2008년 통계청 정책연구용역과제 최종보고

패널자료 생성 및 이용활성화 방안 연구
- 농가경제조사(2003~2007)를 중심으로

2008. 11. 18

사업수행기관 : 성신여대 산학협력단

목 차

Ⅰ. 서론 ····· 2

Ⅱ. 국내외 패널데이타 작성 현황과 패널데이타 분석방법 소개 ····· 4
1. 국내외 패널데이타 작성 현황 ····· 4
2. 기본적인 패널데이타 분석 방법 ····· 12

Ⅲ. 농가경제조사 패널화 과정과 결과 ····· 24
1. 농가경제조사 자료의 특성과 패널구축방법 ····· 24
2. 패널화 작업 결과 ····· 27
3. 표본유지 및 표본변동 분석 ····· 29
4. 종단면 가중치 작성 ····· 30

Ⅳ. 농가경제조사 패널데이타를 이용한 분석결과 ····· 39
1. 소득이동 결정요인 분석 ····· 39
2. 농업소득 및 농업순생산 결정요인 분석 ····· 46

Ⅴ. 한계점 및 추후 연구 방향 ····· 51

부록 1. 표본유지율 결정요인 분석 ····· 53
2. 가중치를 사용하지 않은 경우의
소득이동 및 상태변화 분석 ····· 59

Ⅰ. 서론

최근 가계, 기업관련 다양한 조사 분석을 위해 패널자료를 이용하는 경우가 급증하고 있음. 이러한 연구수요에 부응하기 위해 국내외에서 다양한 패널자료가 작성되고 있으며, 특히 한국의 경우 패널자료의 수집이 급격하게 증가하고 있다. 다음의 <표 1-1>은 현재 실시되고 있는 대표적인 정부출연기관의 패널조사를 열거한 내용이다. 이에 따라 동 표에 제시된 패널조사의 2007년 예산이 68억 원에 달하고 있다.

<표 1-1> 정부출연연구기관의 패널사업 실시현황(단위 : 백만원)

개시연도	주관기관	조 사 명	사업기간	'07년 예산(안)
1998	한국노동연구원	■ 한국노동패널사업	1998~계속	780
2001	산업연구원	■ KIET 제조업 실태조사	2001~2006	300
2002	한국노동연구원	■ 사업체패널사업	2002~계속	500
	한국보건사회연구원	■ 저소득층자활사업실태조사	2002~2006	282
2003	한국청소년개발원	■ 한국청소년패널사업	2003~2009	550
2004	한국직업능력개발원	■ 한국교육고용패널사업	2004~2015	800
2005	한국교육개발원	■ 한국교육종단연구	2005~2023	200
	한국보건사회연구원	■ 차상위·빈곤층패널사업	2005~계속	1,300
	한국직업능력개발원	■ 인적자본기업패널사업	2005~계속	300
2006	한국노동연구원	■ 고령화연구패널사업	2006~계속	700
	한국여성개발원	■ 여성가족패널사업	2006~2010	500
2007	에너지경제연구원	■ '가정·건물부문 상설표본 운영체계 구축'조사	2007	100
	한국보건사회연구원	■ 의료이용 및 의료비 패널구축	2007~계속	500

주 : 2007년부터 한국보건사회연구원의 '차상위·빈곤층 패널사업'와 '저소득층자활사업실태조사'는 통합되었음. 이 외 노동부 산하의 한국고용정보원에서 2001년부터 '청년패널'조사를 수행중이며, 2006년부터 대졸자직업이동패널조사를 실시 중임.

자료 : 경제인문사회연구회

그런데 패널조사를 실시하는 경우 동 자료로부터 일반적인 반복횡단면자료로는 분석하기 어려운 다양한 분석이 가능하다는 장점이 있으나, 새로운 패널조사를 실시하는 데는 비용이

많이 소요된다는 문제점이 있을 수 있다. 국내에서 다양한 패널자료가 생성됨에 따라 통계청에서도 패널자료를 만들어야 한다는 의견이 제시되고 있는 바, 비용측면과 인력측면 그리고 기술적인 측면에서 통계청에서 새로운 패널자료를 만든다는 방안은 비현실적이라는 의견도 없지 않다.

이와 같은 점을 감안할 때 먼저 통계청이 실시한 다양한 반복횡단면자료를 패널자료로 연결하여 패널자료를 새로 수집하는 것과 유사한 효과를 가질 수 있는 지를 검토하는 작업은 매우 의미있는 작업이라고 할 수 있다. 통계청에서 국가예산으로 작성한 자료를 약간의 추가적인 노력으로 가공하여 정책적 시사점을 가지고 있는 새로운 사실을 도출해낼 수 있다면 이는 국가예산의 효율적인 사용이라는 측면에서 매우 바람직할 일이라고 할 수 있을 것이다. 또한 다양한 패널자료가 생성되고 있고 패널자료에 대한 분석필요성이 제기되고 있는 현실을 감안할 때 통계개발원에서 통계청직원들을 위해 사용할 수 있는 패널자료 분석기법을 정리하여 제공하는 작업도 매우 의미 있는 작업이라고 할 수 있다.

이에 본 연구에서는 통계청에서 가지고 있다는 기존의 반복횡단조사를 패널자료로 구축하는 방안을 연구하고자 한다. 통계청에서 동일한 대상을 상대로 특정 주제에 대해서 반복적으로 측정하는 반복횡단조사는 패널조사의 속성을 지닌다. 이미 미국, 영국, 캐나다, 뉴질랜드 통계청에서 반복횡단조사를 패널자료로 구축하여 제공하고 있다.

반복 횡단면자료를 패널자료화 함으로써 다음과 같은 목적을 달성하고자 한다. 기존의 패널자료분석에 대한 연구결과를 토대로 패널자료분석을 통해 얻어질 수 있는 추가적인 분석가능성을 정리하고, 이러한 분석가능성을 농가경제조사 자료를 이용하여 구축한 패널자료에 이용하여 실증적으로 제시하고자 한다. 이와 같은 작업을 통해 농가경제조사 이외의 다른 반복횡단면 자료의 패널자료화 가능성에 대한 시사점을 도출하고 자 한다.

Ⅱ. 국내외 패널데이타 작성 현황과 패널데이타 분석 방법 소개

1. 국내외 패널데이타 작성 현황[1)]

가. 국내패널조사 현황

국내에서는 다양한 패널조사가 시행되고 있으며, 대부분은 국책연구기관에 의해 실시되고 있다. 다양한 패널조사는 경제·인문사회연구회 소속연구기관들의 연구회의 예산관할 하에서 실시하는 패널조사와 각 부처가 각 연구기관에 직접 발주한 패널조사 등으로 구분된다. 먼저 경제·인문사회연구회의 관할 하에서 실시되고 있거나 실시예정인 사업들은 다음과 같다. 이러한 패널조사에 대한 보다 자세한 사항은 이희길(2007)을 참고할 수 있다.

□ 학업중단자의 실태 파악과 교육 지원을 위한 종단조사(한국교육개발원)

학업 중단 청소년은 자퇴, 재입학, 자퇴 등을 경험하며, 이 과정에서 필요한 학습이 제대로 이루어지지 않는다. 이들의 학습, 취업 지원을 위해서는 학업중단 이후의 생활실태를 파악할 필요가 있다. 이러한 실태 파악을 위해서는 학업 중단 이후 성인이 되기까지 변화를 종단적으로 추적조사를 할 필요가 있다. 동 조사는 학업 중단 청소년을 유형화하여 유형별 특성을 파악하고, 학업 중단 청소년을 위한 시설 및 프로그램에 대하여 활용도를 진단하고, 개선방안을 제시하는 데 사용될 수 있다. 동 자료를 통하여 학교 밖 청소년들의 교육복지 수준 제고를 위한 기초자료를 확보하고, 학교 밖 청소년들을 위한 교육 및 훈련 프로그램과 복지 지원체제 구축방안을 제시할 수 있다. 2006년부터 1만명의 표본으로 시작하여 향후 10년 동안 추적조사하고 있다.

□ 한국교육종단조사(한국교육개발원)

한국교육종단조사는 국가수준의 종단적 교육자료 구축을 통해 교육정책 연구에 필요한 기초자료를 확보하고, 교육통계조사자료의 보완(학업성취, 심리적 요소, 진학 등에 대한 개인수준의 시계열 자료의 구축 가능)을 목적으로 하고 있다. 특히 한국교육종단조사는 고교 간 및 지역 간 교육격차의 분석을 가능하게 하고, 평준화 및 계열화 등의 교육정책과 관련된 실

1) 이 부분은 경제·인문사회연구회 협동연구 총서 07-03-01, "사회·복지 프로그램 패널자료 구축에 관한 타당성 연구"를 주로 참고하였다.

증적 분석에 필요한 정보를 제공함으로써 교육연구의 활성화에 이바지하고 있다. 2003년에 7,000명을 대상으로 1차 년도 조사를 시작하였고, 이후 17년을 추적조사할 예정이다. 중학교 표본을 시작으로 고등학교까지 매 1년씩 조사하고, 고등학교 이후 30세까지는 매 2년씩 조사하고, 30세 이후에는 1회 조사하는 것으로 되어 있다.

□ 한국복지패널조사(한국보건사회연구원)

한국복지패널조사는 국민의 경제·사회적 행태변화, 특히 빈곤층 및 차상위층의 규모·실태 변화에 대한 분석을 통해 현재의 사회·복지제도를 효과적으로 개편할 수 있는 방안을 제시하는 데 있어 중요한 자료가 된다. 또한 한국복지패널조사는 정책욕구 및 수요의 체계적인 파악에 일조하고, 정책우선 지원 순위의 결정 및 그에 따른 중장기 재정수요 파악을 통해 합리적인 복지지출계획의 수립을 지원하고 있다. 복지의존성이 한 세대에서 다음 세대로 이전되느냐를 규명할 수 있는 빈곤의 세대 간 연구 등과 같은 심층 분석자료 제공을 통해 학술연구 발전에 기여하고, 장기적인 사회통합성 제고방안에 기여한다. 정부정책 또는 경제적· 인구학적 여건변화에 대한 충격분석 및 각 정책의 효과성 평가를 통해 사회·복지정책의 개선방안을 제시한다. 패널표본은 장기적인 신뢰성·효율성·안정성 등의 원칙에 따라, 국민생활실태조사(표본 30,000가구)에서 표본을 추출하였으며, 2006년에 7,072가구를 1차 웨이브 표본으로 조사 완료하였다. 매 1년 추적조사할 예정이다.

□ 의료 이용 및 의료비 패널조사(한국보건사회연구원)

의료 이용 및 의료비 패널조사는 의료 이용의 원인과 결과에 대한 동태적인 구조관계를 규명하여 급증하고 있는 의료 이용 및 의료비 지출 규모의 적정성을 유도할 수 있는 보건의료 정책을 시현하는 데 기여함을 목적으로 하고 있다. 자료의 분석을 통해 주요 질병의 효율적 관리가 가능하게 되고, 보건재정의 중복 낭비요인 제거를 통해 보건에의 투자 효율성 제고를 기대할 수 있다. 또한 보건·의료 분야 국가 중장기 재정운용계획 수립에 기초자료로 활용된다. 중장기적으로 의료 이용 및 의료비 패널조사는 보건의료정책의 중장기 방향 수립 및 제도개선과 국민의 삶의 질 향상에 기여할 수 있으며 단기적으로 보건의료 관련 DB 등의 효과적인 활용으로 정책구상 및 개선에 기여한다. 2007년도에 예비조사가 시행되었고, 2008년부터 1차 본조사가 시행될 예정이다.

□ 한국여성가족패널조사(한국여성정책연구원)

한국여성가족패널조사는 여성정책의 핵심적 사항인 여성의 경제활동참여 촉진 및 경제적 자립 증진, 여성빈곤화 방지 등의 정책수립을 위한 기초자료로 활용된다. 자료는 가족의 변화 방향을 파악하고 객관적 사항에 의거한 가족정책을 수립하는 데 사용되며, 양성평등한 가족정책의 수립과 현재 가족문제의 핵심적 사항들인 보육, 이혼, 동거, 가족의 다양화, 저출산, 노인부양, 빈곤 문제 등을 파악하고 지원할 수 있는 토대를 마련한다. 조사자료가 구축되면, 향후 다양한 집단의 여성인력 양성과 활용 등에 관련된 제반 문제점을 파악할 수 있는 기초자료로 활용되어 체계적인 인력정책 수립에도 기여하게 될 것이다. 2006년도의 예비조사를 거쳐 2007년에는 1차 본조사가 완료되었다. 2007년도부터 시작하여 약 10,000명의 성인여성과 그 성인여성이 소속된 가구들을 추적조사하고 있다.

□ 한국영유아종단조사(육아정책연구센터)

한국영유아종단조사는 출생에서부터 취학 전까지 아동의 발달 및 생활에 관한 국가수준의 횡단적·종단적 정보의 제공을 목적으로 하고 있다. 아동의 성장·발달에 영향을 미치는 부모, 가정, 육아지원으로부터 제공되는 경험에 대한 기초자료를 제공하고 이를 통해 아동의 긍정적인 발달과 적응에 영향을 미치는 부모의 양육, 가정환경 및 육아지원의 특성 등을 파악하여 바람직한 양육 및 효과적인 육아지원방안을 모색할 수 있으며 수집된 자료를 활용하여 정기적으로 학술대회개최를 개최함으로써 관련 연구를 활성화시킬 것으로 기대되고 있다. 영유아 가족을 대상으로 하는 육아정책에 대한 패널가구의 인식 및 만족도를 점검하고 관련 수요를 조사함으로써 정책 모니터링에 기여한다. 2007년도에 예비조사를 실시하였고, 2008년도부터 본조사가 시행될 예정이다.

□ 한국교육고용패널조사(한국직업능력개발원)

인적자원의 개발과 활용에 관한 정책 수립의 기초자료를 제공하고, 중등-고등교육의 이행과 고등-직업으로의 이행에 관한 자료를 분석하여, 교육에서 노동시장으로의 이행과 활용에 대한 정책수립의 기초자료를 제공한다. 학교교육 이후 직업이동 과정에서의 교육훈련 효과 분석을 통해 직업훈련정책 수립의 기초자료를 제공하며 교육과 노동시장의 연계고리를 규명함으로써 고용정책 수립과 노동시장과 연계된 교육정책 수립의 기초자료로서의 역할을 하고 있다. 특히 대학·연구기관 등의 연구자를 위한 연구자료로 제공되어 교육과 고용에 대한 연관성 분석의 기초가 되고 있다. 2004년도에 1차 조사를 시작하였고, 2008년도에는 5차 웨이브가 진행될 예정이다.

□ 한국청소년패널조사(한국청소년정책연구원)

한국청소년패널조사는 청소년정책의 중장기 비전 설정 및 과학적 정책 수립·시행을 위한 기초자료를 제공하고, 체계적인 통계DB의 구축 및 제공을 통해 청소년 관련 정책연구 활성화에 기여함을 목적으로 한다. 조사를 통하여 청소년들의 잠재적 직업선택, 향후 진로설정 및 준비, 여가, 일탈행위 등의 생활실태에 대한 횡단적 실태를 파악할 수 있고, 종단적 변화양상 및 종단적 변화에 영향을 미치는 제 요인에 대한 정교한 이해를 얻을 수 있다. 횡단적·종단적 데이터를 관련 학계의 연구자들에게 공급하고, 이를 통하여 현실적인 연구여건의 한계를 극복하는 질 높은 논문의 양산을 유도함으로써, 관련 학계의 학문발전에 기여할 수 있을 것으로 기대된다. 2003년에 전국의 중학교 2학년 청소년들 및 학부모 3,500명(중2 패널: 2003년 조사시작~2008년까지 추적조사)과 2004년에 전국의 초등학교 4학년 청소년들 및 학부모 2,500명(초4 패널: 2004년 조사시작~2008년까지 추적조사)을 대상으로 추적조사를 실시하기 시작하였다.

□ 한국노동패널조사(한국노동연구원)

현재 우리나라에서 실시하고 있는 각종 노동시장 관련 통계조사들은 아직 그 내용이 빈약하여 개인의 특성 및 의사결정에 바탕을 둔 미시적이고 동태적인 차원에서의 노동시장의 이해 및 정책결정에 별다른 도움을 주지 못하고 있는 실정이다. 또한 경제와 노동시장의 국제화가 급속히 진행되고 있는 가운데 다른 나라들과의 비교분석이 가능한 노동시장 자료 또한 부재한 상황이다. 이러한 시점에서 한국노동패널조사는 노동시장 관련 기초 조사자료의 미흡 및 부재의 문제를 해소하고 나아가 분석적인 노동시장연구를 활성화시킴으로써, 보다 합리적이고 정확한 노동시장 및 고용정책의 수립과 평가에 기여함을 목적으로 하고 있다. 동 자료는 각종 노동정책의 수립 및 평가, 고용보험제도의 효율적인 운영과 장기적인 발전에 기여할 뿐만 아니라 노동 관련 연구의 활성화를 위한 귀중한 기초자료로 자리 잡을 것으로 기대된다. 1998년에 전국 비농어촌지역에 거주하는 약 5,000가구를 대상으로 조사를 시작하였으며, 2007년에는 10차 조사가 진행되었다.

□ 고령화연구패널조사(한국노동연구원)

우리나라는 고령화에 대한 기초적 자료가 미흡한 실정으로, 체계적인 통계자료 구축이 필요하다. 고령사회를 대비한 제도 개혁 및 정책 결정을 위해서는 고령자의 노동시장 참여, 소

득 및 자산 현황, 개인의 은퇴결정, 사회보장제도의 효과, 건강, 가족 내 소득 이전 등을 파악하기 위한 체계적인 자료 구축이 요구된다. 고령화연구패널조사는 향후 고령사회로 변화해 가는 과정에서 효과적인 사회경제정책을 수립하고 시행하는 데에 활용될 기초자료를 생산하는 데에 목적이 있다. 이를 통해 고령사회의 다양한 측면들을 포착하고, 여러 분야의 학제적 연구가 가능한 자료를 구축하며, 나아가 미국, 유럽 등 이미 중고령자 대상 패널조사를 실시하고 있는 선진국과 비교가 가능한 자료를 생산하여 고령화 정책 및 학술연구의 기초자료로 활용한다. 2006년에 약 10,000명을 대상으로 조사를 시작하였으며, 짝수년에는 동일한 설문지를, 홀수년에는 특정 주제를 정하여 조사를 실시한다. 2007년에는 개인 생애사를 조사하고 있다.

□ 기 타

이상과 같은 조사 이외에도 보건복지부가 주관하여 한국보건사회연구원에서 실시할 예정인 인구패널조사가 있으며, 국민연금관리공단에서 실시하는 노후보장소득패널, 여성가족부에서 주관하여 한국여성정책연구원에서 실시하고 있는 여성인력패널조사, 노동부 산하의 장애인고용촉진공단에서 실시하고 있는 장애인패널조사, 근로복지공단에서 기획하고 있는 산재패널조사가 있다. 또한 한국고용정보원에서 실시하고 있는 청년패널과 대졸자직업이동경로조사가 있다. 청년패널은 2001년부터 시작한 바 있으며, 2007년에 새로운 표본을 구축하여 조사를 실시하였다. 대졸자직업이동경로조사는 2006년부터 조사를 실시하였다. 노동부는 비정규직 패널조사를 기획하고 있다.

<표 2-1>은 이상의 주요 패널조사 현황을 조사주체, 진행상황, 표본 수, 표집방법 및 주요조사 내용을 중심으로 간단히 정리·비교하고 있다. 한편 <표 2-2>은 표본대상별로 주요 패널조사를 재분류하고 있는데, 대부분의 패널이 학력, 연령 등 조사대상의 인적특성에 따라 추적조사되고 있는 것을 알 수 있다.

나. 해외 패널조사 현황

외국의 경우 패널분석의 중요성이 부가됨에 따라 다양한 목적으로 가구 및 개인 패널조사가 시행되고 있으며, 대부분의 패널조사에 대해서는 국내에 이미 널리 소개되어 있다. 각종 패널조사에 대한 구체적인 내용은 해당 조사의 홈페이지를 참조할 수 있다.

각국을 대표하는 가구패널조사에 대해서는 강석훈 외(1994)의 『한국가구경제활동연구』를 참고할 수 있으며, 여성 및 가족과 관련된 패널조사는 박수미 외(2006)의 『여성가족패널조사』에 자세히 설명되어 있다. 한편 교육 및 청년관련 패널에 대해서는 채창균 외(2003)의 『한

국교육고용패널조사』에서 소개되고 있으며, 청소년 및 아동패널조사에 대해서는 한국청소년정책연구원(2002)의 『한국청소년패널조사의 방향과 과제』에서 비교적 상세하게 다루고 있고, 영유아패널조사에 대해서는 장명림 외(2006)에서 다루고 있다.

<표 2-3>는 Haisken-DeNew(2001)에서 소개하고 있는 각국 가구패널조사이다.

