

## 한국의 소득집중도: update, 1933-2016

김 낙 년\*

### 논문 초록

본고의 과제는 World Wealth and Incomes Database에 실려 있는 한국의 소득집중도를 2016년까지 update하는 것이다. 그 동안에 국민계정 통계가 개편되었고, 소득세 자료에 면세자에 관한 정보가 추가되는 등 자료상의 변화가 있었는데 이를 반영하였다. 소득세 자료의 구간별 통계의 보간법도 새로 제안된 generalized Pareto curve라는 방법으로 바꾸었다. 그 결과 1990년대 중엽 이후 소득집중도의 급격한 상승 추세는 다소 완화되는 것으로 수정되었다. 주목되는 것은 근로소득의 불평등이 2010년 이후 지속적으로 개선되고 있다는 점이다. 그 배후에는 하위 50%에 해당하는 근로자의 소득이 상위 10%보다 빠르게 증가했기 때문이다. 이에 비해 주로 사업소득과 금융소득으로 이루어진 비 근로소득의 집중도는 이전보다는 악화되었지만 상승 추세를 이어가고 있다. 그에 따라 양자를 통합한 전체소득의 집중도는 2010년대 전반에는 하락 또는 정체 기미를 보였지만, 최근에는 비 근로소득의 집중도 상승이 근로소득 집중도 하락보다 빨라 다시 상승으로 돌아서고 있다.

핵심 주제어: 소득집중도, 소득세 자료, 보간법, 소득불평등

경제학문헌목록 주제분류: J3, N3

투고 일자: 2018. 3. 12. 심사 및 수정 일자: 2018. 4. 4. 게재 확정 일자: 2018. 4. 11.

\* 동국대학교 경제학과 교수, e-mail: nnkim@dongguk.edu

## I. 머리말

소득집중도(top income shares)란 예컨대 성인 인구 중에서 소득 순으로 상위 0.1%, 1%, 10%에 해당하는 자들이 전체 소득의 몇 %를 차지하고 있는지를 보여주는 지표이다.<sup>1)</sup> 현재 59개국의 추계 결과가 World Wealth and Incomes Database (이하 WID)에 실려 있어 국제비교가 가능하다. 한국에 관해서는 근로소득에 한정해서 구한 소득집중도(김낙년, 2012b)와 거기에 사업소득이나 금융소득을 더한 전체소득의 집중도(김낙년, 2012a)가 추계된 바 있고, 이를 종합한 결과(Kim and Kim, 2014)가 WID의 한국 데이터로 실려 있다. 다만 추계된 계열이 각각 2010년과 2012년으로 그치고 있는데, 본고는 이를 2016년까지 update하여 최근 소득집중도가 어떤 추이를 보이고 있는지 제시하고자 한다.

다만 그 사이에 자료 상황의 변화가 있었기 때문에 이를 반영하고자 한다. 소득집중도는 상위 소득자의 소득(분자)을 전체 소득(분모)으로 나누어 구하는데, 분자는 소득세 자료에서 얻을 수 있고 분모는 국민계정에서 가계로 귀속되는 소득을 추출하여 구한다. 그런데 국민소득 통계가 종전의 1993년 국민계정체계(System of National Accounts, 이하 SNA)에서 2008년 SNA에 의거한 것으로 이행하면서 과거의 계열에도 크고 작은 변화가 생겼는데 이를 반영할 필요가 있다.

한편 소득세 자료에서도 종전에는 과세자에 관한 정보만을 제공하던 것이 2009년 이후 면세자나 일용근로소득자에 관한 정보도 실는 등 보다 충실해졌다. 그로 인해 예컨대 근로소득의 경우 상위 소득자뿐만 아니라 전체소득도 소득세 자료에서 파악할 수 있게 되었다. 종전에는 소득집중도를 구할 때 분자는 소득세 자료, 분모는 국민계정에 각각 의거했는데, 2009년 이후는 분자와 분모를 모두 소득세 자료에 의거함으로써 추계의 일관성을 높일 수 있다. 근로소득과 같이 이러한 접근이 가능한 경우 상위의 소득집중도에 그치지 않고 중하위를 포함한 전체 소득분포를 제시할 수 있다.

그리고 추계방법에서도 개선이 이루어졌다. 소득세 자료는 원 자료가 아니라 통상

1) 본고에서 다루는 소득집중도는 개인의 세전 소득을 기준으로 한 것이다. 이에 대해 가계조사에 의거하여 산출하는 지니계수 등의 소득분배 지표는 개인이 아니라 가구를 대상으로 한다. 그 경우 가구원수가 다른 가구 간의 후생 수준을 비교할 수 있도록 가구원수의 제곱근으로 나누어 구한 균등화 소득을 이용하며 이 가구소득을 소득이 없는 가구원도 공유하는 것으로 보고 소득분배 지표를 구한다. 이 때 시장소득과 거기에 순 공적 이전소득을 더하여 구한 처분가능소득의 두 가지 소득을 기준으로 소득분배 지표를 구한다. 따라서 본고와 가계조사의 소득분배 지표를 비교할 때에는 이러한 차이에 유의할 필요가 있다.

소득구간별로 인원수와 소득금액을 집계한 표로 제시된다. 이를 이용할 경우 예컨대 상위 1%에 해당하는 자가 어느 소득구간에 속하는지를 알 수 있지만, 그 구간 내 정확한 위치를 특정하기 어렵다. 이를 추정하기 위해 Pareto interpolation이나 mean split histogram과 같은 보간법을 적용하여 상위 x%의 소득집중도를 구했다. 그런데 WID에서는 최근 이를 개선한 Generalized Pareto Curve(이하 GPC)라는 보간법을 제안하고 있으며, 본고에서도 이 방법을 적용하기로 한다.

본고는 다음과 같이 구성되어 있다. 뒤이은 제II장에서는 먼저 근로소득을 대상으로 하여 전술한 자료상의 변화와 방법을 반영한 소득집중도의 추계 결과를 제시하고, 기존 추계와 얼마나 달라졌는지를 보인다. 2009년 이후에는 상위뿐만 아니라 중하위 소득자까지 포함한 전체 소득분포도 제시한다. 제III장에서는 근로소득 이외에 사업 소득과 금융소득을 포함한 전체소득을 대상으로 하는 소득집중도의 추계 결과를 제시한다. 그런데 그 사이에 한국의 소득집중도를 추계한 다른 연구 결과(박명호, 2014; 홍민기, 2015a, 2015b)도 나왔는데, 이들 중에서 기존 추계(Kim and Kim, 2014)나 본고의 결과와 큰 차이를 보인 경우도 있다. 제IV장에서는 이러한 차이가 어디에서 기인한 것인지 비판적으로 검토하기로 한다. 마지막으로 제V장은 본고에서 새로 밝혀진 사실을 정리하고 남겨진 과제를 언급하기로 한다.

## II. 근로소득 집중도

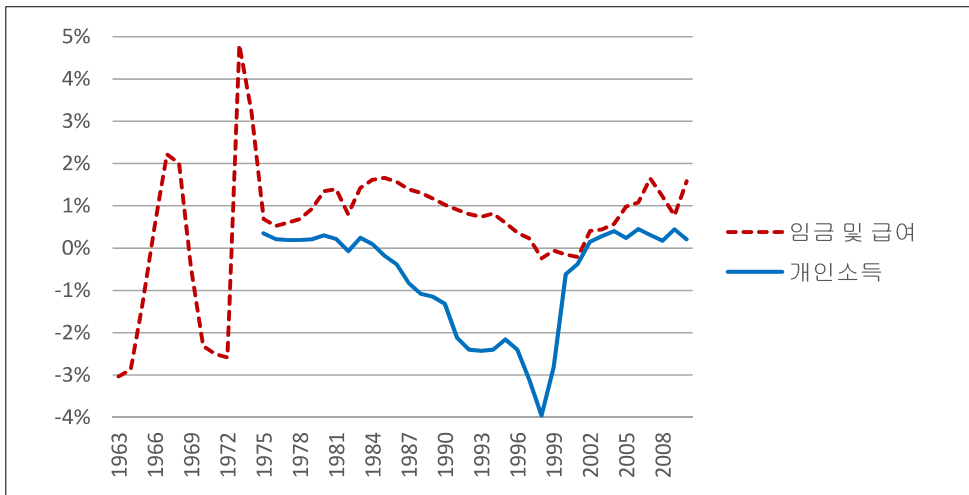
### 1. 근로소득 통계의 보정

근로소득 집중도는 전체 근로자 중에서 상위 x%가 전체 근로소득의 몇 %를 차지하는지를 보여준다. 전체 근로자수는 통계청의 경제활동인구조사에서 가져왔고, 전체 근로소득은 국민계정의 피용자보수 중에서 임금 및 급여를 이용하였다(김낙년 2012b). 상위 소득자의 소득은 소득세 자료에서 얻을 수 있는데, 자료에 따라서는 소득금액이 아니고 과표로 제시되는 경우도 있다. 이 때에는 이 과표에 소득 공제된 부분을 더해주어 소득금액으로 전환해 줄 필요가 있다. 이하에서는 기존 추계에 이용된 소득 통계에 어떠한 변경이 이루어졌는지를 언급하기로 한다.

첫째, 국민계정의 임금 및 급여를 종전의 1993 SNA에서 2008 SNA로 이행한 결과로 대체하였다. 그로 인해 매년의 수치가 얼마나 변했는지를 보이기 위해 <그림 1>을 제시하였다. 1975년을 경계로 그 전후에 자료상의 단절이 있다. 그에 따르면 먼저

1975년 이후의 임금 및 급여는 대체로 1% 전후 높아진 반면, 1998-2001년에는 거꾸로 0.1-0.2% 낮아졌다. 분모에 해당하는 전체 임금 및 급여가 커짐에 따라 소득집중도의 수준은 그만큼 낮아지고 1998-2001년은 거꾸로 높아지는 방향으로 영향을 미칠 것으로 예상된다. 그 결과 기존 추계에서는 외환위기 이후 소득집중도가 급속히 상승한 것으로 나왔는데, 이번 수정으로 그 상승 속도가 다소나마 떨어지게 된다. 1974년 이전은 임금 및 급여 통계가 없어 거기에 고용주의 사회부담금이 포함된 피용자보수의 추이로 연장한 것이다. 당초 1953 SNA에 의거해 작성된 통계가 2008 SNA로 이행한 결과 적지 않은 차이가 생겼음을 알 수 있다.

〈그림 1〉 2008 SNA 이행으로 인한 임금 및 급여와 개인소득의 변화율



주: 1) 1963-74년은 1975년 이후의 임금 및 급여를 피용자보수의 추이로 연장한 것이다. 1974년 이전의 피용자보수는 1953 SNA에서 2008 SNA에 의거한 것으로 이행하였다.

2) 개인소득은 가계에 귀속되는 임금 및 급여, 영업잉여, 재산소득의 합계에서 귀속임료, 금융중개 서비스(FISIM), 투자소득지급을 빼서 구했다.

자료: 한국은행, ECOS; 한국은행(1982, pp. 170-173).

둘째, 2009년 이후에는 전체 근로소득을 국민계정의 임금 및 급여 이외에 소득세 자료에서도 구할 수 있다. 먼저 국세청의 『국세통계연보』에서 근로소득으로 파악되는 것은 근로소득 연말정산 신고자료와 일용근로소득 지급명세서 자료이다. 전자는 과세자와 면세자(즉 결정세액이 있는 자와 없는 자)로 다시 나뉘어져 있다. 그런데 과세당국이 파악한 근로소득이 있는 자들은 취업 근로자수보다 더 많다. 과세 자료에서 파악된 연말정산 신고자와 일용근로소득자를 모두 합치면 2015년에 2,546만 명이

되지만 근로자수는 1,923만 명에 불과하다. 한 해 동안 근로소득의 실적이 있더라도 경제활동인구조사에서 취업자로 파악되지 않는 경우가 적지 않음을 뜻한다. 여기서 말하는 취업자란 조사 기간 중에 수입을 목적으로 1시간 이상 일한 경우로 정의되는데 비해, 과세 자료에서는 1년 중 한번이라도 소득이 발행하면 모두 파악되기 때문에 그러한 차이가 생긴 것이다.

이들 소득세 자료는 고용주가 근로자에게 지급한 임금 및 급여와 원천 징수한 소득세를 과세당국에 신고한 것이다. 고용주가 신고한 근로소득은 비용으로 인정되기 때문에 국세청에 모두 신고한 것으로 볼 수 있으며, 이를 통해 거의 모든 근로소득이 과세당국에 의해 파악되고 있다. 다만 가사도우미나 과외 교사와 같이 가계가 고용한 경우에는 그 인건비가 비용으로 처리되지 않기 때문에 신고가 이루어지지 않는 경우가 있다.

