delta}{\beta} - \frac{1}{2}(\sigma_{\zeta SE}^2 - \sigma_{\zeta EE}^2), \frac{\delta}{\beta} + \frac{1}{2}(\sigma_{\zeta SE}^2 - \sigma_{\zeta EE}^2) \right] \quad (11)$$

여기서는 패널자료를 이용하여 위에서 제시된 비합리적인 가정을 동반한 구간추정 결과를 보완한다. 즉, 동일가구 시계열 자료에서 시간에 대한 평균값을 구해 사용하는 Between Estimation의 적용을 통해 소득의 변동성을 통제할 수 있다. 즉, 시간에 대한 개인소득의 평균값은 다음과 같이 표현된다.

$$\overline{\ln y_{it}^*} = \overline{\ln y_{it}} - \overline{\ln k_{it}} = \ln y_i^P + \overline{\ln g_{it}} - \overline{\ln k_{it}} \quad (12)$$

위에서 $\overline{\ln y_{it}^*}$ 는 $\sum_{t=1}^T \ln y_{it}^* / T$ 이며, 이와 같은 시간에 대한 평균 값은 소득의 변동성의 분산 (위의 고전적인 측정오차해석에서의 확률오차의 분산)을 다음과 같이 축소시킨다.

$$p \lim_{T \rightarrow \infty} \sigma_{u_i}^2 (= \frac{\sigma_u^2}{T}) = 0 \quad (13)$$

즉, T 가 충분히 커서 소득의 변동성을 통제할 수 있다면 자영업자 가구의 평균적인 축소보고율인 \bar{k} 는 소득의 변동성과 축소보고율간의 공분산을 제거하여, 식 (14)와 같이 정리된다.

$$\ln \bar{k} = \mu_{kSE} + \frac{1}{2} \sigma_{vSE}^2 = \frac{\delta}{\beta} + \frac{1}{2} (\sigma_{\zeta SE}^2 - \sigma_{\zeta EE}^2) \quad (14)$$

두 가지의 오차 중, 시간에 대한 평균값은 경기변동 등으로 인한 소득의 변동성을 제거하고, 나아가 소득의 변동성에 따른 축소보고율의 변동도 제거한다. 한편, 자영업자 내에서의 소득축소보고율의 분산은 그대로 유지된다. 따라서, 시간에 대한

평균 값을 적용하는 Between Estimation에서는 축소보고율이 일정하다는 가정 없이도 구간이 아닌, 하나의 추정치를 구하는 것이 가능하다. 이에 대한 실증결과는 다음 절에서 논의한다. 한편, 위의 수요추정식의 추정결과에서 계산된 자영업자의 소득축소보고에 대한 또 다른 해석으로 Hamilton (2001b)에서의 해석을 고려할 수 있다. 즉, 앵겔곡선의 이동원인을 흑인과 백인간의 소비자물가 차이에 귀납하는 Hamilton (2001b)에서와 같이 본 연구에서도 추정된 앵겔곡선의 이동이 자영업자와 근로소득자 두 직업군이 접하는 물가의 상대적 크기의 차이로 해석할 수 있으나, 인종별 거주지분리와 같이 물가지수의 지리적 차이가 확연히 드러나는 Hamilton의 경우와는 달리 두 직업군의 지리적 분포 차이가 명확하지 않은 본 연구에 이러한 해석을 적용하는 것은 개연성이 높지 않을 것으로 보인다.

III. 실증분석

1. 자료

한국노동패널조사(Korean Labor and Income Panel Study: 이하 노동패널)는 국내 패널 중 매년 개인 및 가구의 경제활동을 관찰하기 위해 실시하는 유일한 패널조사로 횡단면 정보와 시계열 정보를 모두 담고 있다. 노동패널은 16개 광역시/도에 거주하는 우리나라의 5,000가구 및 그 구성원을 대상으로 1998년 1차 면접조사를 시작으로 2006년 9차 면접조사까지 진행되었다.