<표 2-1> 국내 주요 패널조사 현황

조사명	조사주체	시작연도	진행상황	표본 수	표집방법	주요 조사내용
한국노동패널조사	한국노동연구원	1998	2007년 10차 조사 진행	6,700가구	2단계 층화집락계통추출	가구특성, 경제활동 및 노동시장 이동, 소득활동 및 소비, 교육 및 직업훈련 등
고령화연구패널조사	한국노동연구원	2006	2007년 개인 생애사 조사	10,000명	집락층화 표집	인구, 가족, 건강, 고용, 소득, 자산, 주관적 기대감 및 삶의 만족도
비정규직패널조사	한국노동연구원	2007	예비조사			
청소년패널조사	한국청소년정책연구원	2003	2007년 5차 조사	2,949명	층화다단계집락표집	학생: 직업선택, 여가, 진로, 일탈, 자아관 부모: 가족구성, 근로형태, 소득 등
한국교육고용패널조사	한국직업능력개발원	2004	2007년 4차 조사	6,000명	층화집락추출법	학교/가정/여가생활, 사교육, 진로, 자아관, 경제의식 등
청년패널 조사	한국고용정보원	2001	2007년 7차 조사	5,000명	산업직업별 고용조사에서 이중추출	학교생활, 직업관 및 향후 진로, 직업이동 및 구직 활동, 직업교육훈련 이수
		2007	2007년 추가	12,000명		※ 새로운 표본으로 시작
대졸자직업이동경로조사	한국고용정보원	2006	2007년 2차 조사	25,000명	다단계 층화 추출법	교육과정, 일자리, 재학 중 취업, 훈련과정, 자격증 등
국민노후보장패널조사	국민연금연구원	2005	2007년 2차 조사	5,000가구	확률비례 층화	고령자의 퇴직, 고용, 지출, 자산, 건강, 주거 등
한국복지패널조사	한국보건사회연구원	2006	2007년 2차 조사	7,000가구	층화이중추출	경제활동상태, 건강, 의료, 장애인, 노인, 아동, 연금 수급, 직업훈련, 생활실태
인구패널 조사	한국보건사회연구원	2007	예비조사	6,000~10,000명	확률비례 층화	교육, 직업, 건강, 소득, 주거 등
의료패널 조사	한국보건사회연구원	2007	예비조사	5,000가구	층화추출	병원이용, 의료비지출 정보, 개인건강상태, 가구 및 개인 특성 등
여성가족패널조사	한국여성정책연구원	2007	2007년 1차 조사	10,000명	다단계 층화계통 추출	여성과 가족, 여성경제활동, 여가 등

여성인력 패널조사	한국여성정책 연구원	2007	2007년 1차 조사	3,000명	비례할당 표집	여성 관리자의 경제활동, 기업내 승진, 교육훈련, 육아, 경력단절 등
장애인패널조사	한국장애인 고용촉진공단	2007	예비조사	5,000명	층화계통 추출	
아동패널 조사	육아정책 개발센터	2007	예비조사	3,000명	27개 원유층 무선표집후 전수조사	양육실태, 성장 발달 특성
한국교육 종단연구	한국교육 개발원	2005	2007년 3차 조사	6,908명	다단계 층화 군집추출법	교육성취, 학생배경, 학교교육 등

<표 2-2> 표본대상별 패널조사 현황 요약

<table>
<tr><th rowspan="2">유아기</th><th colspan="3">청소년기</th><th>대 학</th><th>청 년</th><th>중년</th><th>노 년</th></tr>
<tr><th>초 등</th><th>중 등</th><th>고 등</th><th colspan="4"></th></tr>
<tr><td rowspan="2">육아 패널</td><td rowspan="2">청소년패널</td><td rowspan="2">청소년패널
교육패널
교육고용패널</td><td rowspan="2">교육패널
교육고용패널</td><td>대졸자 패널</td><td>청년패널
대졸자패널</td><td></td><td>고령화패널
국민노후 보장패널</td></tr>
<tr><td colspan="4">노동패널, 복지패널, 여성가족패널</td></tr>
</table>

<표 2-3> 각국의 주요 가구패널조사

국 가	조사명	홈페이지 및 작성기관
미 국	PSID (1968~)	www.isr.umich.edu/src/psid/ ISR, University of Michigan
독 일	GSOEP (1984~)	www.diw.de/gsoep LDM, DIW Berlin
영 국	BHPS (1991~)	www.irc.essex.ac.uk/bhps/index.php ISER, University of Essex
캐나다	SLID (1993~)	www.statcan.ca:80/english/sdds/3889/htm Statistics Canada
러시아	RLMS (1994~)	www.cpc.unc.edu/projects/rlms CPC, University of North Carolina
러시아	RUSSET (1993~1999)	home.pscw.uva.nl/saris University of Amsterdam
호 주	HILDA (2001~)	www.melbourneinstitute.com/hilda Melbourne Institute, Melbourne University
이스라엘	ILS (2002/3~)	www.cbs.cbs.il/engindex.htm Israel Central Bureau of Statistics
스웨덴	HUS (1984~98)	www.handels.gu.se/econ/econometrics/hus/husin.htm Dept. of Economics, Gothernburg University
벨기에	SEP (1985~1992)	www.ufsia.ac.br/~csb/eng/septab.htm CSP, University of Antwerp
룩셈부르크	PSELL (1985~)	www.ceps.lu/psell/pselpres.htm CEPS-INSTEAD
인도네시아	IFLS (1993~)	www.rand.org/labor/FLS/IFLS/intro.html RAND
일 본	JPSC (1993~)	www.kakeiken.or.jp/english/index/html The Institute of Household Economy
한 국	KLIPS (1998~)	www.kli.re.kr/klips/klips_eng.htm Korea Labor Institute
스위스	SHP (1999~)	www.unine.ch/psm.welcome.html University of Neuchatel and SFSO

자료: Haisken-DeNew(2001).

2. 기본적인 패널데이타 분석 방법[2)]

패널데이타는 기본적인 N명의 개인과 각각의 개인(i)에 대응되는 T_i기간 동안의 관측치로 구성된 데이터이다. 이 때 모든 개인에 대하여 T_i가 T로 동일한 경우를 균형패널(balanced panel) 이라고 하고, 가 개인별로 상이한 경우를 불균형패널(unbalanced panel)이라고 한다.

가. 분산분석

일반적으로 패널데이타를 이용한 기본적인 선형 다변량분석모형을 다음과 같이 쓸 수 있다[3)].

$$y_{it}=\alpha_{it}+x_{it}'\beta_{it}+\varepsilon_{it}$$

단, $\beta_{it}'=(\beta_{1it},\beta_{2it},....\beta_{Kit})$, $x_{it}'=(x_{1it},x_{2it},....x_{Kit})$, $i=1,2,,...N$ $t=1,2...T$.

이 식의 경우에는 NT개의 관측치로 $NT(K+1)$개의 파라메터를 추정하여야 한다는 점에서 추정가능한 형태는 아니다. 이에 따라 추정가능한 형태로서 다음과 같은 세가지 경우로 구분해 본다.

<2-1> $$y_{it}=\alpha_{i}+x_{it}'\beta_{i}+\varepsilon_{it}$$

<2-2> $$y_{it}=\alpha_{i}+x_{it}'\beta+\varepsilon_{it}$$

<2-3> $$y_{it}=\alpha+x_{it}'\beta+\varepsilon_{it}$$

<식 2-1>은 상수항과 기울기계수가 모든 개인에 따라 다르다는 형태이며, <식 2-2>는 상수항은 모든 개인에 따라 다르지만, 기울기계수는 모든 개인에 따라 동일하다는 형태이며, <식 2-3>은 상수항과 기울기가 모든 개인에 관계없이 동일하다는 형태이

2) 이 부분은 한국청소년정책연구원에서 개최한 패널데이타 분석방법론 세미나에서 강석훈(2007)이 발표한 내용을 수정, 보완하였다.

3) 편의상 균형패널을 분석대상으로 한다. 불균형 패널데이타를 이용하는 경우 선형분석방법에서는 약간의 수정을 가하여 사용가능하다. 비선형분석의 경우에는 새로운 방법이 요구되는 경우가 많다.

다. 이러한 세 가지 형태 중에서 선택하는 방법으로 다음과 같은 검정방법을 고려해 볼 수 있다.

가설 H_1은 <식 2-1>에서 상수항은 상이하고, 기울기계수는 동일하다는 가설을 의미하고 H_3은 <식 2-1>에서 상수항과 기울기계수가 모든 개인에 따라 동일하다는 점을 의미한다[4].

가설 H_1은 <식 2-1>에서 다음과 같은 $(N-1)K$개의 제약을 부과한 경우라고 볼 수 있다.

$$H_1 : \beta_1 = \beta_2 = \ldots = \beta_N$$

또한 가설 H_3은 <식 2-1>에서 다음과 같은 $(N-1)(K+1)$개의 제약을 부과한 경우라고 볼 수 있다.

$$H_1 : \alpha_1 = \alpha_2 = \ldots = \alpha_N,\ \beta_1 = \beta_2 = \ldots = \beta_N$$

가설 H_3을 검정하기 위해서는 정규분포의 가정 하에서의 다음의 F-통계량을 사용할 수 있다. 이 때 F-분포의 자유도는 각각 $(N-1)(K+1), NT-N(K+1)$이다.

$$F_3 = \frac{(S_3 - S_1)/[(N-1)(K+1)]}{S_1/[NT-N(K+1)]}$$

여기에서 $S_1 = \sum_{i=1}^{N} RSS_i$ (<식 2-1>을 모든 개인에 대하여 회귀분석할 때의 잔차제곱합의 총합)이다. 즉,

$$RSS_i = W_{yy,i} - W'_{xy,i} W_{xx,i}^{-1} W_{xy,i}$$

$$W_{yy,i} = \sum_{t=1}^{T} (y_{it} - \overline{y_i})^2,\quad W_{xx,i} = \sum_{t=1}^{T} (x_{it} - \overline{x_i})(x_{it} - \overline{x_i})',\quad W_{xy,i} = \sum_{t=1}^{T} (x_{it} - \overline{x_i})(y_{it} - \overline{y_i})$$

$$\overline{y_i} = (1/T) \sum_{t=1}^{T} y_{it},\quad \overline{x_i} = (1/T) \sum_{t=1}^{T} x_{it}$$

4) 이외에도 상수항은 동일하지만 기울기가 개인에 따라 다르다는 가설도 고려해 볼 수 있지만 이는 현실적으로는 의미가 없는 가설이다.

$S_3 = T_{yy} - T'_{xy} T_{xx}^{-1} T_{xy}$(<식 2-3>의 잔차제곱합)

$T_{yy} = (1/NT) \sum_{i=1}^{N} \sum_{t=1}^{T} (y_{it} - \bar{y})^2$, $T_{xx} = (1/NT) \sum_{i=1}^{N} \sum_{t=1}^{T} (x_{it} - \bar{x})(x_{it} - \bar{x})'$,

$T_{xy} = (1/NT) \sum_{i=1}^{N} \sum_{t=1}^{T} (x_{it} - \bar{x})(y_{it} - \bar{y})'$

$\bar{y} = (1/NT) \sum_{i=1}^{N} \sum_{t=1}^{T} y_{it}$, $\bar{x} = (1/NT) \sum_{i=1}^{N} \sum_{t=1}^{T} x_{it}$

만약 가설 H_3가 채택되면 자료를 통합(pool)하여 <식 2-3>의 단일 방정식을 추정한다. 만약 H_3가 기각되면 H_1가설을 검정한다.

가설 H_1을 검정하기 위해서는 다음과 같은 F-통계량을 사용한다. 이 때 F-분포의 자유도는 각각 $(N-1)K, NT-N(K+1)$가 된다.

$$F_2 = \frac{(S_2 - S_1)/(N-1)K}{S_1/[NT-N(K+1)]}$$

여기에서 $S_2 = W_{yy} - W'_{xy} W_{xx}^{-1} W_{xy}$ (<식 2-2>의 잔차제곱합).

$W_{xx} = \sum_{i=1}^{N} W_{xx,i}$, $W_{yy} = \sum_{i=1}^{N} W_{yy,i}$ $W_{xy} = \sum_{i=1}^{N} W_{xy,i}$

만약 H_1이 기각되면 <식 2-1>을 이용한다. 만약 H_1이 채택되면 <식 2-2>를 이용한다[5].

나. 정태 선형분석

우리는 다음과 같은 회귀분석 모형을 고려한다.

---

5) 이와 같은 방법은 추정계수가 시간에 따라 변화한다는 모형에서도 동일하게 사용될 수 있다. 즉, 다음과 같은 모형에도 적용된다.

$y_{it} = \alpha_t + x_{it}'\beta_t + \varepsilon_{it}$

$y_{it} = \alpha_t + x_{it}'\beta + \varepsilon_{it}$

$y_{it} = \alpha + x_{it}'\beta + \varepsilon_{it}$

<2-4> $y_{it}=\alpha_i+x_{it}'\beta+\varepsilon_{it},\quad i=1,2,\ldots N,\quad t=1,2,\ldots T$

□ 고정효과모형(Fixed Effects Model)

고정효과모형은 α_i를 고정된 값으로 취급하는 모형이며, 최소자승더미변수모형(least squares dummy-variable model, LDV) 또는 공분산모형(analysis of covariance model)이라고 불린다.

<식 2-4>는 다음과 같이 벡터형태로 쓸 수 있다.

$$Y=\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{bmatrix}=\begin{bmatrix} e \\ 0 \\ \vdots \\ 0 \end{bmatrix}\alpha_1+\begin{bmatrix} 0 \\ e \\ \vdots \\ 0 \end{bmatrix}\alpha_2+\ldots+\begin{bmatrix} 0 \\ 0 \\ \vdots \\ e \end{bmatrix}\alpha_N+\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{bmatrix}\beta+\begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_N \end{bmatrix}$$

단,

$$y_i=\begin{bmatrix} y_{i1} \\ y_{i2} \\ \vdots \\ y_{iT} \end{bmatrix},\quad X_i=\begin{bmatrix} x_{1i1} & x_{2i1} & \cdots & x_{Ki1} \\ x_{1i2} & x_{2i2} & \cdots & x_{Ki2} \\ \vdots & \vdots & \vdots & \vdots \\ x_{1iT} & x_{2iT} & \cdots & x_{KiT} \end{bmatrix},\quad e=\begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix},\quad \varepsilon_i=\begin{bmatrix} \varepsilon_{i1} \\ \varepsilon_{i2} \\ \vdots \\ \varepsilon_{iT} \end{bmatrix}$$

$E(u_i)=0,\quad E(\varepsilon_i\varepsilon_i')=\sigma_i^2 I_T,\quad E(\varepsilon_i\varepsilon_j')=0 \text{ if } i\neq j$

잔차제곱합을 최소화하는 방법으로 다음과 같은 추정량을 얻을 수 있다.

$$\hat{\beta}_{CV}=\left[\sum_{i=1}^{N}\sum_{t=1}^{T}(x_{it}-\overline{x_i})(x_{it}-\overline{x_i})'\right]^{-1}\left[\sum_{i=1}^{N}\sum_{t=1}^{T}(x_{it}-\overline{x_i})(y_{it}-\overline{y_i})\right]$$

$$\hat{\alpha}=\overline{y_i}-\overline{x_i}'\hat{\beta}_{CV}$$

단, $\overline{x_i}=\frac{1}{T}\sum_{t=1}^{T}x_{it},\quad \overline{y_i}=\frac{1}{T}\sum_{t=1}^{T}y_{it}$

$\hat{\beta}_{CV}$은 집단내추정량(Within Estimator), 공분산추정량(Covariance Estimator) 그리고 고정효과추정량(Fixed Effect Estimator)을 불린다. $\hat{\beta}_{CV}$은 다음과 같은 형태로 표기될 수 있다.

$$\hat{\beta}_{CV}=\left[\sum_{i=1}^{N}X_i'Q^{-1}X_i\right]^{-1}\left[\sum_{i=1}^{N}X_i'Q^{-1}y_i\right]$$

단, $Q = I_T - (1/T)ee'$.

공분산추정량은 다음과 같은 성질을 가진다.

$$Var(\hat{\beta}_{CV}) = \sigma_\varepsilon^2 \left[\sum_{i=1}^{N} X_i' Q X_i \right]^{-1}$$

$$plim_{N\to\infty} \hat{\beta}_{CV} = \beta$$

$$plim_{T\to\infty} \hat{\beta}_{CV} = \beta$$

$$plim_{T\to\infty} \hat{\alpha} = \alpha$$

$$plim_{N\to\infty} \hat{\alpha} \neq \alpha$$

β_{CV}는 최소분산불편선형추정량(BLUE)가 된다. 상수항의 경우에는 T가 증가하면 모수에 대한 일치추정량이 되지만, N이 증가하는 경우에는 일치추정량이 되지 않는다. 동 추정량은 기본적으로 개인효과를 차분을 통해 제거하는 방법을 사용하고 있기 때문에 추정모형에 시간불변변수(time-invariant regressors)가 있는 경우에 동 변수의 계수값은 추정되지 않는다.

□ 임의효과 모형

<식 2-4>는 다음과 같이 쓸 수 있다.

$$y_{it} = \mu + x_{it}'\beta + \alpha_i + \varepsilon_{it} = \mu + x_{it}'\beta + u_{it}$$

단, $u_{it} = \alpha_i + \varepsilon_{it}$.

개별 효과를 나타내는 α_i를 확률변수로 취급하는 동 모형은 임의효과모형(Random Effect Model) 또는 분산구성모형(Variance Component Model) 또는 일방향오차구성모형(One-way Error Component Model)으로 불린다.

이 모형에서는 이제 오차항 u_{it}가 두개의 구성요소를 가지게 되는데, 각각에 대하여 다음을 가정하자.

$$E(\alpha_i) = 0$$

$$E(\varepsilon_{it}) = 0$$

$E(\alpha_i \alpha_j)=\sigma_\alpha^2$ if $i=j$, $E(\alpha_i \alpha_j)=0$ if $i \neq j$

$E(\varepsilon_{it} \varepsilon_{js})=\sigma_\varepsilon^2$ if $i=j$ and $t=s$, $E(\varepsilon_{it} \varepsilon_{js})=0$ of *otherwise*.

$E(x_{it} \mu_i)=E(u_{it} x_{it})=0$

이 때 오차항의 분산은 다음과 같다.

$$E(u_i u_i')=\sigma_\varepsilon^2 I_T+\sigma_\alpha^2 ee'=V, \quad V^{-1}=\frac{1}{\sigma_\varepsilon^2}\left[I_T-\frac{\sigma_\alpha^2}{\sigma_\varepsilon^2+T\sigma_\alpha^2} ee'\right]$$

예를 들어 T=3인 경우 이 분산행열은 다음과 같다.

$$V=\begin{bmatrix} \sigma_\varepsilon^2+\sigma_\alpha^2 & \sigma_\alpha^2 & \sigma_\alpha^2 \\ \sigma_\alpha^2 & \sigma_\varepsilon^2+\sigma_\alpha^2 & \sigma_\alpha^2 \\ \sigma_\alpha^2 & \sigma_\alpha^2 & \sigma_\varepsilon^2+\sigma_\alpha^2 \end{bmatrix}$$

임의효과모형을 공분산추정량으로 추정하는 경우에도 동 추정량은 기울기계수에 대하여 불편추정량이고 일치추정량이 되지만, 효율적인 추정량은 되지 못한다. 왜냐하면 공분산추정량은 그룹내 변동(within group variation)만을 반영할 뿐 그룹간 변동(between group variation)을 반영하지 못하기 때문이다.

이제 오차항의 이분산성을 감안한 일반최소자승추정량은 다음과 같다.

<2-5> $$\hat{\delta}_{GLS}=\left[\sum_{i=1}^{N} \tilde{X}_i' V^{-1} \tilde{X}_i\right]^{-1}\left[\sum_{i=1}^{N} \tilde{X}_i' V^{-1} y_i\right]$$

단, $\delta'=(\mu, \beta')$. 이 일반최소자승추정량은 추정하고자 하는 모수의 최소분산선형불편추정량이 된다.

한편, 동 일반자승최소추정량은 다음과 같이 도출될 수도 있다. 먼저 $\overline{y_i}$ 와 $\overline{x_i}$ 의 $(1-\Psi^{1/2})$ 부분만을 차감한다(단, $\Psi=\frac{\sigma_\varepsilon^2}{\sigma_\varepsilon^2+T\sigma_\alpha^2}$). 그리고 차감된 즉, $[y_{it}-(1-\Psi^{1/2})\overline{y_i}]$를 상수항과 $[x_{it}-(1-\Psi^{1/2})\overline{x_i}]$에 대하여 회귀분석한다. 이 때의 추정량은 일반최소자승추정량과 같게 된다.

한편, σ_α^2와 σ_ε^2가 미지의 모수이기 때문에 실제 일반최소자승추정량을 사용하려면 동 파라메터를 추정해야 한다. 먼저, 1단계에서는 공분산추정량과 같은 일치추정량을 이용하여 동 파라메터들을 추정한 다음 2단계에서 동 추정치를 대입한 분산행열을 이용하여 일

반최소자승추정량을 구한다. 적절한 표본크기 하에서($T \geq 3$, $N-(K+1) \geq 9$, for $T=2$, $N-(K+1) \geq 10$)) 이러한 2단계 추정량은 공분산추정량보다 효율적이다.

□ 그룹내 그리고 그룹간 추정량(Within and Between Estimator)

우리는 다음과 같이 세가지 방법으로 통합회귀분석(pooled regression)을 고려할 수 있다.