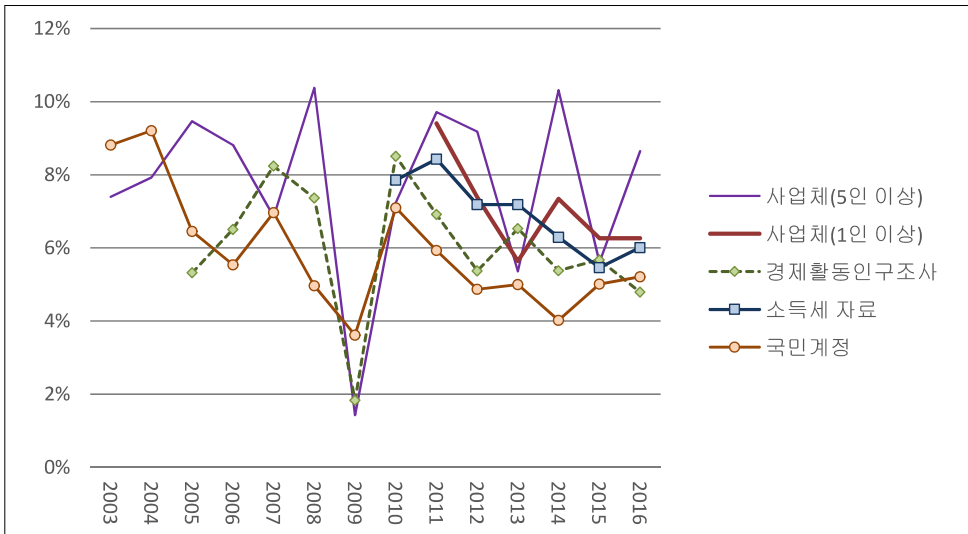
그러면 소득세 자료로 파악된 근로소득을 국민계정의 임금 및 급여와 비교하면 어느 쪽이 더 클까? 2009년에 소득세 자료 쪽이 국민계정에 비해 5% 적었지만, 그 후 격차가 점차 줄다가 역전되어 2016년에는 거꾸로 6%가 많아졌다. 두 자료의 근로소득 동향에 큰 차이가 있었던 셈인데, 어느 쪽을 더 신뢰할 수 있을까? <그림 2>에는 소득세 자료와 국민계정 이외에도 명목 임금총액의 증가율을 보여주는 각종 조사 결과를 제시하였다.<sup>2)</sup> 거기에서 드러나는 것은, 2010년 이후 국민계정의 임금총액 증가율이 소득세 자료에 비해 상당히 낮은 수준에 계속 머물러 있었다는 점이다. 노동부나 통계청의 조사자료와 비교하더라도 국민계정의 증가율이 가장 낮았다. 2009년 이전에는 국민계정의 임금총액 증가율이 다른 조사보다 높은 때도 있지만 낮은 때도 있어 한쪽으로 치우치지는 않았다. 이에 대해 2010년 이후는 국민계정의 과소 파악의 편향이 계속 누적되어 왔음을 알 수 있다.

그 결과 국민계정에서 파악한 임금 및 급여가 최근에는 소득세 자료보다 더 적어졌다. 소득세 자료의 수치는 표본조사에 의한 추정치가 아니라 실제 신고되어 국세청에

2) 소득세 자료나 국민계정은 근로소득의 총액을 제시하고 있으므로 그로부터 직접 증가율을 구할 수 있다. 이에 대해 다른 조사는 임금과 종사자수의 증가율로부터 임금총액 증가율을 구했다. 5인 이상 사업체의 경우는 임금과 근로자수 모두 고용형태별 근로실태조사에서 얻었고, 1인 이상 사업체의 경우 임금은 고용형태별 근로실태조사, 근로자수는 사업체노동력조사에서 얻었다. 이에 대해 통계청의 경제활동인구조사는 고용형태별 부가조사에서 매년 3월과 8월의 임금을 조사(여기서는 8월 임금을 이용)하고 있는데, 이것은 전체 근로자를 대상으로 한다. 조사의 범위는 5인이상<1인이상<경제활동인구조사<소득세자료<국민계정(<그림 2>의 범례의 순서)의 순으로 커진다고 할 수 있다.

서 과세(또는 면세) 대상이 된 소득이며 전수가 집계된 것이다. 더구나 신고 누락이 있음을 감안할 때 소득의 파악 범위가 가장 넓은 국민계정의 수치가 당연히 소득세 자료보다 많을 것으로 기대할 수 있다. 그런데 그것이 오히려 6% 적다는 것은 국민계정의 임금 및 급여가 현재 상당히 과소하게 파악되었음을 뜻한다. 그리고 <그림 2>는 그러한 편향이 2010년대에 들어와 매년 누적되어 온 결과임을 보여준다.

<그림 2> 명목 임금총액의 증가율 비교



주: 1) 소득세 자료는 『국세통계연보』에서 구한 근로소득(연말정산 근로소득과 일용근로소득의 합계)의 증가율이다.

2) 국민계정은 한국은행의 임금 및 급여의 증가율이다.

3) 나머지는 각 조사의 임금상승률과 고용증가율로부터 구한 임금총액 증가율이다. 임금총액증가율 = 임금상승률 + 고용증가율 + 임금상승률 \* 고용증가율.

자료: 국세청, 『국세통계연보』; 한국은행, ECOS; 노동부, 『사업체노동력조사』, 『고용형태별근로실태조사』; 통계청, 『경제활동인구조사』.

여기서는 2008년 이전은 소득세 자료로 전체소득을 알 수 없으므로 국민계정의 임금 및 급여를 이용하지만, 2009년 이후에는 소득세 자료에서 파악된 근로소득을 이용하기로 한다.<sup>3)</sup> 그렇게 한 것은 소득세 자료에서 파악된 전체 근로소득이 국민계정의

3) 전술했듯이 소득세 자료에는 취업자로 파악되지 않는 자들의 소득도 포함되어 있다. 따라서 취업한 근로자를 대상으로 소득집중도를 구할 때에는 분모에 들어가는 전체 근로소득에서 비취업 자들의 소득을 추정하여 제외할 필요가 있다. 여기서는 일용근로소득자 중에서 소득이 낮은 순으로 비취업자일 가능성이 높다고 보아 그들의 소득을 공제하는 방법으로 전체 근로소득을 구했

임금 및 급여보다 더 신뢰할 수 있다는 점 이외에도 소득집중도의 분자와 분모가 모두 동일한 자료에 의거함으로써 일관성을 높일 수 있다는 점을 고려한 것이다. 그 경우 상위의 소득집중도를 뿐만 아니라 중하위 소득자를 포함한 10분위 분포도 구할 수 있다.

셋째, 상위 소득자의 소득은 소득세 자료의 소득구간별 통계로부터 얻을 수 있는데, 1995-2004년은 자료상의 제약이 있다. 즉 이 시기의 소득세 자료에서는 소득금액의 총액을 알 수 있지만, 그 분포는 과표 구간별 통계로만 제시되고 있다. 상위 소득자의 소득을 알기 위해서는 이들 과표에서 공제된 소득을 더하여 과세대상 근로소득(나아가 급여총계)으로 전환해 줄 필요가 있는데,<sup>4)</sup> 기존연구(김낙년, 2012a, 2012b)는 2005년의 소득구간별 전환율(즉 근로소득금액/과표 및 과세대상근로소득/근로소득금액의 각 배율)을 자료가 없는 1995-2004년에 적용하였다. 물론 이 전환율은 소득공제 제도나 소득의 분포가 변하면 달라지기 때문에 이 방법으로 복원한 소득금액이 실제의 수치와 괴리가 생긴다. 다행히 소득세 자료에서 소득금액의 총액을 알 수 있으므로 두 수치가 일치하도록 구간별 전환율을 조정할 수 있으며, 이를 통해 오차를 줄일 수 있다.

다만 전체소득의 전환율은 자료에서 구할 수 있지만, 소득구간별 전환율에 관해서는 추가 정보나 가정이 필요하다. 이 시기는 소득공제가 점차 확대되어 왔고, 소득 중에서 공제되는 소득의 비율은 하위 소득구간으로 갈수록 더 커진다. 따라서 2005년 이전 시기로 거슬러 올라갈수록 그 해의 전환율(=근로소득금액/과표)은 2005년에 비해 낮아지고, 그 영향은 하위의 소득구간에서 더 크게 나타날 것으로 볼 수 있다. 이전의 방식은 이 점을 반영하도록 구간별 전환율을 조정하였다. 그렇지만 전환율이 구간별로 얼마나 차이가 날 것인가에 관한 가정은 뚜렷한 근거가 있는 것은 아니었기 때문에 여기서는 이를 보다 합리적인 방법<sup>5)</sup>으로 개선하였다.

다. 2008년 이전에는 자료의 제약으로 그들에게 귀속되는 근로소득 비율을 알 수 없어 2009년에 파악된 동 비율(7.4%)을 그 이전의 모든 해에 일률적으로 적용하였다.

4) 여기서 『국세통계연보』의 소득의 범위를 정리할 필요가 있다. 근로소득의 경우 과세표준(과표로 줄임), 근로소득금액(=과표+각종 소득공제), 과세대상 근로소득(=근로소득금액+근로소득공제), 급여총계(=과세대상 근로소득+비과세소득)로 통계가 제시된다.

5) 예컨대 2001년의 근로소득의 과표는 62조 원, 근로소득금액은 103조 원, 과세대상 근로소득은 165조 원인데, 먼저 과표에서 근로소득금액으로 전환하는 방법을 예시하면 다음과 같다. 2001년의 과표 구간별 통계에 2005년의 구간별 전환율(근로소득금액/과표)을 적용하여 근로소득금액을 구하면 118조 원이 된다. 즉 15.3%(=118/103) 만큼 더 과대하게 추정되었는데, 그것은 2001년의 실제 전환율이 2005년에 비해 그만큼 낮았음을 뜻한다. 다만 이 비율은 전체 전환율에

넷째, 1985년 이전의 근로소득세 자료는 전술한 1995-2004년과는 달리 모두 소득금액 기준으로 작성되어 있다. 그런데 김낙년(2012b)은 1963-69년의 통계가 과표 기준으로 작성된 것으로 오해하고, 인접한 1970년의 전환율(=소득금액/과표)을 이용하여 이를 소득금액 기준으로 전환하고자 하였다. 그렇지만 1960년대에는 아직 소득공제 제도가 도입되지 않아 소득금액이 곧 과표였기 때문에 이러한 조정은 필요하지 않다. 본고는 이러한 오류를 바로잡았다.

그런데 앞의 셋째와 넷째 요인은 모두 과표를 소득금액으로 전환하는 것과 관련된 것인데, 상위 소득으로 갈수록 소득공제의 비중이 작아지므로 과표와 소득금액 간의 차이가 크지 않다. 예컨대 상위 10%의 소득비중을 추정할 경우에는 전환율의 가정에 포함될 수 있는 오차가 어느 정도 영향을 미칠 수 있지만, 상위 1% 나아가 상위 0.1%로 갈수록 그 영향은 미미해질 것으로 예상할 수 있다.

## 2. GPC 보간법의 적용

소득세 자료는 소득구간별로 인원수와 소득이 어떻게 분포되어 있는지를 보여준다.

〈표 1〉은 2015년의 근로소득 구간별 인원수와 소득의 분포를 예시한 것이다.<sup>6)</sup> 여기

관한 것이어서 각 구간별로 어떠한지에 대해서는 알 수가 없다. 이전의 방식은 하위 구간으로 갈수록 과대 추정의 정도가 더 컸을 것으로 보고 구간별 전환율을 조정하였다. 그렇지만 그 구간별 차이에 관한 가정에는 자의성이 들어가 있어 만족스럽지 못했다. 이에 대해 본고는 다른 방식으로 접근하였다. 즉 2001년의 구간별 과표에 2005년의 구간별 전환율을 적용하여 얻은 근로소득금액과 과표와의 차액은 56(=118-62)조 원으로 실제 값인 41(=103-62)조 원보다 과대하게 추정되었는데, 양자의 비율인 0.723(=41/56)을 각 구간의 추정 근로소득금액과 과표와의 차액에 일률적으로 곱해주는 것이다. 이전과는 달리 차액을 가지고 과대 추정 비율을 구했고, 이 비율을 모든 구간에 동일하게 적용한 것이므로 이전 방식에 비해 가정이 단순하고 알기 쉽다. 그렇게 추정한 2001년의 구간별 전환율을 2005년과 비교해 보면, 이전 방식이 상정한 대로 하위 구간으로 갈수록 전환율이 더 크게 하락한 것을 확인할 수 있지만, 그 하락의 폭이 이전 방식보다 다소 커졌다. 8천만 원 이상과 1천만 원 미만의 두 구간의 수치를 예시하면, 2005년의 근로소득금액/과표의 전환율은 각각 1.119와 2.943이었지만, 이전 방식으로 추정한 2001년의 전환율은 각각 1.079와 2.481, 본고의 결과는 각각 1.088과 2.405로 하락하는 것으로 추정되었다. 이상은 과표에서 근로소득금액으로의 전환을 예시한 것이지만, 근로소득금액에서 과세대상 근로소득으로 전환하는 것도 같은 방법으로 접근하였다.

- 6) 『국세통계연보』에 나오는 근로소득 연말대상자(과세자 및 면세자)와 일용 근로소득자를 모두 합치면 경제활동인구조사의 전체 근로자수를 넘는다는 점은 전술한 바와 같다. 여기서는 일용근로소득자 중에서 소득이 낮은 자들은 취업자로 파악되지 않았을 것으로 보고, 그들과 그들에게 귀속되는 소득을 제외하는 방법으로 전체 근로자수와 근로소득을 구했다. 그리고 과세대상소득에서 빠진 비과세소득을 더하여 급여총계 기준으로 바꾸었다. 〈표 1〉의 소득구간에서 예컨대 1



서 예컨대 상위 1%에 해당하는 자가 1-2억의 소득구간에 들어가 있으며, 그들이 소득비중이 3.77%-14.9%의 범위에 있는 것을 알 수 있지만,<sup>7)</sup> 그 정확한 위치를 특정하기 어렵다. 이를 추정하기 위해서 Pareto 보간법(Pareto interpolation)이나 mean split histogram의 방법이 이용되고 있다. 그런데 전자의 경우 적용 방법에 따라 여러 가지 추정치가 나올 수 있으며, 최근에는 GPC (generalized Pareto curve) 라는 방법이 제안되고 있다. 여기서는 각각의 방법이 어떻게 다르며, 어느 정도 차이가 나는지를 간단히 비교하기로 한다.

Pareto 분포는 다음의 식 (1)과 같이 주어진다.

$$1 - F(y) = (z/y)^a \quad (1)$$

여기서  $y$ 는 소득,  $F(y)$ 는 누적분포함수(cumulative distribution function),  $z$ 는 Pareto 분포가 적용 가능한 소득의 하한,  $a$ 는 Pareto 계수를 가리킨다.  $1 - F(y)$ 는 최상위로부터 누적하여 상위  $x\%$  안에 들어가는 비율(<표 1>의 표기로 나타내면  $1 - p^N$ ) 이 된다. 예컨대 상위 1%에 해당하는 소득( $y$ )을 구할 경우에는  $1 - F(y) = 0.01$ 로 놓으면 된다.