노동패널 자료는 가구를 조사대상으로 하는 가구용 자료와 가구에 속한 만 15세 이상의 가구 구성원을 조사대상으로 하는 개인용 자료로 크게 구분된다. 가구용 자료의 내용은 가구원의 인적 사항, 변동 가구원 관련 사항, 가족관계와 세대간 경제적 자원 교류, 주거상태, 자녀교육과 보육, 가구의 소득과 소비, 가구의 자산과 부채, 가구의 경제상태 등의 내용을 담고 있으며, 개인용 자료는 개인의 경제활동상태, 소득활동 및 소비, 교육 및 직업 훈련, 고용상의 특성, 근로시간, 직무만족 및 생활만족, 구직활동, 노동시장에서의 이동 등의 다양한 내용을 담고 있다. 본 연구에서는 2000년을 기준년도로 하여 2005년까지 6년 간의 노동패널자료를 이용하여 자영업자의 소득 축소보고 크기를 추정한다.⁴⁾

본 절의 실증분석은 노동패널 자료 중 다음의 조건을 만족하는 가구를 분석대상

으로 선정하여 시행하였다. 이렇게 분석대상에 제한을 두어 추출하는 이유는 가구 내 성인의 숫자 및 가구주 또는 배우자의 성별이나 연령 등 가구의 특성에 따른 가구 소비행태의 이질성을 통제하기 위함이다. 첫째, 남성이 가구주이고 여성이 배우자인 가구이며, 20세 이상의 성인이 2명만 존재하는 가구를 선택한다. 이는 현재 우리나라 가구 구조상 이러한 가구가 가장 일반적인 형태를 띠고 있기 때문이다. 둘째, 가구주와 배우자의 연령을 분석기간 내내 만 20세부터 65세까지로 제한하였다. 이는 노인가구와 청소년가구를 제외함으로써 이들 가구를 포함했을 경우 발생 가능한 가구소비행태의 이질성을 통제하기 위함이다. 셋째, 기존의 실증연구에서와 마찬가지로 가구소득 대비 식비 비중이 1% 미만이거나 99% 이상인 비정상적인 가구들은 분석대상에서 제외되었다.⁵⁾ 넷째, 분석기간 중에 가구 구성원의 변화가 있는 가구는 제외하였다. 이는 가구 구성원의 유입과 유출로 인해 가구 소득 및 식비의 변동이 발생할 수 있기 때문이다. 다섯째, 회귀방정식에서 독립변수로 사용될 변수를 생성하는 과정에서 응답이 누락된 경우가 존재하는 가구는 분석 과정에서 제외하였다. 연구결과의 정확도와 다른 실증연구와의 비교도를 높이기 위해 위와 같은 자료의 정제과정을 통해 남은 자료는 6년간 총 4876가구이며, 연 평균 약 800 가구에 대한 정보이다. 본 연구는 실질소득의 산정을 위한 물가정보로 소비자물가지수를 사용한다.⁶⁾ 각각 표본정제조건의 표본대상이 상호간에 배타적이지 않으므로, 본 논문의 실증분석과정에 따라 개별 정제조건의 표본수에 대한 영향을 기술한

4) 1998년에 시작된 노동패널 1차부터 9차년도까지의 자료 중, 2001년 조사된 제4차 노동패널부터 “작년 한해” 식비에 대한 질문을 시행했기 때문에 2000년에 조사된 제3차 노동패널 자료부터 최근 제한적으로 공개된 제 9차 노동패널자료까지를 분석에 포함할 수 있다, 사용된 t 년도의 소득은 $t+1$ 년도에 조사된 “작년 한해” 소득이다.

5) 이는 식비 비중이 과도하게 낮거나 높은 경우 식비나 가구 소득변수에 있어서 측정오차가 존재할 수도 있기 때문이다. 민감성 분석의 일환으로 가구소득 대비 식비 비중이 2% 미만이거나 80% 이상인 개별가구를 제외하여 분석을 시행한 경우에도 추정결과에서 차이가 나타나지 않았다.

6) 자영업자가 조세회피 등을 목적으로 소득을 축소하여 보고한다는 가정은 국세청 자료와 같은 공식 소득자료에서는 쉽게 받아들일 수 있으나, 노동패널과 같은 서베이자료에서는 의문의 여지가 있다. 서베이자료에서의 자영업자의 소득축소보고는 국세청에 보고하는 공식자료와 서베이자료간에 개인들이 보고의 일관성을 유지한다는 가정에서 비롯한다. 설혹 국세청 자료와 같은 공식자료를 이용한다고 해도, 그 세무자료와 결합된 자료가 표본의 대표성을 유지하기 힘든, 제한적으로 가공된 자료라는 점에서 문제가 있다. 또한, 세무조사 대상선정에서의 표본의 대표성 문제도 지적될 수 있다.

