

발 간 등 록 번 호

11-1240000-001337-01



가계금융복지조사에서의 조사자료와 행정자료의 통합방법 이해

2020. 6.



통계청
Statistics Korea

이 보고서를 펴내며

많은 국가들에서 통계목적으로 행정자료를 활용하는 사례들이 점점 늘어나고 있다. 이러한 움직임은 공식통계 분야에서 더 빠르게 나타나는 것 같다. 행정자료는 자료수집 비용이나 응답자의 부담을 줄이고 통계자료의 정확성을 높이기 위한 목적으로 널리 사용되고 있다.

가구단위 통계를 중심으로 하는 사회통계분야에 있어서 자료를 수집하는 과정은 점점 더 복잡하고 어려워지고 있다. 자료를 수집하고 통계를 작성하는 방법은 인구, 가구 구조의 변화 및 개인, 가구 특성의 다양화에 따른 사회적 흐름에 영향을 받을 밖에 없다. 이에 대부분의 국가통계작성 기관에서는 저비용, 고효율의 자료수집을 통한 통계생산을 위해 노력하고 있으며, 행정자료는 새로운 통계 작성에 필요한 수단 중 하나로 자리매김하고 있다.

가계금융복지조사는 대표적인 가구단위 통계조사 중 하나로 우리나라 가구의 소득, 자산, 부채 및 지출 현황 파악에 필요한 기초통계의 역할을 하고 있다. 2017년부터 공식 소득분배지표 작성에 활용하게 됨에 따라 자료에 대한 활용도나 관심도가 증가하고 있다. 이에 통계청은 과세자료 등의 다양한 행정자료를 통해 조사오차를 줄임으로써 통계자료의 정확성을 높이고자 하였다. 특히 다양한 출처자료들을 결합하여 활용할 수 있는 방법을 강구함으로써 조사자료에 행정자료를 결합하여 통계를 작성할 수 있게 되었다.

행정자료를 통한 통계작성은 행정자료가 통계목적으로 수집되는 자료가 아니라는 점에서 다양하고 복잡한 자료들을 결합하고 처리할 수 있는 상당한 방법론적 지식과 숙련된 경험이 필요한 분야로 등장하였다. 이에 본 보고서는 통계청이 먼저 경험하고 축적한 노하우를 공유할 수 있는 계기가 된 것 같다.

마지막으로 본 보고서가 가계금융복지조사 자료의 이용자들로 하여금 유용한 정보가 될 수 있기를 희망한다.

통계청 사회통계국 국장 은순현

들어가며

가계금융복지조사는 우리나라 가구의 소득과 자산 및 소득분배지표 통계 작성에 이용되는 통계로서 특히 경제, 사회, 복지 분야와 관련한 정책관련자 및 연구자를 중심으로 폭넓게 활용되고 있다.

본 자료는 가계금융복지조사에 행정자료를 활용하기까지의 검토내용에 대한 주요내용을 포함하고 있다. 본 보고서의 내용은 내부적으로 추진한 여러 편의 연구과제, 국내외 참고자료 및 내부검토 보고서 등을 바탕으로 작성되었다.

2010년을 들어서면서 소득통계 및 소득분배지표에 대한 관심과 함께 통계자료의 품질을 더 높일 수 있도록 노력해야한다는 목소리도 높아졌다. 이에 2013년 이후 행정자료 활용을 위한 연구 등의 노력이 본격화되었다. 그 결과 2017년에는 행정자료로 보완한 가계금융복지조사 기반 소득분배지표 결과를 제공하였고, 2018년에는 가구소득 및 공적지출통계 대해 행정자료로 보완한 결과와 마이크로 자료를 제공하였다.

본 보고서 작성을 위해 많은 참고자료들이 인용되었지만 일부 출처가 누락 되었을 수 있으며, 참고문헌 목록을 별도로 작성하지 않고 필요할 때마다 본문에 제시하는 것으로 같음하였다. 행정자료는 통계법, 국세기본법, 개인정보보호법 등 여러 법령들을 근거로 사용하였고, 소득과 관련하여 10여개 기관의 30여종의 행정자료가 활용되었다.

그 동안 행정자료 활용기준을 마련하기 위해 방대한 자료를 심층분석하고 검토하는데 함께한 복지통계과 홍경희, 권혜지 주무관에게 감사드린다. 그리고 보고서의 분석표들은 연구자료를 분석한 결과로서 통계청 발표수치 및 공식견해와 다를 수 있음을 밝혀둔다.

본 보고서를 통해 행정자료 활용에 대한 경험과 지식을 이용자들과 공유함으로써 이용자들이 행정자료가 통계작성에 어떻게 활용되었는지를 이해하는 데 조금 이나마 도움이 될 수 있기를 기대한다.

복지통계과 과장 김서영



CONTENTS

I. 개요	1
1. 기본배경	3
2. 행정자료 활용에 대한 필요성	3
3. 가계금융복지조사에서의 행정자료 활용	4
4. 조사자료와 행정자료의 결합 활용에 대한 장단점	6
II. 해외사례	9
1. 행정자료 활용에 대한 국제동향	11
2. 과세자료로 조사자료를 대체하는 국가별 사례	12
3. 소득등록부 구축을 통한 행정자료 활용 사례	17
4. OECD 회원국의 소득분배지표 작성 통계 현황	19
III. 행정자료와 조사자료 결합방법	21
1. 개요	23
2. 조사자료와 행정자료의 연계과정	25
3. 매칭편향 보정방법	32
IV. 행정자료의 전처리 과정	33
1. 기본방향	35
2. 소득원천별 행정통합자료 작성 방법	36
V. 행정자료 활용기준 및 적용결과 요약	53
VI. 결론 및 향후방향	63
[참고분석] 조사자료와 행정자료의 특성 분석	67
〈근로소득〉	69
〈사업소득〉	76
〈금융소득〉	83
〈소득세〉	88

PART

I

개요

1. 기본배경
2. 행정자료 활용에 대한 필요성
3. 가계금융복지조사에서의 행정자료 활용
4. 조사자료와 행정자료의 결합 활용에 대한 장단점





I 개요

1 기본배경

- 가계금융복지조사에서는 가구의 자산, 부채, 소득, 지출을 조사하여, 가계의 경제 상황, 재무건전성, 소득분배, 빈곤현황 등과 같은 가구의 전반적인 경제수준, 규모의 변화 및 구조적 특성을 파악한다.
- 이 조사는 연간조사로서 매년 약 2만 가구를 대상으로, 조사원이 가구를 방문하여 가구의 대표자(가구주) 또는 가구의 경제상황을 잘 아는 가구원을 직접 면접하여 조사표를 완성하는 대면조사 방식이다.
- 조사대상 가구는 시설가구를 제외한 전국의 일반가구를 대상으로 한다. 가구는 직업과 학업 때문에 떨어져 살지만 생계를 같이하는 가구원(주말에 본가로 오는 배우자, 미혼 학생자녀 등)을 포함하여 구성한다. 다만, 경제적으로 독립하였거나 결혼하여 따로 사는 성인자녀는 가구원에 포함하지 않는다.
- 가계금융복지조사가 시작되던 2012년을 기점으로 국내에서도 가구소득 및 소득 분배지표에 대한 사회적 관심이 더욱 높아졌다. 이에 행정자료와 같은 보조자료를 통해 조사자료의 정확성을 높이기 위한 노력이 본격화 되었다.
- 행정자료가 조사자료의 대체재라는 점에 이견은 없었던 것 같다. 다만, 행정자료가 통계목적으로 수집된 자료가 아니라서 어떻게 활용할 것인가에 대한 충분한 검토가 선행되어야 했고, 출처가 다른 다양한 행정자료를 결합하여 통계목적으로 활용하기까지는 생각보다 많은 시간이 소요되었다.

2 행정자료 활용에 대한 필요성

- 1 현장조사를 통해 소득 등 민감한 조사항목을 정확하게 측정하는 것이 점점 더 어려워지고 있다.

- 1인 가구 및 맞벌이 가구 증가, 개인정보보호 및 사생활보호에 대한 인식이 높아지는 인구사회적 구조 변화는 현장조사를 더 어렵게 만들고 있다.
 - 응답거부나 응답착오로 인한 응답오차는 자료의 신뢰성을 떨어뜨릴 수 있으며, 행정자료와 같은 대체자료를 통해 측정오차를 줄일 필요가 있다.
- ② 가구소득은 현재 가계동향조사와 가계금융복지조사를 통해 측정되고 있어 자료의 다양성 측면에서는 좋은 점이 있지만, 두 자료 간의 수치적 불일치성 문제 또는 현실적 체감 차이를 해결할 방법을 찾아야 한다는 의견도 있었다.¹⁾
- ③ 가구의 소득조사에서 무응답 또는 과소·과대 응답 현상은 이미 오래 전부터 다른 나라들에서도 겪고 있는 것으로 알려져 있다. 이에 국제적으로도 2000년대 중반부터 행정자료를 통해 조사자료의 품질을 높이고자 하는 노력이 있어왔다.
- 선행사례들의 경우를 보면 과세자료(tax income) 또는 사회수혜금(social benefit) 과 같은 행정자료들이 조사자료를 대체할 수 있는 하나의 대안으로 검토되고 있다.
 - 행정자료는 국가의 정책 또는 행정자료의 양적·질적 수준에 따라 활용정도나 방법이 다르며, 대체로 등록부(register)와 같은 전수자료, 표본조사의 보정(calibration)을 위한 벤치마크, 조사자료와의 통합을 위해 활용되는 것 같다.

③ 가계금융복지조사에서의 행정자료 활용

- 가계금융복지조사의 소득항목은 근로소득, 자영업자(비법인 사업자)의 사업소득, 재산소득, 공적이전소득, 사적이전소득으로 구성된다. 공적지출항목은 소득세, 재산세, 자동차세 및 기타세금 등의 각종 세금과 국민연금, 기타연금(공무원연금 등) 등의 연금기여금 및 건강보험료와 같은 사회보험료로 구성된다.

1) 2013년부터 소득통계에 대한 관심과 통계품질의 개선 필요성에 대한 관심이 크게 증가하였음

- 관련 회의
 - 소득통계 체감차이 원인 및 대응방안 회의('13.5.23)
 - 국가통계위원회 안건 상정('13.10.8): 국민체감과 통계지표 간 차이 해소방안 논의
- 2013년 이후 국회, 언론 등에서 고소득층 및 금융소득이 잘 잡히지 않아 지니계수가 과소하게 추정된다는 현실성 논란 문제 지속
- 국회 등 국세청 소득과 차이가 있다는 지적(국세청 통합소득과 비교) 등

- 소득통계는 소득분배지표와 같은 국가간 비교지표로도 널리 활용되고 있어 국제 사회에서 소득측정에 대한 표준을 제시하고 있다. 국가마다 소득 관련 정책이 다르고 조사환경도 저마다 다르다는 점에서 가구의 소득을 정확하게 측정해내는 것이 상당히 어려운 것이 사실이다. 이에 대부분의 국가에서 소득을 측정함에 있어 OECD등 국제지침*을 따르며 대체로 순소득(net income)으로 측정한다. 우리나라도 OECD의 지침에 가구소득을 측정하고 있다.

* 캔버라 그룹 소득연구회(Canberra group income study, 2011), OECD(2013)

<가계금융복지조사에서의 소득의 구성 및 개념, 2019년 기준>

조사항목		조사포괄 및 행정자료 관련 내용	
근로소득		- 일용근로소득(원천분) - 근로소득(원천/연말/종합분) - 퇴직수당 (퇴직(연금) 일시금 제외) 등	
사업소득		- 자영업자의 사업소득(원천/종합), 기타소득	
재산소득	금융소득	- 이자, 배당 등의 금융소득	
	임대소득	- 부동산을 임대해주고 받는 임대소득	
	기타소득	- 개인연금, 퇴직연금 등	
공적이전소득	공적연금	- 국민연금 - 기타연금(공무원, 사학, 별정우체국, 군인)	
	기초연금	- 기초연금(65세 이상 개인이 정책대상)	
	맞춤형기초생활	- 주거, 생계, 교육급여 등 현금급여	
	장애수당	- 장애연금, 장애수당, 장애아동수당, 장애인급여(지자체 지원)	
	양육수당(출산장려)	- 양육수당, 아동수당 - 출산장려금(지자체)	
	근로자녀장려금	- 근로·자녀장려금	
	기타이전소득	- 실업급여, 재해, 보훈 등	
공적이전지출	세금	소득세	- 소득세(원천/연말/종합분)
		재산세	- 재산세, 종합부동산세
		자동차세	- 자동차세, 환경개선부담금
		기타세금	- 주민세 등 기타세금(양도세, 부동산세 등 제외)
	공적연금·사회보험료	국민연금기여금	- 국민연금 기여금
		기타연금기여금	- 공무원, 사학, 별정우체국, 군인연금의 각 기여금
		건강보험료	- 건강보험료, 노인장기요양보험료
		고용보험료	- 고용보험료

- 우리나라는 소득과 관련하여 과세자료(국세청), 연금소득자료(각 연금공단), 건강보험료(건강보험공단) 및 사회수혜금(보건복지부 등) 등의 소득과 관련한 다양한 행정자료를 검토해 볼 수 있다.
 - 과세자료에는 원천징수, 연말정산, 종합소득신고에 의한 신고자료가 있으며, 사회수혜금(social benefits)자료에는 일정시점에 각 정책대상자에게 지급되는 현금급여가 있다. 소득 및 소득세 자료는 개인단위의 연간자료이며, 전년도 소득에 대해 연말정산(원천징수 포함)하거나 종합소득으로 신고한다. 따라서 신고자료의 경우 조사자료에 비해 응답오차 측면에서는 정확하다 할 수 있다.
 - 가계금융복지조사는 개인이 지난 한 해 동안에 벌어들인 연간 소득을 측정한다는 점에서 과세자료와 유사한 자료구조를 하고 있으므로 행정자료에 접근하기가 상대적으로 수월하다 하겠다.
- 이탈리아, 프랑스, 영국, 캐나다, 미국(CPS, ACS), 호주 등 많은 국가들에서 조사자료와 행정자료를 통합하여 활용하는데 관심이 많으나 두 자료의 연계방법에 대해서도 많은 연구들이 진행 중인 것으로 파악하고 있다.

* 미국 CPS(공식지표 작성)에 사회수혜금자료 개인단위 대체방법 연구(14~16년)

** CPS: current population survey, ACS: american community survey

4 조사자료와 행정자료의 결합 활용에 대한 장단점

- 조사자료와 행정자료의 통합(integrating of survey and administrative data)은 자료의 정확성, 일관성, 완전성 등의 측면에서 장점이 있다.
 - (정확성) 무응답 및 측정오차(과소오차, 기억오차, 텔레스코핑 효과*)를 줄여서 조사자료의 품질향상에 기여한다.
 - * 최근의 일을 더 과거일로 기억하거나, 반대로 과거의 일을 최근의 일로 인식하는 것
 - (일치성) 동일한 정의, 분류 및 개념적 틀을 사용할 경우에 보다 일관성 있는 자료를 생산할 수 있다.
 - (완전성) 조사의 무응답을 대체함으로써 모집단에 대한 포괄률을 높일 수 있다. 이 외에도 행정자료 활용으로 조사항목을 간소화할 수 있으며, 이를 통해 조사비용, 응답부담 및 현장조사의 비표본오차도 줄일 수 있다.

- 한편, 행정자료는 입수시점에 대한 제약이 있고, 행정목적으로 수집된 자료라는 점에서 통계목적으로 수집되는 자료의 특성과는 다를 수도 있다. 따라서 행정자료의 품질 관리 차원에서 행정제도 및 행정자료의 구조변화를 꾸준히 모니터링해야하는 부담도 따른다.
- 통계는 다단계의 자료처리과정을 거쳐 작성되므로 각 처리단계별로 행정자료 활용과 관련하여 명확한 자료처리기준이 필요하다. 두 자료를 결합할 때 생기는 새로운 현상들을 종합적으로 분석할 수 있는 경험적 요소도 중요하다.

<조사항목별 해당 행정자료 항목>

조사소득항목		입수한 행정자료	입수기관	
근로소득		- 일용근로소득(원천분) - 근로소득(원천/연말/종합분) - 퇴직소득 (퇴직일시금과 분리 불가) - 자활급여	국세청 보건복지부	
사업소득		- 사업소득(원천/연말/종합분) - 기타소득*	국세청	
재산소득	금융소득	- 이자, 배당 등 금융소득(원천/종합분)	국세청	
	임대소득	- 임대소득(임대사업 등록자 자료)		
공적이전소득	공적연금	- 연금소득	각 연금공단	
	기초연금	- 기초연금	보건복지부	
	맞춤형기초생활	- 주거급여, 생계급여, 교육급여, 해산급여, 장제급여	보건복지부, 국토부, 교육부	
	장애수당	- 장애연금, 장애수당, 장애아동수당, 장애인급여	보건복지부	
	양육수당(출산장려)	- 양육수당, 아동수당	보건복지부	
	근로자녀장려금	- 근로자녀장려금	국세청	
공적 이전 지출	세금	소득세	- 소득세(원천/연말/종합분)	국세청
		재산세*	- 재산세, 종합부동산세**	행안부
		자동차세*	- 자동차세, 환경개선부담금**	행안부
	공적연금· 사회 보험료	국민연금기여금	- 국민연금	국민연금공단
		기타연금기여금	- 공무원, 사학, 별정우체국, 군인연금	각 연금공단
		건강보험료	- 건강보험료, 노인장기요양보험료	국민건강보험공단
		고용보험료*	- 고용보험료	근로복지공단

* 입수자료를 검토한 결과 자료구조 및 개념 불일치로 인해 활용하는데 한계가 있음

** 조사개념에 포함되는 자료이나 분석당시 행정자료를 입수하지 못함

PART
II

해외사례

1. 행정자료 활용에 대한 국제동향
2. 과세자료로 조사자료를 대체하는 국가별 사례
3. 소득등록부 구축을 통한 행정자료 활용 사례
4. OECD 회원국의 소득분배지표 작성 통계 현황



PART



II 해외사례²⁾

1 행정자료 활용에 대한 국제동향

- 행정자료와 조사자료의 결합사례에서는 가구는 민감한 소득항목에 대해 응답하지 않거나 사실과 다르게 응답하는 경향이 있으므로 소득통계의 품질 향상을 위해 행정자료를 활용할 필요가 있다고 강조하고 있다. 과세자료, 사회수혜금 등을 좋은 대안자료로 꼽고 있으며, 조사자료에 행정자료를 결합한다는 것은 행정자료로 조사자료를 수정·대체한다는 것을 말한다.
- 조사자료와 행정자료를 통합하여 통계를 작성할 경우 소득 및 소득분배지표의 수준이 변동될 수 있으며, 조사자료와 행정자료를 정확하게 결합하기 위해서는 적합한 연계키(matching key)를 찾는 것이 중요하다.

<조사자료와 행정자료의 연계 활용 사례별 현황>

	주요방향	매칭키	특이사항
이탈리아 (SILC)	· 조사자료의 품질 개선 · 행정자료의 결합 - 연금소득, 사업소득, 임금소득 · 결측값 또는 신뢰할 수 없는 값 대체 (큰값으로 수정)	정확매칭키(PTN)	- 평균소득 10% 증가 - 지니계수 5.2% 감소 (0.330→0.313) - 빈곤율 1.8%p 감소
프랑스 (SILC)	· 측정오차 축소 및 조사품질 개선 · 행정소득이 더 정확하다고 가정 · 통계법 개정 등 행정자료 활용권한 확보	텍스트 키 (성별, 이름 등)	- 처분가능소득 15% 증가 - 지니계수 12.8% 증가 (0.39→0.44) - 빈곤율 3.8% 감소
오스트리아 (SILC)	· 응답자 부담감소 · 행정자료가 더 객관적 & 소득분포의 양극값을 더 커버하는 결과	정확매칭키(PIN)	- 행정근로소득이 조사자료보다 3.3% 더 높음

2) 본 사례는 국내외 여러 보고서를 포함한 참고문헌 등의 행정자료 활용과 관련한 부분을 참고하여 작성하였다.

· The use of registers in the context of EU-SILC: challenges and opportunities(Eurostat working paper, 2013), Family resource survey(FRS), Canadian Income survey(CIS), American community survey(ACS), Current population survey(CPS), Statistics on income and living conditions(SILC), 해외출장보고서(복지통계과, 2018), 복지통계과 내부(2017)

		주요방향	매칭키	특이사항
영국 (FRS)		· 조사에서 사회수혜금 자료에 대해 과소 응답하는 경향 확인 · 사회수혜금 자료의 정확성 향상	텍스트 키 (성별,이름,생년월일, 주소)	- 정보활용 동의 구함 · 60% 동의 및 전체 대상자의 57% 매칭
캐나다 (CIS)		· 과세자료 활용에 대해 면접시 공지. 조사에서는 최소한의 질문	텍스트 키	- 조사대상자의 89%에 대해 과세자료 활용
미국	CPS	· 항목무응답, 과소응답 대체	정확매칭키(PIK)	- 가구별 91% 매칭률 - 빈곤율 감소효과: 5~6%P
	ACS	· 행정자료로 조사항목 수정·대체 - 소득자료 이용(14~16년)연구	정확매칭키(PIK)	- 거처별: 927%, 개인별 921% 매칭률(10년)

* PTN: personal tax number, PIN: sector specific personal identifier, PIK : personal identification key

2 과세자료로 조사자료를 대체하는 국가별 사례

1 덴마크 1994년부터 Household budget survey(HBS)를 통해 소득, 지출, 실물 자산, 연금, 직접세, 수입생산세, 건강, 교육, 사회적 현물 등을 조사한다.

- 가구의 응답부담을 줄이기 위해 행정자료를 이용하여 조사자료를 대체·보완할 수 있는 항목은 조사항목에서 제외하고 조사한다. 덴마크는 모든 개인에 대해 PIN(Personal identifier number)이 있으며, 가구조사에서 PIN을 직접 활용하기 어려운 경우에는 거처주소를 이용하여 조사자료와 행정자료를 연계한다.(인구등록자료는 거처 주소와 개인 PIN를 모두 보유)
- 소득항목, 지출항목 및 세금항목에 대해 조사자료와 행정자료를 연계하여 소득등록부(income register)를 만들고 이를 통해 소득자료를 대체한다.
 - 다만, 자영업소득, 자가소비용 생산에 의한 소득, 금융소득, 자가주거비, 사적 이전, 사회보험료, 유산, 연금수령액, 보험료, 현금지불액, 벌금, 경상이전, 소비지출 등 행정자료 대체가 어려운 항목에 대해서는 직접 조사한다.

② (캐나다) 1997년 이후 가구조사(CIS)에서 혼합모드(mixed mode)로 자료를 수집한다.

