

소득세 신고자료를 활용한 최상위 소득계층의 소득집중도 추정*

박 명 호**

논문 초록

본 연구는 근로소득세와 종합소득세 통계의 중복 집계 문제가 해소된 국세청 신고 자료를 활용하여 2007~2012년 동안 우리나라 최상위계층의 소득집중도 추이를 새롭게 추정하였다. 또한 본 연구에서는 파레토 분포를 따른다는 기존 연구의 가정을 사용하지 않고 Atkinson (2005)의 Mean Split Histogram 방법론을 사용하였다. 그 결과, 최상위계층의 소득집중도는 지니계수나 5분위 배율 같은 기존 소득분배지표와는 달리 글로벌 금융위기 이후 소득분배가 더 악화되었음을 보여준다. 한편 The World Wealth and Income Database에 우리나라의 수치로 수록된, 자료의 제약이 심한 기존 연구의 추정치는 소득계층의 범위가 확대될수록 본 연구의 결과와 무시할 수 없는 차이를 보였다. 이는 국세청의 신고자료를 통해 신뢰할 수 있는 최상위 계층의 소득집중도 추정치를 생산하여 전통적인 소득분배지표를 보완하여야 함을 시사한다.

핵심 주제어: 소득분배, 소득집중도, 소득세 신고자료

경제학문헌목록 주제분류: H24, D31

투고 일자: 2016. 6. 29. 심사 및 수정 일자: 2016. 7. 12. 게재 확정 일자: 2016. 7. 18.

* 이 논문은 한국조세재정연구원의 2014년도 기본연구과제인 『소득분배 변화와 정책과제: 소득집중도와 소득이동성 분석을 중심으로』를 바탕으로 한다.

** 한국조세재정연구원 선임연구위원, e-mail: ecpmh@kipf.re.kr

I. 서론

1997년 외환위기 이후의 소득격차 확대는 우리사회 전체의 소득분배 상태 및 그 추이에 대한 많은 관심을 불러왔다. 특히 글로벌 금융위기 이후 소득격차 확대에 따른 미국내 집단 간 갈등¹⁾이 전세계적인 주목을 받음에 따라 우리나라에서도 소득불평등 심화에 대한 우려가 크게 높아진 상태이다. 이에 우리사회의 소득분배 상태 및 그 추이를 정확하게 파악하는 작업이 중요해졌다.

우리사회의 소득분배의 상태에 관한 대표적인 지표는 통계청에서 발표하는 소득불평등 지표인 지니계수 및 5분위배율이다. 글로벌 금융위기 이후 국내 경기가 좀처럼 회복되지 않음에 따라 국민들이 체감하는 소득불평등은 심화되고 있음에도 통계청의 소득불평등 지표들에서는 이러한 현상을 발견하기 힘든 상태이다. 예를 들면, 가처분소득²⁾ 기준 지니계수는 전체 가구, 2인 이상 비농가 가구, 도시 2인 이상 가구 모두에서 2008년 또는 2009년을 정점으로 하락하는 모습을 보이고 있어서 소득불평등이 완화되고 있음을 보여준다. 또한 가처분소득 기준 5분위 배율 지표에서도 이와 유사한 추이를 보여주고 있다. 이는 소득불평등 지표가 소득분배 실태를 제대로 반영하지 못할 가능성이 있음을 시사한다.³⁾

기본적으로 통계청의 소득불평등 지표는 통계청이 행하는 가계동향조사를 사용하여 작성된다. 가구의 설문 응답에 의존하는 가계동향조사의 특성상 고소득 계층의 정보가 충분하게 반영되지 않을 우려가 존재한다. 그리고 설문 응답시 자신의 실제 소득보다 과장하거나 과소하게 응답할 가능성도 있다. 따라서 기존의 소득불평등 지표의 현실 반영도를 높이는 방안에 대한 필요성이 제기되곤 한다.⁴⁾

1) 그 대표적인 예가 2011년 9월 17일부터 미국 뉴욕의 월가에서 시작된 월가시위(‘월가를 점령하라(Occupy Wall Street)’라고 불림)이다. 월가시위는 글로벌 금융위기를 발생시킨 월가 금융기관들이 오히려 상당한 보너스를 주요 임원들에게 제공하는 등 도덕적 해이가 심각하였고, 일반 국민들은 주택을 압류 당하는 등 빈부격차가 심화된 상황 속에서 발생하였다. 월가시위의 대표적인 구호는 ‘We are the 99%’로 이는 최고 부자 1% 대 나머지 99%의 미국인 간의 소득 및 부의 불평등에 대한 시위자들의 불만을 표현한 것이다. 월가시위는 곧 미국의 주요도시로 확산되었고 전 세계에서 유사한 형태의 시위를 유발하였다. 이는 소득불평등 심화가 전 세계적인 현상임을 시사한다.

2) 가처분소득은 시장소득과 공적이전소득을 합한 금액에서 공적 비소비지출을 차감한 것이다. 동변수들의 정의 및 지니계수, 5분위배율에 대한 정의는 통계청(2011)을 참조하기 바란다.

3) 물론 소득불평등 지표가 사실인 반면 사람들이 체감하는 것이 현실과 다를 수도 있다.

4) 통계청(2008) 및 김낙년·김종일(2013)을 참고하기 바란다.

한편 앞서 언급한 바와 같이 전 세계적으로 소득불평등에 대한 우려가 확산되는 가운데 외국에서는 2000년 초반 이후 소득세 통계를 활용한 최상위 소득계층의 소득집중도를 추정한 후 소득불평등의 추이를 조사하는 연구가 활발하게 진행되고 있다. 물론 우리나라의 경우도 박명호(2012), 김낙년(2012) 및 Kim and Kim(2014)을 통해 최상위 소득계층의 소득집중도에 대한 분석이 이루어졌다. 박명호(2012)는 국세통계연보의 종합소득세 통계와 근로소득세 통계를 단순 결합하여 2006~2010년 동안 과세대상 소득이 1억원 이상에 해당하는 특정 분위수(2006년 1%, 2007년 1.2%, 2008년 1.1%, 2009년 1.3%, 2010년 1.6%)의 소득집중도를 산출하였다. 그리고 통계청의 가계금융조사 자료와 한국조세재정연구원의 재정패널조사 자료를 활용하여 상위 1%의 소득비중을 추계한 후 고소득 가구의 과소 표집 가능성을 제기하였다. 그리고 다른 나라에서 최상위 소득계층의 소득집중도 추정을 위해 사용되고 있는 연구방법론을 적용한 보다 정교한 연구의 필요성을 제시하였다.

김낙년(2012) 및 Kim and Kim(2014)은 최상위 소득계층의 소득집중도 추이에 관한 해외의 연구방법론을 본격적으로 수용하고 분석기간을 가능한 최대로 연장한 연구라는 의의가 있다. 그러나 이들의 연구들은 우리나라의 소득분포 상단에서도 파레토 분포를 따르는 가에 대한 검증과정을 거치지 않고 파레토 보간법을 활용한 한계가 있다. 또한 Atkinson(2007)에서 언급한 것처럼 소득구간의 수가 매우 한정된 국세통계연보 상의 자료를 활용하여 최상위 소득계층의 소득집중도 추이를 추정하여 정확도가 떨어질 가능성이 높다. 특히 우리나라의 경우 2000년대 중반 이전 국세통계연보에 제공된 정보는 일관성이 떨어지고 정보의 양이 빈약하여 매우 자의적인 가정이 도입될 수밖에 없는 상태이다. 2000년 중반 이후의 국세통계연보 상의 통계도 근로소득세 통계와 종합소득세 통계의 중복 문제가 존재하며, 이를 해결하는 과정에서 근거가 부족한 가정을 사용해야 하는 한계가 있다.⁵⁾ 이 밖에도 분석대상 기간을 과거로 확대하려는 과정에서 조세제도 및 과세행정의 변화를 반영하지 못한 측면도 있다.

이에 본 연구에서는 최상위 소득계층의 소득집중도 추이를 보다 객관적으로 추정함으로써 우리나라의 소득분배 구조에 대한 이해를 높이하고자 한다. 특히 Cho, Park, Phillips(2016)의 연구에 따르면, 우리나라의 경우 상위 0.5%보다 더 낮은 소득구간(예: 상위 1%, 상위 5% 등)에서는 파레토분포를 따른다는 가설이 기각된다고 한

5) 특히 근로소득세 통계와 종합소득세 통계의 중복 문제를 해결하는 과정, 국민계정 자료를 통해 개인부문의 소득을 산출하는 과정 등에서 충분한 근거를 가지고 있지 못한 가정들을 사용하는 한계가 있다.