<2-6> $y_{it} = \alpha + x_{it}'\beta + \varepsilon_{it}$

<2-7> $y_{it} - \overline{y}_i = (x_{it} - \overline{x}_i)'\beta + \varepsilon_{it} - \overline{\varepsilon}_i$

<2-8> $\overline{y}_i = \alpha + \overline{x}_i'\beta + \overline{\varepsilon}_i$

OLS방법을 <식 2-6>, <식 2-7>과 <식 2-8>에 적용하여 통상적인 OLS추정량 $\hat{\beta}_{ols}$, 그룹내 추정량 $\hat{\beta}_W$ 그리고 그룹간 추정량 $\hat{\beta}_B$을 얻을 수 있다. 약간의 변형을 통하여 우리는 다음과 같이 OLS추정량이 그룹내 추정량과 그룹간 추정량의 가중평균임을 보일 수 있다.

$$\hat{\beta}_{OLS} = F^W \hat{\beta}_W + (I - F^W)\hat{\beta}_B$$

단, $F^W = [S_{xx}^W + S_{xx}^B]^{-1} S_{xx}^W$, $S_{xx}^W = \sum_{i=1}^{N}\sum_{t=1}^{T}(x_{it} - \overline{x}_i)(x_{it} - \overline{x}_i)'$,

$S_{xx}^B = \sum_{i=1}^{N} T(\overline{x}_i - \overline{x})(\overline{x}_i - \overline{x})'$.

또한 GLS추정량도 다음과 같이 그룹내 추정량과 그룹간 추정량의 가중평균임을 보일 수 있다.

$$\hat{\beta}_{GLS} = H^W \hat{\beta}_W + (I - H^W)\hat{\beta}_B$$

단, $H^W = [S_{xx}^W + \Psi S_{xx}^B]^{-1} S_{xx}^W$.

만약 $T\to\infty$, $\Psi\to 0$가 되고, 이 때 $\hat{\beta}_{GLS}=\widehat{\beta_W}=\hat{\beta}_{CV}$이 성립한다.

□ 교정효과 또는 임의효과 모형의 선택

고정효과모형은 개인간의 차이가 회귀분석의 대상이 되는 파라메터의 변화(parametric shifts)로 확신할 수 있을 때 합리적인 접근방법이다. 이 모형은 표본내의 개인에게 적용되는 것으로 이해할 수 있으며, 표본이외에는 동 모형이 적용되지 않는다는 점을 가정하고 있다. 예를 들러 국가가간 비교에서 모든 국가가 포함되어 있는 경우라면 개별효과가 고정되어 있다고 보는 것이 합리적일 것이다.

이에 반해 개별표본들이 이들을 포함하는 더욱 큰 모집단에서 임의로 추출된 표본이라고 보는 경우에는 개별효과가 임의적으로 분포되어 있다고 판단하는 것이 합리적일 것이다. 이러한 경우에는 PSID 또는 KLIPS와 같은 경우에 해당된다.

실제 실용성의 측면에서만 판단한다면 더미변수접근 방법은 횡단면표본수가 많은 종단면자료의 경우에 자유도의 측면에서 우위가 없어지며, 이 때에는 임의효과모형을 사용하는 것이 직관적으로 선호된다. 하지만 고정효과 모형도 큰 장점을 가지고 있다. 즉. 임의효과모형에서와 같이 개별효과가 다른 설명 변수들과는 모형에 사용하는 다른 설명변수와 상관관계가 없다는 가정이 합리성을 찾기가 쉽지 않다. 따라서 고정효과모형은 생략변수에서 오는 추정량으로 불일치성에서 자유롭기 어렵게 된다.

고정효과와 임의효과 모형의 구분은 T가 많아지면 공분산추정량과 일반최소자승추정량이 같아지기 때문에 의미가 없어진다. T가 고정되어 있고, N이 큰 경우 고정이냐 임의냐의 선택은 쉽지 않다. 실제로 두 모형의 추정치는 매우 큰 차이를 보인다. 많은 개인에 대해서 적은 수의 관측치만 있는 경우에는 적은 수의 관측치로부터의 정보를 효과적으로 사용하는 것이 효율적인 추정량을 얻는데에 매우 중요해진다.

고정효과모형은 연구자가 관계성을 표본내로 한정하는 조건부 분석(conditional inference)을 실시하는 것으로 볼 수 있다. 이에 반해 임의효과모형은 연구자가 모집단 전체에 대하여 비조건부 분석(unconditional inference) 또한 한계분석(marginal inference)을 실시하는 것을 볼 수 있다. 실제로 조건부 분석과 비조건부 분석의 선택은 연구자의 몫이라고 하겠다. 실제로 연구결과를 표본내로 한정한다면 고정효과 모형이 적절하며, 연구결과가 표본의 행태분석결과를 이용하여 모집단의 행태분석에 위한 것이라면 임의효과 모형이 적절하다고 하겠다.

□ Hausman검정

임의효과모형에서 임의효과를 나타내는 항과 독립변수간의 부상관성을 검정하는 방법으로 Hausman검정을 실시할 수 있다. 동 검정의 기본 아이디어는 상관관계가 없다는 가설 하에서는 공분산추정량과 일반최소자승추정량이 모두 일치추정량이지만 공분산추정량이 비효율적이며, 대립가설에서는 공분산추정량은 일치성을 가지지만, 일반최소자승추정량은 일치성을 갖지 못한다는 점에 근거한다. 구체적인 검정 통계량은 다음과 같다.

$$[\hat{\beta}_{CV}-\hat{\beta}_{GLS}]'\Sigma^{-1}[\hat{\beta}_{CV}-\hat{\beta}_{GLS}]$$

단, $\Sigma = Var(\hat{\beta}_{CV}) - Var(\hat{\beta}_{GLS})$. 상관관계가 없다는 귀무가설 하에서 위의 검정통계량은 점근적으로 자유도가 K인 카이제곱 분포를 따르게 된다.

□ 개인효과와 시간효과를 모두 포함하는 모형

고정효과모형은 다음과 같이 쓸 수 있다.

$$y_{it}=\alpha_i+\gamma_t+x_{it}'\beta+\varepsilon_{it} \quad 단,\ i=1,2\ldots N,\ t=1,2,\ldots T-1$$

$$y_{it}=\mu+\alpha_i+\gamma_t+x_{it}'\beta+\varepsilon_{it}$$

단, $i=1,2\ldots N,\ t=1,2,\ldots T-1$ and $\sum_{i=1}^{N}\alpha_i=0,\qquad \sum_{t=1}^{T}\gamma_t=0$

기울기계수의 최소자승추정량 $\hat{\beta}_{CV}$은 종속변수를 $(y_{it}-\bar{y}_i-\bar{y}_i+\bar{y})$, 독립변수를 $(x_{it}-\bar{x}_i-\bar{x}_i+\bar{x})$로 사용하는 회귀분석에서 얻을 수 있다. 단, $\bar{y}_t=(1/N)\sum_{i=1}^{N}y_{it}$.

임의효과모형은 α_i와 γ_t가 모두 확률변수이고, x_{it}와 상관관계가 없다면 다음의 가중치행열을 사용하는 일반최소자승추정량이 최소분산선형불편추정량이 된다.

$$V=\sigma_\varepsilon^2 I_{NT}+\sigma_\alpha^2 I_N\otimes ee'+\sigma_\gamma^2 e_N e_N'\otimes I_T.$$

단, $\sigma_\gamma^2=Var(\gamma_t)$.

□ 불균형 패널

고정효과모형인 경우에는 약간의 변형을 통해 추정가능하다. σ_ε^2 , $Var(\hat{\beta}_{CV})$와 F-통계량의 추정시에 자유도의 조정이 필요하며, 개별 하위의 평균을 구할 때 표본수를 조정하면 된다.

임의효과모형인 경우 분산행열의 대각원소가 더 이상 동일한 크기가 아니기 때문에 더 이상 $I\otimes V$가 되지 않는다. 또한 V가 T에 의존하기 때문에 이분산성의 문제가 발생된다. GLS를 일부분 보완하여 사용하면 추정이 가능한데, 균형패널에서 사용하였던 Ψ대신에 다음과 같은 Ψ_i를 사용하여 추정을 실시한다.

$$\Psi_i = \frac{\sigma_\varepsilon^2}{\sigma_\varepsilon^2 + T_i\sigma_\alpha^2}$$

다. 동태 선형분석

개별효과를 포함하는 동태패널분석에서는 고정효과모형이던 임의효과모형이던간에 공분산추정량은 불일치하게 된다. 다음의 모형을 고려해 보자.

<2-9> $$y_{it} = \Theta y_{i,t-1} + \alpha_i + \varepsilon_{it}$$

단, $i=1,2,\ldots N$ $t=1,2,\ldots T$, $|\Theta|<1$이고 y_{i0}는 관측가능하다. 공분산추정량은 다음과 같이 쓸 수 있다.

<2-10> $$\hat{\Theta}_{CV} = \frac{\sum_{i=1}^{N}\sum_{t=1}^{T}(y_{it}-\bar{y}_i)(y_{y,t-1}-\bar{y}_{i,-1})}{\sum_{i=1}^{N}\sum_{t=1}^{T}(y_{i,t-1}-\overline{y_{i,-1}})}$$

<식 2-10>의 공분산은 다음과 같은 편향(bias)을 가진다.

$$p\lim_{N\to\infty}(\hat{\Theta}_{CV}-\Theta) = -\frac{1+\Theta}{T-1}\left(1-\frac{1}{T}\frac{1-\Theta^T}{1-\Theta}\right)\left[1-\frac{2\Theta}{(1-\Theta)(T-1)}\left(1-\frac{1-\Theta^T}{T(1-\Theta)}\right)\right]^{-1}.$$

이러한 편향이 발생하는 이유는 공분산추정량에서 개별 효과를 제거하는 과정에서

제거한 모형에서의 잔차와 설명변수가 상관관계를 가지기 때문이다. 또한 임의효과모형에서는 최소자승추정량은 모수를 과대추정한다(정확한 과대추정의 정도는 Hsiao, p.73를 참조하라).

동태모형의 추정에는 MLE(Maximum Likelihood Estimator), GLS(Generalized Least Squares Estimator), IV(Instrumental Variable Estimator), GMM(Generalized Method of Moments) 등이 사용된다.

□ 임의효과 모형의 추정

MLE, GLS, IV, GMM 등의 방법이 가능하다. MLE의 경우 최초 조건을 잘못 선택하면 일치성을 보장받지 못한다. IV형태의 추정량은 최초 조건으로부터 자유롭다. IV추정량의 기본 아이디어는 다음과 같다. <식 2-9>에서 다음을 얻을 수 있다.

$$\triangle y_{it} = y_{it} - y_{i,t-1} = \Theta \triangle y_{i,t-1} + \triangle \varepsilon_{it}, \text{ where } t=2,3 \ldots T$$

이 때 다음과 같은 적률조건(moment condition)을 얻을 수 있다.

$$E(y_{i,t-2} \triangle \varepsilon_{it}) = 0$$
$$E(y_{i,t-2} \triangle y_{i,t-1}) = 0$$

따라서 우리는 $y_{i,t-2}$. 를 IV로 이용하여 Θ를 추정할 수 있다. 또한 다음과 같은 적률조건도 구할 수 있다.

$$E(y_{i,t-2} \triangle \varepsilon_{it}) = 0$$
$$E(y_{i,t-3} \triangle \varepsilon_{it}) = 0$$
$$\vdots$$
$$E(y_{i,0} \triangle \varepsilon_{it}) = 0$$
$$E(\triangle y_{i,t-2} \triangle \varepsilon_{it}) = 0$$
$$E(\triangle y_{i,t-3} \triangle \varepsilon_{it}) = 0$$
$$\vdots$$

이러한 적률조건들을 이용하여 GMM추정을 할 수 있다. 독립적인 설명변수가 포함되어 있는 경우에는 다음과 같은 Arellano and Bond의 GMM 추정량을 사용할 수 있다.

단, $\delta=(\Theta,\beta')$.

$$y_{it}=\Theta y_{i,t-1}+x_{it}'\beta+\alpha_i+\varepsilon_{it}$$

$$\hat{\delta}_{GMM,AB}=\left\{\left[\sum_{i=1}^{N}\begin{pmatrix}\triangle y_{i,-1}\\ \triangle X_i'\end{pmatrix}W_i'\right]\left[\sum_{i=1}^{N}W_i\widetilde{A}W_i'\right]\left[\sum_{i=1}^{N}W_i(\triangle y_{i,-1},\triangle X_i)\right]\right\}^{-1}$$
$$\times\left\{\left[\sum_{i=1}^{N}\begin{bmatrix}\triangle y_{i,-1}\\ \triangle x_i'\end{bmatrix}W_i'\right]\left[\sum_{i=1}^{N}W_i\widetilde{A}W_i'\right]^{-1}\left[\sum_{i=1}^{N}W_i\triangle y_i\right]\right\}$$

단, $q_{it}=(y_{i,0},y_{i,1},\ldots y_{i,t-2},x_i')$, $x_i'=(x_{i1}',x_{i2}',\ldots x_{iT}')$

$$W_i=\begin{pmatrix}q_{i1}&0&\cdots&0\\0&q_{i2}&\cdots&0\\\vdots&\vdots&\vdots&\vdots\\0&0&\cdots&q_{iT}\end{pmatrix},\quad \widetilde{A}=\begin{pmatrix}2&-1&0&\cdots&0\\-1&2&-1&\cdots&0\\0&-1&2&\cdots&0\\\vdots&\vdots&\vdots&\vdots&\vdots\\0&0&0&\cdots&2\end{pmatrix}$$

GMM추정량은 (y_i,x_i')가 i에 대해 iid인 경우에는 α_i가 고정이던 확률변수이던 간에 관계없이 일치성을 가지며, 점근적 정규분포를 가진다.

□ 고정효과 모형의 추정

공분산추정량과 정규분포MLE는 일치성을 보장받지 못한다. GMM추정량은 (y_i,x_i')가 i에 대해 iid인 경우에는 α_i가 고정이던 확률변수이던 간에 관계없이 일치성을 가지며, 점근적 정규분포를 가진다. 이 경우에는 또한 변환우도비추정량(transformed likelihood estimator)을 사용할 수 있다.

Ⅲ. 농가경제조사 패널화 과정과 결과

1. 농가경제조사 자료의 특성과 패널구축방법

가. 농가경제조사 자료의 특성

경기 침체 장기화 및 원자재가격의 상승에 따른 물가상승 여파로 기업 및 가계경제는 많은 어려움을 겪고 있다. 농촌경제연구원이 최근 발표한 '농업·농촌경제동향'에 따르면 2008년 2분기의 농산물가격은 전분기 대비 전반적으로 하락한 반면, 비료 등의 농자재가격은 큰 폭으로 상승해 농가경제 또한 더욱 악화된 것으로 나타났다.

농가경제 악화는 농촌사회 뿐 아니라 사회전반에 큰 영향을 미칠 수 있으므로 정부차원의 농가경영 안정 대책 마련이 시급한 시점이다. 이러한 농가경제 안정화 대책을 위한 기초자료로서 제공될 수 있는 가장 좋은 자료는 '농가경제조사'이다. 농가경제조사는 농가경제 동향과 농업 경영실태를 파악하기 위한 목적으로 만들어진 조사로 현재 전국 320개 표본 조사구내 3,200 표본농가를 대상으로 행해지고 있다. 농가소득에서 농가자산 상황에 이르기까지 다양한 항목을 조사하는 만큼 농업정책 수립 및 농업경영 개선책 마련에 실증 자료로서 제공될 수 있다. 다음의 <표 3-1>은 농가경제조사의 개요를 간단하게 설명하고 있다.

<표 3-1> 농가경제조사 개요

조사목적		농가경제의 동향과 농업경영 실태를 파악하여 농업정책 수립과 농업경영개선을 위한 기초자료 제공
조사기간	일계부	매월 1일부터 말일까지 1개월 단위로 조사(조사대상기간 : 1년)
	원부	매년 연초 및 연말을 기준으로 조사(년 2회)
조사대상 가구		전국 320개 표본 조사구내 3,200 표본농가
조사방법		경영주에 의한 자기기입식 조사 원칙(예외, 기장능력이 없는 농가는 조사담당자에 의한 면접조사 방법에 따름.)
조사내용	일계부	작물재배현황, 도정량, 수입·지출, 농업노동 투입내역, 농산물생산비 투입내역, 자가농업생산물 중 자가소비량
	원부	가구원현황, 농가자산, 농가부채

현재 농가경제조사의 원자료 및 매년 발행되는 '농가경제통계' 보고서는 학계를 중심으로 농가경제 분석자료로서 활발히 이용되고 있다. 동일가구를 지속적으로 조사하는 방법으로 설계된 자료의 특성상 앞으로도 다양한 분석 및 활용이 가능할 것으로 보인다.

이러한 조사방법은 더 다양한 분석이 가능한 패널조사의 성격을 포함하고 있기 때문에 조사된 자료를 활용하여 패널데이타화한다면, 새로운 패널조사를 설계하는데 드는 많은 비용과 시간을 절감할 수 있을 것이다. 이러한 인식하에 본 연구에서는 농가경제조사 2003년부터 2007년까지 5개년의 패널구축을 시도해 보고자 하였다.

나. 패널구축 방법

□ 가구명부 작성

농가경제조사 최근 5개년간(2003~2007년)의 자료 3,200가구를 대상으로 패널화를 시도하였다. 먼저 농가의 기본정보가 들어있는 nonga파일에서 시도구분코드가 불분명함을 나타내는 '00'을 제거한 후 농가연계키(hid)를 생성하였다. hid는 시도(sido-2자리), 지구(jigu-3자리), 농가(nonga-3자리)를 이용하여 생성하되, 패널화 대상 자료 중 가장 최근 자료인 2007년 자료에서 농가(nonga)의 3자리 중 세 번째 자리수를 제외한 8자리 별도의 키를 만들어 가구번호의 중복횟수를 체크한다. -농가(nonga)는 3자리로 구성되지만 마지막 세 번째 자리의 경우 농가의 교체횟수를 나타내고 있어, 앞의 2자리만을 이용했다.- 이는 교체되었음에도 여전히 존재하고 있는 기존 가구를 제거하기 위한 작업이다. 중복 가구의 경우, 최종 교체가구를 제외한 가구 데이터를 삭제하면 2,589의 데이터가 남게 되는데, 이는 2003년부터 2007년까지 5개년간 바뀌지 않은 표본가구수를 의미한다.

2,589의 가구에 교체횟수만큼의 가구번호를 새로 생성하면 4,070의 가구가 최종 확인되고, 이렇게 만들어진 4,070의 가구 hid를 이용하여 2003년부터 2007년까지의 데이터 매칭을 시도한다. 다음으로 통계청 지역통계과에서 별도로 받은 가구주 표본대체사유 데이터를 클리닝 한 후[6], 만들어진 농가명부 파일에 매칭하게 된다.

가구명부를 작성한 파일은 다음의 <표 3-2>와 같은 형태로 작성되었다.

<표 3-2> 가구명부 파일의 구조

6) 지역통계과의 자료는 월별로 정리되어 있고 별도의 코드를 사용하지 않고 있으므로 클리닝 작업이 필요하다.

key					연도별 집계 및 교체 사유					전겸업여부					연도별 가중치					5년간 패널여부	연도별.시도별 농기본 자료
시도(2)	지구(3)	농가(3)			03	04	05	06	07	03	04	05	06	07	03	04	05	06	07		
11	010	0	1	0	0	0	0													1	
11	010	0	1	1				1	0											1	
11	010	0	2	0	0	0	0	0	0											0	
11	010	0	3	0	0																
11	010	0	3	1		2	0	0	0												

주 : (1) 농가번호의 셋째자리가 0이면 교체가 없었음을 뜻하고, 1이면 1번 교체, 2이면 2번 교체 등을 의미함.

(2) 연도별 집계에서 Y는 집계, N은 미집계를 의미함.

(3) 교체사유는 1 단독, 2 불능, 3불응, 4 비농가, 5 장기출타, 6 전출, 7 이상치, 8 기타를 의미함.

(4) 통계청 표본관리과의 표본대체 자료 활용하였음.

□ 가구원 명부 작성

현재 가구원의 표본이탈 사유 및 전입 사유는 조사되지 않고 있어 인구통계학적 가구원 분석은 한계가 있다. 또한 가구원의 경우 적합한 연계키를 만들기가 용이하지 않은 문제점을 내포하고 있다. 농가의 경우 명확한 연계키를 만들 수 있었지만, 가구원은 연계키의 역할을 할 대표 변수가 존재하지 않는다. 연계키를 찾기 위한 여러 시도 중 hid, 생년, 성별을 나열한 키가 가장 활용성이 높은 편이나 이 또한 완벽한 일치를 기대하기 어렵다. 생년의 경우, 호적 생년이 아닌 실제 생년을 기재하게 되어 있어 동일 가구원임에도 매칭이 안되는 경우가 일부 발생하기 때문이다.

이러한 문제점에도 hid, 생년, 성별을 연계키로 이용하고자 한다면, 별도의 클리닝 작업이 필요하고, 이 경우에도 데이터 이용에 있어서 한계가 있음을 인식해야 한다. 다음의 <표 3-3>은 가구원명부 파일의 구조를 보여주고 있다.

<표 3-3> 가구원 명부파일의 구조

key					1번가구원										2번가구원										비 고
시도(2)		지구(3)		농가(3)	03		04		05		06		07		03		04		05		06		07		
					여부	사유	여부	사유	여부	사유	여부	사유	여부	사유	여부	사유	여부	사유	여부	사유	여부	사유	여부	사유	

주 : (1) 가구원 명부 작성 방법은 농가번호 셋째 자리 0인 가구를 기준으로 하여, 생년월, 성별 기준으로 일치화를 점검함.