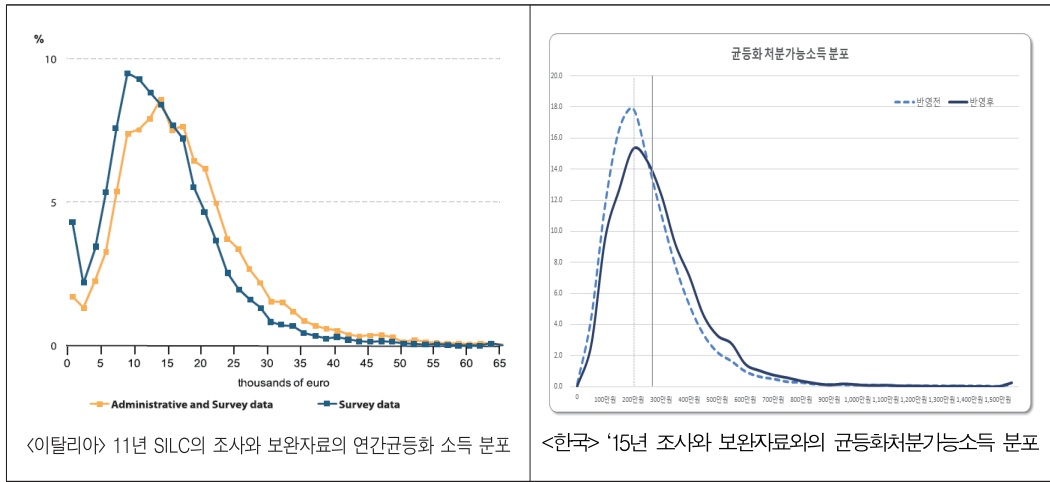
- 캐나다 통계청은 표본가구에 대한 과세자료를 통해 조사 소득자료를 대체하고 있으며, 과세당국에 신고되지 않은 소득 원천에 대해서는 면접조사를 통해 파악한다.
- 통계법(캐나다 Act)에 근거하여 행정자료를 사용한다. 통계청은 표본가구에 행정자료를 사용한다는 것에 대해 공지하며, 행정자료 활용을 동의하지 않은 경우에는 조사자료를 활용한다. CIS 조사 표본 중 약 80%가 행정자료 활용에 대해 동의한다고 보고된 바 있다.

③ (이탈리아) EU-SILC* 조사에서 행정자료와 조사자료를 결합하여 활용한다.

* European Survey of Income and Living Conditions

- 과세자료와 조사자료를 개인단위로 연계하여 두 자료를 결합한다. 2004년에 사업소득과 연금소득에 대해 행정자료를 사용하였고, 2005년에는 근로소득을 행정자료와 결합하여 소득통계를 작성하였다. 현재는 근로소득과 사업소득은 조사자료와 행정자료를 통합하여 사용하고, 연금소득과 일부 사회적 이전소득(social transfers, 실업급여, 퇴직금 및 교육급여)은 행정자료로 완전대체되어 2016년부터 현장조사에서 이 항목은 조사하는 않는다.
- 행정자료는 결측값을 대체하거나 이상치를 수정하거나 하는 등 신뢰할 수 없는 조사값을 보완하는데 사용된다. 두 자료의 결합은 Ptx(Personal tax number)라는 매칭키에 의한 정확매칭법을 사용한다.
- 이탈리아는 사업소득과 근로소득에 대해서 조사자료와 행정자료 중 큰 값을 사용한다. 행정자료로 조사자료를 보완할 경우, 소득평균이 10% 증가하였고, 상대적 빈곤율은 1.8%p 감소, 지니계수는 1.6%p 감소하였다.

< 이탈리아와 가금복의 보완 전과 후의 균등화소득 분포 >



④ (프랑스) EU-SILC를 통해 소득분배지표를 작성한다. 이 조사에서는 현장조사에 의한 소득조사에서 발생하는 측정오차와 조사품질을 관리하기가 어렵다고 판단하여 행정 자료를 활용하고 있다.

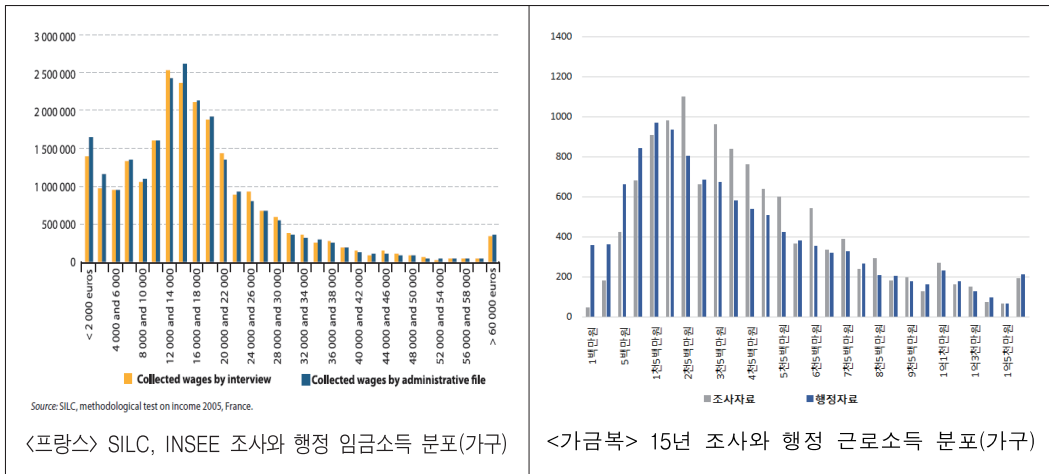
- 프랑스는 소득 행정자료를 활용할 때, 국세청에 신고하는 자료가 조사자료보다 더 정확하고 시점별로도 더 균질한 품질을 갖는다고 가정하였다. 프랑스는 행정자료를 통계목적으로 사용하기 위해 통계법을 개정하는 등의 제도적 장치를 마련함으로써 통계청이 각종 행정자료를 활용할 수 있도록 하였다.

- SILC 조사자료에 행정자료를 연계할 때 개인별 이름, 성별, 생년월일, 주소 등의 개인정보를 매칭키로 사용하였다. 매칭키를 통해 조사대상 가구의 98%가 연계가 된다고 말하고 있다. 다만 주소변경, 가족과 떨어져 사는 학생 중 일부는 조사 자료와 행정자료 간 연계가 실패하는 경우도 있다고 한다.

- 프랑스는 행정자료를 통해 근로소득은 보다 넓은 범위의 근로소득을 포착하거나 표본조사에서 무응답한 사람들의 소득도 파악할 수 있다고 하였다. 한편 사업소득은 두 자료에서 정의하는 대상이 달라 조사자료를 통해 사업소득을 측정하는 것이 더 타당하다고 하였다. 사업소득 자료의 상황은 우리와 비슷한 것 같다.

- 행정자료로 조사자료를 대체한 후, 2007~2008년 처분가능소득은 15% 증가하였고, 지니계수는 0.39에서 0.44로 수준이 올랐으며, 상대적 빈곤율은 13.1%에서 12.6%로 감소하였다.

< 프랑스와 가금복의 근로소득 분포 >



5] (네덜란드) 1977년부터 조사대상자의 과세자료를 통해 소득자료를 수집한다.

- Social Statistical Database(SSD)라는 등록기반 자료시스템을 구축하여 자료 보정 또는 표준화, 여러 마이크로자료를 통합하는데 사용하고 있다. SSD는 개인, 가구, 직업, 사회보장수혜, 연금, 조수 및 거체에 관한 마이크로자료를 포함하는 등록기반 시스템이다. SSD는 'Record Identification Number'라는 개인식별 대체번호를 포함하고 있어, 출처가 다른 자료끼리 연결할 수 있는 핵심키 역할을 한다.
- 네덜란드의 경우 2001년 이후 소득 및 자산에 관한 전체 국민의 행정자료를 이용할 수 있으나, 여전히 표본조사결과를 통해 소득통계를 작성하고 있다. EU-SILC는 네덜란드의 소득통계 작성에 사용되는 기초자료로서, 조사대상자에게 행정자료가 활용될 수 있음을 안내하며 사전안내장(advance letters)을 통해 응답자가 응답한 자료가 행정자료와 같은 다른 자료소스에 결합되어 사용될 수 있음을 공지하고 있다.

- 조사자료와 행정자료 간의 연계는 주소, 성, 생년월일 등의 변수를 조합한 텍스트 정보를 이용한다. 과세자료가 있는 경우의 소득자료는 현장조사하지 않고 행정자료로 완전대체하고, 과세정보가 없는 경우에는 추정한다. 가구이전소득이나 임대소득과 같이 행정자료를 없는 경우에 대해서는 조사를 통해 자료를 수집한다.

⑥ (슬로베니아) SILC 조사에서 대부분의 소득 항목은 행정자료로 대체하고, 농업관련 자영업소득 등 일부 항목에 대해서는 현장조사한다.

- 슬로베니아 통계청은 통계법(National Statistics Acts)에 근거하여 행정자료 및 등록부 자료를 사용한다. 현장조사를 통해 수집한 개인별 이름, 성, 생년월일 등의 텍스트 정보를 조합하여 개인식별번호 PIN(personal identifier number)을 찾고, 이 PIN을 조사자료와 행정자료를 결합하는데 사용한다.
- 슬로베니아 SILC의 주요내용은 소득에 관한 것이다. 대부분 종류의 소득은 행정자료를 통해 수집되면 몇 개의 소득종류에 대해서만 조사표를 통해 수집한다. 이처럼 대부분의 소득항목이 행정자료를 통해 수집되기 때문에 행정자료의 품질에 많은 관심을 갖는다.
- 슬로베니아의 경우 소득은 개인의 고용상태와 밀접한 관련이 있으며 정확한 자료를 수집하기 위해 전혀 다른 행정자료를 통해 수정보완할 수 있어야 한다는 점을 강조하고 있다. 슬로베니아는 EU-SILC뿐만 아니라 지금은 HBS(Household budget survey)에서 행정자료를 활용하는 등 점차적으로 통계작성에 행정자료의 활용을 확대해가고 있다.

⑦ (영국) FRS(Family resources survey)에서 근로·자녀장려금 등 저소득층이 받는 사회수혜금(Social benefit)자료에 대해 행정자료로 대체한다.

- FRS의 응답자들은 사회수혜금에 대해 과소하게 응답하는 경향이 있다고 판단하고 관련 조사항목에 대해 행정자료로 대체하고자 하였다.

- 이 조사에서는 행정자료를 활용한다는 사실을 조사과정에 응답대상자에게 적극적으로 알린다. 조사원이 면접하는 과정에 행정자료를 활용하는 것에 대해 동의하는지 여부를 묻는다. 이 질문은 조사표의 모든 항목에 대한 질문이 끝난 후 즉, 면접 조사를 완료하기 직전에 한다고 한다. 이 조사에서는 행정자료 활용 동의여부를 묻는 것 자체가 응답률 하락에 영향을 주지 않았다고 한다.
- 조사자료와 행정자료는 이름, 생년월일, 주소 간의 함수관계와 6가지 규칙에 의해 연계된다. 한편 영국은 행정자료 활용에 대해 ‘동의자와 비동의자’, ‘매칭키를 통해 두 자료가 연계된 사람과 그렇지 않은 사람’ 간에 차이가 있다고 가정하고 이로 인한 편향을 줄이기 위한 방법도 검토한 것으로 파악되었다.

3 소득등록부 구축을 통한 행정자료 활용 사례

- 소득 관련 행정자료를 등록부 또는 데이터베이스와 같은 형태로 구축하여 소득통계를 작성하거나 다른 통계를 위한 벤치마크로 사용하고 있다. 대체로 행정적 신고체계가 잘 되어 있거나 복지 시스템이 잘 갖춰진 나라에서 행정자료를 통계 목적으로 활용 하는데 관심이 높은 것 같다.

(1) 유럽 국가의 등록부 구축

- (노르웨이) 소득등록부를 만들고 이를 통해 소득통계 작성한다.

- * 소득 등록부(Income Register) 구성을 위한 소득원천
- 소득 및 연금자료 : Tax Returns
 - 세금 : Tax Register
 - 실업급여, 면세자료 : Certificate Register
 - 가족수당, 현금급여 : Labor and Welfare Register
 - 사회수혜금 : Ministry of Labor and Social Inclusion
 - 학자금대출, 장학금 : State Educational Loan Fund 등

- (덴마크) 소득관련 모집단을 구축하고 이를 통해 소득통계를 작성한다.

* 소득모집단(Danish Law Model System)에 포함된 소득원천

- 임시소득 : A Income Statistics
- 가족소득 : Family Income Statistics
- 개인소득 : Tax-related Income Statistics

- (네덜란드) 2001년 이후 소득 및 자산에 대한 과세 행정자료를 통해 전체 국민에 대한 행정자료 등록부를 구축하여 활용하고 있다.

(2) 미국의 SCF 조사의 경우 과세모집단(tax returns)을 통해 표본을 추출하고 사후 보정(calibration)한다.

- SCF*는 2가지 방법(List 추출과 Area 추출)으로 표본을 추출하고 각 방법으로 추출한 두 표본을 결합하는 방식으로 최종 표본을 추출한다.

* Survey of consumer finances

- List추출은 IRS(Internal revenue service) 모집단을 활용하는데, 이때 표본을 추출하기 위해 소득수준을 고려하여 7개 층*으로 나눈다. 순서적으로 소득이 높은 6번째 층(5천만달러 이상)과 7번째 층(3억달러 이상)은 과대추출한다. 이렇게 상위 2개 층에서 과대추출하는 이유는 고소득가구에 대한 표본의 대표성을 높이기 위해서다.

- 2016년의 경우 Forbes라는 잡지에 수록된 400명은 고소득자에 대한 표본추출 명부(list)에서 제외되었다.

* 인용: Federal Reserve BULLETIN(2017.9, 13권, 3호)

- area 확률추출은 지역별 랜덤추출 방식이다. 이 방법은 표본추출시 모집단 커버리지를 높이기 위한 일반적인 방법이다. 2016년 조사의 경우, 6,254 응답가구 중 area 표본에서 4,754가구, list 추출에서 1,500가구가 응답하였다.

- (3) 캐나다의 SFS의 경우 행정자료는 자료대체, 표본추출, 사후조정에 사용된다.
- SFS에서는 자산, 부채는 직접 조사하나, 소득자료는 조사하지 않고 과세자료와 결합하여 활용한다. * Survey of financial security
 - 표본설계는 층화다단계 Dual frame(Stratified multiple-stage Dual frame) 추출방식으로, 이때 T1FF라는 개인 소득세 자료를 이용하여 추출한다.
 - T1FF는 순자산에 의해 층화된 프레임이며, 이 프레임을 사용하는 것은 이 조사가 가족 순자산을 잘 대표할 수 있는 표본추출을 목표로 하기 때문이다. 그러나, T1은 주소불명 등 소득자의 정보가 명확하지 않고, 특히 시골지역(rural area)에서 주소 불명자가 많기 때문에 도시(urban area)지역의 표본추출에는 제한적으로 활용한다.
 - T4는 급여지불현황 자료이며 이 자료는 사후조정에 사용된다. 전체 자료를 급여크기순을 고려하여 25%, 50%, 65%, 75%, 95% 구간으로 나누어 가장 하위 구간은 제외한다. 경우에 따라서는 가장 상위구간은 98%, 99%(표본크기가 충분히 큰 지역에 한함)로 세분화한 후 각 구간의 인구수를 활용하여 고소득층 분포를 조정하는데 활용하기도 한다.
 - 자료들 간의 매칭작업은 보통 4개월 가량 소요되며 매칭이 안된 경우는 자료의 완성도를 높이기 위해 임putation) 처리한다.

4 OECD 회원국의 소득분배지표 작성 통계 현황

- OECD는 36개 국가의 소득분배지표를 'OECD 소득분배 DB'를 통해 제공하고 있다. 각 국가의 소득분배지표 작성에 사용되는 기초 소득자료는 다음과 같다.
- 지표작성에는 연간단위 소득자료가 사용된다. 국가에 따라 조사자료, 행정자료 및 결합자료가 사용되며 덴마크, 노르웨이는 다양한 행정자료로 구축된 소득DB와 같은 등록부자료를 사용하기도 한다.

국가명	조사명	조사주기	자료수집 방법
호 주	Survey of Income and Housing(SIH)	2년	조사
오스트리아	EU-SILC	매년	조사+행정
벨기에	EU-SILC	매년	조사
캐나다	Canada Income Survey(CIS)	매년	조사+행정
칠레	National Characterization SocioEconomic Survey(Casen)	2년,3년	조사
체코	EU-SILC	매년	조사
덴마크	The Danish Law Model System	매년	행정
에스토니아	EU-SILC	매년	조사+행정
핀란드	EU-SILC	매년	조사+행정
프랑스	Enquête Revenus Fiscaux et Sociaux(ERFS)	매년	조사+행정
독일	German Socio Economic Panel(SOEP)	매년	조사
그리스	EU-SILC	매년	조사
헝가리	EU-SILC	매년	조사
아이슬란드	EU-SILC	매년	조사+행정
아일랜드	EU-SILC	매년	조사+행정
이스라엘	Households Expenditure survey	매년	조사
이탈리아	EU-SILC	매년	조사+행정
일본	Comprehensive Survey of Living Conditions	3년	조사
한국	Survey of Household Finances and Living Conditions	매년	조사+행정
라트비아	EU-SILC	매년	조사+행정
룩셈부르크	EU-SILC	매년	조사
멕시코	Household Income Survey(ENIGH)	2년	조사
네덜란드	Income Panel Survey	매년	조사+행정
뉴질랜드	Household Economic Survey(HES)	매년	조사
노르웨이	Income and Wealth Statistics for households	매년	행정
폴란드	EU-SILC	매년	조사
포르투갈	EU-SILC	매년	조사
슬로바키아	EU-SILC	매년	조사
슬로베니아	EU-SILC	매년	조사+행정
스페인	EU-SILC	매년	조사+행정
스웨덴	Income Distribution Survey(HEK)	매년	조사+행정
스위스	SILC	매년	조사+행정
터키	SILC	매년	조사+행정
영국	Family Resources Survey(FRS)	매년	조사
미국	CPS ASEC	매년	조사

* 출처: OECD.Stat(<http://stats.oecd.org>) Income Distribution Database, 국가별 메타데이터('19.9)

* 조사변경국가: (한국) 가계동향조사(~'15년), (캐나다) Survey of Labour and Income Dynamics(~'11), (이스라엘) Income Survey(~'11)

PART

III

행정자료와 조사자료 결합방법

1. 개요
2. 조사자료와 행정자료의 연계과정
3. 매칭편향 보정방법



PART



Ⅲ 조사자료와 행정자료의 결합

1 개요

1] 행정자료의 특성

- 행정자료는 다양한 출처의 행정자료들끼리 결합하거나, 조사자료와 통합하거나 또는 벤치마크(benchmark)와 같은 기준점(gold standard)으로 사용하거나 하는 등의 다양한 방식으로 통계작성에 이용될 수 있다.
- 행정자료와 조사자료의 통합 활용은 조사효율성 및 조사자료의 정확성 제고 측면에서 상당히 매력적일 수 있다. 다만 두 자료를 제대로 통합하기 위해서는 적절한 결합 기준을 찾는 것이 무엇보다 중요하다. 이는 어떤 변수의 경우 조사자료와 행정자료에서 측정하고자 하는 내용과 범위가 서로 다를 수 있기 때문이다.
 - 예를들어, 한 조사표내에서 측정하는 여러 소득 항목 중에도 어떤 항목은 행정자료로 완전히 대체할 수 있는 반면에 어떤 항목은 행정자료가 없거나 포괄 범위가 서로 달라서 행정자료를 부분적으로 활용하거나 아예 활용할 수 없는 경우도 있다.
- 행정자료가 조사자료의 대안이 될 수 있다는 점에 공감하지만 행정자료를 활용할 때 행정자료가 통계목적으로 수집된 것이 아니라는 점에 주목할 필요가 있다. 즉, 행정자료를 활용함으로써 조사자료의 측정오차를 줄이고 현장조사의 어려움을 극복할 수 있다는 장점이 있지만, 신고오류, 포함오류, 개념오류 등과 같이 행정자료에 있을 수 있는 새로운 오류들이 더해질 수도 있다.
 - 행정자료 활용으로 통계자료의 총오차(total error)가 줄어든다고 확신할 수는 없겠지만, 통계의 생산성 및 정확성 측면에서 기여하는 바는 크다 할 것이다.

<왜 행정자료에 관심을 가지게 되는가>

- (1) 현장조사를 통해 민감한 조사항목을 정확하게 측정하기가 점점 어려워지고 있다.
 - 표본가구(또는 개인)를 직접 대면조사하는 경우, 소득, 자산과 같은 민감한 응답항목에 대해 가구는 일부 항목에 대해 응답을 거부하거나 응답을 하더라도 어떤 이유에서든 정확한 응답을 회피하려는 경향이 커지고 있다
- (2) 인구사회적 환경변화와 개인성향의 변화가 자료수집방법에 대한 변화를 요구하고 있다.
 - 인구고령화, 1인 또는 맞벌이 가구화가 빨라지고 사생활과 개인정보에 대한 보호의식이 강해지면서 다양한 소스를 통해 자료를 수집하는 것이 필요해졌다.
- (3) 행정자료가 다양해지고 자료의 품질이 좋아지면서 현장조사를 하지 않고 자료를 수집할 수 있는 기회가 점점 많아지고 있다.

② 조사자료와 행정자료를 결합하기 위한 주요 절차 3가지

- 두 자료의 결합은 방법론적으로나 절차적으로나 그 과정이 꽤나 복잡한 것 같다. 행정자료를 활용할 것인가 또는 어떻게 활용할 것인가를 결정하기에 앞서 다음의 몇 가지 사항을 확인할 필요가 있을 것이다.
 - 첫째, 조사항목을 대체할 만한 행정자료가 있는가? 행정자료는 행정처리목적을 위해 수집된 자료라는 점에서 통계작성에 적합한 자료형태가 아닐 수 있다. 어떤 행정자료는 정책의 결과물일 수 있으며 이러한 정책의 생멸, 변화과정 등을 모니터링하고 그 결과 또한 통계작성에 참고할 수 있어야 한다.
 - 둘째, 조사와 행정자료에서 측정하고자 하는 개념과 범위가 같은가? 또는 측정 대상에 대한 포괄범위가 같은가? 겉으로는 두 자료가 추구하는 바가 동일하게 보일 수 있지만, 행정목적용 행정자료와 통계목적용 조사자료는 각각이 의미하는 바는 다를 수 있다는 점을 이해해야 한다.
 - 셋째, 조사자료와 행정자료를 어떻게 결합할 것인가? 방법론적으로 정확매칭(exact matching), 확률매칭(probability matching), 통계적매칭(statistical matching) 등이 있다. 정확매칭은 두 자료를 결합할 변수(매칭키)가 있어야 하고 매칭 후 매칭률과 매칭편향(matching bias)을 확인해야 한다. 매칭편향이 있을 경우 다출처 자료로 작성한 통계의 신뢰성을 위해 보정할 필요가 있다.

