

Naksungdae Institute of Economic Research

Working Paper Series

한국의 개인소득 분포: 소득세 자료에 의한 접근

김낙년

Working Paper 2014-08

Dec, 2014

Naksungdae Institute of Economic Research

5, Bongcheon-ro 31, Kwanak-gu,

Seoul, 151-843, South Korea

- I. 머리말
- II. 자료와 방법
- III. 소득구간별 분포
 - 1. 통합소득
 - 2. 근로소득
 - 3. 사업소득
 - 4. 금융소득
- IV. 개인소득의 분포와 분위별 소득비중
- V. 맺음말

I. 머리말

우리나라에서 개인소득은 어떻게 분포되어 있을까? 개인의 소득을 소득 순으로 정렬할 경우 어떤 분포가 그려질까? 전체 인원수를 10분위로 나눌 때 각 분위의 소득 비중은 어떻게 될까? 상위 10%를 다시 나누어 상위 1% 또는 0.1%나 0.01%로 세분할 경우 이들 최상층의 소득비중은 어떻게 될까? 그리고 개인소득은 어떤 종류의 소득(근로소득, 사업소득, 재산소득)으로 이루어졌으며 계층에 따라 소득 구성이 어떻게 달라질까?

이상은 소득분포를 따질 때 제기되는 기본적인면서도 흥미 있는 질문이라고 생각한다. 지금까지 이에 대한 답은 가계조사에서 구할 수 있었다. 예컨대 통계청의 가계동향조사는 약 1만 가구의 샘플을 대상으로 가구원의 소득을 조사했다. 이 조사는 개인이 복수의 소득원을 가진 경우 이들을 통합하여 제시하므로 이로부터 개인소득의 분포를 구할 수 있다. 그렇지만 이 자료의 이용에는 샘플조사로서의 한계가 있다. 10분위 정도의 소득비중은 구할 수 있지만 이를 세분하여 예컨대 상위 1% 또는 0.1%로 가면 샘플이 적어져 소득 실태를 보이기 어렵게 된다. 더구나 가계조사에서는 최상층의 소득이 대부분 누락되었다는 점이 지적되고 있다(김낙년·김종일 2013). 그뿐만 아니라 후술하는 바와 같이 그 이하 계층에 대해서도 소득분포가 실태와 상당히 괴리되어 있다는 의문이 제기되고 있다.

여기서는 2010년의 소득세 자료를 이용하여 앞에서 제기한 과제에 접근하고자 한다. 근래에 소득세 자료를 이용하여 상위 소득자의 소득비중(top income shares)을 구하는 연구가 활발하게 이루어지고 있고, 그러한 접근의 유용성이 확인되고 있다¹. 다만 이러한 접근은 상위 10%에 한정되

¹ T. Piketty가 주도한 이 연구는 프랑스(Piketty 2003), 미국(Piketty and Saez 2003), 영국(Atkinson 2005),

고 있어 그 이하 계층의 소득분포가 어떠한지에 대해서는 알 수가 없었다. 여기서는 고찰의 범위를 상층뿐만 아니라 중층과 하층을 포함한 전체 계층으로 확대하고자 한다.

이러한 시도가 가능해진 것은 소득세 자료가 과거와는 달리 충실해졌기 때문이다. 예컨대 국세청의 『국세통계연보』(이하 『연보』로 줄임)는 2008년 이전에는 종합소득세 신고 자료와 근로소득 연말정산 자료를 신고 있지만, 면세점 이하의 소득자 정보가 빠져 있다. 그리고 이들 통계로는 종합소득세 신고를 하지 않는 각종 사업소득자나 연말정산 대상이 아닌 일용근로자 등의 소득은 알 수가 없었다. 금융소득도 금융소득 종합과세 대상이 아닌 경우 마찬가지였다. 이러한 자료 상황으로는 최상층의 소득을 파악하는 데에는 큰 문제가 없지만, 고찰의 대상을 전체 계층으로 확대하기는 어려웠다. 그렇지만 최근의 『연보』는 소득세법과 같이 법령에 의해 파악된 소득을 대부분 신고 있고, 면세점 이하의 소득 통계도 나와 있어 소득의 전모를 상당히 드러내고 있다고 볼 수 있다. 다만 소득세 자료에 나오는 각종 소득은 소득구간별 통계로 제시되는데, 소득의 기준이 다르거나 소득구간이 다르게 설정되어 있는 경우가 많다. 이들은 동일한 기준으로 바꾸고 구간별 통계를 더하거나 뺄 경우 소득구간을 조정할 필요가 생기는데, 그를 위해서는 후술하는 보간법을 이용하기로 한다.

한편, 소득세 통계로 개인소득을 파악하는데 한계도 있다. 개인이 복수의 소득원(예컨대 근로소득, 사업소득, 금융소득)을 가지고 있는 경우 개인소득은 이를 통합해야 하지만, 소득세 자료에서는 이들이 별개의 소득자로 파악되는 경우가 있기 때문이다. 우리나라에서 종합소득 신고자의 경우에는 모든 종류의 소득이 개인별로 통합되어 있어 그러한 문제가 없다. 그렇지만 이러한 종합소득 신고자는 일부에 불과하다. 최근 국세청은 『연보』에는 없지만 국회의원의 자료 요구에 응하는 형태로 소위 '통합소득' 자료를 공개한 바 있다. 그것은 종합소득 신고자 이외에 근로소득 연말정산 대상자(납세자에 한정)의 소득을 개인별로 통합하여 백분위별 통계로 제시한 것이다². 여기서는 이들 자료를 이용하지만, 통합소득 중에서도 대부분의 금융소득(종합과세 대상이 아닌 금융소득 4,000만원 미만)은 제외되어 있으며, 그 자료에서 빠진 소득자 또한 적지 않다. 이들 중에서 복수의 소득원이 있는 경우는 개인소득으로 통합하기 위한 추정이 필요하다³.

일본(Moriguchi and Saez 2008) 등으로 확대되었으며, 현재 28개국의 top income shares에 관한 성과가 World Top Income Shares Database(<http://topincomes.parisschoolofeconomics.eu/>)에 실려 있다. 한국의 연구성과(Kim and Kim 2014)도 최근 여기에 등재되었다.

² 이 자료는 흥종학 의원실 또는 경제개혁연대에서 얻을 수 있다. 김상조(2014)는 이 자료를 이용하여 소득분위별 소득점유 비중을 추정하였고 여러 가지 흥미로운 사실을 도출하고 있다. 다만 그의 방법에는 몇 가지 한계가 있다. 먼저 그의 분석 대상은 국세청의 통합소득 자료에 과세 기준에 미달한 근로소득자를 추가하였으며, 이를 모집단으로 하여 분위별 소득비중을 구하였다. 그렇지만 그 경우 거기에서 빠진 소득이 많으며, 그로 인해 모집단의 의미도 불분명해졌다. 예컨대 근로소득자의 모집단에는 연말정산 대상자만 포함되고 일용근로자는 제외되었는데, 양자는 과세제도상의 구분일 뿐 근로소득임에는 변함이 없기 때문이다. 그리고 분위별 소득비중을 구하기 위해 그가 이용한 보간법은 널리 활용되기 어렵고 엄밀하지도 않다. 그의 방법은 과세 기준 미달자의 소득이 납세자의 소득보다 낮을 것으로 암묵적으로 상정하고 있지만, 이는 과도한 단순화다. 과세 미달자 중에는 소득이 높더라도 소득공제가 많아서 과표가 0이 된 자들이 적지 않게 포함되어 있기 때문이다.

³ 김우철(2014)은 김낙년(2012a)과 김낙년·김종일(2013)이 소득세 자료에 의거하여 가계조사의 문제점

이하의 논의는 다음과 같이 구성되어 있다. Ⅱ절에서는 『연보』의 소득세 자료를 소개하고, 소득구간의 분할이나 조정이 필요한 경우 적용할 보간법(interpolation)을 설명한다. Ⅲ절에서는 2010년을 대상으로 국세청이 제시한 통합소득과, 거기에서 빠진 근로소득이나 사업소득의 각 항목을 하나하나 추가하고 누락된 금융소득을 추정하여 소득구간별로 정렬한다. Ⅳ절에서는 이들 중에서 개인이 복수의 소득원을 갖는 경우 이를 통합한 개인소득의 분포를 제시한다. 그리고 이것과 가계조사의 개인소득 분포와 비교하여 구간별로 어떤 차이가 나는지를 보인다. 나아가 소득구간별 개인소득 분포를 소득분위(10분위뿐만 아니라 상위 10%의 경우 상위 1%, 0.1%, 0.01%로 세분)별로 재편하여 각 분위가 차지하는 소득비중 등을 구하고, 기존 연구결과와 비교하기로 한다. 마지막으로 V절은 이상의 논의를 마무리하고 남겨진 과제를 언급한다.

Ⅱ. 자료와 방법

우리나라 소득세법은 모든 소득자가 자신의 소득을 합쳐 종합소득을 신고하도록 규정하고 있지만, 신고를 하지 않아도 되는 예외를 널리 인정하고 있다. 2010년에 종합소득 신고자는 379만명에 불과하며, 나머지 대부분은 원천징수로 과세가 종결(나중에 연말 정산되는 경우도 포함)된다. 따라서 소득의 전모를 파악하기 위해서는 『연보』에서 종합소득 신고자료 이외에 원천징수의 각 소득항목을 포괄할 필요가 있다. <표 1>은 『연보』에서 파악된 소득항목을 소득종류별로 나누어 정리하고, 각 항목별로 파악된 인원수와 소득금액을 제시하였다.

여기서 소득의 범위는 개인소득으로 하되, 그 해에 창출된 본원소득(근로소득, 사업소득, 재산소득)으로 한정하기로 한다. 그리고 자산가격의 변동에 의해 창출되는 자본이득(capital gain)도 제외하기로 한다. 따라서 소득세법에 규정되어 있는 퇴직소득, 연금소득, 양도소득 이외에 이전소득도 고찰대상에서 제외되고, 과세 이전(before tax) 소득이 대상이 된다.

먼저 종합소득은 사업 및 부동산임대 소득자 이외에 일정 규모 이상의 복수의 소득원이 있는 경우 신고대상이 된다. 금융소득은 4000만원을 넘는 경우 종합과세 신고대상으로 포함된다. 종합소득은 개인별로 복수의 소득원을 통합하여 보여주기 때문에 개인별 소득분포를 보는데 적합하지만, 대상자의 커버리지가 한정되어 있다.

근로소득의 경우 소득세를 원천 징수한 다음, 개인별로 소득공제 등을 반영하여 소득세를 확정하기 위한 연말정산을 하게 된다. 『연보』는 연말정산의 결과 근로소득을 납부한 자들과 과세대상 소득이 면세점 미만인 자(과세미달자)로 나누어서 통계를 제시하고 있다. 이들 통계는 각각 급여총계와 함께 그 내역으로서 비과세소득, 과세대상소득, 각종 소득공제와 산출세액 등을 제시하고 있다. 그런데 소득세법은 근로소득 중에서 고용계약 기간이 3개월(건설업은 1년) 미만인 자들을 일용근로자로 분류하고, 이들에 대해서는 낮은 세율로 소득세를 원천 징수하는 것(연말정산은 생략)으로 과세가 종결되도록 하고 있다. 『연보』에는 일용근로소득을 연말정산 대상 근로소득과 구분하여 별도의 통계표로 제시하고 있다.

을 구체적으로 밝힌 점을 평가하면서도 소득세 자료로는 개인소득의 전모를 드러내기는 어렵다고 비판하고 있다. 그렇지만 최근의 『연보』에는 면세자를 포함하여 근로 및 사업 소득자를 대부분 파악할 수 있으며, 일부 소득 항목이 누락되었거나 개인별로 통합된 소득을 알기 어려운 경우가 있지만, 후술하듯이 다른 자료로 보완하거나 합리적 가정으로 어느 정도 실태에 접근할 수 있다.

<표 1> 소득세 자료와 국민계정에서 파악된 개인소득(2010년) (단위: 명, 10억원)

소득세 자료				국민계정의 개인소득				
소득의 구분				표 번호	인원수	소득금액	소득의 구분	소득금액
종합소득	A=J	사업 및 부동산임대 소득		3-2-3	3,302,071	63,835		
	B=Q	종합과세 대상 금융소득		3-1-5	48,907	9,853	합계(괄호 항목 제외)	638,798
	C	근로소득		3-2-3	773,758	25,587		
	D	기타소득		3-2-3	159,657	1,005		
근로소득	E	근로소득 납세자		4-2-4	9,244,424	356,351	임금 및 급여	466,460
	F	근로소득 과세미달자		4-2-4	5,932,358	44,735	(고용주의 사회부담금)	68,842
	G	비과세소득		4-2-4	2,593,972	4,798		
	H	일용근로소득		4-4-8	7,825,444	46,255		
	I	(조정 근로소득)						
사업 및 부동산임대 소득	J=A	사업 및 부동산임대 소득	종합소득 신고소득	3-2-3	3,302,071	63,835	사업 및 부동산소득	101,385
	K		(미 신고소득)				(귀속 임료)	26,571
	L		(조정 사업소득)				재배업 소득	17,580
	M	연말정산 사업소득		4-3-1	587,774	8,755	임료	2,420
	N	원천징수 사업소득		4-4-5	3,085,626	45,337		
	O	기타소득		4-4-7	2,170,211	6,976		
	P	(재배업 소득)						
금융소득	Q=B	종합과세 대상 금융소득	배당	3-2-3	41,503	7,201	배당	14,403
	R=B		이자	3-2-3	44,177	2,628	이자	36,550
	S	(비 대상 금융소득)					(금융중개서비스)	12,183

주: 1) 왼쪽의 ()는 소득세 자료에서 접근할 수 없는 것이고, 오른쪽의 ()는 개인소득에서 제외되는 항목을 나타낸다.

- 2) 퇴직소득, 연금소득, 양도소득은 제외한다.
- 3) 조정 근로(사업)소득이란 국민계정의 해당 소득과의 차이를 말한다.
- 4) 표 번호란 『국세통계연보』에 실린 표의 번호를 말한다.
- 5) 소득세 자료의 인원수와 소득금액에는 다른 항목과 중복되는 경우가 포함되어 있다.

자료: 국세청 『국세통계연보』 2010년; 한국은행, ECOS.

사업소득(부동산임대소득을 포함)의 경우 모두 종합소득 신고대상이 되며, 전술한 종합소득세 통계에서 파악된다. 다만 영세한 규모의 사업자 중에는 종합소득을 신고하지 않는 경우가 적지 않게 존재한다. 과세당국의 입장에서도 신고하더라도 면세점 미만일 가능성이 큰 경우 신고를 독려할 이유도 없다. 따라서 미 신고 사업소득자는 소득세 통계에서 빠지게 된다. 그 외에 소득세법은 주로 사업체가 없는 자영자의 사업소득을 대상으로 원천징수로 과세가 종결될 수 있는 길을 열어두고 있다. 연말정산 대상 사업소득(보험모집인이나 방문판매원)이나 원천징수 대상 사업소득(부가가치세가 면제되는 의료보건 종사자와 개인용역이나 봉사료)이 그것이다. 다만 의료보건 종사자의 경우에는 대부분 종합소득 신고대상이며 종합소득 통계와 중복되어 있다. 기타소득도 사업소득의 성격이 강한 것으로 보아 여기에 포함하였다. 그리고 재배업(즉 농업소득)은 소득세법의 면세 대상이므로 소득세 자료에서는 빠지지만, 사업소득으로 포함된다.

금융소득의 경우 이자와 배당을 합쳐 4,000만원을 넘는 경우 종합소득 신고대상이 되고, 종합소득에 포함되어 제시된다. 그 외의 비 대상 금융소득은 소득세가 원천 징수되지만, 다른 소득항목과는 달리 『연보』에 총액만 제시될 뿐 소득구간별 통계가 제시되어 있지 않다.

<표 1>은 이들 소득세 자료에서 파악된 소득이 전체 소득의 어느 정도를 커버하는지 가능하기 위해서 국민계정의 개인소득도 함께 제시하였다. 국민계정은 소득자의 인원수나 소득구간별 분포를 알 수 없지만 전체소득에 관해서는 신뢰할 수 있는 정보를 주기 때문이다. 표에는 소득세 자료와 대응하도록 개인소득을 근로소득, 사업소득, 금융소득으로 나누어 제시하였다. 이 중에서 고용주가 부담하는 사회보험료 부담금은 개인소득에 포함되지 않으며, 귀속임료나 금융중개서비스(이상 괄호한 항목)도 통상의 개인소득 범주에 들지 않으므로 제외한다⁴. 그 경우 국민계정에서 파악된 전체 개인소득은 약 639조 원이 된다. 그 중에서 근로소득을 보면, 국민계정에서는 466조 원이지만, 소득세 자료(근로소득 납세자와 과세미달자 및 일용근로소득의 합계)에서는 447조 원으로 나와 약간의 차이가 있다. 이 차이는 소득세 과세대상에서 빠지는 근로소득(예컨대 가사서비스 등)으로 볼 수 있는데, <표 1>에서 이 차이를 조정 근로소득의 항목으로 처리한다면 두 자료의 소득금액을 일치하도록 할 수 있다.

소득세 자료의 사업 및 부동산임대 소득은 국민계정의 해당 소득항목(영업잉여와 준 법인기업 소득인출)에서 귀속임료와 재배업 소득을 뺀 101조 원이 대응한다. 소득세 자료의 원천징수 사업소득 중에는 종합소득 신고대상으로 중복된 부분이 있어 이를 빼줄 필요가 있고, 미 신고 사업소득을 추정하여 보완할 필요가 있다. 이러한 조정을 거쳐 두 자료의 사업 및 부동산임대 소득의 합계를 비교하면 후술하듯이 약간의 차이가 나타난다. 이를 전술한 근로소득과 마찬가지로 조정 사업소득 항목으로 처리할 수 있다.

