

한국노동패널조사(KLIPS)의 패널이탈 분석

-비단조 이탈을 중심으로-

이 상 호*·이 혜 정**1)

<초록>

본 연구는 한국노동패널조사의 비단조 이탈에 초점을 맞추어 98표본 가구와 개인(15세 미만과 15세 이상으로 구분)의 이탈패턴을 사건계열분석과 다항로짓 분석 기법을 사용하여 비교분석하였다. 1-19차년도 가구 및 15세 이상 개인 표본의 이탈 유형은 공통적으로 4가지 집단-(1) 잠정적 이탈, (2) 확정적 이탈, (3) 일시적 이탈과 (4) 반복 이탈 집단-으로 유형화될 수 있다. 반면 15세 미만 개인은 주로 연령 집단에 따라 구획되는 것으로 나타났다.

응답률 추세 및 다항로짓 분석 결과에서 연령의 효과는 일관된 반면 성별의 효과는 서로 상이한 것으로 나타났다. 건강상태가 좋지 않을수록 잠정적 이탈 혹은 자녀의 조사 미진입 가능성을 높이는 것으로 밝혀졌다. 가구 특성과 관련하여 서울 및 수도권 지역일수록 상대적으로 모든 집단에서의 잠정적 이탈확률이 높고, 자녀의 조사 미진입 확률이 높았다. 가구 소득 계층 효과 역시 1차년도 초기 소득은 고소득 가구일수록, 1-19차년도 평균 소득은 저소득 계층으로의 하락인 경우에 가구, 개인의 잠정적 이탈 및 자녀의 조사 미진입 확률을 높이는 것으로 확인되었다. 자녀의 패널이탈에 있어서 부모의 특성이 미치는 효과 역시 대체로 유의미한 것으로 나타났다. 가구 단위의 응답이력이 개인의 응답이력에 미치는 효과, 부모의 응답이력이 자녀세대에 미치는 효과 역시 통계적으로 유의미하게 나타났다.

1. 서론

한국의 종단조사는 지난 20여 년 간 사회과학분야 실증연구의 가장 중요한 발전 중 하나라고 볼 수 있다. 그리고 이러한 발전의 선두에는 1998년에 시작된 이래 20년 동안 이어져온 한국노동패널조사(Korean Labor and Income Panel Study, 이하 KLIPS)가 자리매김하고 있다. 종단조사는 대규모 표본을 바탕으로 다양한 항목들을 조사할 수 있는 횡단면 조사의 장점과, 조사대상을 시간의 흐름에 따라 추적·조사함으로써 시계열 조사의 장점을 모두 갖고 있다. 뿐만 아니라 분석대상의 관측되지 않은 특성을 통제하거나, 장기적인 생애사를 분석할 수도 있다. 그럼에도 불구하고 복잡한 표본설계 과정과 조사관리 측면에서의 금전적-비금전적 비용, 가중치나 임putation과 같은 통계적 처리 과정, 다양한 비표본오차의 발생 가능성, 복잡한 자료 구조에 기인한 분석측면에서의 어려움 등 극복해야 할 과제들도 만만찮게 발생한다.

1) * 한국고용정보원 부연구위원(cclg007@keis.or.kr)

** 한국노동연구원 노동패널팀 전문위원(snp625@kli.re.kr)

이 중에서도 패널이탈(panel attrition)은 가장 중요한 이슈 중 하나이다. 패널이탈은 조사가 반복됨에 따라서 발생하는 특수한 형태의 단위무응답으로, 표집된 조사단위(개인, 가구, 학교, 사업체 등)가 조사에 참여하지 않음으로 발생하는 자료의 ‘영구적인 손실(permanent loss)’을 의미한다(Uhrg, 2008). 그러나 횡단조사의 단위무응답은 가중치를 통해서 해결될 수 있는 문제인 반면, 종단조사의 패널이탈은 보다 더 근본적인 이슈 중 하나이다. 패널이탈은 추정치의 신뢰성을 떨어뜨리고, 심각할 경우 조사의 지속가능성을 위협할 수 있기 때문이다. 더욱이 패널이탈이 무작위적이지 않고 체계적(non-ignorable attrition)일 경우 모집단 추정치의 편의를 가져온다.

이런 이유로 서구에서는 오래전부터 패널이탈의 요인을 규명하고 해결방안을 모색하려는 노력이 활발하게 이루어져왔다. 예컨대 1998년 「Journal of Human Resources」의 특집호에서 패널이탈 요인으로 개인의 인구학적 특성, 가구소득이나 노동시장지위와 같은 경제적 특성, 실업이나 이혼과 같은 개인의 사회경제적 상태 변화, 그리고, 면접원 효과와 조사체계의 효과 등이 다루어졌다(Fitzgerald et al., 1998; Lillard and Panis, 1998; Zabel, 1998 등). 또한 2009년에 발간된 「Methodology of Longitudinal Surveys」에서 패널이탈과 관련된 주요 이론적-실증적 쟁점들이 보다 체계화되었다(Lynn. et al, 2009).

국내에서도 그 동안 패널이탈에 관한 연구들이 적지 않게 축적되었다(이상호, 2005; 이상협 외, 2011; 남기성 외, 2012; 이경희 외, 2016; 홍민기·최효미, 2014; 최효미, 2016 등). 국내 연구들에서 다루고 있는 주제들 역시 위에서 언급한 쟁점들에 대부분 초점을 맞추고 있다. 다만 기존연구들 다수가 KLIPS에 집중되어 있으며 여전히 양적인 측면에서나 다양성 측면에서 부족한 것이 현실이다. 또한 많은 연구들이 초기표본이탈 연구 혹은 단조이탈(monotone attrition) 문제에 중점을 두고 있기 때문에, KLIPS와 같이 조사가 장기화됨에 따라 패널이탈 패턴과 이탈의 요인이 어떻게 달라지는지에 대해서는 상대적으로 소홀했다고 볼 수 있다.

이상의 문제의식에 기반하여 본 연구는 KLIPS의 비단조 이탈에 초점을 맞추어 패널이탈의 특징과 패턴을 살펴보고, 패널이탈에 영향을 미치는 요인을 분석하고자 하며 다음과 같은 점에 초점을 두고 기존 연구와 차별화를 시도한다. 비단조이탈의 경우 단조이탈에 비해 경우의 수가 매우 다양하고 복잡하기 때문에 사건계열분석(event sequence analysis) 기법을 이용하여 패널이탈 유형화를 시도한다. 이때 가구표본과 개인표본의 이탈패턴이 어떤 차이를 보이는지 비교분석한다. 1차년도 원표본이면서도 개인조사에 포함되지 않았던 15세미만 자녀들이 조사에 진입하면서 전체적인 표본이탈에 어떤 차이를 보이는지 함께 살펴본다.

본문의 주요 내용은 다음과 같다. 제2장에서는 패널이탈의 주요 쟁점들을 중심으로 선행연구 결과들을 검토한다. 이때 KLIPS의 패널이탈과 관련된 주요 선행연구들의 특징을 살펴본 후, 패널조사가 장기화됨에 점점 중요성이 커지는 비단조 이탈 문제와 유소년 표본의 조사진입 후 패널이탈과 관련된 논의에 초점을 맞춘다. 제3장에서는 자료 및 변수의 구성과 분석방법에 대한 논의를 전개한다. 제4장에서는 패널이탈 유형화 결과 및 이탈 유형별 특징을 살펴본 후, 제5장에서는 다항로지트모형 분석을 통해 패널이탈에 영향을 미치는 요인을 분석한다. 마지막으로 제6장에서는 분석결과를 요약하고 정책적 함의와 논의 과제를 제시할 것이다.

II. 선행연구 검토

KLIPS의 패널이탈에 관한 연구들을 살펴보면 크게 3가지 주제로 나눌 수 있다.

첫째, 표본 가구 및 개인의 성별, 연령, 교육수준과 같은 인구사회학적 특성이 패널이탈에 미치는 효과이다. 선행연구들은 공통적으로 미혼이거나 기혼이라 하더라도 이혼이거나 별거상태인 경우, 학력이 높을수록 패널이탈 위험이 높은 것으로 나타나고 있다. 연령의 경우 양극화된 형태를 보이는데 연령이 젊을수록 이탈위험이 커지지만, 65세 이상인 경우에도 패널이탈 위험이 커지는 것으로 나타나고 있다(이상호, 2005; 이상협 외, 2011; 남기성 외, 2012; 홍민기·최효미, 2014; 이경희·민인식, 2016).

둘째, 경제활동상태, 가구소득과 같은 경제적 요인이 미치는 효과이다. 선행연구들에서는 공통적으로 가구주가 실업자이거나 비임금근로자인 경우 패널이탈 위험이 커지는 것으로 나타나고 있다. 또한 소득분포의 양극단에 위치하는 저소득 계층과 고소득 계층에서 패널이탈 위험이 높은 것으로 나타난다(이상호, 2005; 홍민기·최효미, 2014).

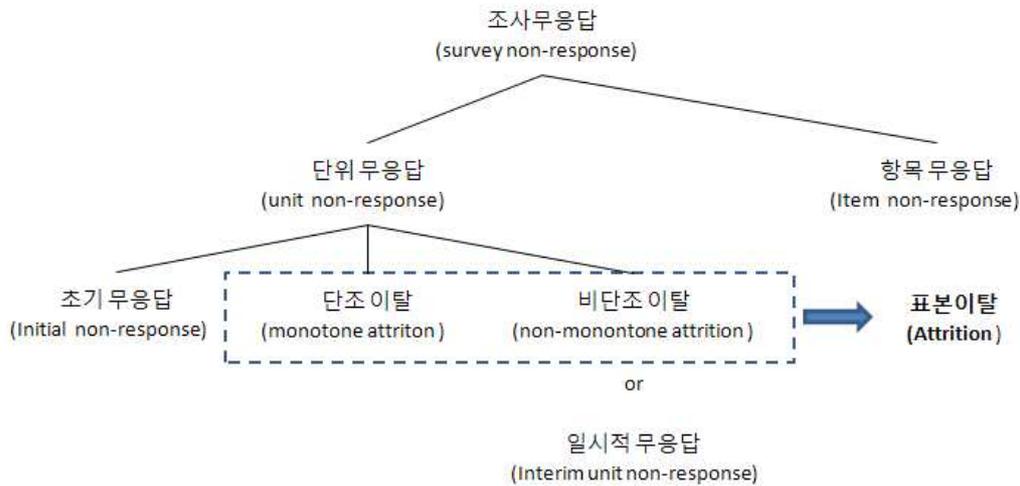
셋째, 이사와 같은 주거지의 변화, 주거의 안정성 역시도 패널이탈에 영향을 미치는 주된 요인 중 하나이다. 로짓모형과 의사결정나무 분석기법을 이용한 남기성·천영민(2012)의 연구에 따르면, 이사에 따른 주거지 변화가 패널이탈 위험을 증가시킨다. 한국의 경우 전월세와 같은 주거형태가 2년 주기로 변화하기 때문에 추적 실패를 감소시키는 조사전략이 중요하다는 점을 지적하고 있다.

넷째, 면접원 특성, CAPI의 도입, 응답자 사례금 등과 같은 조사체계의 특성 혹은 변화가 패널이탈에 미치는 효과이다. 신선옥(2008)은 면접원의 교육수준이 높고 경력이 많을수록 응답의 협조도에 긍정적인 영향을 미친다고 분석하였다. 최효미 외(2012)는 CAPI와 같은 조사방법의 변화가 패널이탈에 미치는 효과, 이른바 모드 효과를 검증하였는데, 실증분석 결과 조사방법의 변화에 따른 체계적인 응답률 차이는 발생하지 않았다. 성재민(2007)의 연구에 따르면 응답자 사례금(현금보상)이 조사 참여에 상당한 영향력을 갖는데, 특히 저소득 가구의 응답률을 높이는데 긍정적인 영향을 미치는 것으로 분석되었다.

이와 같이 그 동안 적지 않은 연구들이 KLIPS의 패널이탈에 영향을 미치는 요인들을 다양한 방법론을 이용하여 규명해왔다. 다만 지금까지의 대부분 연구들은 주로 단조이탈(monotone attrition) 문제에 중점을 두고 있기 때문에, 비단조이탈의 특성과 요인을 규명하고자 하는 노력은 상대적으로 부족했던 것으로 평가할 수 있다. 단조 이탈은 패널이탈을 한 번 발생하면 돌이킬 수 없는 비가역적인 사건으로 간주한다. 반대로 일시적으로 무응답이 발생하더라도 재복귀하거나, 이탈과 재진입을 반복하는 경우도 발생할 수 있다. 따라서 원표본에 대해 사망 등으로 인한 표본소멸을 제외하고 기한을 정하지 않고 반복 추적하는 KLIPS에서는 단조이탈 중심의 접근법이 전체 패널이탈의 특징을 규명하고 요인을 분석하는데 있어서 한계를 가질 수 밖에 없다. 특정 차수에 대해 일시적으로 단위무응답이 발생하더라도 이후의 차수에 다시 재복귀가 가능하기 때문이다. 설령 사망이

나 이민의 경우에도 조사원의 접촉 자체가 불가능한 경우가 많기 때문에 상당한 기간이 지난 이후에 확인되는 경우도 적지 않다(Burkam and Valerie, 1998; Uhrig, 2008, Lugtig, 2014)). 즉 이상적으로는 [그림 1]과 같이 명확하게 패널이탈이 정의된다 하더라도 현실에서의 패널이탈은 훨씬 더 가변적이고 복잡한 양상을 보이게 된다. 이러한 문제와 관련하여 이상호(2013)는 KLIPS 1-5차 원표본 가구 중 단조이탈은 19%에 불과하지만, 재진입을 1회 이상 경험한 비단조이탈 표본은 35%에 이르는 것으로 나타났다.

[그림 1] 조사 무응답(survey non-response)의 구성



더구나 조사차수가 누적될수록 비단조이탈의 비중은 점점 더 증가하며 이탈 패턴 역시 더욱 복잡해지게 된다. 뒤에서 다루겠지만 이 연구에서는 사건계열분석 접근법을 이용하여 비단조 이탈 패턴 몇 가지 유형으로 구분함으로써 이러한 문제에 대응할 것이다²⁾.

비단조 이탈을 분석함에 있어서 두 번째로 고려해야 할 요소는 15세 미만 자녀의 조사 진입에 따른 자료의 대표성과 비단조 이탈 패턴의 변화이다. KLIPS에서 15세 미만 자녀들은 원표본이면서도 15세 이후에 조사에 진입하기 때문에 가구의 패널이탈 혹은 부모의 패널이탈이 자녀에게 누적적인 영향을 미칠 수 있다. 이 주제와 관련하여 해외의 몇몇 연구자들은 패널이탈 효과의 세대 간 연계 문제로 인식하고 부모세대의 이탈이 자녀의 조사 진입 이후의 응답에 어떤 영향을 미치는지를 분석하고 있다(Fitzgerald, 1998b; Fitzgerald, 2011; Schoeni and Wiemers, 2015). 이상의 문제의식에 기초하여 대부분 15세 이상 개인을 대상으로 패널이탈 분석을 시도한 기존 연구와 달리 이 연구에서는 15세 미만 자녀의 패널이탈 패턴 분석과 이탈 요인 분석을 별도로 다룰 것이다.