Pareto 계수를 구하는 방법은 여러 가지가 있다. 첫째, Pareto 분포의 중요한 특징의 하나가 소득이  $y$  이상인 자들의 평균소득은  $y$ 의  $b$ 배가 된다는 점을 이용하는 것이다. 이  $b$ 를 역 파레토 계수(inverted Pareto coefficient)라고 하며,  $b = a/(a - 1)$ 의 관계를 가지고 있다.<sup>8)</sup> Piketty and Saez (2003)는 이러한 방식으로 구간별로  $b$ 를 구했는데, 이를 이용하면 상위  $x\%$ 에 드는 경계소득과 그들의 소득비중을 구할 수 있다. 이 방법을 이하에서는 Pareto 1이라고 하고, <표 1>에는 한국의 2015년 근로소득 데이터를 이용하여 이렇게 구한  $b$ 를  $b_1$ 로 제시하였다. 이를 <표 1>의 표기법으로 나타내면  $b_{1i} = (\sum_i^k Y_i / \sum_i^k N_i) / y_i$  이 된다. 여기서  $i$ 는 최하위 소득구간에서부터  $i$ 번째

억 원이 100,681천 원과 같이 자투리가 붙어 있는 것은 그 때문이다.

7) <표 1>에서 1-2억 원 구간의 인원수의 누적분포( $p^N$ )는 0.9686-0.9967인데, 이것을 최상위로부터 누적(즉  $1 - p^N$ )하면, 상위 3.14%-0.33%에 해당한다. 그에 대응하는 소득의 누적분포( $p^Y$ )는 0.8510-0.9623인데, 이를 최상위로부터 누적(즉  $1 - p^Y$ )하면, 14.9%-3.77%에 해당한다.

8) 지면의 제약으로 수식을 제시하지 않았는데, 구체적인 것은 김낙년(2012a, pp. 9-11)을 참조하기 바란다.

의 소득구간을 말하며 최상위 소득구간이  $k$ 번째 구간이 된다.

〈표 1〉 소득구간별 근로소득의 분포와 역 Pareto 계수(2015년)

하한	상한	인원수	소득	인원수 누적분포	소득 누적분포	역 Pareto 계수		
$y$ (천 원)	(천 원)	$N$ (천 명)	$Y$ (10억원)	$p^N$	$p^Y$	$b1$	$b2$	$b3$
	10, 158	3, 760	19, 490	-	-			6, 690
10, 158	15, 152	2, 662	33, 759	0. 1955	0. 0320	3. 752		3. 202
15, 152	20, 183	2, 298	40, 454	0. 3340	0. 0874	2. 864		2. 617
20, 183	30, 221	3, 337	82, 732	0. 4535	0. 1538	2. 430		2. 183
30, 221	40, 256	2, 196	76, 511	0. 6270	0. 2897	1. 996		1. 878
40, 256	45, 297	791	33, 738	0. 7412	0. 4153	1. 777	3. 147	1. 738
45, 297	50, 343	655	31, 246	0. 7823	0. 4707	1. 700	2. 639	1. 668
50, 343	60, 396	1, 023	56, 457	0. 8163	0. 5220	1. 637	2. 139	1. 586
60, 396	80, 541	1, 328	92, 265	0. 8695	0. 6147	1. 549	1. 618	1. 508
80, 541	100, 681	578	51, 648	0. 9386	0. 7662	1. 497	1. 497	1. 493
100, 681	200, 695	540	67, 783	0. 9686	0. 8510	1. 495	1. 440	1. 644
200, 695	300, 462	38	9, 057	0. 9967	0. 9623	1. 817	1. 787	1. 829
300, 462	500, 280	17	6, 429	0. 9987	0. 9771	1. 839	1. 805	1. 847
500, 280	1, 000, 184	6	4, 059	0. 9996	0. 9877	1. 865	1. 929	1. 848
1, 000, 184		2	3, 447	0. 9999	0. 9943	1. 805		
합계		19, 230	609, 075	1. 0000	1. 0000			

주: 1) 인원수와 소득의 누적분포( $p^N$ ,  $p^Y$ )는 직전 소득구간까지의 누적 인원수(또는 소득)를 전체 인원수(또는 소득)으로 나누어 구한 것이다.

2)  $b1$ ,  $b2$ ,  $b3$ 은 본문에서 설명한 3가지 Pareto 보간법으로 구한 역 Pareto 계수를 말한다.

자료: 국세청, 『국세통계연보』, 2016.

둘째, Feenberg and Poterba(1992)에 의해 이용된 log-linear interpolation의 방법이 있다(이하에서는 이를 Pareto 2라고 함). 위의 식 (1)로부터 인접한 두 소득구간에 대해  $1 - F(y_i) = (z/y_i)^a$ 와  $1 - F(y_{i+1}) = (z/y_{i+1})^a$ 가 성립하는데, 이로부터  $a = \log[(1 - F(y_i))/(1 - F(y_{i+1}))]/\log[y_{i+1}/y_i]$ 가 도출된다. 이를 〈표 1〉의 표기법으로 바꾸어 표현하면,  $a = \log((1 - p_i^N)/(1 - p_{i+1}^N))/\log(y_{i+1}/y_i)$ 가 된다. 이  $a$ 로부터 구한  $b$ 를 〈표 1〉에서는  $b2$ 로 제시하였다.

셋째, 위의 식 (1)을 미분해서 밀도함수(density function)를 구하고, 이를 이용하면 소득이  $y$ 보다 높은 인원수의 비율뿐만 아니라 그들에게 귀속되는 소득금액의 비율

을 구할 수 있다. 이를 이용하면 인접한 구간의 인원수와 소득 정보로부터  $b$ 를 구할 수 있는 산식이 도출되는데,<sup>9)</sup> <표 1>의 표기법으로 나타내면  $\log((1 - p_i^N)/(1 - p_{i+1}^N))/\log((1 - p_i^Y)/(1 - p_{i+1}^Y)) = a/(a - 1) = b$ 가 된다. 이 방법을 Pareto 3이라고 하고, 그렇게 구한  $b$ 를  $b3$ 으로 제시하였다.

이상의 방법으로 구한  $b$ 를 비교해 보면 다소 차이가 보이는데, 그것은 이용한 정보에 차이가 있기 때문이다. Pareto 1의 방법에서는 소득구간의 하한과 달리 상한의 정보가 이용되지 않았고, Pareto 2에서는 소득구간과 인원수만 이용되고 소득 정보는 이용되지 않았으며, Pareto 3의 방법은 인접한 구간의 인원수와 소득 정보만을 이용하였다. Pareto 1에서는 상한이 없는 구간(<표 1>에서는 10억원 이상 구간)에서도  $b$ 를 구할 수 있지만, Pareto 2, 3에서는 하한과 상한이 모두 있는 소득구간에만 적용할 수 있다는 차이가 있다.

<표 1>에 따르면  $b$ 는 구간에 따라 변하고 있음을 알 수 있다. 원래 엄격한 Pareto 분포에서는 계수 값( $a$  또는  $b$ )이 고정되어 있다. 만약  $b$ 가 2라고 한다면 예컨대 1억원 이상 소득자의 평균소득은 2억 원이 되며, 다른 구간에서도 마찬가지로 관계가 성립한다(예컨대 5천만 원 또는 10억 원 이상 소득자의 평균소득은 각각 그 2 배인 1억 원 또는 20억 원이 된다)는 뜻이다. 그렇지만 <표 1>에서 보았듯이 현실에서는 그렇지 않고 구간에 따라  $b$ 가 변함을 알 수 있다. <그림 3>은 이를 다시 그래프로 보인 것인데, X축에는 소득 대신에 인원수의 누적분포인  $p^N$ 으로 바꾸어 제시하였다.

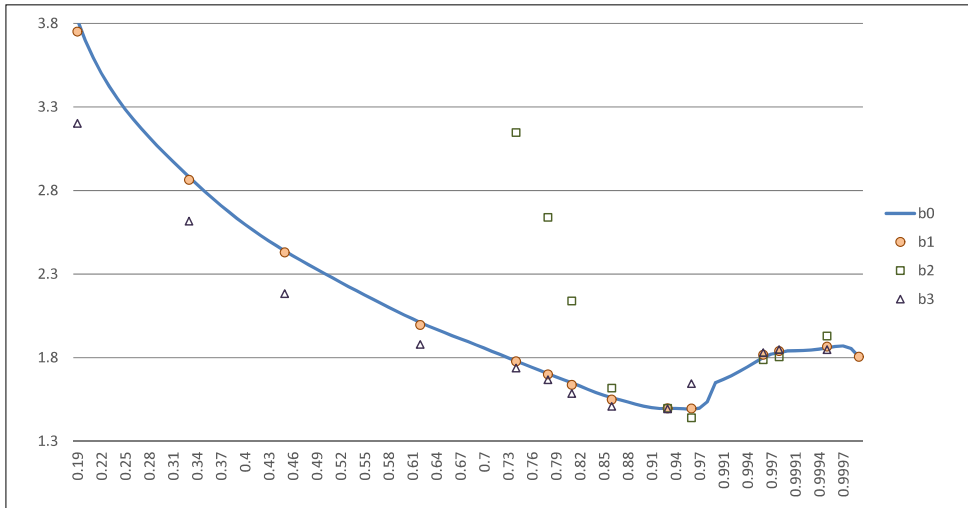
그에 따르면 상위 소득구간으로 갈수록  $b$ 는 점차 하락하다가 누적분포가 0.95(또는 상위 5%) 전후에서 상승으로 돌아서는 U자형을 보이고 있다.  $b$ 가 상승 국면에 있을 때에는 상위 소득구간으로 갈수록 불평등이 커지고 있음을 뜻한다. 경험적으로 보면 다른 연도 또는 다른 나라에서도 다소의 차이가 있지만, <그림 3>과 유사한 U자형의 양상을 보이고 있다.

Blanchet, Fournier and Piketty(2017)은 이러한 소득(또는 부) 분포의 특성을 포착하기 위해  $b$ 가 고정된 Pareto 분포가 아니라 보다 유연한 GPC라는 개념을 제시하였다. 거기에서 역 Pareto 계수의 커브인  $b(p) = E[X > Q(p)]/Q(p)$ 로 정의된다. 여기서  $p$ 는 인원수의 누적분포(<표 1>에서는  $p^N$ )로서  $0 \leq p < 1$ 의 범위에 있고,  $Q(p)$ 란  $p$ 에 상응하는 소득( $y$ )으로 보면 된다. 즉  $b(p)$ 는 인원수의 누적분포  $p$ 에 상응하는 소득  $y(p)$ 보다 더 높은 자들의 평균소득이  $y(p)$ 의 몇 배인가를 보여준다. 이

9) 여기에서 수식의 제시를 생략하였는데, 김낙년(2012a, pp. 9-11, 34-35)를 참조하기 바란다.

점에서는 전술한  $b_1$ 을 구하는 방법과 동일하다. 다만  $b(p)$ 는 연속하는 함수로 정의되었으며, 이를 추정할 때 소득세 자료의 모든 정보가 제약조건으로 반영되고 있다는 점에서 전술한  $b_1 - b_3$ 와 차이가 있다.

〈그림 3〉 보간 방법에 따라 추정된 역 Pareto 계수( $b$ )의 비교



주: 1) X축은 인원수의 누적분포(〈표 1〉의  $p^N$ )를 말한다. 0.99까지는 0.01 단위로 구분하고, 0.99-0.999까지는 0.001, 다시 0.999-1까지는 0.0001로 세분하였다.

2)  $b_1$ ,  $b_2$ ,  $b_3$ 은 〈표 1〉의  $b$ 를 말하고,  $b_0$ 는 GPC 방식으로 구한 것이다.

3) 한국의 2015년 근로소득을 대상으로 한 것이다.

자료: 국세청, 『국세통계연보』, 2016.

〈그림 3〉에서는 한국의 2015년 근로소득을 대상으로 추정한  $b(p)$ 를  $b_0$ 로 제시하였다.<sup>10)</sup> 소득구간에 따라  $b$ 가 계단식으로 바뀌는  $b_1 - b_3$ 와 달리  $b_0$ 는 연속적으로 변함을 알 수 있다. 중하위 소득구간에서는 추정방법에 따라  $b$ 의 차이가 컸지만, 상위 10%에서는  $b$ 가 대체로 접근하고 있음을 알 수 있다. 다만 상위 1%를 포함하는 소득 구간(〈표 1〉의  $p^N$ 으로 말하면 0.9386에서 0.9967의 사이)에서는  $b$ 의 차이가 커져서 어느  $b$ 로 추정하는가에 따라 상위 1%의 소득집중도에 차이가 생길 수 있다는 점에도 유의할 필요가 있다.