이상에서 본 바와 같이 근로소득과 사업소득의 경우 소득세 자료로 파악되는 소득은 약간의 보완을 한다면 국민계정에서 파악된 전체 소득에 근접한 것을 알 수 있다. 이에 대해 금융소득의 경우 소득세 자료는 이자와 배당의 합계가 4,000만원이 넘는 소득자를 대상으로 파악한 것이기 때문에 국민계정의 전체 금융소득과 차이가 크다. 종합과세 비대상 금융소득의 경우는 후술하듯이 가계금융조사와 같은 다른 자료로 접근할 수 있다. 사업소득의 경우에도 미 신고 사업소득과 재배업 소득은 소득세 자료에서 정보를 얻을 수 없기 때문에 부가가치세 자료나 농림어업조사와 같은 다른 자료로 접근할 필요가 있다. 이에 관해서는 다음 절에서 다룬다.

한편, <표 1>의 각 소득항목은 소득구간별 인원수와 소득금액을 제시하고 있는데, 이를 소득구간별로 더하거나 뺄 필요가 있다. 이 때 소득구간이 일정하지 않거나 소득 기준이 다를 경우에는 조정이 필요하다. <표 2>는 소득세 자료의 한 예시로서 『연보』의 일용근로소득 통계를 보였다. 이를 후술하는 <표 3>의 일용근로소득과 같은 형태로 소득구간을 재편할 경우 두 가지 문제를 해결할 필요가 있다. 하나는 1-3(백만원)과 3-5(백만원) 구간을 합쳐서 1-5(백만원) 구간을 만드는 것은 문제가 없지만, 8-12(백만원) 구간(음영 표시)의 경우에는 10(백만원)을 경계로 하여 인원수와 소득금액을 분할할 필요가 있다는 것이다. 또 하나는 30(백만원) 이상인 최상위 구간의 경우 그 평균소득(음영 표시)이 거의 40(백만원)에 달한 데서 알 수 있듯이 30-40(백만원) 구간을 넘는 소

⁴ 국민계정의 개인소득은 한국은행의 제도부문별 소득계정에서 가계 및 비영리단체(원천)의 피용자보수, 영업잉여, 재산소득을 포함한다. 그 중에서 고용주의 사회부담금과 보험과 연금의 투자소득지급은 개인의 실제 소득이 아니므로 제외한다. 그리고 영업잉여 속에 포함되어 있는 귀속임료와 재산소득 중 이자에 포함되어 있는 금융중개서비스(FISIM)도 마찬가지로 이유로 제외해야 하지만, 한국은행은 현재 그 수치를 공표하지 않고 있다. <표 1>의 귀속임료와 금융중개서비스는 김낙년(2012a: 79)의 방법에 따라 추정한 것이다.

득자도 많을 것으로 추론된다⁵. 이 경우 그들이 그보다 상위의 소득구간에 걸쳐 어떻게 분포하고 있는지를 추정할 필요가 있다.

소득구간을 필요에 따라 분할하고자 할 경우 보간법(interpolation)을 사용하는데, 두 가지 접근이 가능하다. 하나는 상위 소득자들의 분포가 Pareto 분포를 따른다고 가정하는 방법이다. 이 방법은 특히 상위 1%와 같은 소득집중도(top income shares)를 추정할 때 널리 이용되고 있다(Piketty 2003; Piketty and Saez 2003; Moriguchi and Saez 2008; 김낙년 2012a; 2012b). 또 하나는 Atkinson(2005: 333-334)에 의해 시도된 것으로서 소득구간별로 인원수가 분포될 수 있는 상한과 하한의 범위(upper and lower bounds)을 가능한 한 좁힌 다음, 그 사이에서 가장 합리적으로 생각되는 분포를 mean split histogram이라는 방법으로 추정하는 것이다⁶. 이 방법은 Pareto 분포와 같은 특정 형태의 분포를 가정하지 않는다는 장점이 있다. 여기서는 mean split histogram의 보간법을 이용하기로 하되, 그 적용이 어려운 최상위 소득구간의 경우 Pareto 보간법으로 접근하기로 한다.

<표 2> 소득세 자료 예시: 소득구간별 일용근로소득(2010년)

하한 l	상한 u	인원수 N	소득금액 Y	평균소득 m=Y/N	누적 인원수 H=∑N	누적 소득금액 G=∑Y	m의 위치 (m-l)/(u-l)
(백만 원)		(천 명)	(10억원)	(백만 원)	(천 명)	(10억원)	
	1	2,131	824	0.4	7,825	46,255	
1	3	1,862	3,482	1.9	5,695	45,431	0.435
3	5	1,058	4,119	3.9	3,832	41,949	0.447
5	8	911	5,778	6.3	2,775	37,830	0.447
8	12	677	6,619	9.8	1,863	32,052	0.443
12	17	476	6,784	14.2	1,186	25,432	0.448
17	20	187	3,435	18.4	709	18,649	0.470
20	24	180	3,926	21.9	523	15,214	0.467
24	30	173	4,625	26.7	343	11,289	0.449
30		170	6,664	39.2	170	6,664	

주: 누적 인원수(소득금액)이란 해당 소득구간을 포함해서 그 상위 구간을 합계한 것이다.

자료: 『연보』의 4-4-8 표.

<그림 1>을 이용하여 mean split histogram의 보간법을 설명해 두고자 한다⁷. 그림은 소득구간의 인원수 분포를 예시한 것인데, X축은 소득을, Y축에는 각 구간까지의 누적 인원수를 보여준다. <표 2>의 8-12(백만원) 소득구간을 예로 들면, 이 구간의 하한(y_i)과 상한(y_{i+1}), 그리고 평균(m_i)은 각각 8, 12, 9.8(백만원)이 된다(여기서 하첨자 i 는 소득구간을 나타낸다). Y축의 하한과 상한의 누적 인원수(H_i , H_{i+1})는 각각 1,863(천명)과 1,186(천명)이 된다. 그런데 여기서 만약 소득이

⁵ <표 1>에 제시된 소득항목은 최고 소득구간이 대부분 5억원 이상이지만, 여기서 든 일용근로소득과 과세미달자의 경우 각각 3,000만원과 1억원 이상으로 낮게 설정되어 있다.

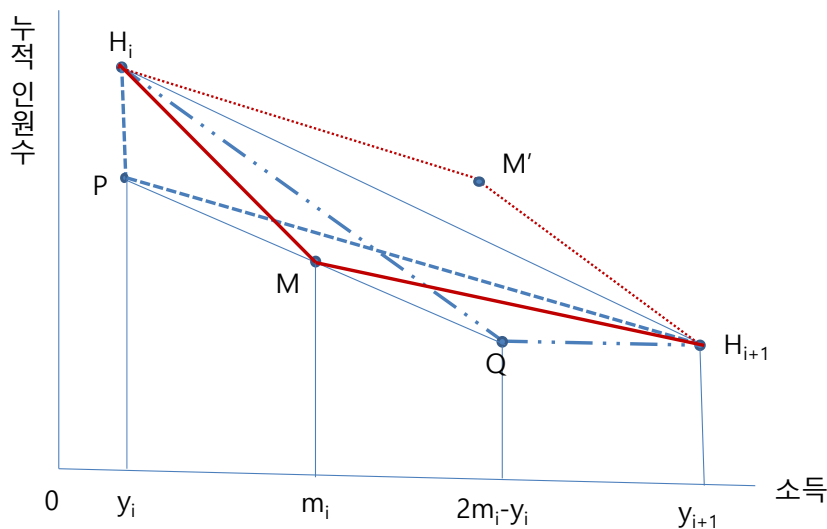
⁶ 이 방법은 Atkinson에 의해 영국의 소득집중도를 추정하는데 사용되었고, 호주, 네덜란드, 핀란드, 노르웨이 등 다른 나라로도 확대되고 있다(Atkinson, Piketty and Saez 2009: 19-25).

⁷ 이 방법은 김낙년(2014)이 계층별 조세 부담률을 추정할 때에도 이용된 바 있다.

높아질수록 인원수가 줄거나 적어도 늘어나지 않는다는 가정을 도입한다면 인원수가 분포될 수 있는 범위를 상당히 좁힐 수 있다. 그림에서 평균 소득(m_i)은 그 구간의 중간점보다 왼쪽에 위치한 것으로 그렸는데, 그것은 이 구간에서 소득이 상대적으로 낮은 경우가 높은 경우에 비해 인원수가 많음을 뜻한다⁸.

이 경우 인원수 분포의 범위는 다음과 같이 주어진다. 소득구간에서 평균소득(m_i)을 유지하면서 상대적으로 저소득자의 분포가 가장 많은 경우를 인원수 분포의 하한(lower bound)이라 하고, 거꾸로 고소득자의 분포가 가장 많은 경우를 상한(upper bound)라고 하자. 먼저 인원수 분포의 하한은 <그림 1>에서 직선 H_iQ 와 QH_{i+1} 로 나타낼 수 있다. 거기서 Q 에 대응하는 소득($2m_i - y_i$)은 평균(m_i)로부터 왼쪽으로 y_i 까지의 거리와 동일한 지점을 m_i 의 오른쪽에 잡은 것이다. 이 경우 이 구간의 인원수는 모두 소득구간 $[y_i, 2m_i - y_i]$ 에 균등하게 분포(따라서 이들의 소득 평균이 곧 m_i 가 된다)하며, 이보다 높은 소득구간 $[2m_i - y_i, y_{i+1}]$ 에서는 인원수가 0인 경우가 된다. 이에 대해 인원수 분포의 상한(upper bound)은 직선 H_iP 와 PH_{i+1} 로 주어진다. 직선 PQ 는 직선 H_iH_{i+1} 와 평행으로 그렸는데, 그 경우 삼각형 H_iQH_{i+1} 와 H_iPH_{i+1} 와 같이 사다리꼴 $H_iH_{i+1}QP$ 의 밑변(H_iH_{i+1})을 공유하는 삼각형은 모두 동일한 면적을 갖는다. 따라서 인원수 분포의 상한인 직선 H_iP 와 PH_{i+1} 도 그 평균소득이 m_i 로 유지함을 알 수 있다. 이 경우는 이 구간에서 소득이 가장 낮은 수준(y_i)에 가능하면 많은 인원수(그림에서 H_iP 만큼)를 집중시킨 다음 나머지를 소득구간 $[y_i, y_{i+1}]$ 에 걸쳐 균등하게 분포하게 만든 것이다.

<그림 1> 소득구간의 보간 방법(mean split histogram)



주: Atkinson(2005: 333)의 그림을 가져오되 필요한 부분을 추가하였다.

그런데 실제 이 구간의 인원수 분포는 위의 상하한의 범위 내에 있을 것으로 터이지만, 이를 어떻게 특정하는 것이 좋을까? 상한과 하한의 평균을 구하는 것도 한가지 방법이지만, 그것보다는 M 을 중심으로 직선 H_iM 와 MH_{i+1} 로 분포하는 것으로 보는 것이 합리적이다. M 은 그

⁸ 평균소득(m_i)이 소득구간의 중앙에 올 경우 <표 2>에서 m 의 위치가 0.5가 될 터이지만, 여기서는 모든 구간이 그보다 작아 왼쪽에 있음을 알 수 있다.

구간의 평균(m_i)에 대응하며, 그 의미에서 이를 mean split histogram의 방법이라고 한다. 그림에서 직선 H_iM 의 기울기는 직선 MH_{i+1} 에 비해 상대적으로 가파른 데에서 알 수 있듯이 그 구간의 평균(m_i) 이하에서는 인원수가 더 많고 그 이상에서는 더 적게 분포하는 양상을 보이고 있다.

그런데 소득구간 중에는 드물게 평균소득(m_i)이 구간의 중간 점의 오른쪽에 오는 경우가 있을 수 있다. 그 경우에는 소득이 높아질수록 인원수가 줄거나 적어도 늘어나지 않는다는 가정이 성립되기 어렵다. 이 때에는 <그림 1>에서 직선 H_iM' 와 $M'H_{i+1}$ 를 이용하여 보간하였다. M' 은 M 에 비해 평균소득이 구간의 중간점의 오른쪽에 있으며, M' 을 도출하는 방법은 M 의 경우와 동일하지만, 그래프가 복잡해지지 않도록 보조선의 표시를 생략하였다. 양자의 차이를 보면, M' 의 경우 직선 H_iM' 보다 직선 $M'H_{i+1}$ 의 기울기가 더 가파른 데에서 알 수 있듯이 소득이 상대적으로 낮은 구간보다 높은 구간의 인원수가 오히려 더 많이 분포하고 있다.

그러면 이 보간법과 <표 2>의 데이터를 이용하여 8-12(백만원) 소득구간의 인원수와 소득금액을 10(백만원)을 경계로 하여 분할하는 방법을 예시해 보기로 하자. <표 2>의 정보를 이용하면 그림의 직선 H_iM 와 직선 MH_{i+1} 의 기울기와 절편(즉 두 직선의 함수)을 구할 있다. 여기에 소득 X 값에 10(백만원)을 대입하면, 거기에 대응하는 누적 인원수(1,455천명)가 산출된다⁹. 여기서 차상위 구간의 누적 인원수(1,186천명)를 빼서 구한 269천명이 이 소득구간의 인원수(677천명) 중에서 10(백만원)을 넘는 인원수가 된다. 그리고 그들의 소득은 이 소득구간의 상한(12백만원)과 경계소득(10백만원)의 평균값에 그 인원수를 곱하여 구해진다. 이 소득을 이 구간의 소득금액에서 빼면 이 구간에서 소득이 10(백만원) 미만인 자들의 소득금액이 계산된다.

이상은 소득구간의 인원수나 소득금액을 새로운 경계소득을 경계로 분할하는 방법이지만, 거꾸로 인원수를 가지고 경계소득을 구하는 데에도 활용된다. 예컨대 상위 10%의 경계소득을 구하고자 할 경우 상위 10%에 해당하는 누적 인원수와 그것을 포함하고 있는 소득구간을 정한 다음, 그 구간의 직선 H_iM 와 직선 MH_{i+1} 의 기울기와 절편에 상위 10%의 누적 인원수를 Y 값에 넣으면 그에 대응하는 경계소득을 구할 수 있다. 그리고 이를 경계로 하여 이 구간의 소득금액을 분할하는 것은 전술한 방법과 마찬가지로이다.

이러한 mean split histogram의 방법은 소득구간의 하한과 상한의 정보가 있을 경우 적용할 수 있다. 그렇지만 최상위 소득구간을 추가로 분할하고자 할 경우 하한의 정보밖에 없기 때문에 이 방법을 적용하기 어렵다. 그 경우에는 Pareto 보간법으로 접근할 수 있다. Pareto 분포의 경우 아래 식 (1)의 특성을 갖기 때문에 그것을 이용하면 Pareto 계수 a 를 구할 수 있다¹⁰. 거기서 $H(y)$ 는 소득 y 를 경계로 하여 그것을 넘는 자들의 누적 인원수를, $G(y)$ 는 그들의 누적 소득금액을 나타낸다. 따라서 $G(y)/H(y)$ 는 그들의 평균소득을 뜻하는데, 그것을 경계소득 y 로 나누어준 것은 일정한 값을 가진다는 특성이다. <표 2>의 최상위 소득구간에서 y , $H(y)$, $G(y)$ 는 각각 30(백만원), 170(천명), 6,664(10억원)이 되며, 이로부터 a 는 4.27로 산출된다. 그리고 아래의 식 (2)는 Pareto 분포로

⁹ 이 때 구하고자 하는 경계소득이 m 보다 아래 또는 위에 위치하느냐에 따라 직선 H_iM 또는 직선 MH_{i+1} 를 적용하면 된다. 이 경우는 경계소득(10백만원)이 m (9.8백만원)보다 위에 있으므로 후자가 된다.

¹⁰ Pareto 분포 함수로부터 식 (1), (2), (3)의 도출은 지면의 제약으로 생략하지만, 김낙년(2012a: 9-11, 34-35)를 참조하기 바란다.

부터 도출된 것인데, 두 경계소득(즉 y_i 과 y_{i+1})에 대응하는 누적 인원수의 비율($H(y_{i+1})/H(y_i)$)은 두 경계소득의 비율(y_i/y_{i+1})의 a 제곱과 같다는 관계가 성립한다. <표 2>에서 만약 새로운 경계소득을 40(백만원)으로 놓고 그 이상의 누적 인원수를 구하고자 할 경우 $y_i, y_{i+1}, H(y_i)$ 는 각각 30(백만원), 40(백만원), 170(천명)이 되므로 이로부터 $H(y_{i+1})$ (즉 40백만원 이상의 누적 인원수)를 구할 수 있다. 50(백만원) 이상의 누적 인원수를 구하고자 할 경우 y_{i+1} 를 50(백만원)으로 바꾸면 된다. 소득금액의 경우는 식 (3)을 이용하는데, 누적 인원수 대신에 누적 소득금액($G(y_{i+1})/G(y_i)$)을 넣고, a 제곱 대신에 $(a-1)$ 제곱으로 바꾸면 된다.

$$G(y)/H(y)/y = a/(a-1) \dots\dots\dots (1)$$

$$H(y_{i+1})/H(y_i)=(y_i/y_{i+1})^a \dots\dots\dots (2)$$

$$G(y_{i+1})/G(y_i)=(y_i/y_{i+1})^{a-1} \dots\dots\dots (3)$$

이상의 방법으로 일용근로소득의 인원수와 소득금액을 새로운 소득구간으로 재편한 결과가 <표 3>의 일용근로소득의 항목으로 제시되어 있다. 이 보간 방법은 다른 소득항목의 구간별 통계를 재편할 때에도 이용되었다.

Ⅲ. 소득구간별 분포

이 장에서는 소득세 자료(그것이 없을 경우는 다른 자료)를 이용하여 각종 소득의 구간별 분포를 정리한다. 이 때 각 자료의 소득 기준을 일치하도록 한다. 예컨대 소득금액에서 비과세소득이 제외되어 있는 경우 이를 추가하여 소득총액 기준으로 통일한다. 소득구간도 예컨대 통합소득이나 근로소득 연말정산 자료의 경우 과세대상소득(비과세소득이 제외된) 기준으로 되어 있지만, 이를 소득총액 기준으로 전환할 필요가 있다. 각종 소득의 소득금액과 소득구간을 모두 소득총액 기준으로 통일한 다음, 각 구간별 통계를 서로 더하거나 뺄 수 있도록 소득구간을 동일하게 설정할 필요가 있다. 여기서는 <표 3>에 제시되어 있는 바와 같이 16개 구간으로 통일하기로 한다. 이하에서는 먼저 국세청의 통합소득 자료를 이 기준에 맞추어 재편한 다음, 거기에서 빠진 소득항목을 하나하나 추가해 가기로 한다.