2) 사건계열분석을 이용한 패널이탈 분석은 아동패널조사의 패널이탈 분석을 실시한 최효미(2016)의 연구에서 활용된 바 있다.

III. 분석 모형 및 자료

1. 분석 모형

분석 방법은 1단계로 사건계열분석(Event Sequence Analysis) 기법을 이용하여 응답력을 유형화한 후, 2단계로 다항로짓 분석을 이용하여 각각의 응답 유형에 영향을 미치는 요인을 규명한다.

사건계열분석은 노동시장 이행과정에서 나타나는 경제활동상태의 변화, 혼인상태의 변화, 주거형태의 변화 등과 같이 일련의 사건들이 시간의 흐름에 따라 연쇄되어 나타나는 계열자료(sequence data)의 분석에 유용한 접근법이다. 종단자료의 응답력 역시도 사건계열의 하나라고 볼 수 있다. 이와 같은 계열자료를 기반으로 여러 시점에서 관찰되는 사건의 연쇄를 하나의 단위로 보고 발생 순서에 따른 특징과 구조를 파악하기 위해 개발된 것이 사건계열분석이다(장지연 외, 2008).

사건계열분석의 핵심은 계열적으로 관찰되는 일련의 사건들을 총체적(holistic)으로 파악한 후 이를 구조화하는데 있다. 따라서 패널이탈분석에서 사건계열분석을 시도할 경우 특정 차수의 응답과 무응답을 각각의 단위로 분석하는 것이 아니라, 전체 응답력을 하나의 분석단위로 분석하게 된다. KLIPS 1-19차 조사자료의 응답이력(응답 =1, 무응답 = 0)은 $2^{19}=524,288$ 개의 조합이 가능하기 때문에, 응답이력의 패턴에 따라 이를 분류하고 유형화하는 과정을 필요로 한다. 따라서 통상적인 통계이론에 근거하여 가설을 검증하는 것이 아니라, 분석대상에 내재된 특징들을 파악하기 위한 탐색적 접근을 취하기 때문에 자료의 시각화 역시 중요한 분석 과정이다.

구체적으로 사건계열분석의 절차는 다음과 같다. 첫 번째 단계는 사건연쇄를 구성하는 요소를 정의한 후 분석대상이 되는 사건연쇄들의 기본적인 특성을 살펴보는 과정으로 그래프를 이용하여 각 사건연쇄를 선(line)으로 나타낸다. 두 번째 단계는 ‘최적일치법’을 이용하여 사건연쇄들을 서로 비교한다³⁾. 세 번째 단계는 앞에서 계산된 두 사건연쇄들 간의 거리(pairwise distance)를 매트릭스로 구성한 후, 군집분석(clustering)을 통해 관찰된 자료들을 몇의 집단으로 묶어내는 과정이다.

사건계열분석을 통해 패널이탈 패턴이 몇 가지로 유형으로 범주화되면, 다항로짓모형을 이용하여 분석할 수 있다. 명목형 반응변수 Y 가 범주 $1, 2, \dots, J$ 를 갖는다고 하면, 이에 대응되는 반응확률은 $\{\pi_1, \pi_2, \dots, \pi_J\}$ 이고 $\sum_j \pi_j = 1$ 을 만족한다. n 개의 관측치를 J 개의 범주에 할당시키는 표본모형은 다항분포(multinomial distribution)을 따르며, 예측변수 x 를 가진 다항로짓모형은 다음과 같다(Agresti, 1990). 여기서 $\log(\pi_j/\pi_J)$ 는 기준 범주 로짓이라고 하는데 임의로 마지막 J 번째 범주를 기준으로 잡고 나머지 반응범주 각각을 고려한 것이다. 분석결과는 기준 범주에 대한 한계효과를 산출하여 제시한다.

$$\log\left(\frac{\pi_j}{\pi_J}\right) = \alpha_j + \beta_j x, \quad j = 1, 2, \dots, J-1 \quad (1)$$

3) 거리를 계산하는 자세한 방법에 대해서는 Macindoe & Abbott(2004)와 Brzinsky-Fay et al(2006) 참조.

2. 자료 및 변수

분석대상 표본은 1차년도(1998년) 원표본 가구 및 개인을 대상으로 한다. 1차년도 가구 원표본⁴⁾은 5000가구이며, 개인 원표본은 전체 17,504명이다. 개인 이때 15세 이상 개인(1차년도 응답자 13,321명 + 비응답자 534명) 및 15세 미만 가구원(3,649명)으로 구성된다. 가구 표본의 경우 정의상 1번 이상 응답이 관찰되지만, 개인 원표본의 경우 1-19차까지 1번도 응답하지 않은 사례수(57건)가 발생하였다. 이들은 분석에서 제외한다. 반면 15세 미만 가구원은 1차년도 개인조사설문의 응답 대상이 아니므로 기존 연구에서는 패널이탈 분석에서 제외되었지만, KLIPS 추적원칙 및 가중치 부여 원리 등의 측면에서 15세 이상 원표본과 동일한 기준이 적용된다. 패널조사가 장기화될수록 자녀들의 원활한 조사진입이 자료의 대표성 및 지속가능성을 높여주기 때문에 오히려 더욱 중요하게 다루어져야 한다. 다항로지트 분석에 사용되는 변수들의 기초통계량을 살펴보면 다음과 같다.

우선 가구표본의 설명변수는 3가지 범주로 구성된다(<표 1> 참조).

첫째 1차 조사 당시 가구주의 인구사회적 특성 변수들로 성별, 연령, 교육수준, 혼인상태가 포함된다.

둘째, 1차년도 가구특성은 거주지역 범주, 가구총소득 범주, 자가여부, 가구원 수 등으로 구성된다.

셋째, 가구의 상태변화와 관련된 변수들로 이사횟수와 평균 소득이 해당된다. 1-19차까지의 가구의 응답 유형을 결정하는데 있어서 초기 상태가 미치는 효과뿐만 아니라 시간에 따른 상태변화 역시 영향을 미칠 수 있다. 다만 사건계열분석 자체가 응답패턴을 시간에 따라 독립적인 단위로 세분화하는 것이 아니라 유형화하는 과정에서 횡단면 자료로 바뀌기 때문에 가구의 상태변화와 관련된 척도들은 각 조사시점에서 파악된 상태들을 기준으로 1-19차까지의 상태들의 출현 횟수를 결합하는 방식으로 구성하였다. 평균 소득의 경우 상태변화는 아니지만, 1차년도 초기 소득 효과가 통제된 상태에서 2차년도의 이후의 평균적인 경제적 계층변화를 의미에서 크게 상태변화 변수로 분류하였다.

가구주의 평균 연령은 약 46.4세이고, 가구주의 교육수준은 고졸미만이 40.58%로 가장 높으며 다음으로 고졸이 37.63%이고 전문대졸 이상이 21.8%를 차지하였다. 가구주의 혼인상태는 대부분이 배우자가 있으며(80.71%), 별거 또는 이혼 등이 12.77%로, 미혼이 6.5%인 것으로 나타났다.

다음으로 거주지역은 광역시(30.45%)가 가장 높은 비율이고, 다음으로 서울(27.21%)과 기타 도지역(25.16%)이 비슷한 비율을 가지며 경기도가 17.18%를 차지하였다. 가구총소득은 절반 이상이 중간소득(60.52%)이고 저소득은 36.86%이며 고소득은 2.62%로 아주 낮은 편이었다. 가구원수는 약 3.6명이고 이사횟수는 약 1.6회로 나타났다.

4) 2008년 표본추가 후 표본 재설계가 이루어졌기 때문에 공식적으로는 '98표본'으로 지칭하는 것이 정확하지만 본 연구에서 1998년 표본을 기준으로 분석이 이루어지므로 편의상 '원표본'으로 사용한다.

<표 2> 가구표본 주요 변수의 기초통계량

(단위: 가구, %)

설명변수	사례	평균 (비중)	S.D (누적비중)	
1차 개인특성 (가주주)	여성더미	4,814	0.133	0.339
	연령	4,814	46.367	12.569
	연령제곱/100	4,814	23.079	12.546
	교육(고졸미만)	1,953	(40.58)	(40.58)
	고졸	1,811	(37.63)	(78.20)
	전문대이상	1,049	(21.80)	(100.00)
	혼인상태(미혼)	312	(6.51)	(6.51)
	기혼 유배우 별거/이혼등	3,867 612	(80.71) (12.77)	(87.23) (100.00)
1차 가구특성	지역(서울=1)	1,310	(27.21)	(27.21)
	광역시	1,466	(30.45)	(57.67)
	경기	827	(17.18)	(74.84)
	기타 도지역	1,211	(25.16)	(100.00)
	소득(중범위)	2,913	(60.52)	(60.52)
	저소득(<1200)	1,774	(36.86)	(97.38)
	고소득(>4800)	126	(2.62)	(100.00)
상태 변화	자가더미	4,814	0.559	0.497
	가구원 수	4,814	3.550	1.335
	이사횟수	4,814	1.561	1.664
	평균소득(중범위)	3,206	(66.60)	(66.60)
	저소득(<1200)	851	(17.68)	(84.28)
고소득(>4800)	757	(15.72)	(100.00)	

주 : 연속 변수 및 더미 변수는 왼쪽부터 사례수, 평균과 표준편차 순으로 제시되었으며, 3개이상의 범주형 변수는 왼쪽부터 빈도와 비중, 누적비중 순으로 제시되었음
 자료 : KLIPS 1~19차(1998~2016년 조사) 학술대회용 자료

<표 2>에 제시된 15세 이상 개인표본의 설명변수는 4가지 범주로 구성된다.

첫째 1차 조사 당시 개인의 인구사회적 특성 변수들로 성별, 연령, 교육수준, 혼인상태, 건강상태 등이 포함된다. 단 개인의 건강상태는 2차년도 가구자료로부터 산출된 것이다. 가구표본의 경우 가구당 가주주 1인의 인구사회학적 특성만 포함되지만, 여기서는 15세 이상 모든 가구원이 각각의 분석단위가 된다.

둘째, 1차년도 가구특성은 가구자료와 동일하게 거주지역 범주, 가구총소득 범주, 자가여부, 가구원 수 등으로 구성된다. 이때 각각의 개인들이 동일한 가구에 포함되었을 경우, 동일한 가구 특성을 할당받게 된다.

셋째, 개인 및 가구의 상태변화와 관련된 변수들이다. 1-19차 까지의 개인의 응답 유형을 결정하는데 있어서 초기 상태가 미치는 효과 뿐만 아니라 시간에 따른 상태변화 역시 영향을 미칠 수 있다. 다만 사건계열분석 자체가 응답패턴을 시간에 따라 독립적인 단위로 세분화하는 것이 아니라 유형화하는 과정에서 횡단면 자료로 바뀌기 때문에 개인의 상태변화와 관련된 척도들은 각 조사시

점에서 파악된 상태들을 기준으로 1-19차까지의 상태들의 출현 횟수를 결합하는 방식으로 구성하였다. 평균 소득의 경우 상태변화는 아니지만, 1차년도 초기 소득 효과가 통제된 상태에서 2차년도의 이후의 평균적인 경제적 계층변화를 의미에서 크게 상태변화 변수로 분류하였다.

마지막으로 가구의 응답이력이 미치는 효과이다. KLIPS와 같은 가구패널조사는 주로 가구응답률을 중심으로 표본을 관리한다. 만일 개인의 응답이력이 대부분 가구의 응답이력에 의해 결정된다면 현재와 같은 가구 중심의 표본관리 정책이 큰 틀에서 효율적이고 효과적인 전략이라고 볼 수 있을 것이다. 그러나 가구의 응답이력이 미치는 효과가 통계적으로 유의미하지 않거나, 통계적으로 유의미하더라도 앞서의 변수들의 효과가 여전히 유의하다면 개인단위의 표본관리 전략이 강화될 필요가 있다는 것을 의미한다.

15세 미만 개인표본의 설명변수는 15세 이상 표본과 비슷한 구조를 갖지만 부모의 특성과 관련된 변수들이 추가되면서 세대간 이전 효과를 측정하는데 초점을 맞춘다.

우선 자녀 당사자의 특성은 성별과 연령, 응답자 본인의 건강상태로 제한된다. 15세 미만은 연령에 따라 교육 수준이 의무적으로 구획되므로 제외하였다. 부모의 특성은 1차년도 당시의 성별, 연령, 혼인상태, 경제활동상태 및 상태변화와 관련된 변수들로 구성하였다. 부모의 정보가 모두 존재하는 경우, 가구주가 취업자가 아니면서 가구주의 배우자가 취업자인 경우에는 가구주의 배우자 정보로 대체하였으며, 그 외 나머지의 경우는 가구주 정보를 사용하였다.

마지막으로 응답이력은 가구의 응답이력이 아닌 부모의 응답이력을 사용하였다. 15세 미만 자녀의 경우 부모세대에서의 가구분가와 자녀 본인의 경제적 독립에 따른 분가 등이 상대적으로 빈번하게 발생하기 때문에 원표본 가구의 응답이력 효과를 매칭하는 것이 적절하지 않다고 판단하였다. 부모의 응답이력을 사용함에 있어서 상대적으로 응답이력이 긴 부모의 정보를 선택하였다.

