

노동정책연구
2023. 제23권 제2호 pp.85~107
한국노동연구원
<http://doi.org/10.22914/jlp.2023.23.2.004>

연구논문

KLIPS의 가족관계행렬 구축과 활용*

민인식**
홍경준***
전예지****

본 연구는 장기간(1~24차) 조사가 진행된 한국노동패널조사(KLIPS)를 활용하여 한국식 FRM(Family Relationship Matrix) 데이터를 구축하는 것을 목적으로 한다. FRM 데이터는 원데이터에서 가족관계를 식별하고 매핑(mapping)하는 시스템을 바탕으로 한다. 기존 데이터는 가구주를 중심으로 가족관계를 정의하지만, FRM 데이터를 활용하면 가구 내 모든 가구원 간의 관계를 살펴볼 수 있다. 또한, 이 데이터는 가구 내 가구원 간 관계뿐 아니라 가구 외 구성원 간 관계 정보를 포함하고 있다. 본 연구에서는 FRM 데이터를 활용한 예제로 가구구성의 시계열적 변화를 시각적으로 제시한다. 가족관계망이 가족관계 만족도에 미치는 영향을 분석한 결과, 평균 촌수가 높을수록 가족관계 만족도는 감소하는 것으로 나타났다. FRM 데이터는 가구주 여부와 관련 없이 동거 혹은 비동거 가족 구성원 간 관계를 통시적·다면적으로 분석할 수 있다는 장점이 있다. 본 연구에서 구축한 FRM 데이터는 가족관계에 관한 연구의 범위를 확장하는 데 기여할 것으로 기대된다.

핵심용어 : 가족관계행렬, 한국노동패널, 가구 외 가족관계행렬, 가족관계 만족도

논문접수일 : 2023년 3월 7일, 심사의뢰일 : 2023년 3월 20일, 심사완료일 : 2023년 5월 11일

* 본 논문은 2021년 대한민국 교육부와 한국연구재단의 지원을 받아 작성하였다(NRF-2021S1A5A2A03066029).

** (제1 저자) 경희대학교 경제학과 교수(imin@khu.ac.kr)

*** (교신저자) 성균관대학교 사회복지학과 교수(zookie@skku.edu)

**** (공동저자) 성균관대학교 사회복지학과 박사수료(yeji.j@skku.edu)

I. 서론

인간은 누구나 가족, 친구, 이웃 등과 사회관계를 맺으며 생활하며, 가족은 개인이 일상생활 속에서 유대 관계를 맺는 일차적인 대상이 된다. 가족관계는 개인의 사회관계망(social network)에서 핵심적인 기능을 한다는 측면에서(김영범·박준식, 2004; 박경숙, 2000) 개인 간 상호작용 혹은 유대를 논의한 연구에서 중요하게 다루어졌다.

가족관계를 논의한 연구는 크게 두 가지 흐름을 보인다. 첫째, 관련 연구(김영범·박준식, 2004; 김유경 외, 2013; 김태현·박숙자, 1992; 윤현숙·유희정, 2006)는 가족 구성원 가운데 배우자 혹은 자녀에 초점을 맞추어 관계를 논의하였다. 20대는 부모, 30대는 부모와 자녀, 40대 이상은 배우자와 긴밀한 관계를 맺으며(김유경 외, 2013), 노인은 자녀 및 배우자와의 관계에 의존하는 정도가 크기 때문이다(김미혜·신경림, 2005; 김영범·박준식, 2004; 김은경, 2002; 박경숙, 2000; 윤현숙·유희정, 2006). 둘째, 가족관계를 논의한 대다수 연구(김영범·박준식, 2004; 김태현·박숙자, 1992; 박경숙, 2000; 이서영, 2015; 이혜자·박경애, 2009; Zheng and Chen, 2020)는 횡단면 자료를 활용하여 분석하였다. 이 연구들은 가족관계에 관한 문항을 바탕으로 설문조사를 실시하여 1차 자료를 구축하고 해당 자료를 분석에 활용하였다.

개인은 배우자나 자녀, 부모 외에 형제, 조부모, 시부모, 친척 등 다양한 가족 구성원과 관계를 맺을 수 있으며, 가족관계는 시간이 흐름에 따라 변할 수 있다. 그러나 기존 연구는 배우자나 자녀, 부모와의 관계를 중심으로 가족관계를 국한하였으며, 하나의 조사 시점에서의 관계를 분석하였다는 한계를 가진다. 선행연구가 이러한 한계를 가진 데에는 여러 원인이 있겠지만, 다양한 가족 구성원과의 관계를 장기간 관찰할 수 있는 데이터가 구축되지 못한 점도 하나의 원인으로 꼽을 수 있다. 현존하는 2차 자료 대다수는 가구주를 중심으로 동거 혹은 비동거 가구원과의 관계 정보를 수집한 까닭에 데이터의 구조상 가구주가 아닌 가구원이 서로 어떤 관계를 맺는지 파악하기 어렵다.

다양한 가족 구성원 간 관계 변화는 “가족관계력(family relationship history)”에 관한 데이터를 활용하여 관찰할 수 있다. 미국에서는 관련 자료를 구축하고 배포하는 움직임이 활발하게 이루어졌다. 미국의 대표적인 패널조사인 PSID (Panel Survey of Income Dynamics)는 원데이터로부터 가족을 식별하고 매핑하는 시스템(Family Identification Mapping System)을 바탕으로 가족관계행렬 (Family Relationship Matrix : FRM) 데이터를 구축하였다(Institute for Social Research and Survey Research Center, 2021b). FRM 데이터는 1968년부터 2019년까지 개인의 형제 혹은 부모, 조부모, 자녀 등에 관한 정보를 포함하고 있으므로 해당 자료를 활용하면 세대 내 관계뿐만 아니라 세대 간 관계를 분석할 수 있다. 그러나 국내에서는 현재까지 가족관계력 데이터를 구축하고자 하는 시도가 매우 부족하였다.