1. 통합소득

국세청의 통합소득은 『연보』에 제시되어 있는 종합소득 신고자료와 근로소득 연말정산 자료(납세자에 한정)를 통합한 것이다. 『연보』의 두 자료에서는 종합소득을 신고한 근로소득자(약 77만명)가 중복으로 포함되어 있어 이를 배제할 필요가 있었는데, 통합소득 자료에서는 이 문제가 해결되었다. 그런데 거기에서 종합소득 신고 대상이 아닌 근로소득자의 경우 그들의 근로소득만이 통합되었고, 금융소득과 같은 다른 소득은 제외되어 있다는 점에 유의할 필요가 있다. 이 문제는 금융소득 항목에서 다룬다. 그리고 통합소득 중에서 근로소득의 경우는 과세대상 근로소득(급여총계에서 비과세소득 제외된) 기준으로, 그 외의 소득의 경우는 소득금액(수입금액에서 필요경비를 제외한) 기준으로 되어 있다. 따라서 근로소득에는 비과세소득을 추가하여 소득총액 기준으로 재편

할 필요가 있다.

그런데 이 통합소득은 <표 2>에서 예시한 바와 같이 소득구간별 통계로 제시되지 않고, 전체 인원수(12,448,203명)¹¹를 소득 순으로 정렬한 후 상위 1%에서 100%까지 백분위(즉 분위당 124,482명)로 나누어 각 분위별로 통합소득, 과세표준, 결정세액 정보를 제공하고 있다¹². 이로부터 각 분위에 속하는 인원수와 그들의 평균소득을 알 수 있지만, 그들이 어떤 소득구간(즉 소득의 하한과 상한)에 속하는지를 직접 알 수 없다. 다른 소득세 통계가 모두 소득구간별 정보를 주고 있기 때문에 이들과 맞추기 위해서는 통합소득의 각 구간별 하한과 상한을 추정할 필요가 있다.

여기서도 전술한 mean split histogram의 방법을 응용할 수 있다. 소득구간의 하한과 상한을 각각 Y_i 와 Y_{i+1} 라고 하고, 그 평균소득을 m , 그리고 하한 Y_i (또는 상한 Y_{i+1})에 대응하는 누적 인원수를 H_i (또는 H_{i+1})라고 하자. 그리고 m 에 대응하는 누적 인원수를 H_m 이라 하자. <그림 1>에서 볼 수 있듯이 소득구간의 평균값 m 을 중심으로 왼쪽 구간의 소득합계(즉 왼쪽 구간의 인원수*해당 구간의 평균소득)는 오른쪽 구간의 소득합계와 일치하며, 그것은 동시에 해당 구간의 전체소득의 1/2이 될 것이다. 이러한 조건을 나타내면 식 (4)와 같다. 식 (4)의 한쪽 방정식에서 H_m 을 구해 다른 쪽에 대입한 후 정리하면 식 (5)와 같은 관계가 도출된다. 즉 소득구간의 상한(또는 하한)은 평균을 제공한 값을 하한(상한)으로 나누어 구할 수 있다는 것이다.

$$(H_i - H_m) * (Y_i + m) / 2 = (H_m - H_{i+1}) * (m + Y_{i+1}) / 2 = (H_i - H_{i+1}) * m / 2 \quad \text{----- (식 4)}$$

$$Y_{i+1} = m^2 / Y_i \quad \text{또는} \quad Y_i = m^2 / Y_{i+1} \quad \text{----- (식 5)}$$

그런데 어느 구간의 상한은 바로 인접한 차상위 구간의 하한에 해당하므로 만약 어느 구간이든 상한(또는 하한)의 값을 임의로 가정하여 부여하면, 그 구간의 하한(상한)을 값을 구할 수 있고, 이를 차상위(또는 차하위) 구간으로 확대하면 위의 산식에 의거하여 모든 구간의 하한과 상한이 도출된다. 여기서는 상위 30%에 해당 하는 소득구간의 상한 값을 상위 30% 소득구간의 평균값과 상위 29% 소득구간의 평균값의 중간에 위치한다는 가정(즉 $Y_{i+1(30\%)} = m_{30\%} + (m_{29\%} - m_{30\%}) / 2$)을 하였다. 통합소득에서 상위 29%와 30%의 평균소득은 40,467(천원)과 39,536(천원)이므로 상위 30%의 상한은 양자의 중간인 40,001(천원)에 위치한 것으로 가정한 것이다. 통합소득이 100분위로 촘촘하게 세분되어 있어 이러한 가정이 실제와 크게 벗어나지는 않은 것으로 생각된다.

이렇게 산출한 소득구간을 검토하기 위해 <그림 2>를 제시하였다(가정 1). 거기서 제시한 지표는 <표 1>에서 소득구간 내의 'm의 위치'에 대응하는 것으로서 각 구간별로 $(m - Y_i) / (Y_{i+1} - Y_i)$ 로 구한 것이다. 이 값이 0.5가 되면 m이 해당 구간의 중앙에 위치한다는 뜻이고 그보다 작으면 m이 중앙값의 왼쪽에 위치한다는 것을 뜻한다. 100개 구간에 걸친 이 지표의 추이를 보이면, 대체로 0.5보다 작지만 0.5에 근접하며, 최상위(또는 최하위) 구간으로 가면 그 값이 더욱 작아진다. 그것은 각 구간의 인원수가 모두 동일하지만 구간이 촘촘하여 상한과 하한의 소득 차이가 크지 않은

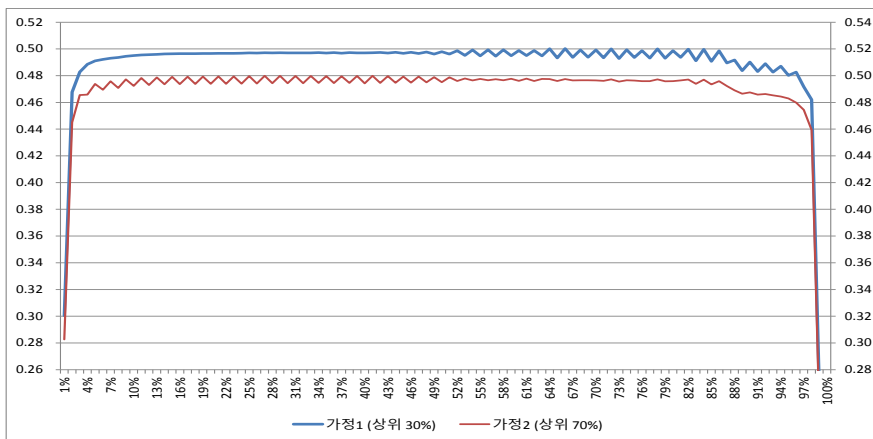
¹¹ 통합소득의 전체 인원수는 종합소득세 신고자와 근로소득 연말정산자 중 납세자를 합한 것이므로 전체 소득자의 일부에 불과하다. 따라서 이를 기준으로 산출한 상위 x%가 국제 비교될 수 있는 것은 아니다.

¹² 상위 10만명에 대해서는 다시 1만명 단위로 세분하고, 상위 1만명은 다시 상위 1000명과 9000명, 그리고 상위 1000명은 다시 상위 100명과 900명으로 세분한 통계를 제시하고 있다.

경우 평균값을 경계로 하는 양쪽의 소득 차이도 크지 않기 때문이다. 이에 대해 최상위(또는 최하위)로 갈수록 상한과 하한의 소득 차이가 크게 벌어지고, 구간 내에서 평균보다 낮은(또는 높은) 소득의 인원수가 훨씬 많아지기 때문이다.

그런데 눈에 띄는 것은 임의로 선택한 상위 30% 구간으로부터 멀리 떨어질 경우 지표가 톱니바퀴 모양으로 물결치는 양상을 보인다. 이것은 이 방식의 소득구간 추정을 모든 구간으로 확대할 경우 일부 구간에서 오차가 점차 누적되기 시작함을 보여준다¹³. <그림 2>에는 상위 70%의 상한을 앞의 상위 30%의 경우와 같은 방법으로 가정한 결과를 보여준다(가정 2). 양자를 같은 눈금으로 제시하면 거의 포개지는 것을 알 수 있다. 어느 구간을 선택하여 가정하더라도 지표의 전체적인 양상은 거의 변함이 없음을 확인하였다. 다만 그림에서 보이듯이 상위 70% 구간을 선택한 경우에도 그로부터 멀어지면 지표가 톱니바퀴 양상을 보이고 있다. 여기서는 이러한 점을 감안하여 상위 50%에 해당하는 구간에서는 상위 30% 구간에 가정한 값(가정 1)을 주고, 상위 51% 이하에 해당하는 구간에서는 상위 70% 구간에 가정한 값(가정 2)을 주어 상한과 하한의 값을 구하기로 한다¹⁴.

<그림 2> 추정된 소득구간 내 평균소득의 위치



주: 1) 소득구간 내 평균소득의 위치는 $(m - Y_i) / (Y_{i+1} - Y_i)$ 으로 구했다.

2) 가정 1, 2는 각각 상위 30%와 70%는 해당 구간의 상한 값을 차상위 구간과의 중간 값으로 가정했을 경우를 보여준다. 전자는 왼쪽 눈금, 후자는 오른쪽 눈금.

3) 두 가정에 의한 결과를 같은 눈금으로 제시하면 거의 포개져서 양자의 차이를 식별하기 어렵기 때문에 눈금을 달리하여 제시하였다.

¹³ 만약 어느 구간의 상한이 실제보다 과대해지기 시작하면 그 차상위 구간의 하한도 그러할 터인데, 그 경우 위의 (식 5)에서 알 수 있듯이 해당 구간의 상한은 과소하게 된다. 그렇게 되면 다시 그 차상위 구간의 하한도 과소하게 되므로 해당 구간의 상한은 과대하게 된다. 이것이 반복되면서 점차 오차가 커지는 것이 톱니바퀴 양상으로 나타난 것이다.

¹⁴ 통합소득에서 상위 15%에서 상위 78% 사이의 구간은 인접한 두 구간의 거리 $((m_i - m_{i+1}) / m_i)$ 가 해당 소득의 3%의 범위를 벗어나지 않을 정도로 촘촘하다. 이들 구간에서는 두 구간 사이의 중간 $(= (m_i + m_{i+1}) / 2)$ 을 각 구간의 하한(또는 상한)으로 간주하는 단순한 방법도 본고의 방식과 차이가 거의 없다. 그렇지만 최상위(또는 최하위) 구간으로 갈수록 차이가 벌어진다.

자료: 국세청의 통합소득.

이상의 방법으로 100분위별로 나누어져 있는 통계에 구간별 상한과 하한 정보를 추가하였다. 그런 다음 이를 <표 3>에 예시한 바와 같이 16개 소득구간으로 통합하였다. 이 때에는 앞에서 예시한 mean split histogram에 의한 보간법을 이용하면 된다. 예컨대 소득구간 5,000-6,000만원에 속하는 인원수와 소득금액을 구할 경우 위에서 추정된 상위 21%의 소득구간(4,873-5,004만원)과 상위 14%의 소득구간(5,918-6,099만원)을 각각 5,000만원과 6,000만원을 경계로 <그림 1>에서 제시한 방법으로 인원수와 소득금액을 나누면 된다¹⁵.

그런데 이상의 소득금액과 소득구간은 모두 비과세소득이 제외된 과세대상 소득금액 기준으로 되어 있어 이를 소득총액 기준으로 바꾸기 위한 추가적인 절차가 필요하다. 통합소득에서 제외된 비과세소득은 3조 9,804억원(통합소득의 0.9%)에 해당하며 구간별로 차이가 있다¹⁶. 예컨대 2,000-3,000만원과 3,000-4,000만원 소득구간(각각의 소득금액은 53조 5,061억원과 53조 3,017억원)에는 각각 5,602억원과 5,783억원의 비과세소득이 추가될 필요가 있다. 이 때 구간별 소득금액에 비과세소득을 더하는 것은 간단하지만, 동시에 소득구간도 소득총액 기준으로 바꾸어 줄 필요가 있다. 소득구간별로 과세대상소득금액 대비 소득총액의 비율을 조정률(위의 두 소득구간의 경우 각각 1.011과 1.012이 된다)이라고 할 경우 이를 어떻게 적용하는 것이 좋을까?

만약 소득구간_i의 조정률(r_i)을 해당 소득구간의 하한(L_i)과 상한(H_i)에 적용하면 각 소득구간별 조정률이 다르기 때문에 예컨대 소득구간_i의 상한($=r_i * H_i$)과 다음 소득구간_{i+1}의 하한($=r_{i+1} * L_{i+1}$)이 일치하지 못하고 벌어지게 된다. 이를 피하기 위해 소득구간_i의 상한과 소득구간_{i+1}의 하한에는 두 구간의 평균 조정률($=(r_i + r_{i+1})/2$)을 적용하는 방법을 생각할 수 있다. 다만 각 소득구간의 평균소득(m_i)이 소득구간의 중앙에 온다면 이 방법에 문제가 없지만, 실제로는 그렇지 않은 경우가 보통이다. 여기서는 소득을 X축, 조정률을 Y축으로 하는 평면 위에서 소득구간_i의 평균소득(m_i)과 그 구간의 조정률(r_i)을 대응시키고, 인접한 두 소득구간의 조정률을 직선 보간하는 방법을 이용하였다. 그로부터 $Y = r_i + (r_{i+1} - r_i) / (m_{i+1} - m_i) * (X - m_i)$ 와 같은 관계식이 구해진다. 이 식에서 구하고자 하는 소득구간_i의 상한(H_i , 이것은 소득구간_{i+1}의 하한 L_{i+1} 과 동일)을 X에 대입하면, 그에 대응하는 조정률 Y를 구할 수 있다. 여기서는 이 방법을 채택하였으며, 앞으로도 소득구간의 조정이 필요한 경우 이 방법을 적용하였다.

그런데 이 방법으로 소득구간과 소득금액을 모두 소득총액 기준으로 바꾸고 나면, 소득구간이 예컨대 2,000-3,000만원과 3,000-4,000만원에서 각각 2,022-3,032만원과 3,032-4,046만원으로 바뀌어진다. 이를 다시 원래의 소득구간으로 되돌리기 위해서는 전술한 mean split histogram의 보간

¹⁵ 국세청은 통합소득 이외에 근로소득과 종합소득에 대해서도 100분위 통계를 제공하고 있는데, 이들 소득에 위의 방법을 적용할 경우 『연보』에 실려 있는 소득구간별 통계와 일치하는 결과를 얻을 수 있는지 검증해 볼 수 있다. 그에 따르면 대부분의 구간에서는 근접한 결과를 얻었지만, 5억원 이상 구간으로 가면 오차가 커짐을 확인하였다. 최상위 구간에서는 구간 내의 소득 차이가 커서 위의 m을 중심으로 2분하는 방법으로는 실태를 반영하는데 한계가 있음을 알 수 있다. 이 점을 감안하여 통합소득의 경우 5억원 이상 구간을 더 이상 세분하지 않기로 한다.

¹⁶ 통합소득에서 제외된 소득구간별 비과세소득은 근로소득 연말정산(납세자) 자료(『연보』의 4-2-4의 비과세소득)에서 얻을 수 있다.

법을 이용하면 된다. 그 결과가 <표 3>에 제시한 소득구간별 통합소득(인원수와 소득금액)이다.

2. 근로소득

국세청이 파악한 근로소득 연말정산 자료 중에서 납세자 소득(거기에 비과세소득을 추가한)은 앞에서 추정된 <표 3>의 통합소득에 포함되어 있다. 따라서 거기에서 제외된 근로소득으로서 과세미달자 소득과 일용근로소득을 추가할 필요가 있다¹⁷. 이 중에서 일용근로소득의 경우 II절의 자료 예시에서 이미 설명한 바 있고 그 결과가 <표 3>에 제시되어 있다. 근로소득 과세 미달자의 경우에도 『연보』(표4-2-4)에서 소득구간별 통계를 얻을 수 있지만, 소득구간이 과세대상근로소득 기준으로 되어 있기 때문에 이를 소득총액 기준으로 바꾸어줄 필요가 있다. 이를 위한 조정률의 계산과 적용은 앞의 통합소득 항목에서 언급한 방법에 따랐다. 그리고 소득구간을 <표 3>의 16개 구간에 맞추기 위해서는 mean split histogram의 방법으로 보간하였다. 다만 최상위 소득구간이 1억원 이상으로 설정되어 있는데, 그 평균이 2억 6,780만원인 데에서 알 수 있듯이 소득구간을 더욱 세분할 필요가 있다. 이와 같이 하한의 정보만 있는 소득구간의 경우 앞의 일용근로소득에서 예시한 바와 같이 Pareto 보간법을 이용하였다. 그 결과가 <표 3>의 과세미달자 항목으로 제시되었다¹⁸.

