

NaSTaB T/R No.1

재정패널조사(NaSTaB)
실사과정 및 품질제고방안 소개

재정패널DB사업

한국조세재정연구원
KOREA INSTITUTE OF PUBLIC FINANCE

1. 재정패널조사

한국조세재정연구원은 조세정책과 행정을 연구하고 분석하는데 활용할 수 있는 실증적인 조사 자료를 수집하고자 재정패널조사를 기획하였다. 재정패널조사 기획 목적은 조세 정책과 복지 정책이 국가와 개별 경제주체인 가계에 미치는 영향을 분석할 기초자료를 수집하는 것이다. 나아가 정부의 조세정책과 재정정책에 대한 장기적인 효과 평가 및 정책 수립에 기여하고자하며, 종단면적 조세·재정관련 통계 인프라의 재편과 체계적인 관리가 가능하리라 기대한다. 마지막으로 학문적 논의를 위한 기초자료의 제공으로 연구 활성화 및 제도발전에 기여할 것으로 기대한다.

재정패널조사는 2008년 1차년도 조사를 시작하여 1년을 주기로 조사하고 있으며, 2014년 7월 현재 1차~6차 횡단면 및 종단면 데이터베이스가 구축되어 있고, 7차년도 현장 조사가 진행되고 있다. 재정패널조사의 모집단은 전국의 일반가구로서, 가구 단위의 조세-지출-복지에 관한 포괄적 데이터를 구축하고자 하였다. 조사 대상 가구의 규모는 최초조사시점에 구축된 원표본 가구(5,634가구)와 분가가구(약300여 가구)를 포함한 약 5,900여 가구 수준이며, 그 가구에 속하며 소득활동이 있는 가구원(약7,000여명)도 조사하고 있다. 이렇게 구축된 가구와 가구원들의 조사 성공률은 6차년도 기준 약 79%로, 국내·외 타패널 조사와 비교하였을 때 우수한 조사 성공률[1]을 유지하고 있다.

재정패널조사의 자료구축화가 안정적인 수준에 이르고, 연구하고자 하는 사용자들이 늘어남에 따라 재정패널조사팀에서는 본 고와 같은 테크니컬 리포트를 작성하여 연구자들에게 제공하고자 한다. 본 고는 테크니컬 리포트의 첫 회 발행본으로, 재정패널조사의 설계에 대한 전반적인 내용을 소개하고, 실사 진행과정 및 그 결과를 소개할 예정이다. 덧붙여, 재정패널조사의 데이터 품질을 향상시키기 위한 방안 및 클리닝 결과를 안내하여 연구자들의 자료에 대한 이해를 높이는 데에 기여하고자 한다.

1) 국내·외 타패널 조사성공률(6차년도 기준) : 한국노동패널 77% , 한국복지패널 75%, 미국PSID 79%, 영국 BHPS 77%

2. 조사 설계 및 실사

가. 조사 설계

1) 표본 설계[2)]

전국 15개 시도에 거주하고 있는 일반 가구와 그 가구에 거주하는 가구원을 모집단으로 정의하는 재정패널조사는 2005년도 인구주택총조사 자료를 표집틀(sampling frame)로 하고 있다. 2007년 예비조사를 수행 하여 모집단에서 추출할 적절한 표본 크기를 5,000가구로 설정하고, 2008년 6월부터 9월까지 본격적인 1차년도 조사를 실시하였다. 1차년도에 추출된 원표본 가구는 총 5,014개 가구로, 이들 가구를 대상으로 동일한 시기인 2009년 6월부터 2차년도 현장 실사가 진행되었다. 그러나 패널조사의 경우 이사·이민·전화번호 변경에 대한 미고지 등으로 추적 실패가 발생하거나, 계속 조사에 대한 부담감, 가구 내부 사정 등에 의한 응답거절 등으로 2차년도 조사시점에서 원표본 가구의 조사실패 비율이 높아지는 현상이 발생한다. 이 같은 상황에서 원표본 패널의 탈락을 최소화시키는 것이 1차적 과제이지만, 조사를 강력하게 거절하는 가구에게 계속 조사요구를 하는 것에는 무리가 있었다. 이에 따라 더 이상 조사가 불가능하다고 판단된 표본(618개)에 대해서는 추가 표본을 추출하여 대체조사를 진행하였다. 그러나 매년 추가표본을 추출한다면, 가중치 보정 과정이 매우 복잡해지고 종단분석에도 어려움이 있기 때문에 3차년도 조사부터는 추가표본을 추출하지 않고 기존에 1차년도(5,014개)와 2차년도에 구축된 표본가구(620개)를 원표본 가구로 정의하여 원가구 및 분가가구에 대한 추적조사만 진행하기로 하였다.

2) 설문지 설계

재정패널조사의 설문지는 가구용과 가구원용으로 설계되어 있다. 가구용 설문지는 가구의 전반적인 경제 상황에 대해 알고 있으며, 소득 및 지출을 주로 관리하고 있는 가구원 한 명이 응답해야 한다. 가구원용 설문지는 만 15세 이상이면서, 소득이 있거나 소득활동을 하는 가구원이 조사대상이 되어 본인이 직접 응답하도록 하였다. 다만 가구원용 설문지의 응답대상이 직접 응답할 수 없는 경우에 한하여 (해당 가구원의 소득 정보를 파악하고 있는 가구원)대리

2) 본 고에서는 표본 설계에 대한 개요를 소개하고자 하며, 2차년도 표본 추출에 대한 자세한 정보는 재정패널조사의 기초분석보고서를 참고하기 바란다.