다. 이에 본 연구에서는 Atkinson(2005)이 사용한 Mean Split Histogram 방법을 사용하여 최상위 소득계층의 소득집중도를 추정한다. 그리고 본 연구에서는 기존 국세통계연보에서 제공되는 통계를 사용하지 않고, 국세통계연보 상의 근로소득세 통계와 종합소득세 통계의 중복 문제를 해결한 국세청의 통합소득세 신고자료를 사용한다.⁶⁾ 다만, 동 자료는 2007~2012년 동안에 대해서만 제공됨에 따라, 본 연구는 최상위 소득계층의 소득집중도 추이를 분석하는 기간이 짧은 한계를 지닌다. 한편 전체 인구(the control population)로는 단순히 20세 이상으로 하는 것이 아니라 영국의 경우처럼 15세 이상의 인구, 병역의무를 감당한 생산가능인구 등을 사용하여 소득집중도 추이에 대한 민감도 분석도 함께 수행한다.

본고의 구성은 다음과 같다. 아래의 제Ⅱ절에서는 본 연구에서 사용하고 있는 분석자료와 분석방법에 대하여 기술한다. 그리고 제Ⅲ절에서는 2007~2012년 동안의 최상위 소득계층의 소득집중도 추정 결과를 제시한다. 마지막으로 제Ⅳ절에서는 앞서의 논의를 요약하며 글을 마치고 있다.

Ⅱ. 분석자료 및 분석방법

1. 분석자료

특정 연도의 최상위 소득계층의 소득집중도 분석에 필요한 정보는 크게 3가지 유형이다. 첫째, 소득집중도의 분모에 해당하는 전체 소득에 대한 정보가 필요하다. 둘째, 특정 분위수에 해당하는 소득계층의 크기를 알기 위해서는 전체 인구에 대한 정보가 필요하다. 마지막으로 특정 분위수에 해당하는 계층의 소득 규모에 대한 정보도 필요하다.

소득집중도 추정에 있어 분모에 해당하는 전체 소득의 측정은 소득세 과세체계에 따라 크게 두 가지 방법이 사용된다. 첫 번째 방식은 소득세 통계를 이용하여 전체 소득을 측정하는 것이다. 동 방식은 소득세 통계에 대부분의 소득이 포착이 되는 포

6) 국세통계연보의 통계는 주로 납세자가 제출한 세무신고 자료를 집계하여 생산한다. 그 결과 연말정산을 하였지만 여러 가지 이유로 종합소득세 신고를 한 납세자는 국세통계연보 상 근로소득세 통계와 종합소득세 통계에 모두 잡히는 중복 문제를 낳는다. 국세청에서는 최근 이런 중복 문제가 해소된 '통합소득 데이터베이스'를 구축하고 있지만 국세통계연보에는 제공하고 있지 않다. 본 연구는 국세통계연보의 통계가 아닌 '통합소득 데이터베이스'에서 추출한 통계를 사용하여 중복 집계 문제를 방지한다.

팔주의 소득세 과세체계를 지닌 미국 같은 국가를 대상으로 한 분석에서 활용된다. 두 번째 방식은 국민계정을 통해 가계부문으로 귀속되는 소득을 추계하여 전체 소득을 측정하는 것이다. 동 방식은 무신고 소득이나 비과세 소득 등으로 인해 소득세 신고 통계상의 과세대상 소득이 전체 소득을 충분하게 포착하지 못할 수 있는 열거주의 소득세 과세체계를 지닌 국가를 대상으로 한 분석에서 사용될 수 있다.

한국의 소득세 체계는 열거주의적 소득세 체계를 바탕으로 하고 있기 때문에 전체 소득의 추정을 위한 방식으로 국민계정을 통해 접근하는 방식을 사용하였다. 따라서 일본을 분석한 Moriguchi and Saez(2010) 및 우리나라를 분석한 김낙년(2012)의 연구처럼 국민계정의 제도부문별 소득계정에서 가계부문으로 실제 귀속되는 소득을 전체 소득으로 간주하였다. 가계부문으로 귀속되는 소득 유형은 크게 세 가지로 피용자보수, 영업잉여, 재산소득으로 분류된다.

먼저, 피용자보수 유형에서는 한국은행의 경제통계시스템에서 제공하는 제도부문별 소득계정에서 가계부문의 원천 소득 중 피용자보수에서 고용주의 사회부담금을 제외한 임금 및 급여를 사용하였다.⁷⁾ 두 번째, 영업잉여 유형은 김낙년(2012)의 연구처럼 자가 주택 귀속임료를 추정⁸⁾한 뒤 이를 가계부문의 영업잉여에서 제외한 금액을 사용하였다. 마지막으로 재산소득 유형 중 이자소득은 김낙년(2012)의 연구와는 달리 국세통계연보의 원천세 통계상의 이자 소득금액을 이용하였다. 이 방법은 제도부문별 소득계정 재산소득 중 이자소득에 포함된 금융중개서비스를 제거하기 위해 필요한 인위적인 가정을 불필요하게 해 주는 장점이 있다. 나머지 재산소득은 국민계정의 제도부문별 소득계정에서 가계부문으로 귀속되는 재산소득을 사용하였다. 아래의 <표 1>은 이상과 같은 방식으로 추계된 전체 소득의 규모를 그 유형별 규모와 함께 제시해 주고 있다.

7) 한국은행 경제통계시스템에서는 가계부문별 소득계정에서 피용자보수의 구성 항목으로 임금 및 급여와 고용주의 사회부담금을 제시하고 있다. 또한, 고용주의 사회부담금은 실제사회부담금과 귀속사회부담금으로 구성되어 있다.

8) 통계청의 인구총조사로부터 점유형태별 가구 통계를 이용하여 연도별 자가 주택 비율을 추산하였고, 이 비율을 각 연도별 산업연관표의 주거서비스 영업잉여에 곱하여 자가 주택 귀속임료를 추정하였다. 한편 인구총조사는 5년마다 작성되지만 사이 연도에 대해서는 동일한 형태의 자료가 없기 때문에 조사시점 사이 연도에 대해서는 선형보간을 통해 자가주택비율을 추산하였다.

〈표 1〉 전체 소득의 규모 및 그 구성

(단위: 조원, %)

	전체 소득	비용자보수 유형		영업잉여 유형		재산소득 유형	
	금액	금액	비중	금액	비중	금액	비중
2007년	560.5	394.0	70.3	90.5	16.1	76.0	13.6
2008년	584.0	415.3	71.1	87.5	15.0	81.2	13.9
2009년	602.6	432.2	71.7	86.7	14.4	83.6	13.9
2010년	635.8	459.2	72.2	88.2	13.9	88.3	13.9
2011년	673.6	487.6	72.4	90.8	13.5	95.2	14.1
2012년	729.2	539.6	74.0	93.1	12.8	96.5	13.2

주: 동 자료들의 출처는 〈부표 1〉을 참조하기 바란다.

한편, 특정 분위수에 해당하는 소득계층의 크기를 찾는데 필요한 전체 인구의 단위는 다른 나라들의 사례와 같이 우리나라의 소득세 과세단위인 개인단위로 한다. 즉, 특정 분위수에 해당하는 소득계층의 소득을 합산하는 인적단위는 개인단위이다. 이와 더불어서 특정 소득분위수의 위치를 결정하는 데 필요한 전체 인구는 20세 초반의 상당수 남성들이 강제적으로 병역의무를 지고 있다는 우리나라의 특성을 반영하기 위해 군인 및 재소자 등을 제외한 20세 이상의 생산가능인구를 기준선으로 사용한다. 또한, 민감도 분석을 위하여 15세 이상 인구, 20세 이상 인구, 15세 이상 생산가능인구를 전체 인구로 전제하여 동일한 분석을 수행하였다.

