

## 연구논문

## 소득분배의 시계열 분석을 위한 한국 소득 데이터의 검토

An Examination of Survey Micro-data in Korea:  
Analysing Trends in Income Distribution

이원진<sup>a)</sup> · 구인회<sup>b)</sup>  
Wonjin Lee · Inhoe Ku

최근 한국의 소득불평등과 빈곤 추이에 대한 관심이 커지고 있지만, 실증 연구에 활용되는 소득분배 데이터의 질에 대해서는 충분한 검토가 이루어지지 않고 있다. 본 연구는 한국 소득분배 추이를 분석할 때 활용되는 여러 데이터를 비교·검토하였다. 검토 결과, 가계동향조사와 비교할 때 1991/1996년 가구소비실태조사와 2005년 가계자산조사는 주로 저소득층의 상대소득을 높게 포착하는 반면, 2000년 가구소비실태조사와 가계금융조사/가계금융·복지조사는 저소득층의 상대소득은 낮게, 고소득층의 상대소득은 높게 포착하는 것으로 나타났다. 이는 가계부기장방식의 소득조사가 무응답양상에 일정한 영향을 미치기 때문일 가능성이 있다. 이러한 분석에 따르면, 가계부기장방식보다는 면접조사방식으로 소득관련 정보를 조사하는 마이크로데이터를 발전시키는 것이 무응답 편의를 줄이는 방안이 될 수 있다. 한편 연구자들은 소득분배의 시계열 분석 시 시계열 일관성을 고려하여 데이터를 선택할 필요가 있다.

**주제어:** 소득분배, 불평등, 빈곤, 마이크로데이터

\* 이 논문은 2014년도 서울대학교 미래 기초학문 분야 기반조성 사업의 지원을 받아 수행된 연구 결과물임.

a) 교신저자(corresponding author): 서울대학교 사회복지학과 박사과정수료 이원진.  
E-mail: lwj4035@naver.com

b) 서울대학교 사회복지학과, 사회과학연구원 교수

Despite a growing interest in income inequality and poverty trends in Korea, the quality of income data for empirical analyses has not been fully scrutinized yet. The present study compared several micro-level data sets containing household income information and investigated their characteristics. The examination reveals that the relative income level of low income group was reported to be higher in the 1991/1996 National Survey of Family Income and Expenditure and the 2005 Survey of Household Wealth than in the Household Income and Expenditure Survey data. On the other hand, the relative income level of low income group was lower and that of high income group was higher in the 2000 National Survey of Family Income and Expenditure and the Survey of Household Finances/Survey of Household Finances and Living Conditions data than in the Household Income and Expenditure Survey data. These results might be attributed to the differential nonresponse rates due to methods collecting information on income. According to the results, collecting income information through interview seems more promising than diary keeping method in reducing nonresponse bias. Also, researchers should make deliberate choice of data sets and check time-series consistency in analysing income distribution trends.

**Key words:** income distribution, inequality, poverty, micro-data

## I. 서론

1990년대 말 외환위기 이후 한국의 소득불평등과 빈곤에 대한 관심이 커지고 있고, 이와 관련된 실증 연구도 활발하게 이루어지고 있다. 기존 연구들은

시계열 측면에서 1990년대 중반 이후 최근까지 주로 노동시장에서의 불평등 증대에 따라 한국의 소득분배가 상당히 악화되었고(장지연·이병희 2013; 전병유 2013; 이병희 2014 등), 국제비교 측면에서 한국의 조세와 공적 이전의 재분배효과가 매우 낮은 수준이라는 점을 밝혀내는 등(유경준 2003; 여유진·김태완 2006 등)의 성과를 거두었다. 이러한 기존 연구의 성과에는 1990년대 이후 한국의 소득분배를 분석할 수 있는 여러 데이터가 축적된 것이 중요한 밑바탕이 되었다. 하지만 여러 데이터가 실제 소득분포를 얼마나 잘 반영하는지, 데이터 간의 차이는 어떻게 설명될 수 있는지 등 데이터 자체에 대한 검토는 아직 충분히 이루어지지 못한 것으로 판단된다.

지금까지 몇몇 연구들이 한국에서의 소득분배 데이터의 질을 검토하였다. 강석훈(2000)은 1996년 가구소비실태조사를 국민계정과 비교하였고, 강석훈·현진권(2003)과 현진권·임병인(2004)은 도시가계조사, 농가경제조사, 가구소비실태조사 등의 특징을 검토하고 소득분배 측정을 위한 데이터의 개선방안을 논의하였으며, 강영욱(2011)은 도시가계조사의 특징과 한계를 집중적으로 검토하였다. 가장 최근에는 김낙년·김종일(2013)이 가구소비실태조사와 가계동향조사를 소득세 자료와 국민계정과 비교하여 고소득층 누락 등의 문제를 보정한 소득분배지표를 추정하였다. 본 연구는 통계청이 조사하여 일반 이용자에게 제공하는 가계동향조사 및 농가경제조사, 가구소비실태조사, 가계자산조사/가계금융조사/가계금융·복지조사 데이터를 함께 검토함으로써 기존 연구의 성과를 확장하고자 한다.

소득분배와 관련된 데이터의 범위는 매우 넓지만, 지면의 제약과 연구의 초점을 고려하여 본 연구는 다음과 같이 범위를 제한한다. 첫째, 소득분배 관련 마이크로데이터가 활발하게 생산되기 시작한 1990년대 이후 시기만을 대상으로 한다. 1980년대까지의 한국 소득분배 추이에 대해서도 중요한 논쟁이 존재하지만(안국신 1995; 이정우·황성현 1998; 이재욱 2000), 본 연구에서는 다루지 않는다. 둘째, 표본조사를 통해 수집된 소득분배 관련 마이크로데이터에 초점을 맞춘다. 최근 소득세 자료를 분석하여 한국 소득분배에 관한 새로운 시각을 제시하는 중요한 연구 성과가 제출되고 있다(김낙년 2012a, 2012b). 하지만 미시단위로 측정된 다양한 개인 및 가구 정보를 활용하는 심층분석을

위해서는 여전히 표본에 대한 단일 서베이로 수집된 마이크로데이터의 필요성이 크다. 본 연구는 가구단위의 소득정보를 포함하는 한국의 여러 마이크로데이터를 이용하여 1990년대 이후 소득분배를 시계열 분석할 때 부딪히는 여러 쟁점을 소득분배 데이터의 질을 중심으로 다룬다. 또한 시계열 분석 시에 경험연구자가 유념해야 할 각 데이터의 특성을 비교·검토하고, 일관된 시계열 분석을 위한 데이터 선택 시의 고려사항을 제시한다.

## II. 한국 소득분배 마이크로데이터의 특성 비교

본 장에서는 가계동향조사와 농가경제조사, 가구소비실태조사, 가계자산조사/가계금융조사/가계금융·복지조사 데이터의 특징을 살펴본다. 가구의 소득과 지출을 조사하는 대표적인 데이터인 가계동향조사는 과거 동지역의 비농어를 조사하는 도시가계조사에서 2003년 읍면지역 비농어를 포함하는 가계조사로 명칭을 변경하였고, 2006년에 1인가구를 포함한 후 2008년에 가계동향조사로 명칭을 변경하여 현재까지 이어져 오고 있다(통계청 2011: 4). 가계동향조사에서 조사되지 않는 농가의 소득을 파악하기 위해서는 농가경제조사가 이루어지고 있다. 통계청은 가계동향조사와 농가경제조사 데이터를 합쳐 공식적인 소득분배지표를 계산하여 공표한 후, 이에 활용된 데이터를 ‘소득분배지표 마이크로데이터’라는 이름으로 일반 이용자에게 제공하고 있다.

2000년대 중반 이전 가계동향조사의 표본 대표성이 제한적이었던 시기에는 전국을 대상으로 소득과 소비를 조사하기 위해 1991년, 1996년, 2001년 세 차례에 걸쳐 가구소비실태조사가 이루어졌다. 그리고 2006년에는 가계자산조사가 실시되었고, 2010년에는 통계청의 가계자산조사, 금융감독원의 가계신용조사, 한국은행의 한은패널조사를 통합하여 제1회 가계금융조사가 실시되었다. 가계금융조사는 2011년 2회 조사까지 이루어진 후 2012년부터는 표본규모와 조사내용을 확대한 가계금융·복지조사로 이어져 오고 있다. 가계자산조사/가계금융조사/가계금융·복지조사는 기본적으로 자산 실태 파악이 주목적이지만, 소득도 함께 조사하여 소득분배 데이터로서도 가치가 있다.

<표 1> 표본 대표성

데이터	조사대상
가구소비 실태조사	1991 - 전국 전 가구
	1996 - 전국 농가를 제외한 전 가구
	2001 - 전국 전 가구 * 조사에서는 농어를 포함했으나 집계 시 농어를 제외했다. 일반에게 공개되는 마이크로데이터에는 주된 활동이 농림축어업경영주인 사례들이 제외됐다.
가계동향조사	1990~2002 - 도시지역 2인 이상 비농가
	2003~2005 - 전국 2인 이상 비농가
	2006~2013 - 전국 1인 이상 비농가
농가경제조사	2006~2012 - 전국 2인 이상 농가
	2013 - 전국 1인 이상 농가
소득분배지표 마이크로데이터	1990~2013 - 가계동향조사+농가경제조사
가계자산조사	- 전국 전 가구
가계금융조사	- 전국 전 가구
가계금융·복지조사	- 전국 전 가구

자료: 통계청 통계설명자료(<http://meta.narastat.kr>, 2015년 3월 2일 접속), 통계청 마이크로데이터 서비스시스템의 각 데이터 설명자료(<http://mdss.kostat.go.kr/mdssex/DataProcessing/extraction/OfferRange.jsp?menu=1>, 2015년 3월 2일 접속)로부터 확인하여 정리하였다. 이하 <표 2>, <표 3>, <표 4> 모두 동일하다.

주: 가계동향조사는 1990년 전, 농가경제조사는 2006년 전 데이터도 존재하지만, 본 연구에서는 통계청의 소득분배지표 작성에 활용된 연도만을 검토한다.

<표 1>~<표 4>에서는 각 데이터의 특성을 몇 가지 측면으로 나누어 정리하였다. 먼저 <표 1>에서는 각 데이터의 표본 대표성을 요약하였다. 가구소비실태조사를 통해서는 기본적으로 전국 농가를<sup>1)</sup> 제외한 전 가구에 대한 소득

1) 임가와 어가는 전체 가구의 약 1%에 불과하여(통계청 국가통계포털 농림어업총조사 및 인구총조사 집계자료를 통해 필자가 계산한 수치, 2005년, 2010년 기준) 통계청의 각 데이터 설명자료에서 ‘농가’라는 표현은 임가와 어가를 포함하는 경우가 많고, 본 연구에서도 대체로 이러한 용법을 따른다. 단, 농가경제조사는 농가

정보를 파악할 수 있다. 1996년에는 조사대상에서 농가가 제외되었고, 2001년에는 농어가 조사에 포함됐으나 이용자에게 공개되는 마이크로데이터에는 주된 활동이 농림축어업경영주인 사례들이 제거되어 있어 농어를 제외한 데이터로 간주할 수 있다. 1991년의 경우에는 전 가구를 조사대상으로 하였고 이용자에게 공개되는 마이크로데이터에도 농업경영자 및 어업경영자 가구가 포함되어 있으나, 국가통계포털의 집계자료에는 해당 가구가 제외되어 있다.