<표 3> 각종 소득의 구간별 인원수와 소득금액 (2010년)

(1) 인원수 (단위: 명)

소득구간 (백만원)	통합소득 A~E, G	제외된 근로소득		제외된 사업소득					제외된 금융소득 R	조정 항목		합계
		과세미달자 F, G	일용근로자 H	연말정산 M	원천징수 N	기타소득 (O-D)	미신고 K	재배업 P		사업소득 L	근로소득 I	
-1	229,507	649,807	2,130,787	130,248	1,044,378	1,455,449	488,802	937				6,129,915
1-5	755,436	2,193,206	2,919,971	132,653	1,006,181	332,084	864,489	205,311				8,409,331
5-10	1,101,906	1,605,420	1,319,833	76,217	254,341	85,377	362,866	530,217				5,336,176
10-20	2,988,771	1,057,161	932,031	90,604	321,021	87,180	102,885	232,662				5,812,316
20-30	2,168,288	301,557	352,778	60,362	33,285	25,443		85,964				3,027,677
30-40	1,531,463	94,737	120,196	35,447		9,098		48,018				1,838,960
40-50	1,114,117	25,443	30,604	24,663		5,323		19,269				1,219,419
50-60	817,989	3,902	10,402	14,654		3,422		12,921				863,291
60-70	574,893	877	4,261	10,110		1,925		9,311				601,377
70-80	349,255	114	1,989	5,408		1,107		7,766				365,639
80-90	252,282	48	1,024	2,738		662		3,901				260,654
90-100	148,421	28	568	1,559		648		3,901				155,124
100-200	335,215	40	1,001	2,899		1,731		15,747				356,633
200-300	44,761	9		176		348		1,192				46,487
300-500	21,404	5		32		233		183				21,858
500-	14,494	4		4		523		17				15,043
합계	12,448,203	5,932,358	7,825,444	587,774	2,659,206	2,010,554	1,819,043	1,177,318				34,459,900

¹⁷ 과세미달자 중에도 다른 소득이 있는 경우 종합소득 신고를 하여 통합소득에 포함될 수 있다(김상조 2014: 17, 그림 1의 "D"). 이 경우 통합소득과 과세미달자 소득을 더하면 그만큼 이중 계산이 될 수 있는데, 이 오차는 여기서는 후술하는 근로소득 조정 항목에서 포착된다.

¹⁸ 과세미달자의 소득이 1억원(심지어는 5억원)을 넘는 것이 의아하게 생각될지 모르겠다. 그렇지만 근로소득을 기부할 경우 기부금 소득공제에 한도가 없어 과세표준이 0이 되는(따라서 과세미달자가 되는) 경우가 생길 수 있다.

(2) 소득금액 (단위: 10억원)

-1	45	325	824	65	522	728	179	0	201	5	341	3,235
1-5	2,528	6,298	7,600	313	1,737	386	2,733	980	1,688	132	3,142	27,537
5-10	8,603	11,530	9,438	557	1,908	704	2,515	3,636	2,907	459	3,901	46,160
10-20	44,138	14,712	13,178	1,323	4,677	1,215	1,219	3,271	6,388	763	5,447	96,331
20-30	53,451	7,257	8,551	1,490	692	611		2,120	5,595	563	3,534	83,865
30-40	53,169	3,203	4,059	1,241		319		1,633	4,255	361	1,678	69,918
40-50	49,711	1,113	1,348	1,103		238		867	3,654	262	557	58,853
50-60	44,881	211	564	806		188		699	3,419	249	233	51,250
60-70	37,060	55	274	653		125		605	2,992	240	113	42,117
70-80	26,078	9	148	406		83		577	2,046	225	61	29,632
80-90	21,431	4	86	232		56		332	1,840	176	36	24,193
90-100	14,064	3	54	148		62		371	1,144	186	22	16,053
100-200	42,455	5	131	364		233		2,137	3,325	894	54	49,598
200-300	10,583	2		41		85		276	537	361		11,886
300-500	8,015	2		11		92		65	220	287		8,693
500-	18,004	6		2		845		11	147	462		19,478
합계	434,215	44,735	46,255	8,755	9,536	5,971	6,647	17,580	40,359	5,626	19,119	638,798

주: 1) 영어 알파벳은 <표 1>의 소득종류에 대응하는 것이다.

2) 통합소득과 과세미달자(근로소득)에는 비과세소득이 포함되어 있다.

3) 제외된 소득이란 국세청의 통합소득 자료에서 빠진 소득을 뜻한다.

4) 조정 항목이란 이상의 국세청 자료 등으로 파악된 소득과 국민계정의 개인소득과의 차이를 말한다.

5) 개인의 소득원이 복수일 경우 개인별로 소득을 통합하지 못한 경우가 있다.

자료: 본문 참조.

그런데 <표 1>에서 보듯이 국세청이 파악한 근로소득의 합계를 국민계정에서 파악된 임금 및 급여와 비교하면 약 19조의 차이가 나타난다. 이것은 국세청에서 포착하지 못한 근로소득으로 볼 수 있다. 사업체에서 임금을 지급하면 비용 처리를 하게 되지만, 파출부 등의 가사서비스와 같이 가계가 임금을 지급하는 경우는 그렇지 않아 소득세 자료에서 빠지는 경우가 생긴다. 여기서는 국민계정의 개인소득을 전체 소득으로 보고 있기 때문에 그 차이를 근로소득 조항항목으로 처리하기로 한다. 다만 그들의 소득이 어떻게 분포되었는지에 관해서는 정보가 없다. 여기서는 이들 소득이 일용근로소득과 같은 분포를 보일 것으로 가정하여 소득구간별로 배분하였다(표 3 참조).

3. 사업소득

사업소득 중에서 종합소득세를 신고한 경우 앞의 통합소득에 포함되어 있지만, 거기서 제외된 사업소득을 하나하나 추가하기로 한다. 『연보』에는 사업소득에 포함될 수 있는 것으로서 연말정산 사업소득, 원천징수 사업소득, 기타소득에 관한 정보를 제시하고 있다(표 1 참조). 이들을 추가할 경우 두 가지 사항을 고려할 필요가 있다. 첫째, 소득세 자료에서 사업소득의 소득기준은 수입금액에서 필요경비를 공제한 소득금액으로 되어 있다. 다만 이 필요경비는 국민계정에서 부가가치를 구할 때 공제되는 중간소비보다 상당히 넓게 인정되고 있다. 예컨대 연말정산 사업소득의 경우 2010년에 전체 수입금액에서 76%가 필요경비로 인정되었으며, 기타소득의 경우에도 수입금액의 80%를 필요경비로 인정하기도 한다. 그렇지만 사업체를 운영하면서 설비를 갖추고 원재료를 구입하는 경우에는 필요경비의 비중이 커질 수 있지만, 그렇지 않은 경우 사업소득이라 해도

인적 용역을 제공하는 성격이 강하다. 이들 소득에 대해 과세 제도상 필요경비가 인정되더라도 국민계정의 기준에서 보면 그것은 중간소비가 아니라 부가가치에 포함되며 소득으로 볼 수 있는 경우가 많다. 여기서는 이 점을 감안하여 이들 소득의 경우 수입금액을 소득으로 간주하기로 한다¹⁹. 둘째, 이들 소득 중에서 종합소득세 신고(따라서 통합소득)에 포함된 경우가 있으며, 이 중복되는 부분을 추정하여 공제할 필요가 있다.

먼저 연말정산 사업소득은 보험모집인과 방문판매원을 대상으로 그들의 수입에 대해 근로소득자와 같은 연말정산을 인정하고 있으며²⁰, 『연보』에는 별도의 통계를 제시하고 있다. 이 중에서 수입금액 기준의 소득구간별 통계를 전술한 방법에 의해 16개 소득구간으로 재편한 결과가 <표 3>에 제시되어 있다. 수입금액은 모집수당이나 판매수당으로 이루어져 있는데, 높은 수입을 올리고 있는 보험모집인의 경우 판촉을 위한 경비로서 중간소비가 적지 않게 포함되어 있을 것으로 생각된다. 다만 이를 추정하여 배제할 필요가 있지만, 그를 위한 정보가 없어 그대로 두었다.

원천징수 사업소득은 부가가치세가 부과되지 않는 사업으로서 의료보건의 종사자, 개인용역이나 봉사료 등의 소득을 대상으로 수입금액에 대해 3%의 소득세가 원천 징수된다. 이들 중에서 의료보건의 종사자의 경우는 대부분 종합소득세를 신고하였기 때문에 두 자료에서 중복되는 부분을 추정하여 공제할 필요가 있다. 『연보』에 따르면, 이들 중 종합소득 신고를 한 자는 426,420명으로 추정되는데²¹, 원천징수 사업소득자 중에서 소득이 높은 순서로 이 인원수만큼을 공제하였다. 그리고 다른 통계의 소득구간에 맞추어 정리한 결과가 <표 3>에 제시되어 있다.

기타소득은 소득세법상 사업소득과 구분되어 다른 세율이 적용되지만, 여기서는 사업소득에 넣어 파악하기로 한다²². 특히 개인용역에 해당하는 사업소득과 기타소득의 경계는 애매하다. 사업소득은 영리를 목적으로 계속적 반복적으로 행해지는 용역에서 얻은 소득이고 기타소득은 일시적 우발적으로 발생하는 소득으로 개념상 구분하고 있지만, 고용관계 없이 자기 책임하에 이루어진다는 점에서 사업소득의 성격을 갖는다고 생각된다. 기타소득은 『연보』에 수입금액 기준의 구간별 통계가 제시되어 있어 이를 이용하면 된다. 다만 그 중에서 종합소득 신고 대상인 경우는 이미 이미 통합소득에 포함되어 있기 때문에 이를 추정하여 중복 부분을 공제할 필요가 있다²³. 그

¹⁹ 이러한 방식은 거꾸로 중간소비로 지출한 부분이 있을 경우 소득이 그만큼 과대 평가될 수 있음에 유의할 필요가 있다.

²⁰ 이 제도는 직전 연도의 수입금액이 7,500만원 미만인 간편장부대상자를 대상으로 시행하는데, 그 결과를 보면 그 기준을 넘어 높은 수입을 올린 자들이 적지 않다. 이들은 다음 해에는 종합소득세 신고 대상자로 바뀌게 된다.

²¹ 『연보』의 종합소득세 신고자 중에서 2010년의 기납부세액의 인원수 1,363,108명(표 3-1-4(20/21))은 원천 징수된 소득자를 말하는데, 거기에서 근로소득(773,758명), 연금소득(3,273명), 기타소득(159,657명)으로 원천 징수된 소득자 수(이상 표 3-2-3)를 제하면, 사업소득 원천징수 대상자로서 종합소득 신고자 수를 구할 수 있다.

²² 기타소득을 몇 가지 예시하면, 상금, 복권, 원고료, 인세, 창작품에 대한 대가, 강연료, 저작권 수입, 자산이나 권리의 양도 및 대여의 대가, 일시적인 인적 용역에 대한 대가 등 다기에 걸쳐 있다. 이 중에서 상금과 복권처럼 본원소득이 아니라 이전소득에 속하는 것이 포함되어 있지만 자료의 부족으로 이를 분리하여 제외하지 못했다.

²³ 종합소득 통계에 포함되어 있는 기타소득을 분리하기 위해 김낙년(2012a: 87-88)이 종합소득 통계에

리고 16개 소득구간으로 정리한 결과가 <표 3>에 제시되어 있다.

그런데 사업소득 중에는 원천징수가 되지도 않고 국세청에 종합소득을 신고하지도 않은 미신고 사업소득이 적지 않게 존재한다. 그들의 규모와 분포에 관해서는 소득세 자료로는 파악할 수 없지만 부가가치세 자료를 이용하면 그 실태를 어느 정도 드러낼 수 있다. 먼저 부가가치세 자료에 의거하여 전체 개인 사업자수가 2010년에 4,694,694명(이중 일반사업자 2,845,843명, 간이사업자 1,848,851명)임을 알 수 있다²⁴. 이 중에서 종합소득을 신고한 사업소득자 수는 3,302,071명이고, 거기에 포함된 전술한 원천징수 사업소득자 426,420명을 뺀 2,875,651명이 소득신고를 한 개인사업자로 볼 수 있다. 따라서 이를 전체 개인사업자수에서 뺀 1,819,043명이 미신고 사업소득자가 된다. 이들은 대부분 매출 규모가 영세한 간이사업자이지만 일반사업자도 포함되어 있을 것으로 생각된다. 여기서는 간이사업자의 등록 기준인 연 매출 4,800만원에 미치지 못하는 간이사업자를 미신고자로 보고, 나머지 미신고자는 일반사업자 중에서 최하위 소득구간에서 나온 것으로 가정하기로 한다²⁵.

일반사업자 및 간이사업자의 부가가치세 통계(『연보』 표 9-3-1와 표 9-4-1)는 반기별 과표를 기준으로 하여 구간이 설정되어 있고, 과표 구간별로 사업자수와 과표 또는 매출액을 보여준다. 이를 소득 기준으로 재편하기 위해서는 몇 가지 조정이 필요하다. 먼저 연간 기준으로 바꾸기 위해서는 매출액의 경우 상반기와 하반기의 수치를 더하고, 동시에 과표 구간도 2배로 늘려준다. 그리고 과표 구간과 매출액은 모두 소득 기준으로 바꿀 필요가 있다. 이를 위해 『경제총조사』(2010년)에서 1인 사업자의 경우 매출액 대비 영업이익 비율을 구하면 28.4%로 나오는데, 이 비율을 과표 구간과 매출액에 모두 곱하여 소득금액 기준으로 전환해 주었다. 이 때 유의할 것은 이 과표 구간별 통계에는 과세대상 매출(표의 '과세분 매출'과 '영세율 매출')만 포함되고 '면세분 매출'은 과세대상이 아니어서 빠져 있다는 것이다. 이것은 세제상의 필요에서 나온 분류이지만 소득규모를 추정할 때에는 면세분 매출도 포함할 필요가 있다. 여기서는 과표 구간별로 면세분 매출을 추가하여 전체 매출액을 구했다. 이 때 과표 구간도 함께 늘려주는 조정이 필요한데, 이를 위한 구간별 조정률은 전술한 방식에 따라 적용하였다. 그 결과가 <표 3>의 미신고소득으로 실려 있다.

마지막으로 소득세 자료에서 누락된 사업소득으로서는 논밭을 작물생산에 이용함으로써 발생하는 농업소득(여기서는 재배업 소득으로 줄인다)이 있다. 이 소득은 소득세법상 비과세소득으로 인정되기 때문이다²⁶. 이들 재배업 소득의 규모는 국민계정에서 부가가치 생산액으로 파악되며 2010년에 17조 5,802억원이다. 문제는 이들 재배업 소득의 분포에 관한 정보이다. 여기서는 통계청의 『농림어업조사』(2010년)에서 얻을 수 있는 판매규모별 농가 수 통계를 이용하기로 한다. 거기에는 판매액을 11개 구간으로 나누어 1,177,318명의 농가가 구간별로 어떻게 분포되어 있는지를 보여준다²⁷. 그런데 이 판매액 이외에도 농가가 자급을 위해 판매하지 않은 생산물도 농업소득

서 근로소득을 분리할 때 이용한 방법을 원용하였다.

²⁴ 이 수치는 2010년의 1기와 2기의 개인 사업자 수의 평균을 취한 것이다(『연보』 표 9-3-1, 표 9-4-1).

²⁵ 간이사업자 중에서 연 매출이 4,800만원(이것은 소득금액 기준으로 환산하면 1,413만원) 미만인 경우는 1,768,437명이며, 나머지 80,415명은 일반사업자에서 나온 미신고자가 된다.

²⁶ 사업소득 중 비과세소득으로 인정되는 것에는 재배업 소득 이외에 농가 부업소득, 전통주 제조 소득, 주택임대소득(1개 주택 소유의 경우) 등이 있다.

²⁷ 이것은 농가 수인데, 종사자 수를 보면 전업 종사자가 1,130,628명, 겸업(1종 및 2종) 종사자가

에 포함할 필요가 있다²⁸. 여기서는 구간별로 농가의 판매액과 자금액의 합계에 작물의 부가가치율(2010년에 0.7, 한국은행의 산업연관표)을 적용하면 평균소득이 도출된다. 그리고 이 구간을 소득 기준으로 재편하고, 최상위 구간을 전술한 방법으로 세분한 결과가 <표 3>에 재배업 소득으로 제시되어 있다.

이상에서 통합소득에서 제외된 사업소득을 추가하였는데, 그것이 국민계정에서 파악된 사업소득과 비교해 보면 5조 6,257억원의 차이를 보인다²⁹. 이 차이가 어디에서 왔는지 몇 가지 생각해 볼 수 있다. 앞에서 개인용역 성격의 사업소득에 대해 인정되는 필요경비가 과대하다는 점을 언급했지만, 개인사업체의 경우에도 필요경비가 국민계정의 중간소비보다 다소 넓게 인정되고 있다. 그로 인해 국세청이 파악하는 사업소득이 다소 과소평가된다. 그리고 부동산 임대소득에서도 국세청의 파악은 충분하지 않은 것으로 생각된다. 한국은행은 국민계정에서 부동산 임대소득을 추정할 때 국세청 자료 이외에 임대료 지출 통계를 이용한 추정으로 보완하고 있다. 여기서는 이 차이를 사업소득의 조정항목으로 잡고, 사업소득이 그만큼 과소 파악된 것으로 본다³⁰. 다만 그 분포에 관한 정보가 없지만, 여기서는 종합소득 중 사업 및 부동산임대 소득의 구간별 분포에 따른 것이라는 가정으로 구간별로 배분하였다. 그 결과가 <표 3>에 제시되어 있다.

4. 금융소득

금융소득에 대한 과세는 이자와 배당을 합쳐 4,000만원이 넘을 경우 종합과세의 대상이 되지만, 그 미만의 금융소득에 대해서는 원천 징수하는 것으로 과세가 종료된다. 전자의 경우 『연보』에서 비교적 상세한 정보를 얻을 수 있는 반면, 후자에 대해서는 원천징수 총액만을 제시할 뿐 그 분포에 관한 정보를 주지 않는다. 전술한 국세청의 통합소득에 포함된 금융소득도 4,000만원 이상에 한정되며, 나머지 금융소득은 모두 빠져 있다³¹. 따라서 이를 추정해서 보완하는 것이 과제가

1,058,871명으로 나온다. 따라서 농가는 평균적으로 전업과 겸업 종사자가 각 1명으로 이루어진 것으로 볼 수 있다. 여기서는 농가의 농업소득을 이 전업 종사자에게 귀속시켜 개인 기준으로 상정하고 있다.

²⁸ 전체 농산물 생산액에서 여기서 구한 판매액을 빼면 판매되지 않은 농가 자금 분을 구할 수 있다. 여기서는 이 자금 분이 모든 농가에게 동일하게 배분된 것(농가당 665만원으로 추정)으로 가정하였다. 따라서 농가의 농산물 수입은 이 자금분과 판매분의 합계로 이루어진다.

