

기 획

재정패널조사 테크니컬 리포트

# 소득증빙서류 수집률이 소득분포에 미치는 영향

■ 세수추계패널센터

\* 한국조세재정연구원은 조세정책과 행정을 연구하고 분석하는 데 활용할 수 있는 조사자료를 수집하기 위해 2008년부터 재정패널조사(NaSTaB: National Survey of Tax and Benefit)를 실시하고 있다. 재정패널조사는 설문조사의 한계인 응답 오차 문제를 완화하기 위해 소득증빙서류를 수집·활용하는데, 본고는 이 소득증빙서류 수집률의 변화가 소득에 대한 응답 오차에 어떠한 영향을 미치는지 살펴본다. <편집자 주>

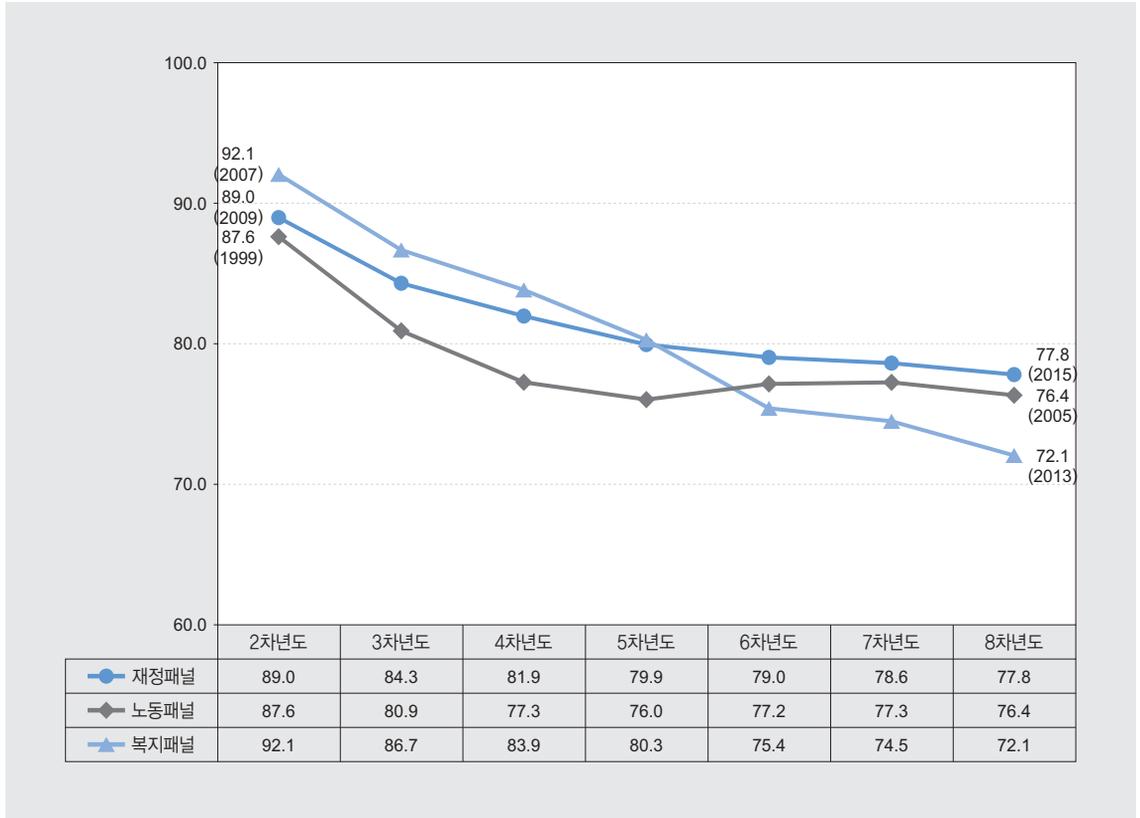
## 들어가는 말

최근 저성장-고령화 사회로의 진입이 가속화되면서 조세·재정제도의 역할이 점점 중요해지고 있다. 특히 복지수요가 증가함에 따라 복지재정을 확충하고 재정정책의 비효율성을 개선하는 정책적 노력이 필요하다. 이와 관련하여 박근혜 정부는 건전재정 기조를 정착시키기 위해 지하경제 양성화하고 비과세·감면제도를 정비하는 등 다양한 제도적 노력을 기울이고 있다. 하지만 이러한 제도적 노력이 결실을 맺기 위해서는 현행 제도의 타당성 및 효과성 분석을 토대로 제도 개선 방안이 모색되어야 한다. 의사도 정확한 진단 없이 환자에게 약을 처방할 수 없듯이 제도의 면밀한 검토 없이 개선방안을 논의할 수 없다. 그렇다면, 현행 제도를 분석함에 있어서 가장 중요한 것은 무엇일까? 그것은 바로 신뢰성 있는 소득, 자산, 부채, 조세, 지출, 복지 등에 대한 포괄적인 통계자료의 확보이다. 우리가 가장 쉽게 생각할 수 있는 통계자료는 행정자료이다. 행정자료는 건강보험자료, 국세 및 지방세 신고자료, 국민연금자료 등이 있다. 행정자료의 가장 큰 장점은 대표성과 정확성이다. 건강보험자료를 예로 들어보자. 우리나라의 모든 의료 공급자와 전 국민은 의무적으로 국민건강보험에 가입해야 한다. 그리고 건강보험자료는 건강보험 운영에 필요한 개인의 성별, 나이, 장애 등의 정보뿐만 아니라 종합소득, 연금소득 등의 정보도 행정전산망을 통해 확보가 가능하다. 따라서 건강보험자료를 통해 건강보험자료 DB에 있는 모든 사람들의 소득 정보는 정확하게 확보할 수 있다. 하지만 행정자료의 가장 큰 단점은 접근성이 상당히 제한되어 있다는 점이다. 개인정보보호 등의 이유로 일반 연구자들이 행정자료를 접근하는 것은 쉽지 않다. 또한 제도를 보다 정확히 분석하기 위해서는 하나의 행정자료만을 사용하기보다는 행정자료 간 또는 행정자료와 다른 설문조사 간의 연계가 필요하다. 최근 통계청이 진행하는 프로젝트 중 가장 흥미로운 것은 다양한 행정자료와 가계금융복지조사와 가계동향조사 자료를 연계하는 프로젝트이다. 이는 신뢰성 있는 자료 확보측면에서 상당히 의미가 있으나 이것이 연구자들에게 공개되기까지는 상당한 시간이 걸릴 것으로 보인다.