15세 이상 인구와 20세 이상 인구 자료는 통계청에서 발표되는 장래인구추계의 값을 이용하였다. 장래인구추계는 인구총조사 실시 다음 해에 작성되어 매 5년마다 발표가 되며, 새로운 추계인구 발표 직전 5개년의 인구 추계치는 새롭게 조사된 인구총조사의 정보를 반영하여 확정인구로 전환된다. 따라서 본 연구에서 이용한 15세 이상 인구와 20세 이상 인구자료는 2014년을 기준으로 2007년부터 2010년까지는 확정인구를 나타내고 2011년과 2012년에는 추계인구를 의미한다.

생산가능인구는 통계청에서 매월 실시하는 경제활동인구조사를 바탕으로 발표되는 15세 이상 생산가능인구 통계를 이용하였다. 생산가능인구는 군인, 사회복무요원, 의무경찰, 형이 확정된 교도소 수감자 등이 제외된 인구로 경제활동 가능성에 초점을 둔 인구 지표이다. 통계청에서는 매월 15세 이상 생산가능인구를 발표하고, 세부적으로 성별, 연령대별(5세 단위별) 정보를 제공한다. 특히, 5세 단위의 연령대별 자료를 통해 15세 이상 생산가능인구로부터 15~19세 생산가능인구를 차감하여 20세 이상

생산가능인구를 생성하였다. 본 연구에서 이용되는 각 기준별 전체 인구의 추이는 다음의 <표 2>와 같다.

<표 2> 인구 기준별 전체 인구 추이

(단위: 천명)

연도	15세 이상 인구 ¹⁾	15세 이상 생산가능인구 ²⁾	20세 이상 인구 ¹⁾	20세 이상 생산가능인구 ²⁾
2007	39,873	39,170	36,641	35,968
2008	40,460	39,598	37,133	36,360
2009	40,950	40,092	37,536	36,808
2010	41,435	40,590	37,968	37,253
2011	42,009	41,052	38,540	37,704
2012	42,445	41,582	39,022	38,284

자료: 1) 통계청, 국가통계포털(KOSIS), 장래인구추계.

2) 통계청, 국가통계포털(KOSIS), 경제활동인구조사.

마지막으로 특정 분위수에 해당하는 소득계층의 소득합계는 특정연도의 종합소득세 신고자와 근로소득 연말정산 신고자의 소득에서 중복을 제거한 ‘통합소득 자료’를 기반으로 한다. 특히 통합소득금액이 6천만원 이상⁹⁾인 경우를 대상으로 매 1천만원 단위로 그 소득구간에 속한 인원과 그들의 종합소득세 및 근로소득세 과세대상 소득합계액 정보를 제공한 통계를 사용한다. 이런 통계를 바탕으로 우리나라의 최상위 소득구간에서의 소득에 대한 인구비중 밀도함수(density function)를 그리면 아래의 <그림 3>과 같다. 동 그림에 따르면 최상위 소득구간에서 소득의 인구비중 밀도함수는 소득이 증가함에 따라 단조적으로 낮아지고 있음을 보여준다.

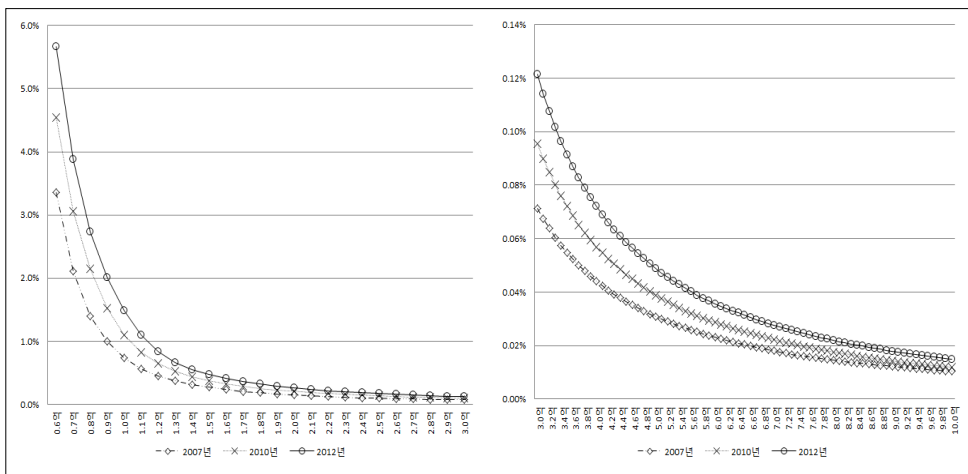
본 연구에서 사용하는 통합소득 통계 자료의 특성을 살펴보면 다음과 같다. 동 통계자료는 일정금액 이상의 과세대상 소득에 대하여 1천만원 단위로 세분화된 과세대상 소득구간별 신고인원과 과세대상 소득에 대한 정보만을 제공하고 있기 때문에 소득유형별 분석이 불가능한 한계가 있다. 그리고 자료가 생산된 기간이 2007~2012년 이어서 외국의 경우처럼 장기간에 걸친 소득집중도 추이를 분석하는 데 한계가 있다.¹⁰⁾ 그럼에도 글로벌 금융위기 전후의 연도들이 포함됨에 따라 글로벌 금융위기가

9) 2007~2010년에 대해서는 5천만원 이상부터 통계가 존재하고 2011~2021년에 대해서는 6천만원 이상부터 통계가 존재한다.

우리나라의 소득분배에 어떤 영향을 주었는지, 그리고 최근에 소득불평등이 더욱 심화되고 있는지를 점검하는 것은 가능하다고 본다. 또한 종합소득세 신고자와 근로소득 연말정산 신고자의 소득에서 중복을 제거한 통계이기 때문에 중복 제거를 위한 인위적인 가정에서 발생할 수 있는 편의가 없다. 그리고 동일한 통합소득 자료를 바탕으로 한 통합소득 100분위 통계의 경우 각 분위별 소득 경계값이 제공되고 있지 않아 이를 추계해야 하는 과정에서 편의가 발생할 수 있지만 본 연구에서 사용하는 통계는 그러한 문제가 없다는 장점이 있다.

〈그림 3〉 최상위 소득구간에서의 소득에 대한 인구비중 밀도함수

(단위: %, 억원)



주: 위쪽 분포는 6천만원에서 3억원, 아래쪽 분포는 3억원에서 10억원까지의 소득분포를 2007년, 2010년, 2012년 등 3개 연도에 대하여 그린 것이다. 또한 인구비중을 구하기 위해 사용한 전체 인구는 20세 이상 생산가능인구이다.

출처: 국세청 내부자료.

한편 소득구간별 제공되는 과세대상 소득은 종합소득세 과세대상인 사업소득, 이자·배당소득, 기타소득, 연금소득, 근로소득으로 구성되어 있다. 따라서 종합소득세의 과세대상이 아닌 분리과세 대상 이자·배당소득 및 분류과세 대상인 퇴직소득,

10) 미국을 분석한 Piketty and Saez (2001)의 경우처럼 기존 국세통계연보에서 제공하는 통계를 기본적으로 사용하여 보다 긴 시계의 소득집중도를 1차적으로 추정한 후 소득구간이 보다 세분화되었지만 시계가 짧은 자료에서 나온 소득집중도 추정치와 비교가 가능할 것이다. 이런 비교를 통해 소득구간이 세분화된 자료가 제공되지 않은 과거 연도들의 소득집중도 추정치를 보정함으로써 보다 정확한 과거 연도들의 추정치를 생산할 수 있을 것이다.

양도소득은 최상위 소득계층의 소득집중도 추정시 포함되지 않았음을 밝힌다. 다만, 연구의 대상이 최상위 소득계층이기 때문에 동 계층의 이자·배당소득은 대체로 금융 소득종합과세 대상일 가능성이 높다고 본다. 따라서 상위 소득구간을 위로 더욱 좁힐 수록 이자·배당소득 중 누락되는 부분은 더욱 줄어들 것으로 기대한다.

2. 분석방법론

특정 분위수에 속한 소득계층의 소득집중도를 추정하기 위해서는 특정 분위수가 속한 소득구간을 식별하고, 특정 분위수에 연계되는 소득경계값 및 그 소득경계값 이상의 소득을 가진 소득계층의 소득을 합산할 수 있어야 한다. 이러한 작업을 위해 사용되는 방법은 주로 파레토 보간법(Pareto interpolation)이나 Mean Split Histogram 방법이다. 파레토 보간법은 해당 소득구간의 소득분포가 파레토 분포를 따른다는 가정 아래에서 사용되는 방법이다. 따라서 파레토 보간법을 사용하기 위해서는 특정 소득구간별로 그 소득분포가 파레토 분포를 따르는지에 대한 검증 작업이 필요하다. 우리나라 자료를 분석한 Cho, Park, Phillips(2015)에 따르면 상위 0.1% 또는 그보다 높은 분위수가 속한 소득구간에서는 파레토 분포 가설을 기각할 수 없었지만 상위 0.5% 및 상위 1% 등 분위수가 낮아질수록 파레토 분포 가설을 기각한다고 한다. 이에 본 연구에서는 파레토 보간법이 아닌 Mean Split Histogram 방법론을 사용하여 소득집중도를 추정하고 있다.