²⁹ 앞에서 연말정산 사업소득과 기타소득을 추정할 때 소득금액이 아니라 수입금액을 이용하였는데, 이들 수입금액에서 중간소비로 인정될 수 있는 필요경비가 일부 포함되었기 때문에 이들 소득의 추정이 다소 과대하게 평가된 점이 있다. 이 점을 감안하면 이 차이는 그만큼 늘어날 것이다.

³⁰ 그런데 사업소득의 경우 무자료 거래 등에 의한 소득의 탈루가 적지 않을 것으로 생각된다. 이 경우는 의거할만한 자료 자체가 없기 때문에 국민계정에서도 이를 반영하기는 어렵다고 생각된다. 따라서 여기서 조정 항목은 탈루 소득까지 커버하기는 어렵고, 국민계정과 갭을 메우는 의미가 있다.

³¹ 최근 최재성 의원실이 국세청에 대한 자료 요청으로 이자 및 배당 소득 100분위 통계를 받아 공개한 바 있다. 그런데 그 통계는 비과세 및 면세 소득이 빠져 있는 불완전한 통계이므로 이용하기 어렵다. 다만 그럼에도 불구하고 이자와 배당 소득자 수가 각각 45,386,750명과 8,515,764명(2010년)으로 나와 있어 대부분의 국민이 소액이라도 금융소득을 얻고 있음을 알 수 있다. 다만 <표 3>의 '제

된다.

이를 위해 여기서는 『가계금융조사』(2010년)를 이용하였다. 이 자료는 1만 가구를 대상으로 경상소득과 금융자산을 포함한 각종 자산 가액을 조사한 것이다. 가구별 금융소득을 직접 조사한 것은 아니지만, 금융자산이 금융소득의 원천이라는 점을 감안하여 추정할 수 있다. 여기서는 국민계정에서 얻은 전체 금융소득(50조 1,879억원)을 가구별 금융자산에 비례한다는 가정으로 16,648,372 가구에 배분하였다. 그리고 이렇게 추정한 금융소득이 4,000만원 미만인 가구를 대상으로 그들의 경상소득 구간별로 가구당 경상소득과 금융소득, 그리고 경상소득 대비 금융소득 비중을 구했다. 그리고 전체 가구를 대상으로 하는 금융소득의 비중도 구해 그 결과를 <표 4>에 제시하였다.

<표 4> 소득구간별 금융소득 비중의 추정(2010년)

	가계금융조사					소득세 자료			추정				전체 비중 (%)
	금융소득 4,000만원 미만 가구				전체 가구	금융소득 종합과세 대상자			전체 소득 자수 (천 명)	금융소득 4,000만원 미만 소득자			
	가구수 (천 가구)	경상소득 (천 원)	금융소득 (천 원)	비중 (%)	비중 (%)	소득자수 (명)	금융소득 (10억 원)	비중 (%)		금융소득 1 (10억 원)	비중 1 (%)	금융소득 2 (10억 원)	
	A	B	C	D=C/B	E	F	G	H	I	J=(I-F)*C	K	L	
-1	50	278	1,028	370.21	543.41	3	0	0.00	4,008	4,121	153.32	201	7.48
1-5	1,003	3,387	379	11.20	11.20	13	0	0.00	7,178	2,723	12.06	1,688	7.48
5-10	1,592	7,306	725	9.92	9.92	20	0	0.00	5,149	3,731	9.59	2,907	7.48
10-20	2,703	14,588	1,092	7.49	8.47	92	0	0.00	5,848	6,388	7.63	6,388	7.63
20-30	2,690	24,068	1,712	7.11	7.26	105	2	0.00	3,268	5,595	7.54	5,595	7.54
30-40	2,456	33,768	2,283	6.76	6.85	53	1	0.00	1,864	4,255	6.69	4,255	6.69
40-50	1,790	43,880	3,008	6.86	6.99	4,770	161	0.30	1,219	3,654	6.72	3,654	7.02
50-60	1,326	53,240	3,975	7.47	7.53	3,182	154	0.32	863	3,419	7.22	3,419	7.54
60-70	988	63,107	5,003	7.93	8.13	3,262	183	0.47	601	2,992	7.72	2,992	8.19
70-80	596	73,536	5,630	7.66	7.79	2,226	172	0.63	366	2,046	7.49	2,046	8.12
80-90	414	83,365	7,131	8.55	8.92	2,633	168	0.76	261	1,840	8.31	1,840	9.07
90-100	258	93,480	7,466	7.99	8.43	1,897	178	1.21	155	1,144	7.78	1,144	8.99
100-200	642	124,213	9,677	7.79	9.82	13,052	1,185	2.61	357	3,325	7.34	3,325	9.95
200-300	52	220,334	13,213	6.00	9.88	5,822	748	6.81	46	537	4.89	537	11.70
300-500	20	326,432	13,082	4.01	6.74	5,003	961	11.74	22	220	2.69	220	14.43
500-	8	604,693	17,830	2.95	4.60	6,774	5,918	31.36	15	147	0.78	147	32.14
합계 / 평균	16,586	37,291	2,766	7.42	7.99	48,907	9,829	1.71	31,219	46,138	8.04	40,359	7.86

주: 1) 가계금융조사는 가구 기준이고, 나머지는 개인 기준의 통계이다.

2) 가계금융조사의 금융소득 비중은 경상소득 대비로 구한 것이고, 그 외의 비중은 전체소득에서 일부 항목(표6의 R, L, I)을 제외한 소득 대비로 구한 것이다.

3) 전체 소득자수는 후술하는 <표 6>의 인원수 합계에서 가져왔다.

4) 음영 부분의 3 구간은 전체 비중이 동일하다는 가정으로 추정된 것이다.

5) 마지막 행에서 기울어진 숫자는 합계이고, 나머지는 평균이다.

자료: 통계청, 『가계금융조사』; 『연보』; <표 6>.

이로부터 금융소득이 차지하는 비중(D, E)이 대부분의 소득구간에서 7% 또는 8% 전후로 일정하다는 사실을 알 수 있지만, 동시에 『가계금융조사』의 문제점도 드러나고 있다. 첫째, 상위 소득구간의 금융소득이 과소하게 파악되었다. 금융소득이 4,000만원인 가구의 금융소득 비중(표의 D)이

외된 금융소득'의 인원수는 따로 추정하지 않고 비워놓았는데, 여기서는 금융소득이 다른 소득 항목이 있는 자에게 귀속되는 것으로 보았다.

최상위 구간으로 갈수록 하락하는 것은 최상위 구간으로 갈수록 금융소득이 4,000만원을 넘는 가구가 많아지기 때문으로 생각되지만, 이들을 포함한 전체 가구의 금융소득 비중(E)도 하락하고 있는 것은 납득하기 어렵다. 소득세 자료에 의해 파악된 금융소득 종합과세 대상자의 금융소득 비중(H)와 비교하면 『가계금융조사』에서 최상위 금융소득자들을 제대로 포착하지 못했음을 알 수 있다³². 둘째, 최하위 소득구간에서도 금융소득 비중(D, E)이 100%를 넘는 등 납득하기 어렵게 나왔다. 경상소득에는 금융자산으로부터 얻은 소득도 포함되어야 하지만, 특히 하위 소득구간에서 그렇지 못한 경우가 적지 않았을 것으로 생각된다.

여기서는 이상의 문제점을 감안하여 다음과 같은 방법으로 소득구간별 금융소득을 추정하기로 한다. 먼저 금융소득 4,000만원 이상인 경우는 9조 8,290억원에 달하는데, 이에 관해서는 다행히 소득세 자료를 이용할 수 있기 때문에 문제는 없다. 다만 금융소득을 포함하고 있는 종합소득 통계(『연보』의 표 3-2-3)의 소득구간은 종합소득 기준으로 되어 있는데 이를 소득총액 기준으로 전환하는 등의 조정이 필요하다³³. 표의 소득구간별 금융소득 분포(F, G)는 그 결과이다.

한편, 금융소득 4,000만원 미만의 경우는 40조 3,590억원으로 그 규모가 훨씬 큰데 그 분포의 추정은 <표 4>의 『가계금융조사』의 정보를 이용하기로 한다. 거기에는 소득구간별로 가구당 금융소득(C)이 제시되어 있는데, 이를 본고에서 추계한 금융소득 4,000만원 미만의 개인에게 적용하기로 한다. 예컨대 가구 소득 60-70(백만원) 구간의 평균 금융소득은 5,003(천원)임을 알 수 있는데, 동일한 소득구간에 속하는 개인의 경우에도 평균 5,003(천원)의 금융소득을 얻고 있는 것으로 가정하고 있는 셈이다. 이 때 소득구간별 전체 소득자수(I)는 후술하는 <표 6>의 인원수 합계에서 가져온 것이다. 이렇게 구한 구간별 '금융소득 1'(J)은 46조원이 넘어 구하고자 수치보다 6조원 정도 과대하게 추정되었음을 알 수 있다. 여기서는 그것이 전술한 『가계금융조사』의 문제가 최하위 3개 구간(음영 표시)에서 나타난 것으로 보고, 그 구간에서 과대 추정 분을 삭감하되 3 구간의 금융소득 비중이 동일하게 되도록 해당 구간의 금융소득을 추정하였다. 그 결과가 '금융소득 2'(L)이고, <표 3>에서 통합소득에서 제외된 금융소득이 곧 그것이다.

이상의 추정 결과에 의거하여 소득구간별로 금융소득이 차지하는 비중을 보면, 금융소득 4,000만원 이상의 경우(H) 최상위 구간으로 갈수록 급속히 상승한 반면, 4,000만원 미만의 경우(K)는 하락하고 있다. 그것은 최상위로 갈수록 금융소득 종합과세 대상이 빠르게 늘어나기 때문이다. 양자

³² 『가계금융조사』는 최상위 구간에서 금융소득뿐만 아니라 경상소득도 과소하게 파악한 것으로 보인다. <표 4>에서 소득구간별 가구수(A)와 소득자수(I)를 비교해 보면, 높은 소득구간에서는 전자가 후자보다 많다가 20-30(백만원) 소득구간부터는 역전되고 있다. 전자는 가구원의 소득을 합한 가구소득이고 후자는 개인소득이므로 이러한 양상이 나타난 것이다. 그렇지만 최상위 2개 구간에서는 거꾸로 가구수가 개인 소득자수보다 적은 것으로 나온다. 그뿐 아니라 500(백만원) 이상 소득구간의 평균 가구소득(6억원)이 <표 3>의 해당 구간의 개인소득 평균의 1/2밖에 되지 않는다. 이들은 모두 『가계금융조사』가 최상위 소득을 제대로 파악하지 못했음을 보여준다.

³³ 소득구간별로 각 구간에 포함된 금융소득은 『연보』에서 얻을 수 있지만, 그 소득구간은 종합소득 기준이므로 이를 소득총액 기준으로 바꾸어줄 필요가 있다. 즉 종합소득에 포함되어 있는 근로소득은 근로소득 공제 등이 제외된 근로소득금액 기준이므로 이를 급여총액으로 바꿀 경우 소득구간도 그 비율만큼 조정될 필요가 있다. 이를 위한 조정률의 적용과 소득구간을 16개 구간으로 재편하기 위한 보간법은 전술한 방법에 따랐다.

를 합한 비중(M)을 보면, 1억원이 넘는 소득구간에서는 전체소득 대비 금융소득의 비중이 빠르게 늘고 있지만, 그 이하 소득구간에서는 그 비중은 어느 정도 일정하게 유지되고 있음을 알 수 있다.

이상의 결과가 <표 3>에 정리되어 있다. 통합소득과 거기에서 제외된 소득을 근로소득, 사업소득, 금융소득으로 나누어 하나하나 추가하였다. 소득세 자료가 일부 미신고 소득과 같은 누락도 있지만 대부분의 소득을 커버하고 있으며, 다른 소득 자료를 보완하면 전체 소득의 분포를 어느 정도 드러낼 수 있다는 점이 주목된다. 그리고 이상의 소득 자료와 국민계정의 개인소득과의 차이를 조정 항목으로 추가함으로써 양자를 일관성 있게 파악할 수 있도록 하였다.

IV. 개인소득의 분포와 분위별 소득비중

<표 3>은 각 소득항목에 대해 소득구간별 분포를 보여주지만, 이것이 개인소득의 분포가 되기 위해서는 추가적인 작업이 필요하다. 표의 각 소득항목 내부에서는 소득이 개인별로 합산되어 집계되기 때문에 중복되는 경우가 없지만, 동일인이 다른 소득항목에 걸쳐 복수의 소득원을 갖는 경우가 적지 않게 존재할 것으로 생각되기 때문이다. 종합소득 자료는 개인별로 통합된 소득을 보여주기 때문에 종합소득세 신고를 한 경우 이 문제는 없다. 금융소득의 경우 소득원이 그것뿐인 경우도 있지만, 근로소득 또는 사업소득과 겹치는 경우가 대부분이며, 이를 통합하는 문제는 후술하기로 한다. 이에 대해 하위 소득자의 경우에는 불완전 취업 등으로 인해 복수의 소득원을 갖지 않을 수 없는 경우가 적지 않을 것으로 생각되며, 여기서는 이 점을 먼저 검토하기로 한다.

하위 소득구간에서 여러 소득원에 걸쳐 소득을 얻는 경우를 예시하면, 임시직이나 일용직으로 일하면서 보험모집인 일을 하거나 파트 타임으로 인적 용역을 제공하면서 대리운전을 하는 등이다. 인적 용역의 제공에서 원천징수 사업소득과 기타소득의 경계가 모호하여 구분하기 어려운 경우도 많다. <표 3>에서 이들은 세제상의 차이 때문에 각각 다른 소득항목으로 파악되지만, 현실에서 그 구분은 큰 의미가 없다. 그렇지만 불행히도 그 실태를 보여주는 자료는 거의 찾기 어렵다. 여기서는 불가피하게 가정에 의한 추정을 시도해 보고자 한다. 그를 위한 실마리가 일용근로소득자 7,825,444명 중에서 다른 근로 또는 사업소득이 없는 순수 일용근로소득자가 5,490,665명이라는 『연보』(표 4-4-8의 각주)의 조사이다. 즉 일용근로자의 29.8%가 다른 소득원을 가지고 있다는 뜻이다. 국세청에 따르면 일용근로자는 동일한 고용주에 의해 3개월(건설업의 경우 1년) 이상 계속 고용되지 아니한 자로 정의된다. <표 3>에서 보듯이 이들의 연소득이 100만원(또는 500만원)에도 미치지 못하는 경우가 200만명(또는 500만명)이 넘는다. 그들 중에는 자발적으로 이루어지는 단기적인 아르바이트 성격의 근로도 다수 포함되어 있겠지만, 불완전 취업으로 인해 다른 소득원을 갖는 경우도 많을 것으로 생각된다. 여기서는 <표 3>의 과세미달자, 연말정산 또는 원천징수 사업소득, 기타소득의 소득항목에서도 이와 유사한 불완전 취업자가 다수 존재할 것으로 보고, 그들의 경우에도 일용근로소득과 동일한 29.8%(다만 기타소득의 경우는 70%³⁴)가 다른 소

³⁴ 기타소득의 경우 다른 소득항목에 비해 일시적 우발적 성격이 짙고 저작권이나 권리의 대여 등에서 발생하는 경우가 많아 다른 소득항목과 중복될 확률이 높다는 점을 감안한 것이다.

득원을 가진 것으로 가정하기로 한다.

<표 5>가 이들 소득항목에서 중복이 있을 경우 이를 개인별로 통합소득을 구하는 절차는 요약한 것이다. 먼저 표의 (1)은 위의 비율을 적용하여 구한 복수의 소득원을 가진 소득자들이 하위의 4개 소득구간에 분포할 것으로 가정³⁵하고, 그들을 각 구간의 소득자수와 비례하도록 배분한 것이다. 그리고 이들은 1개의 다른 소득원을 갖는 것으로 가정한다³⁶. 예컨대 일용근로소득자는 4개의 다른 소득항목 중의 하나와 결합하여 2개의 소득원을 갖는다. 그러면 (1)의 인원수 합계는 640만 명으로 되어 있지만, 실제의 인원수는 그 1/2인 320만명이 된다. 이 때 어떤 소득자가 가지고 있는 두 개의 소득원을 결합하면, 그 통합된 소득이 속하는 구간은 상방 이동할 가능성이 있다. 여기서는 이 점을 반영하여 소득구간별 실제 인원수와 소득액을 추정하는 것이 과제가 된다.

이를 위해서는 먼저 두 구간이 결합되는 경우의 수를 생각해 볼 필요가 있다. 최하위 구간부터 4개 구간을 a, b, c, d로 지칭하기로 하자. 그러면 두 개의 소득원이 결합할 수 있는 경우의 수는 10개(aa, bb, cc, dd, ab, ac, ad, bc, bd, cd)가 된다. 이들 각각의 경우 통합된 소득이 해당 구간에 머물거나 상위 구간으로 올라갈 수 있는데, 어떤 비율로 그렇게 되는가는 추정할 필요가 있다. 예컨대 aa의 경우 두 구간의 하한-평균-상한의 합계는 0-100-200(만원, 이하 생략)이 된다. 차상위인 b 구간의 하한이 100이므로 두 합계 구간의 중앙에 온다. 이 경우는 a 구간에 머물거나 b 구간으로 올라갈 확률은 동일하므로 각 1/2로 본다. bb인 경우는 100-300-500인데 두 구간을 합하면 200-600-1000이 된다. 차상위 구간인 c의 하한이 500이므로 200-1000을 500을 기준으로 나누면 b 구간에 머물 확률은 3/8, c 구간으로 올라갈 확률은 5/8이 된다. cc의 경우는 500-1000을 합하면 1000-2000이 되어 d 구간과 동일하게 되므로 모두 d구간으로 귀속하면 된다. dd인 경우 합하면 2000-4000이 되므로 2000-3000구간(e)과 3000-4000구간(f)에 각 1/2씩 배분하면 된다³⁷.