현재 연구목적으로 활용할 수 있는 행정자료의 사용이 제한적이기 때문에 본원은 조세재정정책을 연구하고 분석할 수 있는 신뢰성 있는 자료를 구축하기 위해 재정패널DB사업을 수행하고 있다. 재정패널자료는 전국 5,000여 가구와 소득활동을 하는 8,000여명의 가구원을 대상으로 인구통계학적 정보(연령, 성별 등)와 함께 소득, 자산, 부채, 지출, 조세, 복지 등에 대해 2008년부터 매년 조사한다. 현재 9차년도까지 설문조사가 완료된 상태로 2007년~2015년 기간 동안의 설문조사자료가 이용 가능하다. 재정패널자료의 원표본 조사 성공률을 살펴보면, 다른 주요 패널 조사와 비교하여 상당히 안정적인 모습을 보여준다(그림 1 참조). 또한 1차년도에 추출된 원표본 가구 중 8차년도까지 계속 응답한 가구 수 비율은 72.4%이고, 2차년도에 추가적으로 추출된 원표본 가구 중 8차년도까지 계속 응답한 가구 수 비율은 79.5%으로 상당히 높은 수준을 보여준다. 안정적 패널가구 유지율은 조세·재정정책이 개인의 경제행위에 어떻게 영향을 미치는지에 대한 신뢰성 있는 동태적 분석을 가능하게 해주기 때문에 이는 재정패널조사에서 상당히 중요한 의미를 가진다.

[그림 1] 국내 패널 원표본 조사 성공률

(단위: %)



출처: 저자 작성

<표 1> 계속응답가구 현황

(단위: 가구)

표본가구 구분	1차년도	2차년도	3차년도	4차년도	5차년도	6차년도	7차년도	8차년도
1차년도 추출 원표본가구	5,014	4,396	4,122	3,954	3,838	3,754	3,697	3,629
2차년도 추출 원표본가구	-	620	575	541	519	509	501	493
합계	5,014	5,016	4,697	4,495	4,357	4,263	4,198	4,122

출처: 저자 작성

하지만 재정패널자료의 단점 중 하나는 응답자들이 회상에 의존하여 응답하기 때문에 분석에 사용되는 변수들이 응답 오차를 가질 수 있다는 점이다. 특히 소득의 경우 원천징수제도로 인해 응답자는 세전소득을 묻는 질문에서 세후소득을 대답할 가능성이 있다. 그뿐만 아니라 일반적으로 제도의 복잡성 등으로 인해 다수의 응답자들이 자신의 공제 내역을 정확히 알지 못할 가능성이 있어 증빙서류 수집은 상당히 중요한 의미를 가진다. 따라서 재정패널조사에서는 응답 오차를 줄이고 궁극적으로 분석결과의 신뢰성을 높이기 위해 개인의 동의를 얻어 연말정산 서류와 종합소득세 신고서류를 수집하고 있다.

본 보고서의 주된 목적은 소득증빙서류 수집률의 변화가 소득에 대한 응답 오차에 어떠한 영향을 미치고 있는지를 살펴보는 것이다. 응답 오차를 분석하기 위해서는 응답자가 회상에 의해 응답한 소득과 소득증빙서류의 소득을 서로 비교하면 된다. 하지만 안타깝게도 재정패널조사는 응답자의 소득 관련 정보를 소득증빙서류의 정보로 대체하고 있다. 따라서 본 보고서는 소득증빙서류 수집률이 증가함에 따라 재정패널자료의 소득분포가 어떻게 변화하는지 살펴보고자 한다. 이때, 소득증빙서류 수집률과 재정패널자료의 소득분포 간의 관계를 좀 더 정확히 살펴보기 위해 본 보고서는 국세청의 소득분포와 재정패널자료의 소득분포의 변화를 서로 비교한다.

본 보고서는 제1장을 포함하여 총 4장으로 구성된다. 제2장은 소득증빙서류 수집률의 현황을 살펴보고, 제3장은 소득증빙서류 수집률과 근로소득에 대한 응답 오차 간의 관계를 살펴본다. 그리고 제4장은 본 보고서의 결과를 요약한다.

## 1. 소득증빙서류 수집

### 가. 수집 자료

설문조사의 한계인 응답 오차 문제를 완화하기 위해 현재 재정패널조사는 주요 조사항목인 개인의 소득과 소득공제현황에 대한 정보가 담겨있는 소득증빙서류를 수집·활용한다. 재정패널조사에서 수집하는 소득증빙서류는 근로자의 원천징수영수증과 종합소득자의 과세표준확정신고 및 자진납부계산서와 소득공제 및 세액공제명세서이다.

〈표 2〉 재정패널자료에서 수집하는 소득증빙서류

구분	소득증빙서류
근로소득자	근로소득 원천징수영수증
종합소득자 <sup>1)</sup>	과세표준확정신고 및 자진납부계산서, 소득공제명세서, 세액공제명세서

주: 1) 방문판매나 보험모집과 같은 사업소득자는 연말에 소득세를 원천징수하는 사례가 있어, 이 경우에는 사업소득 원천징수영수증도 수집서류에 포함 시킴