한편 Pareto 분포와 같은 특정 분포를 전제로 하지 않고 접근하는 방법으로서

10) WID (<http://wid.world/gpinter>)에서는 GPC 방법을 적용할 때 필요한 정보를 입력하면 그 결과를 자동으로 계산해 주어 편리하다.

mean split histogram의 방법이 있다. 이 방법의 아이디어는 다음과 같다. 상위 구간에서는 소득이 높아질수록 인원수가 줄어드는 경향이 있으며, 그 경우 평균값은 해당 구간의 중앙보다 왼쪽에 위치하게 된다. 즉, 소득구간의 하한과 상한을  $y_i$ ,  $y_{i+1}$ , 그 구간의 평균소득을  $y_{mi}$ 이라고 하면,  $y_{mi} < (y_i + y_{i+1})/2$ 가 된다. 그리고 이 구간을  $y_{mi}$ 을 중심으로  $[y_i, y_{mi}]$ 와  $[y_{mi}, y_{i+1}]$ 의 좌우 두 구간으로 나누면, 오른쪽 구간보다 왼쪽 구간에 더 많은 인원수가 분포되어 있다. 그리고 각 구간에서는 소득이 높아짐에 따라 인원수가 감소하는 속도가 일정하다고 가정한다. 이러한 단순한 가정만으로도 소득구간 내 실제의 인원수 분포에 근접한 결과를 얻을 수 있다는 것이다.<sup>11)</sup>

이상의 다양한 interpolation의 방법 중에서 어느 것이 실태에 보다 근접한 결과를 얻을 수 있을까? Blanchet, Fournier and Piketty (2017, pp. 18-23)는 소득세 통계의 원 자료를 얻을 수 있는 미국(1962-2014년)과 프랑스(1944-2012년)를 대상으로 네 가지 방법(전술한 Pareto 1, 2, GPC, mean split histogram)을 비교하고 있다. 먼저 원 자료에서 구한  $p$ 가 10%, 50%, 90%, 99%에 해당하는 소득 정보를 가지고, 거기에 위의 각 interpolation의 방법을 적용하여 상위 5%, 25%, 70%에 들기 위한 경계소득과 소득비중을 추정한 다음, 이를 원 자료의 실제 값과 비교하였다. 그에 따르면 Pareto 1, 2의 방법이 오차가 비교적 컸고, mean split histogram이 그 다음이고, GPC가 가장 양호한 결과를 보였다. Pareto 1은 상위 소득비중의 추정에서는 비교적 양호하였지만 경계소득 추정에서 오차가 컸고, mean split histogram은 중하위의 소득비중에서 실제와 가까운 결과를 얻었다. 이에 대해 GPC는 경계소득이든 소득비중이든 대부분의 경우 오차가 1%를 넘지 않아 다른 방법보다 월등한 결과를 보였다. 불과 네 지점의 매우 제한된 소득 정보만을 가지고 실제에 근접한 결과를 얻었다는 것은 주목할 만하다. 물론 보다 세분된 정보를 이용하였더라면 각 방법에 의한 차

11) 이를 X축이 소득, Y축이 최상위 소득자로부터의 누적분포인  $1-p^N$ 를 나타내는 좌표로 표현하면 다음과 같다. 소득구간의 하한( $y_i$ ), 구간평균( $y_{mi}$ ), 상한( $y_{i+1}$ )과 그에 대응하는  $1-p_i^N$ ,  $1-p_{mi}^N$ ,  $1-p_{i+1}^N$ 으로 이루어지는 3 좌표를 각각  $L(y_i, 1-p_i^N)$ ,  $M(y_{mi}, 1-p_{mi}^N)$ ,  $H(y_{i+1}, 1-p_{i+1}^N)$ 이라고 하자. 이 때  $y$ 가 높아짐에 따라  $1-p^N$ 은 감소하기 때문에 이들 좌표를 잇는 직선의 기울기는 (-)가 된다. mean split histogram은 구간 내에서 평균 아래 쪽의 인원수 분포가 더 많고(즉 직선 LM의 기울기가 직선 MH의 기울기보다 가파르고), 각 구간 내에 인원수 감소 속도가 일정하다(즉 LM과 MH는 직선이다)고 가정하고 있는 셈이다. 그렇지만 실제의 인원수 분포가 평균을 경계로 하여 달라진다거나, 소득이 증가함에 따라 나타나는  $1-p^N$ 의 감소가 직선을 따라 이루어진다는 보장은 없다. 그래프를 이용한 설명은 Atkinson (2005, p. 333)이나 김낙년(2014, pp. 63-73)을 참조.

이는 줄어든다. 여기서는 이상의 점을 고려하여 GPC의 방법으로 접근하기로 한다.

〈표 2〉 보간 방법에 따라 추정된 분위별 소득비중과 경계소득의 비교(한국의 2015년 근로소득)

	분위별 소득 비중(%)					경계소득(천 원)				
	GPC	Pareto 1	Pareto 2	Pareto 3	mean split histogram	GPC	Pareto 1	Pareto 2	Pareto 3	mean split histogram
1분위	0.83			1.56	3.34					
2분위	2.52	3.34		1.81	3.34	5,355			5,178	
3분위	3.84	3.80		3.95	3.84	10,343	10,199		11,459	10,322
4분위	4.99	4.87		5.00	5.00	13,901	14,669		13,096	13,954
5분위	6.33	6.42		6.45	6.34	17,843	16,217		17,690	17,818
6분위	7.96	8.01		7.51	7.95	22,414	21,268		23,574	22,471
7분위	10.10	10.03		10.07	10.10	28,329	29,186		31,082	28,384
8분위	13.09	13.18		13.35	13.10	36,049	37,738		38,673	36,012
9분위	17.96	17.91		18.01	17.96	47,797	46,908		47,770	47,757
10분위	32.37	32.45	34.16	32.30	32.37	67,920	66,367	66,860	67,819	67,692
상위 5%	20.38	20.38	20.38	20.37	20.37	86,409	86,237	86,236	86,437	86,446
상위 1%	7.16	6.94	6.49	7.44	7.27	137,390	147,008	142,747	143,272	128,310
상위 0.5%	4.77	4.76	4.70	4.81	4.82	170,603	165,992	166,636	166,439	182,154
상위 0.1%	1.97	1.97	1.93	1.98	1.97	339,759	304,046	339,117	338,684	341,236
상위 0.05%	1.36	1.36	1.40	1.36	1.36	462,403	460,512	459,057	465,283	469,020
상위 0.01%	0.57	0.57			0.57	996,710	996,880			997,757

주: 1) Pareto 1, 2, 3은 〈표 1〉의 b1, b2, b3에 대응하는 Pareto 보간법을 말한다.

2) 열은 연두색은 GPC와 비교하여 1-3%의 괴리, 짙은 회색 음영은 3% 이상 괴리가 생긴 경우를 뜻한다.

자료: 국세청, 『국세통계연보』, 2016.

우리나라에서는 소득세의 원 자료를 이용할 수 없기 때문에 각 방법의 우열을 위의 같이 검증할 수는 없다. 다만 2015년의 근로소득을 대상으로 각 방법으로 구한 분위별 소득 비중과 경계소득이 얼마나 차이가 나는지를 보이고자 한다. 〈표 2〉에서는 GPC의 방법을 기준으로 그로부터 1% 이상의 오차가 난 경우를 열은 연두색(그 중 3% 이상 오차가 난 경우는 짙은 회색) 음영으로 표시하였다. 먼저 소득 비중을 보면 Pareto 2를 제외하면 대체로 근접한 결과를 보이고 있다. mean split histogram은 상위 1%의 소득 비중을 제외하면 GPC와 가장 근접한 결과를 보였다. Pareto 1, 3도 상위 1%를 제외하면 상위 소득 집중도를 구하는 데에는 큰 차이가 없음을 알 수 있다. 그렇지만 경계소득의 추정에서는 Pareto 1, 3도 소득비중을 추정할 때와 비교하

여 GPC와의 괴리가 더 커졌다. 여기서도 mean split histogram이 10분위 이하 구간에서 GPC에 가장 근접하였지만, 상위 1% 이상에서는 괴리가 커져서 많은 경우 7%나 차이가 벌어졌다. 이와 같이 특히 상위 1% 전후에서 오차가 커진 것은 <그림 3>에서 알 수 있듯이 그 구간에서  $b$ 의 추세가 바뀌고 있지만 소득세 자료의 소득구간은 오히려 넓어져<sup>12)</sup> 이를 추적하기에 충분히 세분된 정보가 제공되지 못했기 때문으로 생각된다.

### 3. 추계 결과

이상에서 설명한 바와 같이 근로소득 통계를 보정하고, GPC의 interpolation의 방법을 적용하여 구한 근로자 상위  $x\%$ 의 경계소득, 평균소득, 소득집중도를 <부표 1>에 제시하였다. <그림 4>는 상위 1%의 소득 비중을 예시하여 이 결과가 기존 추계(Kim and Kim, 2014)와 얼마나 달라졌는지를 비교하였다. 그에 따르면 기존 추계에 비해 비교적 차이가 크게 나타난 시기는 1960년대와 1970년대 초, 1995-2004년, 그리고 2010년대 이후로 나타났다. 먼저 보간 방법을 Pareto 1에서 GPC로 바꾼 데 기인하는 차이는 대부분의 시기에 1% 미만으로 나왔다. 다만 <표 2>의 2015년 사례에서 보았듯이 일부 연도(1963, 2011-16년)에서 그 차이가 벌어져 3% 전후로 나타나기도 하였다.

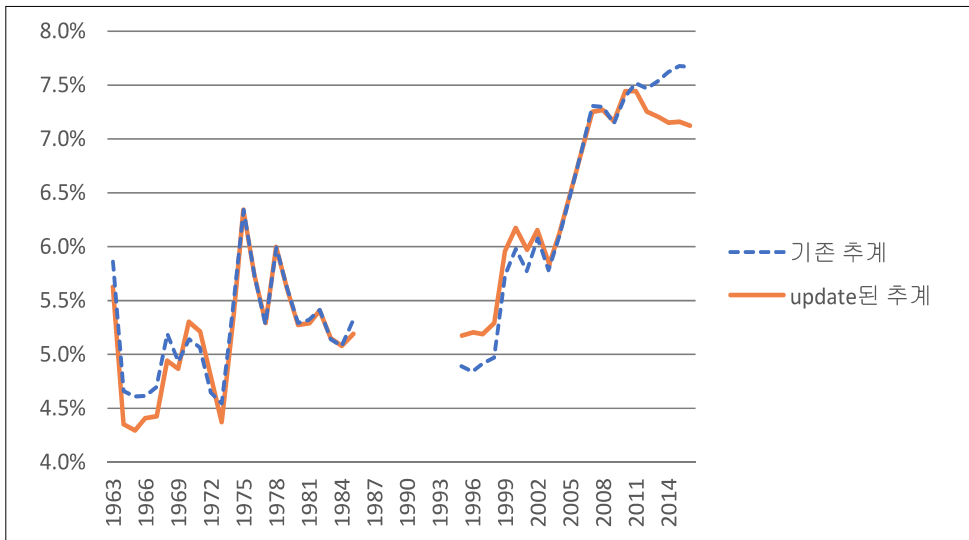
양자의 차이는 대부분 자료의 보정에 기인하는 것이다. 1960년대와 1970년대 초에 나타난 괴리는 앞의 II-1에서 언급한 첫째 및 넷째 요인, 즉 <그림 1>에서 보인 바와 같이 1974년 이전의 임금 및 급여 계열을 1953년 SNA에서 2008년 SNA로 이행한 결과를 적용한 것과, 1960년대의 소득세 통계를 처리하는 과정에서 범한 오류를 바로잡은 데 기인한다. 1995-2004년의 괴리는 II-1에서 언급한 2008년 SNA로 이행한 것 이외에, 셋째 요인으로 언급했듯이 과표 기준으로 되어 있는 소득세 자료를 소득금액으로 전환하기 위한 전환율을 개정한 데 기인한다. 그 결과 이 시기의 update된 소득집중도가 약간 높아졌고, 그로 인해 소득집중도의 상승 추세가 다소 완화된 것으로 나타났다.

주목되는 것은 기존 추계의 소득집중도가 2010년대 이후 계속 상승하고 있는데 반

12) <표 1>에서 알 수 있듯이 1억 원까지의 소득구간은 1천만 원이나 2천만 원 간격인 반면, 그 다음 소득구간이 1억 원 간격으로 뛰는데, 상위 1%는 이 구간에 속한다.

해, update된 추계는 하락으로 돌아서 대조를 보인다는 점이다. 그것은 II-1에서 둘째 요인으로 언급했듯이 기존 추계가 집중도를 계산할 때 분모에 해당하는 전체 근로소득을 국민계정의 임금 및 급여를 계속 이용한 반면, update된 추계는 이를 소득세 자료에 의거하는 방식으로 바꾼 데 기인한다. 국민계정의 임금 및 급여가 2010년대에 들어와 과소평가가 누적되어 왔다는 점과, 집중도의 분모와 분자를 모두 동일한 소득세 자료에 의거하여 일관성을 유지하고 있다는 점에서 update된 추계가 실태를 더 잘 반영하는 것으로 볼 수 있다.

〈그림 4〉 근로소득 상위 1% 소득집중도의 비교: 기존 추계와 update된 추계



주: 기존 추계는 2010년까지 나와 있지만, 그 방식을 2016년까지 연장하여 제시하였다.

자료: Kim and Kim(2014); 〈부표 1〉.

2009년 이후는 소득세 자료를 이용하여 상위 소득자에 한정하지 않고 전체 소득분포를 보일 수 있다. 〈표 3〉은 전체 근로자수를 기준으로 하는 10분위별 근로소득의 비중과 지니 계수를 보여주며, 이를 요약한 것을 〈그림 5〉로 제시하였다. 그에 따르면 하위 50%의 소득 비중이 2010년에 16.1%로 바닥을 친 후 계속 상승하여 2016년에는 19.0%로 높아졌다. 그에 대해 상위 10%는 같은 해에 33.9%에서 32.0%로 하락했고, 그 사이의 중간 40%는 50%에서 49%로 하락하였다. 그 결과, 지니 계수가 2010년 이후 지속적으로 하락하는 양상을 보였다. 〈그림 4〉에서 상위 1%의 소득집중도가 2010년 이후 하락한 것을 보았지만, 그 배후에 하위 50%의 소득증가가 상대



적으로 빨랐던 것이 중요한 요인이었음을 알 수 있다.<sup>13)</sup>

〈표 3〉 근로소득의 분위별 분포(단위: %)

	2009	2010	2011	2012	2013	2014	2015	2016
전체 평균(천 원)	24,501	25,489	27,065	28,569	29,733	30,770	31,673	33,078
1분위	0.592	0.520	0.679	0.774	0.740	0.801	0.828	0.824
2분위	1.967	1.731	2.124	2.364	2.392	2.527	2.516	2.608
3분위	3.382	3.323	3.447	3.681	3.785	3.902	3.842	3.971
4분위	4.585	4.546	4.592	4.773	4.870	5.034	4.992	5.138
5분위	5.967	5.952	5.980	6.116	6.225	6.350	6.333	6.464
6분위	7.814	7.756	7.746	7.820	7.905	7.968	7.964	8.054
7분위	10.178	10.101	10.054	10.074	10.135	10.136	10.102	10.122
8분위	13.534	13.454	13.289	13.210	13.186	13.108	13.094	13.022
9분위	18.784	18.735	18.459	18.227	18.103	17.895	17.958	17.787
10분위	33.197	33.882	33.630	32.961	32.658	32.279	32.372	32.012
지니 계수	0.491	0.499	0.489	0.476	0.472	0.464	0.465	0.459

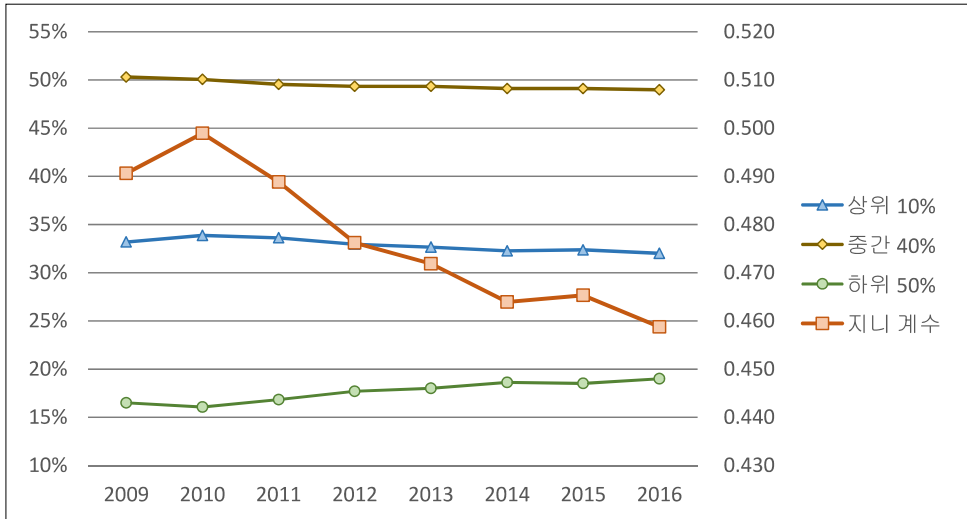
주: 1) 전체 근로자를 소득 순으로 10등분하였을 때 각 분위의 소득 비중이다.