다. 먼저, 가구구성원의 변화가 있는 가구는 전체가구의 약 10% 정도로 변화가 없는 가구에 비해 상대적으로 식료품의 비중이 크고, 소득은 작은 것으로 나타난다. 이들 가구는 성인 2인가구조건 등 다른 조건에 의해 배제되는 표본과 상당한 상관관계를 보인다. 가구구성의 변화가 없는 가구들로 구성된 표본크기는 2005년도 기준으로 소득 등 주요변수 값을 보유해 실증분석이 가능한 상태에서 1978가구이다. 이중 가구주와 배우자의 연령제한, 성인 2인가구로의 제한, 가구주의 성별제한 등 인구동학변수 조건을 통해 1134가구로 축소된다. 나아가 제시된 식료품비중의 구간조건으로 1117가구로 조정되며, 최종적으로 이전소득(사회보험소득 등) 수혜가구 배제조건에 의해 694가구로 축소된다. 이중 식료품비중조건과 이전소득 수혜의 유무조건에 대해서는 민감성분석을 시도하였다.

본 연구에서 사용하는 변수들의 기술통계량은 <표 1>에서 볼 수 있다. 이 변수들을 간략히 설명하자면, 회귀분석의 종속변수인 연도별/가구별 식료품 예산비중은 월간 평균 식료품비에 12개월을 곱한 연간 식료품비를 가구의 연간 총소득으로 나눈 값이다. 자료기간인 2000년에서 2005년까지의 식료품비 지출비중 평균은 <표 1>의 첫째 행과 같이 약 23.8 %이며, 동일한 방식으로 구해진 외식비 예산비중의 6년간 평균은 약 3.6%이다. 가구별 소득은 근로소득과 금융소득 등 기타 각종 소득을 모두 합한 것으로 평균은 약 3천 4백만원이다. 가구주의 평균연령은 41.2세이고, 표준편차는 6.49세이며, 교육연한은 12.73년이다. 배우자는 가구주에 비해 연령에서 약 3년, 교육연한에서는 약 1년 정도 낮은 수준을 보이고 있다. 가구주의 연간 노동시간은 약 2700시간, 배우자는 이의 절반 이하의 노동시간을 보이고 있다. 자영업자의 비율은 약 33.5%이며, 자료분석기간 중 연도별 자영업자의 비율은 30%대 초반에서 큰 변동 없이 일정하게 유지되었다.⁷⁾ 표본의 연도별 비중은 2000년의 20.7%에서 표본감모(attrition)로 점차 줄어 2005년에는 13.1%로 낮아진다.

본 연구의 핵심변수인 식료품비 지출비중과 명목 가구소득 변화를 살펴보면, <표 2>의 첫째 행에서 보듯이 식료품비 지출비중은 지속적으로 감소하여 2000년의 약 30%에서 2005년에는 약 19%로 5년 만에 11% 포인트 정도 낮아진다. 가구소득은 지속적으로 상승해 5년간 63%가 상승했으나, 연간 약 3.5%씩 상승한 소비자물가

7) 연도별 자영업자의 비율은 2000년 33.9%에서 2002년의 31.3%로 감소되었다가 다시 상승해 2005년 35.5%를 보이고 있다.

지수를 감안한 실질소득은 약 40% 상승한 것으로 나타난다. 미국이나 러시아 등 다른 나라의 실증연구에서 실질소득이 약 10% 상승하면 식료품비 지출비중은 약 1~1.5% 포인트 정도 하락하는 것을 감안할 때, 우리나라의 경우 관측된 실질소득의 증가에 비해 상대적으로 더 큰 식료품비 지출비중의 하락을 경험한 것으로 나타나고 있다. 이는 실제 실질소득의 변화가 관측된 것보다 더 크게 증가했을 가능성이 있음을 의미하며 그 원인을 소비자물가지수의 과대계상 등에서 찾을 수 있다.