<표 2> 개인표본 주요 변수의 기초통계량

(단위: 개인, %)

(1) 1차년도 15세 이상 개인 표본					(2) 1차년도 15세 미만 개인 표본				
설명변수		사례	평균 (비중)	S.D (누적비중)	설명변수		사례	평균 (비중)	S.D (누적비중)
1차 개인 특성	여성 더미	12,567	0.524	0.499	1차 자녀 특성	여성 더미	3,651	0.462	0.499
	연령	12,567	38.318	15.260		연령	3,651	7.239	4.230
	연령제곱/100	12,567	17.011	13.245		연령제곱/100	3,651	0.703	0.634
	교육(고졸미만)	5,168	(41.13)	(41.13)		비건강 더미	3,651	0.145	0.352
	고졸	5,161	(41.07)	(82.20)	1차 가구 특성	지역(서울=1)	954	(26.1)	(26.1)
	전문대 이상	2,237	(17.80)	(100.00)		광역시	1,061	(29.1)	(55.2)
	혼인상태(미혼)	447	(3.56)	(3.56)		경기	757	(20.7)	(75.9)
	기혼 유배우	10,778	(85.95)	(89.51)		기타 도지역	879	(24.1)	(100.0)
	별거/이혼 등	1,315	(10.49)	(100.00)		소득(중범위)	2,653	(72.7)	(72.7)
비건강 더미	12,567	0.226	0.419	저소득(<1200)	882	(24.2)	(96.8)		
1차 가구 특성	지역(서울=1)	3,405	(27.09)	(27.09)	고소득(>4800)	116	(3.2)	(100.0)	
	광역시	3,976	(31.64)	(58.73)	자가 더미	3,651	0.486	0.500	
	경기	2,061	(16.40)	(75.13)	가구원 수	3,651	4.271	0.936	
	기타 도지역	3,125	(24.87)	(100.00)	1차 부모 특성	여성 더미	3,651	0.076	0.266
	소득(중범위)	8,105	(64.50)	(64.50)		연령	3,646	39.176	7.904
	저소득(<1200)	4,075	(32.43)	(96.94)		연령제곱/100	3,646	15.972	7.467
	고소득(>4800)	385	(3.06)	(100.00)		부모 별거/이혼	3,651	0.052	0.222
	자가 더미	12,567	0.614	0.487		부모 실업상태	3,651	0.098	0.298
	가구원 수	12,567	3.952	1.287	상태 변화	부모 실업횟수	3,646	1.551	3.229
실업횟수	12,567	5.351	5.743	부모 별거이혼		3,646	0.148	0.355	
별거이혼	12,567	0.185	0.388	가구 이사횟수		3,651	1.739	1.657	
이사횟수	12,567	1.430	1.590	평균소득(중범위)		2,614	(71.6)	(71.6)	
평균소득(중범위)	8,666	(68.96)	(68.96)	저소득(<1200)		209	(5.7)	(77.3)	
저소득(<1200)	1,791	(14.25)	(83.21)	고소득(>4800)		828	(22.7)	(100.0)	
가구 응답 이력	고소득(>4800)	2,110	(16.79)	(100.00)	부모 응답 이력	부모 응답횟수	3,651	13.534	6.661
	가구 응답횟수	12,567	14.649	5.906		계속 응답(기준)	1,640	(44.9)	(44.9)
	계속 응답(기준)	1,492	(8.9)	(8.9)		잠정적 이탈	690	(18.9)	(63.8)
	잠정적 이탈	2,449	(14.7)	(23.6)		반복 이탈	925	(25.3)	(89.2)
	반복 이탈	4,874	(29.2)	(52.8)		일시적 이탈	383	(10.5)	(99.6)
	일시적 이탈	396	(2.4)	(55.2)		확정적 이탈	13	(0.4)	(100.0)
	확정적 이탈	7,484	(44.8)	(100.0)					

주 : 연속 변수 및 더미 변수는 왼쪽부터 사례수, 평균과 표준편차 순으로 제시되었으며, 3개이상의 범주형 변수는 왼쪽부터 빈도와 비중, 누적비중 순으로 제시되었음
 자료 : KLIPS 1~19차(1998~2016년 조사) 학술대회용 자료

IV. 패널이탈 유형화⁵⁾

1. 응답률 추이

[그림 2]는 KLIPS 가구 및 개인의 응답률 추이를 제시하고 있는데, 다음과 같은 몇 가지 특징을 파악할 수 있다.

첫째, 그림의 (a)에서 나타나듯이, KLIPS의 가구응답률은 크게 3가지 국면으로 구분할 수 있다. 우선 초기 4년동안 가파른 하락세를 보여 4차년도 77.3% 까지 응답률이 하락한다. 그러나 5차년도부터 9차까지는 표본이 1%p 이내의 이탈률을 보이며 동일한 수준을 유지하다가 10차년도부터 매년 약 1%p 수준으로 감소하여 19차조사에서는 67.9%를 기록한다. 조사 초기 시행착오를 겪으면서 초기 패널이탈을 최소화하는데 성공적이지 못했지만, 5차년도 이후 기존응답 가구를 유지하면서 이탈가구를 설득하여 재진입시키는데 성공했기 때문이다. 그러나 10차년도 이후 고령화된 표본들이 사망하면서 가구 소멸이 증가하고 이사 등으로 인해 추적에 실패한 가구의 재진입이 어려워지면서 표본의 자연감소가 진행되는 상황으로 요약할 수 있다.

둘째, 개인조사 응답률은 전체적으로 가구 응답률과 비슷한 추이를 보이지만, 가구 응답률보다 체계적으로 낮다. 그림의 (a)에서 1차년도에 응답한 개인 원표본(13,321명)의 응답률은 1-5차년도까지 가파른 하락세, 5-10차년도까지의 표본 유지국면, 11-17차까지의 1-2% 수준의 선형적 감소국면 등으로 구분 된다. 다만 추세적으로 가구응답률과의 격차가 매년 벌어져 19차년도 기준 개인 응답률이 56.2%로 가구 응답률보다 11.7%p가 낮다. 다만 개인응답률을 1차년도 원표본 전체로 확대하면 가구 응답률과의 격차가 소폭이나마 줄어드는 추세이다. 1차년도 당시 15세 미만이었던 자녀들이 성장하여 15세 이상 개인조사대상에 포함되면서 개인조사 응답률을 견인하는 것이다.

셋째, 개인표본 응답률은 여성보다 남성이 체계적으로 낮다. 그림 (b)의 성별 응답률 추이를 살펴보면 여성 표본의 경우 1차년도 78.8%에서 19차조사 55.7%로 하락하였다. 반면 남성 표본의 경우 1차년도 73.4%의 응답률을 기록하였다가 19차조사 51.3%까지 하락하였다. 다만 남녀간 응답률 격차는 대체로 약 5%p 이내로 일정한 수준을 유지하였다.

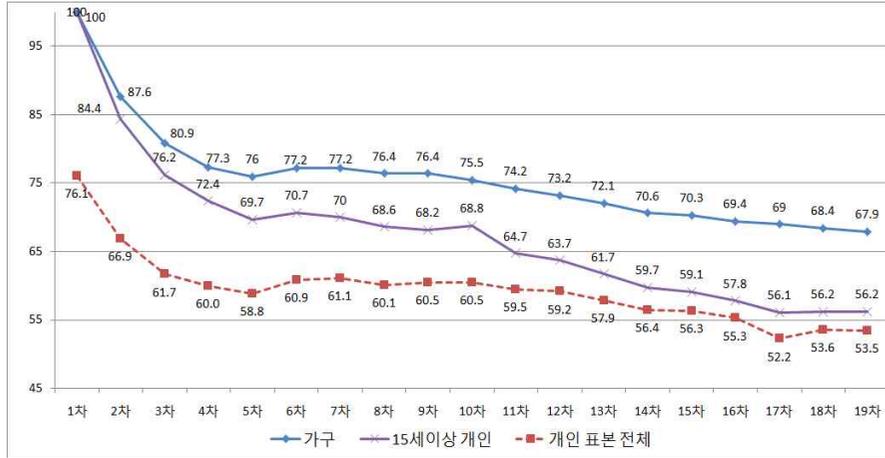
넷째, 개인 응답률은 연령에 따라 체계적인 편차를 보인다. 특히 그림 (c)의 15세 미만 응답률은 15세 미만 응답자들이 15세에 진입하는 비율이 최대 70% 초반에 그치며 연령이 낮을수록 수준이 점차 낮아진다는 점을 확인할 수 있다. 즉 자녀의 연령이 어릴수록 그만큼 개인조사에 진입할 가능성이 낮아지기 때문에 더욱 많은 표본유지 노력이 필요하다는 것을 의미한다. 부록의 [부도 1]은 1세별 응답률을 남녀로 구분하여 보여주고 있는데, 남성의 경우 군입대로 인해 20대 초반에 20% 수준까지 응답률이 떨어지는 것을 확인할 수 있다. 마지막으로 그림 (d)의 15세 이상 집단의 경우 60세 이상 집단은 사망으로 인해 표본이 선형적으로 감소하지만, 나머지 집단들은 연령이 높을수록 응답률이 높으며 시간에 따른 감소추세 역시 완만하게 나타났다.

5) 이 장의 내용 중 가구 표본에 관한 분석결과 일부는 이혜정(2017)에 포함된 것임.

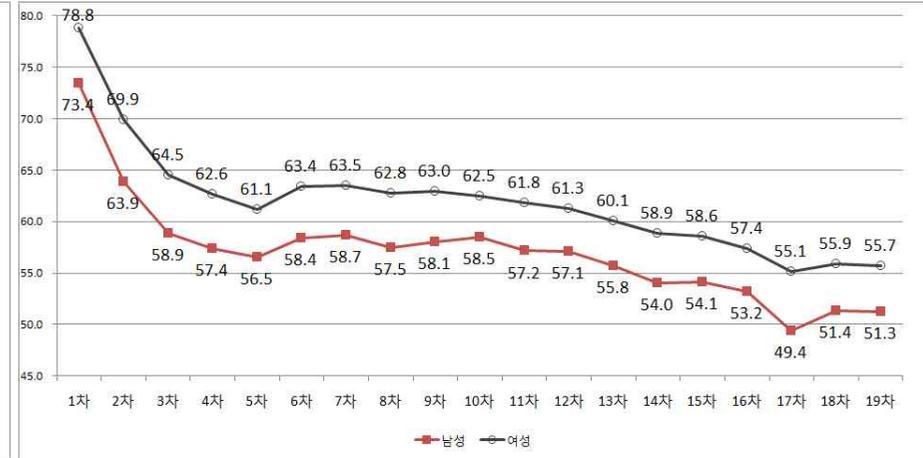
[그림 2] 원표본 가구 및 개인의 응답률 추이

(단위 : %)

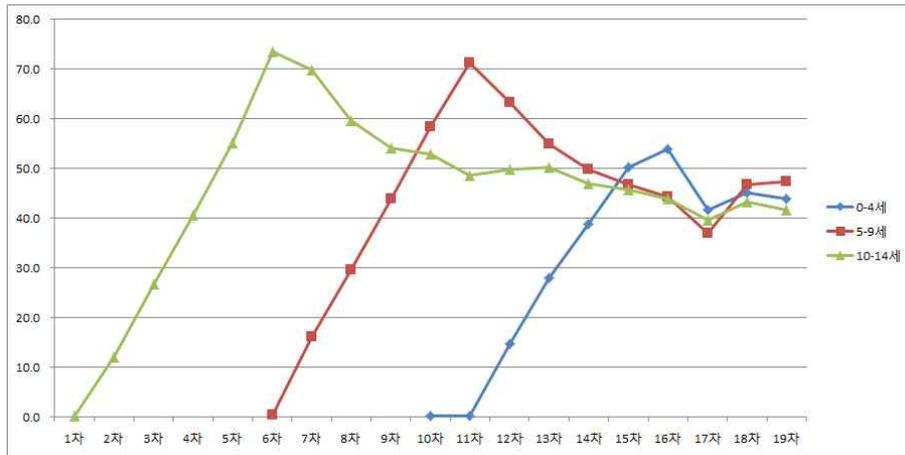
(a) 가구 및 개인 원표본 전체 응답률



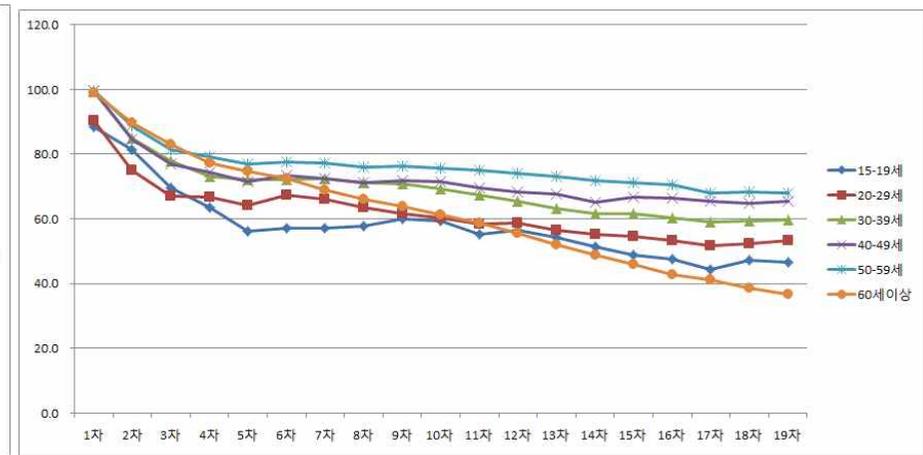
(b) 개인 원표본 성별 응답률



(c) 15세 미만 개인 5세별 응답률



(d) 15세 이상 개인 5세별 응답률



주 1) : 각 차수별 0 = 무응답, 1= 응답, 2 = 소멸
 2) : X축은 조사차수, Y축은 가구를 의미함
 자료 : KLIPS 1~19차(1998~2016년 조사) 학술대회용 자료

2. 패널이탈 유형화

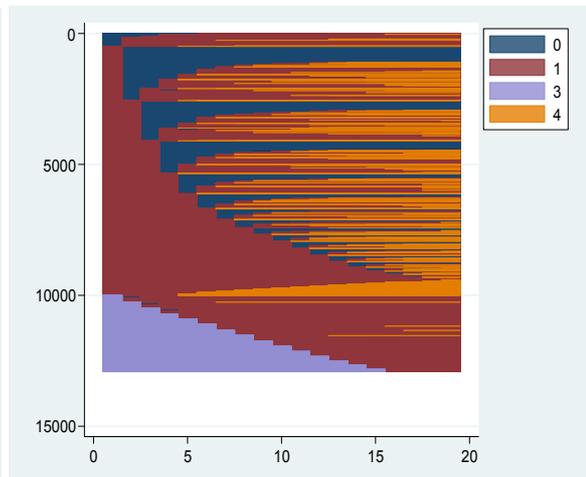
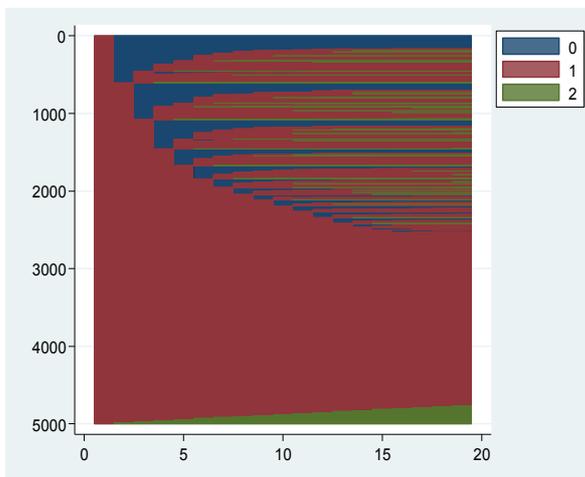
앞서 살펴본 [그림 2]가 평균적인 응답률을 보여주는 반면, 아래의 [그림 3]은 각 개별 표본의 응답이력을 시간 순서에 따라 나열한 후 그림으로 표현한 것이다.

우선 [그림 3]의 (a)의 가구표본은 응답, 무응답, 소멸 등 3가지의 상태로 구분하여 응답이력을 분석한 것이다. 1-19차 기간 동안 관측된 응답이력은 총862가지 형태이며, 가장 주된 형태는 1~19차를 모두 응답한 경우로 43.4%를 차지하고, 1차만 응답하고 나머지 차수는 모두 응답하지 않은 형태가 167가구(3.3%)로 다음으로 많았다. (b)에서 제시된 개인표본은 응답, 무응답, 개인조사 신규 진입, 사망 등 4가지 상태로 구분하여 응답이력을 분석한 것이다. 개인표본의 경우 가구표본보다 훨씬 방대한 12,923개의 응답이력이 관찰되었다.