2) 여기서 전체 근로자는 경제활동인구조사(통계청)에서 파악된 전체 근로자수이다.

자료: 국세청, 『국세통계연보』, 2017; 통계청, KOSIS(경제활동인구조사).

13) 그런데 근로소득 분포에 관해서는 노동부의 고용형태별 근로실태조사로 대표되는 사업체조사가 널리 이용되고 있다. 정혁(2017, pp.64-66)은 이 조사의 마이크로데이터를 이용하여 근로자 10인 이상 사업체를 대상으로 하여 1980-2015년의 근로소득의 불평등 지표를 추계하였다. 그에 따르면 2010년 이후에도 불평등 지표는 계속 높아지고 있어 본고(〈표 3〉과 〈그림 5〉)와 상반된 결과가 나왔다. 이 문제에 대한 본격적 검토는 다음 연구를 기약할 수밖에 없지만, 이 사업체조사가 근로소득의 분포 실태를 제대로 반영하지 못하고 있을 가능성을 지적해 두는데 그치기로 한다. 지면의 제약 때문에 심지 못한 본고의 워킹페이퍼 버전(김낙년, 2018, pp.14-16)에서는 이 사업체조사의 소득구간별 근로자수 분포를 소득세 자료와 비교하고 있다. 그에 따르면 사업체조사는 중위 소득구간(2-3천만 원)에서는 근로자수가 상대적으로 과대하게 파악되어 있는 반면, 그 보다 하위 또는 상위 구간으로 갈수록 포착되는 근로자수가 빠르게 줄어드는 양상을 보이고 있다. 예컨대 소득세 자료에 따르면 연소득 2억 원 이상 구간에 들어가는 근로소득자수가 50,908명인데 비해 1인 이상 사업체조사에서는 그 절반 이하, 5인 이상 사업체조사에서는 15,873명이 파악되는데 불과하였다.

〈그림 5〉 근로소득의 분위별 비중과 지니 계수의 추이



주: 1) 중간 40%는 하위 50%와 상위 10%를 제외한 그룹을 말한다.

2) 소득비중은 좌측 눈금, 지니 계수는 우측 눈금.

자료: 〈표 3〉.

### Ⅲ. 전체소득 집중도

그러면 근로소득뿐만 아니라 금융소득과 사업소득을 포함한 전체소득의 집중도는 어떻게 될까? 이 경우에도 소득집중도의 분자에 해당하는 상위 x%에 속하는 자들의 소득은 소득세 자료에서 구한다. 개인이 수취하는 각종 소득의 합산 금액은 종합소득세 자료에서 알 수 있지만, 종합소득을 신고하지 않는 경우가 많다. 근로소득자의 경우 다른 소득이 일정액을 넘지 않을 경우 종합소득의 신고 없이 연말정산으로 과세가 종결되기 때문이다. 따라서 상위 소득자를 포착하기 위해서는 종합소득세 자료와 근로소득세 연말정산 자료 등을 통합해서 이용할 필요가 있다.<sup>14)</sup>

14) 구체적인 방법은 김낙년(2012a, pp. 78-93)을 참조하기 바란다. 한가지 부연하면, 종합소득세와 근로소득 연말정산 자료를 통합할 때 두 자료에 중복해서 나오는 근로소득자를 추산하여 배제할 필요가 있다는 점이다. 이 점에서는 한국과 일본의 자료상황이 유사한데, 김낙년(2012a, p. 87)은 Moriguchi and Saez(2008)가 일본에 적용한 방법을 한국에 원용하여 이를 추정한다. 그런데 국세청은 국회의원의 자료 요청에 대응하는 형태로 일부 시기(2007-2012년)에 대해 두 자료의 중복을 배제한 통합소득을 공개하였다. 그런데 이 통합소득을 이용해서 구한 소득집중도는 김낙년(2012a)의 방법으로 구한 것과 비교하여 큰 차이를 보이지 않았다(예컨대 2012년에 상위 1%의 소득비중은 11.7%로 동일, 상위 10%의 경우는 43.0%와 42.6%). 여기서는

소득집중도의 분모에 해당하는 전체소득은 종래에는 국민계정에서 가계에 귀속되는 소득을 추출하여 이용하였다. 그렇지만 전술한 근로소득에서 드러난 바와 같이 분모도 가능한 한 소득세 자료에 의거하는 것이 바람직스럽다. 국민계정의 개인소득은 유형에 따라 근로소득(국민계정의 소득 항목 중에서 임금 및 급여가 거기에 해당한다, 이하 동일), 금융소득(이자, 배당), 사업소득(영업잉여, 임료, 준법인기업의 소득인출)으로 나누어 볼 수 있다. 먼저 근로소득의 경우 전술한 과소 파악의 편향을 바로잡을 필요가 있다. 여기서는 2010년 이후의 근로소득이 <그림 2>의 국민계정 자료의 증가율이 아니고 그 보다 높은 소득세 자료의 증가율에 따라 증가한 것으로 보정하였다.

금융소득인 이자와 배당은 모두 원천징수 대상이기 때문에 소득세 자료에서 그 총액을 구할 수 있다. 국민계정의 이자 소득에는 금융중개서비스(FISIM)가 포함되어 있어 이를 추정해서 제외할 필요가 있지만, 그렇게 구한 이자는 소득세 자료의 이자와 거의 일치한다. 배당소득도 두 자료의 수치에 차이가 거의 없는데, 이것은 국민계정의 금융소득이 소득세 자료에 의거하고 있음을 보여준다. 따라서 금융소득은 국민계정 자료를 소득세 자료로 대체해도 무방하다.

그런데 사업소득은 종합소득세를 신고하는 경우가 일반적이지만, 그 이외에도 원천 징수되거나 연말정산 대상이 되는 사업소득이 있으며 기타소득 중에도 사업소득이 포함되어 있다. 이들 각각에 대해 소득세 자료가 있지만, 이들 중 일부는 종합소득세를 신고하여 중복되는 부분이 있다. 이를 제거해야 하지만, 자료상의 제약으로 쉽지 않다. 그리고 영세한 사업체의 경우 미신고 소득도 존재한다. 그뿐만 아니라 사업소득은 탈세되어 소득세 자료가 포착하지 못한 부분도 적지 않을 것으로 생각된다. 따라서 사업소득의 경우는 소득세 자료만 가지고서는 그 총액을 확정하기 어렵다. 여기서는 국민계정 통계에 의존하기로 한다. 국민계정에서 영업잉여, 임료, 준법인기업의 소득인출이라는 항목이 가계의 사업소득에 해당한다. 다만 영업잉여 속에 포함되어 있는 자가 주택에 대한 귀속임대료는 통상의 소득이 아니기 때문에 이를 추정하여 공제할 필요가 있다.<sup>15)</sup>

여기서 국민계정 통계를 이용할 때에는 모두 2008년 SNA로 이행한 결과로 update

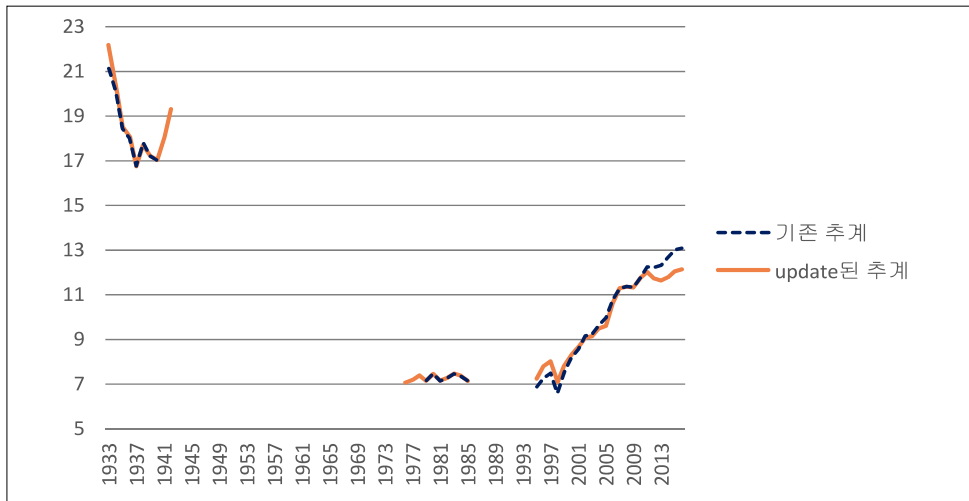
김낙년(2012a)의 방법이 결과를 크게 왜곡하지 않는다고 판단하고, 또 자료가 공개되지 않은 시기와의 일관성을 감안하여 기존의 방법을 유지하였다.

15) 한국은행의 산업연관표에서 주거서비스의 영업잉여를 구할 수 있고, 주택센서스의 자가 및 임대 주택 비율을 이용하여 이를 다시 실제와 귀속 임대료로 구분할 수 있다.

하였고, 소득구간을 보간할 때에는 전술한 GPC의 방법을 적용하였다. 해방 전 계열(1933-40)은 1942년까지 연장하였고,<sup>16)</sup> 마찬가지로 GPC 방법을 적용하였다. 이상의 근로소득, 금융소득, 사업소득의 합계를 전체소득(분모)으로 하고, 20세 이상 인구수를 전체 인원수로 하여 구한 상위 x%의 경계소득, 평균소득, 소득 비중을 <부표 2>에 제시하였다. 그리고 <그림 6>에는 상위 1%의 소득집중도를 기존 추계치(Kim and Kim, 2014)와 update된 결과를 비교하였다. 기존 추계는 2012년으로 끝나지만, 전체소득을 국민계정 통계에 의존하는 종래의 방식으로 2016년까지 연장한 것이다.

2009년 이전에 기존 추계와 update된 추계의 괴리가 나타난 것은 2000년대 이후 인구 통계가 다소 수정된 것 이외에는 국민계정 통계가 2008년 SNA로 이행한 것에 의해 영향을 받았다. <그림 1>에서 2008년 SNA로의 이행으로 개인소득이 얼마나 달라졌는지를 보였는데, 특히 1990년대에 개인소득이 많을 경우 4%까지 줄었고,<sup>17)</sup> 그로 인해 이 시기 소득집중도가 높아지게 된 것을 알 수 있다. 이에 대해 2010년대 이

<그림 6> 전체소득 상위 1% 소득집중도의 비교(단위: %): 기존 추계 vs. update된 추계



자료: Kim and Kim(2014); <부표 2>.

16) 1941-42년은 소득세 자료는 있지만, 전체 개인소득 통계가 없다. 여기서는 2010년 불변가격의 1인당 GDP(김낙년·박기주·박이택·차명수 편, 2018, pp. 734-735)와 소비자물가지수(박기주·김낙년, 2011)를 이용하여 개인소득을 1942년까지 연장하는 방법으로 구했다.

17) 1990년대에는 높은 이자율로 가계의 이자소득이 크게 늘어났는데, 그 중에는 보험회사 등이 운명한 투자자금의 수익이 포함되어 있다. 2008년 SNA에서는 이를 투자소득지급이라는 항목으로 파악하고 있다. 이것은 장차 개인으로 귀속될 것이지만, 당해 연도의 개인소득에서는 제외된다. 이를 조정한 결과 이 시기 개인소득은 <그림 1>에서 보인 것과 같이 줄어들게 되었다.

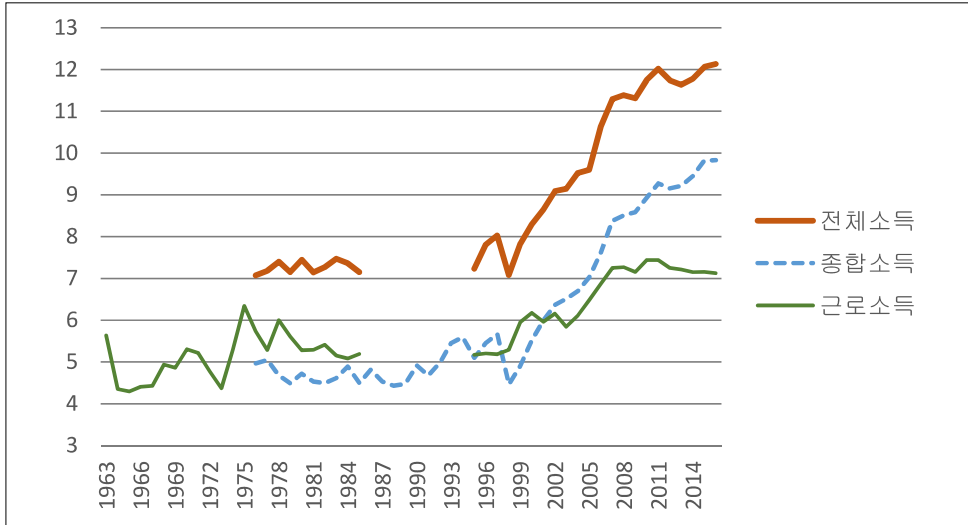
후에 나타난 괴리는 국민계정에서 전술한 근로소득의 과소평가를 보정한 데 기인한다. update의 결과 근로소득과 마찬가지로 1990년대 중엽 이후 소득집중도의 상승 추세가 다소 완화된 것으로 나타났다.