출처: 저자 작성

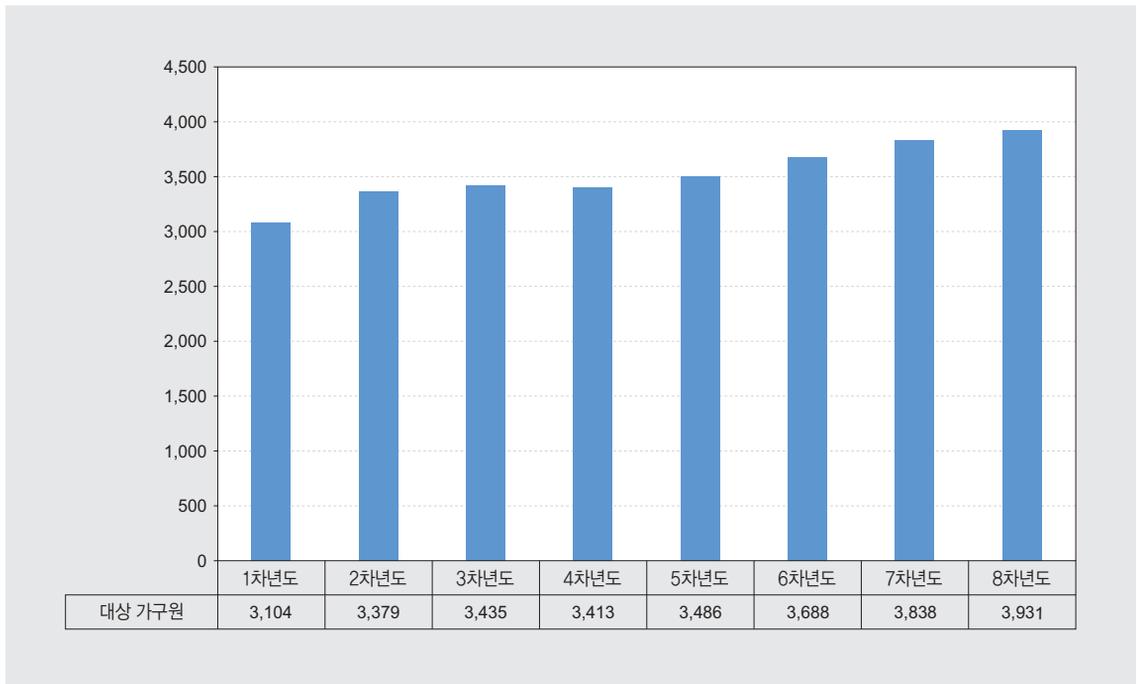
재정패널조사에서 활용하는 소득증빙서류 정보는 가구원 설문지의 연간소득, 소득공제내역, 결정세액이다. 설문지 응답자는 소득증빙서류의 정보에 대한 설문문항에 대해서도 응답을 하지만 이는 최종 응답값을 전산화 하는 과정에서 소득증빙서류의 정보로 대체된다.

### 나. 증빙서류 수집률

재정패널조사의 소득증빙서류 대상자는 근로소득자나 사업소득자 중 실제 소득신고를 한 사람이다. 예컨대, 근로소득자의 경우 연말정산을 하지 않았다고 응답한 근로자는 수집대상에 포함되지 않는다. 소득증빙서류 수집 대상자는 2차년도에 크게 증가하였고, 4차년도에는 그 전년에 비해 소폭 하락하는 추세가 나타났으나, 6차년도부터 다시 그 수가 증가하기 시작하였다. 2차년도에 소득증빙서류 대상자가 증가한 이유는 1차년도에 크게 이탈한 표본을 보완하기 위해 2차년도 조사시점에 620가구를 추가 추출하여 표본을 보완한 영향인 것으로 보인다([그림 2] 참조).

[그림 2] 소득증빙서류 제출 대상 가구원 수

(단위: 명)

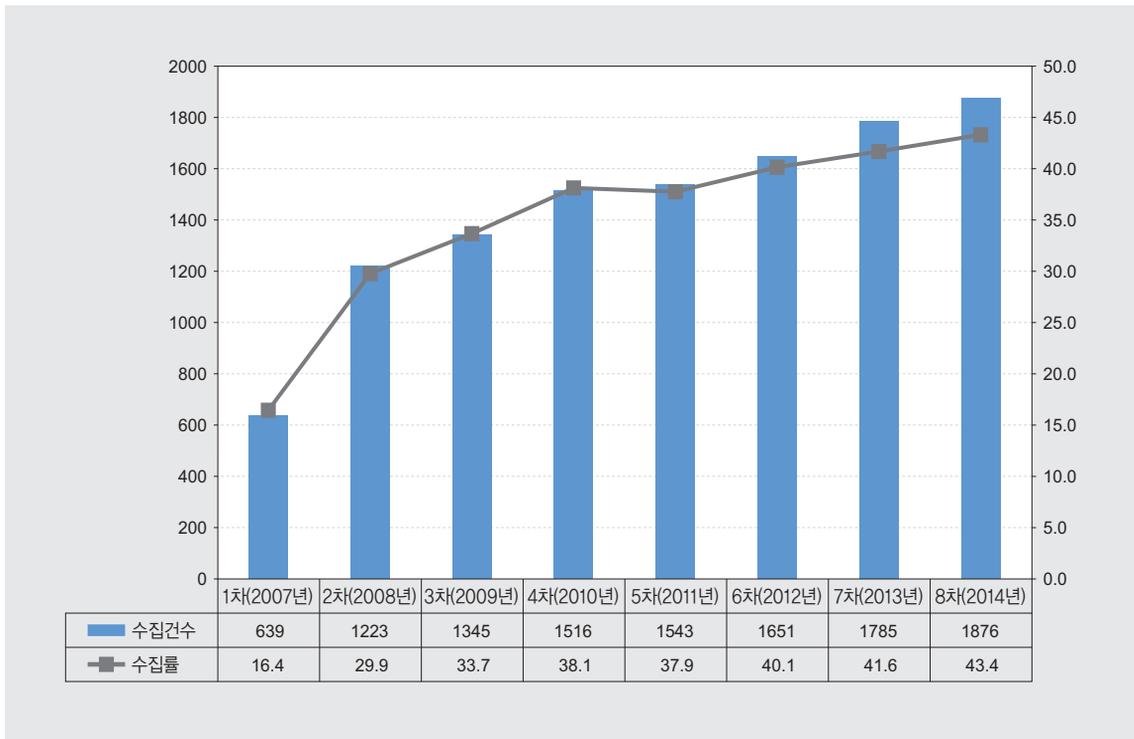


출처: 저자 작성

소득증빙서류 수집률은 1차년도에 26.2%였으나, 2차년도에 47.5%로 크게 증가하였다. 그 이후 증가추세를 계속 유지하고 있으며, 8차년도에는 1,876건의 서류가 수집되어 43.4%<sup>5)</sup>수준으로 증가하였다. 다만 4차년도부터 소득증빙서류 증가율 속도는 다소 감소하고 있다. 소득증빙서류 증가율이 증가하고 있는 이유는 한 응답자에게 동일한 면접원이 설문조사를 지속적으로 함에 따라 면접원과 응답자 간의 친밀감과 신뢰가 형성되었기 때문인 것으로 보인다. 참고로 재정패널조사는 평균적으로 약 90여명의 면접원이 매년 조사에 투입되며 그 중 90% 이상이 동일한 응답자에게 찾아가 설문조사를 실시한다.