응답이 가능하도록 하였다. 기본적인 설문지의 구조는 아래의 표와 같다.

<표 1> 설문지의 구조

가구용 설문지	가구원용 설문지
[표지] 가구 적격 여부 확인	[표지] 가구원 적격 여부 확인
가구 기본 현황	경제활동
주택, 종합부동산세, 자동차 보유 현황	소득 및 연금/보험 관련 지출 현황
가계 지출 현황	소득세 유형 및 소득 공제 현황
개인이전지출 및 개인이전소득	조세제도에 대한 인식
복지현황	
자산 및 부채 현황	
근로장려금	
연말정산환급금 및 종합소득세환급금	
유가보조금	

3) 소득공제 내역 증빙서류 수집 설계

앞서 살펴본 것과 같이 재정패널조사의 가구원 설문지는 개인의 연간 소득, 연금·보험 지출 및 소득, 소득공제 내역, (소득)결정세액 등의 항목을 회상에 의해 응답하도록 되어 있다. 이런 항목들의 조사시점은 '작년'인데, 소득공제 내역의 경우 각 항목별로 복잡한 기준과 계산법이 적용되는 경우가 많아 정확한 금액을 회상하여 응답하기 어렵다는 문제점이 있다. 이에 따라 재정패널조사에서는 개인의 소득공제 내역을 증빙할 수 있는 서류를 함께 수집하여 조사 내용의 정확성을 높이고자 하였다.

수집 대상 서류는 근로소득자 서류와 종합소득세 신고자 서류로 나뉜다. 모든 서류는 현장조사시점 당시 작년에 귀속되는 서류를 기준으로 하며, 원천징수영수증(근로소득자)과 과세표준확정신고·자진납부계산서·소득공제명세서(종합소득 신고자)를 수집하고 있다.

나. 실사 진행 과정

1) 실사 준비

(1) 설문 문항 개선

재정패널 조사의 설문 문항은 기본적으로 작년에 실시한 설문 구조를 골격으로 한다. 여기에 한국조세재정연구원 원내 연구진에서 당해 연도에 수정 및 보완할 사항을 검토하여 수정된 설문지를 조사업체에 전달하고, 조사업체에서는 조사 진행 여건을 고려하여 수정할 부분을 제안한다. 그런 다음 여러 차례의 토론을 거쳐 1차 설문 문항을 확정한 후, 관련 학계의 자문교수로부터 감수를 거쳐 2차 설문 문항으로 완성한다. 이렇게 작성된 설문지는 통계청의 문항변경 승인을 받으면 조사에 사용될 최종 설문지로 확정된다.

원내 연구진과 조사업체, 자문 교수 등 3단계에 걸쳐 진행되는 설문 문항의 개선은 전년도 조사결과를 바탕으로 진행된다. 예년에 비해 응답자의 조사 친숙도가 다소 상승했을 것으로 예상되는 문항에서는 간략화 되었던 질문을 보다 구체화하고, 응답자나 면접원의 수정 건의사항을 받아들여 보다 효율적인 조사가 이루어지도록 하는 데에 주안점을 두고 진행하고 있다. 또한, 재정패널조사에서는 특정 시점에서의 사회현상을 측정하고자 매년 신규로 개발하는 설문영역을 두고 있는데, 이 조사파트를 부가조사로 명명하고 있다. 부가조사란 종단 분석이 가능한 문항들과 함께 조사가 이루어지지만 한 해에만 1회에 걸쳐 조사가 이루어지며, 조사차수마다 그 주제가 변경되는 조사로 4차년도에는 출산장려정책 이용 현황을 조사하였고, 5차년도에는 타가구 거주하는 가구주의 부모 및 가구주 배우자 부모의 정보, 6차년도에는 가구원의 출산이력과 직업이력을 부가조사의 형태로 조사하였다.

(2) 조사원 선발 및 교육

재정패널조사에서 현장조사에 투입되는 조사인원은 약 90명 내외로 재정패널조사를 수행한 경험이 있는 면접원 중 성과가 우수한 면접원을 우선적으로 선발하였다. 면접원 투입이 완료된 후에는 전체 면접원을 대상으로 집체교육을 실시하였다. 1차년도의 경우 1박 2일 기간 동안 전체 면접원 및 조사업체 연구진, 한국조세재정연구원 연구진이 합숙하여 교육을 진행하였고, 2차년도 이후부터는 지역별 집중 교육을 실시하고 있다.

(3) 조사 안내 및 홍보

면접원 교육이 종료된 이후에는 본격 현장조사에 투입하기 전 인터넷 홈페이지, 사전 안내편지, 사전 전화 컨택을 통하여 조사 대상 가구에 현장 조사 상황을 안내하도록 하였다. 특히

사전 안내 편지에는 소득증빙서류 수집에 대해 안내하여 사전에 서류를 준비할 수 있는 시간적 여유를 주고자 하였다.