이러한 갭을 줄이기 위해 본 연구에서는 우리나라에서 최장기간 동안 조사된 한국노동패널조사(Korean Labor & Income Panel Study : KLIPS)를 활용하여 한국식 FRM 데이터를 구축하고자 한다. KLIPS는 1998년(1차)부터 2021년(24차)까지 가구주와 가구원에 관한 정보를 1년 주기로 반복 조사하였으며, 원가구와 더불어 분가한 가구를 추적조사하였다. 물론 FRM 데이터를 만들지 않더라도 KLIPS를 활용하여 가족 구성원 간의 관계를 파악할 수 있지만, 이는 어디까지나 가구주의 아내 혹은 자녀와 같이 가구주를 중심으로 정의된 관계에 한정된다. 반면, FRM 데이터는 가구주 여부와 관계없이 동거 혹은 비동거 가구원을 서로 매칭하기 때문에 KLIPS만으로는 파악하기 어려운 형제간 관계, 조부모와 손자녀의 관계, 며느리와 시부모와의 관계 등을 심찰하는 데 활용될 수 있다.

본 연구의 목적은 KLIPS에 기초하여 FRM 데이터를 구축하는 과정을 소개하고, 해당 자료를 활용하여 가족관계를 분석하는 예시를 제시하는 것이다. 더 나아가 PSID에서 시도하지 않았던 가구 외(out-of-household) 가족관계력을 구축하는 과정을 설명하고자 한다. 논문의 구성은 다음과 같다. 제Ⅱ장에서는 PSID의 가족관계행렬 데이터를 소개한다. 제Ⅲ장에서는 KLIPS를 활용하여 한국식 FRM 자료를 구축하는 과정을 동거하는 가족과의 관계와 비동거하는 가족과의 관계로 나누어 정리한다. 제Ⅳ장에서는 FRM 데이터를 활용한 두 가지

예시를 제시한다. 첫 번째 예시로 가구구성의 시계열적 변화를 살펴보고, 두 번째 예시로 가족관계망이 가족관계 만족도에 미치는 영향을 분석한다. 마지막으로 제V장에서는 실증분석결과를 요약하고 연구가 가지는 함의를 제시한다.

II. PSID의 가족관계행렬

미국의 PSID(Panel Survey of Income Dynamics)는 1968년부터 시작된 가구·가구원 패널조사로 전 세계에서 가장 오랫동안 조사된 데이터이다. PSID는 미국의 5,000가구 내 18,000명 가구원에 대하여 다양한 특성을 조사하였다. 1968년부터 1996년까지는 매년 조사를 하였으며, 그 이후에는 격년제로 패널조사를 진행하였다. 2023년 1월 기준 PSID 홈페이지(<https://psidonline.isr.umich.edu/default.aspx>)에서는 2019년에 조사된 데이터를 공개하고 있다. PSID는 패널조사의 선구자적 역할을 하였는데, 우리나라의 경우 한국노동패널(Korea Labor and Income Panel Survey)이 PSID와 유사한 형태로 1998년부터 현재까지 조사되었다.

PSID는 가구·가구원에 관한 연도별 조사데이터와 더불어 매년 누적된 데이터에서 파생한 데이터(“패키지 데이터”로 호칭)를 추가로 생성하여 제공한다. <표 1>에서는 PSID에서 제공하고 있는 파생 데이터의 종류와 분석기간을 정리한다. marriage history 파일은 “결혼력” 파일로 알려져 있다. KLIPS는 공식적으로 결혼력 데이터를 제공하지 않지만, 혼인여부와 혼인연도 변수가 있으므로 결혼력 데이터를 생성하는 것이 가능하다. childbirth and adoption 파일은 “출생/입양력” 파일로 이해할 수 있으며, 한국의 경우 조세재정연구원의 재정패널 조사에서 이와 유사한 출생력 데이터를 따로 제공한다. 다만, PSID에서는 입양에 관한 정보까지 포함하는 것이 특징이다. parent identification 파일은 “부모·자녀 식별” 파일로써 자녀 표본에서 부모를 식별하고, 부모 표본에서 자녀를 식별할 수 있도록 구성되어 있다.

PSID에서 제공하는 파생 데이터 중 family relationship matrix 파일은 본 연구의 가족관계행렬과 직접적으로 연관되어 있다. PSID의 FRM은 총 12개 변수

〈표 1〉 PSID 파생 데이터 종류와 분석기간

파일이름	파일내용	분석기간
marriage-history file (혼인이력 파일)	각 개인의 혼인상태 변화 이력을 데이터로 생성: 배우자, 혼인, 이혼, 사별시점, 결혼순서, 총 결혼횟수 등을 변수로 생성	1985~2019년
childbirth and adoption history file (출생과 입양 파일)	출생한 자녀 id, 출생시점, 출생순서, 출생 시 몸무게, 키, 인종, 사망일자, 입양의 경우 입양관련 변수	1985~2019년
parent identification file (부모 식별 파일)	1983년 조사 이후 부모-자녀 간 관계를 추적. 조사대상이 자녀인 경우에는 부모 id를 추적. 조사대상이 부모인 경우에는 자녀 id를 추적	2019년
family relationship matrix (가족관계행렬 파일)	특정 개인이 속한 가구 내 모든 가구원과의 관계를 확정	1968~2019년
active savings (개인저축관련 파일)	1989년과 1994년 2개 시점에서 예금, 주식, 연금 등 금융자산(저축) 변화를 추적	1989, 1994년

자료: Institute for Social Research and Survey Research Center(2021a)를 활용하여 저자 작성.

〈표 2〉 PSID의 FRM 데이터 주요 변수

변수명	변수 내용	주요 변수값
Ego relation to reference person	가구 내 특정 가구원(ego)과 가구주(reference person)의 관계	reference person(가구주), spouse(가구주의 배우자), son or daughter(가구주의 자녀) 등
Ego relation to alter	가구 내 특정 가구원(ego)과 가구 내 다른 가구원(alter)의 관계	self(본인), spouse(배우자), child(자녀), sibling(형제자매) 등

자료: PSID 1968~2019 FRM 코드북을 활용하여 저자 작성.