[그림 3] (전체 근로소득자 기준) 소득증빙서류 수집률 및 수집건수

(단위: 건, %)



출처: 저자 작성

1) 재정패널 근로소득자 중에는 연말정산을 하지 않았다고 응답한 가구원들이 포함되어 있으며, 이렇게 소득증빙서류(연말정산신고서류) 제출이 불가능 가 구원을 제외한 후 소득증빙서류 수집비율을 산출하면 1차년도 27.0%, 2차년도 47.6%, 3차년도 52.7%, 4차년도 59.5%, 6차년도 59.3%, 7차년도 60.5%, 8차년도 61.3%로 나타남

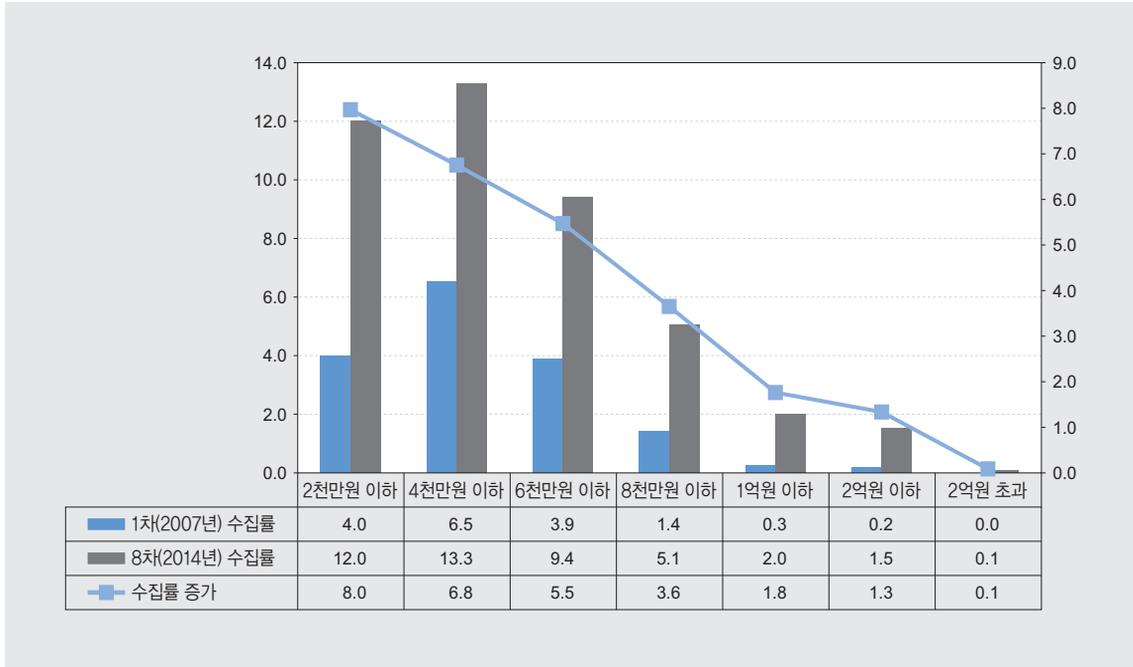
## 2. 소득증빙서류 수집률과 근로소득

재정패널조사에서 사용하고 있는 회상에 의한 조사방식은 응답자들의 조사 거부감을 낮추고 상대적으로 단 시간 안에 응답할 수 있다는 장점이 있는 반면, 조사항목이 세분화되어 있거나 정확한 응답값을 요구하는 조사영역에서는 응답 오차가 발생할 가능성이 높다는 약점이 있다. 재정패널조사에서는 이런 회상에 의한 조사 방법의 문제점을 최소화하고자 소득증빙서류를 수집하여 정보화함으로써 소득관련 응답의 오차를 줄이고자 하였다. 소득증빙서류는 1차년도부터 지속적으로 수집하고 있으며, 소득증빙서류를 제출하는 가구원은 증가하는 추세이다. 이는 재정패널조사가 장기화될수록 면접원과 패널의 친밀도 및 조사 신뢰도 상승하기 때문인 것으로 보인다. 그렇다면 소득증빙서류의 수집률이 높아짐에 따라 소득정보의 응답 오차는 어느 정도 감소하였을까? 이 질문에 정확히 답하기 위해서는 응답자가 설문조사 문항에 작성한 소득과 소득증빙서류상의 소득을 서로 비교하는 것이 필요하다. 하지만 재정패널조사는 소득증빙서류를 제출한 응답자의 경우 소득 정보 등을 소득증빙서류 정보로 대체한 후 응답자가 작성한 소득 정보 등은 전산화되지 않는다. 따라서 본 보고서에서는 국세청이 발간하는 『국세통계연보』의 급여총계(Gross Wage and Salary) 분포와 재정패널조사의 근로소득 분포를 비교하여 응답 오차의 변화를 살펴보고자 한다. 다시 말해, 『국세통계연보』에 제시된 소득 분포가 모집단이라고 보고 재정패널조사의 소득 분포가 모집단의 소득 분포와 비교했을 때 차이가 나타나는지, 그리고 소득증빙서류 수집률이 증가하면서 그 차이가 감소하는지를 살펴보는 것이다. 이 분석의 기본적인 아이디어는 『국세통계연보』의 소득 분포와 재정패널조사의 소득 분포 간의 차이가 응답 오차와 상관관계가 있다고 보는 것이다.

우선 소득구간별로 재정패널의 소득증빙서류 수집률을 살펴보자. 소득증빙서류 수집률이 가장 높은 소득구간은 근로소득 2천만원 초과 4천만원 이하 구간이었다. 이 구간의 1차년도 수집률은 전체의 6.5% 수준이었으며, 8차년도에는 제출자 중 13.3%의 가구원들이 이 구간에 속하는 사람들로 가장 높은 수집률을 보였다. 반면, 수집률의 증가폭이 가장 큰 구간은 근로소득 2천만원 이하 구간이었다. 1차년도에는 4%가 서류를 제출하여 6천만원 이하 구간과 비슷한 수집률을 보였으나, 7년 후인 8차년도에는 수집률이 8%p 증가하여 제출자 중 12%가 2천만원 이하 구간의 가구원들이었다.