[그림 3] 1~19차 가구 및 개인의 응답이력

(a) 가구 표본

(b) 개인 표본



주 1) : 각 차수별 0 = 무응답, 1= 응답, 2 = 소멸
2) : X축은 조사차수, Y축은 가구를 의미함

주 3) : 각 차수별 0=무응답, 1=응답, 3=미진입, 4=사망
4) : X축은 조사차수, Y축은 개인을 의미함

자료 : KLIPS 1~19차(1998~2016년 조사) 학술대회용 자료

다음으로 위에서 정의된 응답상태별 이력에 대한 사건계열분석을 실시하였다. 가구표본의 경우 1-19차를 모두 응답한 2,169가구는 제외하고, 한번 이탈 이탈이 관찰된 2,831가구를 분석 대상으로 선정하였다. 반면 개인표본의 경우 15세 이상 개인표본과 15세 미만 자녀표본으로 구분하여 사건계열분석을 실시하였다⁶⁾. 15세 이상 표본의 경우 가구표본과 마찬가지로 1-19차 모두 응답한 표본은 유형화 과정에서 제외하였으며, 15세미만 자녀표본은 정의상 1-19차 계속응답이 발생할 수 없

6) 시뮬레이션 과정에서 15-64세, 65세 이상 집단 등 다양한 그룹으로 나누어 사건계열분석을 실시(부록의 부도 2와 부도 3 참조)하였으나, 최종적으로 15세를 기준으로 2개의 경우로 구분하는 것이 가장 효과적이라고 판단하였다.

기 때문에 표본 표본을 분석대상에 포함시켰다.

사건계열분석 후, 군집분석을 통해 응달이력을 유형화한 결과는 [그림 4]부터 [그림 6], <표 4>부터 <표 6>과 같다. 우선 계속응달을 제외한 가구표본 및 15세 이상 개인표본의 응달이력은 최종적으로 4가지 집단-(1) 잠정적 이탈, (2) 반복 이탈, (3) 일시적 이탈, (4) 확정적 이탈-으로 유형화되었다. 이중 (1) 잠정적 이탈과 (4) 확정적 이탈은 단조 이탈에 해당되며, (2) 반복 이탈과 (3) 일시적 이탈은 비단조 이탈에 해당된다고 볼 수 있다.

<표 3> 가구표본 및 15세 이상 개인표본의 응달이력 유형화 결과

이탈 유형	특성
(1) 잠정적 이탈	- 3차 이전까지 응달하였으나 다시 복귀하는 경우는 거의 없음. 초기 이탈이나 소멸되지 않는 한 항시 재복귀 가능성은 존재함
(2) 반복 이탈	- 꾸준히 응달하다가 최초 이탈 후 진입과 재진입을 반복하는 형태이거나, 아니면 최초 이탈 후 계속 무응답 형태(재진입 없음)를 나타내고 있음
(3) 일시적 이탈	- 전반적으로 응달이력이 높은 편이고, 잠시 무응답 구간(이탈)이 존재하나 재진입 후에는 지속적인 응달 형태를 보임
(4) 확정적 이탈	- 조사 중반 전 까지 지속적인 응달이력을 보였으나 가구 소멸 혹은 개인의 경우 사망으로 인해 영구적으로 이탈이 확정된 집단임

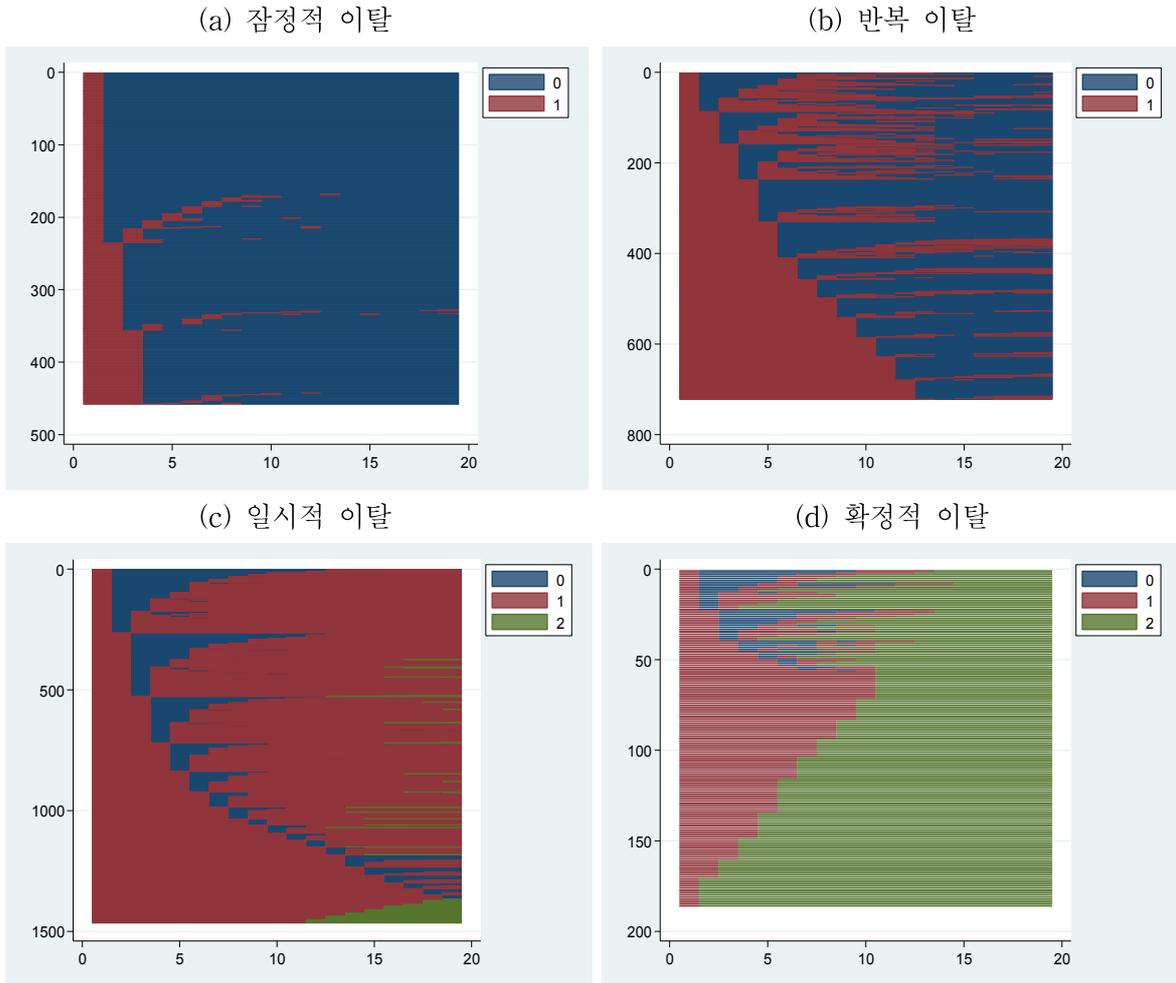
가구표본의 유형화 결과를 보다 구체적으로 살펴보면 다음과 같다([그림 4] 및 <표 4> 참조). 전체 표본을 기준으로 볼 때, 계속응달 집단은 43.4%로 집단 중에서 가장 많으며, 다음으로 일시적 이탈 집단이 29.3%를 차지했다. 가구소멸로 인한 확정적 이탈 집단은 3.7%로 가장 적은 비중을 차지했다. 각 유형별 응달 횟수를 살펴보면 일시적 이탈 15.7회 > 반복 이탈 > 8.1회 > 확정적 이탈 5.3회 > 잠정적 이탈 2회로 나타났다. 일시적 이탈 집단의 경우 최초 이탈 전 평균 응달회수가 5.9년으로 가장 길었으며, 재진입 비중도 85.9%에 이르는 것으로 나타났다. 재진입 후 평균 응달회수는 11.4년으로 응달의 충성도 역시 높았다. 반복이탈 집단은 최초 이탈 후 재진입 비중이 57.1%이며, 재진입 후 평균 응달횟수 역시 4.3회로 일시적 이탈집단 보다 응달 충성도가 절반에 미치지 못하는 것으로 확인되었다. 잠정적 이탈은 이탈전 평균 응달 횟수가 1.7회에 그쳤으며 그림에서도 초기 이탈이 대부분을 차지하였다. 재진입 비중은 24.7%에 불과하였으며, 재진입 하더라도 응달횟수가 1.3회에 그쳤다.

15세 이상 개인표본의 유형화 결과는 [그림 5] 및 <표 5>와 같다. 전체 표본을 기준으로 볼 때, 계속응달 집단은 28.5%로 가구표본(43.4%)보다 계속응달 비중이 14.9%p 낮았다. 일시적 이탈 집단은 29.1%로 가구표본과 거의 비슷한 수준이었으며, 잠정적 이탈은 24.2%로 가구표본보다 15%p 높은 수준을 기록했다. 반복 이탈과 확정적 이탈은 각각 12.5%와 5.6%로 가구표본과 비슷한 수준을 보였다. 각 유형별 응달 횟수를 살펴보면 일시적 이탈 15.6회 > 반복 이탈 > 11.7회 > 확정적 이

탈 8.5회 > 잠정적 이탈 3.9회로 가구표본보다 전반적인 응답회수는 더 많은 것으로 나타났다. 일시적 이탈 집단의 경우 최초 이탈 전 평균 응답회수가 4.8년이었으며, 재진입 비중은 97.0%를 차지했다. 재진입 후 평균 응답회수는 10.74년으로 응답의 충성도 역시 높았다. 반복이탈 집단은 최초 이탈 후 재진입 비중이 60.3%이며, 재진입 후 평균 응답횟수 역시 6.1회로 일시적 이탈집단 보다 응답 충성도가 절반을 약간 넘는 것으로 확인되었다. 잠정적 이탈은 이탈전 평균 응답 횟수가 2.5회에 그쳤으며 그림에서도 초기 이탈이 대부분을 차지하였다. 재진입 비중은 34.8%에 불과하였으며, 재진입 하더라도 응답횟수가 2.4회에 그쳤다. 이상의 이탈 유형 중 확정적 이탈은 가구소멸과 개인의 사망 등에 기인한 것이므로 조사정책적 개입이 불가능한 유형이다. 따라서 다항로지분석에서는 확정적 이탈을 제외하고 분석할 것이다.

반면 15세 미만 자녀표본의 경우, 주로 연령 집단에 따른 진입 시기가 주로 집단 구분을 결정하는 것으로 나타났다. 전체 표본에서 조사 상대적으로 10차년도 이전에 진입한 초기진입자는 크게 일시적 이탈과 잠정적 이탈로 양분되었다. 초기진입-일시적 이탈은 25.2%, 초기진입-잠정적 이탈은 16.2%를 차지하였다. 중기 및 후기 진입자는 각각 22%와 16.8%를 차지하였다. 마지막으로 사건 계열분석에 포함되지 않았지만 조사에 한번도 참여하지 않은 미진입자는 19.8%를 차지하였다. 최종적으로 다항로지분 분석을 위해 응답집단을 총 5개의 집단을 4개로 축소하였으며, 이 가운데 중기 및 후기 진입은 하나의 집단으로 통합하였다.

[그림 4] 가구 표본의 패널리탈 유형



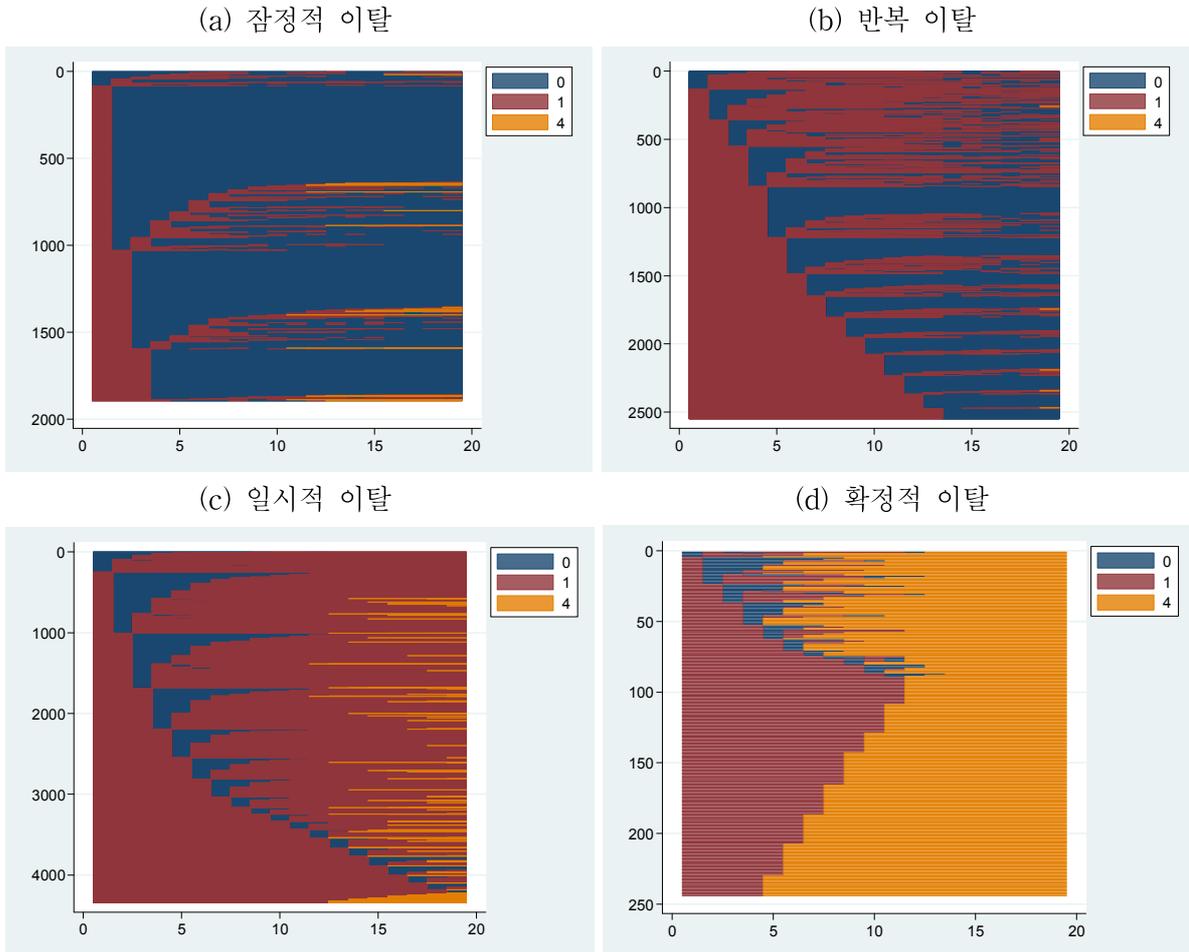
주 1) : 각 차수별 0 = 무응답, 1 = 응답, 2 = 소멸
 2) : X축은 조사차수, Y축은 가구를 의미함
 자료 : KLIPS 1~19차(1998~2016년 조사) 학술대회용 자료

<표 4> 가구표본의 응답 유형 및 집단별 평균 응답기간

	가수 수 (%)	평균 응답 횟수	최초 이탈 전 평균 응답 횟수	재진입 비중	재진입 후 평균 응답 횟수
잠정적 이탈	458 (9.2)	2.0	1.7	24.7	1.3
반복 이탈	722 (14.4)	8.1	5.6	57.1	4.3
일시적 이탈	1,465 (29.3)	15.7	5.9	85.9	11.4
확정적 이탈	186 (3.7)	5.3	4.5	30.1	2.8
계속 응답	2,169 (43.4)	19.0	-	-	-
전체	5,000 (100.0)	14.4	-	-	-