그 다음 각 경우의 수가 어떤 확률로 일어날 것인지를 생각할 필요가 있다. 표의 (2)는 결합할 다른 소득의 인원수가 많을수록 결합될 확률이 높다는 가정으로 그 확률을 구한 것이다. 예컨대 a 구간의 일용근로소득은 39.1%로 나와 있는데, 그것은 일용근로소득이 동일한 a 구간의 다른 소득(그들의 소득종류는 불문)과 결합될 확률을 뜻한다. 그것을 (1)에서 일용근로자를 제외한 다른 4개 소득의 인원수 합계 중에서 a 구간에 속하는 인원수의 비중으로 구한 것이다. 일용근로소득의 b, c, d 구간의 확률이 31.4%, 16.5%, 12.9%로 떨어지는 것은 상위 구간으로 갈수록 인원수가 줄어드는 것을 반영하고 있다.

이상의 정보를 이용하면, 통합된 소득의 구간별 인원수를 구할 수 있다. 예컨대 b 구간의 일용근로소득자는 933,567명인데, 그들이 동일한 b 구간의 다른 소득과 결합될 확률은 31.4%이다. 그리고 bb의 경우는 b 구간에 머물 확률이 3/8이고 c 구간으로 이동할 확률이 5/8이므로 각각을 곱하면 각 구간으로 재배치될 인원수를 구할 수 있다. 이를 모든 경우에 수에 적용한 결과가 (3)이다. 이것을 (1)의 인원수와 비교하면 구간별 인원수가 상위 구간 쪽으로 이동하였고, 그 중의 일부는 20-30, 30-40(백만원) 구간에까지 걸쳐 있음을 알 수 있다. (3)의 합계 인원수를 1/2로 줄이

³⁵ 연소득이 2,000만원을 넘으면 full time job이 될 가능성이 높아진다는 점을 감안하였다.

³⁶ 3개 이상의 소득원을 가진 경우도 있을 수 있겠지만, 현실적으로 많지 않을 것으로 생각되어 이렇게 가정하였다.

³⁷ 같은 요령으로 ab는 b와 c에 4/5, 1/5, ac는 c와 d에 5/6, 1/6, ad는 d, e에 10/11, 1/11, bc는 c와 d에 4/9, 5/9, bd는 d와 e에 9/14, 5/14, cd는 d와 e에 1/3, 2/3으로 배분된다.

면 실제 인원수가 되는데, <표 6>의 '통합소득 2'의 인원수가 그 결과이다. <표 6>은 여기서 수정한 것을 제외하면 <표 3>의 수치를 그대로 가져온 것이다. 과세미달자로부터 기타소득까지의 음영 부분은 복수 소득원을 가진 것으로 추정된 소득자를 제외한 대신에 그들의 소득을 통합한 결과를 '통합소득 2'의 항목으로 추가하였다.

<표 5> 일부 소득항목의 통합소득 추정

	과세미달자	일용근로자	연말정산	원천징수	기타소득	합계
(1) 복수의 소득원을 가진 것으로 가정된 인원수(단위: 명)						
-1	208,903	681,251	53,153	315,548	1,045,045	2,303,899
1-5	705,082	933,567	54,135	304,007	238,443	2,235,234
5-10	516,117	421,974	31,104	76,846	61,303	1,107,344
10-20	339,861	297,987	36,975	96,993	62,597	834,413
합계	1,769,963	2,334,779	175,367	793,394	1,407,388	6,480,890
(2) 각 소득이 해당 구간의 다른 소득을 복수로 가질 확률(단위: %)						
-1	44.5	39.1	35.7	35.0	24.8	
1-5	32.5	31.4	34.6	34.0	39.4	
5-10	12.6	16.5	17.1	18.1	20.6	
10-20	10.5	12.9	12.6	13.0	15.2	
합계	100.0	100.0	100.0	100.0	100.0	
(3) 복수의 소득원을 통합한 구간별 인원수(단위: 명)						
-1	46,451	133,309	9,486	55,158	129,650	374,054
1-5	437,460	706,615	46,676	264,610	541,213	1,996,574
5-10	546,369	657,966	44,946	213,313	377,548	1,840,141
10-20	557,854	633,936	53,929	196,738	297,015	1,739,471
20-30	163,990	183,677	17,992	57,288	57,201	480,147
30-40	17,839	19,277	2,338	6,288	4,761	50,503
합계	1,769,963	2,334,779	175,367	793,394	1,407,388	6,480,890
(4) 패널 (1)에 대응하는 소득금액(단위: 10억원)						
-1	104	263	27	158	523	1,075
1-5	2,025	2,430	128	525	277	5,384
5-10	3,707	3,018	227	576	506	8,034
10-20	4,730	4,213	540	1,413	873	11,769
합계	10,565	9,924	922	2,672	2,178	26,261
(5) 패널 (3)에 대응하는 구간별 소득금액(단위: 10억원)						
-1	12	26	2	14	32	86
1-5	819	953	54	243	309	2,379
5-10	2,599	2,334	153	492	440	6,019
10-20	4,876	4,328	440	1,189	865	11,698
20-30	1,970	1,965	232	628	453	5,248
30-40	290	318	40	107	77	832
합계	10,565	9,924	922	2,672	2,178	26,261

주: 1) 패널 (3)은 패널 (1)의 인원이 두 개의 다른 소득원을 가졌다는 가정으로 구간별 인원수를 재편한 것이다. 따라서 순 인원수는 그 1/2이 된다.

2) 복수의 소득원을 통합하여 인원수와 소득금액을 추정한 방법에 관해서는 본문 참조.

자료: 본문 참조.

이상은 인원수에 관한 것인데, 그에 대응하는 소득금액에 대해서도 같은 요령으로 추정하였다. (4)의 소득금액은 (1)의 인원수가 해당 구간의 전체 인원수에서 차지하는 비율(즉 중복 소득자의 비율)을 구해 해당 구간의 소득금액에 곱해서 구했다. 그리고 (5)의 소득금액도 (3)의 인원수를 산출하는데 이용한 방법을 적용하였다³⁸. 그리고 (5)의 재편된 구간별 소득금액의 합계가 <표 6>의 '통합소득 2'의 소득금액이 된다. 앞의 인원수의 경우에는 실제 인원수를 구하기 위해 1/2를 하였지만, 소득금액의 경우는 그럴 필요가 없다.

이상과 같이 하위 소득구간에서 복수의 소득원을 가진 경우 개인소득으로 통합하고자 하였지만, 가정에 크게 의존하였다는 점에서 한계가 있다. 예컨대 일용근로소득의 중복 비율을 다른 소득항목에도 그대로 적용했는데, 새로운 정보가 나오면 보완할 필요가 있다. 복수의 소득원을 2개로 한정하였고 2,000만원을 넘을 경우 그 가능성을 배제했는데, 실제로는 그렇지 않은 경우도 있을 것이다. 보통 소득이 높아지면 불완전 취업의 가능성이 낮아진다는 점을 감안한 것인데, 기타소득을 가진 경우는 그렇지 않을 수도 있다³⁹. 이상의 추정은 여러 가지 점에서 한계가 있지만, <표 3>과 <표 6>의 음영 부분을 비교해 보면 하위 소득구간에서 개인소득을 통합할 경우 인원수와 소득금액의 구간별 분포에 어떤 변화가 나타나는지를 보여준다는 점에서 일정한 의미가 있다고 생각된다⁴⁰.

한편, 금융소득의 경우 앞에서 구간별로 국세청의 통합소득(표 6의 '통합소득 1')에서 제외된 금융소득이 얼마나 되는지를 추정하였는데, 이를 구간별로 더할 필요가 있다. 그리고 사업소득과 근로소득의 조정항목도 '통합소득 1'에서는 과소하게 파악되었으므로 마찬가지로 추가할 필요가 있다. 그런데 <표 6>에서 '통합소득 1'에서 '통합소득 2'까지의 소득항목은 구간별 개인소득을 보여주는 것이므로 이를 구간별로 더할(즉 horizontal sum) 수 있다. 그에 대해 제외된 금융소득을 포함한 마지막 세 항목은 그렇지 않으며, 이를 추가할 경우 소득구간도 그에 따라 조정이 필요하다. 이를 위한 조정률의 적용과 16개 소득구간으로 재편하기 위한 보간은 앞에서 제시한 방법을 따랐다. 그 결과로 얻은 인원수와 소득금액이 '통합소득 3'으로 제시되어 있다.

³⁸ 각 경우의 수에 따른 소득금액의 배분은 인원수의 배분과 한가지 차이가 있다. 예컨대 bb의 경우 인원수를 b 구간에 3/8, c 구간에 5/8를 배분하였다고 했는데, 소득금액의 경우 b 구간보다 c 구간으로 상방 이동한 인원수의 평균소득이 더 높다는 점을 반영할 필요가 있다. 그 경우 b와 c 구간으로의 배분 비율은 각각 21.9%와 78.1%로 되어 c 구간으로의 배분 비율이 더 커진다.

³⁹ 예컨대 교수의 강연료나 원고료 또는 전문가의 인적 용역 등이 그러하다. 다만 이 경우에도 기타소득이 1,500만원을 넘으면 종합소득 신고 대상이 되므로 거기에서 파악되게 된다.

⁴⁰ 예컨대 가정을 바꾸어 하위 소득구간에서 원천징수 소득항목간 중복 비율을 바꿀 경우 결과가 어떻게 달라지는지를 모의실험을 해 볼 수 있다. <표 3>과 <표 6>의 음영 부분을 비교하면 0-10(백만원) 구간의 인원수가 줄어든 대신에 10-40(백만원) 구간의 인원수가 늘었지만 전체적으로 320만명이 줄어들었다. 만약 일용근로소득 이외의 다른 소득항목의 중복 비율을 29.8%에서 50%(기타소득은 80%)로 높이면, 0-20(백만원) 구간의 인원수는 줄어들고 20-40(백만원) 구간의 인원수는 늘지만 전체적으로는 103만명이 줄어드는 것으로 나온다. 이를 통해 위의 가정이 하위 소득구간의 분포에 어떤 방식으로 영향을 주는지 알 수 있다.

<표 6> 각종 소득의 구간별 인원수와 소득금액과 개인별 통합소득 (2010년)

(1) 인원수 (단위: 명)

소득구간 (백만원)	통합소득1 A~E, G	제외된 근로소득		제외된 사업소득					통합소득2 F, G, H, M, N, (O-D)	제외된 금융소득 R	조정 항목		합계	통합소득3
		과세미달자 F, G	일용근로자 H	연말정산 M	원천징수 N	기타소득 (O-D)	미신고 K	재배업 P			사업소득 L	근로소득 I		
-1	229,507	440,904	1,449,536	77,095	728,830	410,404	483,501	937	187,027				4,007,742	3,302,921
1-5	755,436	1,488,124	1,986,404	78,518	702,174	93,640	869,996	205,311	998,287				7,177,891	6,576,162
5-10	1,101,906	1,089,302	897,858	45,113	177,495	24,074	362,708	530,217	920,071				5,148,745	5,216,319
10-20	2,988,771	717,300	634,044	53,629	224,028	24,583	102,837	232,662	869,735				5,847,590	6,004,182
20-30	2,168,288	301,557	352,778	60,362	33,285	25,443		85,964	240,073				3,267,750	3,579,185
30-40	1,531,463	94,737	120,196	35,447		9,098		48,018	25,252				1,864,212	2,095,830
40-50	1,114,117	25,443	30,604	24,663		5,323		19,269					1,219,419	1,377,227
50-60	817,989	3,902	10,402	14,654		3,422		12,921					863,291	877,043
60-70	574,893	877	4,261	10,110		1,925		9,311					601,377	693,961
70-80	349,255	114	1,989	5,408		1,107		7,766					365,639	433,976
80-90	252,282	48	1,024	2,738		662		3,901					260,654	271,454
90-100	148,421	28	568	1,559		648		3,901					155,124	222,709
100-200	335,215	40	1,001	2,899		1,731		15,747					356,633	463,407
200-300	44,761	9		176		348		1,192					46,487	62,540
300-500	21,404	5		32		233		183					21,858	25,733
500-	14,494	4		4		523		17					15,043	16,806
합계	12,448,203	4,162,395	5,490,665	412,407	1,865,812	603,166	1,819,043	1,177,318	3,240,445				31,219,454	31,219,454

(2) 소득금액 (단위: 10억원)

-1	45	220.45	560.53	38.55	364.42	205.20	179	0	86	201	5	341	2,246	1,434
1-5	2,528	4,273.31	5,170.47	185.22	1,212.46	108.79	2,733	980	2,379	1,688	132	3,142	24,532	18,077
5-10	8,603	7,822.99	6,420.67	329.95	1,331.21	198.60	2,515	3,636	6,019	2,907	459	3,901	44,144	37,915
10-20	44,138	9,982.18	8,964.51	783.36	3,263.91	342.67	1,219	3,271	11,698	6,388	763	5,447	96,260	86,654
20-30	53,451	7,257.37	8,550.73	1,490.42	692.08	611.43		2,120	5,248	5,595	563	3,534	89,113	87,832
30-40	53,169	3,203.36	4,059.02	1,240.66		318.93		1,633	832	4,255	361	1,678	70,750	72,597
40-50	49,711	1,113.24	1,347.69	1,102.52		238.41		867		3,654	262	557	58,853	61,528
50-60	44,881	210.96	563.86	805.99		188.24		699		3,419	249	233	51,250	48,113
60-70	37,060	54.94	274.08	652.97		124.60		605		2,992	240	113	42,117	45,027
70-80	26,078	8.56	148.04	405.58		82.93		577		2,046	225	61	29,632	32,474
80-90	21,431	4.05	86.48	231.76		56.32		332		1,840	176	36	24,193	22,968
90-100	14,064	2.63	53.67	148.14		61.54		371		1,144	186	22	16,053	21,115
100-200	42,455	5.30	130.70	363.50		233.48		2,137		3,325	894	54	49,598	58,484
200-300	10,583	2.12		41.00		84.82		276		537	361		11,886	14,657
300-500	8,015	2.04		11.46		92.16		65		220	287		8,693	9,539
500-	18,004	5.76		2.47		844.97		11		147	462		19,478	20,386
합계	434,215	34,169	36,330	7,834	6,864	3,793	6,647	17,580	26,261	40,359	5,626	19,119	638,798	638,798

주: 1) 통합소득 2는 과세미달자에서 기타소득까지의 소득항목(음영 부분) 중에서 복수의 소득원을 가진 경우를 추출하여 이를 개인별로 통합한 것이다.

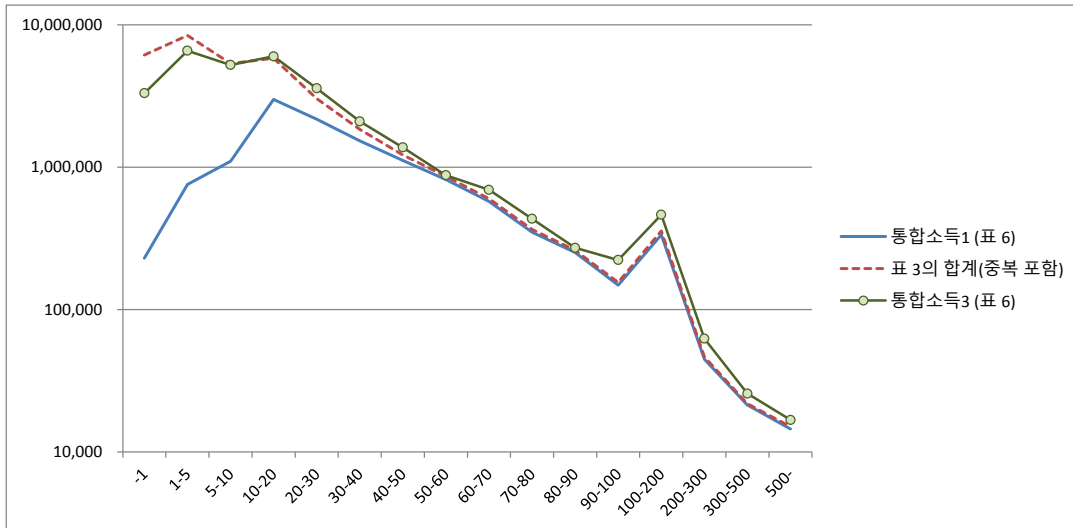
2) 통합소득 3은 제외된 금융소득과 조정항목까지 포함한 개인별 통합소득이다.

자료: 본문 참조.

<그림 3>은 이상의 소득구간별 인원수의 추계 결과를 비교한 것이다. <표 6>의 '통합소득 1'은 국세청이 제시한 것이지만, 거기에서 커버하지 못한 소득자가 다수 존재하며 그 소득에도 4,000만원 미만의 금융소득이 제외되어 있는 등 통합소득이라 부르기에는 한계가 있다. 한편 '표 3의 합계'는 '통합소득 1'에서 빠진 소득자를 원천징수 소득세 자료 등을 이용하여 보완하였지만, 그들 내부의 중복을 포함하고 있고, 4,000만원 미만의 금융소득도 빠져 있는 것이다. '통합소득 1'과 비교해 보면 양자의 차이는 하위 소득구간에서 주로 나타나고 중상위 구간에서의 차이는 그렇게 크지 않았다. 이에 대해 '통합소득 3'은 여기에 두 가지 조정을 추가한 효과가 결합되어 나타난 것이다. 하나는 하위 구간에서 중복을 배제한 결과 0-10(백만원) 구간에서는 인원수가 줄어든 반면 10-40(백만원) 구간에서는 다소 늘어난 전술한 '통합소득 2'의 효과이다. 또 하나는 4,000만원 미

만의 금융소득과 조정 항목이 추가된 효과인데, 이것은 최하위 두 구간을 제외한 전체 구간에 걸쳐 인원수를 늘리는 결과를 가져왔다. 이 '통합소득 3'이 본고가 추계한 개인소득 분포의 최종결과가 된다.

<그림 3> 소득구간별 인원수 분포의 비교



주: Y축은 인원수(단위: 명)이며 로그 눈금으로 표시하였다.

자료: <표 3, 6>.