다음으로 가계동향조사는 시간에 따라 조사대상이 크게 변화해 왔다. 1990~2002년에는 도시지역 2인 이상 비농가를 대상으로 하였으나 2003년에는 읍면지역을, 2006년에는 1인가구를 추가하여 1996, 2001년 가구소비실태조사와 유사하게 전국 1인 이상 비농가를 대상으로 하게 되었다.<sup>2)</sup> 농가경제조사는 2006~2012년에 전국 2인 이상 농가를 대상으로 하였으나 2013년부터는 1인 농가도 포함하였다. 소득분배지표 마이크로데이터는 1990~2005년에는 가계동향조사와 동일한 대상을 포괄하고, 2006년 이후에는 가계동향조사와 농가경제조사의 대상을 함께 포괄한다. 그런데 2006~2012년 농가경제조사에서 1인 농가가 제외되었으므로, 엄격한 의미에서는 소득분배지표 마이크로데이터가 2013년이 되어서야 비로소 전체 가구를 포괄하게 되었다. 하지만 실제로는 1인 농가의 규모가 상당히 작기 때문에<sup>3)</sup> 통계청은 2006~2012년과 2013년 이후를 구분하지 않고 2006년 이후 소득분배지표 마이크로데이터가 전체 가구를 대표하는 것으로 간주한다. 본 연구 역시 불필요한 혼란을 줄이기 위해 이러한 해석을 받아들인다. 한편 가계자산조사/가계금융조사/가계금융·복지조사는 2006년 이후 소득분배지표 마이크로데이터와 유사하게 전체 가구를 대상으로 한다.<sup>4)</sup>

---

만을 조사하고, 어가경제조사와 임가경제조사가 별도로 실시되고 있지만 통계청의 소득분배지표 작성에는 활용되지 않는다는 점은 언급될 필요가 있다.

- 2) 통계청은 가계동향조사 설명자료에서 동지역을 ‘도시’로, 동지역과 읍면지역을 합쳐 ‘전국’으로 표현한다(통계청 2011: 6). 본 연구 역시 이러한 용법을 따른다.
- 3) 2013년 소득분배지표 마이크로데이터 분석 결과, 1인 농가에 속한 인구는 전체 인구의 0.33%이고, 1인 농가 포함 시와 제외 시 지니계수 차이는 약 0.001, 빈곤율 차이는 약 0.1%p로 나타났다.

<표 2> 소득조사방식

데이터	소득조사방식
1991	- 가계수지: 1991년 10월 1일~1991년 11월 30일 2개월간 수입과 지출을 가계부에 직접 기입, 수입은 근로자 가구에 한해 조사
	- 연간소득: 1990년 12월 1일~1991년 11월 30일 기간의 연간소득을 1991년 12월에 조사표로 조사
가구소비 실태조사	- 가계수지: 1996년 10월 1일~1996년 11월 30일 2개월간 수입과 지출을 가계부에 직접 기입, 수입은 근로자 가구에 한해 조사
	- 연간소득: 1995년 12월 1일~1996년 11월 30일 기간의 연간소득을 1996년 12월에 조사표로 조사
2001	- 2000년 1년 기간의 연간소득을 2001년 5월에 조사표로 조사
가계동향조사	- 매월 수입과 지출을 종이가계부 또는 전자가계부로 조사
농가경제조사	- 매월 수입과 지출을 일계부 조사표로 조사
소득분배지표 마이크로데이터	- 가계동향조사+농가경제조사
가계자산조사	2006 - 2005년 1년 기간의 연간소득을 2006년 6월에 조사표로 조사
가계금융조사	2010 - 2009년 1년 기간의 연간소득을 2010년 4~5월에 조사표로 조사
	2011 - 2010년 1년 기간의 연간소득을 2011년 4~5월에 조사표로 조사
가계금융· 복지조사	2012 - 2011년 1년 기간의 연간소득을 2012년 4~5월에 조사표로 조사
	2013 - 2012년 1년 기간의 연간소득을 2013년 4월에 조사표로 조사

주: 조사표 조사는 주로 면접조사로 이루어지며, 인터넷조사나 자기기입 방식이 병행되기도 한다.

4) 참고로 각각의 마이크로데이터에 포함된 가구 수는 가구소비실태조사 1991년 26,943가구(11월), 1996년 24,290가구, 2001년 23,720가구, 가계자산조사 8,275가구, 가계금융조사 1차년도 10,000가구, 가계금융·복지조사 1차년도 19,744가구이다. 가계동향조사와 농가경제조사는 연도별로 차이가 있으나 2006년 이후 각각 대략 10,000~12,000가구, 2,300~3,100가구이다.

<표 2>에 소득조사방식을 정리하였다. 1991년과 1996년 가구소비실태조사의 경우 각 연도 10월과 11월, 두 달동안의 가계수지를 가계부기장방식으로 조사하였고, 추가적으로 각 연도 12월에 지난 1년간의 연간소득을 조사표로 조사하였다. 이때 가계부기장방식으로 조사된 소득은 계절성의 영향을 받을 뿐만 아니라 근로자가구에 대해서만 조사했기 때문에 소득분배 분석에 활용되기 어렵고, 조사표로 조사된 연간소득을 활용하는 것이 타당하다. 2001년 가구소비실태조사와 가계자산조사/가계금융조사/가계금융·복지조사는 모두 조사시점을 기준으로 작년 1년간의 연간소득을 조사표로 조사했다. 따라서 소득년도가 조사년도와 다르다는 점에 주의해야 하고, 본 연구의 분석과 결론 부분(Ⅲ장과 Ⅳ장)에서는 소득년도를 기준으로 데이터를 지칭한다. 한편 가계동향조사는 매월 수입을 가계부기장방식으로 조사한 후 월단위로 소득을 집계하여 데이터를 생성한다. 따라서 1년을 준거기간으로 한 소득을 분석하기 위해서는 월소득 데이터를 연간화한 데이터를 사용해야 한다. 농가경제조사는 매월 수입과 지출을 일계부 조사표로 조사한 후 연간소득으로 집계하여 제공하고 있다.

<표 3>에서는 각 데이터의 소득구성을 정리하였다. 1991년 가구소비실태조사, 가계자산조사/가계금융조사는 소득원천별 변수를 공개하지 않기 때문에 소득원천별 분석이 불가능하다. 하지만 모든 데이터에 총소득 변수는 제공되기 때문에 기본적인 소득분배지표 분석은 가능하다. 가계동향조사는 총소득에서 경상조세와 사회보험료를 제한 가처분소득을 구성할 수 있고, 통계청은 농가경제조사 데이터에도 동일한 기준으로 가처분소득을 구성한 후 소득분배지표 마이크로데이터로 합쳐서 이용자에게 제공한다. 하지만 이용자에게 공개된 농가경제조사 데이터에는 경상조세와 비경상조세가 구분되지 않아 이용자가 직접 농가경제조사 데이터로부터 통계청 소득분배지표 기준 가처분소득을 구성할 수는 없다. 2001년 가구소비실태조사, 가계금융·복지조사도 경상조세와 비경상조세를 구분하지 않아 원칙적으로는 경상적 개념에 기초한 가처분소득을 구성하기 어렵다. 하지만 실제로는 비경상조세를 포함했을 때와 제외했을 때 가처분소득 지니계수 및 빈곤율의 차이가 미미하기 때문에,<sup>5)</sup> 2001년 가구

5) 2006~2013년 가계동향조사 데이터를 활용해 <총소득-경상조세-사회보험료>와 <총소득-경상조세-비경상조세-사회보험료>의 2가지 가처분소득을 분석한 결과, 모든

<표 3> 소득구성

데이터	소득구성	소득원천별 분석	총소득 분석	가처분소득 분석
가구 소비 실태 조사	(연간소득의 경우) - 조사: (세전)근로소득, 사업소득, 재산 및 이전소득 - 데이터: 각 소득원천의 합계로 이루어진 연간소득만 제공 * 비소비지출: 2 개월간 가계부기장방식 조사결과 제공	×	○	△
	(연간소득의 경우) - 조사: (세전)근로소득, 사업소득, 부업소득, 재산소득, 이전소득 - 데이터: 소득원천별 항목 제공 * 비소비지출: 2 개월간 가계부기장방식 조사결과 제공	○	○	△
	2001 - 조사: (세전)근로소득, 사업소득, 농림축어업소득, 재 산소득, 이전소득, 기타소득(비경상 성격), 비소비지출 - 데이터: 소득원천별 항목 제공	○	○	△
가계동향조사	- 조사: (세전)근로소득, 사업소득, 재산소득, 이전소득, 비경상소득, 비소비지출 - 데이터: 소득원천별 항목 제공 * 사회적 현물이전 제공	○	○	○
농가경제조사	- 조사: 농업소득, 겸업소득, 근로수입, 자본수입, 이전 소득, 비경상소득, 비소비지출 - 데이터: 소득원천별 항목 제공	○	○	△
소득분배지표 마이크로데이터	- 조사: 가계동향조사+농가경제조사 - 데이터: 근로소득, 사업소득, 재산소득, 이전소득, 공 적 비소비지출 등 소득원천별 항목 제공	○	○	○
가계자산조사	- 조사: 근로소득(세전 여부 불확실), 사업소득, 재산소 득, 이전소득, 비경상소득 - 데이터: 각 소득원천의 합계로 이루어진 경상소득, 비 경상소득만 제공	×	○	×
가계금융조사	- 조사: (세전)근로소득, 사업소득, 금융소득, 실물자산 소득, 이전소득, 비소비지출 - 데이터: 각 소득원천의 합계로 이루어진 경상소득, 비 소비지출만 제공	×	○	△
가계금융·복지 조사	- 조사: (세전)근로소득, 사업소득, 재산소득, 이전소득, 비소비지출 - 데이터: 소득원천별 항목 제공	○	○	△

주: 데이터에 따라 각 소득항목의 하위범주가 추가로 존재하지만, 편의상 일반적으로 분석에서 활  
용되는 근로소득/사업소득/재산소득/이전소득 수준의 범주까지만 제시했다.  
가계동향조사의 소득구성은 2009년 신분류체계 기준이다.

연도에서 지니계수 차이가 0.001 미만, 빈곤율 차이가 0.1%p 미만으로 나타났다.

소비실태조사와 가계금융·복지조사로 가처분소득을 구성하는 데 큰 무리는 없다. 가계금융조사의 경우에는 세금과 과징금, 공적연금 및 사회보험료, 가구 간 이전, 비영리단체 이전 등을 합한 비소비지출 변수만 제공하기 때문에 조세와 사회보험료 등의 공적 비소비지출만을 제한 가처분소득은 구성할 수 없고, 가계자산조사는 비소비지출을 아예 제공하지 않아 가처분소득 구성이 불가능하다. 1991년과 1996년 가구소비실태조사는 연간소득과 측정 방식은 다르지만 가계부기장방식으로 2개월간 조사된 비소비지출이 제공되기 때문에, 비소비지출의 계절성으로 인한 편의를 감수한다면 연간 총소득에서 2개월의 조세 및 사회보험료를 연단위로 환산한 공적 비소비지출을 제하여 가처분소득을 구성하는 것은 가능하다. 실제로 장지연·이병희(2013: 80)는 이런 방식으로 1996년 가구소비실태조사의 가처분소득을 계산한 바 있다.