(2) 사유는 1 전입, 2 전출, 3 사망을 의미함.

(3) 가구원 현황에 대한 자료는 경영주와의 관계, 가구주, 성별, 교육정도, 농업종사여부, 노동능력, 주업상황(산업, 직업) 취식상태, 전입월, 전출월 등임.

2. 패널화 작업 결과

<표 3-4>는 패널화농가가구의 교체현황을 보여주고 있다. 5개년동안 교체된 경우가 없는 경우는 매년도 조사된 가구 3,200가구의 80.9%인 2,589가구였으며, 이 중에서 이상치인 47가구를 제외하면 최종적으로 5개년도 생존가구수는 2,542가구로 나타났다.

<표 3-4> 5년간(2003~2005) 패널화 농가 교체 현황(단위: 가구)

	전체	교 체 없음	교 체 있음	1번	2번	3번	4번	5번	패널	이상치
가 구 수	3,200	2,589	611	412	149	42	6	2	2,542	47
교체비율	100.0	80.9	19.1	12.9	4.7	1.3	0.2	0.1	79.4	1.5
(교체있음비율)			(100.0)	(67.4)	(24.4)	(6.9)	(1.0)	(0.3)		
교체가구수			870	412	298	126	24	10		

<표 3-5>는 교체사유별로 교체가구를 구분하여 제시하고 있다. 5년 합계를 기준으로 가장 많은 교체사유는 단독가구로의 전환이었으며 이러한 경우는 전체 변동가구수 870가구의 28.0%에 해당되는 244가구였다. 다음으로 비농가로의 전환이 전체변동 중에 23.2%를 차지하였으며, 불응의 경우는 전체의 12.0%로 나타나 상대적으로 매우 적었다. 불응이 경우 2004년

이후에는 지속적으로 동 경우가 감소하고 있는데 이는 조사에 참여하기 시작한 2년차 이후부터는 불응가구가 점차 감소한다는 일반적인 패널조사에서 나타나는 현상과 유사하다고 할 수 있다.

<표 3-5> 교체사유별 교체 현황 (단위:가구)

구분		2003	2004	2005	2006	2007	5년 합계	
조사 가구수(A)		3,200	3,200	3,200	3,200	3,200		
변동 가구수(B)		189	184	209	190	98	870	구성비
	변동율(%)	5.9	5.8	6.5	5.9	3.1	5.4	(%)
변동사유	단 독	51	59	55	47	32	244	28.0
	불 능	27	4	6	3	0	40	4.6
	불 응	26	32	28	13	5	104	12.0
	비 농 가	50	40	43	46	23	202	23.2
	장기출타	2	1	6	2	0	11	1.3
	전 출	29	24	26	21	9	109	12.5
	기 타	4	24	45	58	29	160	18.4
변동없는 가구수 (A-B)		3,011	3,016	2,991	3,010	3,102		

<표 3-6>은 가구원 명부 패널화 작업의 결과를 보여주고 있다. 전체 가구원 9,160명중에서 5년동안 변동이 없는 가구원은 6581명으로 나타나 전체 가구원의 71.8%로 나타났다. 동 비율이 가구의 경우에는 80.9%였음을 감안할 때 가구원의 변동은 가구의 변동보다는 많은 것으로 나타났다.

가구원변동의 사유로는 전체가구원인 경우가 1,713명으로서 변동이 있었던 가구원의 66.4%였으며, 나머지는 전입인 경우로 나타났다.

<표 3-6>가구원 명부 패널화 작업 결과

구분	전체	변동없음	변동있음	전출가구원	전입가구원
가 구 원	9,160	6,581	2,579	1,713	866
변동비율	100.0	71.8	28.2	18.7	9.5
(변동있음비율)			(100.0)	(66.4)	(33.6)

주 : 전출가구원에는 전출, 사망이 포함되며, 전입가구원에는 전입, 출생이 포함되어 있음.

3. 표본유지 및 표본 변동 분석

<표 3-7>은 표본유지율을 제시하고 있다. 인접한 2개년도를 기준으로 계속 표본유지율은 95~96%의 매우 높은 수준을 보여주었다. 예를 들어 2003년 응답가구 중에서 2004년에도 유지된 가구의 비율은 94.9%였으며, 2004년 응답가구 중에서 2005년에도 유지된 가구는 94.6%였다.

한편, 2003년에 응답한 가구는 시간이 갈수록 표본유지율이 떨어진다. 동 유지율은 2004년의 경우 94.9%에서 2005년에는 89.9%, 2006년에는 86.2% 그리고 2007년에는 83.6%를 기록하였다.

<표 3-7> 기본적인 응답률

2003	2004	2005	2006	2007	기준년도 가구수	가구수	응답률1	응답률2
Y						3,042		
Y	Y				3,042	2,888	94.94	
	Y	Y			3,056	2,891	94.60	
		Y	Y		3,048	2,912	95.54	
			Y	Y	3,121	2,979	96.91	
Y	Y				3,042	2,888	94.94	94.94
Y	Y	Y			2,888	2,736	94.74	89.94
Y	Y	Y	Y		2,736	2,622	95.83	86.19
Y	Y	Y	Y	Y	2,622	2,542	96.95	83.56

주 : (1) Y는 응답하였음을 의미함.

(2) 응답률1은 기준년도가구수 대비 당해 연도 응답률이며, 응답률2는 5년간 계속응답한 가구의 2003년 대비 응답률을 의미함.

농가경제조사의 패널유지율은 다른 패널조사 또는 다른 거처패널조사와 비교할 때 매우 높은 표본유지율로 평가된다. <표 3-8>에 나타난 각국의 주요 가구패널조사의 표본유지율을 보면 5차년도를 기준으로 미국 PSID는 81%, 독일 GSEP는 81%, BHPS는 75% 등으로 나타난다. 한국의 가구패널조사의 경우 대우패널의 5년 유지율은 60%, 노동패널의 경우에는 76%로 나타난다. 이와 같은 가구패널조사는 표본을 추적하여 조사한다는 점에서 거처패널보다는 더 높은 표본유지율을 예상할 수 있다. 그럼에도 불구하고 농가경제패널데이타의 5년 표본유

지율이 86%에 달한다는 점은 매우 놀라운 사실이라고 할 것이다.

<표 3-8> 주요 가구패널조사에서의 표본유지율

구분	PSID	GSEP	BHPS	KHPS	KLIPS
2차	89	90	88	79	88
3차	86	86	83	68	81
4차	84	85	79	62	77
5차	81	81	75	60	76
6차	79	79	74	-	77

주 : PSID는 미국가구패널, GSEP는 독일가구패널, BHPS는 영국가구패널, KHPS는 대우경제연구소의 가구패널, KLIPS는 한국노동연구원의 노동패널을 의미함.

다음의 <표 3-9>는 통계청의 경활패널과 도시가계패널의 표본유지율을 제시하고 있다. 경활패널의 경우 가구단위 5년 유지율은 46.5%, 개인단위 5년유지율은 36.7%를 기록하고 있다. 도시가계패널의 경우 가구단위 5년 유지율은 34.5%, 개인단위 5년유지율은 27.9%를 기록하고 있다. 통계청의 기타 거처패널과 비교할 때 농가경제조사는 매우 높은 표본유지율을 기록하고 있음을 확인할 수 있다.

<표 3-9> 경제활동인구조사와 도시가계조사에서의 표본유지율

구분			1998년 1월	1999년 1월	2000년 1월	2001년 1월	2002년 1월	2002년 12월
경활 표본	가구	패널규모	29,271	23,570	19,448	16,902	15,081	13,616
		유지율		80.5	66.4	57.7	51.5	46.5
	개인	패널규모	93,726	72,563	57,125	47,221	39,931	34,386
		유지율		77.4	60.9	50.4	42.6	36.7
도시가계 표본	가구	패널규모	4,612	3,572	2,764	2,233	1,868	1,590
		유지율		77.5	59.9	48.4	40.5	34.5
	개인	패널규모	17,258	12,769	9,506	7,360	5,906	4,818
		유지율		74.0	55.1	42.6	34.2	27.9

자료 : 이지연(2005), 가구조사자료의 종단화 방안, 통계청 통계연구보고서 05-03

4. 종단면 가중치 작성

종단면 가중치를 작성하기 위해 응답률 모형을 구축하였다. 세 가지 모형을 이용하였는데 모두 공통된 설명변수로는 경영주연령, 경영주 연령 제곱, 가구원수, 농업상시종사자, 경지면적, 농가소득, 농업소득, 농가자산, 농가부채, 농업소득의 가계비 충족도 등을 사용하였다. 이러한 모형을 기본모형이라고 하고, 이 기본모형에 전·겸업여부를 나타내는 더미변수를 추가한 모형과 전·겸업의 경우를 구분하여 각각 기본모형을 추정하는 모형 등의 3가지 모형을 고려하였다. 이 때 나타난 결과는 크게 상이하지는 않았으나, 본고에서는 편의상 세 번째 모형, 즉 전·겸업을 구분하고 각각의 경우에 기본모형을 추정하는 세 번째 모형을 사용한 결과를 제시한다[7]. 세 번째 모형에서 구체적인 응답률 추정 결과는 다음의 <표 3-10>과 같다.

<표 3-10> 응답률 추정모형

(1) 2004년

2004년	전업			1종 겸업			2종 겸업		
설명변수	계수	S.E.	z-값	계수	S.E.	z-값	계수	S.E.	z-값
mage2003	0.217	0.090	0.016	0.044	0.227	0.847	0.044	0.227	0.847
mage03s	-0.002	0.001	0.012	0.000	0.002	0.963	0.000	0.002	0.963
g012003	-0.123	0.130	0.342	0.272	0.226	0.230	0.272	0.226	0.230
g022003	0.815	0.232	0.000	0.414	0.381	0.277	0.414	0.381	0.277
g032003s	0.001	0.003	0.601	0.002	0.004	0.534	0.002	0.004	0.534
g052003s	0.010	0.010	0.331	0.007	0.015	0.642	0.007	0.015	0.642
g082003s	-0.029	0.017	0.079	-0.025	0.024	0.282	-0.025	0.024	0.282
g172003s	0.001	0.001	0.355	0.000	0.001	0.988	0.000	0.001	0.988
g182003s	0.000	0.003	0.864	-0.003	0.005	0.586	-0.003	0.005	0.586
g212003	0.008	0.003	0.008	0.008	0.006	0.178	0.008	0.006	0.178
_cons	-4.660	2.734	0.088	-0.983	6.406	0.878	-0.983	6.406	0.878

7) 각각의 경우에 총수를 조정하기 전 가중치의 평균과 분산은 다음과 같다. hwXX는 기본모형인 경우, hwXX_2는 기본모형에 전겸업더미변수를 추가한 경우, hwXX_3는 전겸업별로 구분하여 기본모형을 적용한 경우를 의미한다.

	관측치수	평균	표준편차	최소값	최대값
hw04	2,886	379.21	232.01	1.09	1,431.30
hw05	2,733	399.08	246.32	1.12	1,720.76
hw06	2,618	415.51	256.19	1.14	1,929.67
hw07	2,542	428.81	267.83	1.16	2,024.43
hw04_2	2,886	379.12	231.74	1.10	1,431.07
hw05_2	2,733	399.02	246.11	1.12	1,743.86
hw06_2	2,618	415.48	256.09	1.13	1,992.55
hw07_2	2,542	428.85	268.15	1.14	2,164.50
hw04_3	2,886	379.27	232.43	1.09	1,526.99
hw05_3	2,733	398.94	246.97	1.09	1,840.30
hw06_3	2,618	415.22	256.12	1.12	1,484.71
hw07_3	2,542	428.47	267.92	1.13	1,839.16

관측치수	1,733	554	753

(2) 2005년

2005년	전업			1종겸업			2종겸업		
설명변수	계수	S.E.	z-값	계수	S.E.	z-값	계수	S.E.	z-값
mage2004	0.214	0.115	0.063	0.475	0.149	0.001	0.092	0.159	0.563
mage04s	−0.002	0.001	0.089	−0.004	0.001	0.001	−0.001	0.001	0.576
g012004	0.222	0.186	0.231	0.352	0.228	0.124	0.178	0.145	0.219
g022004	0.380	0.276	0.169	−0.308	0.401	0.443	0.659	0.232	0.004
g032004s	0.011	0.005	0.017	0.001	0.003	0.855	0.006	0.009	0.459
g052004s	0.005	0.010	0.616	0.004	0.014	0.775	0.001	0.008	0.951
g082004s	0.011	0.017	0.539	0.004	0.024	0.866	0.073	0.048	0.129
g172004s	−0.001	0.000	0.000	−0.001	0.000	0.022	0.001	0.001	0.193
g182004s	−0.001	0.002	0.622	−0.005	0.004	0.235	0.011	0.007	0.138
g212004	0.001	0.003	0.765	−0.001	0.006	0.874	−0.002	0.010	0.850
_cons	−5.862	3.678	0.111	−9.200	4.386	0.036	−2.082	4.578	0.649
관측치수			1,388			739			760

(3) 2006년

2006년	전업			1종 겸업			2종 겸업		
설명변수	계수	S.E.	z-값	계수	S.E.	z-값	계수	S.E.	z-값
mage2005	0.197	0.124	0.112	0.311	0.216	0.150	0.256	0.152	0.093
mage05s	−0.002	0.001	0.041	−0.003	0.002	0.155	−0.002	0.001	0.092
g012005	−0.113	0.178	0.524	−0.080	0.235	0.733	0.058	0.153	0.706
g022005	0.920	0.283	0.001	0.915	0.493	0.063	0.336	0.243	0.167
g032005s	−0.001	0.002	0.657	−0.002	0.005	0.741	0.004	0.008	0.649
g052005s	−0.008	0.008	0.324	−0.008	0.021	0.704	0.023	0.012	0.052
g082005s	−0.004	0.014	0.794	0.022	0.035	0.527	−0.017	0.057	0.773
g172005s	0.000	0.001	0.797	0.002	0.002	0.384	−0.001	0.000	0.135
g182005s	−0.003	0.003	0.303	0.000	0.006	0.995	0.002	0.005	0.750
g212005	0.005	0.003	0.137	−0.002	0.007	0.820	0.009	0.015	0.530
_cons	−2.717	4.094	0.507	−6.603	6.288	0.294	−5.537	0.000	0.213
관측치수			1,363			663			709

(4) 2007년

2007년	전업			1종겸업			2종겸업		
설명변수	계수	S.E.	z-값	계수	S.E.	z-값	계수	S.E.	z-값
mage2006	0.146	0.174	0.403	0.523	0.301	0.082	0.134	0.232	0.562
mage06s	−0.001	0.001	0.363	−0.005	0.002	0.056	−0.001	0.002	0.473
g012006	0.203	0.255	0.426	0.149	0.428	0.727	−0.182	0.176	0.302
g022006	0.904	0.315	0.004	2.034	0.549	0.000	0.516	0.353	0.144
g032006s	0.016	0.007	0.013	0.007	0.009	0.431	−0.004	0.011	0.682

g052006s	−0.008	0.011	0.461	−0.023	0.023	0.325	0.005	0.010	0.605
g082006s	0.011	0.017	0.512	0.053	0.048	0.268	−0.059	0.056	0.289
g172006s	0.001	0.001	0.513	0.005	0.003	0.116	−0.001	0.000	0.088
g182006s	−0.003	0.003	0.333	−0.017	0.007	0.017	0.012	0.010	0.231
g212006	−0.005	0.003	0.060	−0.005	0.006	0.416	0.033	0.014	0.016
_cons	−3.449	5.885	0.558	−13.545	9.270	0.144	0.599	6.974	0.932
관측치수			1,225			718			674

주 : 변수명은 다음과 같음. mage는 경영주 연령, mages는 경영주 연령의 제곱, g01은 가구원수, g02는 농업상시종사자수, g03는 경지면적, g05는 농가소득, g08은 농업소득, g17은 농가자산, g18은 농가부채, g21는 농업소득의 가계비 충족률을 각각 의미함.

이제 이와 같이 추정된 응답확률을 이용하여 종단면가중치를 구하였다. 먼저 2003년의 종단면가중치는 2003년의 횡단면가중치를 그대로 사용하였다. 2004년의 경우에는 2003년에 응답하고 2004년에도 응답한 가구에 대하여 위에서 추정된 응답률의 역수를 곱하여 2004년의 종단면가중치를 계산하였다. 이후 연도의 경우에도 동일하게 가중치를 작성하였다. 이렇게 작성된 가중치를 종단면 기본가중치라고 명명한다.

매년도의 종단면기본 가중치에 대하여 지구별, 전·겸업별 종단면 기본 가중치의 합을 구하고, 이를 지구별, 전·겸업별 횡단면가중치의 합과 일치하도록 비율 조정을 실시하였다[8]. 이때 각 셀에서 표본수가 20개 미만인 경우에는 셀을 지구별로 결합하여 하나의 셀을 만든 다음 비율조정을 실시하였다. 비율조정을 실시하였을 때의 비율과 최종적으로 구해진 가중치의 기초 통계량은 다음과 같다. <표 3-11>에 의하면 2004년의 경우 평균적으로는 종단면기본가중치에 0.9804를 곱하여 종단면 최종가중치를 계산하였다. 이러한 비율조정 값은 최소값이 0.5832였고, 최대값이 1.5330으로 나타났다. 다른 년도의 경우에도 동일하게 해석할 수 있다.

<표 3-12>는 이렇게 계산된 종단면 최종가중치(이하 종단면가중치)의 기초통계량을 제시하고 있다. 2004년의 경우 2,886개의 표본가구에 대하여 적용된 가중치는 평균 372가구였으며, 최소값은 0.84, 최대값은 1,272.17로 나타났다[9]. 종단면 가중치의 평균은 해가 거듭할수록 커지게 되는데 이는 웨이브가 지남에 따라 표본탈락이 발생하는 반면 표본이 모집단으로 사용하고 있는 농가규모는 약 107만 가구로 변동이 거의 없기 때문이다. 한편, 가중치의 표준편차도 매년 커지게 되는 데, 이는 응답률의 변동이 가중치에 지속적으로 반영되기 때문이다.

8) 전업과 1종 겸업, 2종 겸업 등 세 가지로 구분하여 보기도 하였으나, 이렇게 구분되는 경우 매우 많은 셀에서 20가구 미만이 나타났다.

9) 통상적인 가중치의 변동보다는 매우 큰 변동이라고 할 수 있다. 이러한 가중치의 변동은 통계청에서 사용하는 기본가중치의 변동폭이 동일 년도의 경우 0.61에서 1,319.53에 달한다는 점에 기인하는 부분이 있다고 하겠다.

<표 3-11> 비율조정변수의 기초통계량

변수	관측치수	평균	표준편차	최소값	최대값
p2004	2,888	0.9804	0.2437	0.5832	1.5330
p2005	2,736	0.9845	0.2139	0.5628	1.4783
p2006	2,622	0.9728	0.2480	0.5679	1.5437
p2007	2,542	0.9744	0.1922	0.6204	1.5066

주 : 횡단면가중치/종단면가중치. pXXXX는 해당년도의 비율조정계수

<표 3-12> 가중치의 기초 통계량

구분	관측치수	평균	표준편차	최소값	최대값	관측치합
hw04_f	2,886	372.09	247.25	0.84	1272.17	1,073,846
hw05_f	2,733	392.92	255.61	0.90	1450.71	1,073,850
hw06_f	2,618	406.12	273.84	0.80	1590.50	1,063,229
hw07_f	2,542	419.09	275.32	0.88	1818.56	1,065,336
w04	3,056	351.39	245.50	0.61	1319.53	1,073,849
w05	3,048	352.31	240.06	0.76	1443.84	1,073,848
w06	3,074	345.88	250.61	0.90	2040.65	1,063,233
w07	3,121	341.34	236.98	0.69	1493.72	1,065,336

<표 3-13>은 가중치를 적용하여 계산한 기본변수의 요약통계량을 제시하고 있다. 먼저 가구원수의 경우에는 종단면가중치를 이용하는 경우와 횡단면가중치를 이용하는 경우에 약 1~2% 정도의 차이가 발생하는 것으로 나타났다. 농업상시종사자의 경우에도 1~2%의 차이가 발생하여 차이가 작게 나타났다. 경지면적의 경우에는 종단면가중치를 사용하는 경우에도 횡단면가중치를 사용하는 경우보다 2~3%가 많은 것으로 나타났다.

농가소득의 경우에는 종단면가중치를 사용하는 경우 횡단면가중치를 사용하는 경우에도 비해 1~4%정도 적게 나타났다. 이러한 차이는 시간이 지남에 따라 점점 커지는 것으로 나타나는데 2007년의 경우에는 종단면가중치를 이용한 추정치가 횡단면가중치를 사용하는 경우에 비해 96.07%수준으로 나타났다. 반면 농업소득의 경우에는 반대로 종단면가중치를 사용하는 경우가 오히려 4~5%정도 많은 것으로 나타났다. 농가자산의 경우에는 종단면가중치를 사용하는 경우가 적게 나타나는데 특히 2006년과 2007년은 양자의 차이가 8~9%에 달하였다. 농가부채의 경우에는 양자가 거의 차이가 없었다. 농업소득의존도는 종단면가중치를 사용하는

경우가 약 3~9%까지 높게 나타났다. 이는 종단면가중치를 사용하는 경우 농업소득은 상대적으로 크게 나타나고 농가소득은 상대적으로 작게 나타나는 데에 기인한다고 하겠다.