[그림 4] (전체 근로소득자 기준) 근로소득 규모별 수집률 변화(2007년, 2014년)

(단위: %, %p)

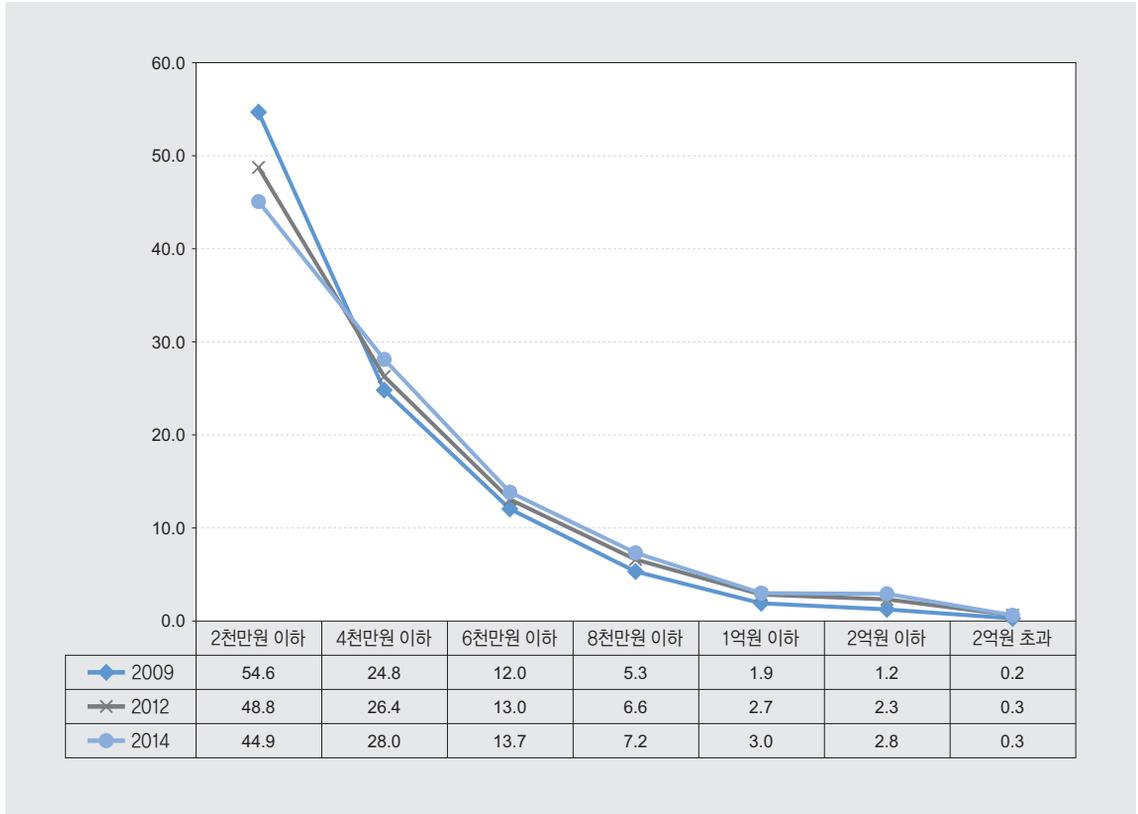


출처: 저자 작성

다음으로 『국세통계연보』 자료의 특징을 간략하게 살펴보고 『국세통계연보』의 소득이 어떻게 분포되어 있는지 살펴보자. 매년 국세청에서는 국세행정에 대한 국민들의 이해를 돕고, 조세정책 수립 및 연구를 지원하기 위해 국세통계자료를 제공하고 있다. 2016년에는 총 418개의 국세통계 항목을 공개하였으며, 원천세, 소득세, 부가세 등 다양한 분야의 통계 정보를 제공하고 있다. 또한 근로소득 연말정산 신고현황 자료를 제공하고 있어 소득규모별 납세대상자 인원과 소득규모별 인원 및 급여총계금액을 파악할 수 있는데, 몇 가지 제약 사항이 있지만 이 정보를 활용하면 재정패널자료의 소득 분포와 비교가 가능하다. 제약 사항이란 다음과 같은 점이다. 먼저, 『국세통계연보』에서 제공하는 소득규모는 귀속년도마다 정의에 차이가 있다. 2007년 귀속자료의 경우 급여총계규모로 소득규모를 정의하였고, 2008년 귀속자료 이후부터는 과세대상근로소득(총급여) 규모로 정의하여 차이가 있다. 다음으로 2007~2008년 귀속자료의 경우 과세미달자의 인원 및 급여총계금액 등은 제공하지 않고 있어, 해당기간 동안 결정세액이 없는 근로소득자의 정보는 확인되지 않는다. 따라서 2008년 이전 『국세통계연보』 정보는 납세대상자 전체의 정보라고 정의하기 어렵다. 이러한 사항을 고려하여 본 장에서는 『국세통계연보』상 2009년 귀속 자료 이후 급여총계 정보를 재정패널자료의 소득 정보와 비교하였다. 먼저 『국세통계연보』 자료의 연도별 소득구간별 납세대상자 인원 분포를 살펴보면 [그림 5]와 같다.

[그림 5] 국세통계연보 소득규모별 납세대상자 인원 분포

(단위: %)