2) 실사 진행

(1) 조사 및 설문응답 과정

현장 실사는 가구에 방문한 면접원이 종이 설문지를 이용하여 설문 내용을 질문하고 그 응답을 면접원이 기록하는 면접원 타계식 조사방법을 사용하였다. 단, 응답자가 요청하거나, 관련 서류를 확인해 응답할 필요가 있는 경우에 한하여 응답자가 직접 설문에 기입하는 유치조사방법을 제한적으로 사용하였고, 그 비율은 6차년도 기준 약 0.6%(27/4,756)로 매년 1% 이내의 수준에서 조사되고 있다.

현장 실사시 면접원은 지난해까지 조사된 데이터를 통해 구축된 가구 기본 정보 리스트(성별, 출생년도, 학력, 결혼 여부, 자동차 보유 현황, 주택 면적, 과년도 소득증빙서류 수집여부 등)를 가지고 방문한다. 기입으로 인한 오류를 방지하고, 인적사항 등 쉽게 변동되지 않는 변수 등에 대한 오류를 예방하고자 하는 목적이다. 그 밖에 설문응답 과정에서 통상적인 수준을 벗어난 응답이 나오거나 현장에서 바로 대처하기 어려운 상황이 발생했을 경우 ①지역 수퍼바이저, ②조사업체 연구진, ③한국조세재정연구원의 순으로 내용을 알리고 지침을 전달받아 처리할 수 있도록 하였다.

(2) 설문지 응답자 선정

가구 설문지는 가구의 경제적인 상황을 잘 알고 있는 가구주나 가구주의 배우자를 대상으로 진행하도록 하였으며, 부득이한 경우 그 외 가계의 의사결정에 관여할 수 있는 가구원이 응답하도록 하였다. 6차년도 기준, 가구주나 가구주의 배우자가 응답한 경우가 96.9%로 매년 95%이상의 수준을 유지하고 있다.

가구원 설문지는 소득이 있거나, 소득활동을 하는 자(4차년도 이후부터 무급가족종사자도 포함)를 대상으로 하며, 가구에 정부지원현금소득 중 국민기초생활보장급여만이 유일한 소득인 경우, 타가구 이전소득이 유일한 소득인 경우에는 가구원 설문지 응답대상이 없는 것으로 처리하였다. 응답자를 선정하는 원칙은 가구원 본인을 직접 만나 조사하는 것이지만, 부재시 월급을 관리하거나 감독하는 가구원 또는 배우자에 한하여 대리응답 할 수 있도록 하였다. 가구의 특성에 따라서는 대리응답자가 정확한 정보를 알고 있는 경우(배우자의 소득을 관리하는 가구원이나 타가구원 명의의 사업장을 관리하는 가구원 등)도 있기 때문이며, 매년 현장 실사 중반쯤 이루어지는 조사원 간담회 과정에서 확인된 바로는 설문지 상에는 대리 응답으로 기

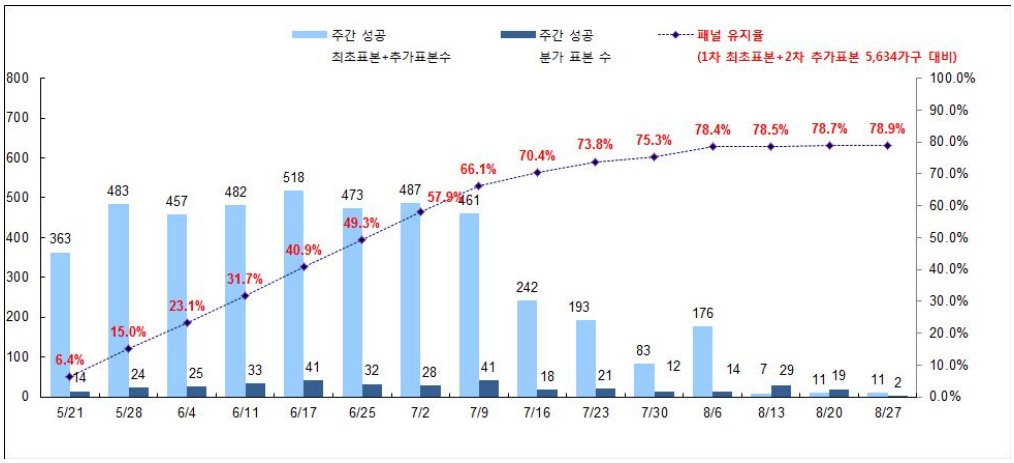
록되어 있더라도 실제 조사 과정에서는 소득자 본인의 동석 하에 조사가 진행되거나 애매한 질문을 전화로 본인에게 확인하는 경우도 상당히 많은 것으로 나타났다. 그러므로 타 가구원이 대신 응답했다고 표기됐다고 해서 조사 품질에 많은 문제가 있다고 보기는 어렵다. 6차년도 기준, 가구원 설문지에 소득자 본인이 직접 응답한 경우는 55.5%이었으며, 매년 55~60%수준에서 소득자 본인이 응답하고 있다.

다. 실사 진행 결과

1) 주별 진행 결과

재정패널 조사의 실사는 통상적으로 5월부터 8월까지 진행된다. 전반적인 실사 진행 상황을 보면, 실사에 돌입하고 7월 첫째 주까지 꾸준하게 조사 성공률이 상승하며 실사 진행의 막바지 시기이자 휴가철인 7월에 들어서면서 성공 가구수가 조금씩 줄어드는 양상을 보인다. 이 시기에는 단순거절 가구에 대한 재설득이 주로 이루어지고, 휴가 등 표본가구의 사정으로 조사 시점이 뒤로 밀린 가구에 대한 조사가 주로 진행되기 때문이다. 2013년도에 진행된 6차년도 주간 진행 상황은 <그림1>과 같다.