로 구성되며, 이 가운데 주요한 2개 변수를 <표 2>에서 정리한다. 첫 번째 변수는 가구주(reference person or household head)와 가구원 간 관계를 정의하는 “ego relation to reference person”이다. KLIPS 데이터 내에는 이미 해당 변수(가구주와의 관계)가 존재한다. FRM의 주요한 기여는 “ego relation to alter” 변수인데, 이 변수를 사용하면 특정 가구원(ego)과 다른 가구원(alter) 간 관계

를 모두 파악할 수 있다. 서론에서 언급하였듯이 KLIPS 데이터는 가구주를 기준으로 가구원의 관계를 정의하기 때문에 가구주가 아닌 가구원 간 서로 어떤 관계인지 파악할 수 없는 문제점이 존재한다. 이에 PSID의 “ego relation to alter” 변수와 유사한 변수를 KLIPS 데이터 내에서 생성하는 것이 연구의 주요한 목적이다.

PSID에서 “ego relation to alter” 변수에 의해 정의되는 관계 수는 비친인척(nonrelative) 관계를 포함하여 모두 53가지 경우가 있다. <표 3>에서는 상대적으로 큰 비중을 차지하는 상위 5가지 관계를 정리한다. PSID의 FRM 데이터에서 전체 관측치는 3,325,609개이다. self는 가구원 본인이 자신과 매칭되는 관계이다. 데이터 구조상 생성될 수밖에 없는 관계이지만, FRM 데이터를 이용한 분석에서는 불필요한 관측치에 해당한다. sibling은 가구원 본인과 매칭되는(같은 가구 내) 형제자매의 관계이다. child는 가구원 본인과 그 자녀 간에 매칭되는 관계이고 parent는 가구원 본인과 그 부모 간에 매칭되는 관계이다. legal spouse는 같은 가구 내에서 법적 배우자와의 관계이다. 가령, 남편-아내가 같은 가구 내에서 거주하고 있다면 남편에게 매칭되는 아내는 legal spouse가 되고, 반대로 아내에게 매칭되는 남편 역시 legal spouse가 된다. 미혼이거나 배우자가 없는 가구에서는 legal spouse 관측치(관계)가 나타나지 않는다.¹⁾

<표 3> PSID FRM에서 가장 많이 나타나는 상위 5가지 관계

순위	관계	비중(%)
1	self(본인 자신): 본인-본인 매칭	25.97
2	sibling(형제자매): 본인-본인의 형제자매 매칭	20.01
3	child(자녀): 본인-본인의 자녀 매칭	17.12
4	parent(부모): 본인-본인의 부모 매칭	17.12
5	legal spouse(법적 배우자): 본인-본인의 배우자 매칭	9.38
합계		전체 관측치: 3,325,609개

1) 주의할 점은 가구주가 미혼이거나 배우자가 없더라도 그 가구주의 자녀가 결혼해서 세 사람이 같은 가구를 구성한다면, 자녀와 자녀의 배우자 간에는 legal spouse 관계가 생성된다. 다만, 가구주에게는 legal spouse 관계가 존재하지 않는다.

Ⅲ. KLIPS에서 가족관계행렬 구축

1. 가구 내 가족관계행렬

PSID에서 제공하는 FRM 데이터는 가구 내(within-household) 가구원 간 관계를 정의한다. 본 연구는 KLIPS 데이터를 이용하여 가구 내 가족관계행렬(within FRM)을 구축하고자 한다. 가구 내 가족관계를 정의하기 위하여 “가구주와의 관계” 변수(h*0261~h*0275) 15개를 활용한다. KLIPS 코드북에 따르면 [가구주와의 관계] 변수는 76가지 관계를 포함한다.²⁾ 가령, “10”은 가구주 본인이고 “20”은 가구주의 배우자를 의미한다. “997”과 “998”은 “기타 친인척” 그리고 “혈연관계가 아닌 동거인”으로 정의한다. <표 4>에서는 가구 내에서 관계를 정의하는 예를 제시한다. 부, 모 그리고 자녀 2명이 있는 4인 가구를 생각해 보자. 가구주 관계 변수에 의해 부는 10, 모는 20, 자녀 2명은 각각 11과 12 번호를 갖게 된다.

<표 4>에서 rel_ego는 가구 내에서 본인의 위치를 의미하고 rel_alter는 매칭되는 가구원의 위치를 뜻한다. code_ego와 code_alter는 노동패널 코드북에서 “가구주와의 관계(p_0102)” 변수에 의해 주어지는 본인과 매칭되는 가구원의 코드번호이다. code_ego와 code_alter 조합을 이용하면 FRM에 해당하는 ea_rel 변수를 정의할 수 있다. code_ego와 code_alter 번호가 같다면 “본인” 매칭을 의미한다. 4인 가구의 경우 4×4=16가지 가구원 매칭이 발생하고, 16가지 중에서 본인(4개), 배우자(2개), 부모(4개), 형제자매(2개), 자녀(4개) 관측치가 나타난다.

2) 자세한 내용은 23차 wave 코드북 참고.

〈표 4〉 4인 가구의 경우 FRM

rel_ego	rel_alter	code_ego	code_alter	ea_rel
부	부	10	10	본인
부	모	10	20	배우자
부	자녀 1	10	11	부모
부	자녀 2	10	12	부모
모	모	20	20	본인
모	부	20	10	배우자
모	자녀 1	20	11	부모
모	자녀 2	20	12	부모
자녀1	자녀 1	11	11	본인
자녀1	부	11	10	자녀
자녀1	모	11	20	자녀
자녀1	자녀 2	11	12	형제자매
자녀2	자녀 2	12	12	본인
자녀2	부	12	10	자녀
자녀2	모	12	20	자녀
자녀2	자녀 1	12	11	형제자매

자료: KLIPS 코드북을 활용하여 저자 작성.