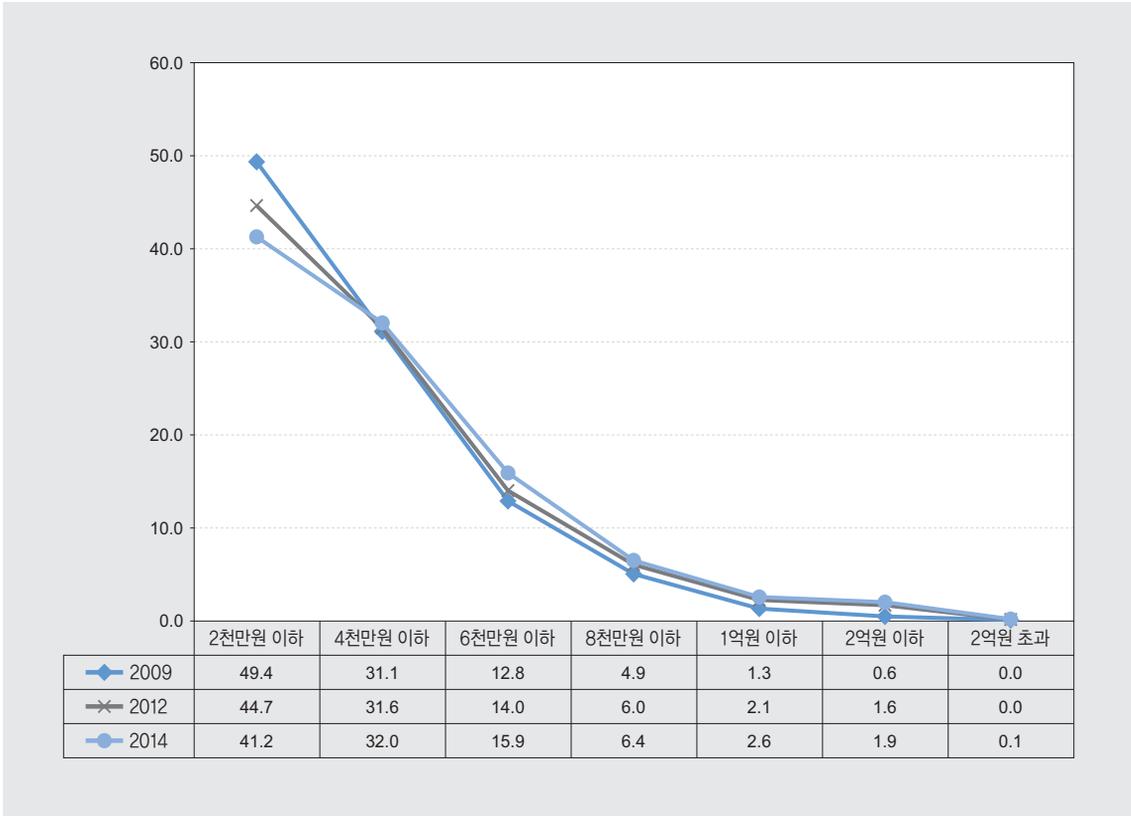


출처: 저자 작성

[그림 5]에 따르면 인원분포가 가장 높은 소득구간은 2천만원 이하 구간으로 2009년 54.6%에서 2014년 44.9%로 9.7%p 감소하였으나 여전히 가장 높은 인원이 분포하는 구간이었다. 2천만원 초과~4천만원 이하 구간의 2009년 인원 분포는 24.8%로 2번째로 높은 비중을 차지하며 지속적으로 증가하여 2014년에는 3.2%p 증가한 28% 분포를 보였다. 동일한 분포를 재정패널 자료로 살펴보면 [그림 6]과 같다. 재정패널자료를 통해 살펴본 소득규모별 근로소득자 인원 분포는 전반적으로 『국세통계연보』의 결과와 유사한 모습이었다. 인원분포가 가장 많은 2천만원 이하 구간과 4천만원 이하 구간의 인원 분포 합은 2014년 기준 73%로, 『국세통계연보』의 72.9%와 매우 유사한 수준이었고, 2천만원 이하 구간 가구원들의 분포가 2009년 49.4%에서 2014년 41.2%로 8.2%p 감소한 것도 비슷한 모습이었다. 다만, 2천만원 이하 구간의 근로소득자 분포가 『국세통계연보』보다 낮고, 4천만원 이하 구간에 해당하는 근로소득자의 분포가 더 많았으며, 이 구간의 연도별 증가폭도 낮은 편(0.8%p)이었다는 차이점이 있었다.

[그림 6] 재정패널 소득규모별 근로소득자 인원 분포

(단위: %)