<그림 1> 재정패널조사 주간 진행상황-6차년도(2013년) 기준



2) 최종 조사 결과

6차년도를 기준으로 하여 재정패널조사의 원표본 유지율은 79%이다. 유효표본 유지율이란 원표본 가구 중 사망 및 해외이민의 사유로 가구가 소멸하여 더 이상 조사를 진행 할 수 없는 가구들을 분모에서 제외한 후 산출한 조사 성공률이며, 6차년도를 기준으로 79.9%이다. 유

효표본에 포함되지 않는 가구는 2개년도 이상 연속하여 '가구원 모두 해외거주로 파악된 가구(해외거주)', '가구원 모두 사망하여 2년 연속 사망으로 확인된 가구(사망)'이다.

<표 2> 원표본 가구 유지율

조사차수	추적해야 하는 원표본 가구		조사 성공	유지율	
	소멸 포함 가구수(A)	소멸 제외 가구수(B)		원표본 유지율A(%)	유효표본 유지율B(%)
6차년도	5,634	5,574	4,451	79.0%	79.9%
5차년도		5,569	4,500	79.9%	80.8%
4차년도		5,585	4,615	81.9%	82.6%
3차년도		5,600	4,747	84.3%	84.8%
2차년도		5,609	5,016	89.0%	89.4%

원표분 가구 이외에 원가구에 속해있던 가구원 중 여러 가지 사정(혼인, 학업 및 직장 등)으로 분가하여 나간 가구 역시 재정패널조사의 조사대상 가구로 인정되는데, 2차년도 23가구가 생성된 것을 시작으로 6차년도 기준 추적해야할 분가가구는 총 316가구로 확인되었다. 그 중 10가구가 사망, 해외 거주 등으로 추적조사 대상에서 제외되었고, 총 276가구가 조사에 성공하였다. 분가가구 유지율은 87.3%, 90.2%로 나타났다.

<표 3> 분가가구 유지율

조사차수	추적해야 하는 분가가구		조사성공	분가가구 유지율	
	소멸 포함 가구수(A)	소멸 제외 가구수(B)		유지율A(%)	유지율B(%)
6차년도	316	306	276	87.3	90.2
5차년도	244	237	212	86.9	89.5
4차년도	150	147	134	89.3	91.2
3차년도	80	79	67	83.8	84.8
2차년도	23	23	23	100.0	100.0

3) 소득공제 증빙서류 수집결과

소득공제 증빙서류 수집결과는 소득신고 대상이 있는 가구를 바탕으로 산출하였으며, <표 4>의 표와 같다. 1차년도의 경우 조사 첫해로 면접원과 응답자간의 신뢰도 형성 비율이 낮아 비교적 증빙서류 수집건수가 적었다가 2차년도 이후 점점 수집건수가 증가하였다. 다만 근로소득 연말정산 가구원의 수집건수에 비해 종합소득신고 가구원의 수집건수는 크게 못 미치는 것으로 나타났다. 6차년도를 기준으로 소득신고를 한 가구원이 있는 가구(2,864가구) 중 증빙

서류를 제출한 가구는 1,785가구로 수집률은 73.3%이며, 1가구에 2명이상 제출한 경우가 있어 근로소득 연말정산 제출 가구원은 1646명, 종합소득세 신고서류 제출 가구원은 360명으로 제출 가구수 보다 많다. 근로소득 연말정산 서류와 종합소득세 신고서류를 동시에 제출한 가구원도 있으며, 매년 2~4건 정도 발생하고 있다.

<표 4> 증빙서류 수집 현황

구분	1차년도(2007년 귀속)			2차년도(2008년 귀속)		
	소득신고 가구원이 있는 가구	소득신고 가구원		소득신고 가구원이 있는 가구	소득신고 가구원	
		근로소득 연말정산	종합소득 신고		근로소득 연말정산	종합소득 신고
해당 case	2,612건	2,386건	721건	2,801건	2,577건	806건
수집건수	681건	638건	76건	1,327건	1,217건	213건
수집률	26.1%	26.7%	10.5%	47.4%	47.2%	26.4%
구분	3차년도(2009년 귀속)			4차년도(2010년 귀속)		
	소득신고 가구원이 있는 가구	소득신고 가구원		소득신고 가구원이 있는 가구	소득신고 가구원	
		근로소득 연말정산	종합소득 신고		근로소득 연말정산	종합소득 신고
해당 case	2,772건	2,561건	876건	2,769건	2,547건	875건
수집건수	1,491건	1,344건	249건	1,661건	1,508건	289건
수집률	53.8%	52.5%	28.4%	60.0%	59.2%	33.0%
구분	5차년도(2011년 귀속)			6차년도(2012년 귀속)		
	소득신고 가구원이 있는 가구	소득신고 가구원		소득신고 가구원이 있는 가구	소득신고 가구원	
		근로소득 연말정산	종합소득 신고		근로소득 연말정산	종합소득 신고
해당 case	2,754건	2,643건	851건	2,864건	2,245건	850건
수집건수	1,695건	1,531건	314건	1,785건	1,646건	360건
수집률	61.5%	57.9%	36.9%	62.3%	73.3%	42.4%

4) 탈락 및 소멸가구 현황

(1) 탈락 및 소멸가구 사유

추적조사에서 탈락한 가구는 각각 다음과 같은 사유로 구분 할 수 있다.