2 조사자료와 행정자료의 연계과정

1 단계1: 활용가능한 행정자료가 있는가

- 가계금융복지조사에서는 국세청의 과세자료, 보건복지부의 사회수혜금, 각 연금 공단의 국민연금, 공무원연금, 사학연금, 별정우체국, 군인연금의 연금소득과 각 공단의 연금기여금 및 건강보험료 지출에 관한 행정자료를 이용한다.
- 본 조사에서 행정자료는 한 변수에 대해 가구원 중 일부 또는 전체 가구원에 대해 대체하는 방식으로 사용된다.
 - 근로소득의 경우 행정자료를 사용할 수 없는 소득원이 있는 가구원에 대해서는 조사자료를 활용한다. 사업소득은 두 자료가 추구하는 개념은 동일하나 식량작물 재배업과 같은 면세사업의 경우는 행정자료에 포함되지 않는 경우 포괄범위가 다를 수 있고 사업소득 경비처리에 대한 인지정도가 두 자료 간에 다를 수 있어서 무응답에 대해서만 행정자료로 대체한다.
 - 금융소득, 공적이전소득(기타 정부보조금 항목 제외), 공적이전지출 중 소득세, 연금 기여금 지출 항목의 경우는 연계키가 없는 극히 일부 자료를 제외하면 가구원 전체에 대해 행정자료로 대체한다.
- 참고로 자산에 대해서는 현재로서는 행정자료를 사용할 수 없다. 부채자료의 경우 자료를 입수하여 보조지표 형태로 활용하고 있지만, 조사자료를 대체하는데 사용하지는 않는다.³⁾ 부채자료만 행정자료로 대체할 경우 주요 경제지표인 순자산^(자산-부채) 통계가 왜곡될 수 있기 때문이다.
- 종합적으로 가계금융복지조사에서는 조사기준 2017년(소득기준 2016년) 자료부터 행정자료로 보완된 소득통계자료를 제공하고 있다. 2012년 이후 모든 조사시점에 대해 소득 행정자료를 입수하여 활용하고 있지만 2017년부터 소득통계 보완결과를 제공하는 이유는 자산, 부채와의 정합성 등의 문제로 과거자료까지 소급하는데 한계가 있기 때문이다.

3) 부채 행정자료는 신용정보원으로부터 입수하여 별도의 보조지표(보도자료)와 별도의 마이크로자료로 제공하고 있다. 다만 자산 행정자료는 법령·제도적 장치 및 집중기관 부재로 입수되지 못하고 있다.

- 한편 소득분배지표는 소득기준 2011년 이후 전체 시계열자료에 대한 행정자료로 보완된 결과를 제공하고 있다.

② 단계2: 조사자료와 행정자료의 개념과 포괄범위가 같은가

- 조사개념에 대해 행정자료의 개념이 얼마나 같은지를 검토한다. 조사자료를 비교 기준으로 삼는 이유는, 행정자료를 통해 조사통계의 품질제고는 물론이고 표본 조사의 효율성을 높이려는 목표가 있기 때문이다.

○ 두 자료의 항목별 개념 비교

	조사자료	행정자료
근로소득	<ul style="list-style-type: none"> - 세금, 각종 부담금 등 세금 공제전 소득 - 근로의 대가로 받은 일체의 현금·현물보수 <ul style="list-style-type: none"> · 임금 및 수당* <ul style="list-style-type: none"> * 기본급, 근속급, 가족수당, 야근수당 등 · 상여금(명절휴가비·가계지원비·정근수당·성과급 등) · 퇴직수당, 명예퇴직수당 포함 <ul style="list-style-type: none"> 단, 퇴직금(퇴직일시금)은 제외 * 연말정산 환급금은 포함하지 않음 	<ul style="list-style-type: none"> - 거의 일치 <ul style="list-style-type: none"> · 원천징수, 연말정산, 종합소득신고에 의한 소득 · 소득하위층 미신고 소득 발생 가능 · 행정자료는 현물성 보수 등은 미포함 · 자활급여⁴⁾ 포함
사업소득	<ul style="list-style-type: none"> - 자영업주 또는 고용주로 직접 사업을 경영하거나 지식 또는 재능을 이용하여 얻은 사업수입 (매출액)에서 인건비, 재료비, 이자 등 비용 차감한 순소득 - 농축림어업 순소득 - 애완견 분양, 텃밭 채소, 나물채취 등 부업소득 - 프리랜서 원고료, 작곡료 등 - 자가소비용 생산품에 대한 소득(16년부터) 	<ul style="list-style-type: none"> - 개념적으로 조사자료와 거의 유사하나 일부 포괄범위 차이 및 경비처리에 대한 인식 차이 · 미등록 소규모 사업체, 현금매출 중심업체의 매출 과소신고 가능성 · 작물재배업, 식량작물 등 농업사업 등 면세사업은 국세청 미신고 대상 · 기타소득에 포함된 사업소득(원고료 등) 일부 분리 수집
임대소득	<ul style="list-style-type: none"> - 주택, 건물, 토지 등 자산을 임대해주고 받은 순소득(현물소득도 포함) <ul style="list-style-type: none"> · 이자·관리비 등 비용은 포함 안함 · 임대료를 현물로 받은 경우 포함 - 사업자등록 여부에 관계없이 모든 임대소득 포괄 	<ul style="list-style-type: none"> - 임대사업소득자 등록한 경우 개념 일치 · 미등록자의 경우 신고 안된 임대소득은 파악할 수 없어 조사에서 측정된 임대소득에 비해 과소
금융소득	<ul style="list-style-type: none"> - (이자) 예·적금, 채권·증권 이자 등 순소득(세후) - (배당) 보험, 신탁, 주식 등에 대한 배당금 수익 <ul style="list-style-type: none"> · 주식 매매 차익은 소득에 포함 안 함 	<ul style="list-style-type: none"> - 일치 <ul style="list-style-type: none"> · 다만, 인정배당, 간주배당은 제외 · 행정자료는 세전소득(자료처리 시 세후)

	조사자료	행정자료	
공적연금	- 국민연금, 공무원연금, 군인연금, 사학연금, 별정우체국연금 · 국민연금은 노령연금, 장애연금, 유족연금 등 포함	- 일치	
기초연금	- 만66세 이상의 수급대상자에게 지급하는 현금급여	- 일치	
맞춤형 기초생활수급금	- 생계급여, 주거급여, 교육급여 등 수혜대상자에 현금으로 지급하는 맞춤형기초생활보장급여	- 일치	
양육수당	- 가정양육수당 등 각종 양육수당, 양육보조금, 아동수당, 지방자치단체 출산장려금 등	- 일치 · 다만, 지자체 출산장려금 등 지자체 지원 관련 행정자료 입수 불가	
장애수당	- 장애인연금, (경증)장애수당, 장애아동수당 등 장애관련 현금급여	- 일치 · 다만, 지자체 사업에 의한 장애인 관련 수당 일부 자료 미입수	
세금	소득세	- 각종 소득세로, 공제 후 결정세액 · 근로소득, 사업소득, 종합소득세 등 · 연금소득세, 이자·배당 소득세	- 일치
	재산세	· 가계용 토지 및 주택 등에 부과 - 재산세, 종합부동산세	- 불일치(조사의 포괄범위보다 넓음) · 사업용 재산의 재산세 포함 · 종합부동산세 자료 미입수
	자동차세	· 가계용 자동차에 부과되는 세금과 및 환경개선부담금	- 불일치(조사의 포괄범위보다 넓음) · 사업용 자동차세 포함 · 환경개선부담금 자료 미입수
기여금·보험료	연금기여금	- 국민연금, 기타연금* 기여금 지출액 * 공무원연금, 사학연금, 군인연금, 별정우체국연금	- 일치
	건강보험료	- 국민건강보험법 따라 근로자 또는 자영자 가구에 부과되는 보험료 · 건강보험료, 노인장기요양보험료	- 일치
	고용보험료	- 고용보험(실업보험료)료로 납부한 금액	- 불일치(조사의 포괄범위보다 넓음) · 자진신고사업장 종사자의 개인별 고용보험료 자료 부재 등 자료구조가 복잡

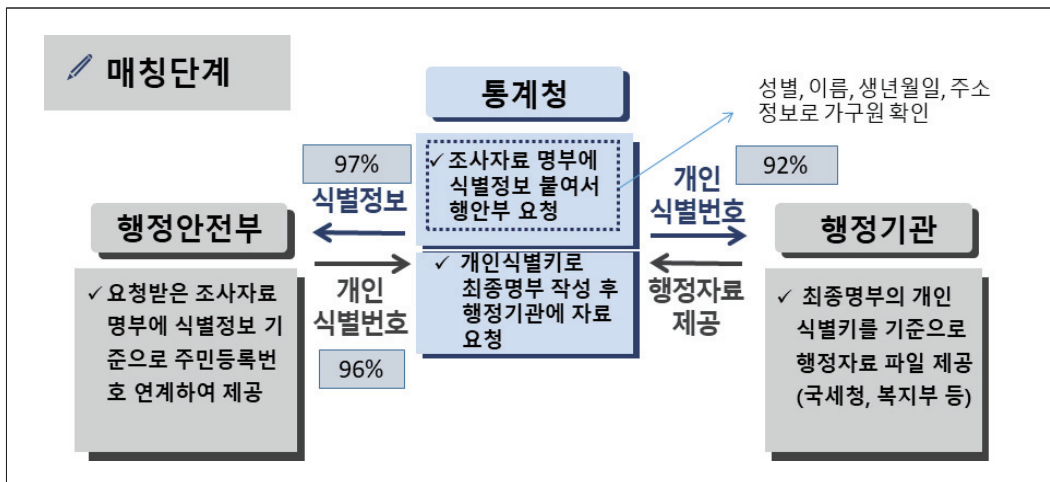
- 4) 자활급여는 국민기초생활보장법에 따라 수급권자에게 근로기회를 제공하여 자립할 수 있도록 돕기 위한 기초생활보장제도의 급여에 해당. 자활일자리는 기초생활보장 수급자 등 근로빈곤층의 자립을 위해 정부가 제공하는 일자리로서, 정부는 자활사업 참여자에 대해서 자활근로에 대한 급여를 지급한다.

- 위의 표에서 ‘일치’ 항목은 행정자료로 조사자료를 대체하고, ‘거의 일치’ 항목은 조사자료와 행정자료를 혼용하고, ‘불일치’이거나 ‘포괄범위가 다른 경우’는 조사자료를 그대로 사용할 수 있다.
- 공적이전소득 관련 항목들은 기본적으로 조사자료와 행정자료의 측정개념이 동일하여 행정자료로 조사자료를 전면대체한다. 항목에 따라 두 자료 간 포괄범위가 미미하게 차이가 있는 경우도 있지만, 그 경우의 수가 매우 적고 금액수준이 전체 소득 분포에 전혀 영향을 주지 못하는 수준이다.

③ 단계3: 두 자료를 어떻게 결합할 것인가

- 조사자료와 행정자료는 개인단위로 결합한다. 가계금융복지조사의 소득자료는 개인단위로 수집하므로, 개인단위에서 두 자료를 결합하여 기본적인 자료셋을 구성한 후 가구단위로 행정자료를 재구성하여 분석한다.
- 두 자료는 다음과 같은 과정을 거쳐 매칭한다(2019년 기준). ① 두 자료를 연계할 연계명부를 작성한다. 연계명부는 개인단위의 성별, 이름, 생년월일, 주소와 같은 텍스트 정보를 포함한다. ② 연계명부에 대한 연계키를 해당 행정기관에 요청하고 해당기관으로부터 명부에 붙은 연계키를 제공받는다. ③ 연계키를 포함한 행정자료 입수명부를 각 행정자료 보유기관에 보내고 각 해당기관으로부터 연계키 단위의 행정자료를 제공받는다.

<두 자료 간 소득자료 연계 과정>



① 두 자료를 연계할 수 있는 적절한 변수(매칭키)를 찾는다.

- 가계금융복지조사에서 정확매칭하는데 필요한 연계변수를 직접 조사하지 않는다. 조사변수로서 개인의 이름, 성별, 생년월일, 주소 등이 측정되며 이 변수들을 행정자료에 대한 매칭변수로 활용한다.

- 연계키는 두 자료를 연계하는 목적으로만 사용되며, 두 자료를 결합한 후에는 비식별처리되어 연계명부에서 삭제한다.

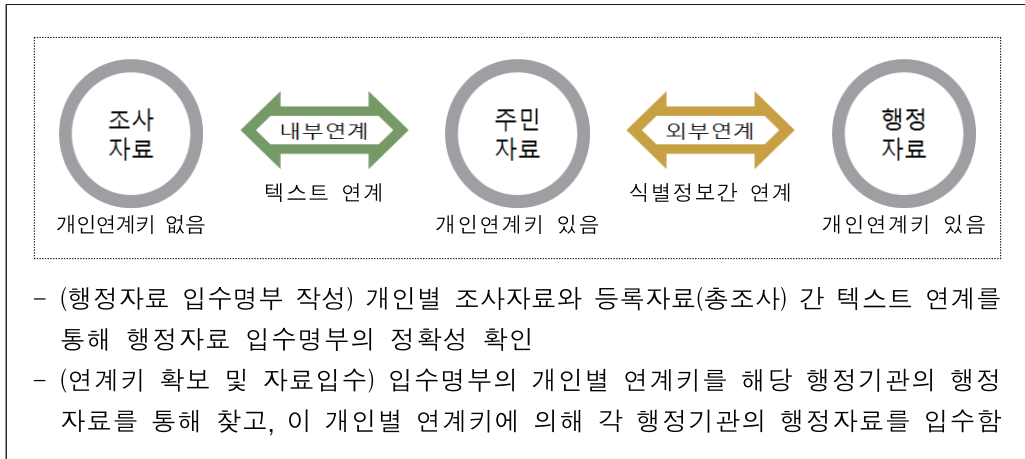
- 두 자료를 정확하게 연계하기 위해서는 정확한 연계키를 찾는 것이 중요하다. 가구는 가구원에 대해 이름, 성별, 생년월일, 거주지주소에 대해 직접 작성하게 되는데 이때 기입오류, 기억오류 등의 응답오차가 발생하기도 한다. 이들 텍스트정보에 오류가 있을 경우 정확한 연계키를 찾을 수 없으므로, 등록기반의 행정자료에 대해 재차 확인하는 과정을 거친다.

- 이렇게 등록자료를 통해 매칭변수를 확인하더라도 모든 가구원에 대한 텍스트 정보를 확인할 수는 없다. 가계금융복지조사의 경우 가구원별 주소지를 조사에서는 직접 파악하는 것이 아니라, 대표 거주지 주소지를 조사하기 때문에 떨어져 사는 가구원의 거주지 정보를 정확하게 확인하는데 한계가 있을 수 있다. 또한 등록자료와 조사자료 간의 시점차이로 주소지 불일치가 있을 수도 있다. 가구원의 텍스트 정보가 등록자료에 대해 확인되는 경우는 전체 조사대상 가구원의 약 97% 정도 된다.

② 연계명부(4개 텍스트 변수를 포함한 개인별 리스트)는 관련 행정기관을 통해 입수되며, 행정자료는 통계법 등 관련 근거에 따라 통계청에 제공한다.

- 이 과정에서도 연계명부와 등록기관의 행정자료 간의 시점차이로 인한 주소의 불일치 등으로 명부에 있는 대상자의 연계키를 100% 찾을 수 있는 것은 아니다. 연계대상자의 약 96% 정도의 연계키를 찾게 되는 것 같다.

- 모든 과정을 종합하면 최종적으로 조사대상 가구원의 약 92~95% 정도에 대해 연계키를 찾을 수 있으며 매년 연계키 확보율은 조금씩 다를 수 있다.



○ 연계명부 대상자의 텍스트 정보를 보다 정확하게 검증하기 위해서는 등록자료에 대해 몇 단계의 확인과정을 거치게 된다.

- 우선 전년도 연계된 가구원이 있는 모든 세대원을 등록기반 총조사자료에서 추출하고, 전년도 미연계 가구원이거나 올해 신규 가구원의 경우는 이름, 성별, 생년월일이 모두 일치하는 가구원의 모든 세대원을 총조사자료에서 추출한다. 상세주소정보가 일치하는 세대원을 모두 추출한다.

- 위에서 추출한 자료 중 1:다(多)로 추출된 자료에 대해 다음과 같은 다단계 절차를 거쳐 1:1로 연계되는 가구원 자료를 확인한다.(2016년 자료 기준)

경우	확인내용(주요과정)	건수	누적비(%)
1	이름, 성별, 생년월일, 법정동코드, 텍스트 상세주소 일치	33,780	66.4
2	이름, 성별, 생년월일, 법정동코드 일치하고 가구원 정보 일치	4,816	75.9
3	이름, 성별, 생년월일 일치하고 가구원 정보 일치	1,725	79.3
4	해당 가구원 정보 불일치 하지만 다른 가구원 정보가 2명 이상 일치	2,724	84.6
5	이름, 성별, 생년월일 전국1대1 일치	3,430	91.4
6	텍스트 상세주소가 같고 이름이 일치	1,718	94.7
7	텍스트 상세주소가 같고 이름이 2자 이상 일치	406	95.5
8	텍스트 상세주소가 같고 연계된 가구원이 2명 이상	30	95.6
9	텍스트 상세주소는 다르지만 연계된 가구원이 2명 이상	650	96.9
전체		50,872	100.0

- 2016년 자료 기준으로 50,872 조사대상 가구원 중 49,288명(96.9%)의 자료가 등록기반 총조사자료에 대해 확인되었다.
- 1인 가구의 일치율은 90%로 2인 이상 가구의 일치율(97.4%)보다 낮다. 2인 이상 가구는 같은 세대정보를 갖는 가구원들 간의 교차 확인이 가능한 반면, 1인가구는 이러한 교차검증이 안되기 때문에 일치율이 상대적으로 낮다.

가구원수	총가구원(명)	구성비(%)		연계(명)	연계율(%)	미연계(명)
		연계	미연계			
계	50,872	100.0		49,288	96.9	1,584
1인	3,583	7.0		3,225	90.0	358
2인 이상	47,289	93.0		46,063	97.4	1,226

- 지역별 연계율은 비슷하다. 연령별 연계율은 30대에서 약간 낮는데, 30대는 혼자 사는 가구가 다른 연령대에 비해 많다는 특성이 있다.

		지역 내 비중		
		전체	연계	미연계
	계	100.0	93.6	6.4
지역	서울	100.0	92.9	7.1
	특광역시	100.0	93.9	6.1
	경기	100.0	94.0	6.0
	경기 이외 도	100.0	93.6	6.4
시도별	서울	100.0	92.9	7.1
	부산	100.0	94.8	5.2
	대구	100.0	93.6	6.4
	인천	100.0	95.7	4.3
	광주	100.0	90.6	9.4
	대전	100.0	93.3	6.7
	울산	100.0	94.4	5.6
	세종	100.0	94.0	6.0
	경기	100.0	94.0	6.0
	강원	100.0	94.3	5.7
	충북	100.0	93.7	6.3
	충남	100.0	92.6	7.4
	전북	100.0	92.6	7.4
	전남	100.0	93.4	6.6
	경북	100.0	93.7	6.3
	경남	100.0	95.5	4.5
제주	100.0	95.5	4.5	

		연령그룹 내 비중		
		전체	연계	미연계
	계	100.0	93.6	6.4
연령그룹	30세 미만	100.0	92.9	7.1
	30~39	100.0	91.9	8.1
	40~49	100.0	95.5	4.5
	50~59	100.0	94.6	5.4
	60세 이상	100.0	93.9	6.1

		종사상지위 내 비중		
		전체	연계	미연계
	계	100.0	93.6	6.4
종사상지위별	상용	100.0	93.1	6.9
	임시일용	100.0	93.2	6.8
	고용원있는 자영업자	100.0	93.4	6.6
	고용원없는 자영업자	100.0	95.0	5.0
	무급가족	100.0	93.9	6.1
	기타종사자	100.0	95.5	4.5
	무직	100.0	93.3	6.7

- 전체적으로 보면 특성 그룹 간 매칭율에 큰 차이는 없어 보인다. 다만, 떨어져 사는 가구원이 있는 경우는 정확한 주소가 없다는 점에서 정확매칭에 다소 어려움이 있기도 하다.
- 두 자료 간의 매칭률이 높다 하더라도 매칭그룹의 특성에 따라 편차가 있을 수 있으므로 적절한 보정방법을 통해 편향을 줄이는 것이 좋다.

3 매칭편향 보정방법

- 가계금융복지조사의 매칭키 확보율은 약 92~95% 정도로 높은 편으로서 매칭편향에 대한 우려는 크게 없을 것이다. 그렇다하더라도 특정 그룹 즉, 혼자 또는 떨어져 사는 특성이 있는 1인 가구 또는 30대 연령그룹에서 연계율에 차이가 있을 수 있으므로 통계적 기법 등에 의해 편향을 보정할 필요가 있다.
 - 매칭편향을 없애고자 하는 이유는, 행정자료로 조사값을 보완할 경우 보완 전에 비해 소득이 상향조정되는 경향이 있으며 매칭키가 없을 경우 조사자료를 사용할 수 밖에 없다. 따라서 이 경우는 행정자료로 보완된 자료에 비해 과소추정될 소지가 있기 때문이다.
- 매칭편향 보정은 조사값이 있는데 매칭키가 없는 가구원에 대해 이루어진다. 아이디어는 행정자료와 조사자료의 상대적 비율만큼 조정한다는 것이며 이러한 개념을 통계적 모형으로 구체화하는 방법을 사용한다⁵⁾.
 - 조사값이 있는 경우 회귀모형 기반 혼합비(**ratio**) 대체방법을 사용한다. 모형에서 행정자료로 보완된 값이 관심변수이며 해당변수의 조사값은 보완값과 강한 양의 상관을 가지므로 조사값이 보완값의 핵심 예측변수가 될 수 있다.

대체모형	비 대체 방법
대체방법	전체 표본을 몇 개의 그룹으로 나누어서 그 셀 내에서 비대체 실시
주요이슈	셀을 몇 개로 나눌 것인가가 중요하며, 이는 혼합모형(mixture model)을 사용

- 연구결과 근로소득과 금융소득항목에서 편향보정 효과가 큰 것으로 나타났다.

5) 행정자료 활용에 따른 시계열 조정 등 편향보정 방법 연구 보고서 (통계청, 2018)

PART IV

행정자료의 전처리 과정

1. 기본방향
2. 소득원천별 행정통합자료 작성 방법



PART



IV 행정자료 활용을 위한 자료의 전처리 과정

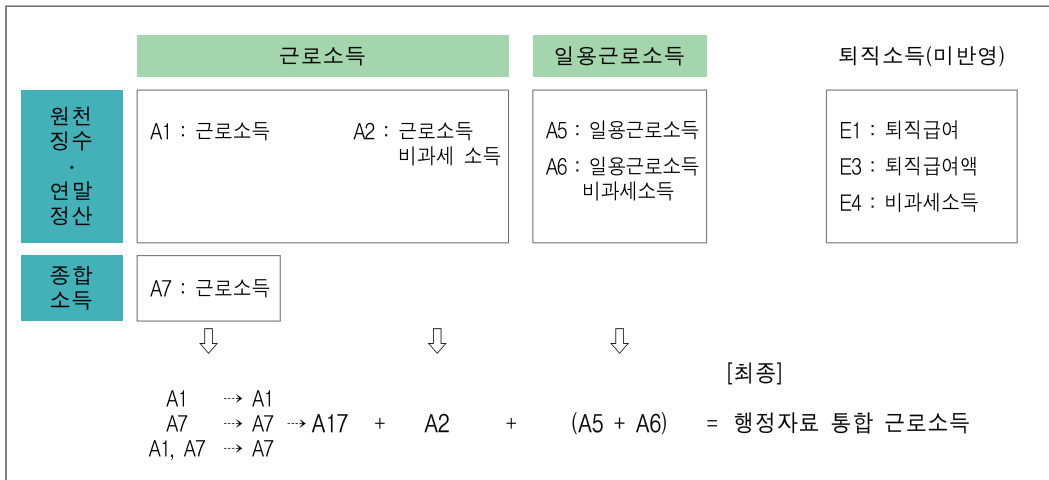
1 기본방향

- (기본원칙) 조사자료와 행정자료의 개념이 일치하거나 포괄범위가 같은 경우에 대해서는 기본적으로 행정자료가 더 정확하다고 가정한다. 조사자료를 기준으로 특정항목에 대해 응답하지 않거나, 응답착오일 것으로 판단되는 경우에는 행정자료로 조사자료를 대체한다. 이를 통해 조사자료의 측정오차를 최소화하고자 한다.
- (대체단위) 개인단위에서 조사자료를 수정하고 대체하는 것으로 원칙으로 한다. 다만, 일부 항목의 경우 개인단위의 응답이라 하더라도 가구단위 검토를 거쳐 최종 개인단위 대체 여부를 결정할 수 있다. 이는 한 가구 안에서 가구원 중 일부에 대해 연계키가 없을 수 있으므로 가구단위에서의 항목특성까지 고려하여 분석할 필요가 있다.
- (포괄범위) 조사항목의 개념 및 포괄범위에 부합하도록 행정자료가 포함하고 있는 다양한 신고서식의 분류별 상세코드를 확인한다. 이를 통해 두 자료의 포괄범위가 일치할 수 있도록 여러 단계의 자료처리과정을 거친다.
- (통합소득 작성) 조사자료와 비교하기 위해 소득원천별로 다출처 행정자료를 통합하여 행정자료 기반 통합근로소득, 통합사업소득, 통합소득세 등을 작성한다.
- (결합분석) 다출처자료들을 통합하고 분석하는 과정으로 자료의 중복성 및 일치성 등을 복합적으로 고려하여 최종적으로 행정자료의 활용 기준을 정한다.