그런데 '통합소득 3'에서 포착된 전체 소득자 수가 31,219(천명)으로 나왔는데, 이를 검토하기 위해 다른 관련 통계와 비교해 보자. 먼저 통계청의 경제활동인구 조사에서 임금근로자와 비 임금근로자를 합한 전체 취업자 수는 23,829(천명)이며 이에 크게 미치지 못한다. 취업자가 아닌 소득자가 700만명이 넘는다는 뜻인데, 이를 어떻게 이해해야 할까? 한편 가계조사에서는 개인별 소득 정보를 얻을 수 있는데, 취업 여부를 불문하고 근로소득이나 사업 및 부동산임대 소득이 있는 소득자 수를 집계하면 27,891(천명)으로 나온다. 가계조사에서는 농업 소득자가 빠져 있으므로 이 들까지 포함하면, 29,069(천명)에 이른다.

두 자료의 소득자가 모두 통계청의 취업자보다 이렇게 많은 것은 재산소득(금융이나 임대료 소득)만 있는 경우도 있지만, 근로소득이나 사업소득을 가지고 있더라도 취업자로 파악되지 않는 경우가 생기기 때문이다. 예컨대 1년 중 3달만 일을 한 경우를 생각해 보면, 그가 1년 중 취업자로 파악될 확률은 3/12에 불과하지만, 소득세 자료는 일의 지속 기간을 불문하고 한 번이라도 소득이 발생하면 소득자로 파악한다. 따라서 본고에서 파악된 소득자 수가 취업자보다 많은 것은 1년 중 상당 기간 비 경제활동인구(또는 실업자)로 존재하다가(따라서 취업자로 파악될 확률이 낮지만) 일시적인 수입 활동으로 소득을 얻는 경우가 그만큼 많다는 것을 뜻한다. <표 6>을 보면 근로 또는 사업 소득으로 번 소득이 연간 100만원이나 500만원에도 미치지 못하는 소득자가 매우 많고, 취업자에 들어가기 위한 경계소득(즉 소득 순으로 23,829천명에 해당하는 소득)이 308만원에 불과한 것을 감안할 때 취업자 중의 하층을 포함하여 상당수가 취업자와 비 경제활동인구의 경계에 위치하는 것으로 생각된다. 결국 취업자를 조사하는 1주일간의 조사기간 동안에는 취업자와 비 경제활동인구의 구분이 명확하다고 할 수 있지만, 1년의 기간을 놓고 보면 양자 사이를 오가는 경

우가 많고 구분이 모호해진다고 할 수 있다. 이들 하위 소득자 중에는 자발적인 파트타이머도 상당수 포함될 것이지만, 불완전 취업형태가 상당히 일반적임을 시사한다.

이 점을 염두에 두고 소득이 포착된 모든 개인(3,122만명)을 대상으로 하는 소득분포를 보면, 1,000만원 미만인 자가 전체 소득자의 절반에 가까운 48.4%를 차지하고 있다. 1,000-4,000만원, 4,000만원-1억원, 1억 이상의 구간은 각각 37.4%, 12.4%, 1.8%로 나온다. 전체 평균소득이 2,046만원이지만, 소득분포의 중앙값(Median)으로 추정된 중위소득은 1,074만원(평균소득의 52.5%)에 불과한 것으로 나온다. 이를 취업자수(2,383만명) 기준⁴¹으로 바꾸어 거기에 들지 못한 연소득 308만원 미만을 제외하고 보면, 상기의 4개 구간의 인원수 분포는 각각 32.3%, 49.0%, 16.3%, 2.4%로 나온다. 그리고 이들의 평균소득이 2,640만원, 중위소득은 1,594만원(평균소득의 60.4%)로 나온다. 취업자 기준으로 보면 다소 나아지기는 했지만, 평균소득과 중위소득간의 갭이 여전히 크게 벌어져 있는 데서 알 수 있듯이 개인소득의 분포는 그만큼 하위 소득구간에 몰려 있음을 알 수 있다.

이하에서는 <표 6>의 결과를 이용하여 두 가지 과제를 더 천착해 보기로 한다. 첫째, 여기에서 얻은 개인소득의 분포는 통계청의 가계조사를 검토하는데 활용될 수 있다. 종래 가계조사가 최상위 소득구간으로 갈수록 소득자를 과소하게 파악하였다는 문제점이 제기(김낙년·김종일 2013)되어 있지만, 그 이하의 소득구간에서는 어떠한지 알 수가 없었다. <그림 4>은 근로소득자와 사업소득자로 나누어 모든 소득구간에서 가계조사와 소득세 자료의 인원수가 어떻게 분포되어 있는지를 보여준다. 가계조사의 경우 원자료(마이크로데이터)에서 근로소득과 사업소득이 0보다 큰 인원수를 16개 소득구간별로 집계한 것이다. 소득세 자료의 경우 패널 (A)와 (B)는 각각 <표 3>과 <표 6>의 인원수를 이용하였는데, 그 때 '통합소득 1'에 포함되어 있는 근로소득자와 사업소득자는 『연보』를 이용하여 분리하였다⁴².

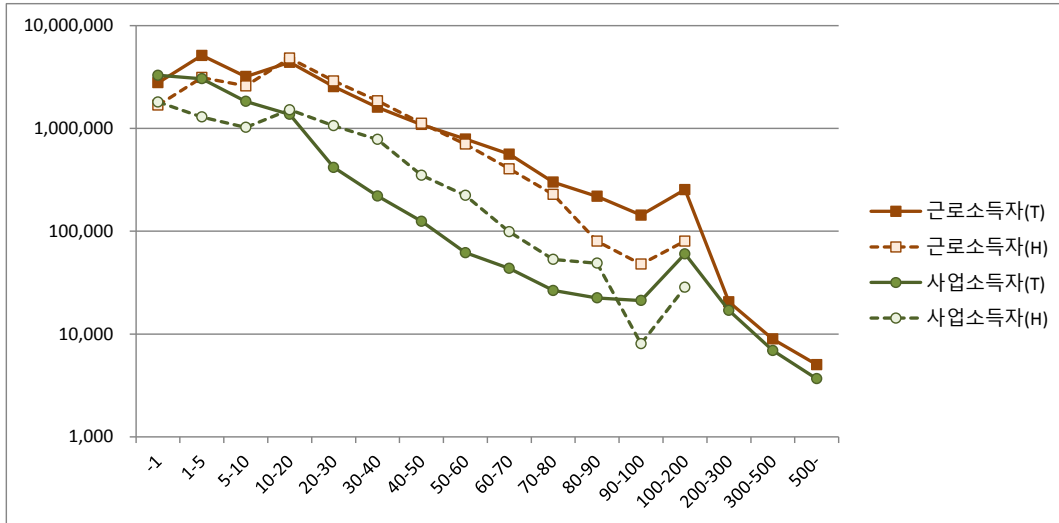
먼저 소득이 0보다 큰 근로소득자와 사업소득자의 총수를 보면, 가계조사의 경우 각각 19,612(천명), 8,280(천명), 소득세 자료의 경우 23,002(천명), 10,538(천명)으로 나왔다. 소득세 자료가 가계조사보다 크게 나온 이유 중의 하나는 전자의 경우 복수 소득원을 가진 경우가 중복 계산되기 때문인데, <그림 4>의 (A)에서 소득세 자료에는 이러한 중복 계산이 포함되어 있다는 점에 유의할 필요가 있다. 그렇지만 (B)에서는 근로소득자와 사업소득자를 통합하는 대신에 중복 계산이 제외된 소득자 수(표 6)를 제시하였다. 그 총수는 30,042(천명)인데, 가계조사의 27,891(천명)보다는 여전히 많지만, 그 차이는 크게 줄었다.

<그림 4> 근로 및 사업 소득자 수의 소득구간별 분포: 가계조사 vs. 소득세 자료(2010년)

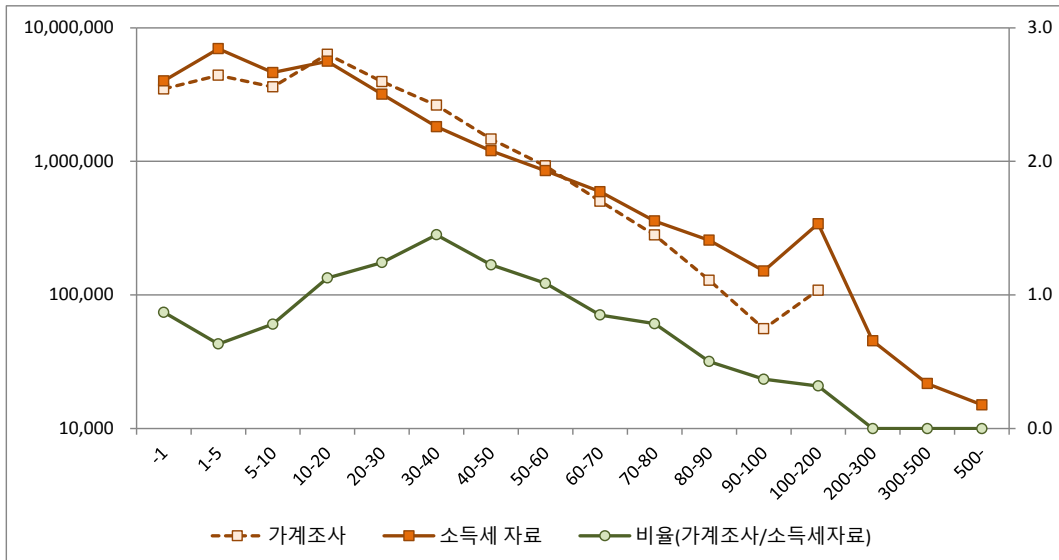
(A) 근로소득자와 사업소득자

⁴¹ 여기서 취업자수 기준이란 전체 소득자 중에서 소득이 높은 순으로 취업자수에 해당하는 인원수를 대상으로 한다는 뜻이다. 따라서 여기에는 취업자가 아니더라도 재산소득이 높은 자가 포함되거나 거꾸로 소득이 낮은 취업자가 제외되는 경우가 있다.

⁴² '통합소득 1'에 포함된 근로소득 납세자는 『연보』(표 4-2-4)에서 얻을 수 있는데, 그 통계의 소득구간은 비과세소득이 제외된 과세대상근로소득 기준으로 되어 있다. 이를 소득총액 기준으로 전환할 필요가 있다. '통합소득 1'에 포함된 사업 및 부동산임대 소득자는 『연보』(표 3-2-3)에서 얻을 수 있지만, 이를 다른 종류의 소득에서 분리할 때 소득구간을 조정할 필요가 있다. 그 방법은 김낙년(2012a: 87-88)이 같은 통계에서 근로소득자를 분리할 때 이용했던 방법을 원용하였다.



(B) 근로 및 사업 소득자 수의 합계와 두 자료의 괴리 비율



주: 1) Y축은 소득자 수(단위: 명, 로그 눈금). 가계조사는 H, 소득세 자료는 T로 구분하였다. 패널 (B)에서 비율은 오른쪽 눈금.

2) 두 자료의 커버리지를 일치시키기 위해 소득세 자료에서는 재배업 소득을 제외하였다.

3) 소득세 자료에서 복수의 소득원이 있는 경우 패널 (A)에서는 중복 집계되었지만, 패널 (B)에서는 이를 개인별로 통합한 결과를 제시하였다.

자료: 통계청, 가계동향조사(마이크로데이터); <표 3, 6>.

이러한 점에 유의하면서 <그림 4>의 두 통계를 비교해 보기로 하자. 먼저 (A)에서 근로소득자의 분포를 비교해 보면 10-50(백만원) 구간에서는 가계조사가 소득세 자료보다 약간 더 많이(3.0%에서 15.6%) 파악된 것으로 나오지만, 이를 전후한 구간에서는 반대로 가계조사의 파악한 인원수가 소득세 자료보다 훨씬 적어진다. 이러한 양상은 사업소득자의 경우 더욱 심해진다. 10-90(백만원) 구간에서는 가계조사의 인원수가 소득세 자료보다 많다가 그 전후 구간에서는 반대로 나타나는데, 그 차이가 훨씬 크다. 예컨대 20-90(백만원) 구간에서 가계조사의 인원수가 소득세 자료에 비해 무려 218%-360%가 더 많다. 소득세 자료에 전술한 중복계산의 문제가 포함된 점을 감안해

도 납득이 가지 않는다⁴³.

<그림 4>의 (B)는 근로소득자와 사업소득자를 합계하면서 복수 소득원의 중복 계산을 배제한 결과를 제시하였다. 그에 따르면 10-50(백만원) 구간에서는 가계조사의 인원수가 소득세 자료보다 더 많고 그 전후 구간에서는 반대로 나타나고 있다. 두 자료의 인원수를 비교할 수 있도록 비율(그림의 오른쪽 눈금)로도 제시하였다. 그에 따르면, 소득세 자료 대비 가계조사의 인원수 비율은 역 U자형을 보이고 있다. 다시 말하면 중간 소득구간에서는 그 비율이 1을 넘어 소득자를 과대 파악하고 있는 반면, 그 전후의 소득구간에서는 1 미만으로 빠르게 떨어지고 있음을 알 수 있다. 이것은 종래 가계조사에 의거한 통계청의 지니계수를 비롯한 소득분배 지표가 왜 실제보다 소득 불평등을 과소하게 파악하게 되는지를 보여주고 있다⁴⁴.

둘째, <표 6>의 '통합소득 3'의 결과는 16개 소득구간별로 제시되어 있지만, 이를 분위별로 재편해 보고자 한다. 소득집중도(top income shares)에 관한 기존 연구(World Top Incomes Database)는 모두 상위 10%에 초점을 맞추고 있지만, 본고의 결과는 이를 상위 10%뿐만 아니라 하위 90%를 세분해서 볼 수 있다. 그리고 본고의 결과가 한국에 관한 기존 연구(Kim and Kim 2014)와 어떤 차이가 나는지도 검토하고자 한다. 분위별 통계를 구할 때 분모에 들어가는 전체 인원수(control total)는 기존 연구와 비교 검토할 수 있도록 먼저 기존 연구와 동일하게 20세 이상 성인 인구(37,967,813명)로 한다. 이와 함께 다른 기준을 적용할 경우 분위별 소득비중이 어떻게 달라지는지를 보기로 한다.

'통합소득 3'의 16개 소득구간별 인원수 및 소득금액 통계를 이용하여 최상위 소득구간으로부터 인원수를 기준으로 10개 분위(즉 분위 당 3,796,781명씩)로 나눈다. 상위 10%의 경우는 이를 다시 상위 1%, 0.1%, 0.01% 등으로 세분하고 거기에 대응하는 인원수를 구한다. 이 때 각 분위의 경계 인원수를 포함하고 있는 소득구간을 대상으로 해당 인원수를 경계로 인원수와 소득금액을 나눌 필요가 있는데, 그를 위한 보간은 전술한 mean split histogram의 방법에 따랐다. 그 결과를 <표 7>에 '성인 인구 기준'으로 제시하였는데, 각 분위의 소득비중뿐만 아니라 경계소득(해당 분위에 들어가기 위한 최소 소득)과 평균소득, 나아가 전체 평균소득 대비 배율도 함께 제시하였다.

그런데 20세 이상 성인 인구를 통계청의 경제활동인구조사(2010년)에 따라 나누면, 취업자 23,625(천명), 실업자 892(천명)이며, 이들을 성인인구에서 뺀 비 경제활동인구는 13,451(천명)이 된다⁴⁵. 다만 취업자가 아니더라도 소득이 반드시 0이 될 필요는 없다. 재산소득만을 가지고 있는 경우도 있을 것이고, 전술했듯이 1년 중 일시적으로 근로소득이나 사업소득을 얻은 실적이 있지만 취업자로 파악되지 않는 경우가 적지 않게 존재하기 때문이다. 이로부터 취업자와 취업자가 아니더라도 재산소득을 갖고 있는 자는 대부분 상위 70%(4-10분위, 26,577천명)에 들어간 것으로

⁴³ 김낙년(2012a: 99-100)은 가계조사의 사업소득이 국민계정이나 『경제총조사』의 해당 통계에 비추어 볼 때 과대하게 파악되었다는 의문을 제기했는데, 그 이유가 사업소득자의 인원수가 20-90(백만원) 구간에서 과대하게 파악된 데 있음이 드러났다.

⁴⁴ 여기서는 가계조사에서 근로소득과 사업소득의 문제를 보여준 것이지만, 금융소득에서 가계조사는 전체소득의 5% 정도밖에 파악하지 못하는 등 문제가 더욱 심각하다(김낙년·김종일 2013).

⁴⁵ 경제활동인구조사의 20세 이상 인구(37,253천명)는 본고에서 이용한 통계청 공식 통계(추계인구)의 20세 이상 인구(37,968천명)에 비해 적는데, 이들은 재소자와 군인으로 추정된다. 그 차이를 여기서는 비 경제활동인구로 넣었다.

생각된다. 하위 20%(1-2분위, 7,594천명)는 소득이 없는 비 경제활동인구(실업자 포함)로, 그 사이의 10%(3분위, 3,797천명)는 취업자로 파악되지는 않지만 소액의 소득이 있는 자들로 구성되어 있는 것으로 생각된다. 이와 같이 성인 인구 기준의 분위별 소득비중을 구할 때 기준이 되는 전체 인원수에는 소득이 없거나 미미한 비 경제활동인구가 다수 포함되어 있다는 점에 유의할 필요가 있다.