그런데 전체소득의 집중도는 근로소득 이외에 사업소득이나 금융소득과 같은 비 근로소득의 영향을 받았을 터인데, 양자를 분리해 볼 수 없을까? 이를 위해 <그림 7>은 전술한 근로소득과 전체소득의 상위 1% 집중도(<그림 4, 6>)와 함께 종합소득세 자료를 이용하여 구한 소득집중도를 제시하였다. 종합소득세 자료에는 근로소득이 일부 포함(2015년에 전체 신고소득의 28%가 근로소득)되어 있지만, 대부분의 사업소득과 일정 규모(예컨대 2,000만 원 이상) 이상의 금융소득이 파악되고 있어 비 근로소득의 집중도를 어느 정도 반영하는 것으로 볼 수 있다.

여기서 종합소득 상위 1%의 소득 비중이란 종합소득 신고자를 대상으로 하되, 전체소득의 집중도와 동일하게 20세 이상의 성인 인구를 기준으로 구했다. 그로 인해 <그림 7>에서 종합소득의 집중도는 전체소득의 집중도보다 낮게 나오며, 양자의 갭은 상위 1% 소득자 중에 종합소득을 신고하지 않은 고소득 근로소득자가 적지 않음을 말한다. 그래프를 제시하지 않았지만, <그림 7>를 상위 0.1%의 소득집중도로 바꾸어 보면 전체소득과 종합소득 집중도의 갭은 크게 줄어 양자가 근접한다. 그것은 성인인구 중에서 소득 상위 0.1%에 드는 자는 대부분 종합소득 신고 대상자가 되기 때문이다.

<그림 7>의 종합소득이 비 근로소득 집중도의 추이를 나타내는 것으로 본다면, 1970년대 중엽 이후 1980년대에 걸쳐 비 근로소득 집중도는 정체하고 있었는데, 같은 시기 근로소득 집중도가 하락하고 있었던 것과는 차이를 보였다. 비 근로소득 집중도는 1990년대 이후 완만하게 상승하다가 외환위기로 일시 하락한 후 빠른 상승을 지속해 오고 있다. 2000년대까지는 근로소득과 비 근로소득의 집중도가 함께 상승하고 있어 소득불평등 심화가 가속되었다. 그에 반해 2010년대 이후에는 근로소득의 집중도는 하락으로 돌아선 데 비해, 비 근로소득 집중도는 계속 상승하는 대조를 보이고 있다. 양자가 상쇄될 때에는 전체소득의 집중도가 정체되었지만, 최근에는 비 근로소득 집중도의 상승이 근로소득 집중도의 하락보다 더 빨라 전체소득의 집중도가 상승 기미를 보이고 있다.

〈그림 7〉 상위 1%의 소득 비중: 근로소득, 종합소득, 전체소득(단위: %)



- 주: 1) ‘근로소득’이란 근로소득을 대상으로 전체 근로자 중에서 상위 1%에 해당하는 자들의 소득이 전체 근로소득에서 차지하는 비중을 말한다.
- 2) ‘종합소득’이란 종합소득 신고자를 대상으로 소득 순으로 20세 이상 인구의 1%에 해당하는 자들의 소득이 전체소득에서 차지하는 비중을 말한다.
- 3) ‘전체소득’이란 양자를 통합한 소득을 대상으로 소득 순으로 20세 이상 인구의 1%에 해당하는 자들의 소득이 전체소득에서 차지하는 비중을 말한다.

자료: 〈부표 1, 2〉. 종합소득 집중도는 국세청, 『국세통계연보』 각 연도로부터 GPC의 보간법을 적용하여 산출하였다.

#### IV. 다른 추계에 대한 검토

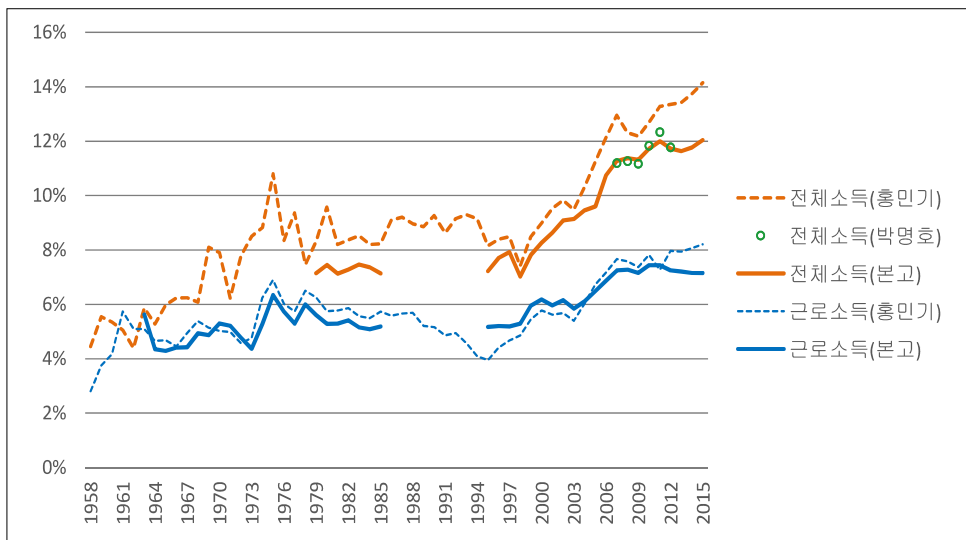
그런데 소득집중도에 관해서는 다른 추계 결과도 나와 있으며, 그 중에는 본고와 큰 차이를 보인 경우도 있다. 이러한 차이가 어디에 기인하는지를 검토하기로 한다. 〈그림 8〉은 전체소득과 근로소득의 상위 1%의 소득비중을 비교한 것이다.

먼저 박명호(2014)는 2007-2012년에 대해 전체소득의 집중도를 추정하였다. 본고보다 소득구간이 훨씬 세분된 국세청 내부 자료를 이용하였고, 종합소득과 근로소득의 중복을 배제한 통합소득 자료를 이용하였다. 이 점에서는 본고보다 더 나은 자료를 이용하였다고 할 수 있다. 소득집중도의 분모에 해당하는 전체소득에서는 미미한 차이가 있다. 본고가 국세청이 파악한 비과세소득도 전체소득에 포함하고 있는 반면, 그는 그것이 모든 비과세소득을 커버하지 못한다는 등의 이유로 제외하고 있다. 그는 매년의 전체소득을 밝히고 있지 않지만, 2007-2009년간은 본고와 차이가 미미할(비

과세소득의 경우 전체소득의 0.58% 정도) 것으로 생각된다. 다만 2010년 이후는 본고가 전술한 근로소득의 과소 파악을 보정하였기 때문에 그만큼 차이가 벌어질 것이다.

그가 차별화한 것으로 특히 강조하고 있는 것은 interpolation의 방법이다. 김낙년(2012a)이나 Kim and Kim(2014)이 소득구간을 분할할 때 Pareto interpolation(전술한 Pareto 1)의 방법으로 접근했는데, 그를 위해서는 먼저 상위 구간의 소득분포가 Pareto 분포로 잘 근사될 수 있는지를 검증해야 한다는 문제를 제기하고 있다. 그리고 일부 구간에서는 그것이 검증되지 않으므로 Pareto 분포와 같은 특정 소득분포를 가정하지 않는 mean split histogram의 방법을 채택하고 있다. 그렇지만 Pareto interpolation의 방법은 계수 값(a 또는 b)이 고정된 엄격한 Pareto 분포를 상정하고 있지 않으며, <표 1>이나 <그림 3>에서 보았듯이 구간별로 b 값이 달라지거나 GPC에서처럼 연속적으로 변하고 있다. 이것은 구간별 소득 정보를 가능한 한 반영할 수 있도록 유연한 Pareto 분포를 상정한 것이라고 할 수 있다.

〈그림 8〉 전체소득과 근로소득의 상위 1%의 소득집중도 추계결과의 비교



자료: <부표 1, 2>; 박명호(2014); 홍민기(2015a, 2015b).

그런데 Pareto interpolation과 mean split histogram의 어느 방법을 취하는가는 실제로는 크게 문제가 되지 않는다. 전술했듯이 두 방법으로 얻은 결과가 통상 미미한 차이밖에 보이지 않으며, 소득집중도를 추정한 연구들을 보면 두 방법이 혼용되고 있

음을 알 수 있다.<sup>18)</sup> 다만 <표 2>에서 보였듯이 한국에서는 자료상의 제약으로 상위 1% 소득 비중이나 경계소득의 추정에서 다소 차이가 날 수 있다는 점에는 유의할 필요가 있다. 실제로 <그림 8>을 보면 본고와 박명호(2014)의 상위 1% 집중도의 추정 결과는 차이가 미미한 것으로 나타났다. 그가 보간 방법의 차이를 강조하고 있지만, 그것이 추정된 결과에 미치는 영향은 한정적임을 알 수 있다.

이에 대해 홍민기(2015a, 2015b)는 근로소득과 전체소득의 집중도를 모두 추정하였는데, 그 수준이나 추이에서 본고와 큰 차이를 보이고 있다. 2010년대 이후 추이가 엇갈리게 나타난 것은 전술했듯이 본고가 국민계정 근로소득의 과소 편향을 보정하였지만, 그는 그렇게 하지 않은 것이 하나의 요인으로 포함되어 있다.<sup>19)</sup> 그렇지만 전체소득의 집중도를 보면 추이에서도 차이가 있지만 그 수준에서도 전체 시기에 걸쳐서 본고보다 상당히 높게 나온다.

그 원인을 추적해 보면 그가 분모에 해당하는 전체소득을 과소하게 추정했기 때문인데, 그것은 소득 개념을 잘못 설정한 데 기인한다.<sup>20)</sup> 그는 전체소득을 가계 부문의 순본원소득잔액(이것은 NNI와 동일)에서 공제 항목(고용주의 사회부담금, 자가 주택의 귀속 임대료, 금융중개서비스, 투자소득지급, 현물급여)을 빼는 방법으로 구했다. 그런데 가계 부문의 순본원소득잔액이란 소득의 원천(피용자보수, 영업잉여, 재산소득)에서 소득의 사용(재산소득)을 공제한 것인데, 여기서 공제된 재산소득이란 가계가 지급한 이자를 말한다.<sup>21)</sup> 통상 이자소득이라고 할 때 수취한 이자를 말하지 거기에서 지급한 이자를 뺀 순 이자를 말하지 않는다. 그 결과 그가 정의한 분모의 전체소득은 가계가 지급한 이자만큼 과소하게<sup>22)</sup> 설정되어 있는 반면, 분자에 해당하

18) Atkinson, Piketty and Saez(2011, Table 4)에 의하면 많은 나라가 Pareto interpolation에 따랐지만, mean split histogram의 방법에 따른 나라도 적지 않았다. WID에 실린 이들 국가의 데이터는 어느 interpolation의 방법에 따랐는지를 구별하지 않고 제시하고 있다.

19) 종합소득세와 근로소득 연말정산 자료를 통합할 때 두 자료에 중복해서 실린 근로소득자를 추산해서 배제해야 하는데, 이를 위해 그는 국세청이 발표한 통합소득 자료(2007-2012)를 이용하였다. 그렇지만 통합소득 자료가 없는 시기에는 중복을 배제하는 절차를 밟지 않았기 때문에 그만큼 상위 소득집중도가 과대하게 추계된다.

20) 그는 소득구간을 보간할 때 전술한 Pareto 3의 방법을 취하고 있다. 이것은 타 연구와 다른 점이지만, 소득 개념의 적용에 포함된 문제에 비하면 미미한 차이에 불과하다.

21) 홍민기(2015a, p. 6)는 “순본원소득잔액은 피용자보수, 개인영업잉여, 재산소득을 합한 것”이라고 언급한 데에서 알 수 있듯이 순본원소득잔액에는 지급한 재산소득(여기서는 지급이자)이 공제된다는 사실을 인식하지 못한 것으로 보인다.

22) 예컨대 국민계정에 따르면 2010년 가계의 이자소득(원천)은 49조원이고 이자소득(사용)은 45조인데, 양쪽에 모두 금융중개서비스가 포함되어 있다. 그의 방식에 따르면 이자소득은 지급이자



는 상위  $x\%$ 의 소득에는 그렇지 않다. 즉 분자와 분모의 소득 범위가 서로 조응하지 않게 설정되었기 때문에 그로부터 산출한 소득집중도가 왜곡되었던 것이다.