2 소득원천별 행정통합자료 작성 방법

1] 근로소득의 구성



① 각 근로소득 관련 신고서식 중에서 입수하는 주요항목

- 근로소득은 일용근로소득, 원천징수(연말정산), 종합소득 형태로 신고된다.

신고서식	입수항목
원천징수 · 연말정산	소득명세서 계_합계, 비과세소득 계_합계, 감면소득 계_합계, 소득세_결정세액
종합소득	총수입금액
일용근로소득	총지금액, 비과세소득, 소득세

② 다양한 서식으로부터 입수한 근로소득 항목을 통합하는 이유

- 조사표에서 '근로소득'은 한 개 항목으로 측정되며, 이 한 항목에는 다양한 개념(형태)의 근로소득을 포함한다. 따라서 각 출처를 달리하여 입수된 행정자료들도 조사자료와 결합하기 위해서는 조사항목과 동일하게 구성해야 한다.
 - 국세청의 소득신고서식은 세금부과를 목적으로 하고 있기 때문에 과세대상소득에 주로 초점을 맞추고 있다. 어떤 사람의 경우 한 해 동안 이직하거나 겸업도 할 수 있으므로 여러 신고양식에 따라 소득신고를 할 수도 있다. 즉, 한 사람에게 여러 개의 근로소득 자료가 있을 수 있다.

③ 자료 내 소득구성 및 다수의 건 자료 처리

□ 원천징수(연말정산) 근로소득

* 여기서 사용하는 번호(예 ②)는 국세청의 '근로소득지급명세서[별지 제24호 서식(1)-1쪽]' 양식에 따르고 있으며 자세한 서식은 국세청 참고

○ 소득구성⑥: (A1)⑩소득명세 계 + (A2)⑳비과세소득 계

- A1(근로소득): 신고서식의 소득명세 ⑩계_합계: 비과세소득 제외한 소득합계
- A2 ⑳비과세소득 계: 비과세소득 합계

○ 한 개인의 다수의 건 자료 처리

- 한 사람이 작년 한 해 동안 이직 또는 겸업하였을 경우 2개 이상의 근로소득 자료가 있을 수 있다.
 - 연말정산은 주(현), 종(전)근무지 모두에 대해 신고하며, 주(현)근무지에 대해 연말정산할 때 종(전)근무지의 근로소득 금액 더해서 정산한다.
 - 서식에서 필요항목은 '⑩계 행의 합계 열'의 금액이며, 이 금액은 주(현), 종(전)근무지의 소득이 모두 더해진 금액이다. *주(현) 자료만 입수된 경우는 합산 처리함

<소득명세서 양식 일부양식>

구 분		주(현)	종(전)	종(전)	⑩-1 납세조합	합 계
I 근 무 처 별 소 득 명 세	⑨ 근 무 처 명					
	⑩ 사업자등록번호					
	⑪ 근무기간	~	~	~	~	~
	⑫ 감면기간	~	~	~	~	~
	⑬ 금 여					
	⑭ 상 여					
	⑮ 인 정 상 여					
	⑮-1 주식매수선택권 행사이익					
	⑮-2 우리사주조합인출금					
	⑮-3 임원 퇴직소득금액 한도초과액					
	⑮-4					
⑯ 계						

- 6) · 임원퇴직금 한도초과액은 근로소득 대상이 아님. 본 조사에서 한도초과액이 있는 경우는 극소수이며 이를 포함시 전체 소득분포에 별 영향을 주지 않음
- 감면소득은 근로소득에 포함되며, 위의 소득명세서의 '계'에 포함됨
 - 인정상여의 경우, 누구에게 귀속됐는지 불분명하여 인정상여 처분을 받았다 하더라도 실제로 대표자가 수령한 경우도 포함될 수 있으며 이 경우는 근로소득으로 보는 것이 타당. 행정자료 연계가구원 중 대표자가 아닌 가구원도 소액의 인정상여액이 있기도 하지만 금액이 매우 작고 항목분류가 어려워 일괄적으로 제외하는데 한계가 있음

- 따라서 전 근무지의 연말정산 자료와 현 근무지의 최종 연말정산 자료가 둘 다 있으면 최종 연말정산자료의 '⑩계_합계' 금액을 선택한다.
- 한 사람에 대해 근로소득 자료가 여러 개 있을 경우 최종 1건의 자료를 선택해야 하며 그 기준은 다음과 같다.

① 근로소득이 가장 큰 값 → ② 동순위인 경우 소득명세 합계가 큰 값 →
③ 동순위인 경우 결정세액이 작은 값

- ②의 근거: 주(현)근무지 소득은 종(현)근무지 소득까지 합산하므로 일반적으로 소득금액이 가장 클 것임
- ③의 근거: 각 자료별 소득이 동일할 경우, 주(현)근무지에서 공제 등이 많이 될 것으로 가정하여 결정세액이 가장 작은 자료를 선택함. 이는 가장 최종적으로 신고된 소득금액을 취하겠다는 의미이기도 함

□ 종합소득신고에 의한 근로소득

종합소득 신고는 사업소득(자영업자)이 있거나 2천만원 이상의 금융소득이 있는 사람들이 하는 소득신고 제도임. 연말정산을 통해 근로소득을 신고한 사람도 종합신고하게 될 경우 기 신고한 근로소득도 포함해서 다시 신고함. 따라서 근로소득의 경우 종합신고상 소득이 가장 최종(정확한) 행정자료일 수 있음

○ 소득구성: 총수입금액을 사용⁷⁾

- (A7) 종합소득 근로소득 ⑤총수입금액
 - 종합소득신고의 총급여액은 비과세소득과 분리과세소득이 포함되지 않는다. 근로소득 중 일용근로소득의 경우만 분리과세대상이 있다.
- 서식에서 '소득금액'은 총급여액에서 근로소득공제·연금소득공제 등을 뺀 금액이므로 '총급여액' 항목이 최종 조사대상 소득에 해당한다.

○ 종합소득은 신고대상에 한해서 한 사람이 한 건만을 신고하므로, 한 사람에 대해 여러 건의 자료처리 상황은 발생하지 않는다.

7) 근로소득명세서[별지 제40호서식(1) 참고

□ 일용근로소득

○ 소득구성방법: (A5)총지급액 + (A6)비과세소득

- (A5) 일용근로소득 신고서상의 ㉔총지급액

- 일용근로자에게 지급한 급여 합계금액으로, 과세대상이 되는 총지급액을 말하며 여기에 비과세소득은 포함되지 않는다.⁸⁾

- (A6) ㉕비과세소득

- ㉔총지급액(과세소득)에 포함하지 않는 비과세대상 소득금액이며 생산직 일용근로자에게 지급한 야간근로수당 등이 여기에 포함된다.

○ 다수의 건 자료 처리

- 한 사람한테 일용근로소득자료가 2건 이상이 있을 수 있으며, 이 경우는 여러 건의 소득을 합산하여 통합 일용근로소득을 만든다.⁹⁾

□ 퇴직소득

○ 퇴직소득은 현 신고서식상 퇴직일시금과 퇴직수당 등이 분리되어 있지 않아 활용하기가 어렵다. 조사에서 퇴직수당(명예퇴직수당)은 근로소득, 퇴직일시금(퇴직일시금(연금)제도)에 의함)은 금융자산에 각각 포함된다.

- 퇴직수당 행정자료를 사용하지 못할 경우 근로소득에 포함되어야 할 퇴직수당 행정자료 금액이 누락될 가능성도 있다.

④ 자료 간 중복 처리 및 합산단계

○ 한 개의 신고서식에 여러 개의 소득항목이 있거나, 한 사람한테서 원천징수, 종합소득, 일용근로소득 등 여러 자료가 발생할 수 있다. 이 경우 서식 내 여러 개 항목을

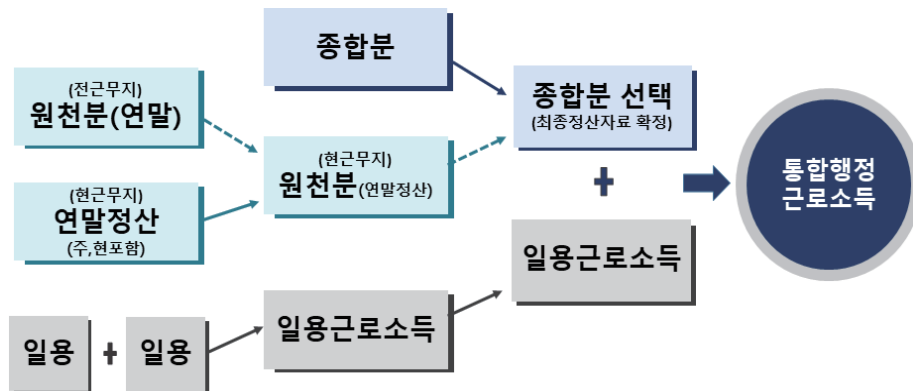
8) 현금영수증발급장치를 통해 신고하면 비과세소득이 포함되며, 이 장치는 납부할 원천징수세액이 없는 일용근로자에게 급여를 지급하고 그 즉시 제출할 때만 사용

9) 일용근로소득의 경우(제한적이긴 하지만) 자영업을 운영하는 가구에서 실제로 일용근로자가 아님에도 불구하고 가공인건비(가칭) 처리 가능성도 배제할 수 없다는 전문가 의견도 있음. 자료검토 결과 미미한 숫자이긴 하지만 가공인건비 소지가 있는 가구도 있을 수 있어, 자영업자가 있는 일부 가구에 대해 이 가구에 포함된 가구원이 받는 일용근로소득은 제외함. 다만 자영업자 본인이 받는 일용근로소득은 포함

통합하거나 불필요한 항목은 없애거나 하는 등의 자료처리 과정을 통해 조사자료와 비교할 수 있는 행정자료를 구성한다.

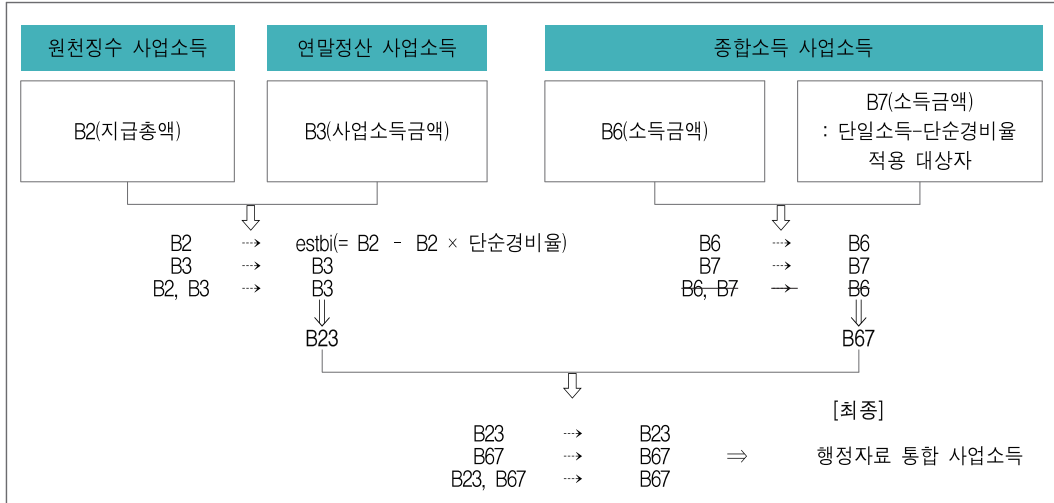
- **종합분과 원천분 근로소득 간 중복 처리는 “종합분의 총수입금액+원천분 비과세 소득 계_합계” 방식으로 처리한다.**
 - 종합소득신고는 시기적으로 가장 늦게 이루어지므로, 정정·추가 과정을 통해 소득변동 사항이 반영된 최종 소득으로 볼 수 있다. 따라서 원천분과 종합분이 모두 있으면 **종합신고상의 근로소득을 최종 근로소득으로 선택한다.**
 - 종합소득신고 시 비과세 근로소득은 신고대상이 아니므로, 원천분 비과세 근로 소득이 있으면 이를 종합분 근로소득에 합산한다.
- **일용근로소득 자료 처리**
 - 한 사람이 원천·종합분의 근로소득과 일용근로소득이 모두 있으면 원천·종합 근로소득과 일용근로소득을 합산한다.
 - 일용근로소득은 원천징수에 의해 납세의무가 종결되며 연말정산 및 종합소득신고 대상이 아니다.(예외인 경우도 있으나 이를 반영하지 않음¹⁰⁾)

<행정자료 활용에 따른 통합 근로소득 작성 과정>



10) 동일사업자에게 3개월 이상(건설업은 1년 이상) 고용된 경우는 해당연도의 일용근로소득도 합산하여 연말정산한다. 이 경우 일용근로소득자료와 연말정산자료가 중복될 수 있지만, 중복숫자는 극히 적고 소득수준도 매우 낮아서 소득분포에 거의 영향을 주지 않을 것으로 예상된다.

② 사업소득



① 사업소득 관련 행정자료 서식과 입수항목

- 사업소득은 종합소득, 연말정산, 원천징수형태로 신고된다.

구분	입수항목
원천징수 사업소득	업종구분코드, 지급총액, 원천징수세액_소득세
연말정산 사업소득	사업소득금액, 결정세액_소득세
종합소득신고 사업소득	소득구분코드, 주업종코드, 총수입금액, 소득금액
단일소득-단순경비율 신고 사업소득	종합소득금액, 총결정세액

② 통합 사업소득 작성 방법

□ 소득의 구성¹¹⁾

- (B2) 사업소득의 ⑮지급총액은 비용을 차감하지 않은 금액이다.
 - 의료보건용역, 인적용역 등 법령에서 정한 일부 사업소득의 경우 지급하는 자가 소득세를 원천징수한다.
 - 원천징수(B2)대해 단순경비율¹²⁾ 기준으로 소득금액 추계하여 반영한다.
 - 소득금액 = 수입금액 - (수입금액 × 단순경비율)

11) 사업소득명세서[별지 제40호서식(1)], 단순경비율 적용대상자의 경우 [별지 제40호서식(4)] 참고

12) 국세청 경비율 고시-제4조(단순경비율 기본율 및 초과율 적용) 인적용역 사업소득자(업종코드 940***)의 귀속 수입금액 4천만원 이하의 단순경비율 기본율을 적용하고 4천만원 초과분은 초과율을 적용함

- (B3) 사업소득 연말정산 ‘사업소득금액’은 수입에서 각 사업별 단순경비율을 적용한 비용을 차감한 금액이다.
 - 보험모집, 방문판매, 음료배달업 등 간편장부 대상 등 조건에 해당하는 경우에는 연말정산으로 납세의무를 종결한다.
- (B6) 종합소득 사업소득명세서 ‘소득금액’은 ‘총수입금액-필요경비’를 말한다.
- (B7) 단일소득, 단순경비율 적용 대상자용 종합소득금액으로, ‘총수입금액-단순경비율에 의한 필요경비’의 형태로 나타난다.
 - * 단순경비율의 경우, 국세청이 대상자에게 해당 내용을 송부하고 해당 내용에 대해 동의하면 서명·제출하고, 이의가 있으면 간편장부로 종합소득신고를 하면 됨

□ 중복자료 처리

○ 자료 내 중복처리

- 원천징수자료는 한 사람에 대해 여러 건의 자료가 입수된 경우는 각 자료의 금액을 합산한다. 연말정산자료가 한 사람에 대해 여러 건이면, 주(현) 지급자 사업소득으로 추정되는 1건만 반영(근로소득 연말정산과 동일)한다.
- “각 자료 내에서 ① 사업소득 금액이 가장 큰 자료, ② ①에서 동순위인 경우 결정세액이 작은 자료” 순서로 1건의 자료를 선택한다.

○ 자료 간 중복처리

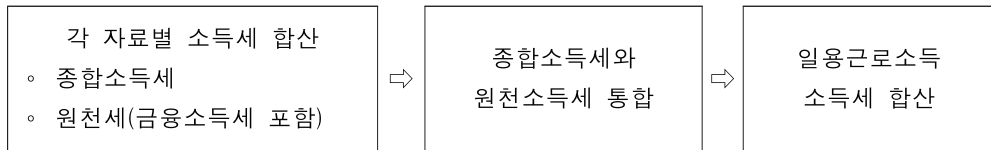
- 연말정산(원천징수)과 종합소득신고에 의한 사업소득이 모두 있으면 종합신고상의 소득자료를 최종적으로 선택한다.
- 연말정산과 원천징수에 의한 사업소득이 모두 있는 경우에는 연말정산 자료로 반영한다.
 - * 참고로 원천징수 사업소득의 업종구분코드 미반영¹³⁾

13) 원천징수와 연말정산을 비교할 때, 원천징수 중 연말정산에 해당하는 업종코드(보험판매, 음료배당 등)로 금액을 분리하지 못하고 전체 합산금액을 비교

③ 소득세

① 자료소스

- 종합소득세 또는 원천세의 근로·사업 등 소득종류별 소득세, 일용근로소득의 소득세 자료를 사용한다.
- 일용근로소득의 소득세 자료는 2015년 자료부터 입수하고 있다. 그 이전 일용근로소득의 소득세에 대해서는 이전자료와 타 정보 등을 보조지표로 활용하여 통계적 모형으로 추정하였다.



② 통합 소득세 처리방법

□ 종합소득세와 단일소득-단순경비율 총 결정세액

- 종합소득세 자료와 각 서식별 원천 소득 자료가 각각의 독립된 파일로 입수되며, 소득세 자료는 세금 금액이 있는 경우만 입수된다. 이는 한 사람에 대해 종합소득세(금) 자료와 종합소득 자료가 쌍을 이루지 못한다는 의미로서, 한 레코드상에서 세금과 소득 간의 관계가 잘 맞지 않는 경우도 발생할 수 있다.
- 연계키가 없는 경우는 소득세를 0으로 간주하며, 최종 소득세는 지방소득세 (10%)¹⁴⁾를 추정하여 작성한다.
- 기타소득은 통합소득 작성에 사용되지 않는다. 다만, 종합소득신고에서 기타소득이 있는 사람은 기타소득의 세금이 종합소득세액에 포함된다. 이는 그 영향은 적지만 한 사람의 최종 소득과 소득세 간 관계 불일치의 원인이 될 수도 있다. 행정소득 중 기타소득만 있는 사람은 기타소득의 소득세도 최종 소득세에 사용하지 않는다. (극소수 해당). 하지만 기타소득외 다른 소득도 있을 경우에는 자료처리의 복잡성을 고려하여 기타소득세를 제외처리하지 않은 종합소득세를 사용한다.(기타소득 외 다른 소득도 있는 경우 약 160여명 수준)

14) 적용 최종 소득세=소득세×1.1; 농어촌특별세는 발생비율이 작아 고려하지 않음

□ 원천징수(연말정산) 소득세 자료 처리

- 원천징수(연말정산)에 의한 소득자료는 종합소득자료와 달리, 한 사람에 대한 소득과 소득세 자료가 쌍을 이루기 때문에 자료처리가 상대적으로 수월하다.
- 소득세의 경우 근로소득, 사업소득, 연금소득, 금융소득에 해당하는 소득세를 모두 더한다. 소득자료에서처럼 연말정산과 원천징수자료에 소득세가 모두 있으면 연말정산의 소득세를 우선적으로 사용한다.
 - 다만, 기타소득·퇴직소득은 통합소득 작성에 사용하지 않으므로 이에 해당하는 소득세도 사용하지 않는다. 이로 인해 기타소득에 포함된 사업소득세는 미미하게나마 누락 될 여지는 있다.
 - 지방소득세(연금소득의 결정세액 외)가 별도로 포함되지 않거나 자료가 입수되지 않은 경우는 지방소득세(소득세의 10%)를 추정한다.

□ 종합소득세와 원천징수(연말정산) 소득세 가구원별 통합

- 소득세 처리방법은 소득처리방법과 동일하다. 종합소득신고와 원천징수(연말정산) 자료에 모두 소득세가 있으면 종합소득세 자료를 선택한다.
- 종합소득으로 신고된 금융소득은 종합소득세와의 중복여부를 고려해야 한다.

④ 금융소득

- ① 금융소득은 원천징수와 종합소득신고 대상 등의 2개 자료로 입수된다.

구분	입수항목
원천세	지급액(소득금액) 합산 금액
종합소득세	이자소득금액, 배당액

② 통합자료 처리

- 원천세와 종합소득세의 금융소득은 과세대상 구분이 다르다. 원천분은 일반과세와 비과세·분리과세 소득으로, 종합분은 원천징수되거나(일반과세) 원천징수안되는(원천분에서 파악되지 않거나 해외 금융소득)소득으로 구분된다.