<표 3>에서 살펴본 가계동향조사의 소득구성은 신분류체계를 기준으로 한 것임에 유의할 필요가 있다. 가계동향조사의 항목분류체계는 여러 차례에 걸쳐 개편되어 왔는데, 2009년에는 캔버라그룹(Canberra Group)의 권고안에 따라 소득분류체계를 개편한 후 개편된 체계를 “신분류”, 과거의 체계를 “구분류”로 부르고 있다(통계청 2011: 3). 그리고 통계청은 1990~2008년 구분류 데이터를 신분류체계에 맞추어 추계한 데이터로 소득분배지표를 분석하고 이용자에게 제공한다. 추계 작업은 여러 변수에 적용되었는데, 예를 들어 구분류 데이터의 기타 사회보장수혜 변수는 신분류체계에서 기초노령연금, 사회수혜금, 사회적 현물이전으로 나뉜다. 이처럼 2008년 이전의 가계동향조사 신분류 데이터의 일부 변수는 추계된 값이기 때문에 평균 분석에는 무리가 없더라도 해당 변수의 값이 0보다 큰 사례의 빈도를 분석하거나 분산 등 산포도를 분석하는 등의 경우에는 왜곡된 결과가 나타날 수 있다. 따라서 2008년 이전의 추계된 신분류 데이터를 분석할 때는 통계청이 제공하는 신분류 개편 연계표를 확인하고 구분류 데이터와 비교하는 등 주의를 기울일 필요가 있다. 본 연구에서 분석하는 가계동향조사 데이터는 신분류 데이터이다.

마지막으로 <표 4>에서는 각 데이터 조사설계의 성격을 정리하였다. 일반적으로 소득분배 추이를 분석하는 데는 조사시점마다 표본을 새로 표집하는 반복횡단조사와 동일한 표본에 대해 조사를 반복하는 패널조사가 모두 활용될 수 있다. 가구소비실태조사는 5년마다 조사가 이루어진 반복횡단조사에 해당한다. 가계동향조사와 농가경제조사 역시 큰 틀에서는 5년마다 표본이 새로 표집되는 반복횡단조사로 볼 수 있지만, 5년 동안에는 동일한 가구에 대해 조사를 지속적으로 실시한다는 점에서 패널조사의 성격을 갖는다. 그리고 가계동향조사의 경우 2004년 이전에는 5년마다 표본이 개편됐지만 2005년부터는 매년 전체 표본의 1/3을 교체하는 연동표본 방식으로 변경되고 동일한 가구에 대한 조사기간은 3년으로 줄었다. 가계금융조사와 가계금융·복지조사는 본격적인 패널조사로 설계되었다. 2010년 전국 10,000가구를 대상으로 제1회 가계금융조사가 실시된 후 2011년 제2회 조사에서는 적정 표본규모 유지를 위해 2,000가구를 추가하는 표본설계가 이루어졌고, 2012년 가계금융·복지조사에서는 20,000가구로 표본을 확대·개편한 후 현재까지 조사가 이루어지고 있다.

**<표 4> 조사설계 성격**

데이터	성격
가구소비실태조사	- 반복횡단조사 * 1991년, 1996년 가계수지의 경우 1가구당 2개월의 2레코드
가계동향조사	- 2004년 이전: 패널조사 성격, 5년마다 표본개편 - 2005년 이후: 연동표본방식, 한 가구에 대해 약 3년간 조사 실시, 매년 전체 표본의 1/3 씩 교체
농가경제조사	- 패널조사 성격, 5년마다 표본 개편
소득분배지표 마이크로데이터	- 가계동향조사+농가경제조사
가계자산조사	- 횡단조사
가계금융조사	- 패널조사
가계금융·복지조사	- 패널조사

한국 소득분배를 분석한 기존 연구에서는 시계열이 길고 소득항목이 자세하게 조사된 가계동향조사가 많이 활용되어 왔다(최근의 예로, 최바울 2013; 김대일 외 2014). 하지만 가계동향조사는 2005년 이전에 1인가구가 제외되고 2002년 이전에 비도시지역이 제외되는 등 표본 대표성의 한계가 크다는 단점이 있다. 이에 일부 연구들은 1990년대의 소득분배 추이를 분석하기 위해 가구소비실태조사를 활용해 왔고(예를 들어, 구인회 2004), 최근에는 소득과 자산을 함께 분석할 수 있는 가계금융·복지조사의 활용도가 높아지고 있다(예를 들어, 김진욱 2015).

### Ⅲ. 소득분배 시계열 분석의 비교·평가

#### 1. 소득의 정의와 분석방법

일반적으로 소득분배 분석에서는 실제 가구의 소비와 저축 능력을 나타내기 위해 총소득보다 조세나 사회보험료와 같은 비재량적 지출을 제외한 가처분소득(disposable income)이 선호된다(UNECE 2011: 17). 통계청 역시 공식적인 소득분배지표 분석 시 총소득에서 경상조세와 사회보험료로 구성된 공적 비소비지출을 제한 가처분소득을 활용한다(통계청 2011: 68~69). 하지만 본 연구의 경우 여러 데이터에 수집된 비소비지출 항목이 서로 달라 가처분소득을 동일하게 구성하기 어려워 일관된 비교가 가능한 총소득을 활용한다. 이때 경조소득이나 퇴직금 등의 비경상소득은 제외한다. 또한 가계동향조사의 경우 통계청의 방식을 따라 사회적 현물이전을 제외한다.

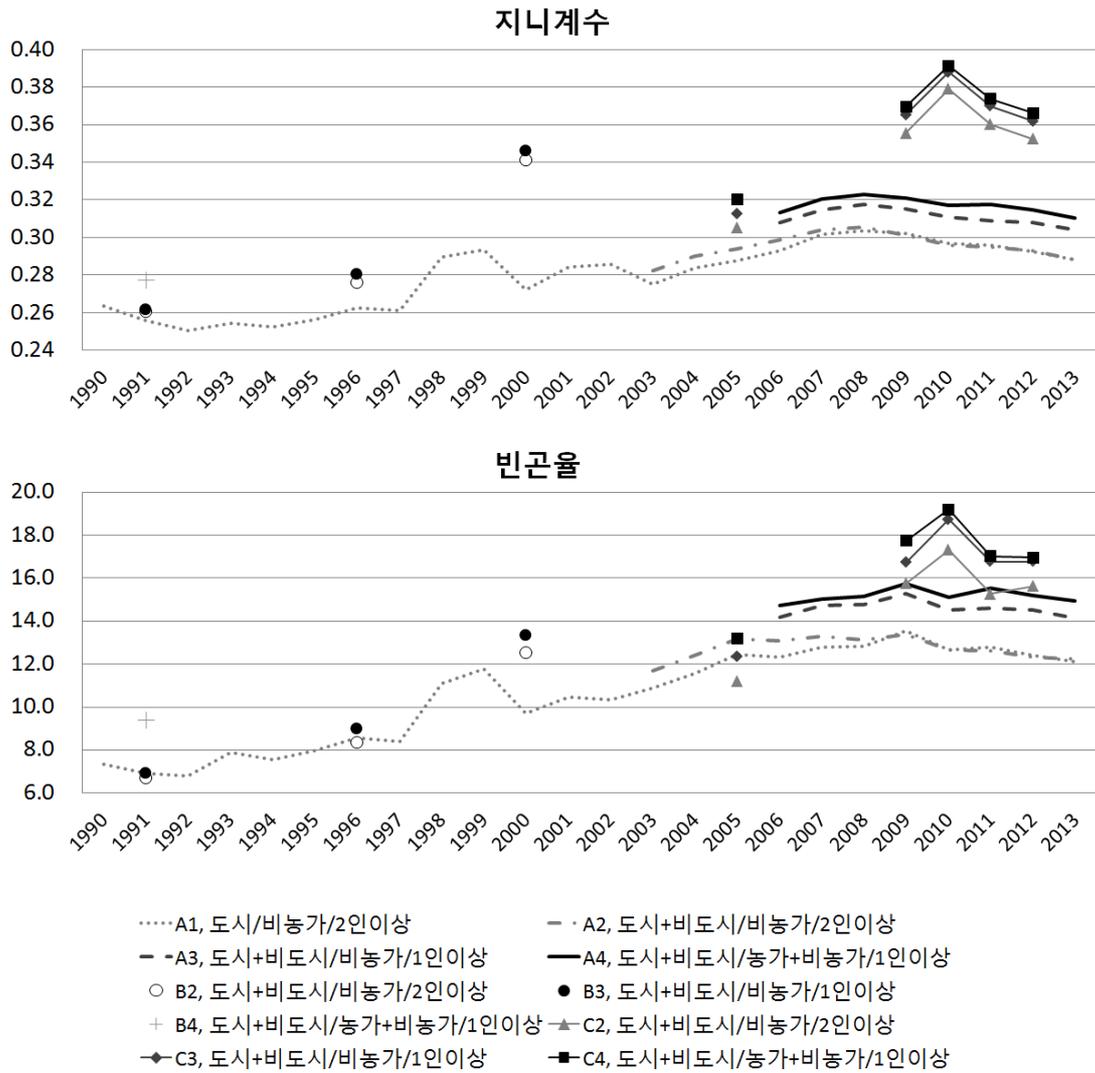
본 연구의 최종적인 관심의 대상은 전체 인구의 소득분포이기 때문에, 소득분배의 분석단위는 개인으로 설정한다. 가구단위로 측정된 소득을 개인단위로 분석하기 위해서는 가구 내 자원의 공유를 전제로 가구의 모든 구성원이 해당 가구의 소득을 균등화한 값을 동일하게 갖는다고 가정하며, 이때 균등화 소득은 가구 내 각 개인에게 주어진 경제적 자원의 수준을 나타내는 지표로 이해될 수 있다(UNECE 2011: 68). 요컨대, 본 연구는 1년을 준거기간으로 측

정된 가구소득을 가구원수 제곱근으로 나누어 균등화하여 개인단위의 소득분배를 분석한다. 이는 OECD 소득분포 데이터베이스(OECD Income Distribution Database)의 방식과 동일한 국제표준으로, 통계청의 소득분배지표도 동일하게 계산된다. 소득분배지표로는 지니계수와 중위소득 50% 기준 상대빈곤율을 활용한다. 지니계수는 0~1의 값을 갖고 불평등이 심할수록 값이 크다. 빈곤율은 소득이 빈곤선 미만인 인구 비율로 백분율로 표시하였다. 본 연구의 모든 분석에는 별도의 언급이 없는 한 데이터에 포함된 가중치가 적용된다. 본 연구는 최종적인 소득분배지표 산출의 기준이 되는 총소득의 음수를 0으로 치환하고, 0과 극단치는 그대로 분석에 포함한다.

## 2. 기초 분석

<그림 1>에는 여러 데이터로 분석한 지니계수와 빈곤율의 추이를 제시하였다. 먼저 소득분배지표 마이크로데이터 자료(A)를 살펴보자. 이 자료는 통계청이 발표하는 공식적인 소득분배지표와 가처분소득이 아닌 총소득 기준이라는 점에서 차이가 있고, 통계청의 소득분배지표와 거의 동일한 추이를 보여준다. 그중 가장 오랜 기간을 포괄하고 있는 도시지역 2인 이상 비농가 자료(A1)를 보면, 한국의 불평등과 빈곤은 1990년대 초중반부터 2000년대 중후반까지 증가하다가 2010년대에 들어 완만한 감소세로 돌아서고 있고, 1998~1999년에는 외환위기의 영향으로 일시적인 증가 양상을 보였다. 2006년부터는 가계동향조사의 조사대상이 전국 1인 이상 비농가로 확대되고 농가경제조사 데이터가 결합되면서 비로소 도시와 비도시, 농가와 비농가, 1인 이상 가구를 모두 포괄하는 소득분배지표 분석이 가능해졌다(A4). 이에 따르면 한국의 불평등과 빈곤은 2008~2009년 정점에 이른 후 최근까지 완만하게 감소하고 있는데, 지니계수에 비해 빈곤율의 감소세는 그다지 두드러지지 않는다.