Mean Split Histogram 방법론은 Atkinson(2005)이 영국의 최상위 소득계층의 소득집중도를 추정할 때 사용한 방법으로 특정 소득분위수가 속한 소득구간에서 해당 소득분위수와 연계되는 소득경계값을 찾는 보간법의 일종이다. 동 방법론을 소개하면 다음과 같다.

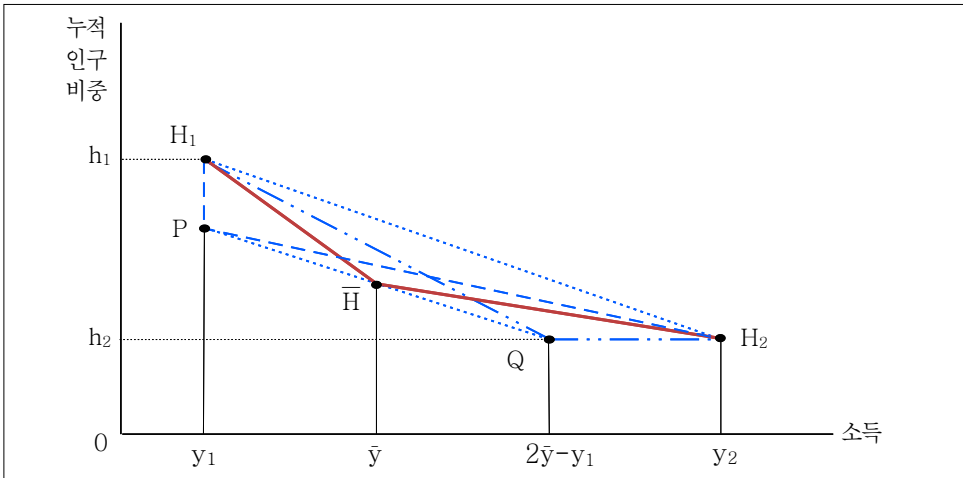
우리가 관심이 있는 특정 소득분위수가 속한 소득구간의 하한 경계값을 y_1 , 상한 경계값을 y_2 라고 하자. 그리고 소득이 y_1 이상인 자들의 인구비중을 h_1 , 소득이 y_2 이상인 자들의 인구비중을 h_2 라 하자. 최상위 소득계층의 소득분포를 보여주는 아래의 <그림 4>에서 보면, 점 $H_1 = (y_1, h_1)$ 과 점 $H_2 = (y_2, h_2)$ 가 이를 나타낸다. 그리고 동 소득구간에 속한 사람들의 평균 소득은 \bar{y} 라고 정의하자.¹¹⁾ 관심 대상인 특정 소

11) 통상 최상위 소득계층의 소득분포에서는 소득이 증가할수록 인구비중은 낮아진다. 다시 말하면 소득이 낮은 사람들의 밀도(density)가 높은 경향이 있다. 이런 경우 소득구간 $[y_1, y_2]$ 에 속

득분위수 h^* 는 세로축의 h_2 와 h_1 의 사이에 있다고 하자. 이때 우리는 특정 소득분위수 h^* 에 대응하는 소득 $y^* = y(h^*)$ 을 찾은 후 소득이 y^* 이상인 자들의 소득 합계가 전체 소득에서 차지하는 비중인 소득집중도를 도출하고자 한다.

이런 상황에서 특정 소득분위수 h^* 에 속한 사람들의 소득집중도의 ‘전체’ 하한(gross lower bound)은 그 구간에 속한 모든 사람들이 평균 소득 \bar{y} 를 갖는 경우가 된다. 반면 이런 소득집중도의 ‘전체’ 상한(gross upper bound)은 평균 소득을 달성하는 가운데 가능한 한 많은 사람들이 동 소득구간의 상한 경계값 y_2 에 분포하고, 나머지 사람들은 소득구간의 하한 경계값 h_1 에 분포하도록 하는 경우이다.

〈그림 4〉 Mean Split Histogram 방법



주: 누적 인구 비중은 오른쪽 상단으로부터의 누적을 말한다. 즉 (1-누적분포함수)이다.

자료: Atkinson (2005)의 〈그림 1〉(p.333)을 저자가 변형하여 그린 것이다.

만약 소득에 대한 인구비중 밀도함수(density function)가 소득이 증가함에 따라 같거나 감소한다면 소득집중도의 하한과 상한의 범위를 더욱 좁히는 것이 가능하다. 소득의 밀도함수가 소득이 증가함에 따라 증가하지 않는다는 전제는 최상위 소득구간에서 비교적 납득 가능한 전제로 생각된다. 앞서의 〈그림 3〉에서 보는 바와 같이 우리나라의 소득 상위구간에서의 소득분포는 이런 전제가 합리적임을 보여준다. 이런 경우 동 소득구간에 속한 사람들의 평균 소득 \bar{y} 는 y_1 과 y_2 의 중앙이 아닌 y_1 에 가까

한 사람들의 평균 소득 \bar{y} 은 y_1 과 y_2 의 중앙이 아니라 y_1 에 가깝게 위치한다.

게 위치한다. 이런 밀도함수에 대한 전제 아래에서 소득집중도의 하한과 상한의 폭을 좁힐 수 있는 새로운 밀도함수 형태를 찾을 수 있다.

먼저, 소득집중도의 새로운 하한을 형성하는 분포 형태를 살펴보자. 소득분포에 관한 위의 〈그림 4〉에서 소득의 위치가 $2\bar{y} - y_1$ 이고, 누적 인구비중은 h_2 인 점 Q 를 찾을 수 있다.¹²⁾ 그리고 원 소득구간의 하단, 즉 구간 $[y_1, 2\bar{y} - y_1]$ 에 대하여 균등분포를 갖는다고 한다면 원 소득구간에 속한 사람들의 평균 소득이 \bar{y} 라는 조건을 만족시킬 수 있다. 이런 밀도함수 형태를 보여주는 것이 바로 선 H_1Q 와 선 QH_2 이다. 이러한 밀도함수 형태를 갖는다면 소득구간 $[y_1, y_2]$ 에 속한 사람들의 평균 소득은 \bar{y} 라는 조건을 만족시키는 가운데 원 소득구간의 상단, 즉 $[2\bar{y} - y_1, y_2]$ 의 밀도함수는 0이 된다.¹³⁾ 이런 분포형태는 소득의 밀도함수가 소득이 증가할 때 증가하지는 않는다는 전제 아래에서 소득집중도의 ‘새로운’ 하한(refined lower bound)으로 작동한다.

다음, 소득집중도의 새로운 상한을 형성하는 분포 형태를 살펴보자. 위의 〈그림 4〉와 같이 선 H_1H_2 과 평행한 선 PQ 를 찾을 수 있다. 여기서 점 P 에서의 소득은 소득구간의 하한 경계값 y_1 이다. 이때 선 H_1P 와 선 PH_2 는 새로운 형태의 밀도함수로서 평균 소득은 \bar{y} 이고, 밀도함수는 증가함수가 아니라는 조건들을 만족한다. 이런 형태의 밀도함수는 위의 두 조건을 만족시키면서 원 소득구간의 하한 경계값에 가장 많은 사람들을 위치시키는 분포이기 때문에 소득집중도의 ‘새로운’ 상한(refined upper bound)으로 작동한다.

이상과 같이 밀도함수가 증가함수는 아니라는 전제를 통해 소득집중도의 하한과 상한의 폭을 좁히는 것이 가능하다. 그리고 소득집중도의 새로운 하한과 상한을 대표하는 대푯값으로 Atkinson (2005)은 Mean Split Histogram 방법론을 사용하여 찾고 있다. 위의 〈그림 4〉에서 이를 나타내는 밀도함수 형태는 선 $H_1\bar{H}$ 와 선 $\bar{H}H_2$ 이다. 여기서 \bar{H} 는 소득이 \bar{y} 에 위치하며 누적 인구 비중(\bar{h})은 선 PQ 상에 있는 점을 나타낸다.