KLIPS에서 가구 내 가구원에 대한 `ea_rel` 변수를 정의하기 위해 다음의 4단계 과정을 거친다(표 5 참고). 1단계에서 76개의 가구주 관계를 매칭하면 $76 \times 76 = 5,776$ 개 관계가 발생한다. 2단계에서는 현실적으로 가구 내에서 동거가 가능할 거라고 예상되는 (친척과 인척 포함) 4촌 이내 관계만 선택한다. 선택된 관계는 1,868개이다. 3단계에서는 KLIPS 가구레벨 데이터에서 각 가구원을 타 가구원에 매칭한 파일을 생성한다.³⁾ 4단계에서는 3단계 파일에 2단계 관계 파일을 병합(merge)하여 최종 데이터를 완성한다.

3) Stata 통계패키지에서 `joinby` 명령어를 이용하면 매칭파일을 생성할 수 있다.

〈표 5〉 가구 내 FRM 구축 과정

단계	설명
1단계	23차 코드북에 정의된 76개 가구주 관계를 매칭하여 5,776개 관계 정의
2단계	5,776개 관계 중 현실에서 관찰되기 어려운 관계는 제외하고 4촌 이내 친인척에 해당하는 1,868개 관계만 선택
3단계	매년 조사된 가구레벨 데이터를 이용하여 가구원-가구원 매칭 파일 생성(표 3 참고)
4단계	3단계 파일에 2단계 파일을 덧붙여 code_ego, code_alter 그리고 ea_rel 변수 추가

본 연구에서는 KLIPS 1차(1998년)~24차(2021년) 가구레벨 데이터와 가구원 관계 파일을 매칭하여 최종 FRM을 완성하였다.⁴⁾ [그림 1]에서는 완성된 FRM 데이터의 주요 변수를 제시하며, 예시로 hhid=1 가구 내에서 각 시점별 가구원 간 매칭을 보여준다. ea_rel 변수는 매칭된 상대방에 대한 본인의 위치를 의미한다. pid_ego는 본인의 가구원 id이고 pid_alter는 매칭된 상대방의 가구원 id이다.

〔그림 1〕 완성된 FRM 데이터

	hhid	wave	ea_rel	code_rel	pid_ego	pid_alter
1	1	1	자녀	11_10	102	101
2	1	2	본인	10_10	101	101
3	1	3	부모	10_11	101	102
4	1	5	부모	10_11	101	102
5	1	6	부모	10_11	101	102
6	1	7	부모	10_11	101	102
7	1	8	본인	10_10	101	101
8	1	9	자녀	11_10	102	101
9	1	10	본인	10_10	101	101
10	1	11	본인	10_10	101	101
11	1	12	본인	12_12	104	104
12	1	13	자녀	12_10	104	101
13	1	14	본인	10_10	101	101
14	1	15	부모	10_12	101	104
15	1	16	자녀	12_10	104	101

4) 완성된 FRM 파일은 논문심사와 한국노동연구원 노동패널팀의 검증을 거쳐 공개할 예정이다.

〈표 6〉 ea_rel 변수 빈도표 : KLIPS 1~24차 FRM

ea_rel	빈도	비율(%)
본인	456,927	28.76
부모	325,450	20.49
자녀	325,450	20.49
배우자	230,682	14.52
형제자매	170,372	10.72
손자녀	22,143	1.39
조부모	22,143	1.39
사위, 며느리	10,637	0.67
시부모, 장인·장모	10,637	0.67
기타 친인척	3,212	0.2
NA	3,084	0.19
삼촌, 이모	2,547	0.16
조카	2,269	0.14
시동생, 시누이, 처제, 처남	1,214	0.08
형수, 올케, 형부	1,214	0.08
사촌	318	0.02
비친인척 동거인	272	0.02
합계	1,588,571	100

완성된 FRM 파일(그림 1)에서 ea_rel 변수의 빈도표는 <표 6>에서 제시한다. 1차~24차 데이터에서 완성한 가구 내 FRM 파일의 관측치는 1,588,571개이다. 본인-본인이 매칭되는 경우를 제외하면 부모-자녀(자녀-부모) 가구원 관계가 가장 많다는 것을 알 수 있다. 가구 내에서는 부모-자녀 관계가 가장 자주 나타나고, 그 다음은 배우자 관계이다. 배우자 관계는 남편-아내가 매칭되는 경우와 아내-남편이 매칭되는 경우를 포함한다. 가구 내에서 형제자매 관계는 전체의 10.7%로 관찰된다. 한 가구 내에서 자녀가 2명 이상이면 형제자매 관계가 나타날 가능성이 크다. 반면, 외동이거나 자녀가 없는 가구에서는 형제자매 관계가 나타나지 않는다. 형제자매 관계는 꼭 자녀에게서만 발견되는 것은 아니다. 가구주와 가구주의 형제자매가 같이 거주하면 형제자매 관계가 발생할 수 있기 때문이다. 그 외 가구원 간 관계는 1%를 조금 넘거나, 1% 미만으로 매우 미미

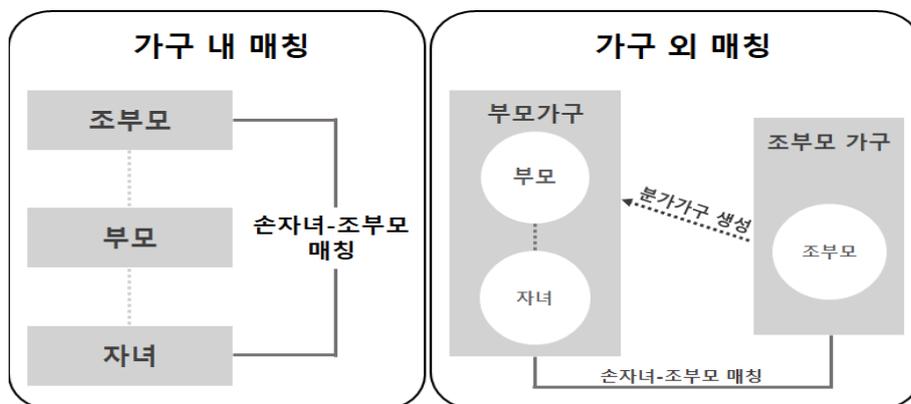
하다. NA는 가구원 간 관계가 1,868개에 속하지 않는 경우이고, 이는 0.19%에 불과하다. 따라서 <표 5>의 2단계에서 정의한 1,868개 관계가 대부분(99.8%)을 차지한다는 것을 알 수 있다.