이처럼 가계동향조사 및 농가경제조사는 한국의 소득분배 추이에 대해 기본적인 그림을 보여주지만, 이는 다른 데이터에 대한 분석결과와 함께 검토될 필요가 있다. B자료는 가구소비실태조사의 분석결과를 보여준다. 같은 시기의 가계동향조사 자료(A1)와 비교하기 위해서는 도시지역 2인 이상 비농가에 대



<그림 1> 총소득 기준 지니계수 및 빈곤율 추이

자료: A-통계청 소득분배지표 마이크로데이터(가계동향조사+농가경제조사), B-가구소비실태조사, C-가계자산조사/가계금융조사/가계금융·복지조사.

주: 가구원수 제곱근으로 나누어 균등화한 총소득을 활용하여 개인단위로 분석하였다. 음의 총소득은 0으로 변환되었다.

가계자산조사/가계금융조사/가계금융·복지조사의 농가는 가구주가 농림어업에 종사하는 가구로 정의하였으며, 이는 농가경제조사의 농가 정의와 다소 차이가 있다. 1991년 가구소비실태조사에서 농가는 농업경영자 및 어업경영자 가구로 정의하였다.

한 분석이 필요하지만, 가구소비실태조사에서 동지역과 읍면지역을 구분할 수 없어 전국 2인 이상 비농가에 대한 분석결과를 제시하였다(B2). 1991년과 1996년을 보면 지니계수는 가구소비실태조사(B2)가 가계동향조사(A1)보다 약간 크고 빈곤율은 거의 유사하다. 하지만 2000년에는 가구소비실태조사의 지니계수와 빈곤율, 특히 지니계수가 가계동향조사에 비해 매우 높게 나타나 두 데이터 간의 불일치가 두드러진다.

2005년 이후 시기에는 가계자산조사/가계금융조사/가계금융·복지조사의 분석결과를 추가했다. 이때 포괄범위가 동일한 C2와 A2, C3과 A3, C4와 A4를 각각 비교할 수 있다. 먼저 2009~2012년을 살펴보면, C2, C3, C4 자료는 각각 A2, A3, A4에 비해 지니계수와 빈곤율 모두 상당히 높게 나타나 두 데이터 간의 불일치가 뚜렷이 확인된다. 한편 2005년의 가계자산조사(C2, C3, C4) 자료를 2005년의 A2, 그리고 인접한 2006년의 A3, A4 자료와 비교해보면, C의 지니계수가 A보다 높긴 하지만 그 정도가 2009~2012년보다 훨씬 작고, 심지어 빈곤율의 경우에는 C가 A보다 낮게 나타난다. 즉, 가계자산조사로 분석한 소득분배는 가계금융조사/가계금융·복지조사와 뚜렷한 차이를 보이는 것이다.

이처럼 여러 데이터 간에 소득분배 수준과 추이가 불일치하는 부분에 대해서는 평가가 필요하다. <표 5>에서는 평가를 위한 기초 자료로 각 데이터의 소득원천별 평균을 분석한 결과를 제시하였다. 이때 모든 시점에 조사된 가계동향조사를 기준으로 각 데이터의 소득과약 수준을 대략 평가해 볼 수 있다. 그런데 1991, 1996, 2000년의 경우 가계동향조사는 도시지역 2인 이상 비농가를 대상으로 한 반면, 가구소비실태조사는 비도시지역을 포함한 전국 2인 이상 비농가를 대상으로 했기 때문에 직접적인 비교에 제한이 있을 수 있다. 그 정도를 판단하기 위해 2005년 가계동향조사 데이터로 도시지역 2인 이상 비농가(a)와 전국 2인 이상 비농가(b)의 소득원천별 평균을 분석한 결과, 근로소득 외의 소득원천은 모두 약 연 3만원 이하의 작은 차이를 보였고, 근로소득과 총소득의 차이가 약 연 37~40만원으로 나타났다. 따라서 1991, 1996, 2000년 가계동향조사에 소득수준이 상대적으로 낮은 비도시지역을 포함한다면 근로소득과 총소득이 좀더 낮아질 것임을 고려하여 분석결과를 해석할 수 있다.

〈표 5〉 소득원천별 평균

(단위: 만원/년, 명목소득 기준)

연도	대상	데이터	근로 소득	사업 소득	재산 소득	공적 이전 소득	사적 이전 소득	총소득
1991년	a	가계동향조사	423	189	23	6	28	669
	b	가구소비실태조사	-	-	-	-	-	690 (103.1)
1996년	a	가계동향조사	786	370	45	14	49	1264
	b	가구소비실태조사	819	438	67	10	45	1379 (109.1)
2000년	a	가계동향조사	840	378	44	25	70	1357
	b	가구소비실태조사	897	514	87	25	58	1582 (116.6)
2005년	a	가계동향조사	1256	449	43	45	97	1889
	b	가계동향조사	1219	448	40	48	94	1849
	b	가계자산조사	-	-	-	-	-	2016 (109.0)
2009년	c	가계동향조사	1462	451	45	86	99	2143
	c	가계금융조사	-	-	-	-	-	2376 (110.9)
2010년	c	가계동향조사	1544	491	45	94	102	2277
	c	가계금융조사	-	-	-	-	-	2518 (110.6)
2011년	c	가계동향조사	1647	517	45	99	111	2418
	c	가계금융·복지조사	1766	689	114	93	41	2703 (111.8)
2012년	c	가계동향조사	1771	521	51	106	125	2573
	c	가계금융·복지조사	1899	658	121	95	36	2808 (109.1)

자료: 가구소비실태조사, 가계자산조사/가계금융조사/가계금융·복지조사, 가계동향조사.

주: a-도시/비농가/2인 이상, b-도시+비도시/비농가/2인 이상, c-도시+비도시/비농가/1인 이상.  
가구원수 제곱근으로 나누어 균등화한 소득을 활용하여 개인단위로 분석하였다.

괄호 안의 수치는 해당 데이터 총소득의 가계동향조사 대비 비율(%)을 의미한다. 2005년은 b 기준이다.

1996년 가구소비실태조사에서 부업소득은 사업소득에 포함하였다. 가계동향조사에서 주택 등 임대소득은 재산소득에 포함하였다. 가계동향조사에서 소득분배지표 마이크로데이터의 구성과 동일하게 가구원수를 8명으로 탑코딩하였다.

표를 살펴보면, 1991년과 1996년 가구소비실태조사의 전국 2인 이상 비농가 총소득 평균은 각각 가계동향조사 도시지역 2인 이상 비농가의 103.1%,

109.1%로 더 높게 나타났다. 1996년의 소득원천별 평균을 살펴보면 가구소비실태조사가 가계동향조사보다 근로소득, 사업소득, 재산소득을 더 많이 포착하는 것을 알 수 있다. 2000년에는 두 데이터 간 차이가 더욱 커져 가구소비실태조사의 총소득은 가계동향조사의 116.6%에 이르고, 특히 재산소득은 2배에 가깝게 조사되었다. 게다가 1991, 1996, 2000년 두 데이터 간 총소득과 근로소득 차이는 가계동향조사에서 비도시지역이 제외되었음을 고려하면 표에 제시된 것보다 더 클 것으로 예측된다. 가계자산조사/가계금융조사/가계금융·복지조사의 총소득 역시 가계동향조사의 약 109~112%로 더 높다. 2011년과 2012년을 살펴보면, 가계금융·복지조사는 가계동향조사에 비해 근로소득, 사업소득, 재산소득을 훨씬 더 많이 포착하는 반면 이전소득, 특히 사적 이전소득을 더 적게 포착한다. 이처럼 전반적으로 가계동향조사는 다른 데이터에 비해 이전소득, 특히 사적 이전소득이 더 높게 조사되지만, 근로소득, 사업소득, 재산소득으로 이루어진 본원소득이 낮게 조사되기 때문에 총소득이 더 낮게 나타난다. 주로 시장에서 결정되는 본원소득이 낮게, 재분배적 성격이 강한 이전소득이 높게 조사되는 것은 가계동향조사의 소득분배수준이 다른 데이터에 비해 양호한 이유를 일정 부분 설명해 준다.

이 분석은 각 데이터의 소득과약 수준에 대한 상대적인 비교로 이해될 수 있지만, 각 데이터가 절대적인 의미에서 소득을 얼마나 정확하게 포착했는지를 판단하기 위해서는 국민계정과 비교가 도움이 될 수 있다. 강석훈(2000: 59)은 1996년 가구소비실태조사의 국민계정 대비 소득원천별 금액이 근로소득 86.4~88.6%, 사업소득 117.6~120.0%, 재산소득 33.9%, 연금 21.6%, 사회보장수혜 16.5%라고 보고하였다. 김낙년(2012b: 99)은 2010년 가계동향조사의 국민계정 대비 소득원천별 금액이 임금 및 급여 84.2%, 이자 6.0%, 배당금 2.2%, 사업소득 121.0%인 것으로 보고하였다. 김낙년·김종일(2013: 14)은 국민계정 대비 소득원천별 금액이 2000년 가구소비실태조사의 경우 임금 및 급여 82.2%, 이자 및 배당금 23.4%, 사업소득 148.2%, 2006년 가계동향조사의 경우 임금 및 급여 81.8%, 이자 및 배당금 5.7%, 사업소득 109.4%로 나타났다고 보고하였다. 이러한 보고를 종합하면, 가구소비실태조사와 가계동향조사의 근로소득은 대체로 국민계정 대비 80%대의 유사한 수준을 보이지만, 두 데이

터의 사업소득이 국민계정에 비해 과다하게 포착되는 가운데 그 정도가 특히 2000년 가구소비실태조사에서 심하게 나타난다. 한편 이자 및 배당금 등의 재산소득은 전반적으로 소득과약률이 상당히 낮은 가운데, 특히 가계동향조사의 재산소득 수준이 매우 낮게 나타난다. 이러한 양상은 <표 5>에서 살펴본 가구소비실태조사와 가계동향조사의 상대적인 비교에 대체로 부합한다.

### 3. 분배지표 차이의 발생요인 검토

지금부터 <그림 1>의 시계열 분석결과를 집중적으로 검토해 보자. 핵심적인 과제는 각 데이터의 시계열 일관성을 판단하는 것이다. 본 연구에서는 분석기간 전체를 포괄하는 가계동향조사가 시계열 일관성을 만족시킨다고 가정하고 이를 기준으로 각 데이터를 평가한다. 가계동향조사는 고소득층의 소득이 낮게 조사되었다는 지적이 여러 차례 제기된 바(김낙년 2012a, 2012b; 김낙년·김종일 2013) 각 시점의 소득분배수준을 정확하게 측정하지 못했을 가능성이 있다. 하지만 동일한 방식으로 꾸준히 이루어진 조사인 만큼 그러한 편위의 내용과 강도가 시간에 따라 비슷하게 유지됐을 것이라고 본다면 시계열 일관성은 확보될 수 있다. 따라서 여러 시점의 가구소비실태조사, 가계자산조사/가계금융조사/가계금융·복지조사 데이터와 동일 시점 가계동향조사와의 소득분포 차이가 일관되게 나타나는지를 분석함으로써 각 데이터의 시계열 일관성을 평가할 수 있다.