이러한 차이가 발생하지만 양자의 방향성은 대체로 일치한다. 평균증가율의 부호가 반대로 나타난 경우는 농가소득의 2006년과 농가자산의 2005년 두 가지 경우 뿐이다[10].

<표 3-13> 가중치를 이용한 통계량 계산결과

구분		횡단면			종단면		
		표본수	평균	표준편차	표본수	평균	표준편차
가구원수	2004년	3,055	2.82	1.20	2,885	2.79	1.18
	2005년	3,047	2.85	1.23	2,732	2.82	1.21
	2006년	3,068	2.83	1.23	2,613	2.79	1.18
	2007년	3,121	2.83	1.23	2,542	2.82	1.20
농업상시종사자	2004년	3,055	1.78	0.68	2,885	1.81	0.66
	2005년	3,047	1.80	0.68	2,732	1.81	0.68
	2006년	3,068	1.78	0.68	2,613	1.79	0.68
	2007년	3,121	1.78	0.68	2,542	1.80	0.67
경지면적	2004년	3,055	4,833	5,516	2,885	4,973	5,592
	2005년	3,047	4,764	5,245	2,732	4,903	5,406
	2006년	3,068	4,855	5,920	2,613	5,010	5,946
	2007년	3,121	4,918	5,916	2,542	5,089	5,886
농가소득	2004년	3,055	29,029,169	25,778,894	2,885	28,931,949	26,285,831
	2005년	3,047	30,315,177	29,518,404	2,732	29,961,518	30,625,904
	2006년	3,068	30,417,933	28,444,889	2,613	29,358,638	26,074,496
	2007년	3,121	31,617,025	30,466,325	2,542	30,374,407	24,566,513
농업소득	2004년	3,055	11,299,795	19,135,204	2,885	11,800,816	19,885,732
	2005년	3,047	11,576,953	22,786,302	2,732	12,176,300	24,190,078
	2006년	3,068	11,348,579	19,054,113	2,613	11,802,365	18,826,853
	2007년	3,121	11,642,276	19,402,520	2,542	12,144,383	18,372,468
농가자산	2004년	3,055	289,577,937	340,148,375	2,885	285,880,910	331,727,361
	2005년	3,047	290,951,098	322,944,829	2,732	276,970,261	288,880,166
	2006년	3,068	310,289,318	379,541,221	2,613	283,034,940	305,001,409
	2007년	3,121	312,224,832	395,956,698	2,542	284,955,235	322,384,976
농가부채	2004년	3,055	27,543,832	56,424,048	2,885	27,461,037	54,231,037
	2005년	3,047	26,867,786	54,715,722	2,732	27,152,749	53,753,122
	2006년	3,068	27,935,469	57,966,025	2,613	28,063,233	55,899,861
	2007년	3,121	28,518,186	56,395,333	2,542	28,494,272	54,897,345
농업소득의존도	2004년	3,055	34.0	235.3	2,885	37.1	221.9
	2005년	3,047	36.5	97.5	2,732	38.9	104.8
	2006년	3,068	50.2	926.1	2,613	53.0	906.1
	2007년	3,121	36.7	116.5	2,542	38.1	86.8

10) 한편, 미국 PSID의 경우 가구소득변수에서 종단면가중치를 사용한 결과와 CPS의 횡단면가중치를 사용한 결과의 평균비율은 0.96~1.03으로 나타난다(Gouskova et al, 2008).

		종단면/횡단면		증가율	
		평균비율	표준편차비율	횡단면	종단면
가구원수	2004년	98.94	98.33		
	2005년	98.95	98.37	1.06	1.08
	2006년	98.59	95.93	−0.70	−1.06
	2007년	99.65	97.56	0.00	1.08
농업상시종사자	2004년	101.69	97.06		
	2005년	100.56	100.00	1.12	0.00
	2006년	100.56	100.00	−1.11	−1.10
	2007년	101.12	98.53	0.00	0.56
경지면적	2004년	102.91	101.38		
	2005년	102.92	103.06	−1.42	−1.41
	2006년	103.19	100.43	1.91	2.17
	2007년	103.49	99.49	1.29	1.58
농가소득	2004년	99.67	101.97		
	2005년	98.83	103.75	4.43	3.56
	2006년	96.52	91.67	0.34	−2.01
	2007년	96.07	80.63	3.94	3.46
농업소득	2004년	104.43	103.92		
	2005년	105.18	106.16	2.45	3.18
	2006년	104.00	98.81	−1.97	−3.07
	2007년	104.31	94.69	2.59	2.90
농가자산	2004년	98.72	97.52		
	2005년	95.19	89.45	0.47	−3.12
	2006년	91.22	80.36	6.65	2.19
	2007년	91.27	81.42	0.62	0.68
농가부채	2004년	99.70	96.11		
	2005년	101.06	98.24	−2.45	−1.12
	2006년	100.46	96.44	3.97	3.35
	2007년	99.92	97.34	2.09	1.54
농업소득의 존도	2004년	109.12	94.31		
	2005년	106.58	107.49	7.35	4.85
	2006년	105.58	97.84	37.53	36.25
	2007년	103.81	74.51	−26.89	−28.11

이러한 차이가 얼마나 큰 차이인가를 알아보기 위해 농가소득과 농업소득에 대하여 각각 횡단면가중치를 사용하기 전후, 종단면가중치를 사용하기 전후의 평균값의 변화를 계산해 보았다.

먼저 농업소득의 경우 횡단면가중치를 사용하면 가중치를 사용하지 않는 경우에 비해 약 2.5~5.8%정도가 크게 나타났다. 종단면가중치를 사용하면 가중치를 사용하지 않는 경우에 비

해 약 4.8~6.9%정도가 크게 나타났다. 종단면가중치를 사용하는 경우 가중치를 사용하지 않는 경우에 비해 평균증가의 크기가 크게 나타났다.

농가소득의 경우 횡단면가중치를 사용하면 가중치를 사용하지 않는 경우에 비해 약 4.4~~6.5%정도가 크게 나타났다. 종단면가중치를 사용하면 가중치를 사용하지 않는 경우에 비해 약 1.5~6.7%정도가 크게 나타났다. 종단면가중치를 사용하는 경우 가중치를 사용하지 않는 경우에 비해 평균증가의 크기가 크게 나타났다.

<표 3-14> 가중치 사용 전후의 변화분석

	무가중치	관측치수	평균	표준편차	최소값	최대값	횡단면/종단면
횡단면	g052004	3,055	30,720,825	27,782,970	-180,000,000	346,682,080	
	g052005	3,047	31,217,233	32,226,339	-366,000,000	971,341,660	
	g052006	3,068	31,550,593	29,344,811	-81,272,071	469,984,940	
	g052007	3,121	32,420,229	35,610,244	-101,000,000	1,220,000,000	
종단면	g052004	2,887	30,730,312	27,796,986	-180,000,000	346,682,080	99.97
	g052005	2,735	31,394,430	32,002,476	-240,000,000	971,341,660	99.44
	g052006	2,617	31,374,123	28,637,739	-81,272,071	469,984,940	100.56
	g052007	2,542	31,817,176	25,831,653	-101,000,000	251,575,188	101.90
횡단면	g082004	3,055	11,912,533	20,429,456	-181,000,000	302,780,427	
	g082005	3,047	12,086,805	25,409,240	-394,000,000	779,317,430	
	g082006	3,068	12,032,385	19,820,349	-122,000,000	206,873,510	
	g082007	3,121	12,397,464	21,720,743	-111,000,000	542,083,281	
종단면	g082004	2,887	11,979,475	20,448,052	-181,000,000	302,780,427	99.44
	g082005	2,735	12,418,446	24,944,959	-259,000,000	779,317,430	97.33
	g082006	2,617	12,289,275	19,903,051	-122,000,000	206,873,510	97.91
	g082007	2,542	12,954,633	19,410,685	-103,000,000	221,784,395	95.70

	가중치	관측치수	평균	표준편차	최소값	최대값	
횡단면	g052004	3,055	29,029,169	25,778,894	-180,000,000	346,682,080	
	g052005	3,047	30,315,177	29,518,404	-366,000,000	971,341,660	
	g052006	3,068	30,417,933	28,444,889	-81,272,071	469,984,940	
	g052007	3,121	31,617,025	30,466,325	-101,000,000	1,220,000,000	
종단	g052004	2,885	28,931,949	26,285,831	-180,000,000	346,682,080	100.34
	g052005	2,732	29,961,518	30,625,904	-240,000,000	971,341,660	101.18

면	g052006	2,613	29,358,638	26,074,496	−81,272,071	469,984,940	103.61
	g052007	2,542	30,374,407	24,566,513	−101,000,000	251,575,188	104.09
횡	g082004	3,055	11,299,795	19,135,204	−181,000,000	302,780,427	
단	g082005	3,047	11,576,953	22,786,302	−394,000,000	779,317,430	
면	g082006	3,068	11,348,579	19,054,113	−122,000,000	206,873,510	
	g082007	3,121	11,642,276	19,402,520	−111,000,000	542,083,281	
종	g082004	2,885	11,800,816	19,885,732	−181,000,000	302,780,427	95.75
단	g082005	2,732	12,176,300	24,190,078	−259,000,000	779,317,430	95.08
면	g082006	2,613	11,802,365	18,826,853	−122,000,000	206,873,510	96.16
	g082007	2,542	12,144,383	18,372,468	−103,000,000	221,784,395	95.87

비율		관측치수	평균	표준편차
횡단면	g052004	3,055	105.83	107.77
	g052005	3,047	102.98	109.17
	g052006	3,068	103.72	103.16
	g052007	3,121	102.54	116.88
종단면	g052004	2,885	106.22	105.75
	g052005	2,732	104.78	104.49
	g052006	2,613	106.87	109.83
	g052007	2,542	104.75	105.15
횡단면	g082004	3,055	105.42	106.76
	g082005	3,047	104.40	111.51
	g082006	3,068	106.03	104.02
	g082007	3,121	106.49	111.95
종단면	g082004	2,885	101.51	102.83
	g082005	2,732	101.99	103.12
	g082006	2,613	104.13	105.72
	g082007	2,542	106.67	105.65

주 : g05XXXX, g08XXXX는 각각 XXXX년의 농가소득과 농업소득을 의미함.

Ⅳ. 농가경제조사패널데이타를 이용한 분석결과

1. 소득이동 결정요인 분석

농가경제활동조사의 패널화자료를 통해 농가 소득의 이동경로를 분석해 본다. 먼저 다음의 <표 4-1>은 농가소득과 농업소득을 5분위로 분석한 결과를 제시하고 있다. 농가소득의 5분위 분배율(상위 20%소득/하위20%소득)은 2003년 9.41, 2004년 11.17, 2005년 12.53을 기록하여 지속적으로 상승하였으나, 이후 2006년 10.18, 2007년 8.69를 기록하여 점차 하락하고 있는 것으로 나타났다.

한편, 농업소득의 경우에는 전 기간에 걸쳐서 하위 20%의 평균소득은 마이너스로 나타났으며, 특히 2005년과 2006년에는 1분위와 2분위의 평균소득도 마이너스를 기록하였다.

<표 4-1> 소득이동 관련 기본 통계량

(1) 가구소득

구분		1	2	3	4	5	전체
2003	평균	7,446,676	16,267,992	23,665,905	35,685,602	70,066,093	30,595,595
	S.E.	6,668,373	1,966,078	2,588,398	4,447,890	38,276,066	27,941,298
2004	평균	5,988,553	15,492,719	23,791,029	34,927,203	66,878,302	29,379,920
	S.E.	10,975,155	2,195,449	2,569,392	4,260,208	30,383,716	25,579,440
2005	평균	5,659,459	16,010,675	24,340,922	36,108,262	70,937,524	30,590,731
	S.E.	18,643,403	2,019,844	2,816,402	4,122,869	48,290,076	32,364,119
2006	평균	6,740,809	15,687,142	23,274,843	34,415,825	68,635,179	29,720,334
	S.E.	6,757,422	2,076,163	2,665,454	3,853,399	32,336,933	26,114,158
2007	평균	7,826,689	16,255,244	23,928,112	35,953,355	67,982,510	30,374,562
	S.E.	6,369,973	1,927,808	2,772,870	4,353,087	27,443,952	24,566,988

(2) 농업소득

구분		1	2	3	4	5	전체
2003	평균	-2,051,200	3,079,775	6,834,922	13,633,703	38,998,851	12,095,105
	S.E.	7,083,659	890,636	1,369,096	2,656,003	26,510,648	18,957,743
2004	평균	-3,338,067	2,854,917	6,617,757	13,688,148	37,345,098	11,426,837
	S.E.	12,027,859	872,057	1,366,557	2,689,044	25,860,355	19,047,458
2005	평균	-4,070,408	3,147,048	6,966,669	13,606,983	41,157,385	12,153,082

05	S.E.	19,549,339	945,694	1,242,655	2,767,108	41,290,424	25,729,161
20	평균	−2,795,706	2,864,468	6,528,244	13,023,363	38,257,271	11,565,332
06	S.E.	7,437,425	857,066	1,365,532	2,530,540	24,791,979	18,438,113
20	평균	−1,985,099	3,217,138	6,960,761	13,574,250	39,005,746	12,144,434
07	S.E.	7,060,901	901,982	1,440,126	2,825,209	24,487,689	18,373,461

농가소득 기준으로 각 분위별 이동 상황을 제시한 표가 <표 4-2>이다. 동 표에 의하면 2003년에 1분위이면서 계속적으로 1분위에 속하는 가구는 2004년에 55.5%, 2005년에 55.3%, 2006년 50.1%, 2007년 53.1%로 나타났다. 2003년에 1분위에 속하는 가구 중 1분위로 이동한 경우는 2004년 5.8%, 2005년 3.8%, 2006년 6.6%, 2007년 4.5% 등으로 나타났다.

2003년에 5분위이면서 동 분위를 유지한 가구의 비율은 2004년 52.7%, 2005년 53.6%, 2006년 52.3%, 2007년 54.7%로서 연도별로 거의 변화가 없었다. 2003년에 3분위이면서 동 분위를 유지한 가구의 비율은 2004년 36.3%, 2005년 33.1%, 2006년 34.8%, 2007년 35.3%로 나타나 역시 큰 변화가 없었다. 2003년에 3분위이면서 2분위이하로 떨어진 비율은 2004년 31.6%, 2005년 31.5%, 2006년 31.1%, 2007년 34.8% 등으로 나타났다.

이상과 같은 분석결과는 1분위와 5분위에 속하는 가구의 경우 소득 이동이 상대적으로 매우 적어서 소득계층이 고착화되는 경향이 있음을 시사하고 있다. 이에 반해 3분위에 속하는 가구의 경우 상위 분위로 이동하거나 하위 분위로 이동하는 경우가 거의 유사하게 나타나 소득이동이 대칭적이었다.

농가소득 뿐만 아니라 농업소득의 경우에도 유사한 패턴이 나타났다. 2003년에 농업소득이 하위 20%에 속하는 가구가 계속 동일 분위에 남아있는 경우는 2004년 55.2%, 2005년 53.8%, 2006년 53.2%, 2007년 54.7%로 나타나 거의 연도별로 변화가 없었다. 또한 동일 분위에 남아있는 비율은 농가소득의 경우와 큰 차이가 없었다. 2003년에 농업소득이 하위 20%에 속하는 가구가 상위 20% 계층으로 이동하는 비중은 2004년 5.4%, 2005년 5.8%, 2006년 7.3%, 2007년 7.4%로 나타나 해가 거듭될 수록 소폭 비중이 증가하는 것으로 나타났다.

2003년에 농업소득이 상위 20%에 속하는 가구의 경우 동일 분위에 남아있는 비중은 2004년 65.1%, 2005년 66.3%, 2006년 63.4%, 2007년 64.7%로 나타나 년도에 따라서 거의 변화가 없었다. 동일한 분위에 속하는 비중의 차이가 크지 않다는 점은 농가소득과 유사하였으나, 동일 분위에 남아 있는 비중은 농가소득에 비해 농업소득이 크게 나타났다. 2003년에

농업소득이 3분위에 속하는 가구의 경우 동일 분위에 남아있는 비중은 2004년 42.1%, 2005년 43.9%, 2006년 42.3%, 2007년 43.2%로 나타나 년도에 따라서 거의 변화가 없었다. 다만 이러한 비중은 농가소득에 비해 상대적으로 높은 비중이었다. 2003년에 농업소득이 3분위에 속하는 가구의 경우 1, 2분위로 떨어지는 비중은 2004년 29.7%, 2005년 32.1%, 2006년 28.3%, 2007년 30.3%로 나타나 연도별로 큰 차이가 나타나지는 않았다. 그러나 동 분위에서 4,5분위로 올라 가는 비중은 모두 1, 2분위로 떨어지는 비중보다 낮은 것으로 나타났다.

이와 같은 분석결과를 종합하면 농가소득의 경우 1분위와 5분위는 약 50~55%% 정도가 5년 후에도 동일분위에 있으며, 농업소득의 경우 1분위는 약 50~55%%, 5분위는 65%내외가 동일 분위에 머무르는 것으로 나타났다. 농업소득이나 농가소득의 고착화현상이 있음을 암시하는 대목이다. 이에 반해 3분위에 속하는 경우에는 비교적 소득이동이 활발하게 일어나고 있는데, 대체로 농가소득의 경우에는 동일분위에 머무르는 비중, 상위 분위로 이동하는 비중, 하위 분위로 이동하는 비중이 유사하게 나타났다. 농업소득의 경우에는 농가소득에 비해 상대적으로 3분위에 머무는 비중이 높은 것으로 나타났으며, 상위 계층으로 이동하기 보다는 하위 계층으로 이동하는 비중이 높은 것으로 나타났다.

<표 4-2> 농가소득이동 분석(단위 : 가구, %)

(1) 2003->2004년

구분	1	2	3	4	5	전체
1	118,444	47,255	24,707	10,640	12,376	213,422
	55.5	22.14	11.58	4.99	5.8	100
2	47,126	93,704	46,529	16,684	9,196	213,239
	22.1	43.94	21.82	7.82	4.31	100
3	25,490	40,822	77,201	44,828	24,492	212,833
	11.98	19.18	36.27	21.06	11.51	100
4	13,055	18,644	43,111	84,566	54,045	213,421
	6.12	8.74	20.2	39.62	25.32	100
5	9,455	12,216	21,601	57,151	111,995	212,418
	4.45	5.75	10.17	26.9	52.72	100
전체	213,570	212,641	213,149	213,869	212,104	1,065,333
	20.05	19.96	20.01	20.08	19.91	100

(2) 2003->2005년

구분	1	2	3	4	5	전체
1	117,935	50,522	22,193	14,566	8,206	213,422
	55.26	23.67	10.4	6.82	3.84	100
2	49,651	85,177	46,975	21,367	10,069	213,239
	23.28	39.94	22.03	10.02	4.72	100
3	21,539	45,839	70,454	49,781	25,220	212,833
	10.12	21.54	33.1	23.39	11.85	100
4	10,234	22,683	52,024	72,859	55,621	213,421
	4.8	10.63	24.38	34.14	26.06	100
5	14,337	9,284	20,454	54,547	113,796	212,418
	6.75	4.37	9.63	25.68	53.57	100
전체	213,696	213,505	212,100	213,120	212,912	1,065,333
	20.06	20.04	19.91	20.01	19.99	100

(3) 2003->2006년

구분	1	2	3	4	5	전체
1	108,372	58,074	20,327	12,655	13,994	213,422
	50.78	27.21	9.52	5.93	6.56	100
2	50,706	82,623	44,165	27,225	8,520	213,239
	23.78	38.75	20.71	12.77	4	100
3	26,021	40,249	73,960	50,390	22,213	212,833
	12.23	18.91	34.75	23.68	10.44	100
4	13,042	20,425	52,740	72,035	55,179	213,421
	6.11	9.57	24.71	33.75	25.85	100
5	14,982	11,660	22,042	51,440	112,294	212,418
	7.05	5.49	10.38	24.22	52.86	100
전체	213,123	213,031	213,234	213,745	212,200	1,065,333
	20.01	20	20.02	20.06	19.92	100

(4) 2003년->2007년

구분	1	2	3	4	5	전체
1	113,414	51,170	22,528	16,790	9,520	213,422
	53.14	23.98	10.56	7.87	4.46	100
2	51,544	77,169	46,880	25,209	12,437	213,239
	24.17	36.19	21.98	11.82	5.83	100
3	27,743	46,469	75,201	39,803	23,617	212,833
	13.04	21.83	35.33	18.7	11.1	100

4	14,070	27,413	45,991	74,881	51,066	213,421
	6.59	12.84	21.55	35.09	23.93	100
5	6,579	11,129	22,167	56,292	116,251	212,418
	3.1	5.24	10.44	26.5	54.73	100
전체	213,350	213,350	212,767	212,975	212,891	1,065,333
	20.03	20.03	19.97	19.99	19.98	100

<표 4-3>농업소득이동 분석 (단위 : 가구, %)