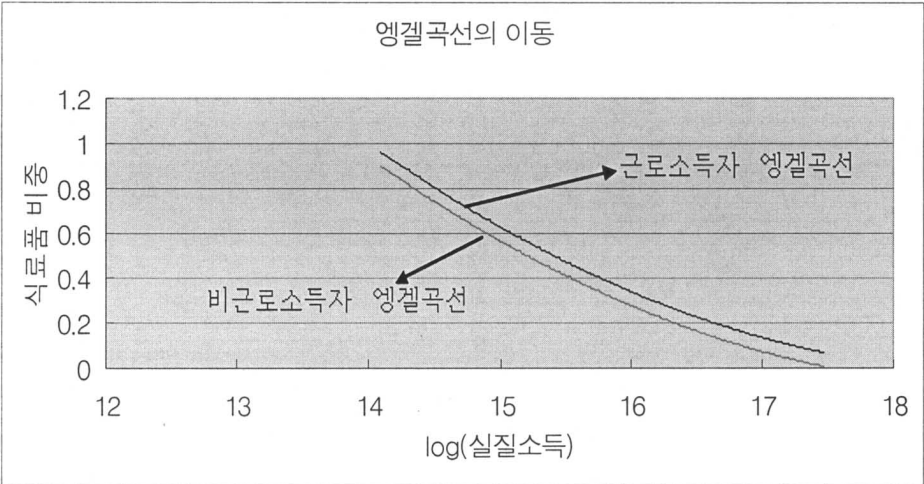
〈표 1〉 기술통계량 (노동패널, 2000-2005), 표본 수=4867

변수	평균	표준편차	최소값	최대값
w (식료품비 지출비중)	.2386	.1059	.0132	.9
X_{res} (외식비 지출비중)	.0362	.0369	0	.347
$\ln(Y)$ (로그 변환된 가구소득)	17.12	.6328	13.12	19.85
$\ln(Y/P)$ (로그 변환된 실질소득)	17.05	.6210	13.12	19.78
가구의 연령	41.22	6.49	23	65
배우자의 연령	38.26	6.38	20	65
가구의 교육 연한	12.73	2.99	0	25
배우자의 교육 연한	11.91	2.62	0	25
가구의 연간 노동 시간	2762.52	823.80	0	8400
배우자의 연간 노동 시간	1218.60	1404.87	0	8400
자영업자 비율	.3350	.4722	0	1

두 직업군의 연도별 식료품비 비중과 실질소득을 살펴보면, 보고된 평균 실질소득은 자료분석기간 중 근로소득자가 자영업자보다 많은 것으로 나타나는 반면, 실질소득을 반영하는(식비 등 지출항목에서는 보고 오차가 없을 것이라는 가정하에) 지출 중 식료품비 비중은 오히려 자영업자가 큰 것으로 나타나고 있다. 소비자물가지수의 측정오차 추정과 유사한 방식으로 이루어지는 자영업자의 소득축소보고율 추정은 다음 절의 실증분석에서 자세히 설명할 것이나, 그 전에 〈표 2〉의 시사점을 engel곡선의 직업군별 이동을 나타내는 그래프로 보다 직관적으로 설명할 수 있다. 직업군별 engel곡선의 행태는 자료분석기간내에 유사한 양태를 보이므로, 그 중 한 해인 2003년의 식료품비 지출비중과 실질소득의 관계를 나타내는 engel곡선을 직업군별로 각각 추정하면, 〈그림 1〉에서 보듯이 비근로 소득자의 engel곡선이 좌측으로 이동함을 관측할 수 있다. 따라서 동일기간 내 소비자들의 선호에 변화가 없었다고

가정하면, 이 곡선의 좌측 이동한 정도가 자영업자 명목소득의 축소보고 정도를 나타낸다고 해석할 수 있다.