① 강력거절 : 표본가구가 조사 참여를 완강히 거절한 경우

② 추적실패(주소 불명) : 전화번호 결번 등으로 사전 컨택이 되지 않으며, 주소지를 방문했으나 이사로 추정되어 더 이상 추적이 불가능한 경우

③ 사망/해외거주/입원 등 : 가구 응답자가 각각의 사유로 조사 참여가 불가능한 경우

④ 기타 : 장기 부재, 약속 반복 위반 등

추적조사에서 더 이상 추적하지 않는 가구는 소멸가구라고 명명하는데, 2개년도 이상 연속하여 동일한 소멸사유로 확인된 경우 조사에서 제외하였으며, 전 가구원 모두 사망 또는 해외이민을 간 경우가 해당된다.

(2) 원표본 가구 중 탈락·소멸 가구

6차년도 조사 참여를 거부한 가구를 사유별로 보면, 강력거절로 인한 탈락이 633가구로 탈락 가구 중 56.4%를 차지해 가장 많았다. 추적실패는 438가구로 39.0%였다. 2년 연속 사망과 해외거주 등으로 소멸된 가구는 총 60가구로 사망가구가 55%, 해외거주 가구가 45%이었다. 6차년도를 기준으로 더 이상 조사되지 않는 소멸가구 리스트는 부록의 표를 참고하기 바란다.

<표 5> 원표본 가구 중 탈락 및 소멸가구 유형

탈락 가구							소멸가구		
강력 거절	추적 실패	병원 입원	사망	해외 거주	기타	소계	사망	해외 거주	소계
633	438	11	18	5	18	1,123	33	27	60
56.4%	39.0%	1.0%	1.6%	0.4%	1.6%	100.0%	55.0%	45.0%	100.0%

3. 실사 품질 제고 방안

패널 조사는 결과의 정확성과 논리성 확보하는 것은 물론 높은 패널 유지율 달성하고 응답의 일관성을 유지하는 것 또한 매우 중요하다. 재정패널 조사는 이러한 점을 고려하여 실사 전후로 다양한 업무를 수행하고 있다. 따라서 본 절에서는 실사 품질 제고를 위한 노력들을 구체적으로 살펴보고자 한다.

가. 사전 검증

1) 재정패널 조사 유경험 면접원 위주 투입

재정패널 조사의 경우, 평균적으로 약 90여명의 면접원이 조사에 투입되며 그 중 90%이상이 과년도 패널조사 경험이 있는 면접원으로 구성되어 있다. 이러한 유경험 면접원들은 조사내용에 대한 이해가 매우 뛰어나 좀 더 정확한 응답 값을 얻는데 도움을 줄 수 있다. 뿐만 아니라 매년 같은 패널가구를 접촉하기 때문에 면접원과 응답자간 친밀감과 신뢰도가 높아 조사 정확성은 물론 패널 유지율을 높이는데 기여하고 있다.

<표 6> 연도별 면접원 투입 현황

조사차수	투입 면접원	재정패널 유경험자	신규 충원
1차년도	120	-	-
2차년도	99	96	3
3차년도	90	85	5
4차년도	88	80	8
5차년도	86	81	5
6차년도	84	81	3

2) 패널 유지율 제고를 위한 노력

패널조사의 경우 일반적인 횡단조사와 달리 종단면분석이 가능하기 때문에 높은 패널 유지율을 달성하는 것이 매우 핵심적인 부분이다. 따라서 재정패널 조사에서는 안정적이고 높은 패널 유지율을 달성하기 위하여 응답사례금을 지급하는 것 외에도 아래와 같은 여러 가지 패널관리 프로그램을 진행하고 있다.

<표 7> 패널 유지율 제고를 위한 방안

구분	패널 관리 프로그램
1월 ~ 3월	패널 대상 경품행사 진행 , 설날 연하장 발송, 조사 사전 안내 전화
4월 ~ 6월	면접원 대상 패널 유지율 중요성 교육 실시, 조사 안내 편지 발송, 조사대상 패널 사전 전화 접촉
5월 ~ 9월 (현장실사기간)	강력거절 가구 리스트 파악 및 별도의 컨택 요원 투입, 응답내용 확인 콜백 중 조사참여 감사 인사 및 주소확인 병행, 추석 연하장 발송
10월 ~ 12월	기념품 발송 및 응답자 명의 사회단체 기부 행사
상시	이사가구 변경 주소지 신고시 사례품 지급, 생일축하 이벤트

나. 사후 검증

1) 가중치 산출 및 보정

재정패널 조사는 데이터 추정 시 특성치인 모수에 대한 비편향 추정량(unbiased estimator)을 얻을 수 있도록 가중치를 제공한다[3]. 또한 패널조사에서는 웨이브가 진행되면서 조사 자료가 누적되면 연도별 자료를 이용한 횡단면 분석을 위한 가중치과 함께 동적 변화를 파악하기 위한 종단면 분석을 위한 가중치를 제공하여 패널자료를 활용해 대표성 있는 통계분석 결과를 얻을 수 있도록 하고 있다.