<표 7> 각 분위의 경계소득과 소득 비중 (2010년)

소득분위	성인 인구 기준					(참고) 취업자 기준				
	인원수 (명)	경계소득 (천원)	평균소득 (천원)	소득비중 (%)	평균소득 대비 배율	인원수 (명)	경계소득 (천원)	평균소득 (천원)	소득비중 (%)	평균소득 대비 배율
(1분위 미만)						7,390,454	-	1,307	1.51	0.05
1분위	3,796,781		-	-	-	2,382,900	3,078	3,915	1.46	0.15
2분위	3,796,781	-	0	0.00	0.00	2,382,900	4,909	5,929	2.21	0.22
3분위	3,796,781	207	814	0.48	0.05	2,382,900	6,937	7,961	2.97	0.30
4분위	3,796,781	1,637	2,640	1.57	0.16	2,382,900	9,198	10,913	4.07	0.41
5분위	3,796,781	3,762	5,391	3.20	0.32	2,382,900	12,609	14,277	5.33	0.54
6분위	3,796,781	7,004	8,892	5.28	0.53	2,382,900	15,945	18,534	6.91	0.70
7분위	3,796,781	11,351	14,024	8.34	0.83	2,382,900	21,460	24,414	9.11	0.92
8분위	3,796,781	16,847	21,557	12.81	1.28	2,382,900	27,433	32,548	12.14	1.23
9분위	3,796,781	26,264	34,056	20.24	2.02	2,382,900	37,897	46,511	17.35	1.76
10분위	3,796,781	44,326	80,851	48.05	4.81	2,382,900	57,657	99,020	36.94	3.75
상위 1%	379,678	116,213	218,219	12.97	12.97	238,290	132,561	273,894	10.22	10.37
상위 0.1%	37,968	317,676	750,963	4.46	44.63	23,829	393,442	986,047	3.68	37.35
상위 0.01%	3,797	1,201,489	2,919,691	1.74	173.54	2,383	1,581,112	3,860,288	1.44	146.21
상위 10-5%	1,898,391	44,326	52,911	15.72	3.14	1,191,450	57,657	66,064	12.32	2.50
상위 5-1%	1,518,713	64,027	81,434	19.36	4.84	953,160	76,761	96,497	14.40	3.65
상위 1-0.5%	189,839	116,213	127,188	3.78	7.56	119,145	132,561	150,656	2.81	5.71
상위 0.5-0.1%	151,871	138,163	198,822	4.73	11.82	95,316	188,853	249,904	3.73	9.47
상위 0.1-0.05%	18,984	317,676	372,776	1.11	22.16	11,915	393,442	487,111	0.91	18.45
상위 0.05-0.01%	15,187	464,337	681,516	1.62	40.51	9,532	612,755	891,156	1.33	33.75

주: 1) 여기서 취업자 기준이란 전체 소득자 중에서 소득 순으로 취업자 수만큼을 취하고 이를 대상으로 분위별 소득비중을 구한 것이다. '1분위 미만'이란 거기에서 빠진 소득자를 말한다.

2) 20세 이상 성인 인구의 평균소득은 16,825 천원, 취업자 기준의 평균소득은 26,402 천원이다.

자료: <표 6>.

상위 소득자의 소득비중을 국제 비교할 때에는 성인 인구 기준을 이용한다. 다른 나라 통계도 이 기준에 맞추고 작성되고 있기 때문이다. 다만 소득이 없거나 미미한 비 경제활동인구까지 포함하여 상위 소득자의 소득비중이 구할 경우 <표 7>에서 보듯이 하위 20-30%의 소득비중은 0에 접근하게 되고, 그만큼 상위 소득자의 소득비중은 더 커지게 된다. 문제는 소득비중의 분모에 들어가는 전체 소득자 수(control total)를 정의하기가 쉽지 않다는 것이다. 취업자 수로 한정해 볼 수 있지만, 그 경우 재산소득만 가지고 있는 소득자들이 배제된다. 전술했듯이 1년 중 일시적인 취업으로 소득을 얻은 경우도 많아 취업자와 비 취업자의 경계도 모호하다. 국제비교에 성인 인구 기준을 이용하는 것은 이와 같이 전체 소득자 수를 정의하기 어렵기 때문이다.

<표 7>에는 참조 자료로서 취업자 기준의 소득비중을 함께 제시하였다. 여기서 취업자 기준이

란 전체 소득자 중에서 소득 순으로 취업자 수(23,829 천명)만큼을 취하고 이를 대상으로 분위별 소득비중을 구한 것이다. 따라서 취업자가 아니더라도 재산소득만을 가진 자가 여기에 포함될 수도 있고, 취업자이지만 소득이 낮아 여기에서 배제된 자도 존재한다. 이 기준으로 보면 연소득이 308 만원(1분위의 경계소득) 미만인 자들이 빠지게 된다. 이들은 표에서 '1분위 미만'으로 잡혀 있는데, 739만명이며 그들의 평균소득은 131만원이 된다. 이들은 대부분 비 경제활동인구에 해당하며 소액의 재산소득이 있거나 1년 중 일시적으로 근로소득이나 사업소득을 얻은 자들로 볼 수 있다.

표에 따르면 상위 1%와 10%의 소득비중은 성인 기준과 취업자 기준에서 각각 12.97%→10.22%와 48.05%→36.94%로 낮아진다. 8분위 이상에서는 이처럼 소득비중이 낮아지지만 7분위 이하에서는 거꾸로 소득비중이 높아짐을 알 수 있다. 상위 1%와 10%에 들어가기 위한 경계소득도 각각 119,213(천원)→132,561(천원)과 44,326(천원)→57,657(천원)으로 높아졌다. 그리고 이 취업자 기준의 수치는 근로소득의 상위 소득비중(김낙년 2012b: 부표 1, 2)과 비교될 수 있다. 이 연구는 취업자 중 근로자를 대상으로 근로소득의 상위 소득비중을 구한 것인데, 상위 0.1%와 1%와 10%는 각각 2.16%와 7.45%와 34.18%로 나온다. 근로소득보다 전체 소득의 소득집중도가 높고 최상위로 갈수록 그 차이가 더 벌어지는 것은 금융소득과 사업소득에서 최상위로의 집중이 더욱 높기 때문이다.

그런데 <표 7>의 추계 결과는 기존 연구와 비교하여 어떤 차이가 있을까? 이를 위해 <표 8>은 Kim and Kim(2014)과 본고가 추계한 상위 소득자의 소득비중을 비교하였다. 그와 함께 국세청의 통합소득 자료를 이용한 소득비중도 추계하여 제시하였다. 앞에서 이용한 국세청의 통합소득 자료는 종합소득 신고자료와 근로소득 연말정산 중에서 납세자 자료를 통합한 것이다. 이를 전술한 방법으로 상한과 하한을 추정하였으며, 소득구간을 분할하여 소득비중을 구할 때 mean split histogram의 방법으로 접근하였다. 이에 대해 Kim and Kim(2014)은 동일한 자료를 이용하였지만, 두 자료를 통합할 때 중복되는 부분을 추정하여 배제하는 추가적인 절차를 밟았다. 그리고 소득구간을 분할할 때에는 Pareto 보간법을 이용하였다. 이러한 방법상의 차이에도 불구하고 두 결과를 비교하면 차이가 미미한 것을 알 수 있다. 보간법으로서 이용된 두 방법(mean split histogram과 Pareto interpolation)이 추정결과에서 차이가 거의 없다는 점을 감안하면, Kim and Kim(2014)이 두 자료의 중복을 배제하기 위해 이용된 추정방법이 실태를 거의 왜곡하지 않았음을 알 수 있다.

그런데 본고에서 추계된 소득비중은 Kim and Kim(2014)에 비해 상당히 높아졌다. 예컨대 상위 10%(또는 1%)에 들기 위한 경계소득은 36,662천원→44,326천원(또는 106,193천원→116,213천원)으로, 소득비중은 43.30%→48.05%(또는 11.76%→12.97%)로 높아졌다. 이 차이는 어디에 기인한 것일까? 그것은 <표 6>에서 알 수 있듯이 본고에서 분위별 소득비중 추계에 이용한 '통합소득 3'에는 4,000만원 미만의 금융소득을 비롯하여 '통합소득 1'에는 제외되어 있는 소득 항목들이 포함되어 있기 때문이다. 따라서 본고의 결과가 기존 연구가 감안하지 못한 소득까지 포괄하고 있다는 점에서 실태에 더 근접한 것으로 볼 수 있다.

<표 8> 추계 결과의 비교: 분위별 경계소득과 소득비중(2010년)

	Kim and Kim(2014)		국세청 통합소득 자료		본고	
	경계소득	소득비중	경계소득	소득비중	경계소득	소득비중
	(천원)	(%)	(천원)	(%)	(천원)	(%)
상위 10%	36,662	43.30	38,618	43.02	44,326	48.05
상위 5%	57,666	29.17	57,081	29.09	64,027	32.33
상위 1%	106,193	11.76	102,038	11.77	116,213	12.97
상위 0.5%	137,340	8.36	131,687	8.37	138,163	9.19
상위 0.1%	288,186	4.19	293,137	4.17	317,676	4.46
상위 0.05%	429,935	3.16	432,553	3.15	464,337	3.36
상위 0.01%	1,121,247	1.65	1,149,255	1.62	1,201,489	1.74

자료: Kim and Kim(2014); 국세청, 통합소득; <표 7>.

V. 맺음말

소득세 자료를 이용한 기존의 소득분배 연구는 상위 10%의 소득자에 초점을 맞추었다. 장기추이를 보이게 할 경우 과거로 거슬러올라가면 소득세 자료가 포착하는 납세자의 범위가 좁아지기 때문이다. 이에 대해 본고는 최근의 한 시점(2010년)에 초점을 맞추는 대신에 상위 10%뿐만 아니라 하위 90%를 세분하여 소득이 어떻게 분포되어 있는지를 보이게 하였다. 최근의 자료는 소득세법상 포착되는 모든 소득을 면세자 정보까지 포함하여 제시하고 있다. 나아가 거기에서 빠진 미신고 사업소득과 비과세 소득(재배업 소득), 그리고 4,000만원 미만의 금융소득도 다른 자료를 이용하여 보완할 수 있었다. 이를 통해 파악된 전체 개인소득이 국민계정에서 파악된 개인소득에 근접함을 확인하였다.

소득세 자료에서는 개인이 복수의 소득원을 가질 경우 이를 통합할 필요가 있다. 국세청의 통합소득 자료를 이용하여 이에 어느 정도 대처할 수 있었지만, 거기에서 빠져 있는 각종 원천징수 소득 항목의 경우 특히 하위 소득구간에서 이러한 문제가 적지 않을 것으로 생각된다. 여기서는 나름의 합리적인 가정을 도입하여 이를 통합하였지만, 그에 따른 오차가 하위 소득자를 중심으로 어느 정도 있을 수 있다는 점에도 유의할 필요가 있다.

본고에서 추정된 개인소득자 수는 3,122만명에 달하며, 구간별 인원수 분포를 보면 1,000만원 미만, 1,000-4,000만원, 4,000만원-1억원, 1억 이상의 구간이 각각 48.4%, 37.4%, 12.4%, 1.8%로 나온다. 1,000만원 미만의 소득자 중에서 2/3가 500만원 미만이며, 100만원 미만도 적지 않다. 전체 평균소득이 2,046만원이지만, 중위소득은 1,074만원(평균소득의 52.5%)에 불과한 것으로 나온다. 이 수치는 취업자 수(2,383만명)보다 훨씬 많은데, 그 차이는 재산소득만을 가진 경우도 있지만, 1년 중 평소에는 비 경제활동인구로 있으면서 일시적인 취업으로 수입을 얻는 경우가 상당히 많은 것으로 설명된다. 취업자라 하더라도 그들의 소득의 하한은 308만원에 불과하여 취업자와 비 경제활동인구의 경계가 모호함을 알 수 있다. 취업자수 기준으로 한정해서 보면, 4개 구간의 인원수 분포는 각각 32.3%, 49.0%, 16.3%, 2.4%로 나오고, 이들의 평균소득과 중위소득은 각각 2,640만원과 1,594만원(평균소득의 60.4%)으로 다소 높아진다. 그렇지만 양자의 갭이 여전히 크게 벌어져

있어 개인소득의 분포는 하위 소득구간에 몰려 있음을 알 수 있다.

본고가 추정한 개인소득의 분포는 가계조사 결과를 검증하는데 이용될 수 있다. 종래 가계조사가 상위 소득자를 과소 파악했다는 문제가 제기되었지만, 그뿐만 아니라 중위 소득자를 과대하게 파악하고 하위 소득자를 과소하게 파악한 실태가 드러났다. 가계조사의 이 문제는 근로소득자의 경우도 그러하지만, 사업소득자의 경우 특히 심각하였다. 본고의 근로 및 사업 소득자를 기준으로 가계조사가 파악한 소득자 분포를 보면 10-60(백만원)의 5개의 중위 소득구간에서 인원수가 8.7%에서 45.0% 더 많고, 하위와 상위에서는 거기에 못 미치는 역 U자형의 양상을 보인다. 가계조사에 의거한 통계청의 불평등 지표인 지니계수가 왜 과소하게 나타났는지를 알 수 있다.

소득구간별 개인소득 분포를 이용하여 분위별 소득비중을 추정하였다. 그 결과는 기존연구(Kim and Kim 2014)에 비해 예컨대 상위 1%와 상위 10%의 소득비중이 각각 11.76%→12.97%와 43.30%→48.05%로 올라가는 등 모든 그룹에서 소득비중이 더 높아졌다. 그것은 기존연구가 이용한 종합소득 신고자료와 근로소득 연말정산(납세자) 자료에는 4,000만원 미만 금융소득을 비롯하여 거기에서 누락된 소득을 보완하여 추가했기 때문이다. 그런데 이상은 성인 인구를 기준으로 한 것이므로 하위 구간에 소득이 없거나 미미한 비 경제활동인구까지 포함하게 되며 그만큼 상위 소득비중이 높게 나타난다. 여기서는 소득비중을 계산하는 기준인 전체 인원수(control total)를 성인 인구 대신에 취업자수로 바꾸어 추정한 결과도 함께 제시하였다. 그에 따르면 예컨대 상위 1%와 10%의 소득비중이 각각 10.22%와 36.94%로 낮아지는 등 상위 30%의 소득비중이 낮아진 대신에 하위 70%의 소득비중은 높아졌다. 이것은 또한 근로소득자에 한정된 소득비중(상위 1%와 10%는 각각 7.45%와 34.18%)과 비교될 수 있는데, 그것이 더 높은 것은 근로소득보다 금융소득이나 사업소득이 상층으로 더욱 집중되어 있기 때문이다.

마지막으로 본고의 한계와 남겨진 앞으로의 과제를 간단히 언급해두고자 한다. 본고는 개인소득에 초점을 맞추고 있지만, 그것으로 가구소득의 실태를 알기 어렵다는 점이다. 나아가 소득세 자료에는 공적 및 사적 이전소득에 관한 정보가 없어 가처분소득 기준의 소득분포가 어떠한지 드러내기 어렵다. 이 분야는 본고에서 그 문제점을 지적한 가계조사가 소득세 자료에 대해 비교우위를 가지고 있다. 따라서 두 자료의 장점을 살려 보완하는 접근이 필요하다.

본고는 2010년의 결과를 보여주는데 그쳤는데, 이를 다른 연도(현재의 자료상황으로는 2009년 이후)로 확장할 수 있다. 이를 통해 비록 기간을 짧지만 개인소득 분포의 추이를 보일 수 있을 것이다. 그리고 본고는 소득세 자료 중에서 소득에 초점을 맞추었는데, 그에 대응하는 세액 자료를 활용할 수 있다. 이를 통해 소득구간별 또는 분위별로 소득세 부담률이 얼마나 다르고 또 어떻게 추이했는지를 보일 수 있다. 이들은 앞으로의 과제이다.

<참고 문헌>

국세청(2010), 『국세통계연보』.

국세청, "통합소득 100분위 자료"(홍중학 의원실).

국세청, "금융소득 100분위 자료"(최재성 의원실).

통계청(2010), 『가계금융조사』.
통계청(2010), 『가계동향조사』.
통계청(2010), 『경제총조사』.
통계청(2010), 『경제활동인구조사』.
통계청 KOSIS (http://kosis.kr/abroad/abroad_01List.jsp)
통계청 마이크로데이터 서비스 시스템 (<http://mdss.kostat.go.kr/mdssex>)
한국은행 ECOS (<http://ecos.bok.or.kr>)

김낙년(2012a), 「한국의 소득집중도 추이와 국제비교, 1976-2010: 소득세 자료에 의한 접근」, 『경제분석』 18(3), 75-114.
김낙년(2012b), 「한국의 소득불평등, 1963-2010: 근로소득을 중심으로」, 『경제발전연구』 18(2), 125-158.
김낙년(2014), 「2013년 소득세제 개편과 계층별 소득세 부담률」, 『재정학연구』 7(2), 59-93.
김낙년·김종일(2013), 「한국 소득분배 지표의 재검토」, 『한국경제의 분석』 제19권 제2호, 1-64.
김상조(2014), 「국세청의 통합소득 자료를 이용한 소득분배 및 실효세율 추이 분석: 모집단 기준 전환 100분위 자료를 기초로」, 『경제개혁리포트』 2014-08호, 1-35.
김우철(2014. 10), 「김낙년 교수의 추계에는 문제가 없는가?」, 『Economy』 21, 45-51.

Alvaredo, Facundo, Anthony B. Atkinson, Thomas Piketty and Emmanuel Saez, The World Top Incomes Database (<http://topincomes.g-mond.parisschoolofeconomics.eu>)
Atkinson, Anthony B.(2005), "Top Incomes in the UK over the 20th Century," *Journal of the Royal Statistical Society*, 168(2), 325-343.
Atkinson, Anthony B., Thomas Piketty, and Emmanuel Saez(2009), "Top Incomes in the Long Run of History," NBER Working Paper No. 15408.
Kim, Nak Nyeon and Jongil Kim(2014), "Top Incomes in Korea, 1933-2010: Evidence from Income Tax Statistics," WTID Working Paper 2014-2. Forthcoming in *Hitotsubashi Journal of Economics*.
Moriguchi, Chiaki, and Emmanuel Saez(2008), "The Evolution of Income Concentration in Japan, 1886-2005: Evidence from Income Tax Statistics," *The Review of Economics and Statistics*, 90(4), 713-734.
Piketty, Thomas(2003), "Income Inequality in France, 1901-1998," *Journal of Political Economy*, 111(5), 1004-1042.
Piketty, Thomas, and Emmanuel Saez(2003), "Income Inequality in the United States, 1913-1998," *Quarterly Journal of Economics*, 118(1), 1-39.