- ※ 금융소득종합과세 : 비과세 및 분리과세 되는 금융소득을 제외한 연간 이자소득과 배당 소득 세전 합계가 2천만원을 초과하는 경우 다른 종합소득과 합산하여 종합과세
- 국외 금융소득처럼 원천징수 되지 않은 금융소득은 2천만원 이하라도 종합소득세 신고
 - 금융소득 구성: 비과세·분리과세 소득+원천징수되는 일반과세 소득+원천징수되지 않은 소득

	비과세·분리과세 소득	원천징수된 일반과세 소득	원천징수 되지 않은 소득
원천	A	B	-
종합	-	B	C

- 두 자료가 모두 입수된 사람의 경우 두 자료 중 최대금액을 사용한다.
- 원천분은 소득 구분코드가 입수되지 않으며, 종합분은 소득 구분코드는 입수되나 중복소득을 분리할 수가 없어 둘 중 큰 금액을 사용한다. 다만 큰 금액 사용시 원천분의 비과세·분리과세 소득 또는 종합분의 원천징수 안된 소득이 누락될 수 있으나 그 숫자는 무시할 수준이다.¹⁵⁾

<원천분과 종합분 금융소득 중복가구 처리(예시)>

	원천분 금융소득(만원)			종합분 금융소득(만원)			금융소득 통합	비고
	합계	일반과세	비과세·분리과세	합계	원천징수 소득	원천징수 안된 소득		
A가구	2,100	2000	100	2,000	2,000	0	2,100	
B가구	2,500	2000	500	2,300	2,000	300	2,500	종합분 소득 300누락
C가구*	900	500	400	1,500	500	1,000	1,500	원천분 소득 400 누락

* 총 금융소득 900만원이지만 원천징수되지 않은 소득이 있을 경우 종합소득 신고대상임

15) 종합분의 소득구분코드 활용 검토

- 종합분 소득구분코드를 통해 원천징수 되지 않은 금융소득(C)만 분리하고 원천분 금융소득(A, B)과 합산하여 중복처리하는 방안이 있음
- '15년 기준 종합분에서 '원천징수 되지 않은 소득'으로 명시된 코드로 입수된 소득(12, 15등 3개)을 분리하면, 가구원 5명이 해당되며, 소득금액은 만원 미만에 해당함
- 위 코드 이외는 원천징수된 소득으로 간주하면, 종합분의 원천징수된 소득보다 원천분 금융소득이 크거나 같아야 하나, 14가구는 종합분이 크게 나타남
- 이에 종합분 구분코드로 분리되지 않는 소득이 있을 것으로 판단되어 활용하는데 한계

5 공적이전소득

- 공적연금, 기초연금, 장애수당, 장애인연금, 가정양육수당, 아동수당, 주거·생계·교육급여, 자활급여(근로소득으로 간주), 근로·자녀장려금 등 연간으로 정부정책에 의해 지급되는 현금급여액 자료를 말한다.
 - 정부정책상 복지지원은 현금급여와 바우처와 같은 현물급여 방식도 있다. 다만, 가계금융복지조사에서는 현금소득 측정이 목표이므로 현물급여 방식¹⁶⁾에 의한 소득 행정자료는 여기서는 다루지 않는다.
- 공적이전소득과 관련한 가계금융복지조사 측정항목은 아래 표와 같다.
 - 공적연금소득은 국민연금과 기타연금으로 구성된다. 여기서 기타연금은 공무원연금, 사학연금, 별정우체국연금, 군인연금 4개 연금을 말한다.
 - 양육수당, 기초생활보장급여 및 근로자녀장려금은 가구내 개인에게 지급되는 급여이다. 이 경우 급여수령자와 급여대상자가 다를 수 있으므로 행정자료는 가구의 연계특성을 고려하여 활용기준을 정한다. 가구원 마이크로자료에서는 해당 가구소득을 가구주의 소득으로 입력하는 방식으로 처리한다.

조사항목	행정자료를 통해 입수하는 자료	소관 부처
공적연금	<ul style="list-style-type: none"> ◦ 국민연금 ◦ 공무원연금, 사학연금, 별정우체국연금, 군인연금 	국민연금공단 등 각 연금공단
기초연금	<ul style="list-style-type: none"> ◦ 기초연금 	보건복지부
장애수당	<ul style="list-style-type: none"> ◦ 장애인연금 ◦ 장애수당 ◦ 장애아동수당 ◦ 장애인급여 및 장애아동수당 	보건복지부
양육수당	<ul style="list-style-type: none"> ◦ 가정양육수당, 가정위탁아동양육보조금지원 ◦ 입양아동양육수당, 장애입양아동양육보조금 ◦ 한부모가족양육수당 ◦ 아동수당('18년 이후, 7세이하 미취학 아동) 	보건복지부 보건복지부 여성가족부
맞춤형 기초생활보장급여	<ul style="list-style-type: none"> ◦ 생계급여, 교육급여, 주거급여(현물제외), 해산·장제급여 ◦ 자활급여(근로소득으로 처리함) 	보건복지부
근로·자녀장려금	<ul style="list-style-type: none"> ◦ 근로·자녀장려금 	국세청

16) 정부의 복지지출측면에서 볼 때 현금지급보다 현물지급방식에 의한 지출비중이 더 크다. 그러나 현물급여지출이 가구소득에 미치는 영향을 파악할 만한 통계자료는 아직 없는 것 같다. 통계청은 2018년 '사회적현물을 반영한 가구소득' 측정연구를 바탕으로 한 시험작성 결과를 발표한 바 있다.

① 공적연금

- 공적연금소득은 국세청과 각 연금공단 자료가 후보자료가 될 수 있다. 국세청 연금소득은 소득세를 부과할 목적으로 연금공단에서 입수된 자료이다. 즉, 국세청 연금소득자료는 과세대상 소득에 국한될 소지가 있다.
- 가계금융복지조사에서는 과세 또는 비과세대상의 공적연금소득을 모두 파악하므로 각 공단의 해당자료가 더 적합하다고 판단한다.
 - 아래 표는 공단자료와 국세청자료를 비교한 것이다. 대상가구원은 거의 대부분 두 자료에 모두 포함되지만, 공단자료에만 있는 경우가 약 330에서 350여명 정도로 더 많다. 국세청자료에만 있는 10여건의 경우는 두 자료 간의 입수시점이 다르기 때문일 것으로 생각한다.

< 연금소득에 대한 공단자료와 국세청 자료 현황 >

(단위: 명, %)

	계	둘다있음				공단만있음			국세청만		
		가구 원수	가구 원수	조사 평균	공단 평균	국세 평균	가구 원수	조사 평균	공단 평균	가구 원수	조사 평균
2016	46,870	4,884	524	615	600	345	2,253	2,507	14	1,633	1,839
2017	44,887	5,207	546	635	623	333	2,285	2,548	10	2,041	2,151

② 기초연금 등 사회수혜금 항목

- 기본적으로 조사자료와 행정자료 간 포괄범위가 일치한다. 공적이전소득에 대해 행정기관은 현금급여의 지급결과를 기록하며, 조사에서는 가구가 연간으로 지급받고 있는 현금소득을 직접 물어보는 방식이다.
- 공적이전소득은 자료 간 개념이 동일하고 개별 행정자료를 입수한다. 따라서 같은 사람에게 대해 여러 자료가 있으면 **조사항목별 각 자료를 합산하면 된다.**
 - (예) 장애수당 행정자료는 장애인연금, 장애수당 등 입수자료 합산¹⁷⁾

17) 공적연금, 기초연금 및 장애수당은 수급자와 조사 대상자가 동일하므로 가구원 단위로 자료를 처리할 수 있다. 양육수당, 맞춤형기초생활보장급여 및 근로자녀장려금의 경우는 실제 급여 수령자와 수급대상자가 다른 경우도 있으므로, 가구원 단위의 자료처리 후 가구단위로 한번 확인하는 단계를 거쳐 최종 대체보완을 결정한다.

⑥ 공적이전지출

- 조사표에서 세금항목은 소득세, 재산세, 자동차세, 기타세금으로 구성되며, 이중 행정자료 검토대상 항목은 소득세, 재산세, 자동차세가 해당한다.¹⁸⁾
 - 세금항목 중 소득세만 행정자료를 활용하고 나머지 항목은 활용하지 않는다.
- 연금 기여금 및 사회보험료 지출항목 중 각 공단에 제출하는 공적연금 기여금과 건강보험료, 고용보험료 등 경상적인 이전지출 비용이 검토대상이다.
 - 공적이전지출항목 중에서 국민연금 기여금, 기타연금 기여금, 건강보험료에 대해서는 행정자료를 활용하지만, 고용보험료에 대해서는 사용하지 못한다.

□ 소득세

- 소득세 자료처리방법은 소득처리에서와 유사하므로 별도로 설명하지 않는다.

□ 재산세와 자동차세

- 재산세와 자동차세의 경우는 이용가능한 행정자료에 대해서 활용가능성 여부를 검토해 보았다. 그러나 두 자료는 자료구조 등 항목구성 방식에 차이가 있어서 행정자료를 활용하기가 사실상 어렵다.
 - 조사개념에서는 재산세와 자동차세는 가계용으로 지출한 세금에 대해서만 측정하는데, 행정자료 구조에서는 사업용과 가계용을 따로 분리하고 있지 않다. 이에 가계용과 사업용을 분리하지 않은 채 행정자료를 사용할 경우 세금이 과대 추정될 수 있어서 행정자료의 활용을 보류한다.

18) 조사에서 상속, 증여, 취득, 양도에 의한 소득세 등 부정기적인 세금은 조사대상이 아니며, 또한 사업상 납부한 부가가치세 등 간접세, 과징금 및 벌금도 조사대상이 아니다.

< 분석자료 >

구분	재산세	자동차세
조사 자료	<ul style="list-style-type: none"> - 조사항목 : 부동산 유무 * 주거용 주택 및 비주거용 건물 (상가, 빌딩), 토지(논밭, 임야, 대지), 해외부동산 등 모두 조사 - 재산세 · 재산세(7, 9월) + 종합부동산세(12월) 	<ul style="list-style-type: none"> - 조사항목 : 자동차 소유여부, 차종, 용도(가계용/사업용) * 개인사업용 소유 포함(택시, 트럭 등) * 업무용 회사(법인) 소유차량 제외 - 자동차세 · 자동차세(분납 6,12월, 일시납 1,3,9월) + 환경개선부담금(3, 9월)
행정 자료	<ul style="list-style-type: none"> · 가계용/사업용 구분 없음 · 재산세(7, 9월) 이용, 종합부동산세 미이용 	<ul style="list-style-type: none"> · 가계용/사업용 구분 없음 · 자동차세(6, 12월) 이용, 환경개선부담금 미이용

□ 건강보험료

① 연계기준 및 가입자격 분류

- 건강보험료는 직장가입자와 지역가입자로 구분된다. 직장가입자의 경우 개인단위 자료이므로 두 자료를 연계하는데 문제는 없다. 하지만 지역가입자의 경우는 세대 단위로 개인, 세대, 가구 특성을 고려하여 자료처리해야한다.
 - * 지역가입자에는 고용원이 없는 자영자, 무직자 등이 포함될 수 있음
- 직장가입자는 가구원의 보험자격이 중복이더라도 실제 본인의 보험료 내역과 연계할 수 있으며, 조사자료와 행정자료에서의 측정개념도 동일하다. 임의가입자, 기타 가입자의 경우도 일대일 매칭할 수 있다.
- 지역가입자는 세대주 기준의 자료로서 보험증번호 기준으로 세대의 월별 보험료의 최대값을 세대주에 연계한다. 이때 조사응답자와 행정자료상 연계자(세대주) 간에 연계오류도 발생할 수 있다.
 - 지역가입자의 복잡한 자료구조를 감안하여, 연간 가입자격 중 세대주 또는 세대원의 자격과 월별자료를 입수함으로써 두 자료 간 연계를 보다 정확하게 하고자 했다.

- 지역가입자는 가구원의 소득과 재산이 건강보험료 산정기준에 포함되며, 가구 사정상 보험료 감액을 위해 가구원을 타 가구의 지역가입자나 직장피부양자로 전입할 가능성도 자료분석 내용에 포함하였다.
- 결론적으로 직장가입자와 달리 지역가입자는 자료구조가 복잡하여 자료처리과정에서 처리오차가 발생할 수 있으나, 조사에서의 응답착오 가능성에 비하면 그 수준은 매우 작을 것으로 예상된다.
- 이에 상세코드 등의 세보정보를 사용하여 보다 정확한 연계기준을 마련함으로써 복잡한 자료구조에 의한 자료처리 과정에서 발생할 수 있는 오차를 최소화하고자 하였다.

② 연계가구원 특성 : 조사자료와 행정자료 연계 분석

- 두 자료를 연계한 후에 행정자료에만 보험료금액이 있는 경우가 있는데,
 - 이 경우는 실제로는 보험료를 납부했지만 조사 사각지대(따로 사는 가구원 등)에 있는 가구원에 대해 조사응답이 누락되었을 가능성도 있다.
 - 자료분석결과에 의하면 무직이거나, 근로소득과 사업소득이 모두 없거나, 소득 1천만원 미만자 그룹에서 행정자료에만 보험료가 있는 사람들의 비율이 높고, 직장가입자격자 중 2,072명이 행정자료에만 보험료가 있는 것으로 나타났다.
- 결합자료에서 조사자료에만 보험료금액이 있는 경우에 대해 여러 가지 이유가 있을 것이다.
 - 실제로는 가구에서 보험료를 납부하지 않았는데 응답과정에서 기억착오가 있거나 조사원의 기입과정에서 조사착오 등도 완전히 배제하기는 어렵다.
 - 이러한 가능성은 고용원이 없는 자영자, 근로소득 또는 사업소득만 있거나 소득 1천만원 이상~5천만원 미만에서 조사자료에만 보험료가 있는 가구원 비율이 높다는 점에서 가늠해볼 수 있다.

< 2017년 자료의 비교(% , 만원) >

		전체		연계키 있는 경우			
				행정만있음		조사만있음	
		구성비	차이평균*	구성비	차이평균*	구성비	차이평균*
전체		100.0	(3.9)	100.0	55.9	100.0	(113.4)
상용근로자		24.9	(8.8)	9.3	65.1	24.3	(104.3)
임시일용		9.2	(2.6)	19.0	41.6	17.3	(71.9)
자영자		14.7	(21.1)	14.3	84.0	44.0	(134.2)
무직 및 기타		51.2	3.2	57.4	52.0	14.4	(115.3)
경상소득 구간	1천미만	53.6	4.2	66.8	45.8	13.5	(73.5)
	3천미만	22.6	(6.5)	25.3	57.2	44.5	(80.4)
	5천미만	13.3	(20.2)	5.3	95.4	26.8	(131.8)
	7천미만	5.3	(26.2)	1.4	147.5	9.0	(177.8)
	1억미만	3.2	(25.8)	0.6	253.3	4.1	(253.3)
	1억이상	2.0	13.4	0.6	355.7	2.2	(280.8)
근로소득 사업소득	둘다 있음	0.7	(15.6)	0.7	65.3	1.4	(117.1)
	근로만 있음	35.2	(6.9)	29.1	49.7	43.6	(88.9)
	사업만 있음	11.5	(28.1)	9.4	84.4	42.3	(133.7)
둘다 없음		52.6	3.6	60.8	54.2	12.8	(129.5)

* 괄호로 표시된 수치는 음수인 경우임

□ 고용보험료

- 고용보험료는 개인이 납부해야 할 금액을 보험공단에서 개인에게 고지한 금액(고지 금액)자료를 검토해 볼 수 있다. 단, 고용보험 자진신고 사업자의 종사상 자료는 입수할 수 없었다.
 - 고용보험료의 경우 납부금액과 고지금액에 차이가 있을 수 있으며, 납부시점, 일괄 납부 및 당해 연도 납부누락 사항 등에 따라 자료구조도 복잡한 것 같다.
- 분석결과 행정자료상의 고지금액보다 조사자료가 과대추정되며, 조사에서 응답한 대상자의 특성도 건강보험료 등 다른 자료들에 비해 두 자료 간의 특성 차이가 꽤 큰 것으로 파악되었다.¹⁹⁾
 - 한편, 자진신고 사업장 종사자의 고용보험료 자료(고용보험 본인부담분의 11%)를 사용할 수 없다는 점에서 조사자료와 비교하기가 어려웠다. 따라서 고용보험료의 경우 행정자료 활용을 보류하고 조사자료를 사용한다.

19) 예를 들면, 행정자료상 고용보험 가입대상자가 아닌 공무원, 자영자(고용주), 주식회사 대표 및 이사 등에서도 고용보험료가 조사되며, 이들 중 고소득자가 많다. 이에 대해 향후 조사착오 또는 행정기록과 조사자료 간의 시점차이 등 다양한 이유를 검토해야 할 것이다.

PART

V

행정자료 활용기준 및 적용결과 요약



PART



V 행정자료 활용기준 및 적용결과 요약

- 가계금융복지조사에서는 2017년에 행정자료를 반영한 소득분배지표를 발표하였고, 2018년에는 가구소득에 대해서도 행정자료 활용 결과를 발표하였다.
- 여기서는 지금까지 설명한 소득원천별, 이전지출 항목별 행정자료 활용기준을 요약하고, 행정자료를 반영하여 분석한 결과를 소개한다. 분석자료에는 2015~2016년 자료가 주로 사용되었고, 활용기준은 2019년 기준이다.

1 행정자료 활용현황

- 2019년 조사항목에 대한 행정자료의 활용현황은 아래표와 같다.

* 아래 표에서 'X'는 행정자료 미활용, 음영은 무응답에만 이용되었음을 의미함

조사 항목		2019년 행정자료 활용 내역	
소득	근로소득	○ 근로소득(종합소득세, 원천세), 일용근로소득, 자활급여	
	사업소득	○ 사업소득(종합소득세, 원천세)	
	재산소득	금융소득	○ 금융소득(종합소득세, 원천세)
		임대소득	○ 사업소득 중 부동산 임대소득(종합소득세)
	공적 이전소득	공적연금	○ 연금지급액
		기초연금	○ 기초연금
		양육수당 (출산장려금 포함)	○ 양육수당, 양육보조금, 한부모가족아동양육수당, 아동수당('18년 정책) 등
		장애수당 (장애아동수당, 장애인연금 등)	○ 장애인연금, 장애수당, 장애아동수당·장애인급여
		기초생활보장지원금	○ 생계급여, 주거급여, 교육급여, 해산·장제급여
	근로·자녀장려금	○ 근로·자녀장려금	
지출	세금	소득세	○ 소득세(종합소득세, 원천세, 일용근로소득세)
		재산세	X
		자동차세	X
	공적연금· 사회보험료	국민연금기여금	○ 국민연금(연금공단)
		기타연금기여금	○ 공무원연금, 사학연금, 별정우체국연금, 군인연금(공단)
건강보험료	○ 건강보험료, 노인장기요양보험료		
고용보험료	X		

- 행정자료를 통한 조사자료 수정·보완 기준은 아래표 같이 정리할 수 있다.
 - 일부항목을 제외한 대부분의 소득 및 이전지출 항목에 대해 행정자료로 조사자료를 대체한다. 다만 근로소득의 경우 미과세 소득자 등 저소득자에 대해서는 조사자료가 활용될 수 있으며, 사업소득(국세청에 신고된 임대사업소득 포함)은 무응답에 대해서만 행정자료로 대체한다.

항목	항목별 보완 기준 (2019년 조사결과 공표시 적용 기준)
근로 소득	<ul style="list-style-type: none"> ○ (기본방향) 행정자료로 조사자료 대체 또는 수정 보완 ○ (예외사항) 조사자료가 미과세 금액이하인 가구원의 경우, 조사자료와 행정자료 중 큰 값 사용. 또한 조사자료만 있는 경우는 행정자료 사업소득과의 오분류 가능성을 교차검증결과에 따라 조사자료 사용 가능 ○ (고려사항) 소득미신고자는 국세청 신고소득 외의 소득에 대해 조사표에 응답할 수 있으므로, 두 자료 중 큰 값을 선택함. 다만, 행정자료 사용을 원칙으로 하므로 미과세기준 이상의 소득이 있는 경우 조사에서 응답한 비과세 또는 미과세 소득이 누락될 소지가 있음
사업/ 임대 소득	<ul style="list-style-type: none"> ○ (기본방향) 기본적으로 조사자료 사용하며, 무응답의 경우만 행정자료로 대체 ○ (고려사항) 가구원 간 사업운영자와 명의자가 다를 수 있어 가구단위에서 보완
금융 소득	<ul style="list-style-type: none"> ○ (기본방향) 행정자료로 조사자료 전면 대체 보완 ○ (예외사항) 빌려준 돈이 있는 경우는 조사자료와 행정자료 중 큰 금액 사용
공적 이전 소득	<ul style="list-style-type: none"> ○ (기본방향) 행정자료로 조사자료 전면 대체 보완 ○ (해당항목) 공적연금, 기초연금, 장애수당, 맞춤형기초생활보장, 근로·자녀 장려금, 양육수당 (출산장려금) * 출산장려금은 지자체 지원으로 행정자료 미입수 ○ (예외사항) 양육수당 중 출산장려금이 있는 경우는 두 자료 중 큰 값 사용
공적 이전 지출	<ul style="list-style-type: none"> ○ (기본방향) 행정자료로 조사자료 전면 대체 보완 ○ (해당항목) 소득세, 공적연금기여금(국민연금, 공무원연금 등 5개 연금), 건강보험료

* 2019년 가계금융복지조사 참고보도자료에 수록

- 소득원천별 행정자료 활용 정도
 - 매년 약 92~95%의 가구원에 대해 연계기를 확보하며 이 연계기에 의해 두 자료를 연계한다. 소득원천별 행정자료의 활용수준도 조금씩 다르다.
 - 2016년 자료에서 근로소득은 입수한 행정자료의 약 89.1%, 사업소득과 임대 소득은 각각 약 41.6%와 28.1%가 활용되었다.