구체적인 평가 방법은 다음과 같다. <표 6>에서는 가계동향조사로 전국 2인 이상 비농가의 소득분위 경계값을 분석하였다. 이때 10분위 경계값은 절대값이 아닌 중위소득 대비 비율로 표시하여 지니계수와 빈곤율과 같이 단위불변성을 갖는 소득분배지표와의 관련성을 이해하는 데 초점을 맞추었다. 예를 들어, 표의 P10은 소득하위 10% 경계값의 중위소득 대비 비율을 의미한다. 이는 각 사례의 소득을 일괄적으로 중위소득으로 나누어 모든 데이터의 중위소득을 100(%)이라는 동일한 값으로 조정된 후의 상대소득분포를 분석하는 것으로 해석될 수 있고, 이하에서는 서술의 편의상 중위소득 대비 소득분위 경계값 비율을 상대소득분위 경계값으로 부른다.

〈표 6〉 전국 2인 이상 비농가 총소득 기준 상대소득분위 경계값, 가계동향조사 (단위: %)

	1991년	1996년	2000년	2005년	2009년	2010년	2011년	2012년
P10	54.1	50.0	48.4	43.8	43.7	44.5	44.7	45.4
P20	68.3	65.4	64.6	60.9	60.6	62.8	61.5	62.3
P30	79.0	77.4	77.2	75.0	74.5	76.3	76.6	75.4
P40	89.5	88.6	88.3	87.6	87.4	88.1	88.3	87.6
P50	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
P60	112.9	112.6	112.6	112.9	112.9	113.9	113.2	113.4
P70	126.5	128.1	128.2	128.6	130.4	129.5	127.7	129.1
P80	144.6	147.5	150.2	148.9	153.6	152.4	151.4	149.3
P90	178.5	180.7	184.7	185.4	190.0	188.9	185.9	186.6

자료: 소득분배지표 마이크로데이터.

주: 상대소득분위 경계값은 중위소득 대비 소득분위 경계값의 비율을 의미한다.

1991, 1996, 2000년은 가계동향조사에 비도시지역을 포함했을 때의 가상적인 값이다.

그런데 1991, 1996, 2000년 가계동향조사는 도시지역만 조사하였으므로, 비도시지역을 포함한 전국 2인 이상 비농가에 대한 소득분포를 얻기 위해서는 일정한 가정에 따른 조정이 필요하다. 본 연구는 도시지역과 비도시지역이 모두 조사된 2003년 가계동향조사 데이터로 도시지역 P10과 전국 P10을 구한 후, 두 수치의 비(ratio)를 1991, 1996, 2000년 도시지역 P10에 곱하는 방식으로 가상적인 전국 P10을 추정하였다. 그 외 P20~P90도 동일한 방식으로 추정하였다. 이때 비도시지역 포함 여부에 따른 소득분포 차이가 1991, 1996, 2000년과 2003년이 동일하지는 않을 것이기 때문에 이러한 추정에는 한계가 있지만, 대략적인 경향을 파악하는 데는 도움이 될 수 있다.<sup>6)</sup> 최종적으로 구성

6) <그림 1>에서 2003년 이후 가계동향조사로 분석한 도시지역 2인 이상 비농가 (A1)와 전국 2인 이상 비농가(A2) 자료를 살펴보면, 비도시지역 포함 여부에 따른 차이가 2008~2009년 이후 거의 사라지는 것이 확인된다. 그 이유를 찾기 위해서는 도시지역과 비도시지역의 인구규모, 두 지역 간 평균적인 소득격차, 각 지역 내 소득불평등 수준의 변화를 살펴봐야 한다. 2인 이상 비농가에 대한 분석 결과, 전체 인구 중 비도시지역의 비율은 2003~2013년 동안 대체로 14~15% 수준을 유지하는 것으로 나타났다. 도시지역 대비 비도시지역의 평균 총소득은 일부 변동이 있

된 표를 살펴보면, 1990년대와 2000년대에 상대소득 P10, P20이 하락하고 P80, P90이 상승하여 소득분배가 악화되는 양상이 확인된다.

다음으로 <표 7>에서는 가구소비실태조사, 가계자산조사/가계금융조사/가계금융·복지조사 데이터를 활용하여 동일한 분석을 실시한 결과를 제시하였다. 그런데 각 데이터의 상대소득분포를 가계동향조사를 기준으로 평가하기 위해 각 데이터의 상대소득분위 경계값을 동일 시점 가계동향조사의 상대소득분위 경계값에 대한 비율로 나타낼 수 있다.

**<표 7> 전국 2인 이상 비농가 총소득 기준 상대소득분위 경계값, 가구소비실태조사/가계자산조사/가계금융조사/가계금융·복지조사** (단위: %)

	1991년	1996년	2000년	2005년	2009년	2010년	2011년	2012년
P10	57.0	53.5	45.2	47.5	39.1	36.0	40.4	39.9
P20	70.0	68.0	62.0	63.5	56.6	55.0	57.5	57.1
P30	80.4	78.6	74.8	76.9	70.8	71.4	71.7	71.6
P40	90.2	89.3	87.0	88.9	85.4	85.2	85.5	84.7
P50	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
P60	111.9	112.2	113.0	112.7	117.4	118.4	117.7	117.0
P70	125.2	126.8	130.4	126.9	135.9	138.4	137.3	136.1
P80	145.3	147.6	154.1	149.3	162.9	166.7	165.8	164.6
P90	178.9	181.7	195.0	188.1	212.2	214.8	215.8	206.8

자료: 가구소비실태조사(1991, 1996, 2000년), 가계자산조사(2005년), 가계금융조사(2009, 2010년), 가계금융·복지조사(2011, 2012년).

지만 2003년 82.4%에서 2013년 90.2%로 증가하는 것으로 나타났다. 비도시지역의 지니계수는 2003~2006년에는 도시지역보다 0.029~0.038 정도 높았으나 점차 차이가 줄어들어 2013년에는 오히려 도시지역보다 0.004 낮게 나타났다. 즉, 비도시지역과 도시지역 간 평균소득격차와 불평등 차이가 감소하는 것이 도시지역과 전국의 소득분배지표 값이 수렴하는 데 영향을 미친 것으로 보인다. 그런데 이러한 변화는 2006~2007년 이후에 나타난 것이므로, <표 6>에서 1991, 1996, 2000년의 분포를 조정할 때는 가장 가까운 시점인 2003년의 데이터를 활용하는 것이 타당하다. 물론 1991, 1996, 2000년의 경우 2003년에 비해 도시지역과 비도시지역 간 소득분포 차이가 더 컸을 가능성도 충분하다는 점에서 본 추정방식에 한계가 있다.

이렇게 구성된 <표 8>이 시계열 일관성 평가의 핵심이며, 표의 수치는 다음과 같이 해석된다. 예를 들어 1991년 하위 10%의 경계값이 가계동향조사에서는 중위소득의 54.1%였지만 가구소비실태조사에서는 중위소득의 57.0%로 나타나 가구소비실태조사의 상대소득 P10이 가계동향조사의 105.2%(57.0 ÷ 54.1)에 해당한다. 즉, 1991년 가구소비실태조사는 가계동향조사에 비해 하위 10% 경계에 해당하는 상대소득을 5.2% 높게 조사한 것이다. 이러한 방식으로 나머지 경계값도 살펴보면, 1991년 가구소비실태조사는 가계동향조사에 비해 상대소득 P10과 P20을 높게, P30~P90은 대체로 비슷하게 포착했음을 알 수 있다. 그런데 이러한 양상은 1996년 가구소비실태조사에서도 놀라울 정도로 유사하게 나타난다. 즉 1991년과 1996년 가구소비실태조사의 상대소득분포는 각각 동일 시점 가계동향조사와 차이가 있지만 그 차이의 형태가 상당히 비슷하다. 따라서 가계동향조사 시계열이 일관된다고 가정하면 1991년과 1996년 가구소비실태조사 역시 비교적 일관된 시계열 데이터로 간주할 수 있다.

**<표 8> 가구소비실태조사/가계자산조사/가계금융조사/가계금융·복지조사의 총소득 기준 상대소득분위 경계값의 가계동향조사 대비 비율** (단위: %)

	1991년	1996년	2000년	2005년	2009년	2010년	2011년	2012년
P10	105.2	106.9	93.4	108.3	89.6	80.8	90.3	88.0
P20	102.4	104.0	96.0	104.2	93.4	87.6	93.5	91.6
P30	101.8	101.5	96.9	102.5	95.0	93.6	93.7	94.9
P40	100.8	100.8	98.5	101.5	97.7	96.7	96.9	96.6
P50	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
P60	99.1	99.7	100.3	99.8	104.0	104.0	104.0	103.2
P70	99.0	99.0	101.7	98.6	104.3	106.9	107.5	105.4
P80	100.5	100.1	102.6	100.3	106.1	109.4	109.5	110.2
P90	100.2	100.5	105.6	101.4	111.7	113.7	116.1	110.8

자료: <표 6>과 <표 7>의 상대소득분위 경계값을 각 분위별로 나누어 계산하였다.

주: 1991/1996/2000년-가구소비실태조사, 2005년-가계자산조사, 2009/2010년-가계금융조사, 2011/2012년-가계금융·복지조사.

이러한 방식으로 표를 살펴보면, 각 데이터와 가계동향조사의 상대소득분포 차이가 크게 두 가지 유형으로 뚜렷하게 구별되는 흥미로운 패턴이 관찰된다. 1991, 1996, 2005년에는 상대소득분위 경계값의 가계동향조사 대비 비율이 P10에서 105.2~108.3%로 높고 분위가 증가할수록 조금씩 낮아져 P70에서 98.6~99.0%로 저점에 이르렀다가 P80과 P90에서 100.1~101.4%로 다시 조금 높아지는 양상이 나타난다. 반면 2000, 2009~2012년에는 P10의 80.8~93.4%에서 P90의 105.6~116.1%까지 분위가 증가할수록 단조증가하는 양상이 나타난다(이하에서는 서술의 편의상 두 유형을 '91/96/05유형', '00/09~12유형'으로 부른다). 즉, 91/96/05유형은 가계동향조사에 비해 저소득층의 상대소득을 상당히 높게, 중상소득층은 조금 낮게, 최고소득층은 조금 높게 조사하는 반면, 00/09~12유형은 저소득층의 상대소득은 낮게, 고소득층의 상대소득은 높게 조사하는 것이다. 00/09~12유형을 91/96/05유형과 직접 비교할 때에도 저소득층의 상대소득은 낮게, 고소득층은 높게 나타난다.

이러한 결과는 데이터를 산출한 조사연구가 동일하거나 유사한 경우에도 연도별 데이터를 연결하여 추정된 소득분포의 시계열이 일관성을 보이지 않음을 시사한다. 기존 연구에서는 1996, 2000년 가구소비실태조사의 표본 대표성이 일치한다는 점에서 두 데이터를 연결한 분석이 현실을 잘 설명한다고 보는 견해가 있지만(현진권·임병인 2004: 61), 본 연구의 분석결과에 따르면 1991/1996년과 2000년 가구소비실태조사는 일관된 시계열로 보기 어렵다. 또한 2009~2012년 가계금융조사/가계금융·복지조사는 조사 내용상 2005년 가계자산조사와 연결되지만, 실제로 조사된 소득분포는 2005년과 2009~2012년 사이에 뚜렷한 시계열 단절이 존재한다. 오히려 1991/1996년 가구소비실태조사는 2005년 가계자산조사와, 2000년 가구소비실태조사는 2009~2012년 가계금융조사/가계금융·복지조사와 유사하게 나타난다.