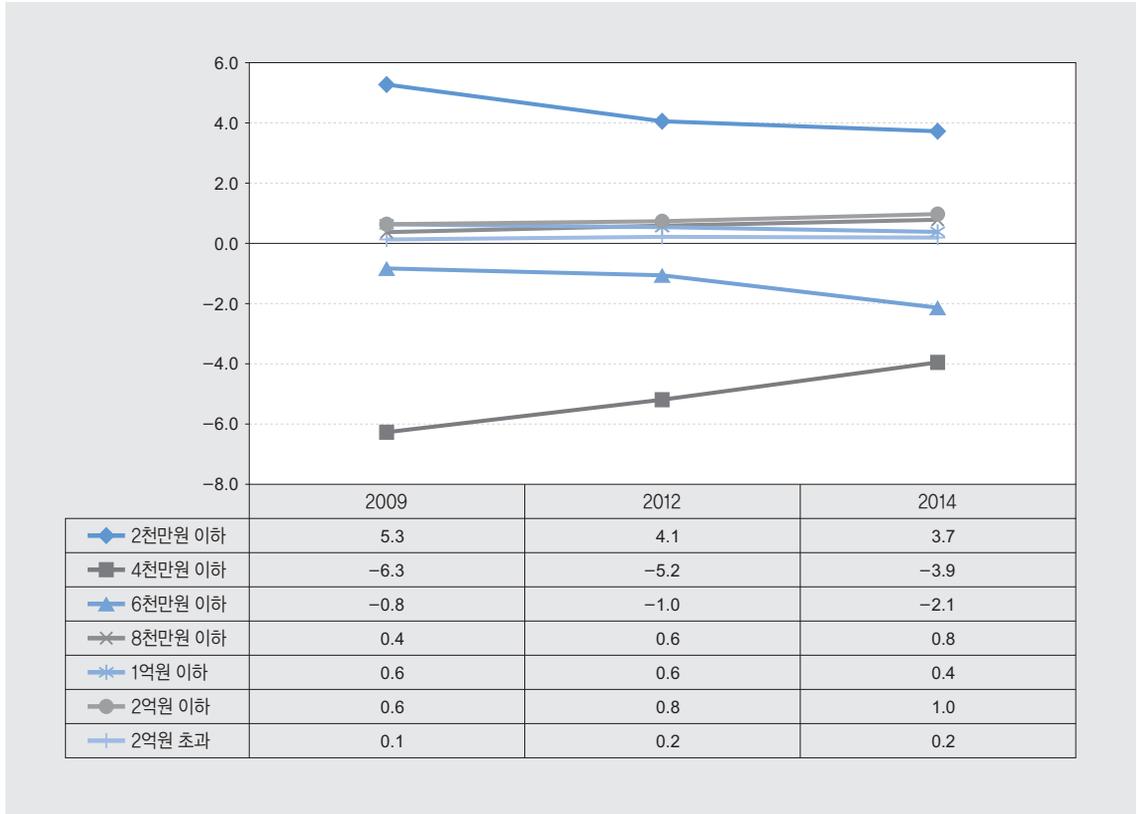


출처: 저자 작성

『국세통계연보』와 재정패널자료 간 소득규모별 인원분포의 차이를 연도별로 비교하여보면 [그림 7]과 같은 결과를 보인다. 먼저, 분석대상기간 중 약간의 차이는 있으나 시간이 지날수록 『국세통계연보』와 재정패널자료의 소득규모별 인원분포 차이는 대부분의 구간에서 0에 가까워지는 모습으로 변화한 것을 알 수 있었다. 특히 증빙서류 수집률이 가장 높게 나타난 구간인 2천만원 이하 구간과 4천만원 이하 구간에서는 그 차이가 눈에 띄게 감소하는 모습을 보여 소득증빙서류 수집에 따른 응답 오차가 감소하고 있음을 알 수 있었다. 다만, 2천만원 이하 소득구간에는 50%에 가까운 소득자들이 분포하고 있는 만큼 소득구간을 세분화하여 살펴본다면 응답 오차 감소정도를 더욱 잘 살펴볼 수 있었을 것으로 판단되나 자료의 한계로 살펴볼 수 없다는 점이 본 연구의 한계점일 것이다.

[그림 7] 근로소득 구간별 인원분포 차이(국세통계연보-재정패널)

(단위: %p)

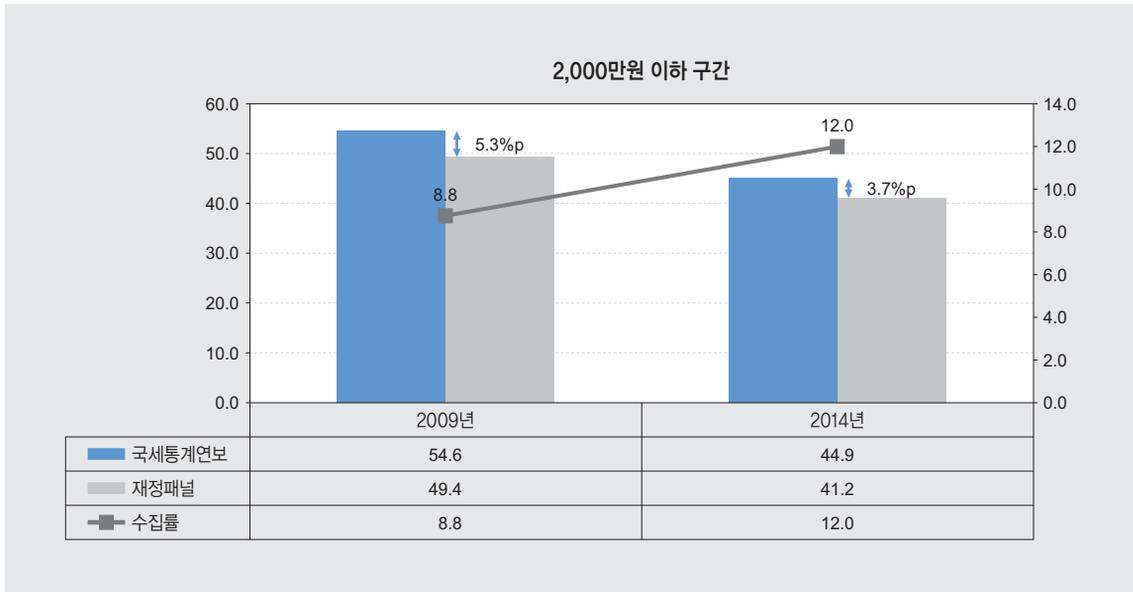


출처: 저자 작성

마지막으로 소득규모를 2천만원 이하, 2천만원 초과~4천만원이하, 4천만원 초과 구간으로 재분류하여 2009년 대비 2014년 인원분포 차이와 소득증빙서류 수집률 정도를 [그림 8]~[그림 10]과 같이 살펴보자. 그 결과, [그림 8]과 같이 근로소득 2천만원 이하 구간의 경우 소득증빙서류 수집률은 2009년 8.8%이었으나, 5년 뒤인 2014년에는 3.2%p가 증가한 12%로 높아졌고, 그 기간 동안의 『국세통계연보』와 인원 분포차이는 5.3%p에서 3.7%p로 소득 분포 격차가 1.6%p 감소한 것을 알 수 있었다. 근로소득 2천만원 초과~4천만원 이하 구간에서는 소득증빙서류 수집률이 0.3%p 증가하고, 인원분포는 3.9%p 차이로 줄어들어 더욱 유사해졌으며([그림 9]), 근로소득 4천만원 초과 구간에서는 소득증빙서류 수집률은 12%에서 18.1%로 6.1%p 증가하였고, 인원분포 차이는 1%p에서 0.2%p로 감소하여 수집률은 증가가 소득정보의 응답 오차를 감소시키는 데 기여했음을 알 수 있었다.

[그림 8] 근로소득 2,000만원 이하 구간 인원분포와 수집률

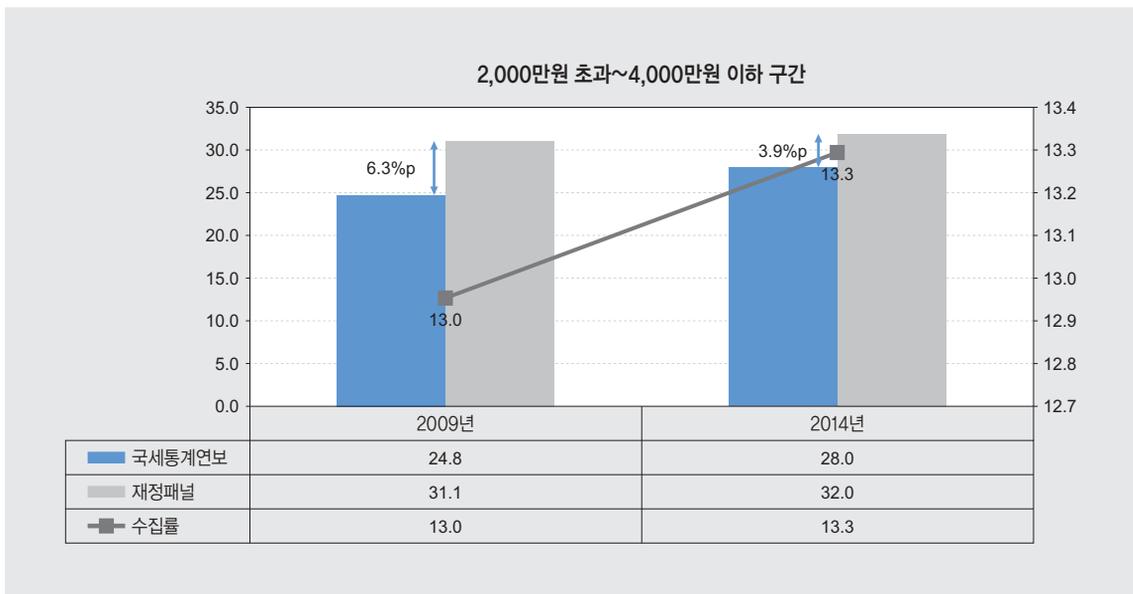
(단위: %, %p)