- 공적연금, 맞춤형기초수급금의 행정자료 활용비율은 각각 96.9%, 93.6%로 매우 높다. 기초연금과 장애수당의 경우 두 자료에서의 금액 일치율이 높은 항목들에 해당한다. 아래의 표에서 두 항목의 행정자료 활용률이 약 63% 정도로 낮아 보이지만 금액이 같은 경우를 감안하면 거의 대부분 행정자료가 사용된다고 볼 수 있다.
- 금융소득 및 소득세는 다른 항목들에 비해 무응답 비율이 높은 편으로 무응답이 대체되는 비율은 각각 51%와 21% 수준이다. 이 두 항목은 전체적으로 5천원 미만의 행정자료를 제외하면 행정자료의 약 99%가 조사자료를 대체하는데 활용된다고 할 수 있다. * 가계금융복지조사는 만원단위로 조사하고 집계함

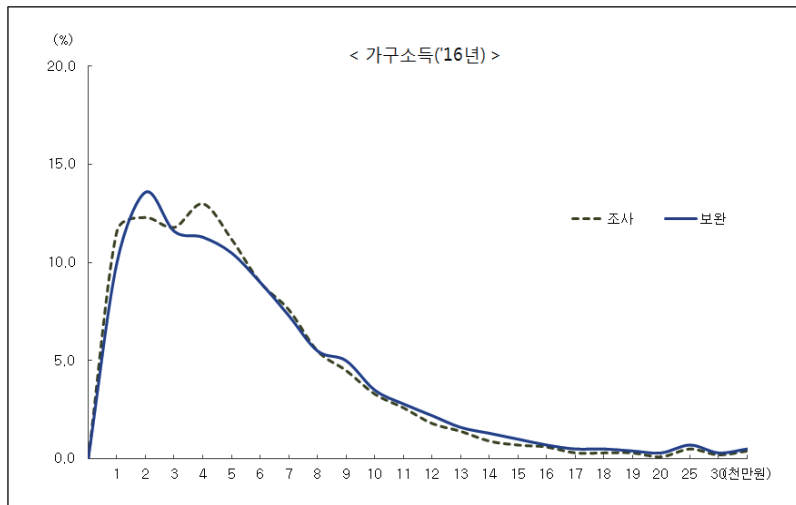
< 소득관련 항목별 입수 및 활용 비율, 2016년 기준 >

보완단위	조사항목명	무응답대체비율(%)	행정자료 활용비율(%)
가구원	근로소득	10.4	89.1
	사업소득	7.0	41.6
	금융소득	51.1	99.6
	임대소득(수입)	0.6	28.1
	공적연금	2.9	96.9
	기초연금	5.7	62.8
	장애수당	0.3	62.7
	소득세	21.2	99.4
	국민연금기여금	9.1	97.3
	기타연금기여금	0.5	96.4
가구	양육수당	2.9	69.2
	맞춤형기초생활보장	0.9	93.6
	근로자녀장려금	10.5	98.5

- ▶ 대상자수 : 입수대상자수로서, 조사응답자 중 입수명부에 포함된 가구원(또는 가구)수
 - 근로소득, 사업소득, 임대소득, 공적연금은 15세 이상 가구원으로 한정하여 계산
 - 기초연금은 65세 이상으로 제한하여 계산하고, 나머지 항목은 전 연령 대상으로 계산함
- ▶ 무응답대체비율: 조사무응답 중 행정자료로 대체된 가구원수(가구) /조사대상 가구원(가구)
- ▶ 행정자료 활용비율: 행정자료로 보완된 대상자수(금액>0)/행정자료가 입수된 대상자수
 - 조사자료에는 금액이 있는데 행정자료가 0이어서 0값으로 보완된 경우는 미포함
- ▶ 금융소득: 전 가구원 연령대상으로 어린이들의 5천원미만도 포함되어 있음
- ▶ 소득세: '15년까지는 가구단위, '16년부터 개인단위 집계함('16년부터 개인단위로 자료수집)
 - * 건강보험료는 '17년부터 활용하고 있으며 연계키가 있는 경우 행정자료로 전면대체

2 행정자료 보완결과

- ※ 이하 통계표상의 숫자는 연구자료를 분석한 결과가 포함되어 있어 국가통계포털에서 제공되는 숫자와 다를 수 있음
- 행정자료로 조사자료를 대체·보완하면, 보완 전에 비해 전체적인 가구소득이 상향 조정되는 모습이며 평균값과 중위값도 상승 이동한다. 소득구간별로는 소득구간 2~3천만원미만, 8천만원이상 구간에 속한 가구비율이 보완 전에 비해 증가하며, 이러한 현상은 연도별로 유사한 것으로 나타났다.



- 2016년 자료의 경우 행정자료로 보완 후의 가구당 평균소득은 9.2%, 중위소득은 6.1% 각각 증가하였다.
 - 소득원천별로는 근로소득이 223만원(3,276→3,498만원)으로 가장 많이 증가했고, 재산소득은 67.3%(209→350만원)로 가장 큰 비율로 증가했다. 재산소득의 경우 금융소득이 행정자료로 보완됨으로써 크게 상승한 것으로 보인다.
 - 공적이전소득을 구성하고 있는 대부분의 조사항목들은 두 자료의 개념이 거의 일치하므로 근로소득이나 금융소득처럼 소득이 크게 증가하는 현상은 상대적으로 덜하다. 다만, 무응답 또는 응답착오에 대한 조사값들이 수정되면서 조사자료의 품질이 상당히 개선되었을 것으로 본다.

< 2016년 기준 행정자료 활용 전과 후의 가구소득의 평균 (단위: 만원) >

소득원천	조사(만원)	보완(만원)
경상소득	5,020	5,478
근로소득	3,281	3,499
사업소득	1,149	1,217
재산소득	212	353
금융소득	30	153
임대수입	166	183
공적이전소득	295	325
공적연금	175	195
기초연금	54	55
소득세	175	245

- 조사자료가 보완됨에 따라 소득분배지표 중 지니계수는 2016년 처분가능소득 기준으로 0.342에서 0.355로 증가하였고, 상대적 빈곤율은 16.1%에서 17.6%로 증가하였다.
- 행정자료를 통해 저소득층 가구에서 누락하였거나 실제보다 낮게 응답한 소득이 고소득층에서보다 더 보완되는 경향이 있다. 전반적으로 저소득층에서는 일용 근로소득, 사업소득 등의 작은 소득금액에 대해 무응답하는 경향이 있고, 고소득층에서는 실제보다 과소하게 응답하는 경향이 있다.
- 또한 행정자료는 가구주가 60세 이상인 가구에서의 소득을 보완하는데 크게 기여하는 것으로 나타났다. 이 연령대 가구에서 근로소득, 사업소득, 금융소득은 조사에서 거의 '0'이거나 아주 소액으로 응답된 경우가 많았으나 이에 대해 행정자료로 크게 보완되었다.(다음 표①)
- 특히 60세 이상의 고연령층 가구에서 실제로 금융자산은 있음에도 불구하고 금융소득이 없다고 응답한 가구가 많았는데, 행정자료를 통해 이러한 무응답들이 크게 보완되었다.(다음 표 ①)

< ① 가구소득이 0에서 행정자료로 보완된 후의 연령별 소득변화 >

소득원천별	가구주 연령별	조사소득 0에서 0이상으로 보완된 가구수	
		'15년	'16년
근로소득	계	977	1,054
	30세미만	7	12
	30대	86	74
	40대	192	175
	50대	197	232
	60세 이상	495	561
사업소득	계	1,743	1,871
	30세미만	54	60
	30대	310	330
	40대	493	515
	50대	511	558
	60세 이상	375	408
금융소득	계	12,679	12,684
	30세미만	257	288
	30대	1,997	1,903
	40대	2,994	2,810
	50대	3,161	3,138
	60세 이상	4,270	4,545

- 다음 표 ②를 통해 금융소득 보완이 소득분위별, 가구주 연령대별로 전체가구 소득 증가에 어떤 영향을 미치는지 이해할 수 있다. 표에서 '금융소득 효과'란 여러 소득항목 중 금융소득만 행정자료로 보완되었을 때의 전체 가구소득의 증감율을 의미한다. * 금융소득 효과 = (보완후 평균-보완전 평균)/보완전 평균
- 금융소득은 소득분위가 낮을수록, 또는 가구주 연령이 높을수록 보완 효과가 더 크게 나타난다. 이는 누락된 소액의 자료들이 행정자료로 대부분 보완된 효과라 할 수 있다.
- 또한 소득 4, 5분위에 속한 60대 가구주 가구의 경우 다른 연령대에 비해 금융소득 보완 효과가 상당히 높다. 이는 이 연령대에서 조사 금융소득이 상당히 누락되었을 가능성을 시사한다 (다음 표③).

< ② 2015년 금융소득보완효과*에 따른 경상소득 보완비율* >

소득 5분위	가구주 연령대별 금융소득 효과(%)					
	전체	30세 미만	30대	40대	50대	60세 이상
전체	2.9	0.9	1.4	1.8	2.8	6.3
1분위	8.9	1.8	5.9	3.3	4.9	10.9
2분위	4.4	0.5	1.3	2.1	3.4	9.0
3분위	2.5	0.7	1.2	1.5	2.5	5.3
4분위	2.3	0.9	1.5	1.6	2.6	4.9
5분위	2.6	2.0	1.3	1.9	2.8	4.6

* 금융소득보완효과: (보완후 평균 - 보완전 평균) / 보완전 평균 × 100. 여기서 보완후는 금융소득만 행정자료로 보완한 가구소득, 보완전은 조사자료 가구소득을 의미함

< ③ 소득원천별 가구주 연령별 보완 전과 후의 평균소득 >

소득 원천	가구주 연령대	2015년 소득(만원)		2016년 소득(만원)	
		보완 전	보완 후	보완 전	보완 후
근로소득	계	3,199	3,401	3,276	3,498
	30세미만	2,943	2,880	2,939	2,876
	30대	4,208	4,216	4,369	4,323
	40대	4,210	4,431	4,423	4,637
	50대	4,011	4,358	4,228	4,659
	60세 이상	1,232	1,417	1,238	1,448
사업소득	계	1,119	1,177	1,149	1,218
	30세미만	202	235	184	211
	30대	778	825	822	877
	40대	1,465	1,536	1,476	1,579
	50대	1,663	1,739	1,731	1,805
	60세 이상	641	682	683	729
금융소득	계	31	172	30	153
	30세미만	2	32	1	24
	30대	10	81	15	68
	40대	25	130	22	119
	50대	37	204	35	179
	60세 이상	43	239	40	206

PART
VI

결론 및 향후방향



PART



VI 결론 및 향후방향

1 행정자료를 통한 조사효율성 제고

- 인구사회구조가 변해감에 따라 가구단위 통계조사도 영향을 받을 수밖에 없다. 민감한 조사항목의 경우 직접적으로든^(대면조사) 간접적으로든^(전화조사, 인터넷조사 등) 가구의 응답을 통해 직접 자료를 수집하는 것은 점점 더 어려워질 것이다.
 - 국제사회는 2000년 중반을 전후로 행정자료를 통해 어려운 조사현실을 극복하는데 관심을 가졌다. 2020년 현재 유럽, 북미, 호주, 한국 등 국가들에서 행정자료를 활용하여 국가통계를 작성하는 사례가 많아지고 있다.
- 가계금융복지조사에서도 가구소득에 대한 무응답 또는 응답착오 현상이 나타나고 있으며, 가구특성에 따라 응답 참여 및 성향도 달라지는 경향이 있다.

2 행정자료 활용시 주요 검토사항

- 그 동안의 경험에서 볼 때, 표본조사의 조사자료와 행정자료를 결합할 경우 조사자료에 연계하고자 하는 행정자료에 대해 충분히 이해하는 것이 무엇보다 중요한 것 같다.
 - 다출처 자료를 결합하는데 있어서, 각 자료가 갖는 속성과 구조가 다를 수 있으므로 각 자료의 특성이 제대로 반영될 수 있도록 명확한 자료처리 기준이 마련되어야 한다.
- 행정자료 활용시 행정자료 자체의 시계열 안정성과 행정자료를 결합한 통계자료의 시계열 안정성을 확인할 필요가 있다.
 - 행정자료는 정책 또는 제도변화에 민감하며 통계자료에도 영향을 줄 수 있으므로, 관련 정책 동향을 지속적으로 모니터링해야 하는 수고가 필요하다. 또한 행정자료의 활용은 과거자료에 대한 소급 및 시계열 단층 가능성도 고려할 필요가 있다.²⁰⁾

20) 신규 정책·제도로 행정자료를 조사자료에 결합해야 하는 경우나 자동차세나 재산세 등을 나중에 활용하게 될 경우에 전체 시계열에 대해 소급보완이 어려운 경우도 있을 수 있다.

3 행정자료를 통한 통계품질 개선 및 응답부담 경감 기대

- 가계금융복지조사에서는 2019년 기준으로 과거 시계열자료에 대해 동일한 행정자료 활용기준을 적용하여 조사자료를 보완하였다. 이에 따라 행정자료 활용으로 인한 시계열자료상의 불규칙성은 없다.
 - 사업소득, 재산세, 자동차세, 고용보험료 등과 아직 행정자료를 완전하게 사용하지 못하는 항목에 대해서는 자료처리방식 및 분류코드 불일치성 등으로 인해 당분간 활용하기는 쉽지 않을 것 같다. 이에 대해서도 가능한 시점에 보완하는 것을 전제로 꾸준히 검토해 나갈 예정이다.
- 가계금융복지조사의 경우 행정자료를 활용함으로써 정확성과 신뢰성 측면에서 통계품질이 상당히 개선되었다. 향후 행정자료 활용항목을 조사표에서 제외하거나 간소화하는 방법을 통해 응답부담도 줄일 수도 있을 것으로 기대한다.

4 다각적 틀에서 행정자료 활용방안 모색

- 행정자료는 현재 조사에 응답한 가구에 대해서만 사용되고 있다. 따라서 단위 무응답자 또는 표본추출시 커버하기 어려운 초고소득자에 대한 소득측정 문제는 여전히 숙제로 남는다.
 - 초고소득가구 포괄문제에 대해서는 소득DB 구축, 고소득자 밀집지역을 고려한 표본추출, 고소득층 사후보정 방법 등을 생각해 볼 수 있겠다. 초고소득자의 소득 포괄문제는 미시소득통계에서 오래된 이슈임에 틀림없다. 그러나 초고소득 가구의 조사협조도 및 응답가능성을 고려할 때 표본추출문제로 해결하기는 쉽지 않을 것으로 보인다.
- 마지막으로 ① 행정자료를 통한 통계품질은 행정자료의 품질이 얼마나 높은가와 밀접한 관련이 있고, ② 행정자료와 조사자료를 결합하여 통계작성을 작성할 경우 다출처 자료의 구조 및 속성에 대한 충분한 이해를 바탕으로 한 자료처리가 매우 중요하다는 점을 한 번 더 강조하고 싶다.

참고분석

조사자료와 행정자료의 특성 분석

- 근로소득
- 사업소득
- 금융소득
- 소득세



PART



참고분석 조사자료와 행정자료의 특성 비교 분석(요약)

- 행정자료 활용기준을 마련하기까지 탐색적 자료분석과 전문가적·경험적 노하우를 바탕으로 다양한 형태의 심층분석을 실시하였음. 여기서는 이용자로 하여금 조사자료와 행정자료의 구조 및 분포특성을 이해할 수 있도록 분석결과 중 일부를 설명하고자 함
 - 특히 근로소득, 사업소득, 금융소득, 소득세는 가구소득과 지출의 큰 부분을 차지하며 자료처리 과정도 타 항목들에 비해 상대적으로 복잡하여 이 항목들을 중심으로 설명함

○ 근로소득

① 조사자료와 행정자료가 갖는 자료의 특성 비교

- 조사 근로소득은 현물보수, 퇴직수당 등을 포함하여 측정하나, 행정자료는 현물보수는 포함하지 않음. 또한 퇴직소득 중 퇴직수당은 근로소득으로 퇴직일시금은 금융자산으로 조사되지만, 행정자료에서 두 항목이 분리되지 않음
 - 조사자료에 행정자료를 연계하는 경우 행정자료를 활용함으로써 조사오차를 줄일 수 있는 장점이 있지만 행정자료 활용에 따른 또 다른 오차도 발생할 수 있음에 유의

② 조사자료와 행정자료 간의 근로소득 유무 차이

- * 이하, 전체 가구원을 대상으로 결과에 대해 연계키가 있는 가구원 중심으로 설명
- 조사자료와 행정자료에 모두 근로소득이 있는 가구원은 전체 가구원 중 27%이며, 8.1%는 조사자료에만, 8.9%는 행정자료에만 근로소득이 있는 경우임
 - 2015년 자료에서 조사 근로소득(이하 조사근로소득)이 있는 가구원 중 행정 근로소득(이하 행정근로소득)이 있는 가구원 비중은 약 74.7%임
- 조사 근로소득은 없는데 행정 근로소득만 있는 가구원 중 약 49%는 일용근로소득만 있는 경우임. 이는 소액의 일용근로소득에 대해 가구에서의 무응답 가능성도 있음

< 전체 가구원의 근로소득 존재 여부(2015년 기준) >

(단위: %)

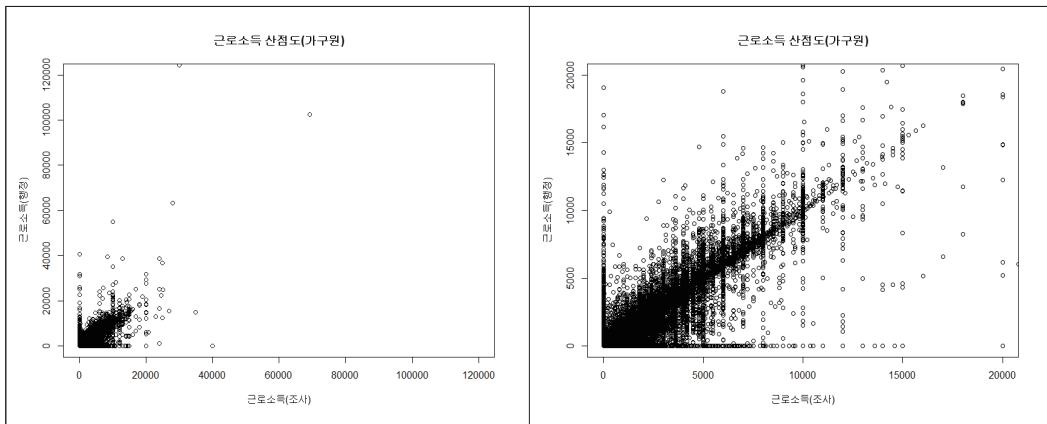
조사 근로소득	전체	입수명부 포함		입수 명부 미포함	전체	입수명부 포함		입수 명부 미포함
		행정 근로소득				행정 근로소득		
		있음	없음			있음	없음	
계	100.0	35.9	60.9	3.2	100.0	35.9	60.9	3.2
있음	36.2	27.0	8.1	1.0	100.0	74.7	22.5	2.8
없음	63.8	8.9	52.8	2.2	100.0	13.9	82.7	3.4

③ 조사자료와 과세자료의 특성 분석

① 두 근로소득(한쪽에만 있는 경우 제외)은 강한 양의 상관관계

- 두 자료의 근로소득 분포가 대체로 유사하며 직선을 벗어나서 산재해 있는 자료는 각 자료만의 특성이 보여주는 현상으로 이해됨. 즉, 과세자료 기준에서 보면 조사에서의 과대·과소 응답 가능성을, 조사기준에서 보면 신고누락 등의 오류 가능성을 가늠해 볼 수 있음

< 근로소득 산점도(가구원 기준으로 전체(왼쪽)와 2억미만(오른쪽)) >



② 양쪽 자료의 소득 유무에 따라 소득구간별 가구원 분포가 다름

- 2016년 자료의 각 근로소득 기준으로 볼 때, 전체적으로 1천만원 미만의 저소득 구간과 7천만원 이상의 고소득 구간에서 행정자료 기준의 가구원 비중이 높게 나타남

< 소득유무별 근로소득구간별 가구원 분포(2016년) >

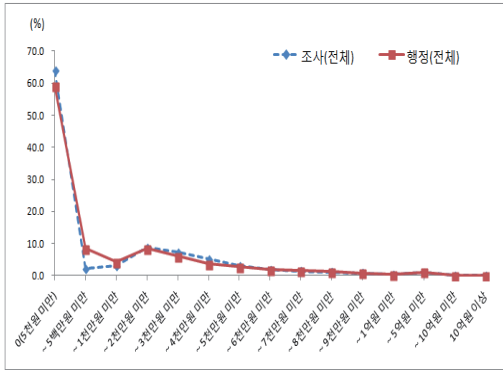
(단위: %)

근로소득구간	조사 전체	입수명부 포함 계		입수명부 포함				입수명부 미포함 조사있음
		조사	행정	둘다 있음		조사만 있음	행정만 있음	
				조사	행정			
비중 계	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
0(5천원 미만)	64.2	64.2	59.3
~5백만원 미만	2.1	2.1	8.5	4.1	9.2	16.3	54.9	6.0
~1천만원 미만	3.0	3.0	4.3	6.2	8.6	20.2	15.9	7.9
~2천만원 미만	8.7	8.8	8.5	22.7	22.4	35.2	16.8	20.8
~3천만원 미만	7.2	7.2	6.1	20.9	17.9	15.7	6.9	21.4
~4천만원 미만	5.2	5.1	3.7	15.7	11.3	6.8	2.3	18.8
~5천만원 미만	3.1	3.0	2.7	9.6	8.5	2.3	1.1	9.7
~6천만원 미만	1.9	1.9	1.9	6.1	6.0	1.1	0.7	5.5
~7천만원 미만	1.4	1.4	1.5	4.6	4.7	0.9	0.4	3.6
~8천만원 미만	1.1	1.1	1.2	3.6	4.0	0.3	0.2	2.4
~9천만원 미만	0.7	0.8	0.8	2.4	2.5	0.3	0.2	1.4
~1억원 미만	0.3	0.3	0.4	1.1	1.3	0.2	0.1	0.7
~5억원 미만	0.9	1.0	1.1	3.1	3.6	0.5	0.3	1.6
~5억원 이상	0.0	0.0	0.0	0.0	0.0	.	.	0.1

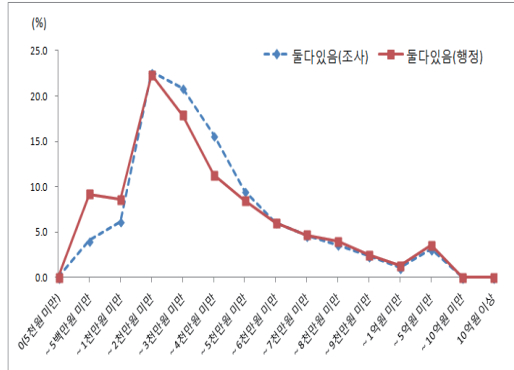
- (둘 다 있는 경우) 전체 가구원 분포에서와 같이, 행정자료 기준의 1천만원 미만과 7천만원 이상 구간 각각에서 조사자료의 동 구간에서보다 가구원 비중이 더 높게 나타남
- (행정 근로소득만 있는 경우) 약 55%의 가구원이 5백만원 미만에 속하며, 2천만원 미만 구간에 속한 가구원 비중이 전체의 약 90%로 소득이 낮은 구간에 다수의 가구원이 분포한 것으로 나타남
- (조사 근로소득만 있는 경우) 2천만원 미만 소득구간에 속한 가구원 비중이 약 70%로, 행정 근로소득만 있는 경우에서와 마찬가지로 높은 소득구간보다 낮은 소득구간에 포함된 가구원 비중이 더 높음

< 근로소득 구간별 가구원 분포(2016년) >

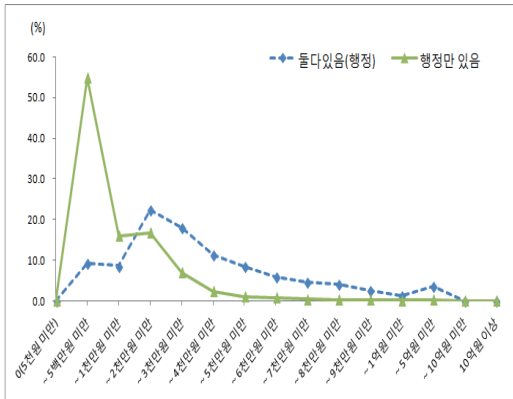
▶ 전체 가구원 대상



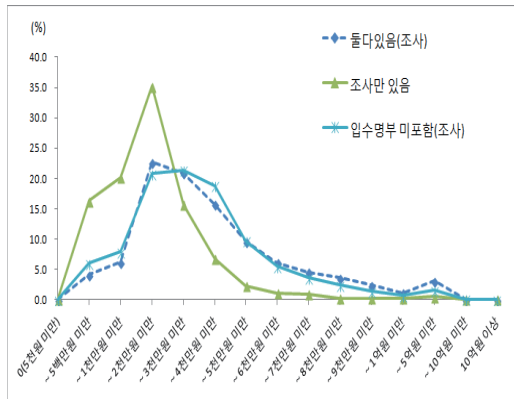
▶ 조사자료와 행정자료 둘 다 있는 경우



▶ 행정 근로소득만 있는 경우



▶ 조사 근로소득만 있는 경우



③ 자료원천에 따라 소득구간별 가구원 특성 분포가 다름

- 전체적으로 낮은 소득구간에서는 행정 소득구간별 가구원 비중이 조사 소득구간별 가구원 비중보다 높게 나타남
- (행정 근로소득만 있는 경우) 가구원의 종사상지위가 기타무직이거나 자영업자인 비중이 약 90% 차지하고, 연령대별로는 20대, 40~50대의 비중이 높고, 60세 이상인 경우도 약 21% 차지
- (조사 근로소득만 있는 경우) 가구원의 종사상 지위가 상용직이거나 임시일용직인 비중이 높고, 30~50대에서 높은 비중을 차지. 이들 그룹의 경우는 현물소득이 있거나 미신고 대상 소득이 있어 신고되지 않은 소득이 조사에서는 포착되었을 가능성도 있어서 경우에 따라 행정자료의 완전대체 활용을 조심할 필요가 있음
- (둘 다 있는 경우) 행정자료 기준 1천만원 미만 또는 초고소득 구간에 속한 가구원 비중이 조사자료 기준의 동일 구간에서보다 약간 높다. 이는 다른 국내외 사례에서도 보듯이 저소득 또는 고소득자일수록 소득을 과소하게 응답하는 성향에서 비롯된 것일 수 있음