〈그림 1〉 추정된 엔겔곡선의 좌측 이동(노동 패널, 2003년 자료)



〈표 2〉 주요 변수의 직종별 자료기간내의 변화 (노동패널, 2000-2005), 표본 수=4867

변수	근로소득자					
년도	00	01	02	03	04	05
w (식료품비 지출비중)	.306	.265	.235	.220	.204	.193
X_{res} (외식비 지출비중)	.040	.041	.038	.031	.033	.033
$\ln(Y/P)$ (로그변환된 실질소득)	16.87	17.07	17.10	17.19	17.29	17.38
	자영업자					
년도	00	01	02	03	04	05
w (식료품비 지출비중)	.291	.240	.212	.206	.195	.182
X_{res} (외식비 지출비중)	.035	.039	.036	.029	.029	.031
$\ln(Y/P)$ (로그변환된 실질소득)	16.84	17.13	17.12	17.24	17.26	17.30

2. 실증분석

엔겔곡선의 이동을 이용하여 자영업자의 소득 축소보고를 추정하는 본 연구는 기본적으로 소비자들의 식료품/비식료품에 대한 선호가 근로소득자나 자영업자나 다

르지 않다고 가정한다. 이 가정은 소비자 물가지수의 측정오차를 연구하는 관련 연구들에서 소비자들의 식료품/비식료품에 대한 선호가 일정한 기간 내에 변하지 않는다는 가정과 그 맥락을 같이한다. 소비자물가지수의 측정오차에 대한 연구와 달리 본 연구는 횡단면 자료를 사용하여서도 가능하며, 실제로 소비와 소득의 관계식을 통해 자영업자의 축소보고문제를 접근하는 기존의 실증연구들은 횡단면자료를 사용하고 있다. 그러나 동일가구의 정보를 시계열자료의 형태로 수집·기록하는 패널자료를 사용하면, 자영업자의 연도별 소득변동성에 따른 상품소비선호의 변동을 통제할 수 있다. 또한 자영업자 소득의 축소보고가 직종의 특성에서 비롯된 것인지, 아니면 자영업을 선택한 개인에게 내재한 축소보고 경향의 분포가 근로소득자의 그 분포와 다름에서 비롯한 것인지를 여부에 관해서도 정보를 얻을 수 있다. 더 나아가, 시간에 따른 가구구성의 차이를 통제함으로써 직업군별 식료품/비식료품에 대한 선호가 동일하다는 가정과 가구별 상품소비의 시간에 따른 불변성의 가정을 보다 쉽게 유지할 수 있을 것이다. 본 연구는 OLS 추정방법과 더불어 가구별 고정효과(Fixed Effects) 방법을 추가하여 식 (5)의 선형모형을 추정하였다.

〈표 3〉은 식료품 Engel곡선의 회귀분석결과를 요약하고 있다. 〈표 3〉 (1) 열의 6년간의 자료를 결합한 Between Estimation (BE) OLS 추정방법에 따른 회귀분석결과를 식 (14)에 적용한 결과와 연도별 횡단면분석에 따른 구간추정 결과는 〈표 4〉에서 볼 수 있다. 먼저 Between Estimation 결과를 이용하여 구한 보고소득 대비 항상소득의 추정치는 1.49로 항상소득이 보고소득의 1.49배이며, 이를 축소보고율로 전환하면 자영업자가 평균적으로 33.2%의 소득을 근로소득자에 비해 축소 보고하는 것으로 나타난다. 축소보고율의 상세한 계산과정은 다음과 같다. Between Estimation의 추정결과를 예로 들어 설명할 때 식 (14)에서 전반부인 δ/β 의 추정비율은 〈표 3〉에서의 각각 추정된 파라미터들의 비율인 $0.33 (= \hat{\delta}/\hat{\beta})$ 이며, 식 (14)의 후반부인 $0.5(\sigma_{\zeta SE}^2 - \sigma_{\zeta EE}^2)$ 는 식 (8)의 잔차항의 분산을 자영업자와 근로소득자 집단별로 나누어 구해진 $0.5 * (0.59 - 0.44)$ 이다. 횡단면 분석의 연도별 결과는 후반부를 전반부에서 더하고 빼서 나온 구간추정치이다.

동일한 방법론을 이용하여 연도별로 나누어 횡단면 분석을 하면, 추정된 소득 축소보고율이 최저 20.8% (2001년)에서 최고 49.9% (2002년)까지 이르는 등, 시간에 따른 변동이 매우 큰 것으로 나타난다. 연도별 횡단면 분석에서 추정된 소득축소보

고율의 연도별 변동은 식 (7)에서 표현되었듯이 연도별 소득의 변동성이 구간 추정
에 포함된 결과이다. 축소보고율을 일정하게 유지할 것이라는 가정에 따른 추정구
간 자체도 연도별로 적지 않은 소득변동성 및 이로 인한 축소보고율의 변동을 감안
할 때 그 신뢰도는 더욱 낮아질 수 밖에 없다.