(1) 2003->2004년

구분	1	2	3	4	5	전체
1	117,688	50,020	20,654	13,255	11,523	213,140
	55.22	23.47	9.69	6.22	5.41	100
2	50,767	100,048	43,184	13,750	5,432	213,181
	23.81	46.93	20.26	6.45	2.55	100
3	14,764	48,440	89,737	49,180	10,986	213,107
	6.93	22.73	42.11	23.08	5.16	100
4	13,935	12,373	47,382	92,732	46,485	212,907
	6.55	5.81	22.25	43.56	21.83	100
5	15,984	3,180	11,266	43,976	138,592	212,998
	7.5	1.49	5.29	20.65	65.07	100
전체	213,138	214,061	212,223	212,893	213,018	1,065,333
	20.01	20.09	19.92	19.98	20	100

(2) 2003->2005년

구분	1	2	3	4	5	전체
1	114,554	45,564	27,128	13,572	12,322	213,140
	53.75	21.38	12.73	6.37	5.78	100
2	51,422	101,731	40,621	15,130	4,277	213,181
	24.12	47.72	19.05	7.1	2.01	100
3	19,279	49,112	93,450	39,313	11,953	213,107
	9.05	23.05	43.85	18.45	5.61	100
4	11,193	9,667	44,408	104,539	43,100	212,907
	5.26	4.54	20.86	49.1	20.24	100
5	17,297	6,420	7,565	40,385	141,331	212,998
	8.12	3.01	3.55	18.96	66.35	100
전체	213,745	212,494	213,172	212,939	212,983	1,065,333

	20.06	19.95	20.01	19.99	19.99	100

(3) 2003->2006년

구분	1	2	3	4	5	전체
1	113,291	55,494	18,502	10,237	15,616	213,140
	53.15	26.04	8.68	4.8	7.33	100
2	47,394	103,545	43,395	13,637	5,210	213,181
	22.23	48.57	20.36	6.4	2.44	100
3	22,347	37,913	91,588	46,524	14,735	213,107
	10.49	17.79	42.98	21.83	6.91	100
4	15,040	11,104	47,866	97,711	41,186	212,907
	7.06	5.22	22.48	45.89	19.34	100
5	15,340	4,838	12,251	44,395	136,174	212,998
	7.2	2.27	5.75	20.84	63.93	100
전체	213,412	212,894	213,602	212,504	212,921	1,065,333
	20.03	19.98	20.05	19.95	19.99	100

(4) 2003->2007년

구분	1	2	3	4	5	전체
1	116,574	47,176	21,084	12,462	15,844	213,140
	54.69	22.13	9.89	5.85	7.43	100
2	55,897	95,020	45,742	13,483	3,039	213,181
	26.22	44.57	21.46	6.32	1.43	100
3	17,828	46,775	92,103	45,472	10,929	213,107
	8.37	21.95	43.22	21.34	5.13	100
4	12,700	18,122	44,448	92,422	45,215	212,907
	5.97	8.51	20.88	43.41	21.24	100
5	10,088	5,999	10,123	49,037	137,751	212,998
	4.74	2.82	4.75	23.02	64.67	100
전체	213,087	213,092	213,500	212,876	212,778	1,065,333
	20	20	20.04	19.98	19.97	100

다음의 <표 4-4>는 전겸업별 농가의 이동 상황을 제시하고 있다. 2003년에 전업농가인 경우가 계속 전업농가로 남는 비중은 2004년 71.2%, 2005년 73.5%, 2006년 70.3%, 2007년 72.1%로 나타나 연도별로 큰 변화가 없었다. 2종 겸업의 경우에도 지속적으로 2종 겸업으로 남는 비중이 높게 나타났다. 2003년에 2종 겸업을 하던 농가의 경우 지속적으로 2

종 겸업을 하는 비중은 2004년 85.2%, 2005년 78.2%, 2006년 76.7%, 2007년 70.7%로 나타났다. 2종 겸업의 경우 높은 지속 확률을 보이고 있지만 추세적으로 지속확률이 낮아지고 있음을 알 수 있다. 이에 반해 1종 겸업의 경우에는 상대적으로 이동 비중이 높게 나타났다. 2003년에 1종 겸업을 하던 가구 중에서 지속적으로 1종 겸업을 하는 비중은 2004년 57.9%, 2005년 46.1%, 2006년 45.9%, 2007년 37.0%로 나타났다. 다른 업종에 비해 상대적으로 지속확률이 매우 낮으며, 동 비율조차도 점차 하락하고 있다.

2003년에 1종 겸업을 하던 농가 중에서 전업농가로 변환한 비율은 2004년 33.5%, 2005년 39.4%, 2006년 40.1%, 2007년 48.2%로 나타나 상당 수의 1종 겸업가구가 전업농가로 전환하고 있으며, 동 비율이 해마다 증가하는 것으로 나타났다. 한편, 전업농가에서 1종 겸업으로 변화한 가구의 비율도 적지 않다. 2003년에 전업농가이다가 1종 겸업 농가로 변화한 비율은 2004년 21.6%, 2005년 19.6%, 2006년 22.8%, 2007년 19.3% 등으로 나타났다. 전업농가와 1종 겸업농가 사이에 상호 이동이 활발하게 일어나고 있음을 알 수 있다.

<표 4-4>전·겸업별 변화 분석

(1)2003->2004년

구분	1	2	3	전체
1	481,136	144,217	43,285	668,638
	71.96	21.57	6.47	100
2	45,719	78,969	11,653	136,341
	33.53	57.92	8.55	100
3	12,005	26,492	221,857	260,354
	4.61	10.18	85.21	100
전체	538,860	249,678	276,795	1,065,333
	50.58	23.44	25.98	100

(2) 2003->2005년

구분	1	2	3	전체
1	491,185	131,158	46,295	668,638
	73.46	19.62	6.92	100
2	53,631	62,896	19,814	136,341
	39.34	46.13	14.53	100
3	21,433	35,359	203,562	260,354
	8.23	13.58	78.19	100

전체	566,249	229,413	269,671	1,065,333
	53.15	21.53	25.31	100

(3) 2004->2006년

구분	1	2	3	전체
1	470,132	152,570	45,936	668,638
	70.31	22.82	6.87	100
2	54,713	62,598	19,030	136,341
	40.13	45.91	13.96	100
3	27,211	33,535	199,608	260,354
	10.45	12.88	76.67	100
전체	552,056	248,703	264,574	1,065,333
	51.82	23.35	24.83	100

(3) 2003->2007년

구분	1	2	3	전체
1	482,431	129,046	57,161	668,638
	72.15	19.3	8.55	100
2	65,746	50,384	20,211	136,341
	48.22	36.95	14.82	100
3	47,307	28,925	184,122	260,354
	18.17	11.11	70.72	100
전체	595,484	208,355	261,494	1,065,333
	55.9	19.56	24.55	100

2. 농업소득 및 농업순생산 결정요인 분석

농가경제활동조사의 패널화자료를 이용하여 농업소득 및 농업순생산의 결정요인을 분석하였다. <표 4-5>은 분석에 사용된 변수의 기초통계량을 제시하고 있으며, <표 4-6>은 분석결과를 제시하고 있다.

<표 4-6>의 (1)은 전 표본을 대상으로 농업종사가 1인당 농업소득(이하 1인당 농업소득)의 결정요인을 분석한 결과를 제시하고 있다. 임의효과 패널분석모형에 의하면 1인당 농업소득을 설명하는 유의한 변수로는 농가자산, 자영노동시간, 각종 1인당 자영 경작지의 면적, 각종 1인당 차용지의 면적, 그리고 일부 1인당 가축변수였다. 먼저 1인당 자산이 많을 수록,

그리고 자영노동시간이 길수록 1인당 농업소득이 높게 나타났다. 1인당 자작지가 많을 수록 농업소득이 많은 것으로 나타났는데, 계수의 상대적 크기는 1인당 논면적이 가장 크게 나타났으며, 다음으로 1인당 과수원 면적이 컸고, 1인당 밭 면적이 상대적으로 가장 작게 나타났다. 1인당 차용지 면적도 클수록 농업소득이 높은 것으로 나타났는데, 다만 계수의 크기는 논이나 밭인 경우에는 자작지가 차용지에 비해 크게 나탔으며, 과수원의 경우에는 자작지에 비해 차용지가 높게 나타났다[11]. 가축류의 경우 한우나 젖소는 통계적으로 유의하게 1인당 농업소득을 높이는 요인으로 나타났지만, 돼지나 닭인 경우에는 통계적으로 유의한 요인이 아니었다.

한편, 통상적으로 1인당 농업소득의 경우 경영주 연령이 많을 수록 낮을 것으로 예상 할 수 있는바, 본 분석에서 고려한 요인들을 통제한 경우 연령은 통계적으로 유의미한 변수가 아닌 것으로 나타났다. <표 4-6>의 (2)는 5년간 지속표본을 대상으로 1인당 농업소득의 결정요인을 분석한 결과를 제시하고 있다. 대부분의 분석결과는 전 표본을 대상으로 한 경우와 거의 차이가 없다.

1인당 농업순생산의 경우에도 1인당 농업소득과 유사한 결과를 보였다. 경영주 연령이 통계적으로 유의하지 않다는 점도 동일하였다[12]. 다만, 농업소득을 결정하는 요인이 아닌 것으로 판명되었던 돼지나 닭의 경우에도 농업순생산의 경우에는 통계적으로 유의하게 나타났다. 주목되는 점은 농업소득의 경우와는 달리 경작지의 면적 중에서 1인당 과수원의 면적이 가장 계수가 크게 나타났다는 점이다. 또한 1인당 밭면적의 경우 차용지의 계수값이 자작지의 계수값보다 큰 것으로 나타났다. 5년간 지속표본을 대상으로 동일한 분석을 실시한 결과 대부분의 결과는 유사하였으나, 돼지의 계수가 유의하게 나타났으며, 경영주 연령의 경우가 통계적으로 유의하게 1인당 농업소득을 낮추는 요인으로 나타났다.

<표 4-5> 변수의 기초통계량

(1) 전 표본

11) 이 부분은 해석하기 어려운 부분이다. 추가적인 검토가 필요한 것으로 판단된다.

12) pooled regression에서는 경영주 연령이 유의한 변수로 나타나고 있으나 개인특화요인(individual specific factor)을 고려한 임의효과 패널모형에서는 유의하지 않게 나타나고 있다.

변수명	설명	관측치수	평균	표준편차
per_inc	1인당 농업소득	15,331.0	5,515,502.0	10,100,000.0
per_pro	1인당 농업순생산(천원)	14,660.0	10,600.0	15,200.0
g17p	1인당 자산(천원)	14,660.0	173,000.0	229,000.0
g04	자영노동시간	15,342.0	1,535.0	1,581.3
mage	경영주연령	15,341.0	60.8	10.5
mages	경영주연령제곱	15,341.0	3,804.0	1,256.0
f07p	1인당 자작지 논 면적	14,651.0	860.4	1,446.7
f08p	1인당 자작지 밭 면적	14,651.0	437.3	746.0
f09p	1인당 자작지 과수원면적	14,651.0	220.1	716.7
f12p	1인당 차용지 논면적	14,651.0	759.6	1,898.9
f13p	1인당 차용지 밭면적	14,651.0	543.1	1,599.2
f14p	1인당 차용지 과수원면적	14,651.0	59.0	391.1
c16p	1인당 한우	14,649.0	1.2	6.2
c17p	1인당 젖소	14,649.0	0.4	3.3
c19p	1인당 돼지	14,649.0	2.7	15.5
c21p	1인당 닭	14,649.0	30.6	950.0

(2) 5년 유지표본

변수명	설명	관측치수	평균	표준편차
per_inc	1인당 농업소득	12,710	5,711,038.0	9,534,314.0
per_pro	1인당 농업순생산(천원)	12,215	10,900.0	14,400.0
g17p	1인당 자산(천원)	12,215	167,000.0	208,000.0
g04	자영노동시간	12,710	1,585.7	1,536.8
mage	경영주연령	12,710	60.6	10.4
mages	경영주연령제곱	12,710	3,777.6	1,233.8
f07p	1인당 자작지 논 면적	12,215	876.3	1,285.5
f08p	1인당 자작지 밭 면적	12,215	441.3	731.5
f09p	1인당 자작지 과수원면적	12,215	231.5	738.4
f12p	1인당 차용지 논면적	12,215	798.3	1,939.6
f13p	1인당 차용지 밭면적	12,215	559.8	1,618.8
f14p	1인당 차용지 과수원면적	12,215	63.2	408.8
c16p	1인당 한우	12,215	1.3	6.6
c17p	1인당 젖소	12,215	0.4	3.2
c19p	1인당 돼지	12,215	2.8	16.5
c21p	1인당 닭	12,215	23.1	883.9

<표 4-6> 회귀분석 결과

(1) 1인당 농업소득 : 전 표본

all data	pooled			random effect		
per_inc	계수추정치	표준오차	t-값	계수추정치	표준오차	t-값
g17p	0.005	0.001	0.000	0.006	0.001	0.000
g17ps	-0.000	0.000	0.000	-0.000	0.000	0.000
g04	2,986.120	83.243	0.000	2,821.880	104.320	0.000
g04s	-0.097	0.006	0.000	-0.077	0.007	0.000
mage	96,540.480	69,810.380	0.167	111,992.900	96,754.850	0.247
mages	-700.928	586.070	0.232	-883.127	810.772	0.276
f07p	997.689	55.179	0.000	864.063	71.790	0.000
f08p	319.601	103.794	0.002	252.866	137.463	0.066
f09p	827.810	110.890	0.000	666.589	152.666	0.000
f12p	439.343	41.212	0.000	368.397	55.689	0.000
f13p	169.143	49.590	0.001	183.492	65.712	0.005
f14p	903.612	189.709	0.000	958.545	244.320	0.000
c16p	170,801.900	12,164.920	0.000	133,152.800	17,074.730	0.000
c17p	462,432.500	22,787.160	0.000	405,487.800	31,747.640	0.000
c19p	9,912.970	4,746.266	0.037	3,147.551	6,216.295	0.613
c21p	143.283	77.237	0.064	106.320	77.505	0.170
_cons	-4,537,239	2,042,121	0.026	-4,577,972	2,838,435	0.107
sigma_u				5,190,701		
sigma_e				7,538,753		
rho				0.322		

(2) 농업소득 : 5년 유지표본

pan5=1	pooled			random effect		
per_pro	계수추정치	표준오차	t-값	계수추정치	표준오차	t-값
g17p	0.005	0.001	0.000	0.006	0.001	0.000
g17ps	-0.000	0.000	0.000	-0.000	0.000	0.000
g04	3,083.014	94.817	0.000	2,967.791	115.005	0.000
g04s	-0.124	0.009	0.000	-0.109	0.011	0.000
mage	5,908.571	73,065.870	0.936	4,627.479	93,150.030	0.960
mages	13.234	615.602	0.983	-14.511	783.938	0.985
f07p	1,040.079	63.701	0.000	894.952	79.179	0.000
f08p	240.425	108.190	0.026	214.164	132.651	0.106
f09p	884.640	112.060	0.000	723.198	141.142	0.000
f12p	442.445	41.127	0.000	390.744	51.210	0.000
f13p	208.322	50.959	0.000	192.123	63.127	0.002
f14p	906.794	185.612	0.000	945.841	223.600	0.000
c16p	161,871.600	11,785.070	0.000	133,043.300	14,863.070	0.000
c17p	428,506.000	23,874.770	0.000	392,215.100	30,604.070	0.000
c19p	10,916.060	4,549.978	0.016	6,322.954	5,518.662	0.252
c21p	-63.524	84.584	0.453	-79.812	84.328	0.344
_cons	-1,739,556	2,129,138	0.414	-1,417,209	2,719,840	0.602

sigma_u	3,603,132		
sigma_e	7,286,080		
rho	0.196		

(3) 1인당 농업순생산 : 전 표본

all data	pooled			random effect		
per_inc	계수추정치	표준오차	t-값	계수추정치	표준오차	t-값
g17p	0.013	0.001	0.000	0.016	0.001	0.000
g17ps	-0.000	0.000	0.000	-0.000	0.000	0.000
g04	4,300.047	106.130	0.000	4,125.201	142.324	0.000
g04s	-0.094	0.008	0.000	-0.080	0.009	0.000
mage	-209,479.000	89,003.870	0.019	-119,447.400	137,960.000	0.387
mages	1,291.474	747.203	0.084	549.658	1,155.019	0.634
f07p	1,808.468	70.349	0.000	1,863.882	99.859	0.000
f08p	632.084	132.330	0.000	668.621	190.994	0.000
f09p	2,338.490	141.377	0.000	1,964.103	216.599	0.000
f12p	1,639.740	52.543	0.000	1,336.532	77.903	0.000
f13p	884.020	63.224	0.000	876.448	91.169	0.000
f14p	1,979.830	241.867	0.000	2,332.442	335.325	0.000
c16p	297,509.200	15,509.520	0.000	245,196.700	24,241.720	0.000
c17p	633,044.700	29,052.200	0.000	564,521.200	45,357.300	0.000
c19p	26,223.680	6,051.193	0.000	20,755.030	8,534.159	0.015
c21p	790.502	98.472	0.000	373.430	99.053	0.000
_cons	5,121,749	2,603,576	0.049	2,512,532	4,053,248	0.535
sigma_u				8,368,430		
sigma_e				9,371,515		
rho				0.444		

(4) 1인당 농업순생산 : 5년 유지표본

pan5=1	pooled			random effect		
per_pro	계수추정치	표준오차	t-값	계수추정치	표준오차	t-값
g17p	0.015	0.001	0.000	0.017	0.001	0.000
g17ps	-0.000	0.000	0.000	-0.000	0.000	0.000
g04	4,207.922	119.810	0.000	4,148.851	144.162	0.000
g04s	-0.100	0.011	0.000	-0.096	0.013	0.000
mage	-326,042.900	92,324.700	0.000	-317,315.000	116,333.800	0.006
mages	2,188.383	777.863	0.005	2,102.738	979.143	0.032
f07p	1,525.043	80.491	0.000	1,438.331	99.071	0.000
f08p	530.765	136.707	0.000	551.664	166.162	0.001
f09p	2,258.519	141.597	0.000	2,021.537	176.408	0.000

f12p	1,693.363	51.967	0.000	1,574.514	64.074	0.000
f13p	879.519	64.390	0.000	865.274	79.012	0.000
f14p	1,921.559	234.536	0.000	2,102.093	280.413	0.000
c16p	281,895.600	14,891.390	0.000	257,035.200	18,577.360	0.000
c17p	569,106.100	30,167.720	0.000	528,427.700	38,205.530	0.000
c19p	24,339.720	5,749.269	0.000	21,725.500	6,918.852	0.002
c21p	73.249	106.879	0.493	−49.555	106.985	0.643
_cons	8,960,596	2,690,339	0.001	8,758,533	3,396,191	0.010
sigma_u				4,395,305		
sigma_e				9,297,124		
rho				0.183		

V. 한계점 및 추후 연구 방향

본 연구는 농가경제활동조사를 패널자료화한 결과와 패널화된 자료를 이용하여 실제로 자료를 분석하는 예를 제시하였다. 본 연구의 결과에 의하면 농가경제활동조사의 경우 비록 표본을 추적하는 패널조사가 아니라는 점에서 한계가 있을 것이나, 실제로는 패널유지율 면에서 추적 조사하는 일반패널조사에 비해 표본유지율이 상대적으로 높음을 보여주었다. 이와 같은 결과는 농가경제활동조사를 패널자료로 만들어서 사용하는 데 큰 제약이 없다는 점을 의미하는 것이라고 하겠다. 또한 통계청의 다른 반복 횡단면 자료의 경우에도 패널화를 통해 다양한 이점을 얻을 수 있다는 점을 시사하는 것이라고 하겠다.

반복횡단면 자료를 이용하여 패널데이터를 생성하는 경우 가중치의 문제가 제기된다. 대부분의 경우 패널탈락은 체계적인 요인에 의해 이루어지는 경우가 많으며, 이를 보정하지 않는 경우 자료분석의 신뢰성을 확보하기 어렵기 때문이다. 그런데 대부분의 패널자료에서 가중치를 작성하는 경우에 봉착하는 문제가 신규 진입표본의 가중치를 부여하는 방안의 모호성이다. 본고의 분석에 의하면 농가패널조사와 같은 반복 횡단면 자료의 경우 이미 횡단면가중치가 존재하기 때문에 신규 진입표본의 가중치설정의 문제가 발생하지 않는다는 점을 알 수 있다. 또한 기존의 횡단면가중치를 이용하여 종단면가중치를 사후층화할 수 있다는 장점도 있다. 결국 본고의 분설결과는 횡단면가중치가 존재하는 반복횡단면자료의 경우 종단면가중치가 쉽게 도출될 수 있다는 점을 보여주는 것이라고 하겠다.

농가경제활동조사의 패널화자료를 이용하여 분석한 결과에 의하면 농가의 경우 1분위와 5분위에서는 소득의 고착화현상이 발견되며, 이러한 고착화현상은 농가소득보다 농어소득에서 더욱 크게 나타났다. 이에 반해 3분위에 속하는 가구의 경우에는 상대적으로 소득이동이 활발

하였는데, 상위계층이나 하위계층으로의 이동비중이 유사하게 나타났다.

또한 한국 농가의 경우 전업농가나 2종 겸업의 경우에는 동일한 형태를 유지하는 비중이 높지만, 1종 겸업의 경우에는 동일한 형태를 유지하는 비중이 상대적으로 낮으며, 전업농가로 변화하는 비중이 높게 나타났다. 전업농가의 경우에도 적지 않은 경우 1종 겸업으로 전환하는 경우가 나타나고 있어서 전업농가와 1종 겸업농가간의 상호이동성이 큰 것으로 나타났다.