3) 재정패널 가중치의 경우 일반적인 복합표본조사와 마찬가지로 첫째, 설계가중치, 둘째, 무응답 조정, 셋째, 모집단 정보를 이용한 사후층화 또는 레이팅에 의한 조정 등 3단계에 걸쳐 산출 된다.

2) 주요변수 비교

재정패널조사는 자료의 신뢰성을 검증하는 한 방법으로 통계청 및 타기관 조사와 주요 변수 분석 결과를 비교 분석한다. 대표적으로 통계청 가계동향조사[4] 결과와 비교 분석이 가능하며 월평균 가구소득을 비교하면 <표 10> 와 같다. 단 가계동향조사의 경우, 농어가를 제외한 2인 이상 가구를 대상으로 하기 때문에 재정패널 월평균 가구소득과 단순비교 할 수 없음을 유의해야한다.

<표 7> 재정패널과 통계청 가계동향조사 월평균 가구소득 비교(단위: 원)

조사차수	통계청 가계동향조사 (농어가 제외 2인 이상 가구 대상)	재정패널조사	
		시지역, 2인 이상 가구 대상	전체 가구 대상
1차년도	3,062,027	3,097,235	2,709,705
2차년도	3,254,087	3,110,554	2,708,488
3차년도	3,293,928	3,269,125	2,859,211
4차년도	3,491,149	3,595,806	3,127,072
5차년도	3,705,284	3,704,446	3,172,120
6차년도	3,934,716	3,831,449	3,239,932

3) 데이터 클리닝(Data Cleaning)

재정패널 조사는 조사과정에서 발생하는 다양한 오류를 최소하기 위하여 지속적으로 데이터를 수정 또는 보완하는 데이터 클리닝(Data Cleaning) 작업을 진행하고 있다. 데이터 클리닝은 여러 단계를 거쳐 진행하게 된다. 그 첫 번째 단계로 입력 오류(inpuet error)를 방지하기 위한 입력시스템 구축이 있다. 재정패널 조사는 종이설문지를 이용하기 때문에 조사된 값을 자료화 하는 과정에서 입력자의 실수 또는 표기의 부정확 등으로 오류가 발생할 수 있다. 따라서 이런 입력 오류를 방지하기 위해서 자료 입력과 동시에 로직을 자동 검증하는 프로그램을 적용하여 설문의 논리적 오류, 응답 이상치를 입력단계에서 차단 할 수 있도록 관리하고 있다. 두 번째 단계로 논리성 및 각 문항별 응답값의 일관성을 점검하는 횡단 클리닝과 이전 조사 응답값과 일관성이 있는지 확인하는 종단 클리닝을 진행한다. 재정패널조사는 응답자의 회상에 의존하는 조사이기 때문에 응답자의 왜곡 및 종단적 응답 오류가 나타날 가능성이 있

4) 관련 통계자료는 통계청 홈페이지 KOSIS DB 가계동향조사에서 확인 가능하다.

다. 따라서 조사의 앞뒤 문항 및 전년도 자료를 활용하여 비교검증을 실시하여 자료의 오류를 점검하고, 수정 및 보완 하고 있다. 다음 <표 8>은 조사 항목별 주요 클리닝 사항을 나타낸 것이다.

<표 9> 조사 항목별 데이터 클리닝 사항(가구 조사)

구분		클리닝 내역
가구	가구 기본 현황	가구원 신규 또는 분가 변동 여부 횡단 및 종단 비교
		가구원 출생년도, 출생월, 성별, 학력 종단 비교
		직업, 종사상지위, 공적연금 응답값 로직 확인
		취업여부와 구직사유, 미구직 사유 응답 로직 확인
	가계지출	각 지출 항목별 지출 금액 응답값 이상치 확인
		지출 대상 가구원 인적사항과 가구기본현황 가구원 인적사항 비교
	개인이전지출 및 개인이전소득	이전지출 및 이전 소득 금액 이상치 확인
		이전지출 대상 가구원 인적사항과 가구기본현황 가구원 인적사항 비교
	복지현황	서비스 대상자 인적사항과 가구기본현황 가구원 인적사항 비교
		서비스 대상자 조건과 인적사항 비교 확인
		복지서비스 받는 횟수 이상치 확인
	자산 및 부채	자산 및 부채 금액 이상치 확인
		부동산 관련 자산 및 부채 응답과 주택 현황 응답 비교
	출산이력	출산 가구원 인적사항과 가구기본현황 가구원 인적사항 비교
	그 외	근로장려금 신청 자격과 대상과 인적사항 비교 확인
		연말정산 환급금 및 종합소득세 환금금액과 소득금액 비교
		출산 비용 이상치 클리닝
		출산 가구원 인적사항과 가구기본현황 가구원 인적사항 비교
		가구원 설문 조사 대상 확인
가구원	경제활동	가구기본현황 가구원 인적사항 중 근로상태 관련 응답과 종단 비교
	소득 및 연금·보험 관련 지출	연금·보험 소득 수급개월 수 및 금액 이상치 확인
		연금·보험 소득, 정부지원 현금소득이 있는 가구원 인적사항 확인
		근로소득, 사업소득 유무와 가구원 직업상태 비교