자료 : KLIPS 1~19차(1998~2016년 조사) 학술대회용 자료

[그림 5] 개인 15세 이상 표본의 패널리탈 유형



주 1) : 각 차수별 0=무응답, 1=응답, 3=미진입, 4=사망
 2) : X축은 조사차수, Y축은 가구를 의미함
 자료 : KLIPS 1~19차(1998~2016년 조사) 학술대회용 자료

<표 5> 개인 15세 이상 표본의 응답 유형 및 집단별 평균 응답기간

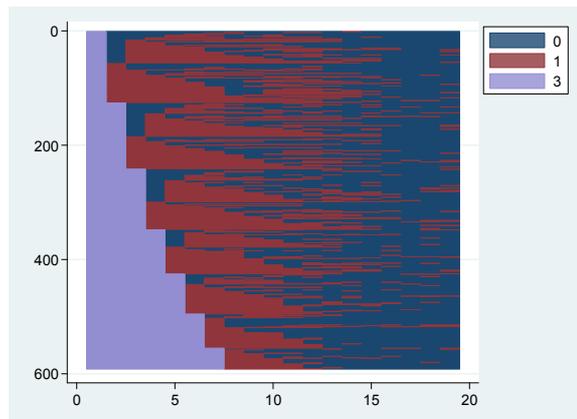
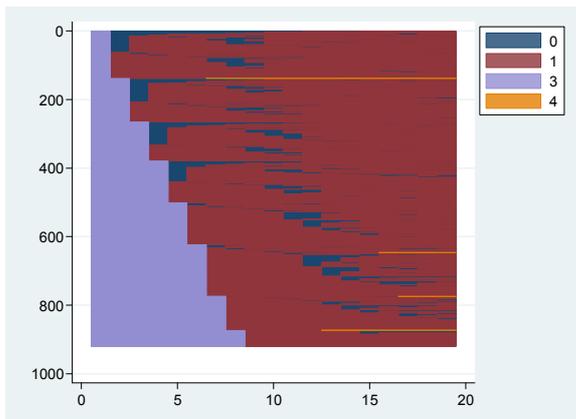
	표본 수 (%)	평균 응답 횟수	최초 이탈 전 평균 응답 횟수	재진입 비율	재진입 후 평균 응답 횟수
잠정적 이탈	3,225 (24.2)	3.9	2.5	34.8	2.4
반복 이탈	1,663 (12.5)	11.7	4.9	60.3	6.1
일시적 이탈	3,878 (29.1)	15.6	4.8	97.0	10.7
확정적 이탈	752 (5.6)	8.5	3.5	20.6	4.8
계속 응답	3,801 (28.5)	19.0	-	-	-
전체	13,319 (100.0)	12.8	-	-	-

주 : 15세 이상이면서 1번도 응답하지 하지않은 57명은 분석대상에서 제외함
 자료 : KLIPS 1~19차(1998~2016년 조사) 학술대회용 자료

[그림 6] 개인 1차년도 15세 미만 자녀표본의 패널이탈 유형

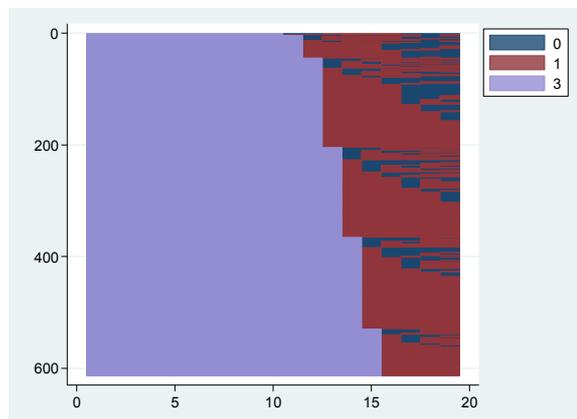
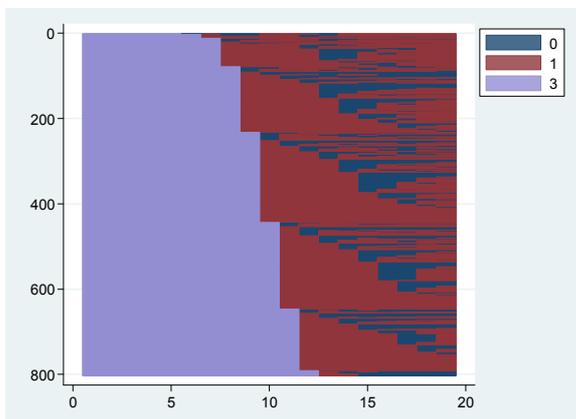
(a) 초기진입-일시적 이탈

(b) 초기진입-잠정적 이탈



(c) 중기진입

(d) 후기진입



주 1) : 각 차수별 0=무응답, 1=응답, 3=미진입, 4=사망
 2) : X축은 조사차수, Y축은 가구를 의미함
 자료 : KLIPS 1~19차(1998~2016년 조사) 학술대회용 자료

<표 6> 1차년도 15세 미만 자녀 표본의 응답 유형분류

		응답 유형 재분류				합계 (비중)
		조사 미진입	초기 진입 후 잠정적 이탈	계속 응답 및 일시적 이탈	중후기 진입	
조사 미진입		723	-	-	-	723 (19.8)
사건 계열 분석	초기진입-일시적이탈	-	-	920	-	920 (25.2)
	초기진입-잠정적이탈	-	592	-	-	592 (16.2)
	중기진입	-	-	-	803	803 (22.0)
	후기진입	-	-	-	613	613 (16.8)
합계		723	592	920	1,416	3,651 (100.0)
(비중)		(19.8)	(16.2)	(25.2)	(38.8)	(100.0)

자료 : KLIPS 1~19차(1998~2016년 조사) 학술대회용 자료

V. 다항로지 분석결과

1. 가구 표본

지금부터 다항로지 분석을 통해 패널리탈에 영향을 미치는 요인을 살펴본다. 분석 대상은 확정적 이탈 집단을 제외한 잠정적 이탈, 반복 이탈, 일시적 이탈, 계속 응답 4개의 집단에 해당되는 표본으로 구성하였다. 이때 기준 집단은 계속 응답군으로 설정하였다. 사용한 설명변수는 1차 조사시점에서의 가구주 특성(가구주의 성별, 연령, 학력과 혼인상태), 1차 조사시점에서 파악된 가구특성(거주지역, 가구소득, 주거형태와 가구원수), 그리고 1-19차 전체 조사차수에 걸친 가구의 상태변화(이사횟수와 평균 가구소득)이다.

<표 7>은 가구표본에 대한 다항로지 분석결과 도출된 계수값으로부터 편미분을 통해 한계효과를 추정한 값이다. 여기서 각 계수의 값은 기준집단인 계속 응답에 비해 해당 유형에 속할 확률을 의미하는 것으로 해석할 수 있다.

우선 1차 조사 시점 가구주의 특성이 미치는 효과를 살펴본다. 가구주가 여성인 경우 잠정적 이탈 확률이 계속응답에 비해 2.8% 증가하였지만, 반복 이탈 및 일시적 이탈은 계속응답을 선택할 확률과 통계적으로 유의미한 차이가 확인되지 않았다. 가구주의 연령이 높을수록 계속 응답 확률이 높지만 이는 증가율 자체는 체감하는 것으로 분석되었다. 교육수준별로는 가구주의 교육수준이 높을수록 이탈 확률이 높은 것으로 나타났는데, 전문대 이상 가구주의 경우 고졸미만보다 잠정적 이탈 확률이 10.3% 높았으며, 일시적 이탈 4.1%, 반복 이탈이 3.2% 순으로 높았다. 1차년도 당시 가구주가 별거이혼인 경우 반복이탈 확률이 10.5% 높았지만, 잠정적 이탈 가능성은 -4.2%로 오히려 계속응답 집단에 속할 확률보다 낮았다.

거주지역은 대체로 서울 > 경기 > 광역시 > 기타 도지역 순으로 이탈 확률이 낮은 것으로 분석되었다. 특히 기타 도지역의 경우 서울보다 반복이탈 확률 -5.9%, < 잠정적 이탈확률 -4.6% < 일시적 이탈확률 -4.4% 순으로 분석되었다. 1차년도 가구소득이 미치는 효과는 고소득가구(4,800만 원 이상)일수록 잠정적 이탈 확률이 높지만, 저소득 가구(1,200만 원 미만)는 중범위 집단보다 오히려 잠정적 이탈 확률이 낮았다. 또한 1차년도 당시 주거형태가 자가인 경우 잠정적 이탈 확률 -6.5% < 반복 이탈 확률 -3.1% 순으로 나타났으며, 일시적 이탈은 계속응답 집단과 주거형태에 따른 유의미한 통계적 차이가 확인되지 않았다.

마지막으로 상태변화와 관련된 변수들의 효과이다. 이사횟수가 많을수록 이탈 가능성을 높이지만, 잠정적 이탈이나 반복이탈은 오히려 계속응답보다 더 낮은 것으로 확인되었다. 남기성·천영민(2012)에 따르면 KLIPS의 주거형태 변화와 이사가 응답확률에 부정적인 영향을 미치는 것으로 나타나는데, 이사의 경우 주소변경을 파악할 수 있다면 일시적 이탈이 발생하더라도 재복귀가 가능하다는 것을 의미한다. 평균소득은 1차년도 가구 소득과 반대되는 영향을 미치는 것으로 나타났다. 초기소득이 동일한 가구 내에서 평균소득이 낮은 저소득 가구, 즉 소득계층 하향 이동한 가구는 상대적으로 잠정적 이탈 확률이 13% 높은 반면, 고소득 가구 집단으로 소득계층이 상향 이동한

가구는 잠정적 이탈 확률이 6.2% 낮은 것으로 확인되었다.

전체적으로 설명변수들의 통계적 영향은 잠정적 이탈과 계속응답 집단 간에 가장 뚜렷하게 나타났다. 일시적 집단의 경우 계속 응답 집단과 설명변수들의 통계적 영향이 유의미 하지 않거나, 일부변수들의 경우 통계적으로 유의미하더라도 대체로 잠정적 이탈이나 반복이탈에 비해 차이가 크지 않았다. 개별 설명변수들의 효과 역시 대체로 선행연구에서 밝혀진 것과 유사한 결과를 보였다. 다만 이어나 가구소득의 변화, 가구주의 혼인상태 변화와 같은 변수들의 효과는 단조이탈 여부를 다룬 기존 연구들과 달리 이탈 유형별로 다양한 차이를 보였다.

<표 7> 1차년도 가구표본 다항로짓 분석결과 (기준 = 계속 응답)

		집단 1 (잠정적 이탈 집단)		집단 2 (반복 이탈 집단)		집단 3 (일시적 이탈 집단)	
		dy/dx	z	dy/dx	z	dy/dx	z
1차 개인 특성 (가구주)	여성 더미	0.028	(1.79)*	-0.026	-(1.10)	-0.006	-(0.19)
	연령	-0.001	-(0.40)	-0.013	-(4.11)***	0.000	-(0.03)
	연령제곱/100	-0.002	-(0.57)	0.009	(2.78)***	0.003	(0.88)
	교육(고졸미만)						
	고졸	0.042	(4.94)***	0.008	(0.64)	0.020	(1.23)
	전문대 이상	0.103	(8.18)***	0.032	(1.99)**	0.041	(1.99)**
	혼인상태(미혼)						
	기혼 유배우 별거/이혼 등	-0.007	-(0.37)	0.044	(2.32)**	-0.056	-(1.48)
1차 가구 특성	지역(서울=1)						
	광역시	-0.030	-(2.96)***	-0.034	-(2.43)**	-0.035	-(2.04)**
	경기	-0.025	-(1.95)*	-0.020	-(1.19)	0.035	(1.68)*
	기타 도지역	-0.046	-(4.35)***	-0.059	-(4.02)***	-0.044	-(2.38)**
	소득(중범위)						
	저소득(<1200)	-0.045	-(5.62)***	-0.014	-(1.08)	0.036	(2.23)**
	고소득(>4800)	0.084	(2.25)**	0.018	(0.50)	-0.018	-(0.47)
	자가 더미	-0.065	-(7.49)***	-0.031	-(2.62)***	-0.004	-(0.24)
가구원 수	-0.003	-(0.81)	-0.001	-(0.21)	0.003	(0.52)	
상태 변화	이사횟수	-0.106	-(19.56)***	-0.014	-(3.87)***	0.058	(15.33)***
	평균소득(중범위)						
	저소득(<1200)	0.135	(7.16)***	0.032	(1.50)	-0.050	-(2.53)**
	고소득(>4800)	-0.062	-(9.34)***	-0.060	-(4.90)***	0.037	(1.83)*
표본 수		4,789					
Wald chi2		950.76***					
Log likelihood		-5128.09					

주 : * p < 0.10, ** p < 0.05, *** p < 0.01.

자료 : KLIPS 1~19차(1998~2016년 조사) 학술대회용 자료

2. 15세 이상 개인 표본

<표 8>은 15세 이상 개인 표본에 대한 다항로지 분석결과이다. 가구표본의 결과와 마찬가지로 제시된 계수값은 한계효과를 추정한 결과이다. 우선 1차 조사 시점 개인의 인구사회학적 특성이 패널리탈에 미치는 효과를 살펴본다. 우선 성별의 효과를 살펴보면, 여성인 경우 남성에 비해 잠정적 이탈 확률이 10.2% 높은 반면, 일시적 이탈확률은 9.5% 낮은 것으로 분석되었다. 연령효과는 가구표본과 마찬가지로 연령이 낮을수록 잠정적 이탈 성향이 더 높았다. 교육 수준 역시 전문대 이상의 고학력 집단의 잠정적 이탈 확률이 6.2% 높았으며, 반복이탈 확률은 2.6% 더 높게 나타났다. 혼인상태 역시 별거/이혼인 경우 반복 이탈 확률이 4.8% 잠정적 이탈 확률이 4.5% 더 높았으나, 일시적 이탈 확률은 5.8% 낮은 것으로 분석되었다. HRS 등에 대한 패널리탈 연구에서는 개인의 건강상태 좋지 않을 경우 표본이탈에 부정적인 영향을 미친다는 점을 밝힌바 있는데 KLIPS에서도 2차년도 시점에서 개인의 건강상태가 ‘좋지 않다’ 혹은 ‘매우 좋지 않다’고 응답한 경우에도 잠정적 이탈 확률이 13.1% 더 높은 것으로 분석되었다.

1차년도 가구 특성의 효과 역시 대체로 선행연구에서 밝혀진 것들과 비슷한 결과를 보였다. 서울에 비해 도 지역은 잠정적 이탈확률이 -9.4% < 광역시 -4.2% < 경기 -4.0% 낮게 나타났다. 반복이탈과 일시적 이탈의 경우 서울과 경기 간에 유의한 통계적 차이가 확인되지 않았지만, 광역시와 기타 도 지역은 이탈 확률이 모두 낮게 나타나 수도권 지역의 이탈확률이 상대적으로 높다는 점에서는 일관된 결과를 보였다. 1차년도 가구소득이 고소득 집단일수록 잠정적 이탈 확률이 높은 반면, 저소득 집단일수록 잠정적 이탈 확률이 낮았으나, 반복 이탈과 일시적 이탈에는 1차년도 가구소득의 효과가 통계적으로 유의미한 차이를 보이지 않았다. 자가인 경우 잠정적 이탈 확률 -5.6%, 반복 이탈 확률 -1.5%로 낮았으나, 일시적 이탈의 경우에는 주거형태에 따른 이탈 확률의 차이가 확인되지 않았다.