만약 우리가 관심이 있는 특정 소득분위수(h^*)가 \bar{h} 보다 크다면 그에 대응하는 소

12) 여기서 평균 소득 \bar{y} 와 점 Q 의 소득인 $2\bar{y} - y_1$ 간의 거리는 평균 소득 \bar{y} 와 y_1 간의 거리와 동일하다.

13) 즉 동 상단에 위치한 사람들은 아무도 없음을 뜻한다.

소득 y^* 는 선 $H_1 \bar{H}$ 를 이용하여 도출할 수 있다.¹⁴⁾ 원 소득구간 $[y_1, y_2]$ 에 해당하는 사람들의 소득 합계는 알고 있기 때문에 소득이 y^* 보다 작은 사람들의 소득 합계를 차감하고 남은 금액과 상위 소득구간에 속한 사람들의 소득 합계를 더하면 특정 소득분위수(h^*)에 해당하는 사람들의 소득 합계, 즉 소득집중도의 분자가 도출된다. 이때 소득이 y^* 보다 작은 사람들의 소득 합계는 특정 소득분위수(h^*)가 \bar{h} 보다 큰 경우 소득구간 $[y_1, y^*]$ 의 중앙값에 누적 인구비중 차이($h_1 - h^*$)와 전체 인구를 곱함으로써 산출된다.

III. 분석결과

본 절에서는 Atkinson (2005)이 소개한 Mean Split Histogram 방법론을 국세청에서 제공한 2007~2012년간의 과세대상 소득에 대한 도수분포 자료에 적용하여 최상위 소득계층의 소득집중도 추정 결과 및 그 추이에 대하여 논의한다.

본 절에서는 분석의 편의상 최상위 소득계층 중 상위 5%, 상위 1%, 상위 0.5%, 상위 0.1%, 상위 0.01%에 초점을 두어 분석한다.¹⁵⁾ 특정 분위수를 지정하기 위해서는 전체 인구에 대한 가정이 필요하다. 본 연구에서는 영국을 분석한 Atkinson (2005)처럼 15세 이상 인구, 김낙년(2012) 및 일본을 분석한 Moriguchi and Saez (2010)처럼 20세 이상 인구, 20세 초반의 남성 중 다수가 국방의 의무를 지고 있는 우리나라의 특성을 반영하여 15세 이상 생산가능인구 및 20세 이상 생산가능인구를 전체 인구로 사용한다. 여기서 최근의 높은 대학 진학률을 감안할 때 20세 이상 생산가능인구를 기준으로 한 소득집중도를 기준선으로 한다.

먼저, 소득집중도 추정 과정에서 필요한 소득분위수별 경계소득 및 동 경계소득과 같거나 큰 소득을 가진 사람들의 평균소득을 소비자물가를 반영한 실질 개념으로 살펴본다. 20세 이상의 생산가능인구를 기준으로 경계소득을 보면, 분석 중인 소득분위수에서 대체로 2009년까지 하락하는 추세를 보이다가 2010년부터 상승세로 전환하여 2012년에 가장 높은 수준에 도달하는 것으로 나타났다. 예를 들면, 아래의 <표 5>에

14) 반대로 특정 소득분위수(h^*)가 \bar{h} 보다 작다면 그에 대응하는 소득 y^* 는 선 $\bar{H}H_2$ 를 이용하여 도출할 수 있다.

15) 각 분석 연도마다 상위 5% 이내의 모든 소득분위수에 대한 소득집중도를 추정하는 것이 가능하다.

〈표 5〉 최상위 소득계층의 경계소득 및 평균소득

(단위: 백만원)

구분		20세 이상 생산가능인구		20세 이상 인구	
		경계소득	평균소득	경계소득	평균소득
2007년	상위 5%	60.4	101.4	59.90	100.65
	상위 1%	105.8	203.3	105.12	201.48
	상위 0.5%	135.7	288.4	134.68	285.57
	상위 0.1%	290.7	721.2	287.68	713.24
	상위 0.01%	1,199.9	2,836.5	1,187.05	2,806.29
2008년	상위 5%	60.5	101.2	59.97	100.37
	상위 1%	105.5	201.2	104.85	199.16
	상위 0.5%	134.3	284.9	133.25	281.80
	상위 0.1%	289.4	709.8	286.06	700.97
	상위 0.01%	1,154.4	2,789.2	1,139.15	2,754.98
2009년	상위 5%	59.5	99.5	59.05	98.71
	상위 1%	103.1	198.0	102.58	196.16
	상위 0.5%	132.2	281.0	131.14	278.07
	상위 0.1%	288.8	694.1	285.72	686.19
	상위 0.01%	1,098.6	2,700.9	1,087.95	2,669.69
2010년	상위 5%	61.2	104.9	60.70	104.03
	상위 1%	109.4	212.4	108.73	210.49
	상위 0.5%	140.5	302.8	139.44	299.73
	상위 0.1%	310.8	753.9	307.60	745.57
	상위 0.01%	1,179.5	2,994.7	1,166.15	2,960.39
2011년	상위 5%	62.1	108.0	61.56	107.04
	상위 1%	112.3	222.7	111.53	220.29
	상위 0.5%	144.6	320.2	143.21	316.42
	상위 0.1%	335.2	804.5	330.81	794.27
	상위 0.01%	1,279.9	3,073.0	1,264.40	3,033.94
2012년	상위 5%	63.2	108.8	62.74	107.93
	상위 1%	113.3	222.0	112.61	219.93
	상위 0.5%	145.5	318.0	144.31	314.70
	상위 0.1%	332.7	787.4	329.07	778.73
	상위 0.01%	1,252.8	2,998.6	1,238.95	2,965.49

주: 경계소득 및 평균소득은 2012년 소비자물가지수를 100으로 놓고 물가를 반영한 실질 개념이다.

서 보여주듯이, 상위 1%의 경계소득은 2007년 1억 577만원, 2008년 1억 550만원, 2009년 1억 315만원으로 하락하다가 2010년 1억 939만원, 2011년 1억 1,229만원, 2012년 1억 1,326만원으로 상승하고 있는 것으로 나타났다. 상위 0.01%의 경우

2007년 11억 9,994만원에서 2009년 10억 9천 860만원까지 하락하였다가 2010년 11억 7,946만원으로 상승 전환한 후 2012년 12억 5,276만원에 도달하였다. 이런 추세는 최상위 소득계층의 평균소득 추이에서도 확인된다. 예를 들면, 상위 1%의 평균소득은 2007년 2억 327만원에서 2009년 1억 9,800만원으로 하락한 후 상승하기 시작하여 2012년 2억 2,199만원에 도달하였다. 이런 추세는 다른 인구기준을 사용해도 유지되는 것으로 나타났다.¹⁶⁾

〈표 6〉 소득집중도 추정 결과 및 인구기준에 따른 격차

상위소득 분위수	인구 기준	구분	단위	연도					
				2007	2008	2009	2010	2011	2012
5%	15세 이상 인구	A	%	29.40	29.87	29.60	30.75	31.50	30.33
	20세 이상 인구	B		27.96	28.38	28.10	29.23	29.97	28.88
	20세 이상 생산가능인구	C		27.65	28.03	27.77	28.90	29.59	28.56
	차이	A-C	%p	1.75	1.85	1.83	1.85	1.91	1.77
		B-C		0.31	0.35	0.33	0.32	0.38	0.32
1%	15세 이상 인구	A	%	11.70	11.79	11.69	12.38	12.89	12.29
	20세 이상 인구	B		11.19	11.26	11.17	11.83	12.33	11.77
	20세 이상 생산가능인구	C		11.08	11.14	11.05	11.71	12.20	11.66
	차이	A-C	%p	0.62	0.65	0.64	0.67	0.69	0.64
		B-C		0.11	0.12	0.11	0.12	0.14	0.11
0.50%	15세 이상 인구	A	%	8.26	8.30	8.25	8.77	9.21	8.75
	20세 이상 인구	B		7.93	7.97	7.92	8.42	8.86	8.42
	20세 이상 생산가능인구	C		7.86	7.89	7.84	8.35	8.77	8.35
	차이	A-C	%p	0.39	0.41	0.41	0.43	0.44	0.41
		B-C		0.07	0.08	0.07	0.07	0.09	0.07
0.10%	15세 이상 인구	A	%	4.10	4.11	4.05	4.34	4.61	4.32
	20세 이상 인구	B		3.96	3.96	3.91	4.19	4.45	4.17
	20세 이상 생산가능인구	C		3.93	3.93	3.87	4.16	4.41	4.13
	차이	A-C	%p	0.17	0.18	0.18	0.19	0.20	0.18
		B-C		0.03	0.03	0.03	0.03	0.04	0.03
0.01%	15세 이상 인구	A	%	1.62	1.61	1.57	1.72	1.76	1.64
	20세 이상 인구	B		1.56	1.56	1.52	1.66	1.70	1.59
	20세 이상 생산가능인구	C		1.55	1.54	1.51	1.65	1.68	1.57
	차이	A-C	%p	0.07	0.07	0.07	0.07	0.08	0.07
		B-C		0.01	0.01	0.01	0.01	0.02	0.01

16) 소비자물가수준을 반영하지 않은 명목 개념의 경계소득 및 평균소득은 매년 증가하는 추세를 보이고 있다.