2. 가구 외 가족관계행렬

PSID의 FRM은 가구 내에서만 가구원 관계를 정의한다. KLIPS의 FRM에서도 PSID와 유사하게 가구 내에서 가구원 관계를 정의할 수 있지만, 본 연구의 새로운 기여는 가구 외 FRM(out-of-household FRM)을 구축하는 작업을 시도하는 것이다. KLIPS는 동일 가구로 조사되었던 부모(A)-자녀(B)가 추후 wave에서 자녀(B)가 독립하여 새로운 분가가구를 생성하는 경우를 추적조사한다. 자녀(B)가 결혼으로 분가하여 새로운 가구를 구성하고 그 가구 내에서 다시 자녀(C)를 갖게 되면, 자녀(B)의 자녀인 C는 최초 가구의 부모(A)와 손자녀 관계로 맺어진다. 그러나 부모(A)와 손자녀(C)는 동일 가구에 속한 적이 없으므로 A와 C의 관계는 KLIPS 내에서 정의될 수 없다. 본 연구에서는 A와 C의 관계를 “가구 외 손자녀(조부모) 관계”로 정의하고자 한다. [그림 2]은 “가구 내 손자녀(조부모) 관계”와 “가구 외 손자녀(조부모) 관계”의 예시를 보여준다.

가구 외 혈연관계에는 손자녀-조부모 외 다양한 관계가 있을 수 있지만, 본 연구에서는 현실적인 코딩작업의 어려움을 고려하여 손자녀-조부모 관계에 초

[그림 2] 가구 내와 가구 외 FRM 예시



점을 맞춘다. 조부모 유형은 친조부모와 외조부모로 분류할 수 있다. 손자녀와 친조부모가 동거한 적이 없더라도 손자녀의 아버지 id를 이용하면 이들을 매칭할 수 있다. 이와 유사하게 손자녀-외조부모가 동거한 적이 없더라도 손자녀의 어머니 id를 이용하면 손자녀와 외조부모의 매칭이 가능하다.

<표 6>에서 ea_rel 변수값이 “손자녀” 또는 “자녀”에 해당하는 표본에 대하여 조부모가 매칭되는 여부를 파악한 결과를 <표 7>에서 정리한다.⁵⁾ KLIPS 데이터 내에서 조부모를 찾을 수 없어 “매칭 없음”에 해당하는 경우는 112,320개로 전체의 67.2%를 차지한다. 노동패널의 조사표본은 동거하는 부모가 가구원으로 포함되는 경우를 제외하고는 부모를 확인할 수 없는 구조이기 때문이다. 한편, “가구 내 친조부모 매칭”은 12.7%를 차지한다. 이는 친조부모와 동거하는 손자녀를 매칭한 경우이다. “가구 외 손자녀-친조부모 매칭”은 비록 손자녀-친조부모가 동거한 적은 없지만 중간에 있는 부모 id를 이용하여 가구 외 매칭을 하는 경우이며, 전체 가구 중 10.2%를 차지한다. “가구 외 손자녀-외조부모 매칭”은 8.7%에 해당한다.

<표 7> 가구 내 & 가구 외 손자녀-조부모 매칭

손자녀-조부모 매칭유형	가구 내/가구 외	빈도(개)	비율(%)
매칭 없음		112,320	67.2
손자녀-친조부모	가구 내	21,364	12.7
손자녀-친조부모	가구 외	17,132	10.2
손자녀-외조부모	가구 외	14,599	8.7
손자녀-친조부모 손자녀-외조부모	가구 내 가구 외	980	0.6
손자녀-외조부모	가구 내	269	0.16
손자녀-친조부모 손자녀-외조부모	가구 내 가구 내	174	0.1
손자녀-친조부모 손자녀-외조부모	가구 외 가구 내	166	0.1
손자녀-친조부모 손자녀-외조부모	가구 외 가구 외	47	0.03
합계		167,051	100

5) “자녀”의 경우 가구주의 자녀로 한정한다.

<표 7>에 따르면 “손자녀-친(외)조부모 동시 매칭” 케이스는 1% 미만이다. 즉, KLIPS 데이터 내에서 손자녀의 친조부모와 외조부모를 동시에 찾아서 매칭할 수 있는 경우는 거의 없다. 결혼을 통해 분가하는 경우, 배우자의 부모가 노동패널 조사표본에 포함될 가능성은 매우 적기 때문이다.⁶⁾ “가구 내 손자녀-친조부모”(12.7%)의 매칭이 “가구 내 손자녀-외조부모”(0.16%) 매칭보다 훨씬 많이 나타난다. 가구 내에서 조부모가 동거하는 경우 친조부모일 가능성이 크다고 해석할 수 있다.