〈표 3〉 식료품 앵겔곡선 추정 (노동패널, 2000-2005), 표본 수=4867

변수	(1) Between OLS 노동패널 (2000-05)	(2) 고정효과 노동패널 (2000-05)
Intercept	1.3468 (.0631) ***	.6786 (.0809) ***
Log (실질소득)	-.0589 (.0039) ***	-.02584 (.0037) ***
Log (Food CPI/Non-food CPI)	-.7060 (.05705) ***	-.7652 (.0366) **
가구주 교육년한	-.0030 (.0009287) ***	-.0038 (.0027)
배우자 교육년한	-.0014 (.0010)	-.0014 (.0010)
가구주 연간 근로시간	-6.09e-10 (2.75e-09)	-4.46e-11 (2.26e-09)
배우자 연간 근로시간	-7.95e-10 (1.62e-09)	2.13e-10 (1.55e-09)
가구내 아동 수	.0058 (.0023) ***	.0096 (.0037) ***
외식비 지출비중	-.2175 (.0619) ***	-.1072 (.0468) **
자영업자 더미	-.0195 (.0044) ***	-.0065 (.0065)
R ²	.2656	.1753

참조: ***, **, * 는 각각 1%, 5%, 그리고 10% 기준으로 통계적으로 유의함을 나타낸다. 노동패널
은 16개 광역시도의 지역적 정보를 가진다.

식 (14)에서 보듯이 소득의 변동성을 통제하는 Between Estimation의 보다 정확
한 추정치는 연도별 횡단면 분석에서의 중간값이 아니라 상한값(Upper Bound)에
해당하는 결과가 주목된다. 이는 결국 Between Estimation의 추정치가
($\sigma^2_{uSE} - \sigma^2_{uEE}$)인 경우가 되어 소득변동성이 제거됨을 의미한다. 따라서 위와 같은

접근방법은 기존연구에서 소득변동성에 따른 소비행태의 이질성 (preference heterogeneity) 을 소득의 축소보고라고 잘못 식별할 수 있다는 문제점을 직접적으로 해결한다. 즉, 자영업자의 고정소득은 생필품 등에 대한 지출에 주로 사용하고 사치품 구입은 상대적으로 변동성이 큰 사업소득에 의존한다는 문제를 소득의 변동성을 통제하여 해결한 것이다.

한편, 자가(自家) 자영업자의 경우 외식비 등을 사업비용으로 처리하는 관행과 같이 종속변수인 식료품비 지출비중에 영향을 미치는 잠재적인 소비행태의 이질성 (preference heterogeneity) 을 통제하기 위해 본 연구는 회귀분석식에 외식비를 추가적인 설명변수로 넣어 분석한다. 자영업자 더미계수의 추정치와 실질소득의 계수추정치는 각각 통상적인 유의도상에서 통계적으로 유의함을 볼 수 있다.

〈표 3〉 (2) 열의 고정효과(Fixed Effects) 모형 회귀분석결과를 이용하여 구한 소득의 축소보고는 25.3%로 Between Estimation 추정치보다 약간 낮은 것으로 나타났다. 만약 소득 축소보고에 대한 개인별 성향의 분포가 근로소득자나 자영업자나 동일하고, 축소보고는 자영업자의 직업군별 특성에서 비롯한다고 하면, 고정효과의 추정치나 횡단면분석 추정치나 유사한 크기를 가질 것이다. 반면에, 축소보고 경향이 큰 사람들이 자영업을 선택하고 근로소득업은 그 반대의 사람들이 선택하며 직종의 전환과정에서 개인적 성향은 일정하게 유지된다면, 고정효과의 추정치는 횡단면분석 추정치에 비해 크게 작은 값으로 도출될 것이다. 본 연구의 추정결과는 소득축소보고가 자기선택(self selection) 보다는 자영업자의 직업군별 특성에서 비롯됨을 시사한다. 본문에서 제시한 고정효과모형 도입이유에 추가하여 Hausman 검정방법을 통해, 랜덤효과(Random Effects) 모형보다 고정효과모형이 적합함을 확인하였다.