< 소득유무별 가구원 종사상지위별 비중(2015년 기준) >

(단위: %)

구분	계	상용	임시 일용	자영업 (고용원有)	자영업 (고용원無)	무급	기타 종사자	기타 (무직 등)
행정소득만	100.0	3.9	2.7	6.3	16.9	1.7	2.2	66.3
조사소득만	100.0	37.7	40.8	0.4	1.7	0.5	0.3	18.6
둘 다	100.0	73.6	18.0	0.3	0.9	0.1	0.1	7.1

< 소득유무별 가구원 연령별 비중(2015년 기준) >

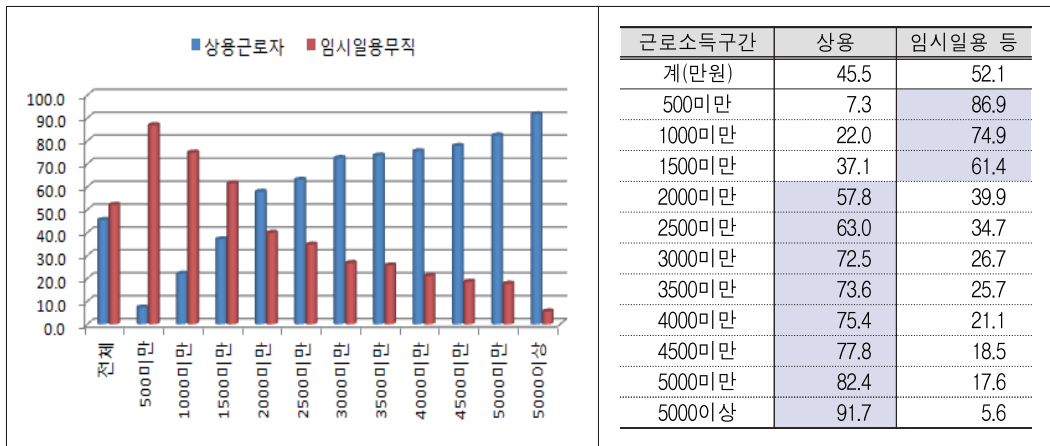
(단위: %)

구분	계	20세 미만	20세~ 24세	25세~ 29세	30세~ 39세	40세~ 49세	50세~ 59세	60세~ 64세	65세~ 69세	70세 이상
행정소득만	100.0	9.9	17.0	8.1	12.3	15.1	16.3	6.8	5.9	8.7
조사소득만	100.0	0.9	12.8	10.4	22.0	20.0	18.8	5.7	4.3	5.1
둘 다	100.0	0.6	4.7	8.9	24.6	27.5	22.6	5.5	3.2	2.4

④ 두 자료 모두 근로소득이 있는 경우 저소득과 고소득자의 응답성향이 다름

- 저소득자와 고소득자 간의 응답성향에 차이가 있는 것으로 보임. 이러한 현상에 대해 조사자료 관점에서 설명하자면 응답오차가 큰 요인일 수 있음. 분석에 따르면 조사 소득이 높을수록 행정소득보다 낮게 응답하고, 소득이 낮을수록 행정소득보다 높게 응답하는 경향이 있음

- 근로소득의 경우 상용직을 중심으로 한 근로소득(일용소득을 제외)은 국세청에 신고한 소득이 더 정확하다고 가정
- 분석결과에 의하면 개인특성별로 ①과소·과대응답자 간 성향, ②과세 미달자 특성 ③본인 및 대리응답자 간 성향 차이가 있는 것으로 나타남
 - 과소 또는 과대 응답은 가구 또는 개인의 특성과 무관하게 랜덤하게 발생하는 경향이 있음. 따라서 두 자료의 일관성 있는 결합결과를 기대한다고 할 때 더 정확하다고 할 수 있는 과세자료로 조사자료를 대체하는 것이 더 바람직해 보임. 이를 통해 두 자료 결합에 따른 총오차를 줄일 수 있을 것으로 기대함
 - 한편, 개인에 따라 급여 또는 인건비를 현금대신 현물로 받을 수 있는데, 이러한 현상은 고소득보다 저소득 구간에 많이 분포하는 임시일용 및 무직·기타 종사자들을 중심으로 일어날 가능성이 높음
 - 즉, 과세 미달소득 구간(약 1,500만원)²¹⁾에서 조사소득이 더 높고, 이 구간에는 임시일용직이 차지하는 비율이 높다는 점을 고려할 때 이 구간에 대해서는 두 자료 중 큰 값을 선택하는 것이 더 타당할 것으로 판단



21) 1,500만원 기준선정은 분석에서 과세미달자 중 조사값이 행정값보다 큰 과세미달자 중 상용직 비율이 임시일용직 등의 비율보다 높아지는 구간

< 분석결과 요약 >

- 조사대상자 중 조사값이 행정값보다 큰 경우는 6,733명이며 이 중 71.2%가 상용직, 27.4%가 임시일용 또는 무직·기타에 해당. 2,000만원이상 구간에서는 상용직 비율이 높고, 1,500만원 미만에서는 임시일용직 등의 비율이 60%이상으로 상용직 비율보다 높음
- 1,500만원 미만에 속한 가구의 응답성향은 ① 본인이 응답하는 가구 비율이 다른 구간에 비해 높고, ② 본인응답 가구는 대리응답 가구에 비해 조사값과 행정값 간의 차이가 작으며, 조사값이 행정값보다 더 크게 측정되는 경향
- 이처럼 저소득구간에서 임시일용직 비율과 본인응답 비율이 높다는 점에서 볼 때, 저소득구간에 속한 사람들에서 신고되지 않거나 현금 등 본인만 아는 소득까지 정확하게 응답할 가능성이 높음
- 가계금융복지조사는 대리응답을 허용하며 이 경우 다른 가구원 소득에 대해 과소하게 응답되는 경향이 있으며, 한편 저소득층에 속한 임시일용직에서 본인응답 비율이 더 높게 나타남

* 전체 본인응답비율 37%, 과세미달자 중 본인응답비율 54%임

○ 분석결과를 바탕으로 한 근로소득 보완 기준 요약

- (행정자료만 있는 경우) 조사에서 무응답한 것으로 간주하고 행정자료로 조사자료 대체. 다만, 조사 사업소득이 0보다 크고 행정 사업소득이 없는 경우에는 조사 무응답을 대체하지 않음. 이는 조사에서 근로소득을 사업소득으로 잘못 기입했을 가능성을 고려한 것임
- (조사자료에만 있는 경우) 소득 미신고 대상이거나 신고제외 대상(또는 소득)으로 간주하여 조사자료를 사용하는 것을 원칙으로 함
- (둘 다 금액이 있는 경우) 행정자료로 조사자료를 대체하는 것을 원칙으로 함. 다만, 행정자료가 1,300만원 미만이고 조사자료가 1,500만원 이하(과세미달구간으로 미신고 대상으로 간주) 구간에서는 두 자료 중 큰 값을 사용한다. 이렇게 하는 이유는 행정자료로 조사자료를 완전대체하게 될 경우 저소득층에서의 미신고 대상소득이 누락됨으로써 전체적인 가구소득 분포가 왜곡될 가능성을 고려한 것임

○ 사업소득

① 행정자료와 조사자료 간 차이

- (개념차이) 두 자료 모두 순(net)소득을 측정하지만 행정 사업소득은 소득신고시 절세목적이 반영된다거나 현금 소득이 덜 신고될 수 있다는 특성이 있음
- (포괄범위 차이) 조사에서 임대소득은 재산소득으로 분류되나, 소득신고시 임대사업으로 등록된 임대소득은 사업소득으로 신고함
 - 또한 조사에서 원고료는 사업소득에 포함되고 인세는 재산소득에 포함되나, 행정자료에서는 이들 소득은 모두 기타소득으로 간주되며 사업소득과 재산소득으로 분리할 수 없음

행정자료(원천, 종합신고)		조사 사업소득	조사 재산소득
임대의 사업소득		사업소득	
임대 사업소득		-	임대소득
기타소득(75)	원고료, 인세 등	원고료는 사업소득	인세는 재산소득
(76)	강의료 등	사업소득	-

② 기타소득을 사업소득으로 반영하기가 어려움

- 기타소득은 양도소득, 사업소득, 일부 재산소득이 서로 다른 분류코드로 분류됨. 일부 코드의 기타소득은 조사에서 측정하는 것처럼 각각의 소득을 분리하기 어려움.²²⁾
- 기타소득은 종합신고와 원천징수로 신고되며 대부분은 원천징수로 신고됨. 기타소득의 경우 종합신고소득과 원천징수신고 소득 금액의 거의 동일하며, 다만 기타소득의 경우 수입 금액의 80~90%를 경비처리²³⁾하여 신고하기 때문에 실제 개인이 받는 소득금액에 비해 행정자료에 신고된 금액은 상대적으로 작을 수 있음

22) 기타소득 75코드는 원고료, 인세 등을 포함한다. 원고료는 조사에서 사업소득에 해당하지만, 인세는 재산소득에 해당하여 제대로 분리하여 반영하기 어려움. 또한 76코드는 강의료, 자문료 등을 포함함. 이들은 개념분류상 차이는 없으나 건수가 많지 않고, 금액은 크지 않지만 자료처리구조가 복잡하여 조사자료 대체에 사용하지 않는 것으로 판단

23) · 64코드: 필요경비 80% 적용[다만, 서화·골동품 보유기간이 10년 이상인 경우 90%를 적용하고, 실제 필요경비가 80(90)%를 초과하는 경우는 그 초과하는 금액도 포함하여 작성
· 71~76코드: 필요경비 80%를 적용(실제 든 필요경비가 80%를 초과하는 경우에는 그 초과하는 금액도 포함)하여 코드별로 구분하여 작성

③ 조사 사업소득 보안을 위한 기초분석

① 사업소득 유무별

- 2015년 자료에서 행정자료와 조사자료에 모두 사업소득이 있는 가구원수는 2,651명이며, 조사 사업소득이 있는 사람 중 행정 사업소득이 있는 경우는 약 45%정도임
- 전체 가구원 중 조사 사업소득이 없는데 행정 사업소득이 있는 경우는 7.7%이며, 조사 사업소득이 있는데 행정사업소득이 없는 경우는 약 6.1%로 나타남.

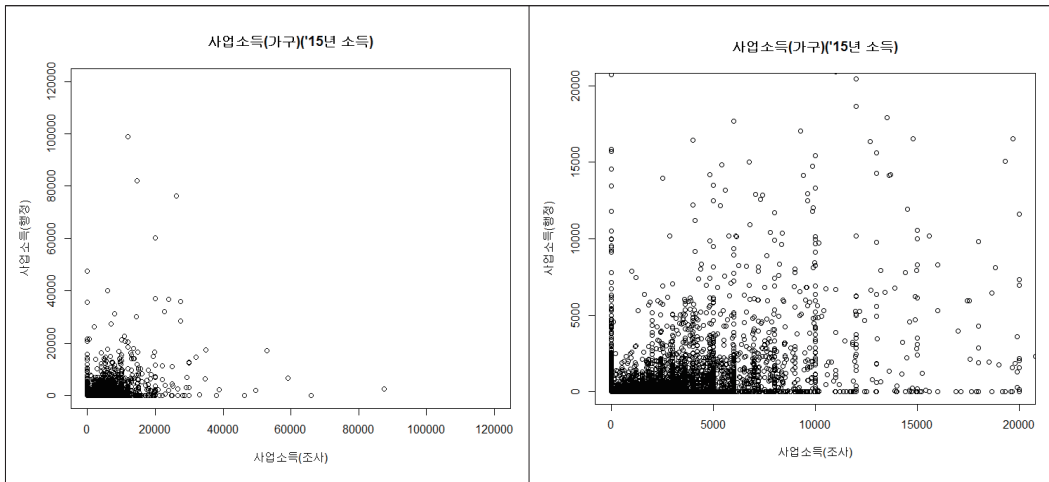
< 전체 가구원의 사업소득 존재 여부 >

(단위: 명, %)

소득 연도	가금복 사업소득	가구원수				비중			
		계	입수명부 포함		입수명부 미포함	계	입수명부 포함		입수명부 미포함
			국세청 사업소득 있음*	없음			국세청 사업소득 있음*	없음	
'15년	계	50,872	6,556	42,700	1,616	100.0	12.9	83.9	3.2
	있음	5,905	2,651	3,126	128	100.0	44.9	52.9	2.2
	없음	44,967	3,905	39,574	1,488	100.0	8.7	88.0	3.3

* 행정자료 사업소득의 '있음'은 매출은 있으나 소득이 0인 경우가 포함된 수치임

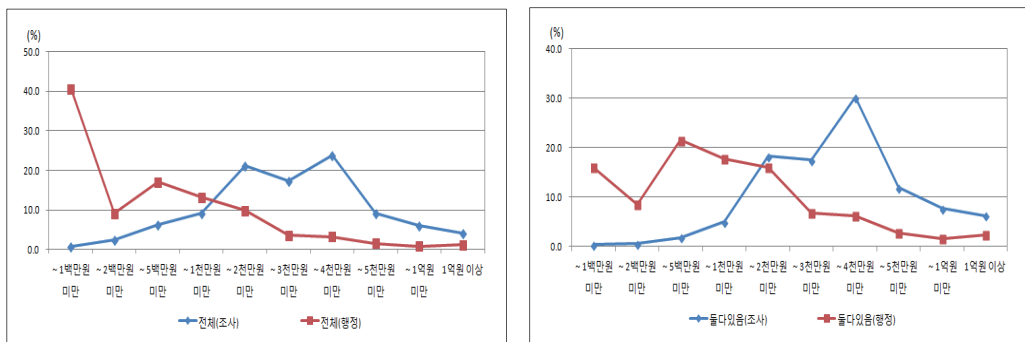
- 아래 산점도에서 가로축은 조사 사업소득, 세로축은 행정 사업소득을 의미함. 사업소득은 각 자료에서 꽤 큰 값들이 있으며, 한쪽 자료에만(축) 있는 경우도 많음



② 각 자료별 사업소득의 가구원 분포 비교

- 행정자료의 소득구간별 가구원은 1백만원 미만 구간에 집중적으로 분포하고, 조사자료는 2천만원~5천만원 구간에 가구원이 집중 분포함
- 조사와 행정자료에 모두 사업소득이 있는 경우는 조사소득 2천만원 이상 구간에서의 가구원 비중이 행정소득의 동 구간에서보다 월등히 높게 나타남.

< 사업소득 구간별 분포: 전체(가구원), 둘다 있음(가구원) >



○ 개인특성별 소득분포

- (연령별) 둘 다 있는 경우 70세 이상을 제외한 모든 연령구간에서 조사소득 구간의 가구원 비중이 행정소득 구간에서보다 높고, 조사값만 있는 경우는 50, 60, 70세 이상에서 비중이 높게 나타남

(단위: %)

	전체	입수명부 포함					
		계	둘다 있음		조사만 있음	행정만 있음	둘다 없음
			조사<행정	조사>=행정			
계	100.0	100.0	100.0	100.0	100.0	100.0	100.0
20세 미만	21.2	0.0	0.0	0.0	0.0	1.1	26.3
20~29	12.2	1.9	1.6	1.9	1.6	15.4	13.2
30~39	13.7	16.9	15.7	17.0	7.8	21.6	13.0
40~49	15.6	34.3	35.4	34.2	19.1	25.3	13.3
50~59	15.2	31.3	31.9	31.3	28.6	19.8	12.8
60~69	10.9	12.7	10.6	12.9	24.9	11.2	9.8
70세 이상	11.2	2.9	4.7	2.7	18.0	5.5	11.7

- (학력별) 행정값만 있거나 행정값이 더 큰 경우에는 대졸이상에 속한 가구원 비율이 높고, 조사값만 있는 경우에는 중졸이하에서 가구원 비율이 높음

(단위: %)

가구원 학력수준	전체	입수명부 포함					
		둘다 있음			조사만 있음	행정만 있음	둘다 없음
		계	조사<행정	조사≥행정			
계	100.0	100.0	100.0	100.0	100.0	100.0	100.0
초졸이하	28.0	4.8	3.1	5.0	23.9	5.8	31.8
중졸	11.4	7.4	6.7	7.4	15.2	7.7	11.8
고졸	30.6	43.9	39.8	44.3	36.5	39.7	28.6
대졸이상	30.0	43.9	50.4	43.3	24.3	46.9	27.8

- (종사상지위별) 행정값이 큰 경우는 고용원이 있는 자영업자 비중이 56.3%로 가장 높고, 조사값이 큰 경우는 고용원이 없는 자영업자 비중이 58.8%로 가장 높음. 조사값만 있는 경우는 자영업자 비중이 75.1%로 가장 높고, 행정값만 있는 경우는 기타·무직/상용/임시일용 비중이 41.1%로 가장 높게 나타남

(단위: %)

가구원 종사상지위	전체	입수명부 포함					
		둘다 있음			조사만 있음	행정만 있음	둘다 없음
		계	조사<행정	조사≥행정			
계	100.0	100.0	100.0	100.0	100.0	100.0	100.0
상용	24.3	1.7	1.6	1.7	2.0	32.4	26.9
임시일용	9.0	0.5	0.4	0.5	1.6	17.9	9.3
자영업 (고용원 有)	2.4	26.7	56.3	23.6	14.1	0.5	0.1
자영업 (고용원 無)	8.0	56.1	30.7	58.8	75.1	1.3	0.2
무급	2.8	0.4	0.8	0.4	0.7	6.4	2.8
기타종사자	1.0	12.9	7.5	13.4	3.5	0.4	0.0
기타(무직 등)	38.1	1.7	2.8	1.6	3.1	41.1	42.8
15세 미만	14.4	0.0	0.0	0.0	0.0	0.0	17.8

- (가구주와의 관계) 조사값만 있는 경우는 가구주인 경우가 82%, 행정값만 있는 경우는 배우자인 경우가 43%로 각각 높게 나타남

(단위: %)

가구주와의 관계	전체	입수명부 포함					
		둘다 있음			조사만 있음	행정만 있음	둘다 없음
		계	조사<행정	조사≥행정			
계	100.0	100.0	100.0	100.0	100.0	100.0	100.0
가구주	35.9	72.6	73.6	72.5	82.4	32.9	30.4
배우자	24.5	22.3	20.9	22.4	13.3	43.4	23.6
미혼자녀	33.6	3.6	5.1	3.4	2.7	19.3	39.5
기혼자녀	0.9	0.8	0.4	0.8	0.6	1.6	0.8
손자녀	1.0	0.0	0.0	0.0	0.0	0.3	1.2
부모	3.2	0.3	0.0	0.3	0.7	1.5	3.7
조부모	0.1	0.0	0.0	0.0	0.0	0.0	0.1
미혼형제자매	0.5	0.4	0.0	0.4	0.2	0.9	0.5
기타	0.3	0.2	0.0	0.2	0.1	0.2	0.3

- (가구주 산업분류) 조사만 있는 경우는 가구주가 농림어업에 종사하는 비율이 34%로 높고, 행정값이 큰 경우는 도소매업 종사비율이 높으며, 행정값만 있는 경우는 산업별로 골고루 분포하는 것으로 나타남

(단위: %)

가구주 산업분류	전체	입수명부 포함					
		둘다 있음			조사만 있음	행정만 있음	둘다 없음
		계	조사<행정	조사≥행정			
계	100.0	100.0	100.0	100.0	100.0	100.0	100.0
1 농림어업	2.5	2.1	2.4	2.0	34.1	0.1	0.4
2 광업,제조업	5.3	6.4	13.8	5.6	4.8	4.2	5.4
3 건설업	2.7	6.2	7.1	6.1	5.2	2.6	2.3
4 도소매업	3.1	16.4	19.7	16.0	9.8	3.0	1.7
5 운수업	1.9	9.3	2.0	10.1	7.0	1.6	1.1
6 숙박및음식점업	1.6	7.2	3.5	7.6	6.3	1.4	0.9
7 금융및보험업	1.5	4.6	2.4	4.8	0.9	2.3	1.3
8 부동산및임대업	0.7	3.5	0.8	3.8	1.5	1.4	0.4
9 전문,과학및기술서비스업/ 사업 시설관리및사업지원서비스업	2.3	3.8	7.9	3.3	1.6	2.9	2.2
10 공공행정,국방및사회보장행정/ 교육서비스업	2.9	3.7	1.2	3.9	2.2	2.9	2.9
11 보건및사회복지서비스업/ 예술,스포츠및여가관련서비스업	1.4	3.5	8.7	3.0	1.6	1.9	1.2
12 협회및단체,수리 및 기타개인서비스업	1.3	4.4	2.4	4.6	5.0	1.4	0.8
13 전기,가스,증기및수도사업 등	0.4	0.3	0.0	0.3	0.5	0.3	0.5
99 가구주(무응답), 비가구주	72.4	28.6	28.3	28.7	19.4	74.2	79.1

< 분석결과 요약 >

- 조사소득만 있는 경우는 미등록 사업자, 식량, 채소등 작물재배업 등이 주로 해당될 수 있으며, 행정소득만 있거나 행정값이 큰 경우는 조사누락 또는 과소응답 가능성이 있다.
- 고용원 있는 자영업자, 상용/임시일용/기타무직자에서 행정값이 크거나 행정값만 있는 비율이 높은데, 이는 이들 그룹에서 과소 또는 무응답 가능성이 있다.
- 고용원 없는 자영업자에서 조사값이 더 큰 것은 이 그룹에서 미등록 사업자 또는 농어업 등에 종사하는 영세사업자가 다수 포함될 가능성이 높다.
 - 영세자영업자 가구주는 조사에서 포착될 가능성이 높고, 일부 가구원의 사업소득 또는 근로자이면서 자영업을 한 사람의 사업소득은 조사에서 미포착 가능성도 있다.
- 가구에서 명의를 달리하여 자영업한 경우, 조사와 행정자료에서 차이가 날 수 있음에 유의