IV. 가족관계행렬의 활용

1. 가구구성의 시계열적 변화

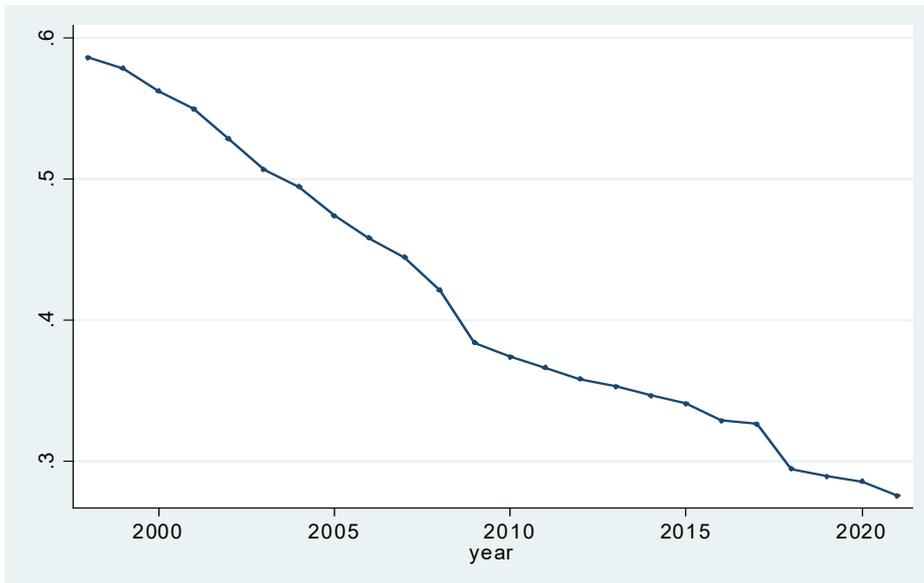
본 소절에서는 가족관계행렬 데이터에 기초하여 KLIPS 조사 내에서 관찰 가능한 가구구성의 시계열적 변화를 제시한다. 앞서 논의하였듯이 KLIPS 원데이터는 가구주(household head)를 중심으로 가구원 관계를 정의하기 때문에 기타 가구원의 가족관계망을 확인할 수 없는 단점이 있다. 그런데 본 연구에서 구축한 가족관계행렬 데이터를 이용하면 가구원을 기준으로 가족구성을 정의하고 그 관계를 살펴볼 수 있다.

가구구성의 첫 번째 주제는 “가구 내 형제자매 관계”이다. “형제자매”는 가구의 자녀가 2명 이상 있는 경우에 정의될 수 있다. “형제자매” 관계는 가구의 자녀뿐만 아니라 가구의 형제자매 또는 가구주 배우자의 형제자매가 동거하고 있는 경우에도 정의할 수 있다. 미국 PSID(표 3 참고)와 한국 KLIPS(표 6 참고)에서 “형제자매” 관계로 나타나는 비율은 각각 20%와 11%로 2배 정도 차이가 난다. 미국 가구구성에 비해 우리나라 가구구성에서 형제자매가 동일가구를 구성하는 경우가 적다고 예상할 수 있다. 혼인가구의 출산율 감소로 무자녀 혹은 외동자녀, 1인 가구가 증가하여 나타난 결과로 보인다.

6) 손자녀-친(외)조부모 가구 내 매칭은 0.1%로 나타난다. 이 경우는 손자녀, 친조부모, 외조부모와 동시에 동거한 시점(wave)이 있다고 해석할 수 있다.

제III장에서 구축한 가구 내 가족관계행렬 데이터에서 가구별로 ea_rel=“형제자매”가 한 번이라도 나타나면 1, 한 번도 나타나지 않는다면 0의 값을 갖는 더미 변수를 생성한다. 그 변수의 평균을 연도별로 계산하면 “형제자매” 관계를 포함한 가구의 비율을 얻을 수 있다. 이 결과를 시계열 그래프로 표현하면 [그림 3]과 같다.⁷⁾

[그림 3] 전체 가구 중 형제자매 관계 포함 가구 비율



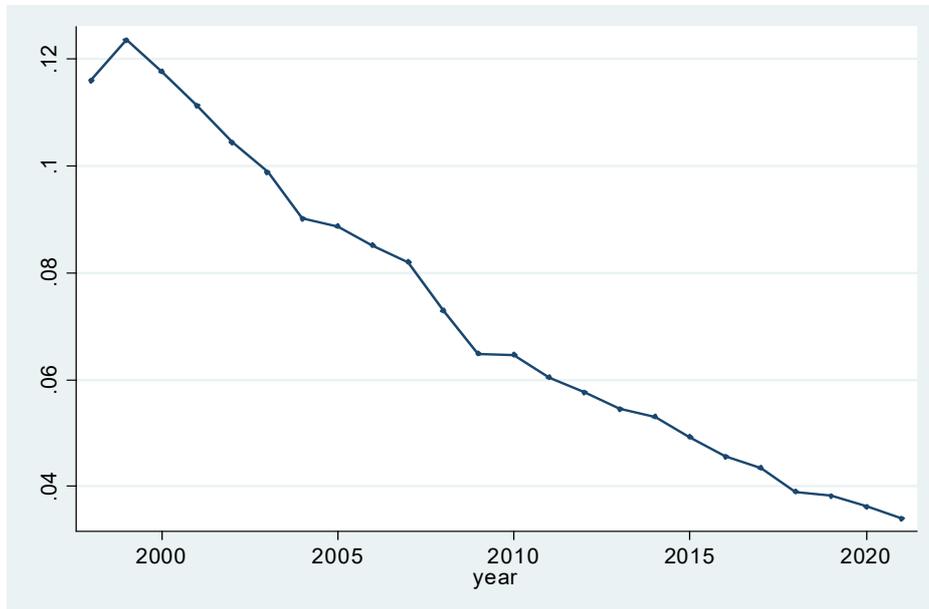
[그림 3]에서 확인할 수 있듯이 KLIPS 조사기간인 1998년부터 2021년까지 형제자매가 포함된 가구의 비율은 꾸준히 감소한다. 1998년에는 “형제자매” 관계가 포함된 가구 비율이 58.6%지만, 그 비율은 지속적으로 감소하여 2021년에는 27.5%로 줄어든다. 즉, 2021년에는 대략 4가구 중 1가구에서만 “형제자매” 관계가 관찰된다.