다음으로 상태변화와 관련된 변수들의 효과이다. 실업횟수, 별거/이혼, 이사횟수와 같은 상태변화의 계수값은 잠정적 이탈에서는 모두 음(-)의 값을 보였지만, 반복 이탈 및 일시적 이탈에서는 모두 (+)의 값을 보였다. 즉 이상과 같은 상태변화들은 단조 이탈에 영향을 미친다기보다 비단조 이탈 확률을 높이는데 기여하는 것으로 볼 수 있다. 평균소득 역시 앞서 가구 표본의 이탈 확률에 대한 분석과 마찬가지로 1차년도 가구 소득과 반대되는 영향을 미치는 것으로 나타났다. 초기소득이 동일한 가구 내에서 평균소득이 낮은 저소득 가구, 즉 소득계층 하향 이동한 가구는 상대적으로 잠정적 이탈 확률이 16.4% 높은 반면, 고소득 가구 집단으로 소득계층이 상향 이동한 가구는 잠정적 이탈 확률이 9.6% 낮은 것으로 확인되었다.

마지막으로 가구의 응답이력이 개인의 응답에 미치는 효과는 모두 통계적으로 유의하였다(<표 9> 참조). 그럼에도 불구하고 일부 변수들을 제외하고는 여전히 개인 수준의 특성들이 여전히 통계적 유의미한 결과를 유지하였다. 다만 가구의 형태는 분가 등으로 인해서 변경될 수 있기 때문에 각 변수들의 통계적 해석에 대해서는 보다 정교한 자료 구축이 필요할 것으로 보인다.

<표 8> 1차년도 15세 이상 개인표본 다항로짓분석 결과: 가구 응답력 제외 (기준 = 계속 응답)

		집단 1 (잠정적 이탈)		집단 2 (반복 이탈 집단)		집단 3 (일시적 이탈 집단)	
		dy/dx	z	dy/dx	z	dy/dx	z
1차 개인 특성	여성 더미	0.102	(15.43) ***	-0.009	-(1.38)	-0.095	-(11.08) ***
	연령	-0.030	-(28.29) ***	-0.012	-(13.26) ***	0.001	(0.47)
	연령제곱/100	0.037	(28.71) ***	0.013	(11.19) ***	-0.007	-(3.61) ***
	교육(고졸미만)						
	고졸	0.048	(6.42) ***	0.002	(0.32)	-0.003	-(0.31)
	전문대 이상	0.062	(6.24) ***	0.026	(2.66) ***	0.010	(0.82)
	혼인상태(미혼)						
	기혼 유배우	0.036	(2.38) **	0.046	(3.43) ***	-0.053	-(2.21) **
	별거/이혼 등	0.045	(2.56) **	0.048	(3.14) ***	-0.058	-(2.18) **
비건강 더미	0.131	(19.26) ***	-0.024	-(3.11) ***	0.097	(9.87) ***	
1차 가구 특성	지역(서울=1)						
	광역시	-0.042	-(4.83) ***	-0.022	-(2.80) ***	-0.034	-(3.26) ***
	경기	-0.040	-(4.00) ***	-0.002	-(0.20)	0.009	(0.69)
	기타 도지역	-0.094	-(10.59) ***	-0.025	-(2.93) ***	-0.026	-(2.35) **
	소득(중범위)						
	저소득(<1200)	-0.034	-(4.31) ***	0.009	(1.20)	0.008	(0.77)
	고소득(>4800)	0.086	(4.31) ***	-0.020	-(1.28)	-0.004	-(0.17)
	자가 더미	-0.056	-(7.95) ***	-0.015	-(2.24) **	0.051	(5.72) ***
	가구원 수	-0.006	-(1.96) **	-0.008	-(2.78) ***	0.020	(5.45) ***
상태 변화	실업횟수	-0.044	-(49.32) ***	0.003	(6.05) ***	0.021	(27.77) ***
	별거이혼	-0.141	-(12.23) ***	0.047	(5.38) ***	0.105	(8.55) ***
	이사횟수	-0.047	-(17.73) ***	0.003	(1.74) *	0.033	(13.04) ***
	평균소득(중범위)						
	저소득(<1200)	0.164	(12.61) ***	-0.002	-(0.24)	-0.090	-(6.97) ***
	고소득(>4800)	-0.096	-(11.45) ***	0.011	(1.27)	0.044	(3.76) ***
가구 응답 이력	가구 응답횟수						
	기준(계속응답)						
	잠정적 이탈						
	반복 이탈						
	일시적 이탈						
확정적 이탈							
표본 수		12,537					
Wald chi2		3853.92***					
Log likelihood		-13607.4					
Pseudo R2		0.192					

주 : * p < 0.10, ** p < 0.05, *** p < 0.001.
 자료 : KLIPS 1~19차(1998~2016년 조사) 학술대회용 자료

<표 9> 1차년도 15세 이상 개인표본 다항로짓분석 결과: 가구 응답력 포함 (기준 = 계속 응답)

		집단 1 (잠정적 이탈)		집단 2 (반복 이탈)		집단 3 (일시적 이탈)	
		dy/dx	z	dy/dx	z	dy/dx	z
1차 개인 특성	여성 더미	0.040	(7.81)***	-0.071	-(10.10)***	-0.019	-(3.18)***
	연령	-0.020	-(25.22)***	0.002	(1.41)	-0.010	-(10.94)***
	연령제곱/100	0.025	(25.88)***	-0.006	-(4.12)***	0.010	(9.30)***
	교육(고졸미만)						
	고졸	0.019	(3.31)***	-0.008	-(1.11)	0.001	(0.18)
	전문대 이상	0.018	(2.31)**	-0.004	-(0.38)	0.018	(2.01)**
	혼인상태(미혼)						
	기혼 유배우	0.046	(3.83)***	-0.036	-(1.89)*	0.039	(2.91)***
	별거/이혼 등	0.063	(4.56)***	-0.041	-(1.93)*	0.037	(2.46)**
비건강 더미	0.031	(4.78)***	0.053	(6.16)***	-0.026	-(3.50)***	
1차 가구 특성	지역(서울=1)						
	광역시	-0.011	-(1.65)*	-0.018	-(2.05)**	-0.011	-(1.48)
	경기	-0.013	-(1.73)*	-0.009	-(0.94)	-0.003	-(0.31)
	기타 도지역	-0.021	-(3.05)***	-0.008	-(0.87)	-0.012	-(1.46)
	소득(중범위)						
	저소득(<1200)	0.003	(0.51)	-0.010	-(1.34)	0.003	(0.50)
	고소득(>4800)	0.029	(1.90)*	0.016	(0.84)	-0.010	-(0.61)
	자가 더미	0.014	(2.52)**	0.030	(4.03)***	-0.014	-(2.22)**
가구원 수	0.008	(3.43)***	0.010	(3.43)***	-0.007	-(2.61)	
상태 변화	실업횟수	-0.022	-(27.03)***	0.012	(16.27)***	0.004	(6.95)***
	별거이혼	-0.084	-(9.67)***	0.058	(5.80)***	0.034	(4.19)***
	이사횟수	0.003	(1.86)*	0.004	(2.00)**	-0.002	-(1.11)
	평균소득(중범위)						
	저소득(<1200)	0.055	(4.95)***	-0.046	-(3.60)***	0.012	(1.24)
	고소득(>4800)	-0.033	-(5.01)***	0.012	(1.42)	0.009	(1.12)
가구 응답 이력	가구 응답횟수	-0.020	-(19.30)***	0.019	(6.93)***	0.003	(2.15)**
	기준(계속응답)						
	잠정적 이탈	-0.141	-(3.12)***	-0.224	-(3.18)***	0.372	(9.64)***
	반복 이탈	-0.191	-(3.88)***	0.104	(1.38)	0.099	(2.75)***
	일시적 이탈	-0.192	-(4.12)***	-0.001	-(0.02)	0.154	(3.09)***
	확정적 이탈	-0.185	-(3.28)***	-0.358	-(4.69)***	0.004	(0.10)
표본 수		12,539					
Wald chi2		4348.61***					
Log likelihood		-8420.7					
Pseudo R2		0.4992					

주 : * p < 0.10, ** p < 0.05, *** p < 0.001.
 자료 : KLIPS 1~19차(1998~2016년 조사) 학술대회용 자료

3. 1차년도 15세 미만 자녀 표본

<표 10>은 15세 미만 자녀 표본에 대한 다항로짓 분석결과이다. 앞서 가구 및 개인 표본에 대한 분석에서는 기준집단이 계속 응답이었던 반면, 15세 미만 자녀표본에서는 중기이후 진입한 표본을 기준 집단으로 설정하였다. 표에서 제시된 계수들은 조사에 한 번도 진입하지 않은 조사 미진입, 조사 초기에 진입한 후 잠정적으로 이탈한 집단, 그리고 조사 초기에 진입한 후 일시적으로 이탈했지만 다시 재복귀한 집단 등 세 그룹에 대한 한계효과이다.

우선 1차조사 시점 자녀 본인인의 인구사회학적 특성이 패널리탈에 미치는 효과를 살펴본다. 우선 성별의 효과를 살펴보면, 여성인 경우 남성에 비해 일시적 이탈 확률이 3.6% 높고, 조사 미진입 확률 역시 2.6% 더 높은 것으로 분석되었다. 그러나 초기 잠정적 이탈 집단은 중기이후 진입한 집단과 통계적으로 유의미한 차이가 확인되지 않았다. 1차년도 시점에서 자녀의 연령은 높을수록 조사 미진입 확률이 낮은 반면, 초기 잠정적 이탈 확률은 더 높은 것으로 확인되었다. 이러한 결과는 앞서 4장의 응답률 분석결과에서도 확인된 것과 같이 연령이 낮을수록 진입 후 응답률의 정점이 하락하는 것과 동일한 맥락에서 해석될 수 있다. 자녀의 건강상태 역시 이탈 확률에 영향을 미치는 것으로 확인되었다. 건강 상태가 좋지 않을수록 조사 미진입 확률이 12.4% 높은 반면 초기 일시적 이탈 확률이 7.3% 낮은 것으로 확인되었다. 초기-일시적 이탈 집단이 전체적으로 계속 응답 집단과 유사한 특성을 보인다는 점에서 일관된 결과로 해석될 수 있다.

1차년도 가구 특성이 미치는 효과는 역시 가구 및 개인 표본의 결과와 유사하게 나타났다. 기타 도지역의 경우 조사 미진입 확률이 가장 낮고 잠정적 이탈 확률 역시 낮았다. 반면 기타 도지역의 경우 서울 보다 일시적 이탈 확률은 상대적으로 높았다. 1차년도 가구소득이 낮을수록 자녀의 조사 미진입 확률 역시도 높은 것으로 확인되었다. 1차년도 고소득 가구 집단은 자녀의 조사 미진입 확률이 18.3% 높은 반면, 저소득 집단은 8.1% 낮았다. 주택형태가 자가인 경우, 가구원 수가 많은 경우에도 조사 미진입 확률이 낮은 것으로 확인되었다. 그러나 전체적으로 거주지역을 제외한 다른 가구 특성은 잠정적 이탈 및 일시적 이탈에서는 중기이후 진입과 대체적으로 통계적인 차이가 유의미하지 않았다(고소득 가구일수록 일시적 이탈 가능성이 높다는 점은 제외).

1차년도 당시 부모의 특성은 대체로 자녀의 조사 미진입에 강한 통계적 영향을 미쳤다. 주된 경제활동을 수행하는 부모가 조사 미진입 확률이 상대적으로 낮았으며, 부모의 연령이 젊을수록 역시 조사 미진입 확률이 낮았다. 반대로 1차년도 당시 부모가 모두 실업상태인 경우 자녀의 조사 미진입 확률이 4.9% 높은 것으로 확인되었다. 1차년도 당시 부모가 별거/이혼 상태인 경우 통계적으로 유의미한 결과를 보이지 않았지만, 이부 부모의 별거이혼 상태가 출현한 경우 잠정적 이탈 확률은 높이는 것으로 확인되었다. 부모의 초기 실업상태는 조사미진입확률을 높였던 반면, 조사 초기 이후 부모의 실업 횟수 자체는 오히려 조사 미진입 확률을 낮추는 것으로 확인되었다. 이사횟수 역시 조사 미진입 확률을 낮추는 반면 일시적 이탈 확률을 높이는 것으로 나타나 15세 이상 개인 표본의 결과와 동일하게 나타났다. 마지막으로 평균소득 효과 역시 앞서 가구 및 개인표본과 마찬가지로 저소득 계층으로의 하향이동이 조사 미진입 확률을 높이는 반면, 고소득 계층으로의 상향

이동이 조사 미진입 확률을 낮추는 것으로 확인되었다. 전체적으로 상태변화와 관련된 변수들의 효과는 초기 상태 효과와의 연관관계 하에서 해석상의 주의가 필요한 것으로 판단된다.

마지막으로 부모의 응달이력이 자녀의 응달에 미치는 효과는 대체로 통계적으로 유의하였다(<표 11> 참조). 부모의 응달횟수가 많을수록 자녀의 조사 미진입 확률이 낮아졌으며, 일시적 이탈 확률은 높아지는 것으로 분석되었다. 부모가 잠정적 이탈 유형에 속하는 경우 자녀의 조사 미진입 확률과 잠정적 이탈에도 정(+)¹의 영향을 미치는 반면 일시적 이탈에는 부(-)²의 영향을 미치는 것으로 나타났다. 전체적으로 부모의 응달 이력이 자녀의 응달에 미치는 효과는 유의미하지만, 이를 통제 한 후에도 다른 설명변수들의 효과는 독립적인 영향을 여전히 유지하는 것을 확인할 수 있다.