다음은 최상위 소득계층의 소득분위수별 경계소득 및 평균소득 등의 정보를 활용하여 추정된 해당 분위수의 소득집중도를 살펴본다. 전체 인구의 기준을 20세 이상 생산가능인구로 하여 추정된 상위 5%의 소득집중도 추이는 아래의 <표 6>에서 보여주듯이 2007년 27.65%에서 서서히 증가하여 2011년 29.59%로 정점에 도달한 후 2012년 28.56%로 다소 감소한 것으로 나타났다. 그리고 상위 1%의 소득집중도도 상위 5%와 같은 추세(2007년 11.08%; 2008년 11.14%; 2009년 11.03%; 2010년 11.71%; 2011년 12.20%; 2012년 11.66%)를 보이는 것으로 나타났다. 이러한 소득집중도 추세는 인구 기준을 20세 이상 인구, 15세 이상 인구 등으로 전환하여도 유지되는 것으로 조사되었다. 한편 최근 6년 동안 소득집중도 추정치의 변동성은 분석 대상 소득분위수 중 상위 5%의 소득계층에서 가장 낮은 것으로 분석되었고, 상위 0.1% 소득계층에서 가장 높은 것으로 나타났다.

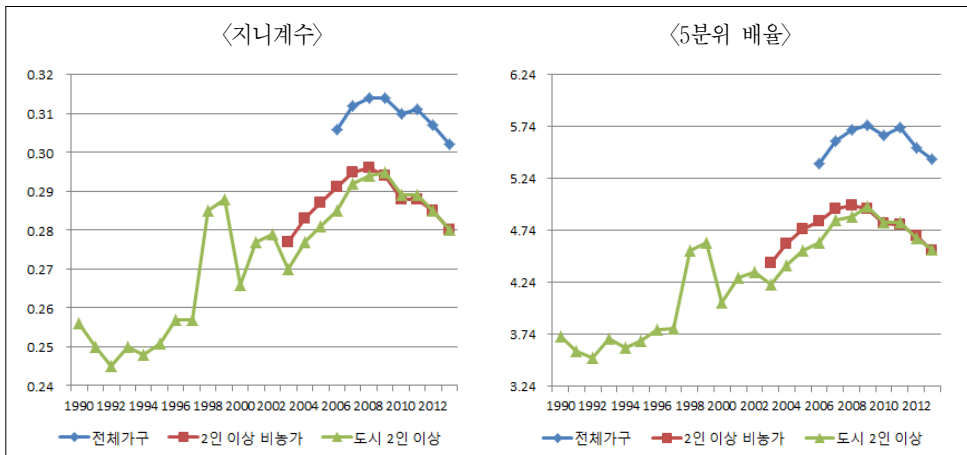
한편 글로벌 금융위기가 최상위층 소득집중도에 미친 부정적인 영향은 일시적이고 그 크기도 미약했던 것으로 판단된다. 이는 글로벌 금융위기가 국내 경제로 전이된 시점인 2009년에 들어서 최상위 소득계층의 소득집중도가 2008년에 비해 소폭 하락하였지만, 이러한 소득분배 개선효과는 곧 사라져서 2010년부터 소득집중도가 다시 증가하다가 글로벌 금융위기 이후 계속된 경기침체로 인해 2012년에는 2010년 수준 정도로 하락한 상태이기 때문이다.

또한 글로벌 금융위기 직전인 2007년 이후의 소득집중도 추이를 보면, 우리나라의 소득분배가 글로벌 금융위기 이후로 개선되지 않았음을 알 수 있다. 이는 2007년도 소득집중도 추정치가 분석기간 중 가장 낮았기 때문이다. 이러한 소득집중도의 추세는 통계청의 소득분배지표인 지니계수나 5분위 배율과 소득분배 변화에 대한 다른 시사점을 제공해 준다. 아래의 <그림 7>이 보여주듯이, 지니계수나 5분위 배율에 따른 우리나라의 소득분배는 2008~2009년까지 악화되었고 그 이후에는 지속적으로 개선되고 있는 것으로 나타났다. 따라서 소득분배 구조의 실태를 정확하게 파악하려면 소득세 신고자료를 이용한 소득집중도처럼 전수조사 자료 및 그에 근거한 행정통계 자료들로 생성된 지표들을 통해 기존 소득분배지표를 보완할 필요가 있음을 알 수 있다.

한편 전체 인구를 잡는 기준에 따라 상위 소득분위수로 사용되는 분위수 중에서 소득집중도 추정 결과에 상당한 차이가 있을 수 있는 것으로 분석되었다. 예를 들면, 20세 이상 인구와 20세 이상 생산가능인구를 전체 인구로 간주하여 상위 0.01%의 소득집중도를 비교한 바 그 차이는 0.02%p 미만으로 미약하였다. 20세 이상 인구와 20

세 이상 생산가능인구의 차이는 대략 67~84만명 사이인 것으로 나타났다. 따라서 상위 0.01%에 속하는 인구의 차이는 대략 67~84명 수준으로 상위 0.01%에 추가되는 인구의 규모가 크지 않아 소득집중도에 미약한 차이만을 유발하는 것으로 나타났다.

〈그림 7〉 우리나라의 소득분배지표 추이



주: 소득분배지표는 가계동향조사에서 생성하는 변수인 가처분소득을 기준으로 작성된 것이다.

출처: 통계청, 국가통계포털 (<http://kosis.kr>).

그러나 상위 5%의 소득집중도를 비교하면, 그 차이는 0.30~0.38%p 사이에 있어서 무시할 수 없는 격차를 유발한다. 이는 인구기준을 20세 이상 생산가능인구에서 20세 이상 인구로 전환할 때 상위 5%에 추가되는 인구의 차이가 3.3~4.2만명으로 상당하기 때문이다. 따라서 인구기준의 설정에 따른 소득집중도 차이는 상위 소득분위수를 5%, 10% 등으로 확대해 가면 더욱 벌어질 것이다. 위의 〈표 6〉에서 알 수 있듯이 영국을 분석한 Atkinson (2005) 처럼 전체 인구를 15세 이상 인구로 한다면 20세 이상 생산가능인구 대비 상위 5%의 소득집중도 추정치는 1.75~1.91%p의 커진다. 따라서 그 나라의 과세체계 및 사회·경제적 상황에 맞는 인구기준을 설정하고 일관되게 유지하는 것이 중요한 것으로 나타났다.