그렇다면 이와 같이 유사한 조사에서 시계열적으로 서로 다른 패턴이 나타나는 이유는 무엇일까? 그 이유로는 소득 관련 정보를 조사하는 방식의 차이를 생각할 수 있다. 먼저 가구소비실태조사의 경우를 살펴보면, 2000년과 비교할 때 1991/1996년에 발견되는 중요한 차이점은 2개월에 걸쳐 가계부기장 방식의 가계수지 조사를 병행했다는 점이다. 본 연구의 분석에서는 1991/1996

년 가구소비실태조사에서 일회성 면접조사로 수집된 연간소득 변수를 활용하였지만, 가계부기장을 병행했다는 사실 자체가 표본의 구성에 영향을 미쳤을 가능성이 있다. 고소득층은 많은 시간이 요구되는 가계부기장에 따르는 기회비용이 크고 소득을 숨기려 하는 성향이 강하기 때문에 소득이 자세하게 조사되는 가계부기장방식 조사에서 누락되는 무응답 편이가 더 강하게 나타날 수 있다. 실제로 통계청은 1991년 가구소비실태조사 보고서에서 소득수준이 상대적으로 높은 서울, 부산, 경기, 제주 지역의 가계부 기재 불응 및 불능률이 전국 평균보다 높다고 보고하면서 “고소득층이거나 고학력자일수록 수입의 노출을 꺼려 가계부 작성거부 등 조사에 불응하는 비율이 높아 실제보다 소득구조가 낮은 쪽으로 치우치게 되었을 가능성이 크다”고 지적하였다(통계청 1993: 10).

그런데 고소득층의 누락은 고소득층의 상대소득이 낮게 조사된 것을 상당부분 설명할 수 있지만, 저소득층의 상대소득이 높게 조사된 것을 설명하는 데는 한계가 있다. 만약 가계부기장방식이 고소득층뿐만 아니라 저소득층을 누락시키는 효과도 있다면 2000년과 1991/1996년 가구소비실태조사 간의 상대소득분포 차이가 좀 더 설명될 수 있다. 분명한 증거를 찾기는 어렵지만, 가계부기장은 일회성 조사에 비해 응답부담이 상당히 크기 때문에<sup>7)</sup> 당장의 생계에 어려움을 겪는 빈곤층이 더 큰 부담을 느껴 이들의 응답률이 낮을 가능성이 있다. 기존 연구에서도 가계동향조사의 경우 가계부기장의 부담 때문에 최상층과 최하층의 응답률이 낮을 가능성이 높다는 진단이 이루어진 바 있다(김낙년 2012b: 101).

가계자산조사가 가계금융조사/가계금융·복지조사와 다른 패턴을 보이는 이유도 가계부기장방식의 영향으로 어느 정도 설명될 수 있다. 2006년에 실시된 가계자산조사는 기본적으로 가계금융조사/가계금융·복지조사와 동일하게 전국 전 가구를 모집단으로 하지만, 실제 표본추출 과정에서는 가계동향조사(당

7) 1998~2002년 동안 일주일의 경제활동만을 조사하는 경제활동인구조사에 비해 한 달의 가계수지를 모두 조사하는 도시가계조사의 표본소실률이 더 높게 나타났다는 사실은 응답부담이 무응답 편이에 영향을 미칠 수 있음을 보여준다(이지연 2005: 21~22).

시 ‘가계조사’)와 농가경제조사의 표본가구 중에서 표본을 선정하였다(통계청 2007: 11). 따라서 가계자산조사의 표본가구는 가계동향조사와 농가경제조사에 참여하면서 가계부기장을 병행한 셈이고, 이로 인한 무응답 편이가 가계자산조사에도 나타났을 가능성이 있다. 1991/1996년 가구소비실태조사와 2005년 가계자산조사는 모두 가계부기장에 참여한 표본을 대상으로 일회성 면접조사방식의 연간소득을 조사했다는 특이한 구조를 공유하고 있고, 이것이 세대데이터의 가계동향조사 대비 상대소득분포가 유사하게 나타나는 데 일정한 영향을 미쳤을 것으로 예상된다. 단, 91/96/05유형이 가계동향조사보다 저소득층의 상대소득을 높게 포착하는 이유는 짐작하기 어렵다.

가계동향조사 역시 가계부기장을 실시하므로 91/96/05유형에서처럼 고소득층이나 저소득층이 누락되는 무응답 편이의 가능성이 있고, 이는 가계동향조사에서 일회성 면접조사방식의 00/09~12유형에 비해 저소득층의 상대소득이 높게, 고소득층의 상대소득이 낮게 조사되는 이유를 부분적으로 설명해준다. 특히 1991/1996년 가구소비실태조사에서는 2개월 동안만 가계부기장을 실시한 반면, 가계동향조사는 몇 년에 걸쳐 가계부를 작성하는 패널조사이기 때문에 응답부담이 훨씬 클 뿐만 아니라 주거가 불안정한 저소득층의 패널탈락률이 높게 나타날 수 있다는 점도 언급될 필요가 있다. 실제로 1998~2002년 동안 가계동향조사(당시 ‘도시가계조사’)에서는 자가 가구에 비해 전세, 사글세, 월세 가구의 표본소실률이 더 높았고, 주된 가구소득이 연금인 가구에 비해(주로 빈곤층일 것으로 예상되는) 수증·보조인 가구의 표본소실률도 높게 나타났다(이지연 2005: 29).

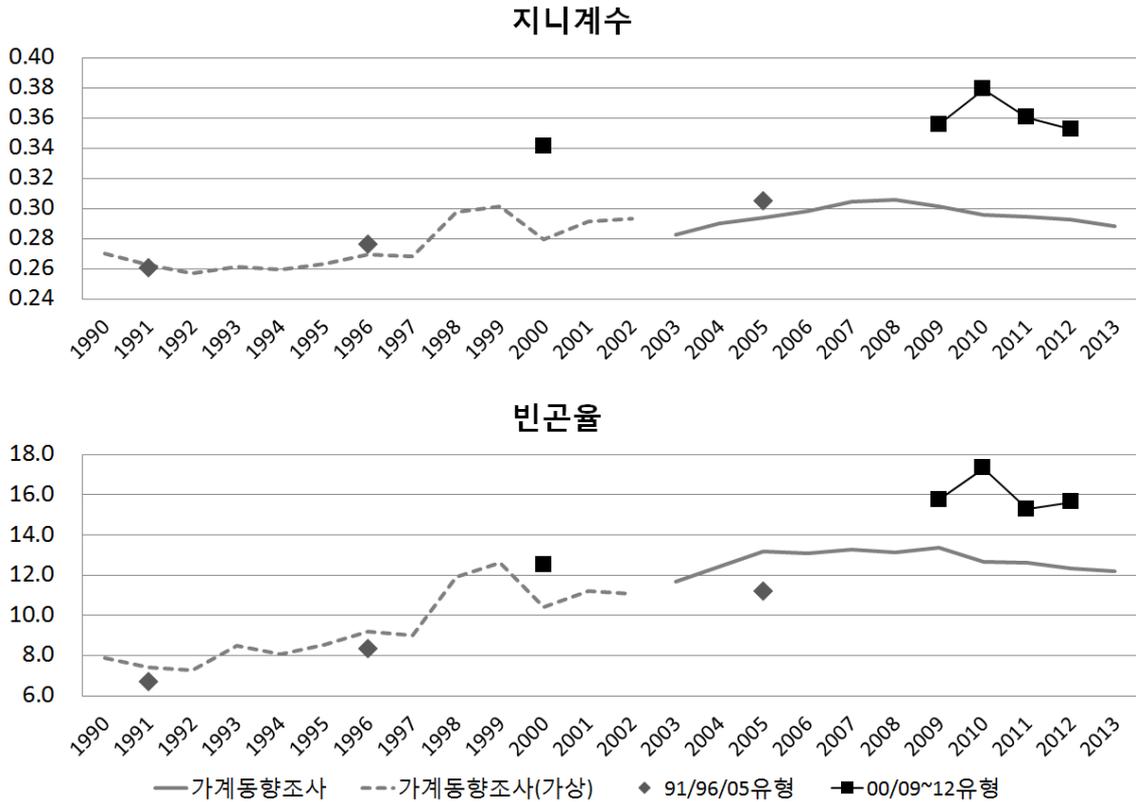
한편 가계동향조사에서는 무응답 편이와 별도로 소득계층별 소득보고 행태에 차이가 있을 가능성도 있다. 91/96/05유형에서는 가계부기장이 병행되긴 하지만 00/09~12유형과 마찬가지로 일회성 면접조사방식의 연간소득이 조사되는 반면, 가계동향조사에서는 가계부기장으로 조사된 월간소득을 연간화하여 사용해야 하기 때문이다. 앞서 <표 5>에서는 가계동향조사가 가계금융·복지조사보다 근로소득, 사업소득, 재산소득은 낮고 이전소득은 높게 조사하는 것을 확인하였는데, 저소득층일수록 총소득에서 이전소득이 차지하는 비중이 크기 때문에 가계동향조사에서 상대적으로 저소득층의 소득이 높게 조사될

가능성이 있다. 그리고 확인하기는 어렵지만, 반대로 소득의 노출을 꺼리는 고소득층의 경우에는 소득을 자세하게 조사하는 가계부기장방식에서 오히려 소득을 낮게 보고할 가능성도 배제할 수 없다.

이처럼 여러 데이터들의 상대소득분포가 몇 가지 유형으로 나뉘는 데는 가계부기장 등 조사방식의 차이가 영향을 미쳤을 것으로 짐작되지만, 현재로서는 잠정적인 예측일 뿐임을 언급해 두고자 한다. 가계부기장방식에서 고소득층이 누락되는 편익이 발생하는 것은 분명해 보이지만 그 편익이 일회성 면접조사방식의 경우보다 반드시 더 크게 나타나는지는 분명하지 않다. 동일한 표본에 대해 일회성 면접조사와 가계부기장방식 조사를 실시했을 때의 소득계층별 응답률과 과소보고 행태를 보여주는 자료는 없기 때문에, 위에서 살펴본 몇몇 근거에도 불구하고 가계부기장방식이 데이터 간의 차이를 낳는 핵심 요인이라고 단언할 수는 없다. 다만 본 연구의 검토 결과는, 그 이유가 확실하지는 않더라도, 여러 데이터 간에 소득계층별 소득 파악 정도가 다르기 때문에 시계열 분석 시 이를 고려해야 한다는 점은 비교적 분명하게 보여준다.

이상의 논의를 바탕으로 <그림 2>에서는 전국 2인 이상 비농가에 대한 소득분배지표를 가계동향조사, 91/96/05유형, 00/09~12유형의 세 가지 시계열로 재구성하였다. 이때 비도시지역이 제외된 1990~2002년 가계동향조사의 경우 <표 6>에서와 마찬가지로 2003년 가계동향조사로 분석된 도시지역 2인 이상 비농가와 전국 2인 이상 비농가의 지니계수 및 빈곤율비(ratio)를 적용하여 추정하였다. 즉, 1990~2002년에 2003년과 마찬가지로 도시지역에 비해 비도시지역을 포함한 전국 지니계수와 빈곤율이 조금 더 높을 것이라고 가정한 것이다. 그림에서 세 가지 시계열이 모두 소득분배가 악화되는 추이를 보여주는 가운데, 각 시점의 소득분배수준은 자료에 따라 다르게 나타난다. 가계동향조사에 비해 91/96/05유형은 빈곤율은 약간 낮고 지니계수는 약간 높거나 유사하게 나타나고, 00/09~12유형은 지니계수와 빈곤율이 모두 높게 나타난다.<sup>8)9)</sup>

8) 세부적으로 살펴보면 2009~2012년 가계금융조사/가계금융·복지조사의 경우 가계동향조사와 달리 유독 2010년의 지니계수와 빈곤율이 높게 나타나는 점이 보인다. 그런데 특별한 경기변동이 관찰되지 않는 2009년과 2010년 사이에 소득분배수준이 실제로 크게 악화되었을 가능성은 작다고 판단된다. 따라서 분명한 이유



<그림 2> 전국 2인 이상 비농가 총소득 기준 지니계수 및 빈곤율 추이

자료: <그림 1>의 도시+비도시/비농가/2인이상 자료를 재구성하였다.