<표 10> 15세 미만 자녀표본 다항로짓분석 결과: 부모 응답력 제외 (기준 = 중기 이후 진입)

		집단 1 (조사 미진입)		집단 2 (잠정적 이탈)		집단 3 (일시적 이탈)	
		dy/dx	z	dy/dx	z	dy/dx	z
1차 자녀 특성	여성 더미	0.026	(2.44) **	-0.009	-(0.87)	0.035	(3.15) ***
	연령	-0.177	-(21.21) ***	0.145	(5.64) ***	0.042	(1.68) *
	연령제공/100	1.144	(19.46) ***	-0.514	-(4.31) ***	-0.126	-(1.05)
	비건강 더미	0.124	(9.69) ***	-0.008	-(0.48)	-0.073	-(4.08) ***
1차 가구 특성	지역(서울=1)						
	광역시	-0.060	-(3.99) ***	-0.030	-(2.11) **	0.047	(3.17) ***
	경기	-0.038	-(2.18) **	-0.030	-(1.87) *	0.061	(3.64) ***
	기타 도지역	-0.076	-(4.92) ***	-0.037	-(2.45) **	0.075	(4.67) ***
	소득(중범위)						
	저소득(<1200)	-0.081	-(6.08) ***	0.010	(0.71)	0.015	(0.97)
	고소득(>4800)	0.183	(4.48) ***	0.023	(0.66)	-0.093	-(3.14) ***
	자가 더미	-0.045	-(3.63) ***	-0.005	-(0.41)	0.004	(0.36)
가구원 수	-0.032	-(4.12) ***	0.002	(0.30)	0.006	(0.92)	
1차 부모 특성	여성 더미	-0.082	-(2.95) ***	-0.034	-(1.61)	0.041	(1.79) *
	연령	-0.013	-(2.62) ***	-0.012	-(1.70) *	0.013	(1.68) *
	연령제공/100	0.015	(2.72) ***	0.013	(1.88) *	-0.015	-(1.86) *
	별거/이혼 등	0.038	(0.81)	-0.040	-(1.59)	0.049	(1.55)
상태 변화	부모 실업상태	0.049	(2.04) **	0.015	(0.67)	-0.038	-(1.63)
	부모 실업횟수	-0.023	-(6.30) ***	-0.003	-(1.45)	0.010	(4.44) ***
	부모 별거이혼	-0.043	-(1.57)	0.045	(2.55) **	-0.007	-(0.37)
	이사횟수	-0.080	-(15.53) ***	-0.005	-(1.29)	0.029	(7.52) ***
	평균소득(중범위)						
	저소득(<1200)	0.273	(6.78) ***	0.014	(0.57)	-0.131	-(5.77) ***
고소득(>4800)	-0.119	-(9.66) ***	-0.028	-(2.09) **	0.057	(4.02) **	
부모 응답 이력	부모 응답횟수						
	기준(계속응답)						
	잠정적 이탈						
	반복 이탈						
	일시적 이탈						
	확정적 이탈						
표본 수		2923					
Wald chi2		761.4***					
Log pseudolikelihood		-1335.5					
Pseudo R2		0.666					

주 : * p < 0.10, ** p < 0.05, *** p < 0.001.
 자료 : KLIPS 1~19차(1998~2016년 조사) 학술대회용 자료

<표 11> 15세 미만 자녀표본 다항로짓분석 결과: 부모 응답력 포함 (기준 = 중기 이후 진입)

		집단 1 (조사 미진입)		집단 2 (잠정적 이탈)		집단 3 (일시적 이탈)	
		dy/dx	z	dy/dx	z	dy/dx	z
1차 자녀 특성	여성 터미	0.011	(1.44)	-0.014	-(1.55)	0.046	(4.75) ***
	연령	-0.123	-(17.24) ***	0.141	(6.29) ***	0.016	(0.68)
	연령제공/100	0.697	(15.41) ***	-0.486	-(4.69) ***	0.036	(0.31)
	비건강 터미	0.009	(0.90)	-0.048	-(3.09) ***	0.035	(2.00) **
1차 가구 특성	지역(서울=1)						
	광역시	-0.012	-(1.17)	-0.019	-(1.48)	0.026	(1.96) **
	경기	0.000	(0.02)	-0.038	-(2.78) ***	0.057	(3.97) ***
	기타 도지역	-0.010	-(0.91)	-0.017	-(1.28)	0.039	(2.82) ***
	소득(중범위)						
	저소득(<1200)	-0.021	-(1.96) **	0.005	(0.40)	0.002	(0.17)
	고소득(>4800)	0.029	(1.23)	0.008	(0.27)	-0.023	-(0.72)
	자가 터미	-0.008	-(0.91)	0.006	(0.63)	-0.011	-(1.03)
가구원 수	-0.003	-(0.53)	0.004	(0.62)	-0.002	-(0.37)	
1차 부모 특성	여성 터미	-0.053	-(2.75) ***	-0.015	-(0.72)	0.025	(1.25)
	연령	-0.002	-(0.56)	-0.012	-(1.78) *	0.010	(1.74) *
	연령제공/100	0.003	(0.82)	0.013	(1.95) *	-0.011	-(1.90) *
	별거/이혼 등	-0.115	-(4.08) ***	0.008	(0.22)	0.092	(1.87) *
상태 변화	부모 실업상태	0.000	-(0.01)	0.007	(0.36)	-0.007	-(0.37)
	부모 실업횟수	-0.003	-(1.70) *	0.000	-(0.10)	0.002	(1.31)
	부모 별거이혼	0.031	(1.74) *	0.048	(2.91) ***	-0.033	-(2.13) **
	이사횟수	-0.016	-(4.15) ***	0.001	(0.34)	0.005	(1.56)
	평균소득(중범위)						
	저소득(<1200)	0.042	(2.33) **	0.016	(0.67)	-0.055	-(2.11) **
고소득(>4800)	-0.023	-(2.09) **	0.004	(0.30)	0.005	(0.44)	
부모 응답 이력	부모 응답횟수	-0.015	-(10.46) ***	0.003	(1.22)	0.008	(2.96) ***
	기준(계속응답)						
	잠정적 이탈	0.209	(4.04) ***	0.330	(13.46) ***	-0.308	-(14.90) ***
	반복 이탈	0.030	(1.19)	0.039	(2.58) ***	-0.060	-(3.09) ***
	일시적 이탈	0.091	(3.24) ***	0.271	(11.17) ***	-0.226	-(7.74) ***
	확정적 이탈	-0.099	-(2.43) **	-0.033	-(0.81)	0.071	(1.14)
표본 수		2,923					
Wald chi2		1186.8***					
Log pseudolikelihood		-1072.3					
Pseudo R2		0.732					

주 : * p < 0.10, ** p < 0.05, *** p < 0.001.
 자료 : KLIPS 1~19차(1998~2016년 조사) 학술대회용 자료

VI. 결론

1. 분석결과 요약

이 글에서는 사건계열분석 기법을 이용하여 KLIPS 1-19차 자료의 패널리탈을 유형화 한 후, 다항로지모형을 이용하여 패널리탈에 영향을 미치는 요인들을 규명하고자 하였다. 이때 1차년도 조사 당시 원가구 표본과 15세 이상 개인표본, 그리고 15세 미만 자녀 표본을 서로 구분하여 공통 점과 차이점을 찾고자 하였다. 본문의 주요 분석결과를 요약하면 다음과 같다.

첫째, 1-19차년도 가구 및 15세이상 개인 표본의 이탈 유형은 공통적으로 4가지 집단- 단조 이탈에 해당되는 잠정적 이탈과 확정적 이탈, 그리고 비단조 이탈에 해당되는 일시적 이탈과 반복 이탈 집단-으로 유형화될 수 있다. 반면 15세 미만 개인은 15세가 되어야 개인 조사에 진입하기 때문에 응답 패턴 자체가 부모 세대와 구조적으로 상이하다. 즉 15세 이전까지는 조사 진입 여부가 확인되지 않으며, 연령이 어릴수록 조사 진입 시기가 늦어지기 때문에 잠재적 이탈 가능성이 더욱 높아진다. 따라서 유형화 결과 역시도 주로 연령 집단에 따라 구획된다.

둘째, 응답률 추세 분석의 결과와 다항로지 분석결과에서 연령의 효과는 일관된 반면 성별의 효과는 서로 상이한 것으로 나타났다. 즉 연령이 젊을수록 이탈확률이 높다는 점은 일관된 결과를 보였으나, 여성의 이탈 확률, 특히 잠정적 이탈확률은 사회경제적 효과들을 통제할 때 오히려 더 높아지는 것으로 나타났다. 그 동안 국내에서는 다루지 않았던 개인의 건강상태가 응답에 미치는 효과를 규명했는데, 건강상태가 좋지 않을수록 잠정적 이탈 혹은 자녀의 조사 미진입 가능성을 높이는 것으로 밝혀졌다.

셋째, 가구 특성과 관련하여 서울 및 수도권 지역일수록 상대적으로 모든 집단에서의 잠정적 이탈확률이 높고, 자녀의 조사 미진입 확률이 높았다. 가구 소득 계층 효과 역시 초기에는 고소득 가구일수록 가구, 개인의 잠정적 이탈 및 자녀의 조사 미진입 확률을 높였으며, 저소득 계층으로의 하락 역시 잠정적 이탈 및 자녀의 조사 미진입 확률을 높이는 것으로 확인되었다.

넷째, 상대변화와 관련된 변수들의 효과에 대해서는 몇몇 변수들에서 통계적으로 유의미한 결과가 관찰되었다. 다만 이들 변수들의 효과를 보다 정확하게 측정하기 위해서는 본 연구에서 사용한 횡단적 접근 보다는 특정 시점의 상대변화가 해당 시점 혹은 다음 시점에서의 이탈 확률에 어떤 영향을 미치는지를 측정하는 동태패널모형 방식의 접근이 필요할 것으로 판단된다.

다섯째, 자녀의 패널리탈에 있어서 부모의 특성이 미치는 효과 역시 대체로 유의미한 것으로 나타났다. 자녀의 조사 진입에 상대적으로 시차가 존재하기 때문에 이탈 여부를 확인하기 어렵지만, 장기적으로 부모의 패널리탈은 자녀세대에 누적적인 영향을 미치게 된다. 마지막으로, 가구단위의 응답 이력이 개인의 응답이력에 미치는 효과, 부모의 응답 이력이 자녀세대에 미치는 효과 역시 통계적으로 유의미하게 나타났다. 이로 인해 개인별 특성들의 통계적 유의미성이 하락하기도 하지만, 여전히 많은 변수의 설명력은 유지되었다.

2. 정책적 함의

이상의 결과로부터 몇 가지 패널 조사 전략 수립을 위한 몇 가지 정책적 함의를 도출할 수 있다.

첫째, 패널조사가 장기화되면 될수록 조사의 지속가능성을 확보하기 위해 표본이탈을 최소화하는 것이 더욱 중요해진다. 그런데 조사 초기에는 단조 이탈 중심의 패턴을 유지하다가, 시간이 흐를수록 비단조 이탈의 비중이 증가하며 그 패턴 역시도 복잡하고 다양해질 수 밖에 없다. 따라서 패널이탈을 효과를 관리하기 위해 다양한 형태의 비단조 이탈을 유형화하여 맞춤형 이탈 억제 노력이 더욱 중요해진다. 본문의 분석결과에서도 나타나듯이, 일시적 이탈의 경우 전체적인 조사의 충성도 측면에서 계속 응답 집단보다 크게 떨어지지 않기 때문에, 이탈방지에 투입되는 노력 못지않게 이탈 발생 후 조기 대응을 통해 단기간에 조사에 재복귀 시키려는 다양한 인센티브가 필요할 것으로 판단된다.

둘째, 지금까지 가구 중심으로 구축된 패널 관리 전략 외에 개인 단위의 패널 관리전략을 체계적으로 고민할 필요가 있다. 이를 위해서는 개인 응답률을 제고하기 위한 개인별 응답 인센티브, 다양한 측면에서의 개인별 조사성공률 관리, 이와 연동된 면접원 인센티브의 변화 등 조사 설계 및 현상 실사 단계에서의 다양한 혁신이 검토될 필요가 있다.

셋째, 지금까지 상대적으로 간과되었던 유소년 표본에 대한 이력관리가 세밀하게 이루어질 필요가 있다. 15세 미만 자녀의 경우 개인 조사대상에 포함되지 않기 때문에 별도의 직접적인 표본관리 대상에 포함되지 않는다고 볼 수 있다. 영국의 BHPS나 미국의 PSID와 같은 다른 가구패널조사들은 별도의 아동 조사를 실시하고 있기 때문에 직접적인 패널 관리가 가능하다. KLIPS에서도 자녀의 출생 시점부터 건강, 교육, 성장 등과 같은 정보들을 체계적으로 축적하면서 노동시장 참여 시기 까지 패널 관리를 연계하는 방안을 검토할 수 있을 것이다.

넷째, 조사 관리 측면에서의 노력 못지않게 통계적 측면에서 표본이탈에 따른 세분화된 검증이 필요하며, 필요시 가중치 부여와 임putation 등을 통해 이를 보완하여 개선하는 노력 필요하다. 특히 기존의 가중치 부여 방식이 비단조 이탈 패턴을 적절하게 교정해주는 모형인지, 패널 장기화에 따라 패널이탈에 따른 초기 오차가 누적되지는 않는지 등에 대한 전반적인 검토가 필요하다. 비단조 이탈의 경우 imputation을 통해 누락된 정보를 보완하는데 효과적일 수 있지만, 비단조 이탈 내에서도 다양한 유형에 따라 자료의 왜곡이 발생하는지를 분석할 필요가 있다.

마지막으로 자료의 분석 측면에서 표본이탈에 대한 처리 방식이 결과의 차이에 어떤 영향을 미치는지를 면밀하게 살펴볼 필요가 있다. KLIPS와 같은 장기패널에서 무비판적인 균형패널의 사용 혹은 불균형패널의 사용은 핵심 변수의 통계적 유의성 뿐만 아니라 방향까지 바꿀 수 있다. 주요 핵심 변수의 경우 패널 자료 구축 형태에 따른 결과 차이 등에 대한 검증을 통해 자료의 신뢰성을 높여주는 노력이 필요하다.