본 연구의 결과와 기존의 Kim and Kim (2014)의 결과를 비교하면, 대체적으로 상위 1% 이내의 소득구간에서는 두 결과간의 차이는 무시할 수 있는 수준으로 보인다. 그렇지만 상위 1%를 벗어나면서 소득집중도 추정치의 오차가 무시할 수 없는 수준으로 커지고 있는 것으로 보인다. 예를 들면, 동일한 인구기준을 사용하더라도 상위 5%의 소득집중도에서 Kim and Kim (2014)의 결과는 2007년 0.53%p, 2008년

〈표 8〉 Kim and Kim(2014)의 결과와의 비교

연도	소득 분위	경제소득 (백만원)			평균소득 (백만원)			소득집중도 (%)			소득집중도 차이 (%p)	
		20세 이상 생산가능 인구	20세 이상 인구	WTID (한국)	20세 이상 생산가능 인구	20세 이상 인구	WTID (한국)	20세 이상 생산가능인구 A	20세 이상 인구 B	WTID (한국) C	C-A	C-B
2007	5%	60.4	59.9	62.3	101.4	100.7	102.4	27.65	27.96	28.49	0.84	0.53
	1%	105.8	105.1	103.6	203.3	201.5	202.9	11.08	11.19	11.28	0.20	0.09
	0.50%	135.7	134.7	134.6	288.4	285.6	287.4	7.86	7.93	7.99	0.13	0.06
	0.10%	290.7	287.7	288.8	721.2	713.2	719.0	3.93	3.96	4.00	0.07	0.04
	0.01%	1,199.9	1,187.1	1,152.0	2,836.5	2,806.3	2,895.5	1.55	1.56	1.61	0.06	0.05
2008	5%	60.5	60.0	61.9	101.2	100.4	101.6	28.03	28.38	28.77	0.74	0.39
	1%	105.5	104.8	104.8	201.2	199.2	200.8	11.14	11.26	11.37	0.23	0.11
	0.50%	134.3	133.2	134.7	284.9	281.8	284.1	7.89	7.97	8.05	0.16	0.08
	0.10%	289.4	286.1	291.7	709.8	701.0	706.3	3.93	3.96	4.00	0.07	0.04
	0.01%	1,154.4	1,139.1	1,132.0	2,789.2	2,755.0	2,785.7	1.54	1.56	1.58	0.04	0.02
2009	5%	59.5	59.1	60.5	99.5	98.7	99.9	27.77	28.10	28.57	0.80	0.47
	1%	103.1	102.6	103.5	198.0	196.2	198.1	11.05	11.17	11.33	0.28	0.16
	0.50%	132.2	131.1	133.8	281.0	278.1	280.7	7.84	7.92	8.03	0.19	0.11
	0.10%	288.8	285.7	291.6	694.1	686.2	692.3	3.87	3.91	3.96	0.09	0.05
	0.01%	1,098.6	1,088.0	1,102.2	2,700.9	2,669.7	2,713.1	1.51	1.52	1.55	0.04	0.03

(표 8)의 계속

연도	소득 분위	경제소득 (백만원)			평균소득 (백만원)			소득집중도 (%)			소득집중도 차이 (%p)	
		20세 이상 생산가능 인구	20세 이상 인구	WID (한국)	20세 이상 생산가능 인구	20세 이상 인구	WID (한국)	20세 이상 생산가능인구 A	20세 이상 인구 B	WID (한국) C	C-A	C-B
2010	5%	61.2	60.7	61.3	104.9	104.0	104.3	28.90	29.23	29.17	0.27	-0.06
	1%	109.4	108.7	112.9	212.4	210.5	210.3	11.71	11.83	11.76	0.05	-0.07
	0.50%	140.5	139.4	146.0	302.8	299.7	298.9	8.35	8.42	8.36	0.01	-0.06
	0.10%	310.8	307.6	306.3	753.9	745.6	748.6	4.16	4.19	4.19	0.03	0.00
	0.01%	1,179.5	1,166.1	1,191.7	2,994.7	2,960.4	2,946.7	1.65	1.66	1.65	0.00	-0.01
2011	5%	62.1	61.6	61.7	108.0	107.0	107.3	29.59	29.97	29.86	0.27	-0.11
	1%	112.3	111.5	108.1	222.7	220.3	220.0	12.20	12.33	12.25	0.05	-0.08
	0.50%	144.6	143.2	153.8	320.2	316.4	313.0	8.77	8.86	8.71	-0.06	-0.15
	0.10%	335.2	330.8	331.8	804.5	794.3	795.3	4.41	4.45	4.43	0.02	-0.02
	0.01%	1,279.9	1,264.4	1,273.3	3,073.0	3,033.9	3,028.3	1.68	1.70	1.69	0.01	-0.01
2012	5%	63.2	62.7	62.4	108.8	107.9	108.3	28.56	28.88	30.09	1.53	1.21
	1%	113.3	112.6	111.1	222.0	219.9	220.1	11.66	11.77	12.23	0.57	0.46
	0.50%	145.5	144.3	156.6	318.0	314.7	310.1	8.35	8.42	8.62	0.27	0.20
	0.10%	332.7	329.1	331.4	787.4	778.7	782.8	4.13	4.17	4.35	0.22	0.18
	0.01%	1,252.8	1,239.0	1,250.8	2,998.6	2,965.5	2,974.7	1.57	1.59	1.65	0.08	0.06

주: WID는 The World Wealth and Income Database에 수록된 한국의 자료로 Kim and Kim (2014)의 연구를 최신화한 자료이다. 경제소득과 평균소득은 2012년 소비자 물가지수를 기준으로 하여 실질소득화한 값이다.

자료: 저자 계산 및 Alvaredo, F., A. B. Atkinson, T. Piketty and E. Saez, The World Wealth and Income Database, <http://www.wid.world/#Database>:

0.39%p, 2009년 0.47%p, 2012년 1.21%p 과대 추정하고 있는 것으로 나타났다. 반면, 2010년과 2011년에는 매우 미약하게 과소 추정하고 있지만 무시할 수 있는 수준으로 보인다. 최근 6년(2007~2012년) 간은 국세통계연보의 통계표가 비교적 소득구간을 세분화하여 정보를 제공하고 있지만, 2006년 이전의 경우 세분화되지 않은 소득구간별 정보만을 제공한다. 따라서 2006년 이전에 대한 기존 연구의 추정결과는 Atkinson(2007)의 지적처럼 정확성 측면에서 신뢰성이 낮기 때문에 기존 연구의 추정치를 사용함에 있어서 신중한 접근이 필요하다.

IV. 결 론

본 연구에서는 최근 전 세계적으로 주목받고 있는 최상위 소득계층의 소득집중도 추이를 분석하여 우리나라의 소득분배 구조의 실태를 파악하였다. 지니계수나 5분위 배율로 대표되는 기존 소득분배 지표는 상위 1% 등의 최상위 소득계층의 정보를 조사 표본에 충분히 반영할 수 없기 때문에 이런 문제를 보완할 지표로서 소득집중도를 추정할 필요가 있다.

본 연구에서는 기존의 국내 연구와는 달리 국세통계연보에서 제공하는 소득세 통계의 한계를 뛰어넘는 국세청의 소득세 신고자료를 사용하였다. 먼저, 동 자료는 종합 소득세 통계와 근로소득세 통계에 중복하여 잡힌 인원을 제거한 자료라는 특징이 있다. 따라서 중복 문제를 해결하기 위한 가정 및 이에 근거한 작업들이 불필요하다. 다음으로는 소득구간이 1천만원 단위로 구성되었고, 소득구간의 숫자가 약 2,760~4,240개에 달할 정도로 세분화되었다는 특징이 있다.

이런 특색이 있는 자료의 장점을 살려서 본 연구에서는 Atkinson(2005)이 제시한 Mean Split Histogram 방법론을 적하여 2007~2012년 동안의 최상위 소득계층의 소득집중도를 추정하였다. 그 결과 20세 이상 생산가능인구를 기준으로 볼 때 상위 1%의 소득집중도는 11.08% (2007년), 11.14% (2008년), 11.05% (2009년), 11.71% (2010년), 12.20% (2011년), 11.66% (2012년)로 2009년 일시적으로 하락한 후 계속 상승하다가 2012년 다시 하락한 추세를 보인다. 이런 추세를 볼 때, 글로벌 금융위기가 최상위층 소득집중도에 미친 부정적인 영향은 일시적이고 그 크기도 미약했던 것으로 판단된다. 또한 글로벌 금융위기 이후 계속된 경기침체로 인해 2012년에는 2010년 수준 이하로 소득집중도가 하락한 것으로 보인다.

한편, 이러한 소득집중도의 추이는 소득분배 변화에 대하여 통계청이 발표하는 소

득분배 지표인 지니계수나 5분위 배율과는 다른 시사점을 제공해 준다. 지니계수나 5분위 배율에 따르면 2008~2009년까지 소득분배가 악화되었다가 그 이후 개선된 것으로 나타났기 때문이다. 이런 사실은 소득분배 실태 및 그 추이를 제대로 파악하기 위해서는 현행 소득분배지표를 보완하는 지표로서 소득세 신고자료를 이용한 소득집중도 등의 지표가 필요함을 시사한다.