그가 공제한 항목 중에서 선행 연구(김낙년, 2012a; 박명호, 2014)와 다른 것은 현물 급여를 공제하고 있다는 것이다. 그는 국민계정의 임금 및 급여 중에는 현물이나 기타노동비용이 포함되어 있다고 보고, 이를 공제하기 위해 그 비중을 『기업체노동비용조사』에 의거하여 추정하고 있다(홍민기, 2015a, p. 6). 그렇지만 고용주의 입장에서 노동비용으로 파악되는 것이 모두 국민계정의 임금 및 급여에 산입되는 것은 아니다. 이들 중에서 그 혜택이 특정된 개인으로 귀속되는 경우는 그러하지만, 그렇지 않은 경우 국민계정에서는 대체로 중간소비로 분류되어 소득으로 산입되지 않기 때문이다.<sup>23)</sup> 그는 현물 급여를 공제함으로써 소득세 자료와 국민계정의 소득을 일치시켰다고 주장하지만, 노동비용 중 개인의 소득으로 볼 수 있는 경우에는 두 자료 모두 소득으로 파악하고 있으며, 소득세의 과세 대상이 되지 않는 경우는 국민계정에서도 대부분 중간소비로 파악되기 때문에 두 자료의 소득의 범위는 그가 상정하듯이 그렇게 다르지 않다. 따라서 국민계정에서 그가 현물 급여라고 생각하여 공제한 것은 분모의 전체소득을 과소하게 산정하게 만든 또 하나의 요인이 된다.<sup>24)</sup>

〈그림 8〉에서 볼 수 있듯이 그는 본고에 비해 1958년까지 소급한 추정 결과를 제시하고 있다. 근로소득세 자료가 없는 1986-94년뿐만 아니라,<sup>25)</sup> 종합소득세가 실시(1975년)되기 이전 시기에 대해서도 소득집중도의 추정 결과를 제시하고 있다. 그에 따르면 전체소득 집중도가 1975년까지 급등한 것으로 나온다. 여기서 먼저 드는 의문은 종합소득세가 아직 시행되지 않은 시기에 전체소득의 집중도를 어떻게 구했을까 하는 점이다. 이 시기는 분리과세가 시행되었기 때문에 소득원이 여럿 있을 경우 개

(45조원) 만큼 과소하게 파악된다. 더구나 거기에서 다시 금융중개서비스(그의 추정에 따르면 8조원)도 공제해야 하므로 그의 이자소득은 -4조원 (=49-45-8), 즉 마이너스가 되는 모순에 빠진다.

23) 그는 『기업체노동비용조사』에 의거하여 노동비용 중의 일부(예컨대 자녀의 학비보조비와 같은 복리후생비)는 과세 대상으로서 소득세 자료에서 포착되지만, 나머지 노동비용(부정기적인 식사나 교통 비용 지원 또는 휴양 시설 유지, 교육훈련비 등)은 과세 대상이 아니지만 국민계정에서 현물 급여로 파악되기 때문에 이를 제외해야 한다고 주장한다. 그렇지만 후자의 경우 국민계정에서는 소득이 아니라 중간소비로 파악된다.

24) 금융중개서비스와 귀속임대료의 추정에서도 오차가 생기고 있다. 예컨대 금융중개서비스를 추정할 때 그는 실제의 이자소득이 소득세 자료에서 파악된 이자소득의 1.15배로 가정하였지만, 한국은행은 양자를 동일한 것으로 보고 있기 때문에 여기에서도 차이가 있다.

25) 이 시기의 소득집중도에 대해서는 단순한 가정으로 도출한 것이므로 검토하지 않는다.

인별로 통합한 소득을 알 수 없기 때문이다. 그의 방법은 근로소득, 사업소득, 부동산소득의 각 소득구간별 통계를 단순 집계하고, 그로부터 소득집중도를 구한 것으로 보인다. 만약 위의 3개 소득을 각각 100만원씩 번 개인이 있을 경우 그의 소득은 300만원이 아니라 100만원을 번 3명의 소득자로 파악된다. 이것은 소득집중도를 실제보다 훨씬 낮게 추정한다. 즉 그의 추계가 고도성장기에 소득불평등이 급속히 심화된 것으로 나온 것은 자료를 잘못 이용한 결과일 뿐이며 실태를 왜곡한 것이다.<sup>26)</sup> 김낙년(2012a)이나 본고가 종합소득세가 시행되지 않은 시기에 대해 소득집중도를 추정하지 않은 것은 이러한 자료상의 한계를 고려했기 때문이다.

## V. 맺음말

이하에서는 본고에서 드러난 사실을 요약하고, 소득불평등의 실태를 구명하기 위해 앞으로 어떤 과제가 남아 있는지를 간단히 언급하기로 한다.

Kim and Kim(2014) 이후 한국의 소득집중도를 추계하는데 이용되는 자료 상황에 적지 않은 변화가 생겼다. 전체 개인소득을 추출하는데 이용되었던 국민계정 통계가 2008 SNA로 이행하였고, 소득세 자료가 2009년 이후 보다 더 충실해졌다. 소득구간의 interpolation도 GPC라는 새로운 방법이 제안되었다. 본고는 이들 자료와 방법의 변화를 반영하여 기존의 소득집중도 추계를 update하고, 계열을 2016년까지 연장하였다.

update된 소득집중도는 기존 추계에 비해 몇 가지 특징을 보이고 있다. 먼저, 근로소득이든 전체소득이든 1990년대 후반의 소득집중도가 기존 추계보다 다소 높아졌다. 이것은 소득 자료의 보정에 주로 기인하는 것인데, 그로 인해 1990년대 이후 소득불평등의 급격한 상승 추세는 다소 완화되는 것으로 수정되었다.

26) 그런데 이 시기 집중도가 급등한 것은 왜일까? 종합소득세를 본격 시행하기 전인 1969년부터 연 소득 500만 원을 넘는 최상층 소득자에 한정해서 합산한 소득에 과세하기 시작하였다. 그리고 1973년에 그 한도를 300만 원으로 낮추었고, 1975년에 종합소득세가 시행됨에 따라 그 한도가 없어졌다. 이 시기는 인플레이와 함께 명목소득이 급등한 시기였기 때문에 이 한도를 넘는 자들이 급속히 늘어났다. 그의 소득집중도가 계속 상승하고 특히 1969, 1973, 1975년에 급등한 것으로 나타난 것은 실태와는 무관하며 이러한 제도상의 변화를 반영하는 것에 불과하다. 그 결과 <그림 8>에서 보이듯이 이 시기 그의 전체소득 집중도가 급등함에도 불구하고 근로소득의 집중도는 그렇지 않아 모순되는 양상을 보이고 있다. 근로소득의 경우는 분리과세 자료를 이용하더라도 전체소득에 대해 전술한 문제는 발생하지 않기 때문이다.

그런데 여기서 주목되는 것은 2010년대 이후 소득불평등의 추세에 변화가 생겼다는 점이다. 먼저 근로소득의 경우 소득집중도가 2010년 이후 계속해서 떨어지고 있다. 그 배후에는 근로자의 하위 50%의 소득 증가가 상위 소득자보다 빨랐기 때문이다. 이에 비해 사업소득이나 금융소득으로 이루어진 비 근로소득의 집중도는 이전보다는 악화되었지만 상승 추세를 이어가고 있다. 그에 따라 양자를 통합한 전체소득 집중도는 2010년대 전반에는 하락 또는 정체 기미를 보였다가 최근에는 다시 높아지고 있다. 근로소득의 불평등이 완화되고 있음에도 불구하고 이를 상쇄하고 남을 정도로 비 근로소득에서 불평등이 커지고 있기 때문이다.

2010년대 이후 소득불평등 추세가 변했다는 본고의 추계 결과는 어느 자료를 더 신뢰할 수 있는가라는 문제를 제기하고 있다. 근로소득 집중도를 계산할 때 분모에 들어가는 전체소득을 종래와 같이 국민계정의 임금 및 급여를 이용할 경우 소득집중도가 계속 상승하여 본고와 대조된 결과가 나온다. 그것은 국민계정의 임금 및 급여가 2010년대 이후 과소 파악의 편향이 누적되어 왔으며, 그 결과 현재에는 전수 집계된 국세청의 근로소득에 비해서도 6%나 적다는 납득하기 어려운 문제를 안고 있기 때문이다. 소득집중도 추계는 분자와 분모를 모두 동일한 소득세 자료에 의거함으로써 추계의 일관성을 높일 필요가 있다. 나아가 상위의 소득집중도뿐만 아니라 중하위 소득자를 포함한 전체 분포로 확장하는 것이 가능하다. 본고에서는 근로소득에 대해 그러한 시도를 하였지만, 사업소득과 같은 비 근로소득의 경우는 자료의 한계로 난관이 많다. 앞으로 전체소득에 대해서도 소득세 자료에 의거하여 추계의 일관성을 높이는 것이 과제라고 할 수 있다.

또 하나 본고의 결과는 노동부의 고용형태별 임금실태조사와 같은 사업체를 대상으로 하는 임금 조사가 임금소득의 분포를 얼마나 잘 드러내고 있는가에 대해서도 의문을 제기하고 있다. 일정 규모 이상의 사업체로 한정할 경우 하위 소득의 근로자들이 조사 대상에서 누락될 것이라는 점은 예상할 수 있지만, 문제는 거기에 그치지 않는다. 본고와 사업체조사의 근로소득의 구간별 분포를 비교해 보면, 상위 구간으로 갈수록 사업체조사의 포착률이 급속히 떨어지고 중위 소득자가 상대적으로 과대 파악되는 샘플링 문제가 드러났다.

이것은 정도의 차이는 있지만, 통계청의 가계조사에 대해서도 마찬가지로 지적되어 온 문제이다(김낙년·김종일, 2013). 최근 통계청이 가계금융복지조사를 개편할 때 개인의 소득 등에 관해서는 국세청의 과세 자료로 보완하는 시도를 하고 있다. 가계조사, 사업체조사, 소득세 자료 등은 소득분배의 한 측면을 보여주지만, 그 결과가

서로 어긋나는 경우가 적지 않다. 그 차이가 어디에서 온 것인지를 추적해서 이해하고, 각 자료의 장단점을 보완해서 활용하는 것이 소득분배의 실태를 밝히는데 중요한 과제가 된다.

## ■ 참 고 문 헌

1. 고용노동부, 『고용형태별 근로실태조사』, (마이크로 데이터), 각년도.
2. \_\_\_\_\_, 『사업체노동력조사』, 각년도.
3. 국세청, 『국세통계연보』, 각년도.
4. 김낙년, “한국의 소득집중도 추이와 국제비교, 1976-2010: 소득세 자료에 의한 접근,” 『경제분석』, 제18권 제3호, 2012a, pp. 75-114.
5. \_\_\_\_\_, “한국의 소득 불평등, 1963-2010: 근로소득을 중심으로,” 『경제발전연구』, 제18권 제2호, 2012b, pp. 125-158.
6. \_\_\_\_\_, “2013년 소득세제 개편과 계층별 소득세 부담률,” 『재정학연구』, 제7권 제2호, 2014, pp. 59-93.
7. \_\_\_\_\_, “한국의 소득집중도: update, 1933-2016,” 낙성대경제연구소 워킹페이퍼 WP2018-01, 2018.
8. 김낙년 · 김종일, “한국 소득분배 지표의 재검토,” 『한국경제의 분석』, 제19권 제2호, 2013, pp. 1-64.
9. 김낙년 · 박기주 · 박이택 · 차명수 편, 『한국의 장기통계』, 해남, 2018.
10. 박기주 · 김낙년, “해방 전(1907-1939) 소비자물가지수의 추계,” 『경제분석』, 제17권 제1호, 2011, pp. 131-168.
11. 박명호, “소득집중도 추이 분석,” 박명호 · 전병목, 『소득분배 변화와 정책과제: 소득집중도와 소득이동성 분석을 중심으로』, 한국조세정책연구원, 2014.
12. 정 혁, “한국의 소득불평등 구조의 실증적 이해와 포용적 성장전략,” 황수경 외, 『소득분배와 경제성장』, 경제 · 인문사회연구회, 2017.
13. 통계청, 『경제활동인구조사』, 각년도.
14. \_\_\_\_\_, KOSIS (<http://kostat.go.kr/portal/index/statistics.action>).
15. 한국은행, ECOS (<http://ecos.bok.or.kr/>).
16. \_\_\_\_\_, 『한국의 국민소득』, 1982.
17. 홍민기, “최상위 소득비중의 장기 추세(1958-2013),” 『경제발전연구』, 제21권 제4호, 2015a, pp. 1-34.
18. \_\_\_\_\_, “최상위 임금 비중의 장기추세(1958-2013),” 『산업노동연구』, 제21권 제1호, 2015b, pp. 191-220.
19. Atkinson, Anthony B., “Top Incomes in the UK over the 20<sup>th</sup> Century,” *Journal of the Royal Statistical Society*, Vol. 168, No. 2, 2005, pp. 325-343.
20. Atkinson A., T. Piketty and E. Saez, “Top Incomes in the Long Run of History,” *Journal of Economic Literature*, Vol. 49, No. 1, 2011, pp. 3-71.

21. Blanchet, Thomas, Juliette Fournier, and Thomas Piketty, "Generalized Pareto Curves: Theory and Applications," WID.world Working Paper 2017/3, 2017.
22. Feenberg, Daniel and James Poterba, "Income Inequality and the Incomes of Very High Income Taxpayers: Evidence from Tax Returns," NBER Working Paper No. 4229, 1992.
23. Kim, Nak Nyeon and Jongil Kim, "Top Incomes in Korea, 1933-2010: Evidence from Income Tax Statistics," WID.world Working Paper 2014/2, 2014.
24. Moriguchi, C. and E. Saez, "The Evolution of Income Concentration in Japan: Evidence from Income Tax Statistics, 1886-2005," *The Review of Economics and Statistics*, Vol. 90, No. 4, 2008, pp.713-734.
25. Piketty, Thomas, and Emmanuel Saez, "Income Inequality in the United States, 1913-1998," *Quarterly Journal of Economics*, Vol. 118, No. 1, 2003, pp.1-39.
26. The World Wealth and Incomes Database, (<http://wid.world>).