참고 문헌

- 남기성, 천영민(2012). “An Analysis of Response Pattern and Panel Attrition in KLIPS”, 『응용 통계연구』 25(6), pp. 933-945.
- 성재민(2007). “응답자 사례금, 조사에 영향이 있는가.”, 『2007 노동리뷰 10월호(통권 제34호)』, 한국노동연구원. pp.67~75.
- 신선옥(2008). “한국노동패널조사의 응답자 태도에 면접원이 미치는 효과.”, 『2008 노동리뷰 1월호(통권 제37호)』, 한국노동연구원. pp.74~82.
- 이경희·민인식(2016). “패널 조사 응답 지속성에 관한 연구 - 한국노동패널조사를 중심으로”, 『조사연구』. 17(3), pp.1~24.
- 이상호(2005). “한국노동패널(KLIPS)의 표본이탈 분석”, 『2005 노동리뷰 11월호(통권 제11호)』, 한국노동연구원. pp.66-80.
- _____(2013). “한국노동패널의 표본이탈: 재진입 표본을 중심으로”, 『제3회 국가통계방법론 심포지엄 자료집』.
- 이상협·박찬용·정성석·최혜미(2011) “한국노동패널 탈락 분석”, 『한국데이터정보과학회지』. 22(1), pp. 1-8.
- 이혜정(2017) “한국노동패널조사 98표본 원가구의 표본이탈 유형화”, 『2017 패널브리프 제9호』, 한국노동연구원.
- 장지연·이혜정(2008) “중·고령자의 근로생애 유형-사건계열분석기법을 이용한 취업력 분석-”, 『2008 노동리뷰 6월호(통권 제42호)』, 한국노동연구원. pp.76~86.
- 최효미(2016) “아동패널조사의 원표본 이탈에 관한 연구”, 『제7회 한국아동패널 국제학술대회 자료집』, 육아정책연구소.
- 최효미·이상호·성재민·김기민·배기준(2012) “패널자료의 품질 개선을 위한 조사기법 연구-컴퓨터를 이용한 조사(CAPI) 방법-”, 한국노동연구원
- 홍민기, 최효미(2014). 노동패널 표본이탈과 불평등 지표. 『조사연구』. 15(4), pp.95~122.
- Agresti(1990) *Categorical Data Analysis*, New York:Wiley.
- Brzinsky-Fay, C., U. Kohler & M. Luniak(2006), 'Sequence Analysis with Stata', *The Stata Journal*, Volume 6, Number 4, pp.435~460.
- Burkam, David T. and Lee, Valerie E. 1998. "Effects of Monotone and Nonmonotone Attrition on Parameter Estimates in Regression Models with Educational Data: Demographic Effects on Achievement, Aspiration, and Attitudes." *The Journal of Human Resources, Special Issue: Attrition in Longitudinal Surveys*, 33(2): 555-574.
- Christian Brzinsky-Fay, Ulrich Kohler, and Magdalena Luniak. 2006. "Sequence analysis with Stata." *The Stata Journal* 6:435-460.
- Fitzgerald, J., P. Gottschalk, and R. Moffitt. 1998. "An Analysis of Sample Attrition in Panel

- Data: The Michigan Panel Study of Income Dynamics.” *Journal of Human Resources* 33:251-99.
- Hill, D. H. and R. J. Willis. 2001. “Reducing Panel Attrition: A Search for Effective Policy Instruments.” *The Journal of Human Resources* 36:416-38.
- Lillard, L. A. and C. W. A. Panis. 1998. “Panel Attrition from the Panel Study of Income Dynamics—Household Income, Marital Status and Mortality.” *The Journal of Human Resources* 33:437-57.
- Lepkowski, J. M. and Couper, M. P. 2002. “Nonresponse in the Second Wave of Longitudinal Household Surveys” in R. M. Groves, et al. (eds) *Survey Nonresponse*, New York: John Wiley & Sons.
- Lynn, p. Buck, N., Burton, J. Jackle, A. and Laurie, H.(2005). "A Review of Methodological Research Pertinent to Longitudinal Survey Design and Data Collection", *ISER Working Paper Series*.
- Lynn et al.(2009) *Methodology of Longitudinal Surveys*, Chichester, England: John Wiley.
- Macindoe, H. & A. Abbott(2004), 'Sequeunce Analysis and Optimal Matching Techniques for Social Science Data', Handbook of Data Analysis, Sage Publication.
- Schoeni, R. E. and Wiemers, E. E., (2015) “The Implications of Selective Attrition for Estimates of Intergenerational Elasticity of Family Income”, *Journal of Economic Inequality*, 13:351 - 372.
- Uhrig, S. N. 2008. “The Nature and Causes of Attrition in the British Household Panel Survey.” *ISER Working Paper*, 5. Colchester, England: Institute for Social and Economic Research.
- Watson, N. and M. Wooden. 2009. “Identifying Factors Affecting Longitudinal Survey Response.” Pp. 157-82 in *Methodology of Longitudinal Surveys*, edited by P. Lynn. Chichester, England: John Wiley.
- Zabel, J. E. 1998. “An Analysis of Attrition in the Panel Study of Income Dynamics and the Survey of Income and Program Participation with an Application to a Model of Labor Market Behavior”, *Journal of Human Resources*, 33(2): 479-506.

<부 록>

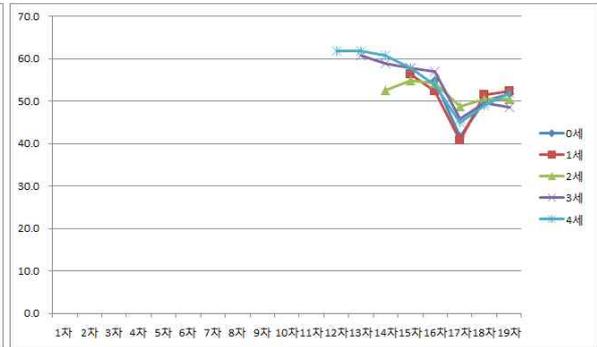
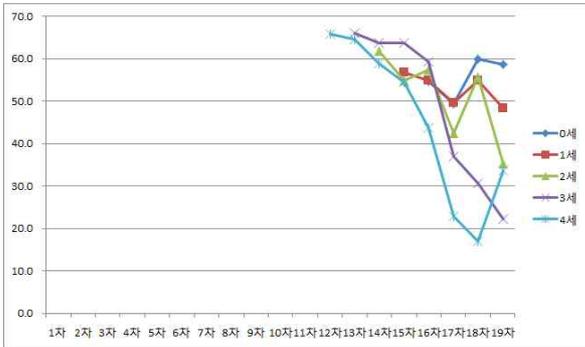
[부도 1] 15세 미만 개인의 성별 응답률 추이

(a) 남성

(b) 여성

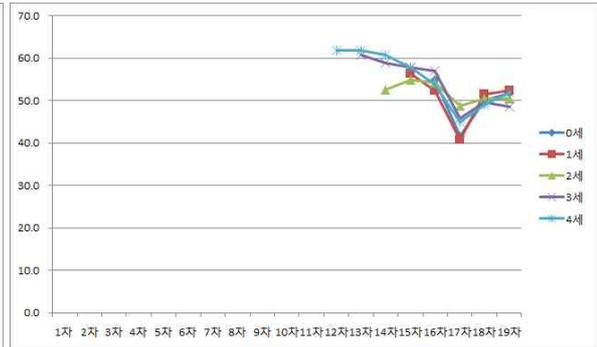
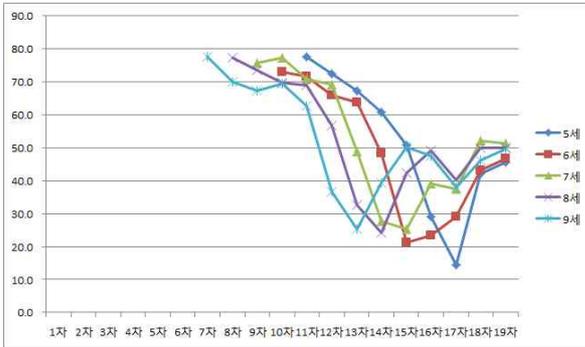
(1) 0-4세

(1) 0-4세



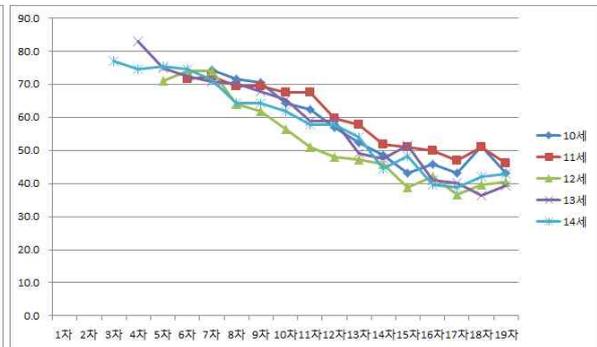
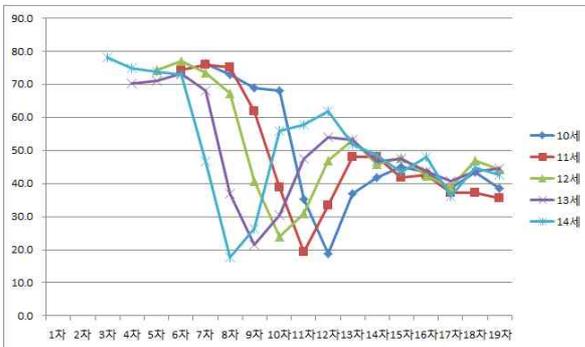
(2) 5-9세

(2) 5-9세



(3) 10-14세

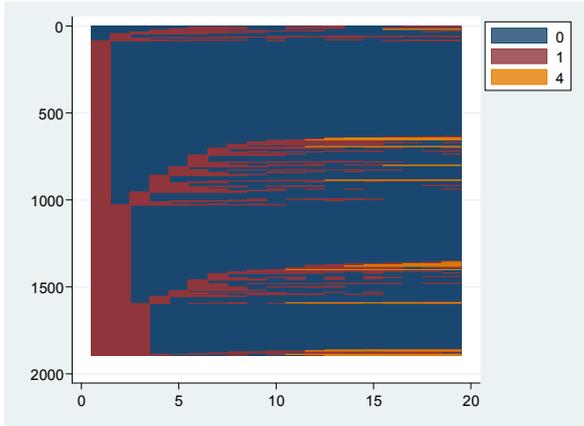
(3) 10-14세



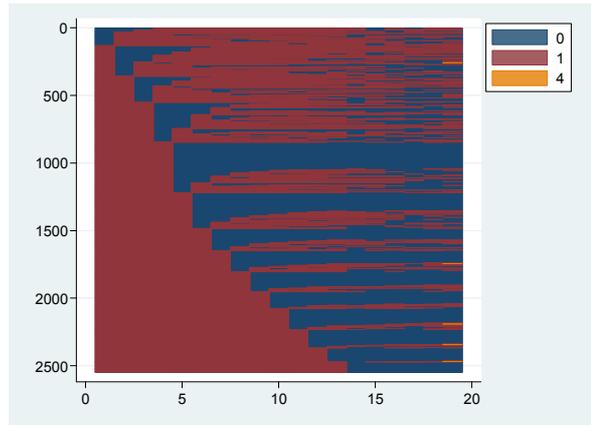
자료: KLIPS 1~19차(1998~2016년 조사) 학술대회용 자료

[부도 2] 개인 15-64세 표본의 패널이탈 유형

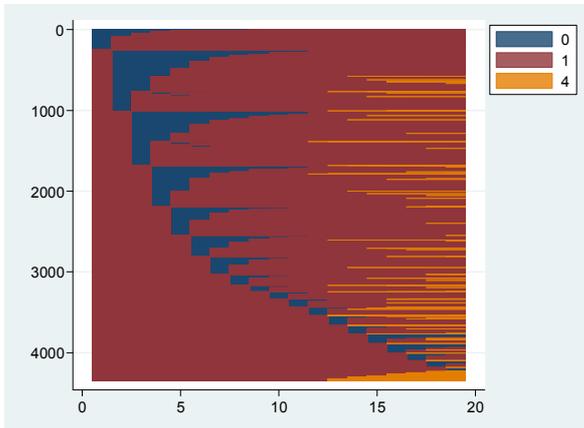
(a) 잠정적 이탈



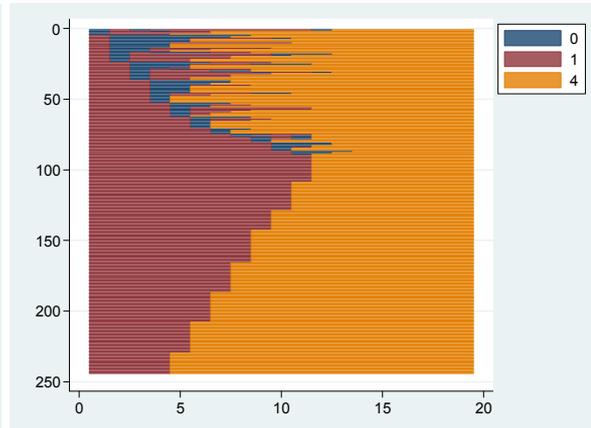
(b) 간헐적 이탈



(c) 일시적 이탈



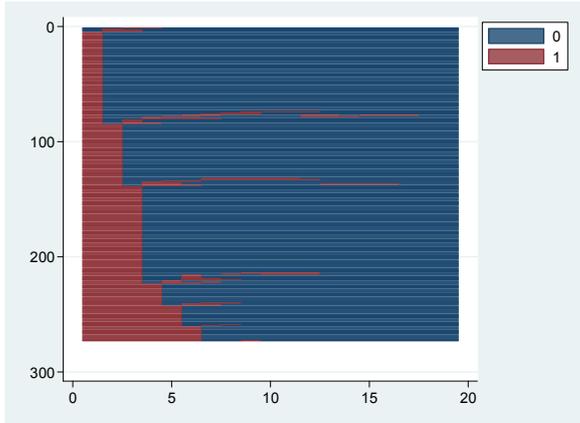
(d) 확정적 이탈



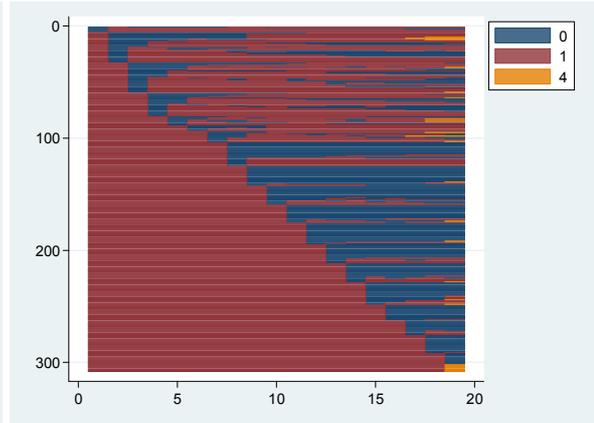
- 주 1) : 각 차수별 0=무응답, 1=응답, 3=미진입, 4=사망
 2) : X축은 조사차수, Y축은 가구를 의미함
 자료 : KLIPS 1~19차(1998~2016년 조사) 학술대회용 자료

[부도 3] 개인 1차년도 65세 이상 표본의 패널이탈 유형

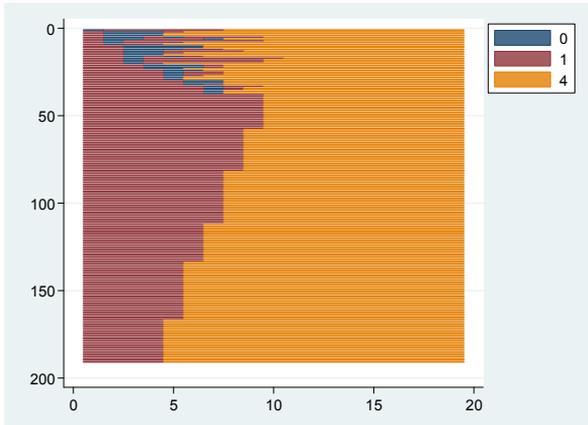
(a) 초기 잠정적 이탈



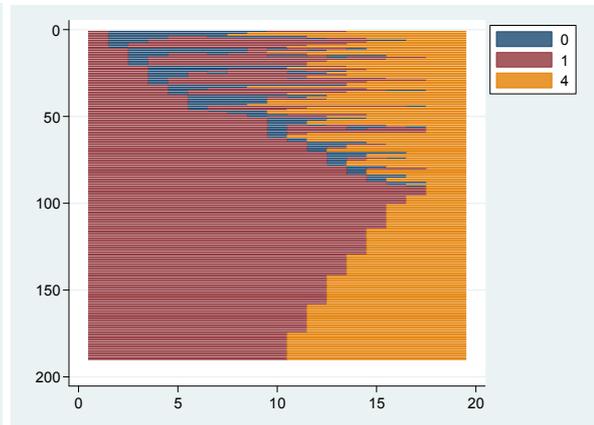
(b) 초기 반복이탈-중후기 잠정적 이탈



(c) 초기 확정적 이탈



(d) 중후기 확정적 이탈



- 주 1) : 각 차수별 0=무응답, 1=응답, 3=미진입, 4=사망
 2) : X축은 조사차수, Y축은 가구를 의미함
 자료: KLIPS 1~19차(1998~2016년 조사) 학술대회용 자료