주: 91/96/05유형 - 가구소비실태조사와 가계자산조사, 00/09~12유형 - 가구소비실태조사와 가계금융조사/가계금융·복지조사.

1990~2002년 가계동향조사(가상)는 비도시지역을 포함했을 때의 가상적인 값이다.

를 확인할 수는 없지만 가계금융조사의 2009~2010년 시계열 일관성을 보장할 수 없고, 2만 가구로 표본이 확대·개편되면서 새로 패널조사가 시작된 2011년 이후 가계금융·복지조사에 좀 더 무게가 실린다고 평가할 수 있다.

- 9) 가계금융·복지조사는 패널조사이므로 패널탈락이 소득분배 추이의 분석결과에 영향을 미칠 가능성이 있다. 그 가능성을 탐색하기 위해 1차년도에 조사된 19,744가구와 1, 2차년도에 모두 조사된 18,076가구에 대해 1차년도 소득의 지니계수와 빈곤율을 분석한 결과(가중치 미적용), 지니계수는 0.384에서 0.380으로 소폭 감소하였고 빈곤율은 17.4%에서 17.5%로 미미하게 증가한 것으로 나타났다. 따라서 2차년도까지는 패널탈락이 소득분배 추이에 그다지 큰 영향을 미치지 않는 것으로 판단된다. 향후 패널차수가 축적되는 과정에서 지속적인 검토가 필요할 것이다.

#### 4. 1990년대 이후 장기 시계열 분석 검토

한국 소득분배의 시계열 분석에서 1990년대 이후 지난 20여 년간의 변화를 분석하는 것은 중요한 과제이다. 이러한 분석에서 1990년대 시점의 데이터 선택에 대해서는 큰 쟁점이 있을 수 없다. 가계동향조사의 경우 도시지역 비농가 2인 이상 가구를 대상으로 하는 시계열 분석에 이용될 수 있지만, 전국의 1인 이상 가구를 포괄하는 데이터로는 표본의 전국적 대표성을 가진 가구소비실태조사가 유일하기 때문이다. 더욱이 1991/1996년 가구소비실태조사 데이터는 같은 연도의 가계동향조사 분석결과와도 비교가능성이 높다는 점이 앞에서 확인되었다.

그런데 2011년 이후 현재는 전국의 전 가구를 포괄하는 데이터로 가계동향조사/농가경제조사 결합데이터와 가계금융·복지조사 데이터의 두 가지가 수집되고 있다. 문제는 두 데이터로 분석한 소득분배지표가 크게 다르다는 점이다. 2011, 2012년 가계금융·복지조사의 지니계수와 빈곤율은 0.366~0.374, 17.0%로 가계동향조사/농가경제조사 결합데이터의 0.315~0.318, 15.2~15.5%보다 상당히 높다(<그림 1>). <표 9>를 보면, 두 데이터의 중위소득이 거의 유사한 가운데 가계금융·복지조사가 가계동향조사/농가경제조사 결합데이터보다 저소득층의 소득을 낮게, 고소득층의 소득을 높게 조사하고 있음을 알 수 있다. 두 데이터의 소득분포가 다른 것은 앞서 논의한 바와 같이 가계동향조사와 농가경제조사에서는 오랜 시간에 걸쳐 가계부기장을 실시하는 반면 가계금융·복지조사에서는 일회성 면접조사로 연간소득을 조사한다는 차이가 주된 요인일 것으로 짐작된다.<sup>10)</sup>

10) 통계청(2013)은 그 외에도 두 데이터 간의 가구 포괄범위의 차이, 사업소득 개념의 차이를 지적하였다. 가계금융·복지조사는 가계동향조사와 달리 취업이나 학업 등으로 일시적으로 분리된 가구원을 가구에 포함하고, 또한 가계금융·복지조사는 순수익을 사업소득으로 조사하는 반면 가계동향조사는 순수익에서 유보분을 제외한 가구전입소득을 사업소득으로 조사한다는 차이가 있다(통계청 2013). 이러한 차이가 소득분배지표 분석에 어떤 영향을 미치는지에 대해서는 그다지 알려진 바가 없다.

〈표 9〉 2011, 2012년 총소득 기준 소득분포

(단위: 만원/년, %, 명목소득 기준)

	2011년		2012년	
	가계금융 ·복지조사	가계동향조사 +농가경제조사	가계금융 ·복지조사	가계동향조사 +농가경제조사
지니계수	0.374	0.318	0.366	0.315
빈곤율	17.0	15.5	17.0	15.2
P10	798 (36.7)	829 (38.3)	844 (36.3)	894 (38.8)
P20	1200 (55.2)	1254 (58.0)	1273 (54.7)	1359 (59.0)
P30	1521 (70.0)	1586 (73.3)	1617 (69.5)	1686 (73.2)
P40	1836 (84.5)	1881 (87.0)	1980 (85.1)	1992 (86.5)
P50	2173 (100.0)	2163 (100.0)	2326 (100.0)	2302 (100.0)
P60	2563 (117.9)	2453 (113.4)	2750 (118.3)	2617 (113.7)
P70	3024 (139.1)	2798 (129.4)	3200 (137.6)	3023 (131.3)
P80	3683 (169.5)	3337 (154.3)	3882 (166.9)	3511 (152.5)
P90	4808 (221.2)	4134 (191.1)	5000 (215.0)	4401 (191.2)

자료: 소득분배지표 마이크로데이터, 가계금융·복지조사.

주: 도시+비도시/농가+비농가/1인이상에 대한 분석결과이다.

괄호 안의 수치는 중위소득 대비 비율이다.

그렇다면 두 데이터 중 실제 소득분포에 더 가까운 것은 무엇일까? 가계동향조사에서 고소득층이 많이 누락되는 경향이 있다는 사실은 기존 연구에서 여러 차례 지적되어 왔다(김낙년 2012a, 2012b; 김낙년·김종일 2013). 김낙년·김종일(2013: 10~16)은 소득세 자료와 국민계정과의 비교를 통해 가계동향조사가 고소득층을 과소하게 파악하고 금융소득을 거의 포착하지 못하는 반면 가구소비실태조사는 사업소득이 과대보고되기는 했지만 가계동향조사보다는 실태에 가까운 것으로 평가하였다. 본 연구 역시 유사한 맥락에서 가계동향조사/농가경제조사 결합데이터보다 고소득층의 소득이 더 높게 포착된 가계금융·복지조사가 고소득층의 실태와 지니계수를 파악하는 데 우위에 있다고 평가한다.

그런데 가계금융·복지조사에서 저소득층의 소득이 더 낮게 포착된 것에 대해서는 별도의 평가가 필요하다. 가계동향조사/농가경제조사 결합데이터에서 저소득층의 소득이 더 높게 포착된 주된 원인이 응답부담이 크고 주거가 불안정한 저소득층의 무응답률이 높은 데 있다면, 저소득층의 실태도 무응답 편이가 덜 나타난 가계금융·복지조사가 더 잘 파악한다고 볼 수 있을 것이다. 하지만 만약 가계부기장으로 이전소득 등 저소득층의 소득을 더 정확하게 파악했기 때문이라면, 저소득층의 실태와 빈곤율을 파악하는 데는 가계동향조사/농가경제조사 결합데이터가 우위에 있다고 볼 수 있을 것이다. 이에 대해 현재로서는 분명한 답을 내리기 어렵다. 참고로, 김낙년·김종일(2013: 46)은 2010년 가계동향조사 데이터에서 국민계정과 소득세 자료를 근거로 누락된 고소득층을 추가하고 금융소득과 소득세 등을 보정했을 때의 가처분소득 지니계수와 빈곤율을 각각 0.371, 14.5%로 계산한 바 있다. 이를 2011년 가계동향조사/농가경제조사 결합데이터와 가계금융·복지조사로 측정된 가처분소득의 분포와 비교하면,<sup>11)</sup> 지니계수는 가계동향조사/농가경제조사 결합데이터의 0.311보다 가계금융·복지조사의 0.364에 훨씬 가깝고 빈곤율은 가계금융·복지조사의 16.8%보다 가계동향조사/농가경제조사 결합데이터의 15.2%에 더 가깝다. 즉 김낙년·김종일(2013)의 추정이 실제에 근접한다고 가정한다면, 불평등은 가계금융·복지조사가 훨씬 잘 파악하고 빈곤은 가계동향조사/농가경제조사 결합데이터가 조금 더 잘 파악한다고 볼 수 있다.

가계금융·복지조사가 가계동향조사/농가경제조사 결합데이터보다 2010년대의 소득분배 실태를 잘 보여주고 있다고 하더라도 1990년대 이후 장기 시계열 분석에서 1990년대의 가구소비실태조사와 2010년대의 가계금융·복지조사를 연결하여 분석하는 것이 적절한 선택인지에 대해서는 검토의 여지가 있다. 1991/1996년 가구소비실태조사가 면접조사방식으로 연간소득을 조사했다는 측면에서 가계금융·복지조사와 유사하다는 점에 주목한다면 가계금융·복지조사와 연결한 시계열 분석의 일관성에 문제가 없다고 볼 수 있다. 실제로 이 두

11) 가계동향조사/농가경제조사 결합데이터의 가처분소득은 총소득에서 경상조세 및 사회보험료를 제하여 계산하였고, 가계금융·복지조사의 가처분소득은 총소득에서 세금 및 과징금과 사회보험료를 제하여 계산하였다.

조사는 가계동향조사에 비해 소득수준을 높게 파악하는 공통점을 보이기도 한다(<표 5>). 하지만 1991/1996년 가구소비실태조사는 조사표를 통한 연간 소득 조사와 함께 2개월간의 가계부기장방식 가계수지 조사를 병행하고 있어 가계부기장방식의 가계동향조사와 유사한 문제를 가지고 있을 가능성이 있다. 앞에서 보여준 소득분위별 분석은 이러한 가능성을 강하게 시사하고 있다. 이러한 점에 주목한다면 1991/1996년 가구소비실태조사와 연결하여 시계열 분석을 할 때 2006년 이후 가계동향조사/농가경제조사 결합데이터를 이용하는 것이 시계열 일관성을 유지하는 데 도움이 될 것이다.

#### IV. 결론

지난 수십 년간 개인과 가족을 대상으로 한 마이크로 수준의 서베이 데이터는 사회과학 연구에서 매우 중요한 역할을 하였고, 이는 소득분배 분석에서도 마찬가지였다. 그런데 근래에 들어 이러한 마이크로데이터의 질에 대한 우려가 제기되고 있다. 조사 가구가 서베이 조사에 응하지 않는 단위 무응답(unit nonresponse)과 일부 문항에 대해 응답하지 않는 항목 무응답(item nonresponse)이 증가하였고, 응답자의 경우에는 응답의 정확성이 떨어지는 측정오차(measurement error)가 증가하고 있다(Meyer et al. 2015). 한국의 경우 마이크로데이터를 이용하여 추정된 소득분배지표가 소득세자료나 국민계정에 근거한 추정치와 큰 차이가 있다는 지적이 있다(김낙년·김종일 2013). 근래에는 분석 데이터에 따라 소득분배지표 추정치에 큰 차이가 나타나기도 하였다(통계청 2013).