구분		클리닝 내역
		타가구 이전 소득 유무 가구 설문 타가구 이전 소득 유무 비교
		부동산 임대소득 유무와 가구 설문 주택현황 응답 비교
		공적연금 및 건강보험, 고용보험, 민간보험 납부급액 이상치 확인
		공적연금, 건강보험, 고용보험 납부 가구원 인적사항 확인
	소득공제현황	소득 공제 현황, 연금보험, 항목별 지출 여부, 신용카드 사용 여부 비교
		결정세액과 근로소득 비교
	직업이력	직업이력 응답과 가구설문 및 가구원 조사 중 가구원 인적사항 근로 관련 응답 횡단 및 종단 비교

재정패널 조사 자료의 정확성을 위하여 가급적 발생할 수 있는 모든 로직상의 오류에 대하여 점검하고 있으나, 간혹 오류의 원인을 파악하기 어렵거나 오류로 판단하기 어려운 특이사항들이 발생한다. 문항간의 로직을 따져보면 오류가 있는 것처럼 보이지만 실제 현실상황에서는 발생 가능한 상황들을 일컫는 것인데, 이런 경우 케이스별 상황을 입체적으로 파악하여 수정할지 그대로 남겨 둘지를 결정하게 된다. 또한, 과도한 데이터 수정은 자료의 편의를 발생시킬 수 있어 입력에러 또는 명백한 오류를 제외하고는 데이터에 그대로 남겨두는 것을 원칙으로 한다.

이런 특이사항들은 사용자의 입장에서는 오류로 판단 할 수 있으나 명백한 오류로 분류하기는 어렵기 때문에 이에 대한 처리는 사용자들의 연구목적 및 연구방법에 따라 개별적으로 처리하기를 권장한다. 재정패널 조사에서 나타난 특이사항은 다음 <표 10>의 내용 등이 있으며, 이에 대한 내용은 수시로 업데이트하여 홈페이지에 공지할 예정이다.

<표 10> 조사 항목별 특이사항

구분	특이사항	점검결과
교육수준	동일한 사람의 교육수준이 조사차수가 높아질수록 낮아진 경우	교육수준의 경우 응답자의 왜곡 가능성이 있는 문항으로 2014년 조사현재 재확인하고 있으나 정확한 사유를 파악하기 어려운 경우 응답값 그대로 처리하였음.
재산세	자가 주택 보유 가구원 중 재산세 미납부자	재산세 면세자 이거나 조사시점에 재산세를 고지받지 못한 경우도 있으므로 오류로 판단할 수 없음.
부채현황	주택 점유형태는 '기타'인데, 부채현황에서 돌려줘야할 전세금을 보유하고 있는 경우	타가구 명의의 주택에 거주하나, 본인 명의의 주택을 별도로 보유하고 있어 그 주택을 전세주고 있는 경우가 있어 오류로 보기 어려움
혼인상태	1~3차 미혼상태였는데 4차 조사이후 기혼(사별, 이혼)으로 되어 있는 경우	4차년도 이후 혼인상태의 선택문항[5]이 변경되었기 때문에 1~3차의 미혼이 사별 또는 이혼인지 판단하기 어려움.
근로월수	12월 말 기준으로 직업이 있는데 근로기간은 1개월 미만인 경우	취업상태는 작년 말(12월말)을 기준으로 하기 때문에 12월 중 취업을 하였더라도 취업상태는 직업이 있는 것으로 판단됨. 다만 12월 중 근로를 시작하였다면 근로기간은 1개월 미만이기 때문에 직업은 유지, 근로월수, 근로시간은 0으로 입력되어 있음. 다만 가구설문종료파트에서는 소득활동을 하지 않는 것으로 처리하고 있음.
국민연금	국민연금 수급연령인데 국민연금을 납부하지도 않고, 수급하고 있지도 않다고 응답한 경우 (수급을 유예하고 있다고 응답한 경우)	국민연금은 만60세 이상자 중 5년 동안 본인의 의사에 따라 임의가입이 가능함. 그러나 이 경우 수급을 유예한 것인지, 임의계속 가입상태인지 파악하기 어려워 연금 가입 중이되 미납부 상태인 것으로 처리
국민연금	국민연금 납부 연령이 초과했는데 설문지에 국민연금을 납부하고 있다고 응답한 경우	종단면으로 확인해보았으나, 마땅한 오류 상황을 발견하지 못해 그대로 '납부 중'으로 처리
민간연금	만 50대 미만의 응답자가 민간퇴직연금을 수령하였다고 응답한 경우	퇴직금을 지금 받을 때, 연금형태로 받는 경우도 있어 오류로 판단하기 어려움
직업이력	종단면적으로 직업이 없던 사람인데, 2013년 조사된 직업이력에는 직업이 있다고 응답한 경우(혹은 그 반대의 경우)	직업이력은 2013년도에 시행한 부가조사로, 10년간의 직업 상태에 대해 회상에 의해 응답한 조사파트임. 따라서 기존의 데이터와 별개의 데이터로 분석하는 것이 바람직 함.
출산이력	출산이력 상에는 출산한 가구원이 있으나, 당시 가구원 중에는 신생아가 없는 경우(혹은 그 반대의 경우)	출산이력은 2013년 조사당시 가구원으로 인정되는 여성의 출산 이력을 파악한 것으로 과거 가구원 중 신생아를 반드시 포함해야하는 것은 아님.