〈부표 1〉 근로소득의 집중도와 경계소득 및 평균소득

	소득 비중(%)					경계소득(천 원)					평균소득(천 원)				
	P99.9-100	P99.5-100	P99-100	P95-100	P90-100	P99.9	P99.5	P99	P95	P90	P99.9-100	P99.5-100	P99-100	P95-100	P90-100
1963	1.221	3.670	5.630	15.685	26.361	27,105	12,004	10,313	6,082	5,916	34,313	20,628	15,824	8,816	7,409
1964	0.814	2.682	4.349	13.905	25.472	17,191	11,281	8,431	6,739	6,702	23,646	15,585	12,636	8,080	7,400
1965	0.843	2.656	4.292	13.205	23.671	18,339	10,998	8,683	6,182	6,134	24,740	15,592	12,600	7,753	6,949
1966	0.963	2.796	4.407	14.055	25.753	21,483	12,251	9,638	7,815	7,766	32,065	18,624	14,677	9,362	8,577
1967	0.960	2.791	4.428	13.817	24.609	23,361	14,092	10,904	8,505	7,831	36,313	21,124	16,757	10,458	9,313
1968	1.124	3.084	4.939	14.337	24.432	25,156	16,156	12,987	8,076	7,819	44,243	24,274	19,440	11,286	9,617
1969	1.007	3.006	4.865	15.020	24.401	27,557	18,005	15,384	8,615	8,005	44,369	26,488	21,437	13,236	10,751
1970	1.048	3.204	5.304	16.574	27.216	31,368	21,019	17,368	9,913	9,308	46,830	28,640	23,700	14,813	12,162
1971	1.124	3.228	5.215	16.036	25.619	34,297	20,161	16,546	10,232	8,304	52,081	29,920	24,167	14,863	11,873
1972	0.978	2.941	4.775	15.164	24.969	31,808	20,326	16,884	10,402	9,476	48,267	29,044	23,575	14,975	12,328
1973	0.918	2.728	4.371	13.602	23.991	36,998	21,823	18,280	12,621	12,483	55,430	32,929	26,381	16,418	14,480
1974	1.242	3.451	5.282	15.021	25.777	58,584	25,622	19,633	13,309	13,136	76,310	42,411	32,462	18,463	15,841
1975	1.256	4.191	6.343	16.080	25.364	24,779	34,474	22,110	11,895	11,369	77,740	51,886	39,271	19,909	15,702
1976	1.037	3.579	5.732	15.136	23.978	48,829	37,583	24,001	12,939	11,907	71,934	49,661	39,767	21,004	16,636
1977	1.060	3.230	5.289	15.350	24.358	54,992	34,711	28,532	14,401	13,493	80,437	49,007	40,116	23,287	18,477
1978	1.241	3.767	5.998	17.681	27.586	73,770	44,412	34,525	20,041	15,712	107,367	65,179	51,882	30,591	23,863
1979	1.201	3.564	5.608	16.867	27.406	77,153	46,116	35,584	21,364	19,299	113,592	67,439	53,063	31,917	25,930
1980	1.004	3.234	5.279	16.118	25.725	65,315	42,526	34,585	20,360	15,819	91,762	59,129	48,256	29,465	23,514
1981	1.104	3.341	5.790	15.957	25.728	72,376	39,634	32,600	19,778	17,027	100,675	60,917	48,224	29,094	23,455
1982	1.028	3.375	5.413	16.533	27.037	72,834	44,300	35,284	21,892	19,216	98,063	64,382	51,629	31,538	25,788
1983	1.049	3.206	5.156	15.771	25.375	73,210	47,609	38,176	23,692	19,120	112,462	68,715	55,257	33,806	27,197
1984	0.964	3.158	5.083	15.290	24.491	74,178	51,792	38,201	23,881	19,190	109,194	71,549	57,581	34,638	27,742
1985	1.083	3.230	5.188	15.729	25.093	82,582	51,739	40,459	25,140	19,275	125,141	74,672	59,970	36,360	29,004
1995	1.273	3.231	5.172	16.578	28.032	145,187	95,663	80,698	57,254	46,148	287,018	145,713	116,607	74,754	63,200
1996	1.022	3.128	5.204	16.404	27.789	160,190	111,307	90,188	58,056	52,076	245,271	150,145	124,898	78,737	66,693
1997	0.922	3.043	5.187	16.676	28.423	157,526	113,122	92,613	59,114	53,490	219,566	144,936	123,527	79,421	67,684
1998	1.163	3.225	5.293	16.239	27.498	155,260	105,501	85,618	54,236	49,689	266,690	147,859	121,343	74,460	63,043
1999	1.337	3.699	5.957	17.015	28.760	198,705	116,520	93,607	55,815	53,719	311,500	172,333	138,766	79,276	66,998
2000	1.397	3.850	6.176	17.849	29.510	210,761	120,632	104,713	58,400	55,410	338,448	186,605	149,671	86,505	71,511
2001	1.398	3.723	5.965	17.819	29.400	201,383	118,145	104,853	60,167	55,678	346,139	184,348	147,665	88,223	72,779
2002	1.430	3.880	6.157	18.276	30.011	220,232	125,446	111,732	62,818	59,143	367,645	199,503	158,279	93,970	77,154
2003	1.330	3.687	5.842	17.927	29.351	216,819	123,984	110,150	64,143	59,280	353,881	196,213	155,444	95,408	78,101
2004	1.470	3.873	6.113	19.017	31.507	214,405	137,530	110,537	75,935	57,835	398,815	210,095	165,826	103,171	85,466
2005	1.748	4.193	6.482	19.418	31.910	234,991	139,431	114,716	75,619	59,411	481,699	231,163	178,667	107,050	87,957
2006	1.886	4.468	6.872	20.413	33.729	267,810	160,749	128,664	77,018	67,990	523,944	248,272	190,945	113,433	93,711
2007	2.168	4.828	7.250	20.800	33.537	277,115	152,626	125,525	80,070	64,353	611,944	272,511	204,618	117,414	94,655
2008	2.133	4.846	7.271	20.989	33.751	285,826	150,883	124,274	80,410	63,878	594,898	270,297	202,811	117,085	94,136
2009	2.078	4.771	7.155	20.622	33.197	280,133	148,053	120,863	78,586	62,497	575,589	264,303	198,195	114,247	91,956
2010	2.138	4.971	7.442	21.279	33.882	301,104	155,582	126,403	80,466	63,014	598,414	278,296	208,347	119,137	94,850
2011	2.156	4.984	7.440	21.204	33.630	309,468	157,608	128,401	80,975	63,350	616,013	284,847	212,582	121,180	96,095
2012	2.050	4.850	7.253	20.739	32.961	312,141	160,188	129,497	82,142	64,357	605,186	286,292	214,090	122,429	97,291
2013	2.037	4.823	7.210	20.556	32.658	320,776	163,438	132,124	83,510	65,478	617,749	292,523	218,656	124,674	99,036
2014	2.021	4.792	7.152	20.343	32.279	328,891	165,254	133,405	84,193	66,007	626,266	296,978	211,619	126,077	100,026
2015	1.974	4.774	7.155	20.377	32.372	339,759	170,603	137,390	86,409	67,920	625,291	302,405	226,634	129,083	102,531
2016	1.981	4.761	7.126	20.167	32.012	348,612	175,216	141,162	87,912	69,496	649,067	311,929	233,462	132,134	104,870

주: 1) 소득금액은 2015년 불변가격이다.

2) 예컨대 P99.9-100은 상위 0.1%에 든 근로자의 평균소득 또는 소득비중을 뜻한다.

3) P99.9는 상위 0.1%에 들기 위한 경계소득(threshold income)을 뜻한다.

〈부표 2〉 전체소득의 집중도와 경계소득 및 평균소득

	소득 비중(%)					경계소득(천 원)					평균소득(천 원)				
	P99.9- 100	P99.5- 100	P99- 100	P95- 100	P90- 100	P99.9	P99.5	P99	P95	P90	P99.9- 100	P99.5- 100	P99- 100	P95- 100	P90- 100
1933	8.376	16.135	22.168			50,679	19,040	14,997			116,961	45,061	30,955		
1934	8.316	15.612	20.436			46,074	15,719	12,857			115,793	43,479	28,456		
1935	7.580	14.332	18.553			53,328	16,518	12,154			117,741	44,521	28,818		
1936	7.365	13.967	18.111			49,983	15,874	11,653			112,686	42,741	27,710		
1937	7.102	13.117	16.769			55,491	17,838	10,843			128,336	47,407	30,302		
1938	7.750	13.905	17.726			57,560	16,804	10,265			136,342	48,925	31,185		
1939	7.424	13.476	17.214			41,936	14,841	9,587			114,038	41,401	26,442		
1940	7.421	13.382	17.036			59,014	17,991	10,739			135,696	48,938	31,151		
1941	7.668	14.262	17.987			55,610	21,434	11,998			146,452	54,477	34,355		
1942	7.517	14.929	19.314			52,885	19,938	11,381			120,954	48,046	31,079		
1976	2.230	4.944	7.075	16.176	25.683	51,613	27,105	17,723	9,756	9,657	113,751	50,437	36,088	16,501	13,099
1977	2.318	5.001	7.181	16.578	25.982	55,999	28,879	20,955	10,873	10,177	129,733	55,977	40,187	18,554	14,540
1978	2.179	5.033	7.404	18.441	28.815	66,542	34,848	26,602	13,506	13,022	138,517	63,988	47,070	23,446	18,318
1979	2.043	4.773	7.148	18.608	28.078	66,247	35,885	28,016	13,713	12,102	134,276	52,727	46,970	24,455	18,450
1980	1.996	4.927	7.444	19.731	29.183	58,483	33,736	25,954	13,498	9,650	114,422	56,498	42,679	22,624	16,731
1981	1.881	4.698	7.144	19.243	29.306	60,556	31,942	25,360	13,167	10,954	108,076	53,995	41,060	22,118	16,843
1982	1.788	4.738	7.272	20.055	29.884	63,547	34,523	27,078	14,453	10,213	106,807	56,599	43,435	23,958	17,850
1983	1.838	4.866	7.469	20.440	30.198	67,025	38,138	29,594	15,494	10,127	116,541	61,694	47,348	25,915	19,143
1984	1.859	4.839	7.363	19.761	29.894	76,711	42,467	31,381	16,439	13,160	129,454	67,406	51,281	27,525	20,820
1985	1.847	4.694	7.148	19.400	29.236	75,848	41,497	32,410	17,258	13,305	136,102	69,195	52,680	28,595	21,547
1995	1.975	4.730	7.234	20.554	31.825	156,929	87,990	70,614	43,413	29,454	310,035	148,486	113,552	64,524	49,955
1996	2.147	5.101	7.811	21.217	35.030	185,673	104,098	79,026	49,464	42,809	359,340	170,718	130,723	71,014	58,623
1997	2.209	5.199	8.026	21.690	35.683	185,056	105,356	83,290	49,893	43,350	369,593	173,958	134,273	72,575	59,696
1998	1.822	4.564	7.080	20.109	33.817	158,920	91,457	68,801	45,900	39,919	285,539	143,051	110,952	63,021	52,992
1999	2.144	5.140	7.828	21.079	35.150	166,693	99,265	73,600	47,982	41,460	341,914	163,930	124,829	67,222	56,048
2000	2.347	5.459	8.295	22.107	35.963	200,081	103,384	85,847	49,774	43,132	394,612	183,595	139,470	74,344	60,471
2001	2.564	5.787	8.647	23.076	36.805	214,300	103,033	89,149	49,916	42,396	430,056	194,152	145,058	77,420	61,740
2002	2.743	6.174	9.086	24.144	37.751	224,768	109,161	94,083	51,057	42,306	467,337	210,419	154,820	82,285	64,330
2003	2.812	6.263	9.142	24.637	38.385	222,478	109,659	93,603	52,196	42,555	479,956	212,805	156,044	84,107	65,520
2004	3.055	6.513	9.523	25.428	38.561	211,283	120,544	93,152	57,367	38,249	530,411	226,128	165,324	88,286	66,941
2005	3.255	6.686	9.600	25.116	37.099	248,893	126,337	98,777	51,786	37,728	578,847	237,751	170,701	89,317	65,965
2006	3.589	7.381	10.637	28.067	41.828	271,708	130,772	101,396	65,247	45,533	654,899	269,335	194,087	102,425	76,322
2007	3.996	7.985	11.292	28.458	42.352	293,673	140,508	110,124	63,595	43,074	741,592	296,409	209,590	105,638	78,605
2008	4.006	8.049	11.385	28.798	42.874	294,920	139,026	109,489	63,594	42,180	728,974	292,911	207,154	104,799	78,013
2009	3.958	8.012	11.313	28.522	42.474	294,062	137,226	107,410	62,473	40,955	714,696	289,321	204,279	103,003	76,694
2010	4.175	8.355	11.745	29.074	43.032	315,664	144,339	112,455	63,246	41,878	772,648	309,212	217,333	107,598	79,629
2011	4.340	8.620	12.018	29.246	43.128	340,024	147,393	115,504	63,715	42,628	820,759	325,986	227,251	110,604	81,553
2012	4.179	8.394	11.740	28.851	42.742	338,231	148,212	116,846	64,596	43,886	808,329	324,681	227,050	111,598	82,665
2013	4.082	8.300	11.632	28.649	42.545	347,177	152,460	119,711	66,148	45,624	813,379	330,746	231,762	114,165	84,772
2014	4.128	8.430	11.779	28.831	42.695	362,893	156,669	121,893	67,514	46,195	835,755	341,386	238,483	116,751	86,446
2015	4.288	8.679	12.061	29.155	43.099	387,073	164,241	126,196	70,237	48,384	896,078	362,754	252,037	121,850	90,066
2016	4.294	8.741	12.130	29.189	43.196	400,977	169,003	128,468	72,003	49,795	915,265	372,658	258,557	124,438	92,075

주: 〈부표 1〉과 동일.

## Top Income Shares in Korea: The Update, 1933-2016

Nak Nyeon Kim\*

### Abstract

This study updates Korea's top income shares in the World Wealth and Income Database until 2016. In the meantime, the national account statistics were revised in accordance with the SNA in 2008 and income tax data have become more substantial, including information on tax exemptions. A new interpolation method of income statistics by interval called GPC was proposed. Top income shares are updated to reflect these changes in data and method. Some interesting findings are as follows. Since the mid-1990s, the sharp rise in income concentration has been somewhat moderated. What is noteworthy is that inequality in earned income has been improved since 2010. This is because the income of workers in the bottom 50% has risen faster than the top 10%. On the other hand, the concentration of income other than earned income, which is mainly composed of business income and financial income, is weaker than before but continues to rise. As a result, the concentration of total income stagnated in the first half of the 2010s, but recently turned to increase, because the increase in income concentration other than earned income has been faster than the fall in concentration of earned income.

**Key Words:** top income shares, income tax statistics, interpolation, inequality

**JEL Classification:** J3, N3

---

*Received: March 12, 2018. Revised: April 4, 2018. Accepted: April 11, 2018.*

\* Professor, Department of Economics, Dongguk University. 30, Pildong-ro 1 gil, Jung-gu, Seoul 04620, Korea, Phone: +82-2-2260-3273, e-mail: nnkim@dongguk.edu