임의효과 패널모형을 이용하여 1인당 농업소득이나 농업순생산을 결정하는 요인을 분석한 결과에 의하며 농가자산과 근로시간이 많을 수록 농업소득이나 농업순생산이 많은 것으로 나타났다. 다양한 종류의 1인당 경작지 면적도 중요 요인으로 나타났는데, 1인당 과수원이 가장 영향을 많이 미치는 요인으로 나타났다. 또한 젖소, 한우 등의 가축류의 경우에도 중요한 결정요인으로 나타났다.

그러나 본고는 다양한 한계점을 가지고 있다. 먼저 본고에서 계산한 종단면가중치는 응답률모형에 근거하고 있는데, 동 응답률 모형은 모형설정의 오류(mis-specification)의 문제를 가질 수 있다. 향후 보다 다양한 응답률 모형을 이용하여 본고의 결과에 대한 강건성(robustness)을 검토할 필요가 있다. 실증분석에 사용한 1인당 농업생산이나 농업소득의 경우에도 모형설정의 문제점을 검토할 필요가 있다.

<부록 1> 표본유지율의 결정요인 분석

농가경제조사의 경우 비록 표본유지율이 높다고 하지만 표본탈락이 체계적으로 발생하는 경우 동 조사를 이용하여 모집단을 분석하는 데에 한계가 있다. 여기에서는 인접한 두 년도간의 표본탈락한 가구와 표본유지한 가구의 특성을 비교함으로써 체계적인 표본탈락이 발생하였는지를 검토한다. <부표 1-1>은 인접한 두 년도의 경우에 표본유지가구와 표본탈락가구의 특성을 비교하고 있다.

평균가구원수는 표본탈락가구와 표본유지가구 사이에 커다란 차이는 없었다. 농업상시종사자수의 경우 표본탈락가구가 평균적으로 0.2명 정도 적은 것으로 나타나 농업상시종사자수가 적을 수록 표본탈락가구로 전환된 확률이 높아지는 것으로 나타났다.

경지면적의 경우에는 표본탈락가구가 표본유지가구보다 체계적으로 작은 것으로 나타났다. 예를 들어 2003년도에 응답한 가구 중에서 2004년에도 계속 응답한 가구는 평균경지면적이 5,164㎡에 달했는데 표본탈락가구는 3,913㎡로 나타나 표본유지가구에 비해 1,251㎡가 적은 것으로 나타났다. 이와 같은 현상은 다른 년도에도 동일하게 나타나는데 2006년의 표본유지가구 중에서 2007년에도 응답한 가구는 평균경지면적이 5,136㎡였지만 표본탈락가구의 평균경지면적은 3,128㎡에 불과하였다.

경지면적이 적은 것을 반영하여 표본탈락가구가 표본유지가구보다 자영농업노동시간이 적은 것으로 나타났다. 예를 들어 2003년의 응답가구 중에서 2004년의 응답가구는 자영농업노동시간 평균이 1,602시간이었지만 비응답가구는 1,320시간으로 300시간 이상 적게 나타났다. 이와 같은 현상은 다른 년도에 비슷하게 나타나는데 2006년 응답자 중 2007년의 응답가구는 평균 자영농업노동시간이 1,611시간인 반면 비응답가구의 경우에는 993시간에 불과하였다.

농가소득의 경우에도 표본탈락가구의 소득이 표본유지가구의 소득에 비해 적은 것으로 나타났다. 이러한 차이는 2004년도의 경우 189만원, 2005년도의 경우 467만원, 2006년도의 45만원, 2007년의 경우 307만원으로 나타났다. 이러한 경향은 경상소득의 경우에도 유사하게 나타났다.

일반적으로 가계지출의 경우에도 표본탈락가구의 가계지출이 표본유지가구의 가계지출보다 작은 것으로 나타났다. 예를 들어 표본탈락가구의 가계지출은 계속표본유지가구에 비해 2005년의 겨우 242만원이 작았으며, 2006년은 217만원, 2007년은 352만원이 적은 것으로 나타났다. 다만 2004년의 경우에는 표본탈락가구의 가계지출이 표본유지가구에 비해 오히려

268만원이 적었다. 소비지출의 경우에도 가계지출과 유사한 경향을 보였다. 농가가처분소득의 경우에는 소득 및 지출과 다른 패턴을 보였는데 2006년도에는 계속표본유지가구의 가처분소득이 표본유지가구보다 적은 것으로 나타났고, 다른 년도에는 더 많은 것으로 나타났다.

농가자산의 경우에는 전반적으로 표본탈락가구가 표본유지가구에 비해 자산이 적은 것으로 나타났다. 2004년의 경우 표본탈락가구의 자산은 계속표본유지가구의 자산에 비해 2,940만원이 작았으며, 2005년의 경우에는 무려 6,242만원이 적은 것으로 나타났다. 2006년의 경우에는 1,989만원이 적었다. 다만 20007년의 경우에는 표본탈락가구의 평균자산이 표본유지가구의 평균자산에 비해 337만원 많은 것으로 나타났다.

농가부채의 경우에는 뚜렷한 패턴을 보이고 있지 않았다. 2004년의 경우에는 표본탈락가구의 농가부채가 표본유지가구의 부채에 비해 436만원이 적었고, 2005년의 경우에는 745만원이 적었으나, 2006년의 경우에는 818만원이 많았고, 2007년의 경우에는 921만원이 많았다. 표본탈락가구는 표본유지가구에 비해 전반적으로 농업순생산이 적었고, 농업의존도가 높았으며, 농업소득률은 높은 것으로 나타났다.

이와 같은 결과를 개략적으로 종합하면 표본탈락가구는 표본유지가구에 비해 농업상시종사자수가 적고 경지면적도 적으며, 이에 따라 농업근로시간과 농가소득 및 경상소득도 적은 것으로 나타났다. 특이한 점은 농가자산과 부채의 경우에는 표본탈락가구가 더 많은 경우가 나타났다.

<부표 1-1> 인접한 년도간 지속표본과 탈락표본의 특성 비교

(1) 2003년과 2004년의 비교

구분	비응답			응답			평균비율
	표본수	평균	표준편차	표본수	평균	표준편차	
가구원수	130	3.1	1.5	2,912	3.0	1.3	104.3
농업상시종사자수	130	1.6	0.7	2,912	1.8	0.7	86.2
경지면적	130	3,813.1	5,068.3	2,912	5,164.2	5,831.7	73.8
자영농업노동시간	130	1,319.6	1,401.7	2,912	1,601.9	1,609.5	82.4
농가소득	130	25,328,079.9	28,320,194.1	2,912	27,221,484.0	24,297,444.1	93.0
경상소득	130	20,608,822.0	25,999,932.1	2,912	22,637,194.5	23,051,173.7	91.0
가계지출	130	26,553,474.1	35,941,595.7	2,912	23,885,861.8	19,601,447.2	111.2
소비지출	130	19,348,020.3	13,484,050.3	2,912	18,117,424.6	10,194,436.6	106.8
농가처분가능소득	130	18,122,626.0	25,441,870.4	2,912	21,453,046.7	26,933,941.6	84.5
농가자산	130	235,353,866.2	301,991,574.1	2,912	205,906,243.5	198,588,846.0	114.3
농가부채	130	32,709,268.5	49,743,946.0	2,912	28,348,714.5	47,844,814.4	115.4
농업순생산	130	13,910,975.4	27,028,597.9	2,912	17,789,725.4	23,939,179.2	78.2
농업의존도	130	-2.7	388.5	2,912	31.0	817.3	-
가계비충족률	130	30.6	71.0	2,912	55.0	83.1	55.5
농업소득률	130	44.7	139.9	2,912	5.5	2,437.0	817.7

(2) 2004년과 2005년의 비교

구분	비응답			응답			평균비율
	표본수	평균	표준편차	표본수	평균	표준편차	
가구원수	138	2.8	1.2	2918	3.0	1.3	94.0
농업상시종사자수	138	1.6	0.8	2918	1.8	0.7	89.5
경지면적	138	3,694.0	4,596.1	2918	5,178.5	5,788.5	71.3
자영농업노동시간	138	1,465.6	3,014.4	2918	1,602.4	1,519.7	91.5
농가소득	138	22,518,723.4	29,814,668.1	2918	27,188,651.2	24,816,166.0	82.8
경상소득	138	17,910,287.7	28,688,949.4	2918	22,569,717.9	23,608,702.8	79.4
가계지출	138	21,529,235.6	12,894,297.0	2918	23,954,714.9	19,584,889.2	89.9
소비지출	138	17,451,554.3	10,158,680.4	2918	18,153,556.4	10,188,822.9	96.1
농가처분가능소득	138	18,441,042.2	29,132,105.2	2918	21,387,492.7	27,444,724.5	86.2
농가자산	138	265,484,312.0	335,826,384.1	2918	203,064,467.6	188,401,209.5	130.7
농가부채	138	36,224,122.8	93,661,835.8	2918	28,771,386.5	48,680,956.7	125.9
농업순생산	138	11,264,031.6	30,172,242.6	2918	17,782,213.0	24,480,034.1	63.3
농업의존도	138	25.5	135.4	2918	31.6	816.0	80.8
가계비충족률	138	19.5	130.3	2918	55.5	85.1	35.1
농업소득률	138	140.5	1,264.4	2918	−0.8	2,420.2	−

(3) 2005년과 2006년의 비교

구분	비응답			응답			평균비율
	표본수	평균	표준편차	표본수	평균	표준편차	
가구원수	120	2.9	1.3	2928	3.0	1.3	95.4
농업상시종사자수	120	1.6	0.7	2928	1.8	0.7	87.3
경지면적	120	3,941.8	4,776.1	2928	5,164.7	5,796.1	76.3
자영농업노동시간	120	1,160.7	1,303.4	2928	1,607.7	1,523.0	72.2
농가소득	120	26,583,694.0	25,515,901.4	2928	27,035,176.4	25,210,553.2	98.3
경상소득	120	22,290,325.8	25,105,488.0	2928	22,353,167.2	24,310,015.0	99.7
가계지출	120	21,992,177.0	17,727,473.4	2928	24,160,277.8	20,655,942.0	91.0
소비지출	120	17,139,250.1	10,454,331.9	2928	18,255,174.1	10,350,757.7	93.9
농가처분가능소득	120	21,730,767.2	28,438,318.4	2928	21,130,072.6	27,562,395.9	102.8
농가자산	120	223,526,359.5	243,530,170.5	2928	204,544,196.2	195,863,067.5	109.3
농가부채	120	21,387,292.0	42,326,740.5	2928	29,555,430.4	50,792,842.8	72.4
농업순생산	120	15,342,868.7	26,817,771.5	2928	17,621,026.6	24,914,187.9	87.1
농업의존도	120	29.9	30.8	2928	30.3	818.7	98.6
가계비충족률	120	47.1	69.5	2928	54.6	86.1	86.3
농업소득률	120	36.5	41.3	2928	−0.4	2,416.2	−

(4) 2006년과 2007년의 비교

구분	비응답			응답			평균비율
	표본수	평균	표준편차	표본수	평균	표준편차	
가구원수	59	2.9	1.2	3002	3.0	1.3	97.0
농업상시종사자수	59	1.5	0.7	3002	1.8	0.7	81.2
경지면적	59	3,127.9	2,839.4	3002	5,136.3	5,795.2	60.9
자영농업노동시간	59	992.8	875.0	3002	1,610.7	1,621.2	61.6
농가소득	59	23,841,508.1	20,506,882.2	3002	26,915,447.2	25,513,301.6	88.6
경상소득	59	20,016,890.2	19,664,826.0	3002	22,242,411.9	24,635,261.8	90.0
가계지출	59	20,603,608.2	13,044,678.1	3002	24,122,539.8	20,484,713.2	85.4
소비지출	59	16,238,517.8	8,890,125.7	3002	18,269,130.8	10,361,016.1	88.9
농가처분가능소득	59	19,476,417.8	20,692,532.8	3002	21,062,038.1	27,758,669.9	92.5
농가자산	59	204,298,501.5	166,446,166.9	3002	207,673,488.4	205,939,494.6	98.4
농가부채	59	20,767,128.2	34,094,528.3	3002	29,980,793.1	52,760,769.6	69.3
농업순생산	59	11,663,143.2	15,726,960.8	3002	17,436,192.4	25,376,103.4	66.9

농업의존도	59	-17.1	382.7	3002	30.7	807.1	-
가계비충족률	59	47.5	76.2	3002	53.1	89.3	89.6
농업소득률	59	84.6	320.5	3002	4.4	2,401.2	1909.9

주 : 가계비충족률은 농업소득의 가계비 충족률을 의미함. 평균비율은 응답가구의 평균 대비 비응답가구의 평균의 비율을 의미함.

이제 이와 같은 일변량 분석결과를 다변량 분석으로 확대함으로써 표본유지를 결정하는 요인들을 점검해 본다. <부표 1-3> (1)은 2003년에 응답한 가구 중에서 2004년에도 응답을 한 경우에 1의 값을 가지고, 2004년도에는 표본탈락한 가구의 경우에는 0의 값을 가지는 이진변수(binary variable)를 종속변수로 하여 표본유지확률을 로짓모형으로 추정한 결과를 제시하고 있다. 분석결과에 의하면 가구원수가 많을 수록 표본유지확률은 적어지지만 통계적으로는 유의하지 않았다. 농업상시종사자수가 많을 수록 통계적으로 유의하게 표본유지확률이 높았다. 경지면적도 많을수록 표본유지확률이 높았으며, 동 변수는 통계적으로 유의한 설명변수로 나타났다. 농가소득의 경우 소득이 많을수록 표본유지확률이 높았지만 동 변수는 통계적으로 유의하지 않았다. 이외에 가계지출이나 가구자산, 가구부채 등의 경우에는 모두 많을 수록 표본유지 확률이 떨어지는 것으로 나타났으나 모두 통계적으로 유의하지는 않았다. <부표 1-4>의 기타 항목들도 유사하게 해석할 수 있다.

그런데 흥미로운 점은 모든 응답년도에 걸쳐 일관되게 통계적으로 유의하게 나타나는 변수는 농업상시종업원수였으며 이는 양수로 추정되었다. 2003년과 2004년 동안의 표본유지를 제외하면 가구원수가 기타 모든 경우에 유의하게 나타났으며, 가구원수가 많을 수록 표본유지확률이 높은 것으로 나타났다. 모든 경우에 경지면적이 많을수록 표본유지 확률이 높은 것으로 나타났으며, 2005년과 2006년 사이를 제외하면 모든 경우에 통계적으로 유의하게 나타났다. 가구소득이나 가구자산 및 부채 변수는 계수부호도 기간에 따라 상이하게 나타났으며, 통계적으로 유의한 경우도 거의 없는 것으로 나타났다.

이와 같은 다변량 분석결과는 가구소득이나 소비, 부채 및 자산 등의 변수에서 체계적인 표본탈락요인은 발견되지 않으며, 다만, 가구원수 또는 농업상시종사자수와 경지면적 등이 표본탈락과 표본유지를 결정짓는 중요한 요인으로 나타났다[13].

한편, <부표 1-4>의 로짓분석에 근거하여 표본유지확률을 추정한 결과는 다음의 <부표 1-5>에 제시되어 있다. <부표 1-5>에 의하면 본고의 응답모형을 사용하는 경우 예측치

13) 본고의 분석결과는 추후 보다 다양한 모형설정을 통해 추정결과의 강건성(robustness)에 대한 검토가 이루어져야 할 것이다. 한편, 본고에서 나타난 결과 중에서 소득이나 부채, 자산변수의 경우에는 상대적으로 측정오차가 크다고 알려진 항목들임에 유의할 필요가 있다.

의 평균치는 95%정도이지만, 최소값은 p23hat의 경우 0.202에 불과하여 동 모형의 응답확률을 이용하여 가중치를 설정하는 경우에 매우 큰 가중치가 도출될 수 있음을 보여 주고 있다.

<표 1-3> 로짓분석 변수의 기초통계량

Variable	변수설명	Obs	Mean	Std. Dev.	Min	Max
p45	2006년도 응답자중 2007년도 응답자더미	3,074	0.98	0.15	0.00	1.00
g012006	2006년의 가구원수	3,074	2.84	1.22	2.00	10.00
g022006	2006년의 농업상시종사자수	3,074	1.78	0.67	0.00	5.00
g032006s	2006년 경지면적/1000	3,074	5.16	6.15	0.00	89.96
g052006s	2006년 농가소득/십만	3,074	336.30	295.85	-921.61	5,281.97
g122006s	2006년 가계지출/백만	3,074	29.19	26.50	2.32	394.22
g172006s	2006년 농가자산/백만	3,074	348.99	384.97	9.66	5,230.45
g182006s	2006년 농가부채/백만	3,074	31.35	60.34	0.00	1,033.11

주 : 편의상 p45인 경우만 제시하였음.

<부표 1-4> 인접 연도간 응답여부의 로짓분석 결과

(1) 2003년->2004년

구분	계수	표준오차	t-값	z-값
g012003	-0.060	0.071	-0.848	0.396
g022003	0.404**	0.129	3.130	0.002
g032003s	0.059**	0.027	2.192	0.028
g052003s	0.052	0.036	1.441	0.150
g122003s	-0.033	0.037	-0.886	0.375
g172003s	-0.006	0.004	-1.521	0.128
g182003s	-0.027	0.019	-1.451	0.147
_cons	2.506***	0.318	7.871	0.000

(2) 2004년->2005년

구분	계수	표준오차	t-값	z-값
g012004	0.162*	0.086	1.888	0.059
g022004	0.403**	0.130	3.101	0.002
g032004s	0.082**	0.028	2.931	0.003
g052004s	0.000	0.000	0.774	0.439

g122004s	0.002	0.005	0.296	0.767
g172004s	0.000**	0.000	−2.461	0.014
g182004s	−0.005***	0.001	−3.673	0.000
_cons	1.698***	0.336	5.049	0.000

(3) 2005−>2006년

구분	계수	표준오차	t−값	z−값
g012005	0.186*	0.093	1.996	0.046
g022005	0.588***	0.139	4.229	0.000
g032005s	0.019	0.023	0.803	0.422
g052005s	0.000	0.000	−0.064	0.949
g122005s	0.001	0.004	0.337	0.736
g172005s	−0.001***	0.000	−2.707	0.007
g182005s	0.002	0.002	0.979	0.328
_cons	1.733***	0.362	4.790	0.000

(4)2006년−>2007년

구분	계수	표준오차	t−값	z−값
g012006	0.216*	0.131	1.647	0.100
g022006	0.416**	0.189	2.206	0.027
g032006s	0.105**	0.045	2.352	0.019
g052006s	0.001	0.001	1.167	0.243
g122006s	−0.005*	0.003	−1.844	0.065
g172006s	0.000	0.000	0.387	0.699
g182006s	−0.002	0.002	−0.702	0.483
_cons	1.986***	0.487	4.076	0.000

<표 1−5> 응답확률 예측치의 기초통계량

변수	관측치수	평균	표준편차	최소값	최대값
p12hat	3042	0.957	0.022	0.672	0.999
p23hat	3056	0.955	0.031	0.202	1.000
p34hat	3048	0.961	0.023	0.543	0.999
p45hat	3074	0.977	0.015	0.738	1.000

주 : p12hat은 2003년에 응답한 가구 중에서 2004년에도 응답할 확률에 대한 추정치이며,
p23hat은 2004년에 응답한 가구 중에서 2005년에도 응답할 확률에 대한 추정치이며,
p34hat은 2005년에 응답한 가구 중에서 2006년에도 응답할 확률에 대한 추정치이며,

p45hat은 2006년에 응답한 가구 중에서 2007년에도 응답할 확률에 대한 추정치를 의미함.

<부록 2> 가중치를 사용하지 않은 경우의 소득이동 및 상태변화 분석결과

1. 소득계층의 이동분석

<부표 2-1>은 인접년도간 공통된 응답자를 대상으로 5분위 분포상의 계층이동상황을 보여주고 있다. 2003년도에 소득 1분위에 속했던 가구들 중 52.2%는 2004년에도 소득 1분위에 속하는 것으로 나타났다. 2004년도에 소득 1분위에 속했던 가구들 중 53.4%는 2005년에도 소득 1분위에 속하는 것으로 나타났다. 2005년에 소득 1분위에 속했던 가구들 중 53.5%는 2006년에도 소득 1분위에 속하는 것으로 나타났다. 2006년에 소득 1분위에 속했던 가구들 중 56.2%는 2007년에도 소득 1분위에 속하는 것으로 나타났다.

이와 같이 소득 1분위에 속하는 가구들이 소득 1분위에 계속 머무를 확률은 최근에 올수록 미세하나마 증가하는 것으로 나타났다. 이에 반해 소득 1분위에 속했던 가구들이 다음 면도에 소득 5분위에 속할 확률은 소폭이나마 점차 감소하는 것으로 나타났다. 2003년도에 소득 1분위에 속했던 가구들 중 6.0%는 2004년에 소득 5분위에 속하는 것으로 나타났다. 2004년도에 소득 1분위에 속했던 가구들 중 5.4%는 2005년에는 소득 5분위에 속하는 것으로 나타났다. 2005년에 소득 1분위에 속했던 가구들 중 4.5%는 2006년에는 소득 5분위에 속하는 것으로 나타났다. 2006년에 소득 1분위에 속했던 가구들 중 5.5%는 2007년에는 소득 1분위에 속하는 것으로 나타났다.