5) 4차년도부터 기혼이 기혼(배우자 있음)과 기혼(사별·이혼)으로 나누어 응답하도록 설문문항이 수정되었음.

4. 맺음말

본 고는 재정패널 조사 자료를 이용하고자 하는 연구자들에게 자료의 이해를 돕기 위한 목적으로 작성된 테크니컬 리포트이다. 올 해 처음으로 발행된 테크니컬 리포트에서는 재정패널 조사의 실사과정 및 품질 향상을 위한 클리닝 내용을 소개하였다. 특히 클리닝 과정에서 발생한 특이사항들과 가구 소멸로 더 이상 조사되지 않는 가구들의 HID(가구고유번호)를 소개하여 사용자들이 연구목적 및 연구방법에 맞추어 자료를 조작할 수 있는데 도움이 되고자 하였다.

현재 재정패널조사는 2014년 학술대회를 위한 6차년도 베타버전 데이터가 구축되어 있으며, 홈페이지(panel.kipf.re.kr) 무료회원가입을 통해 1차~5차 공식버전 데이터를 제공받을 수 있다. 재정패널 조사의 공식버전 데이터는 매년 3월~4월경 홈페이지를 통해 공개되며, 현재 조사 중인 7차년도 자료는 2015년 학술대회 베타버전은 내년 7월경 공개될 예정이다. 재정패널조사 데이터 사용시에는 본 고와 같은 테크니컬 리포트 뿐만 아니라 사용자들의 자료 사용법에 대해 자세하게 안내된 유저가이드(User's Guide) 및 통합설문지, 통합코드북을 참고하기 바라며, 앞으로도 재정패널 조사를 이용한 많은 연구가 생산되기를 기대한다.

<부록표> 소멸 가구(60가구) 리스트-사망, 해외거주

HID	2차년도	3차년도	4차년도	5차년도	6차년도
10113	성공	성공	성공	사망	사망
10294	성공	성공	성공	해외거주	해외거주
10626	기타	강력거절	강력거절	해외거주	해외거주
10679	해외거주	해외거주	해외거주	해외거주	해외거주
10721	해외거주	해외거주	해외거주	해외거주	해외거주
10745	성공	사망	사망	사망	사망
10907	해외거주	해외거주	해외거주	해외거주	해외거주
11015	성공	기타	기타	해외거주	해외거주
11221	병원입원	사망	사망	사망	사망
11222	성공	성공	병원입원	사망	사망
11224	사망	사망	사망	사망	사망
11249	성공	사망	사망	사망	사망
11310	성공	성공	해외거주	해외거주	해외거주
11351	추적실패	추적실패	추적실패	사망	사망
11356	해외거주	해외거주	해외거주	해외거주	해외거주
11497	성공	성공	성공	사망	사망
11530	성공	사망	사망	사망	사망
11558	성공	해외거주	해외거주	해외거주	해외거주
11823	해외거주	해외거주	해외거주	해외거주	해외거주
11939	사망	사망	사망	사망	사망
11950	성공	추적실패	사망	사망	사망
12001	사망	사망	사망	사망	사망
12144	성공	사망	사망	사망	사망
12324	성공	성공	성공	해외거주	해외거주
12426	성공	성공	해외거주	해외거주	해외거주
12472	성공	추적실패	해외거주	해외거주	해외거주
12557	해외거주	해외거주	해외거주	해외거주	해외거주
12647	강력거절	강력거절	강력거절	해외거주	해외거주
12691	성공	사망	사망	사망	사망
12872	사망	사망	사망	사망	사망
13080	성공	성공	기타	해외거주	해외거주
13209	성공	성공	성공	해외거주	해외거주
13351	성공	성공	성공	사망	사망
13539	사망	사망	사망	사망	사망
13579	강력거절	사망	사망	사망	사망
13613	성공	성공	해외거주	해외거주	해외거주
13627	성공	성공	해외거주	해외거주	해외거주
13834	성공	성공	사망	사망	사망
13913	성공	성공	사망	사망	사망
14074	성공	성공	추적실패	사망	사망
14083	성공	사망	사망	사망	사망
14164	사망	사망	사망	사망	사망
14224	성공	성공	성공	사망	사망
14235	성공	사망	사망	사망	사망
14279	사망	사망	사망	사망	사망
14391	성공	성공	성공	사망	사망
14462	성공	추적실패	사망	사망	사망

14511	성공	성공	사망	사망	사망
14572	성공	성공	성공	해외거주	해외거주
14622	성공	성공	성공	사망	사망
14697	성공	사망	사망	사망	사망
14710	성공	성공	성공	해외거주	해외거주
14711	성공	성공	성공	해외거주	해외거주
14991	사망	사망	사망	사망	사망
90589	성공	해외거주	해외거주	해외거주	해외거주
90674	성공	성공	성공	해외거주	해외거주
90758	성공	해외거주	해외거주	해외거주	해외거주
90939	성공	해외거주	해외거주	해외거주	해외거주
91024	성공	추적실패	해외거주	해외거주	해외거주
91214	성공	성공	성공	사망	사망