〈표 4〉 보고소득 대비 항상소득의 추정 구간(노동패널, 2000-2005), 표본 수=4867

	2000	2001	2002	2003	2004	2005	BE 추정치 식 (14)
\bar{k} 구간 추정 (하한, 상한), 소득 축소보고율 구간 중간값	(1.41, 1.63) .419	(1.14, 1.32) .208	(1.56, 1.73) .499	(1.18, 1.35) .237	(1.14, 1.43) .248	(1.26, 1.50) .320	1.49 .332

IV. 결론 및 시사점

본 연구는 Hamilton (2001a)의 앵겔곡선 추정방법론을 이용하여 우리나라 자영업자의 소득축소보고 정도를 추정하였다. 이 추정을 위해 본 연구는 2000년에서 2005년까지 6년간의 노동패널을 사용하였으며, 기존 실증연구에서 채택하는 방법론의 문제점을 보완하여 분석한 결과, 평균적으로 자영업자의 항상소득이 보고소득의 약 1.5배에 해당함을 추정할 수 있었다. 또한, 고정효과모형을 통해 자영업자의 축소보고 경향이 자영업자의 개인적인 특성이 아니라 직종의 특성에서 비롯한다는 것을 알 수 있었다. 한편, 조세회피 목적이 아닌 명목소득 자체의 측정오차 등에 관한 연구는 노동패널자료 자체의 정확도에 대한 연구와 함께 추후 병행되어야 할 것이다.

■ 참 고 문 헌

1. 유일호, “우리나라의 탈세규모 추정: 소득세와 부가가치세,” 『재정논집』, 제9집, 1995, pp. 39-66.
2. 이철인, “패널자료를 이용한 탈루규모의 추정,” 『공공경제』, 3(1), 1998, pp. 67-96.
3. 정 철 · 김봉근 · 박명호, “한국의 소비자물가지수 편차 추정과 국제물가비교에 대한 시사점,” 『대외경제연구』, 11(2), 2007, pp. 261-84.
4. Beatty, T. and Larsen, E., “Using Engel Curves to Estimate Bias in the Canadian CPI as a Cost of Living Index,” *Canadian Journal of Economics*, 38(2), 2005, pp. 482-499.
5. Deaton, A., and Muellbauer, J., “An Almost Ideal Demand System,” *American Economic Review*, 70(3), 1980, pp. 312-326.
6. Gibson, J., Stillman, S. and Le, T., “CPI Bias and Real Living Standards in Russia,” *Journal of Development Economics*, forthcoming, 2007.
7. Hamilton, B., “Using Engel’s Law to Estimate CPI Bias,” *American Economic Review*, 91(3), 2001a, pp. 619-630.
8. Hamilton, B., “Black-White Difference in Inflation: 1974-1991,” *Journal of Urban Economics*, 50(1), 2001b, pp. 77-96.
9. Pissarides, C. and Guglielmo W., “An Expenditure-Based Estimate of Britain’s Black

Economy," *Journal of Public Economics*, 39(1), 1989, pp. 17-32.

10. Lyssiotou, P., Pashardes, P. and Stengos T., "Estimates of the Black Economy Based on Consumer Demand Approaches," *The Economic Journal*, 114(497), 2004, pp. 622-640.

Shift of Engel Curve for the Self-Employed and Income Under-Reporting

Bonggeun Kim^{*} · Chul Chung^{**} · Myung-Ho Park^{***}

Abstract

This paper adapts an Engel curve methodology developed by Hamilton (2001b) for estimating income under-reporting by the self-employed. We examine discrepancies between food shares and reported incomes of the self-employed and other households. With a panel data methodology, we differentiate income under-reporting and transitory income fluctuations of the self-employed. Using the KLIPS data, the under-reporting rate of the self-employed is 33.2%.

Key Words: income, under-reporting, self-employed

^{*} Associate Professor, School of Economics, Sungkyunkwan University

^{**} Research Fellow, Korea Institute for International Economic Policy

^{***} Research Fellow, Korea Institute of Public Finance