이 밖에도 전체 인구를 잡는 기준에 따라 소득집중도 추정 결과에 유의미한 영향을 미치는 것으로 분석되었다. 따라서 그 나라의 과세체계 및 사회·경제적 상황에 맞는 인구기준을 설정하고 일관되게 유지하는 것이 필요하다고 본다.

한편, 본 연구의 최상위층 소득집중도 추정 결과와 The World Wealth and Income Database(WDI)¹⁷⁾에 등재된 Kim and Kim(2014)의 수치를 비교하였다. Kim and Kim(2014)의 추정치는 대체로 상위 1% 이내에서는 큰 차이가 없어 보이지만, 소득 상위 5% 및 1%에서는 무시할 수 없는 수준의 격차가 존재할 수 있는 것으로 판단된다. 특히 상위 10% 등 소득상위의 포괄범위를 확대할수록 추정치에 대한 신뢰는 더욱 낮아질 것으로 보인다. 또한 소득구간의 숫자가 매우 빈약하고 세원투명성이 상당히 낮았던 2006년 이전의 기간들에서는 추정치의 신뢰 측면에서 우려를 낳을 수 있다고 본다.

마지막으로 최상위 소득계층의 소득집중도가 현행 소득분배 지표를 보완할 수 있는 장점이 있음을 감안할 때, 보다 신뢰할 수 있는 추정치를 생산하고, 소득원천별 분석을 가능하게 하며, 분석 기간을 과거로 더 확장시키는 방향으로 국세청 신고자료를 활용할 필요가 있다. 이를 통해 우리나라 경제가 발전하는 과정에서 소득불평등도가 어떻게 진화하여 왔고, 그 원인에 대한 연구가 가능해질 것이라 본다.

■ 참 고 문 헌

1. 국세청, 『국세통계연보』, 각 연도.
2. 김낙년, “한국의 소득집중도 추이와 국제비교, 1976-2010: 소득세 자료에 의한 접근,” 낙성대경제연구소 워킹페이퍼 WP 2012-03, 2012.
3. 김낙년·김종일, “한국 소득분배 지표의 재검토,” 낙성대경제연구소 워킹페이퍼 WP 2013-08, 2013.
4. 박명호, “초고소득층의 특성에 관한 국제비교,” 『조세·재정BRIEF』, 한국조세연구원, 2012.
5. 박명호·전명목, 『소득분배 변화와 정책과제: 소득집중도와 소득이동성 분석을 중심으로』, 한국

17) 기존에는 The World Top Income Database로 불렸다.

- 조세재정연구원, 2014.
6. 통계청, “2008년 1/4분기 가계수지동향 및 소득분배지표 개선안,” 통계청 보도자료, 2008.
7. ———, 『가계동향조사 이용자 가이드』, 2011.
8. Atkinson, A. B., “Top Incomes in the UK over the 20th Century,” *Journal of the Royal Statistical Society*, 168 (2), 2005, pp. 325-343.
9. ———, “Measuring Top Incomes: Methodological Issues,” In: Atkinson, A. B. and Piketty, T. (Eds.), *Top Incomes over the Twentieth Century: A Contrast between Continental European and English-speaking Countries*, Oxford University Press, Oxford, UK, 2007.
10. Cho, J., M. Park and P. C. B. Phillips, “Minimum Distance Testing and Top Income Shares in Korea,” *Journal of Business and Economic Statistics*, Forthcoming.
11. Kim, N. and J. Kim, “Top Incomes in Korea, 1933-2010: Evidence from Income Tax Statistics,” Naksungdae Institute of Economic Research Working Paper 2014-03, 2014.
12. Moriguchi, C. and E. Saez, “The Evolution of Income Concentration in Japan, 1886-2005: Evidence from Income Tax Statistics,” *The Review of Economics and Statistics*, 90 (4), 2008, pp. 713-734.
13. Piketty, T. and E. Saez, “Income and Wage Inequality in the United States, 1913 - 1998,” NBER Working Paper no. 8467, 2001.
14. <http://ecos.bok.or.kr/>
15. <http://kosis.kr/>
16. <http://www.wid.world/#Database>:

〈부표 1〉 전체 소득 추정에 사용된 자료 목록

전체 소득 유형		이용 자료
피용자보수 유형		제도부문별 소득계정 중 가계부문의 임금 및 급여 - 2007년~2012년, 한국은행 경제통계시스템
영업 잉여 유형	영업 잉여	제도부문별 소득계정 중 가계부문의 영업 잉여 - 2007년~2012년, 한국은행 경제통계시스템
	자가 주택 귀속임료	산업연관표의 주거서비스 영업 잉여 - 2007년~2012년 생산자가격 기준 기본분류
	자가 주택 비율	점유형태별 가구 통계 - 1975년~2010년, 통계청, 인구총조사 매 5년 단위
재산 소득 유형	이자 소득	원천세 통계 중 이자 소득 - 2007년~2012년, 국세통계연보
	배당 소득	원천세 통계 중 배당 소득 - 2007년~2012년, 국세통계연보

〈부표 2〉 인구 기준별 소득집중도 추이

인구 기준	소득 분위	2007	2008	2009	2010	2011	2012	평균	표준 편차	변동 계수
생산 가능 인구 (20세)	5%	27.65	28.03	27.77	28.90	29.59	28.56	28.42	0.68	0.0240
	1%	11.08	11.14	11.05	11.71	12.20	11.66	11.47	0.42	0.0365
	0.50%	7.86	7.89	7.84	8.35	8.77	8.35	8.18	0.34	0.0419
	0.10%	3.93	3.93	3.87	4.16	4.41	4.13	4.07	0.18	0.0450
	0.01%	1.55	1.54	1.51	1.65	1.68	1.57	1.58	0.06	0.0392
20세 이상 인구	5%	27.96	28.38	28.10	29.23	29.97	28.88	28.75	0.70	0.0243
	1%	11.19	11.26	11.17	11.83	12.33	11.77	11.59	0.43	0.0367
	0.50%	7.93	7.97	7.92	8.42	8.86	8.42	8.25	0.35	0.0420
	0.10%	3.96	3.96	3.91	4.19	4.45	4.17	4.11	0.19	0.0453
	0.01%	1.56	1.56	1.52	1.66	1.70	1.59	1.60	0.06	0.0394
생산 가능 인구 (15세)	5%	29.09	29.49	29.23	30.39	31.08	29.97	29.87	0.70	0.0233
	1%	11.59	11.65	11.56	12.24	12.74	12.16	11.99	0.43	0.0359
	0.50%	7.86	7.88	7.83	8.33	8.76	8.33	8.17	0.34	0.0417
	0.10%	3.93	3.93	3.87	4.15	4.40	4.13	4.07	0.18	0.0448
	0.01%	1.55	1.54	1.51	1.65	1.68	1.57	1.58	0.06	0.0391
15세 이상 인구	5%	29.40	29.87	29.60	30.75	31.50	30.33	30.24	0.72	0.0238
	1%	11.70	11.79	11.69	12.38	12.89	12.29	12.12	0.44	0.0362
	0.50%	8.26	8.30	8.25	8.77	9.21	8.75	8.59	0.36	0.0414
	0.10%	4.10	4.11	4.05	4.34	4.61	4.32	4.25	0.19	0.0456
	0.01%	1.62	1.61	1.57	1.72	1.76	1.64	1.66	0.07	0.0394

Measuring Top Income Shares in Korea Using Income Tax Return Data

Myung-Ho Park*

Abstract

This paper re-estimates top incomes in Korea using detailed income tax return data from 2007 to 2012. The mean-split histogram method is used as the interpolation method. The new estimates indicate income distribution in Korea has deteriorated since the global financial crisis, but conventional measures of income distribution, such as the Gini index, fail to capture this trend. In addition, the gaps between the new estimates using better data and the existing estimates in the WID are not negligible. These imply that more reliable estimates of top incomes measured using tax return data should supplement the conventional income distribution indexes.

Key Words: income distribution, top incomes, income tax return data

JEL Classification: H24, D31

Received: June 29, 2016. Revised: July 12, 2016. Accepted: July 18, 2016.

* Senior Research Fellow, Korea Institute of Public Finance, 1924 Hannuri-daero, Sejong 30147, Korea, Phone: +82-44-414-2258, e-mail: ecpmh@kipf.re.kr