출처: 저자 작성

[그림 9] 근로소득 2,000만원 초과~ 4,000만원 이하 구간 인원분포와 수집률

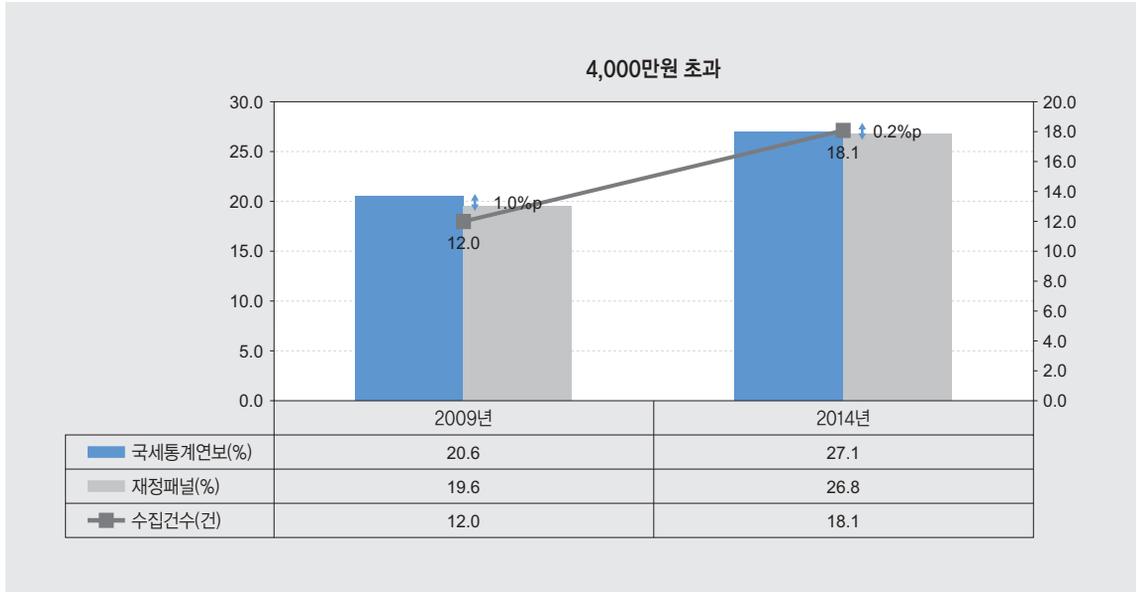
(단위: %, %p)



출처: 저자 작성

[그림 10] 근로소득 4,000만원 초과 구간 인원분포와 수집률

(단위: %, 건, %p)



출처: 저자 작성

### 3. 맺음말

본고는 재정패널조사에서 사용하고 있는 특징적인 조사방법인 소득증빙서류 수집에 대해 설명하고 그 결과에 대해 소개함으로써 재정패널 자료에 대한 연구자들의 이해를 도모하기 위해 작성되었다. 먼저 소득증빙서류 수집률은 1차년도 16.4%에서 계속 증가하여 8차년도에는 전체 근로소득자 중 43.4%가 서류를 제출하였다. 수집률이 가장 높게 나타난 구간은 2천만원 초과~4천만원 이하 구간이었으며, 8차년도 기준 13.3%의 수집률을 보였다. 『국세통계연보』자료와 비교를 통해 인원분포 격차를 살펴보았을 때, 재정패널자료에서 이들 집단의 인원분포가 상대적으로 높게 나타나긴 하였으나 소득증빙서류 제출률이 증가한 8차년도에는 두 자료의 인원분포차가 6.3%p에서 3.9%p로 줄어들어 소득증빙서류 수집이 응답 오차 감소에 긍정적으로 영향을 미쳤을 것으로 판단된다. 그리고 소득증빙 수집률이 두 번째로 높게 나타난 구간은 2천만원 이하 구간으로 1차년도에서 8차년도까지 증빙서류 수집률 증가 수준이 가장 높았다(8%p). 이 구간에서도 『국세통계연보』와 비교한 인원분포 비율의 차이가 5.3%p 격차에서 3.7%p로 감소하여 증빙서류 수집률의 증가가 재정패널조사의 소득 정보에 대한 응답 오차 감소와 관련이 있는 것으로 보인다. 다만, 2천만원 이하 집단의 경우 포함되는 인원이 전체 근로소득자의 약 44.9%(『국세통계연보』, 2014년 귀속 기준)를 차지하는 수준으로 많아 단순한 소득 증가에 의한 소득구간 이동인지, 소득증빙서류 제출에 의한 소득구간 이동인지를 정확하게 판단하기 모호한

부분이 있다. 따라서 근로소득 2천만원 이하에 포함되는 근로자들을 더욱 세분화된 소득규모별로 살펴볼 필요가 있을 것인데, 비교 자료(『국세통계연보』)의 한계로 살펴보지 못한 것이 본 원고의 한계점이라 할 수 있을 것이다.

그럼에도 불구하고, 본고를 통해 소득증빙서류 수집률 증가가 재정패널의 근로소득에 대한 응답 오차 감소와 어느 정도 관련이 있는 것으로 판단할 수 있다. 앞으로도 소득증빙서류 수집률이 지속적으로 증가하여 재정패널조사의 소득 정보가 더욱 신뢰성 있는 정보로 발돋움하길 기대해본다.