두 번째 가족구성의 시계열적 변화로 조부모-손자녀 동거 가구에 대한 분석을 제시한다. 손자녀의 조부모가 가구주가 아닌 경우, KLIPS 데이터에서 손자

7) 2008년과 2018년에 “형제자매 관계 포함 가구 비율”이 상대적으로 크게 감소한 이유는 패널마모를 고려하여 추가 표본조사를 실시하였기 때문으로 예상된다.

녀-조부모 관계를 찾기는 쉽지 않다. 그러나 III장에서 제시한 가족관계행렬 데이터에서 가구 내 조부모-손자녀 관계에 해당하는 가구원을 쉽게 찾을 수 있다. 매칭 가구원 중 `ea_rel=="손자녀"`가 있는 케이스는 조부모와 동거하고 있는 가구에 해당한다. [그림 3]과 유사하게 [그림 4]에서는 가족관계행렬에서 `ea_rel=="손자녀"`가 한 번이라도 포함된 가구 비율을 시계열 그래프로 제시한다.⁸⁾

[그림 4] 전체 가구 중 손자녀 관계 포함 가구 비율



[그림 4]에서 확인할 수 있듯이 조부모-손자녀 동거 가구 비율은 1998년과 1999년 사이를 제외하고 꾸준히 감소한다. 1999년 12.3%를 최고점으로 2021년에는 3.4%까지 감소한다. 조부모와 손자녀가 함께 사는 가구는 2021년 100 가구 중 약 3가구에 해당한다는 것을 알 수 있다.

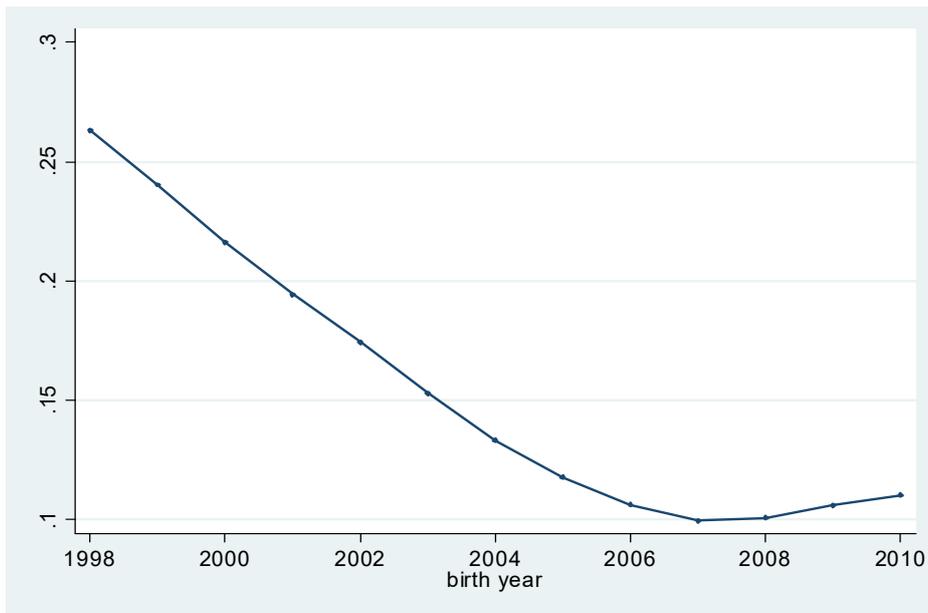
[그림 5]에서는 손자녀의 출생연도에 따른 조부모와의 동거경험 비율을 시계열 그래프로 제시한다. 제III장에서 구축한 가족관계행렬은 가구원 id를 하나의

8) “조부모”는 친조부모와 외조부모를 모두 포함하는 것으로 정의한다. 그리고 할아버지와 할머니 중 1명이라도 같은 가구에 포함되면 “조부모 동거”로 정의한다.

케이스로 가지고 있기 때문에 KLIPS 개인레벨 데이터와 병합할 수 있다는 장점이 있다. 가구원 출생연도를 FRM 데이터와 병합하면 그래프를 작성할 수 있다.

각 가구원 id의 출생연도는 1998년생~2010년생으로 한정한다. 1998년생은 KLIPS 1차 조사인 1998년에 1세, 2010년생은 KLIPS 24차 조사인 2021년에 12세에 해당한다. 해당 표본이 1세~12세(초등학교 졸업) 시점에 해당하는 조사 wave만 남긴다. [그림 5]는 1998년생~2010년생에 해당하는 가구원 중 조사기간 동안(1세~12세 기간) 한 번이라도 조부모와 동거한 경우의 비율을 보여준다.⁹⁾

[그림 5] 출생연도별 조부모 동거경험 비율



[그림 5]에 따르면 1998년생~2010년생 손자녀가 조부모와 동거한 비율은 꾸준히 감소하지만, 2007년생~2010년생 손자녀의 경우 그 비율이 다시 반등한다.¹⁰⁾ 최근 손자녀 양육을 위해 조부모와 동거하는 가구가 다시 증가하고 있다