본 연구는 한국 소득분배 데이터의 문제점을 소득분배 시계열 분석을 중심으로 검토하였다. 먼저 시계열 차원에서 1990년 이후 현재까지 생산되고 있는 가계동향조사를 기준으로 각 데이터의 소득분포를 비교하였다. 분석 결과, 가계동향조사와 비교할 때 1991/1996년 가구소비실태조사와 2005년 가계자산조사는 주로 저소득층의 상대소득을 높게 포착하는 반면, 2000년 가구소비실태조사와 가계금융조사/가계금융·복지조사는 저소득층의 상대소득은 낮게, 고소

득층의 상대소득은 높게 포착하는 것으로 나타났다. 이러한 차이가 나타난 것은 가계부기장방식이 무응답 양상에 일정한 영향을 미쳤기 때문일 가능성이 있다. 가계부기장방식의 조사에서 고소득층이 많이 누락된다는 사실은 비교적 분명해 보이고, 저소득층이 누락될 가능성도 존재한다. 특히 가계동향조사처럼 비교적 장기간의 패널조사로 가계부기장을 실시하는 경우에는 주거가 불안정한 저소득층이 탈락할 가능성이 있다. 한편 확실하지는 않지만 가계부기장방식으로 소득을 조사할 때 이전소득 등 저소득층의 소득이 상대적으로 높게 포착되어 과소보고와 같은 측정오차가 줄어들 가능성도 있다.

이러한 분석결과에 따르면, 한국에서 가계부기장방식보다는 면접조사방식으로 소득 관련 정보를 조사하는 마이크로데이터를 발전시키는 것이 소득분배 실태분석에서 무응답 편향(nonresponse bias)을 줄이는 방안이 될 것이다. 사실 가계동향조사와 같은 소비자 조사에서는 가계부기장방식의 가계수지 조사를 채택하는 것이 필요하지만, 소득조사에서 가계부기장방식의 필요성은 적다. 한편, 한국의 마이크로데이터는 소득의 원천이 되는 경제활동에 대한 조사가 취약하여 분석에 제약이 크고 근로소득 등 소득원의 정확한 파악에도 제한이 있다. 이러한 소득과 경제활동에 대한 면접조사가 반드시 패널조사일 필요는 없다. 한 시점에서의 전국 가구와 개인의 소득실태와 분배상태를 파악하고 시간적 추이를 분석하는 데는 표본규모가 큰 횡단조사를 주기적으로 반복하는 것이 효과적일 것이다.<sup>12)</sup>

한편 주어진 현실에서 연구자들이 한국 소득분배를 시계열적으로 분석할 때 어떤 데이터를 선택해야 할지는 매우 중요하게 검토해야 할 문제이다. 기존 연구에서 많이 활용되어 온 가계동향조사는 시계열이 길고 소득항목이 자세하게 조사되었다는 장점이 있지만, 2005년까지는 표본 대표성에 제한이 있고 대체로 고소득층의 소득이 낮게 포착된다는 단점이 있다. 이러한 단점을

12) 단, 본 연구가 모든 소득조사를 하나로 통합해야 한다고 주장하는 것은 아님을 분명히 해두고자 한다. 반복횡단조사는 소득분배 추이를 분석하는 데 장점이 있지만, 개인수준의 소득변화나 소득이동을 추적하는 데는 패널조사가 필요하다. 또한 조사내용의 측면에서도 소득과 함께 소비나 자산, 경제활동 등을 자세하게 파악하는 특화된 조사의 필요성도 크다.

보완하기 위해서는 다른 데이터가 필요하다.

본 연구의 검토에 따르면, 1991/1996년 가구소비실태조사와 2005년 가계자산조사를 연결하면 1인 이상 비농가에 대해 1991~2005년의 소득분배 추이를 비교적 일관되게 분석할 수 있다. 단, 표본 대표성을 일치시키기 위해서는 1991년 가구소비실태조사와 가계자산조사에서 농가에 해당하는 가구를 제외해야 한다. 세 데이터는 표본이 가계부기장에 참여하였지만 일회성 면접조사 방식의 연간소득을 조사하였다는 점에서 유사한 구조를 갖고 있다. 하지만 가계자산조사는 한 차례밖에 이루어지지 않았고 소득원천별 정보가 공개되지 않는다는 점에서 활용도가 그다지 높지 않다.

한편 2000년 가구소비실태조사는 2011년 이후 가계금융·복지조사와 연결하는 것이 가능하다. 이들은 가계부기장의 영향을 받지 않았고 가계동향조사보다 불평등과 빈곤 수준을 높게 추정한다는 점에서 유사하다. 한 예로, 이병희(2014)는 2000년 가구소비실태조사와 2011년 가계금융·복지조사를 연결하여 2000년대 소득불평등 변화를 분석한 바 있다. 단, 가계금융·복지조사에는 2000년 가구소비실태조사와 달리 농어가가 포함되었다는 차이가 있는데, 이 문제는 가계금융·복지조사에서 가구주가 농림어업에 종사하는 가구를 제외하는 것으로 어느 정도 해결 가능하다. 가계금융·복지조사는 가계동향조사에 비해 표본 대표성이 우월하고, 표본규모가 더 크며, 자산이나 경제활동상태 등 다양한 정보를 수집한다는 측면에서 향후 활용도가 높을 것으로 전망된다.

1990년대에서 2010년대에 이르는 보다 장기적인 시계열 분석에서는 1991/1996년 가구소비실태조사와 2010년대의 가계금융·복지조사를 연결하는 분석과 1991/1996년 가구소비실태조사와 2006년 이후 가계동향조사를 연결하는 분석의 두 가지 대안이 존재한다. 가계금융·복지조사는 같은 시기의 가계동향조사에 비해 소득불평등이 심화한 현실을 잘 반영하고 있다는 장점을 가지고 있고 면접조사방식의 소득조사를 한다는 점에서 1991/1996년 가구소비실태조사와 공통점이 있다. 한편 1991/1996년 가구소비실태조사가 (아마도 가계부기장의 영향으로 인해) 같은 시기 가계동향조사와 유사한 소득분배 실태를 보여 비교가능성이 높다는 점을 고려한다면 1991/1996년 가구소비실태조사와 2006년 이후 가계동향조사를 연결하는 분석이 시계열적 변화를 파악하는 데는 더

적절한 선택이 될 수 있다. 한 예로, 장지연·이병희(2013)는 1996년 가구소비실태조사와 2011년 가계동향조사를 연결하여 소득불평등 변화를 분석한 바 있다.

마지막으로 한국 소득분배 시계열 분석의 정확도를 높이기 위해서는 근본적으로 현실의 소득분포를 제대로 반영할 수 있도록 소득조사의 질이 개선되어야 한다는 점이 지적될 필요가 있다. 이와 관련하여 가장 우선적인 접근은 서베이 데이터의 질을 평가할 수 있는 기준(benchmark)이 되는 자료와의 비교·검토이다. 현재 수집되고 있는 대표적인 소득 데이터들과 국민계정이나 소득세 자료, 행정데이터 등과의 비교 등을 통해 개인과 가구단위 서베이 데이터의 문제점을 밝히는 한편, 나아가 데이터 간 불일치의 원인을 파악하고 그 차이를 해소하려는 노력이 필요하다. 소득조사의 질을 높이기 위해서는 많은 연구에서 지적한 바와 같이 소득세 자료나 사회보험료 관련 자료 등 행정데이터를 활용하는 것이 중요한 방안이 될 수 있다(강석훈·현진권 2003; 김낙년·김종일 2013; 통계청 2013). 실제로 많은 국가들은 행정데이터로 가구소득을 파악하거나 행정데이터와 서베이 데이터를 결합하여 가구소득을 조사하고 있다(UNECE 2011: 22~24). 이러한 과정에서 통계청과 관련 학계의 활발한 논의와 협력이 이루어지길 기대한다.

## 참고문헌

- 강석훈. 2000. “서베이데이터와 집계데이터의 비교연구: 가구소비실태조사와 국민계정을 중심으로.” 《계량경제학보》 11(1): 41-70.
- 강석훈·현진권. 2003. “소득분배 관련 미시자료를 어떻게 개선할 것인가?” 《한국정책학회보》 12(4): 201-224.
- 강영욱. 2011. “소득불균등 지표 산출에 쓰이는 기초 통계자료의 한계점에 관한 고찰.” 《한국정책학회보》 20(4): 163-188.
- 구인회. 2004. “한국의 빈곤, 왜 감소하지 않는가? 1990년대 이후 빈곤 추이의 분석.” 《한국사회복지학》 56(4): 57-78.
- 김낙년. 2012a. “한국의 소득불평등, 1963-2010: 근로소득을 중심으로.” 《경제

- 발전연구》 18(2): 125-158.
- 김낙년. 2012b. “한국의 소득집중도 추이와 국제비교, 1976-2010: 소득세 자료에 의한 접근.” 《경제분석》 18(3): 75-114.
- 김낙년·김종일. 2013. “한국 소득분배 지표의 재검토.” 《한국경제의 분석》 19(2): 1-50.
- 김대일·이석배·황운재. 2014. “가구구성과 소득 불평등.” 《노동경제논집》 37(3): 1-44.
- 김진욱. 2015. “한국의 가처분소득 및 순자산 결합분포 특성에 관한 연구.” 《민주사회와 정책연구》 27: 34-68.
- 안국신. 1995. “한국의 경제발전과 소득분배.” 《경제발전연구》 1: 53-76.
- 여유진·김태완. 2006. “한국의 소득불평등 동향과 정책방향: 외국과의 비교분석을 중심으로.” 《보건사회연구》 26(1): 95-134.
- 유경준. 2003. “소득분배의 국제비교를 통한 복지정책의 방향.” 《KDI 정책연구》 25(2): 55-88.
- 이병희. 2014. “2000년대 소득불평등 증가요인 분석: 단계적 분석방법.” 《경제발전연구》 20(1): 73-103.
- 이재욱. 2000. “1980년대 소득분배 추이에 대한 상반된 견해의 평가.” 《한국경상논총》 17(3): 89-108.
- 이정우·황성현. 1998. “한국의 분배문제: 현황, 문제점과 정책방향.” 《KDI 정책연구》 20(1·2): 153-214.
- 이지연. 2005. 《가구조사자료의 종단화 방안》. 통계청 통계연구보고서 05-03.
- 장지연·이병희. 2013. “소득불평등 심화의 메커니즘과 정책 선택.” 《민주사회와 정책연구》 23: 71-109.
- 전병유. 2013. “한국 사회에서의 소득불평등 심화와 동인에 관한 연구.” 《민주사회와 정책연구》 23: 15-40.
- 최바울. 2013. “부부의 노동소득과 취업상태가 소득불평등 변화에 미치는 효과.” 《노동경제논집》 36(3): 97-128.
- 통계청. 1993. 《1991 가구소비실태조사보고서 제1권 가계수지편》.
- 통계청. 2007. 《2006 가계자산조사보고서》.
- 통계청. 2011. 《가계동향조사 이용자 가이드》. 2011년 8월.
- 통계청. 2013. “2013년 가계금융·복지조사 결과 보도 참고자료: 소득분배지표 현

- 황 및 개선방안.” 2013년 11월 19일 통계청 보도자료.
- 현진권·임병인. 2004. “우리나라 소득분배 실증연구의 한계.” 《응용경제》 6(1): 49-68.
- Meyer, B.D., W.K.C. Mok, and J. Sullivan. 2015. “Household Surveys in Crisis.” National Bureau of Economic Research Working Paper No.21399.
- UNECE. 2011. *Canberra Group Handbook on Household Income Statistics*. 2nd Edition. ECE/CES/11.

<접수 2015/8/10, 수정 2015/9/21, 게재확정 2015/9/30>