이에 반해 2003년도에 소득 5분위에 속했던 가구들 중 60.9%는 2004년에도 소득 5분위에 속하는 것으로 나타났다. 2004년도에 소득 5분위에 속했던 가구들 중 57.4%는 2005년에도 소득 5분위에 속하는 것으로 나타났다. 2005년에 소득 5분위에 속했던 가구들 중 60.7%는 2006년에도 소득 5분위에 속하는 것으로 나타났다. 2006년에 소득 5분위에 속했던 가구들 중 59.2%는 2007년에도 소득 5분위에 속하는 것으로 나타났다.

이와 같이 소득 1분위와 2분위가구들은 같은 분위에 머무를 확률이 55~60%정도였지만 다른 분위에서는 비교적 분위이동이 상대적으로 많은 것으로 나타났다. 소득 3분위를 기준으

로 볼 때 2003년에서 2004년에 걸쳐 동일 분위에 머무를 확률은 34.7%, 2004년과 2005년 동안은 37.8%, 2005년과 2006년 동안은 35.9%, 2006년과 2007년 동안은 36.2%로 기간별로 거의 차이가 없었다.

소득 3분위에서 직전 1년 동안 1 또는 2분위로 추락할 확률과 4 또는 5분위로 상승할 확률은 2004년의 경우 각각 27.2%, 38.2%였으며, 2005년의 경우에는 각각 30.3%, 31.9%, 2006년의 경우에는 각각 31.8%, 32.3% 그리고 2007년의 경우에는 33.5%, 30.3%로 나타났다. 이는 소득 3분위에 위치하던 가구들이 하위소득분위로 떨어질 확률이 점점 더 커지고 있는 반면, 소득상위분위로 올라갈 확률이 점점 더 하락하고 있음을 보여주는 결과이다.

<부표 2-1> 5분위 분포상의 변화 : 인접년도간 분석(단위:가구,%)

(1) 2003년->2004년

분위	1	2	3	4	5	전체
1	307	139	67	40	35	588
	52.2	23.6	11.4	6.8	6.0	100.0
2	108	221	130	58	32	549
	19.7	40.3	23.7	10.6	5.8	100.0
3	54	104	202	150	72	582
	9.3	17.9	34.7	25.8	12.4	100.0
4	41	37	126	220	173	597
	6.9	6.2	21.1	36.9	29.0	100.0
5	29	23	52	129	363	596
	4.9	3.9	8.7	21.6	60.9	100.0
합계	539	524	577	597	675	2912
	18.5	18.0	19.8	20.5	23.2	100.0

(2) 2004년->2005년

구분	1	2	3	4	5	합계
1	288	123	60	39	29	539
	53.4	22.8	11.1	7.2	5.4	100.0
2	126	207	111	48	23	515
	24.5	40.2	21.6	9.3	4.5	100.0
3	59	118	221	143	43	584
	10.1	20.2	37.8	24.5	7.4	100.0

4	23	51	146	239	138	597
	3.9	8.5	24.5	40.0	23.1	100.0
5	37	29	65	160	392	683
	5.4	4.3	9.5	23.4	57.4	100.0
합계	533	528	603	629	625	2918
	18.3	18.1	20.7	21.6	21.4	100.0

(3) 2005년->2006년

구분	1	2	3	4	5	합계
1	284	123	66	34	24	531
	53.5	23.2	12.4	6.4	4.5	100.0
2	123	212	120	44	28	527
	23.3	40.2	22.8	8.4	5.3	100.0
3	53	140	218	128	68	607
	8.7	23.1	35.9	21.1	11.2	100.0
4	32	60	129	274	142	637
	5.0	9.4	20.3	43.0	22.3	100.0
5	23	26	53	144	380	626
	3.7	4.2	8.5	23.0	60.7	100.0
합계	515	561	586	624	642	2928
	17.6	19.2	20.0	21.3	21.9	100.0

(4) 2006년->2007년

구분	1	2	3	4	5	합계
1	286	120	46	29	28	509
	56.2	23.6	9.0	5.7	5.5	100.0
2	125	246	153	48	17	589
	21.2	41.8	26.0	8.2	2.9	100.0
3	68	135	219	133	50	605
	11.2	22.3	36.2	22.0	8.3	100.0
4	28	56	146	258	159	647
	4.3	8.7	22.6	39.9	24.6	100.0
5	36	22	68	140	386	652
	5.5	3.4	10.4	21.5	59.2	100.0
합계	543	579	632	608	640	3002

	18.1	19.3	21.1	20.3	21.3	100.0

<부표 2-2>는 5년간 지속적으로 표본을 유지한 가구들을 대상으로 5년간 소득분위의 변동을 보여주고 있다. 소득 1분위를 기준으로 할 때 2003년에 1분위에 속하는 가구 중 50.4%는 2004년에도 계속해서 1분위에 머물렀으며, 2년 후에 2005년에는 44.4%가 1분위에 머물고 있었고, 3년 후인 2006년에는 41.5%가 1분위에 머물러 있었고, 4년 후인 2007년에는 44.2%가 1분위에 머물러 있었다. 이와 같은 상황은 소득 1분위에 속하는 가구는 4년 후에도 무려 44.2%가 역시 소득 1분위에 머물러 있다는 점을 의미하며, 저소득 농가가구의 고착화현상을 보여주는 결과라고 하겠다.

소득 5분위를 기준으로 할 때 2003년에 5분위에 속하는 가구 중 60.2%는 2004년에도 계속해서 5분위에 머물렀으며, 2년 후에 2005년에는 54.5%가 1분위에 머물고 있었고, 3년 후인 2006년에는 50.7%가 5분위에 머물러 있었고, 4년 후인 2007년에는 48.9%가 5분위에 머물러 있었다. 이와 같은 상황은 소득 5분위에 속하는 가구는 4년 후에도 무려 반 정도가 역시 소득 5분위에 머물러 있다는 점을 의미하며, 고소득 농가가구의 안정화현상을 보여주는 결과라고 하겠다.

소득 1분위와 5분위를 제외한 다른 분위에서는 분위간 이동이 상대적으로 활발하였다. 소득 3분위를 기준으로 할 때 2003년에 3분위에 속하는 가구 중 35.1%는 2004년에도 계속해서 3분위에 머물렀으며, 2년 후에 2005년에는 33.9%가 3분위에 머물고 있었고, 3년 후인 2006년에는 31.4%가 3분위에 머물러 있었고, 4년 후인 2007년에는 33.3%가 3분위에 머물러 있었다. 동일 분위에 머무를 확률이 3분위의 경우에 1분위와 5분위보다 낮다는 점을 보여주고 있다.

이에 반해 2003년에 3분위에 속하는 가구 중 26.6%는 2004년에는 2분위 이하로 떨어졌으며, 2년 후에 2005년에는 31%가 2분위 이하로 떨어졌으며, 3년 후인 2006년에는 34.4%가 2분위 이하로 떨어졌으며, 4년 후인 2007년에는 32.6%가 2분위 이하로 떨어졌다. 3분위의 농가들이 2분위 이하로 떨어질 확률이 4년차까지 증가하는 것으로 나타났다.

<부표 2-2> 5분위 분포상 변화 : 5년간의 변화(단위:가구,%)

(1) 2003년->2004년

구분	1	2	3	4	5	합계
1	244	117	61	32	30	484
	50.41	24.17	12.6	6.61	6.2	100

2	91	198	115	54	29	487
	18.69	40.66	23.61	11.09	5.95	100
3	47	88	178	131	63	507
	9.27	17.36	35.11	25.84	12.43	100
4	32	33	107	199	155	526
	6.08	6.27	20.34	37.83	29.47	100
5	27	19	47	121	324	538
	5.02	3.53	8.74	22.49	60.22	100
합계	441	455	508	537	601	2542
	17.35	17.9	19.98	21.13	23.64	100

(2) 2003년->2005년

구분	1	2	3	4	5	합계
1	215	125	61	47	36	484
	44.4	25.8	12.6	9.7	7.4	100.0
2	116	167	108	67	29	487
	23.8	34.3	22.2	13.8	6.0	100.0
3	48	109	172	124	54	507
	9.5	21.5	33.9	24.5	10.7	100.0
4	31	37	129	187	142	526
	5.9	7.0	24.5	35.6	27.0	100.0
5	28	33	57	127	293	538
	5.2	6.1	10.6	23.6	54.5	100.0
합계	438	471	527	552	554	2,542
	17.2	18.5	20.7	21.7	21.8	100.0

(3) 2003년->2006년

구분	1	2	3	4	5	합계
1	201	115	71	50	47	484
	41.5	23.8	14.7	10.3	9.7	100.0
2	108	178	106	68	27	487
	22.2	36.6	21.8	14.0	5.5	100.0
3	55	119	159	120	54	507
	10.9	23.5	31.4	23.7	10.7	100.0

4	36	39	120	172	159	526
	6.8	7.4	22.8	32.7	30.2	100.0
5	26	33	64	142	273	538
	4.8	6.1	11.9	26.4	50.7	100.0
합계	426	484	520	552	560	2542
	16.8	19.0	20.5	21.7	22.0	100.0

(4) 2003년->2007년

구분	1	2	3	4	5	합계
1	214	117	58	57	38	484
	44.2	24.2	12.0	11.8	7.9	100.0
2	113	162	106	74	32	487
	23.2	33.3	21.8	15.2	6.6	100.0
3	53	112	169	112	61	507
	10.5	22.1	33.3	22.1	12.0	100.0
4	41	50	118	170	147	526
	7.8	9.5	22.4	32.3	28.0	100.0
5	36	41	81	117	263	538
	6.7	7.6	15.1	21.8	48.9	100.0
합계	457	482	532	530	541	2542
	18.0	19.0	20.9	20.9	21.3	100.0

<부표 2-3>은 연도별로 인접 연도간에 주업, 겸업의 변화를 보여주고 있다. 2004년의 경우 직전 1년 동안 주업을 계속하는 비중은 70.3%였으며, 1종 겸업을 계속하는 비중은 57.3%, 2종 겸업을 계속하는 비중은 81.6%였다. 2005년의 경우 직전 1년 동안 주업을 계속하는 비중은 84.4%였으며, 1종 겸업을 계속하는 비중은 60.2%, 2종 겸업을 계속하는 비중은 83.0%였다. 2006년의 경우 직전 1년동안 주업을 계속하는 비중은 79.7%였으며, 1종 겸업을 계속하는 비중은 65.2%, 2종 겸업을 계속하는 비중은 84.3%였다. 2007년의 경우 직전 1년동안 주업을 계속하는 비중은 82.3%였으며, 1종 겸업을 계속하는 비중은 62.6%, 2종 겸업을 계속하는 비중은 85.3였다. 이와 같이 최근에 올수록 주업과 겸업을 하는 농가는 각각 주업과 겸업을 계속하는 비중이 높아졌다.

주업에서 1종겸업으로 바뀌는 비중은 2004년 이전 1년 동안은 21.7%에 달했으나, 2005년 이전 1년 동안은 12.5%, 2006년 이전 1년 동안은 16.9%, 2007년 이전 1년 동안은

13.8%로 나타나 대체적으로 하락하는 추세를 보여주었다.

<부표 2-4>은 5년간 표본을 지속한 가구를 대상으로 주업, 겸업의 5년간 변화를 보여주고 있다. 2003년에 주업으로 하는 가구 중 69.6%는 1년 후인 2004년도에 주업으로 하고 있었으며, 70.3%는 2년 후인 2005년도 주업을 계속하고 있었고, 65.6%는 3년 후인 2006년에도 주업을 계속하고 있었으며, 65.2%는 4년 후인 2007년에도 주업을 계속하고 있었다. 이와 같은 현상은 인접연도간의 분석과는 달리 5년간 지속된 표본가구를 대상으로 볼 때 주업을 지속하는 가구의 비중이 지속적으로 하락하고 있음을 알 수 있다.

한편, 2003년도에 1종 부업을 하던 가구의 45.0%만이 4년 후인 2007년에도 1종 부업을 하고 있었으며, 2003년에 2종 부업을 하던 가구의 71.3%가 4년 후인 2007년에도 2종 부업을 하고 있었다. 이와 같은 결과는 2종 부업을 하는 가구가 2종 부업을 계속할 확률이 주업이나 1종 부업을 계속할 확률보다 높음을 보여주는 것이다.

<부표 2-3> 주겸업별의 변화 : 인접연도간 분석(단위:가구,%)

(1) 2003->2004년

구분	1	2	3	합계
1	1172	361	134	1667
	70.31	21.66	8.04	100
2	174	305	53	532
	32.71	57.33	9.96	100
3	54	77	582	713
	7.57	10.8	81.63	100
합계	1400	743	769	2912
	48.08	25.52	26.41	100

(2) 2004->2005년

구분	1	2	3	합계
1	1175	174	43	1392
	84.41	12.5	3.09	100
2	211	450	87	748
	28.21	60.16	11.63	100
3	58	74	646	778
	7.46	9.51	83.03	100

합계	1444	698	776	2918
	49.49	23.92	26.59	100

(3) 2005->2006년

구분	1	2	3	합계
1	1145	243	48	1436
	79.74	16.92	3.34	100
2	172	465	76	713
	24.12	65.22	10.66	100
3	45	77	657	779
	5.78	9.88	84.34	100
합계	1362	785	781	2928
	46.52	26.81	26.67	100

(4)2006->2007년

구분	1	2	3	합계
1	1141	191	55	1387
	82.26	13.77	3.97	100
2	226	503	74	803
	28.14	62.64	9.22	100
3	62	57	693	812
	7.64	7.02	85.34	100
합계	1429	751	822	3002
	47.6	25.02	27.38	100

주 : 1은 전업, 2는 1종 겸업, 2는 2종 겸업을 의미함.

<부표 2-4> 주겸업별의 변화 : 5년간의 변화(단위:가구,%)

(1) 2003->2004년

구분	1	2	3	합계
1	1008	327	114	1449
	69.57	22.57	7.87	100
2	153	290	48	491
	31.16	59.06	9.78	100

3	37	68	497	602
	6.15	11.3	82.56	100
합계	1198	685	659	2542
	47.13	26.95	25.92	100

(2) 2003->2005년

구분	1	2	3	합계
1	1020	318	111	1449
	70.39	21.95	7.66	100
2	175	236	80	491
	35.64	48.07	16.29	100
3	55	85	462	602
	9.14	14.12	76.74	100
합계	1250	639	653	2542
	49.17	25.14	25.69	100

(3) 2003->2006년

구분	1	2	3	합계
1	950	378	121	1449
	65.56	26.09	8.35	100
2	172	244	75	491
	35.03	49.69	15.27	100
3	57	85	460	602
	9.47	14.12	76.41	100
합계	1179	707	656	2542
	46.38	27.81	25.81	100

(4) 2003->2007년

구분	1	2	3	합계
1	944	356	149	1449
	65.15	24.57	10.28	100
2	191	221	79	491
	38.9	45.01	16.09	100
3	90	83	429	602

	14.95	13.79	71.26	100
합계	1225	660	657	2542
	48.19	25.96	25.85	100

주 : 1은 전업, 2는 1종 겸업, 2는 2종 겸업을 의미함.

<부표 2-5>은 인접년도별로 주부업별 농가의 변화를 보여주고 있다. 2003년에 전문주업을 하던 농가 중에서 87.8%는 2004년에도 전문주업을 하고 있었으며, 2004년에 전문주업을 하던 농가 중에서 84.8%는 2005년에도 전문주업을 하고 있었으며, 2005년에 전문주업을 하던 농가 중에서 87.3%는 2006년에도 전문주업을 하고 있었으며, 2006년에 전문주업을 하던 농가 중에서 87.0%는 2007년에도 전문주업을 하고 있었다. 따라서 전문주업을 계속하는 비중은 80%대 후반에서 거의 변화가 없었다.

2003년에 일반주업을 하던 농가 중에서 77.9%는 2004년에도 일반주업을 하고 있었으며, 2004년에 일반주업을 하던 농가 중에서 79.8%는 2005년에도 일반주업을 하고 있었으며, 2005년에 일반주업을 하던 농가 중에서 77.5%는 2006년에도 일반주업을 하고 있었으며, 2006년에 일반주업을 하던 농가 중에서 81.7%%는 2007년에도 일반주업을 하고 있었다. 따라서 일반주업을 계속하는 비중은 전문주업에 비해서는 소폭 낮으나 80% 주위에서 거의 변화가 없었다.

2003년에 부업을 하던 농가 중에서 73.6%는 2004년에도 부업을 하고 있었으며, 2004년에 부업을 하던 농가 중에서 78.9%는 2005년에도 부업을 하고 있었으며, 2005년에 부업을 하던 농가 중에서 79.2%는 2006년에도 부업을 하고 있었으며, 2006년에 부업을 하던 농가 중에서 80.3%는 2007년에도 부업을 하고 있었다. 따라서 부업을 계속하는 비중은 일반주업과 마찬가지로 전문주업에 비해서는 소폭 낮으나 80% 주위에서 거의 변화가 없었다. 이와 같은 결과는 자급의 경우에도 거의 유사하게 나타났다.

<부표 2-6>은 동일한 분석을 5년간 표본유지를 하였던 가구를 대상으로 실시한 결과이다. 2003년을 기준으로 4년 후에도 동일한 형태를 계속하는 비중은 전문주업의 경우 81.1%였으며, 일반주업의 경우 67.2%, 부업의 경우 61.3%, 자급의 경우에는 66.7%였다.

<부표 2-5> 주, 부업 자급별의 변화 : 인접연도간 변화(단위:가구,%)

(1) 2003->2004년

구분	1	2	3	4	합계
1	870	81	38	1	990

	87.88	8.18	3.84	0.1	100
2	127	791	87	10	1015
	12.51	77.93	8.57	0.99	100
3	50	131	583	28	792
	6.31	16.54	73.61	3.54	100
4	0	6	16	93	115
	0	5.22	13.91	80.87	100
합계	1047	1009	724	132	2912
	35.95	34.65	24.86	4.53	100

(2) 2004->2005년

구분	1	2	3	4	합계
1	909	112	50	0	1071
	84.87	10.46	4.67	0	100
2	82	787	96	21	986
	8.32	79.82	9.74	2.13	100
3	36	88	579	31	734
	4.9	11.99	78.88	4.22	100
4	0	7	21	99	127
	0	5.51	16.54	77.95	100
합계	1027	994	746	151	2918
	35.2	34.06	25.57	5.17	100

(3) 2005->2006년

구분	1	2	3	4	합계
1	902	80	49	2	1033
	87.32	7.74	4.74	0.19	100
2	113	776	86	26	1001
	11.29	77.52	8.59	2.6	100
3	39	90	595	29	753
	5.18	11.95	79.02	3.85	100
4	1	8	18	114	141
	0.71	5.67	12.77	80.85	100
합계	1055	954	748	171	2928
	36.03	32.58	25.55	5.84	100

(4) 2006->2007년

구분	1	2	3	4	합계
1	936	103	37	0	1076
	86.99	9.57	3.44	0	100
2	76	799	79	22	976
	7.79	81.86	8.09	2.25	100
3	37	84	618	31	770
	4.81	10.91	80.26	4.03	100
4	1	16	19	144	180
	0.56	8.89	10.56	80	100
합계	1050	1002	753	197	3002
	34.98	33.38	25.08	6.56	100

주 : 1은 전문주업, 2는 일반주업, 3은 부업, 4는 자급을 의미함.

<부표 2-6> 주, 부업 자급별의 변화 : 5년간의 변화(단위:가구,%)

(1) 2003->2004년

구분	1	2	3	4	합계
1	804	75	32	1	912
	88.16	8.22	3.51	0.11	100
2	118	676	70	9	873
	13.52	77.43	8.02	1.03	100
3	40	107	501	22	670
	5.97	15.97	74.78	3.28	100
4	0	5	13	69	87
	0	5.75	14.94	79.31	100
합계	962	863	616	101	2542
	37.84	33.95	24.23	3.97	100

(2) 2003->2005년

구분	1	2	3	4	합계
1	760	100	51	1	912
	83.33	10.96	5.59	0.11	100
2	119	646	96	12	873

	13.63	74	11	1.37	100
3	51	121	465	33	670
	7.61	18.06	69.4	4.93	100
4	0	6	17	64	87
	0	6.9	19.54	73.56	100
합계	930	873	629	110	2542
	36.59	34.34	24.74	4.33	100

(3) 2003->2006년

구분	1	2	3	4	합계
1	756	95	57	4	912
	82.89	10.42	6.25	0.44	100
2	137	609	101	26	873
	15.69	69.76	11.57	2.98	100
3	57	119	452	42	670
	8.51	17.76	67.46	6.27	100
4	0	7	19	61	87
	0	8.05	21.84	70.11	100
합계	950	830	629	133	2542
	37.37	32.65	24.74	5.23	100

(4) 2003->2007년

구분	1	2	3	4	합계
1	740	108	61	3	912
	81.14	11.84	6.69	0.33	100
2	134	587	108	44	873
	15.35	67.24	12.37	5.04	100
3	53	152	411	54	670
	7.91	22.69	61.34	8.06	100
4	0	8	21	58	87
	0	9.2	24.14	66.67	100
합계	927	855	601	159	2542
	36.47	33.63	23.64	6.25	100

주 : 1은 전문주업, 2는 일반주업, 3은 부업, 4는 자급을 의미함.