③ 각 자료에서 사업소득 유무에 따른 비교

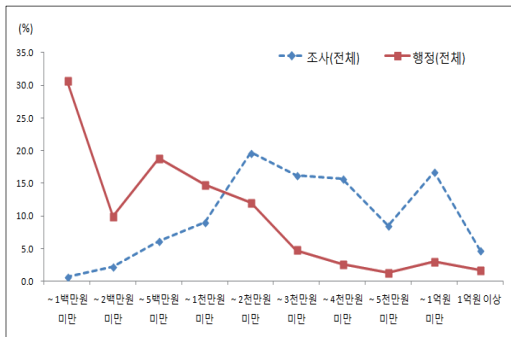
- 2016년 기준 조사 사업소득이 있는 가구원(5,676명) 중 50.4%가 행정 사업소득에 연계됨
 - 조사 사업소득만 있는 가구원은 2,816명, 행정 사업소득만 있는 가구원은 4,056명임
- 가구의 사업소득 구간별 분포를 보면, 1천만원 미만에서는 행정소득구간에서, 1천만원 이상에서는 조사소득구간에서 가구원 비중이 각각 높게 나타남.
 - (둘다 있는 경우) 1천만원 미만구간에서는 행정소득구간에서, 2천만원 이상 구간에서는 조사소득구간에서 가구원 비중이 높게 나타남
 - (행정자료만 있는 경우) 1천만원 미만 구간에 가구원의 약 85%가 포함됨. 행정자료만 있는 경우는 사업소득이 낮은 구간에 대부분의 가구원이 포함되며, 1억원 이상 구간에도 2.9%의 가구원이 포함되는 것으로 파악
 - (조사자료만 있는 경우) 2천만원 미만 구간에 약 53%의 가구원이 포함되며, 양쪽 모두 값이 있는 경우에 비해 저소득 구간에 속한 가구원 비중이 더 높게 나타남

④ 사업소득(임대)에 대한 조사자료 대체 기준

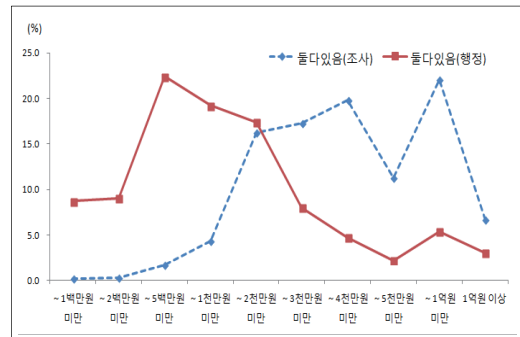
- 조사 사업소득은 없고 행정 사업소득만 있는 경우에는 조사에서 사업소득을 무응답한 것으로 간주하여 행정자료로 대체
 - 단, 응답자가 조사소득 항목 간의 오분류로 기입착오 가능성으로 인한 가구소득의 중복 집계 현상을 방지하기 위해, 조사 근로소득이 있으나 행정 근로소득이 없는 가구원은 조사에서 사업소득 항목이 무응답이라 하더라도 행정자료로 대체하지 않음
- 한편, 가구 내 조사 사업소득만 있는 가구원과 행정 사업소득만 있는 가구원이 동시에 있을 경우, 이 가구는 가구원 모두 사업소득 보완 대상에서 제외
 - 자영업을 운영하는 경우 실제 운영자와 사업자 등록자가 다를 수 있음, 이 경우 실제 운영자가 조사에서는 응답했으나, 행정자료는 사업등록자한테서 발생할 수 있으므로 가구 내 사업운영자와 등록자가 다를 경우 가구소득의 중복 가능성 방지

< 사업소득 구간별 가구원 분포(2016년) >

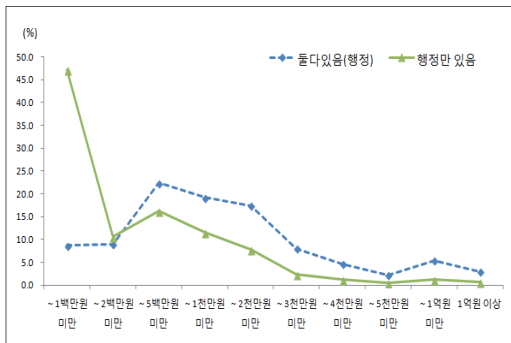
▶ 전체 가구원 대상



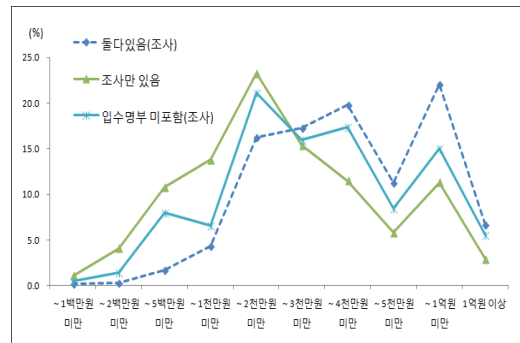
▶ 조사, 행정 둘다 있으며 0보다 큰 경우



▶ 행정자료의 가구원 분포



▶ 조사자료의 가구원 분포



○ 금융소득

① 행정 금융소득의 구성 및 개념

- 은행 등에서 받는 예·적금, 예탁금 등의 이자소득과 주식 및 출자금에서 발생하는 배당 소득으로 구성
 - 국세청 금융소득은 원천징수 금융소득과 종합신고 금융소득²⁴⁾으로 구성되며, 두 자료 간에는 중복가구(개인)가 발생할 수 있음

	비과세·분리과세 소득	원천징수된 일반과세 소득	원천징수 되지 않은 소득
원천분	A	B	-
종합분	-	B	C

- * 금융소득 = 비과세·분리과세소득+원천징수 일반과세 소득+원천징수 안되는 소득
- * 원천분 = 일반과세 소득 + 비과세·분리과세 소득 포함
- * 종합분 = 원천징수된 일반과세 소득 + 원천징수 되지 않은 소득(해외금융소득 등)
- * 소득세법, 조세특례제한법에 의해 비과세 또는 분리과세되는 이자소득은 제외 신고

○ 조사와 행정자료 상에서 이자와 배당소득의 개념은 대부분 일치

- 다만, 개인 간 거래(타인에게 빌려준 돈)에 의한 이자는 조사자료에서만 포착됨
- 또한 행정자료의 간주배당(57코드), 법인세법에 따른 일부배당(54코드) 등은 개념상 조사포괄범위에 포함되지 않음

[가계 금융복지조사 이자소득(지침서)]

- 금융회사에 맡긴 예·적금 이자
- 국가나 지방자치단체, 법인이 발행한 채권·증권의 이자(세금전 이자) 등
- 회사나 개인에게 빌려준 돈에 대한 이자

24) · 금융소득 종합과세 제도는 비과세 및 분리과세되는 금융소득을 제외한 연간 이자소득과 배당소득에 대한 세전 합계금액이 2천만원을 초과하는 경우 다른 종합소득과 합산하여 종합과세함.
· 국외금융소득 등과 같이 원천징수되지 않은 금융소득은 2천만원 이하이라도 종합소득세 부과대상으로 종합신고 해야 함

[국세청 이자소득(소득세신고)]

- 국내에서 받는 예금(적금·부금·예탁금 등 포함)의 이자
- 국외에서 받는 예금의 이자
- 국가·지방자치단체, 내국법인이 발행한 채권·증권의 이자와 할인액
- 신용계·신용부금으로 인한 이익
- 외국법인, 외국법인 국내지점 등의 회사채의 이자와 할인액
- 환매조건부 매매차익, 채권대차거래에 따른 이자상당액
- 저축성보험의 보험차익
- 직장공제회 초과반환금
- 비영업대금의 이익
- 이자소득을 발생시키는 상품과 결합된 파생상품의 이익
- 외국법인의 이자소득으로 상기 이외에 대금의 이자 및 신탁의 이익

[가계금융복지조사 배당소득]

- 보험, 신탁, 주식 등에 대하여 받은 배당금 수익(주식배당금, 신탁배당금, 생명보험 이익배당금)
- ※ 주식배당이 아닌 매매하여 얻은 수입(매매차익)은 해당되지 않음

[국세청 배당소득]

- 내국법인 배당·분배금, 건설이자의 배당
- 법인으로 보는 단체로부터 받는 배당·분배금
- 의제배당
- 「법인세법」에 따라 배당으로 처분된 금액
- 집합투자기구로부터의 이익
- 외국법인 배당·분배금, 건설이자의 배당, 이와 유사한 성질의 배당
- 「국제조세조정에 관한 법률」 제17조에 따라 배당받은 것으로 간주된 금액
- 「국제조세조정에 관한 법률」 제9조 및 제14조에 따라 배당으로 처분된 금액
- 출자공동사업자의 손익분배비율에 해당하는 금액
- 주가연계증권, 기타 파생결합증권
- 주식대차거래에 따른 배당상당액
- 배당소득을 발생시키는 상품과 결합된 파생상품의 이익
- ※ 채권이나 주식 등 유가증권의 매매에 따른 매매차익은 금융소득이 아님

② 조사자료와 행정자료 비교 분석

① 두 자료 간 금융소득 유무별 가구수(비율)²⁵⁾

- (둘 다 있는 경우) 조사자료와 행정자료에 모두 금융소득이 있는 가구의 비율은 12.7%로 높지 않음. 이들 가구 중 행정 금융소득이 큰 가구의 비율은 77.4%이며, 전반적으로 행정 금융소득이 조사 금융소득보다 더 큼
- (행정 금융소득만 있는 경우) 행정 금융소득만 있는 가구의 비율은 72.2%로 높는데, 이는 조사에서 금융소득을 무응답했을 가능성이 높음
- (조사 금융소득만 있는 경우) 조사 금융소득만 가구의 비율은 0.2%로 매우 낮음. 일부 가구원에서 사인 간 거래에 의한 이자소득이 조사에서만 포착되었을 가능성이 높음

< 금융소득 유무별 가구 분포, 2016년 기준 >

(단위: %)

		계	입수명부		입수명부 미포함
			있음	없음	
조사자료 금융소득	계	100.0	84.9	13.8	1.2
	있음	13.0	12.7	0.2	0.1
	없음	87.0	72.2	13.6	1.2

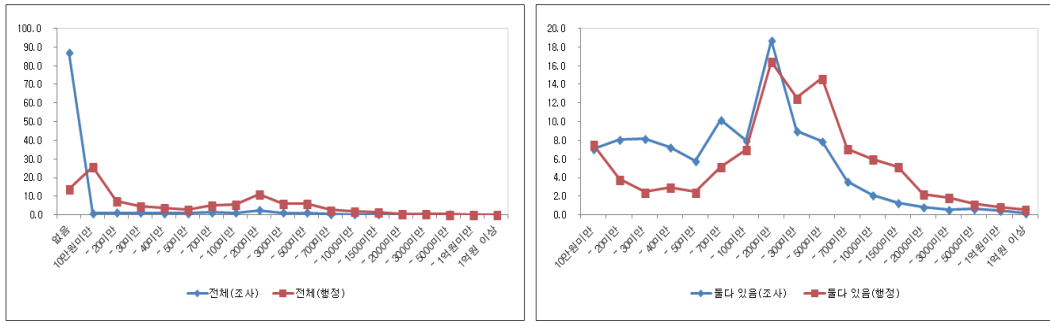
② 두 자료의 금융소득구간별 가구분포

- 위의 표의 경우 가구의 87%가 금융소득이 '0'원으로 조사에 응답했고, 가구의 14%가 5천원 미만(만원미만 자료처리시 '0'처리)의 행정 금융소득이 있는 것으로 나타남
 - 행정 금융소득기준으로 가구의 약 70%는 100만원 미만의 금융소득이 있음. 1천만원이상 금융소득이 있는 가구도 3.2%정도로 이는 조사자료(0.6%)의 5배 이상에 해당
- (둘 다 있는 경우) 100만원 미만 구간에서는 조사자료 구간의 가구 비율이 높지만, 100만원 이상 구간에서는 행정자료 구간의 가구 비율이 더 높음
 - 1000만원 이상 금융소득이 있는 가구는 조사자료에서 4.1%, 행정자료에서 12%로 행정 소득 기준 고액구간에 속한 가구 비중이 조사에 비해 더 높음
- (행정 금융소득만 있는 경우) 행정 금융소득만 있는 가구의 약 70%는 100만원 미만의 금액을 보유하고 있으며, 약 2.5%의 가구는 1천만원 이상을 보유함

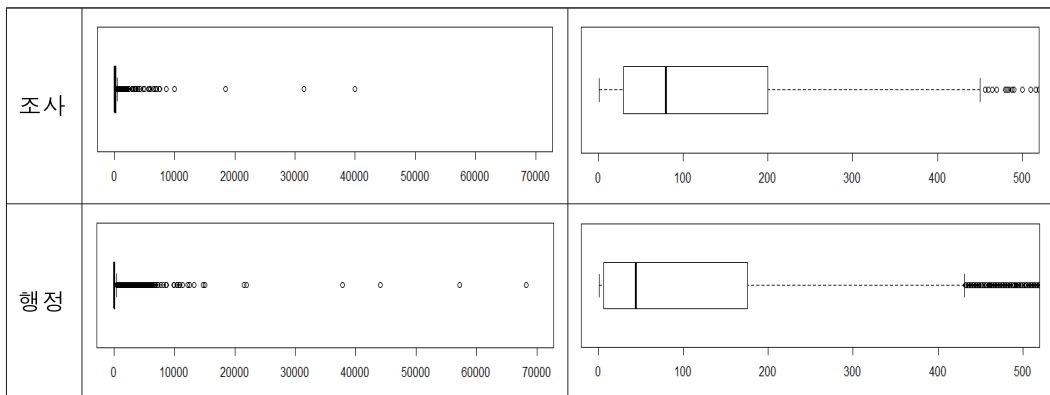
25) 정확한 금액비교를 위해 일부가구원만 연계된 가구(명부작성시)는 미포함된 가구원의 금융소득을 제외한 조사 소득과 행정소득을 비교·분석

- 행정금융소득만 있는 가구의 약 34%는 10만원 미만의 소액을 보유하고, 이는 소액의 금융소득에 대한 가구의 응답누락 가능성이 높음

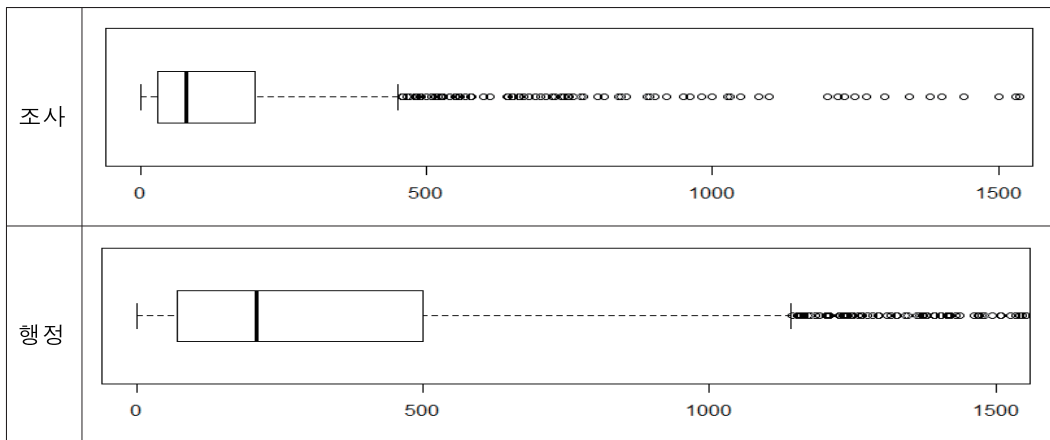
< 전체 또는 둘다 있는 가구의 조사, 행정 금융소득 구간별 가구분포(단위: 만원) >



< 금융소득 분포: 각 자료별 소득이 있는 가구 전체(단위: 만원) >



< 금융소득 분포: 조사와 행정자료 둘다 금융소득이 있는 가구(단위: 만원) >



③ 두 자료 간 차이 원인 분석

- 양쪽 모두 금액이 있는 가구의 약 80%는 행정소득이 조사소득보다 높고, 가구의 약 89.5%는 두 자료의 금융소득금액 간에 500만원 정도 차이 있음
 - (조사 금융소득이 큰 경우) 둘 다 있는 가구의 약 22%로 대부분은 500만원 이하의 금액 차이. 이 경우는 개인이 ‘빌려준 돈’에 대한 이자소득이 있는 가구도 포함되며, 조사금액이 더 큰 가구(521가구)중 약 12.7%는 ‘빌려준 돈’이 있는 가구에 해당

< 빌려준 돈 보유여부 비교 >

(단위: 가구, %)

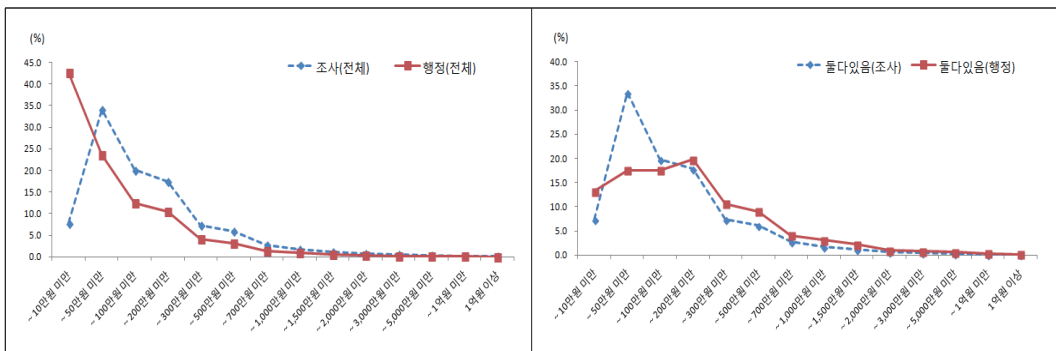
구분	가구수			비율		
	전체	빌려준 돈 있음	빌려준 돈 없음	전체	빌려준 돈 있음	빌려준 돈 없음
계	2,328	165	2,163	100.0	7.1	92.9
조사<행정	1,801	98	1,703	100.0	5.4	94.6
조사=행정	6	1	5	100.0	16.7	83.3
조사>행정	521	66	455	100.0	12.7	87.3
1천만원 이상 차이	33	14	19	100.0	42.4	57.6

- (행정 금융소득이 큰 경우) 조사에서 금융소득을 응답하지 않았거나, 금액을 축소해서 응답했거나, 또는 세후금액을 응답 가능성도 있음
 - 위의 표에서 행정자료가 더 큰 가구에도 ‘빌려준 돈’이 있는 98가구 정도 있으나 ‘빌려준 돈’에 대한 이자소득을 따로 조사하지 않으므로, 행정소득과 합산처리 곤란
 - 행정 금융소득만 있는 가구 중 20가구를 제외한 모든 가구는 금융자산 있는 것으로 볼 때 이 경우는 응답누락 가능성이 상당히 높음

< 금융소득 구간별 가구원 분포(2016년 기준) >

▶ 전체

▶ 조사, 행정자료 둘다 있음(> 0)



③ 금융소득 행정자료 활용 기준 요약

- 금융소득 행정자료는 원천분과 종합분이 있으며, 자료의 중복 활용을 피하기 위해 두 자료 중 최대금액을 사용함
- 금융소득은 조사에서 응답을 누락하거나 정확한 응답을 기대하기 어려운 항목 중 하나로서 행정자료로 대체하는 것이 더 정확할 것으로 보임
 - 다만, 조사값이 크거나 조사소득만 있는 가구 중 ‘빌려준 돈’이 있는 가구의 이자소득은 행정자료에서 포착할 수 없으므로 조사자료 활용

○ 소득세

① 종합소득세(단일소득-단순경비율 포함)와 원천세 자료 이용

- 소득세는 근로소득, 사업소득(원천, 연말), 공적연금, 금융소득에서 발생하는 세금 포함 (기타, 퇴직소득세는 제외). 소득세 중 종합소득세와 근로소득세가 큰 비중 차지함
- 개인에 따라 소득은 있지만 소득세는 없을 수 있는데, 이는 연말정산 후 소득세 공제 후 세금, 즉 결정세액을 측정하기 때문임

< 소득자료별 소득세 입수 현황 >

(단위: 명, 가구, %)

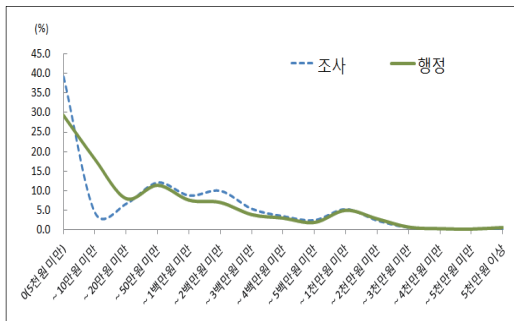
구분	소득 또는 수입 있음(a)	소득세 없음(b)	비율 (b/a)
종합소득세	4,172	2,284	54.7
단일소득 단순경비율 사업소득	846	535	63.2
원천세			
근로소득	14,062	6,898	49.1
사업소득(원천)	3,614	399	11.0
사업소득(연말정산)	458	313	68.3
연금소득(공적연금)	4,373	4,309	98.5
금융소득(가구)	18,020	486	2.7
기타소득	2,198	876	39.9
퇴직소득	1,533	416	27.1

② 조사자료와 행정자료 간 차이 발생 근거 및 활용기준

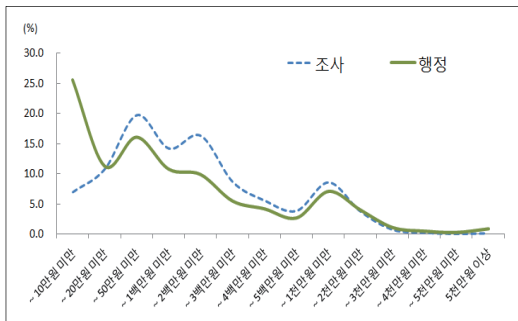
- 행정 소득세와 조사 소득세는 두 금액 간에 차이가 있으며, 이 차이는 응답착오 가능성이 높아 보임
 - 조사값이 행정값보다 더 큰 경우는, 행정자료는 연말정산 공제 후 자료를 사용하는 반면에 가구에서는 공제전의 원천징수 세액을 응답했거나 반올림 응답 가능성
 - 행정값이 더 큰 경우는, 가구가 반내림하여 응답하거나 일용근로소득세에 대한 무응답 등의 금액 착오 가능성
 - 행정값만 있는 경우는, 조사에서의 무응답 가능성이 가장 큼
- 소득세 구간별 가구분포(아래 그림)를 보면 소득세가 낮은 구간에서 두 자료 간 가구 분포가 차이가 있음

< 소득세 구간별 가구 분포(2016년 기준) >

▶ 전체가구의 구간별 분포

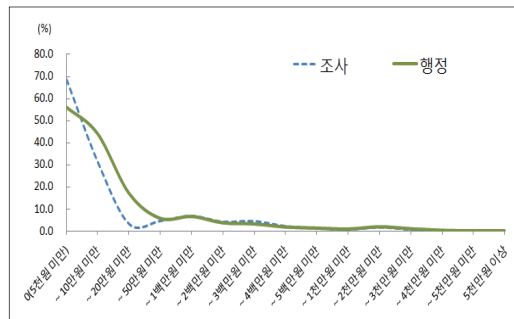


▶ 소득세가 0보다 큰 가구의 구간별 분포



< 소득세 구간별 가구원 분포(2016년 기준) >

▶ 전체가구원의 구간별 분포



▶ 소득세가 0보다 큰 가구원의 구간별 분포