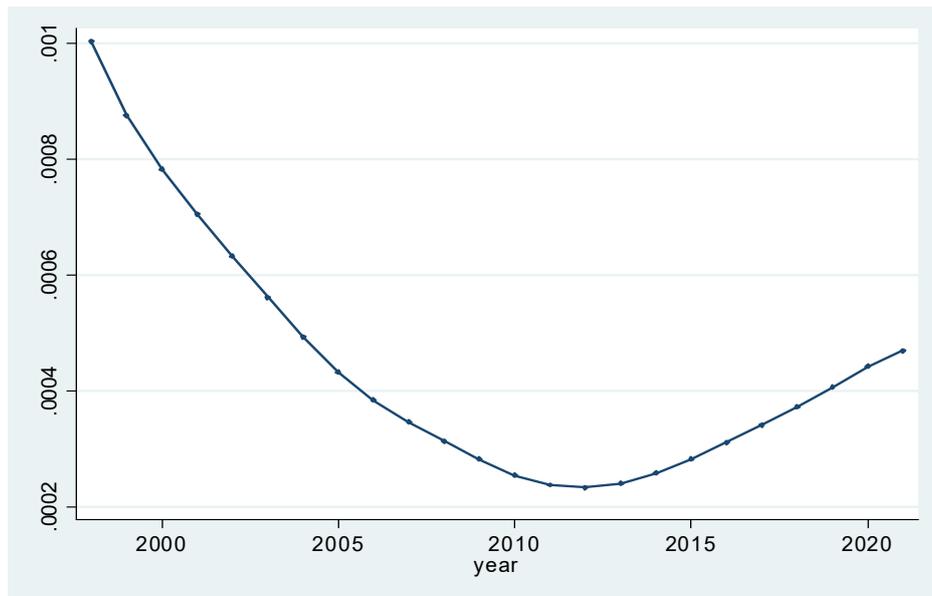
9) 출생연도별 추세를 파악하기 위해 Lowess smoothing을 통해 평활화한 결과를 제시한다.

10) <표 7>의 손자녀-조부모 매칭에서 “매칭없음” 비율은 67%로 높기 때문에 조부모와의 동거경험 비율은 1998년생인 경우에도 26%에 불과하다.

고 예상할 수 있다.

최근 “생활동반자”라는 개념이 언론에서 자주 등장하고 있다. 골자는 혈연이나 혼인 여부와 상관없이 생계와 주거를 공유한다면 “가족”으로 받아들이고, 이들에게 일상가사, 돌봄 등 생애 전 과정에서 가족으로의 권리를 보장하자는 것이다.¹¹⁾ 생활동반자를 비친인척 동거인으로 정의한다면, FRM에서 `ea_rel==` “비친인척 동거인” 관계에 해당한다. [그림 6]에서는 비친인척 동거인을 포함하는 가구 비율을 연도별로 계산하여 제시한다. 2012년까지 그 비율은 꾸준히 감소하지만, 이후 다시 증가 추세를 보인다는 것을 확인할 수 있다.

[그림 6] 비친인척 동거인 포함 가구 비율



2. 가족관계망과 가족관계 만족도의 관계

본 소절에서는 FRM 데이터를 활용하여 가족관계망을 정의하고, 가구원의 가족관계 만족도에 미치는 영향을 분석하고자 한다. KLIPS는 (15세 이상) 가구

11) 자세한 내용은 “생활동반자관계에 관한 법률안”(용혜인 기본소득당 의원 대표발의)을 참고.

원 조사 패널에서 각 가구원의 가족관계 만족도(p_6504 변수)를 조사항목에 포함하고 있다. 가족관계 만족도는 가구원 레벨에서 생활만족도의 한 가지 항목으로 측정되며, 1점(매우 불만족)~5점(매우 만족)과 같이 5점 척도로 구성된다.

FRM을 활용하면 가구주뿐 아니라 개별 가구원이 타 가구원과 어떠한 혈연관계(또는 비혈연 관계)를 가지는지 구체적으로 확인할 수 있다. 본 연구에서는 가족관계망을 가구원 간 촌수로 정의한다. 촌수 값이 클수록 서로 거리가 먼 친인척이라고 예상할 수 있다. 구체적으로 각 가구원과 다른 모든 가구원의 촌수를 계산하고, 촌수의 평균값을 가족관계망으로 측정한다. 촌수는 다음과 같이 배정한다. 친척과 인척을 구별하지 않고 촌수를 배정하며, 기타 친인척은 5촌 이상이라고 가정하고 5촌을 배정한다. 비친인척 동거인은 8촌으로 정의한다. 가구원 간 관계가 정의되지 않은 “NA”는 평균 계산에서 제외한다.

〈표 8〉 가구원 간 촌수

가구원 간 관계	촌수
본인-배우자	0촌
본인-자녀, 부모	1촌
본인-형제자매	2촌
본인-손자녀, 조부모	2촌
본인-시부모, 장인·장모, 사위, 며느리	1촌
본인-사촌	4촌
본인-삼촌, 이모, 조카	3촌
본인-시동생, 시누이, 처남, 처제	2촌
본인-형수, 올케, 형부	2촌
기타 친인척	5촌
비친인척 동거인	8촌

실증분석을 위해 다음과 같이 패널 선형회귀모형을 가정하자.

$$y_{it} = \alpha + \beta x_{it} + \gamma z_{it} + u_i + e_{it} \quad (1)$$

i 는 가구원 id이고 t 는 조사 wave이다. y_{it} 는 1점~5점으로 측정된 가구원 i 의 가족관계 만족도이다. x_{it} 는 가구원 i 가 가구 내에서 접하는 가구원과의 평균 촌수로 회귀모형의 독립변수에 해당한다. z_{it} 는 시간가변 통제변수로 가구원의 나이, 거주지역, 가구주 여부, 배우자 여부, 균등화 가구소득, 가구원 수 등을 포함한다. u_i 는 고정효과(fixed effects)로 가정하고 추정하고자 한다. 이에 따라 가구원 성별과 같은 시간불변 통제변수는 모형 설정에서 제외하더라도 β 에 대한 일치추정량을 얻는 데 문제가 없다고 판단된다.

<표 9>에서는 평균 촌수로 계산한 가족관계망의 기초통계량을 정리한다. 분석의 목적을 고려하여 1인 가구는 분석에서 제외한다. 가구원의 평균 촌수는 0.75촌으로 1촌 미만이다. 부부만 있는 2인 가구에서는 평균=0이 도출되며, 자녀가 있는 가구는 평균=1에 가까울 것이다. 최솟값은 0촌이고 최댓값은 8촌으로 나타난다.

<표 9> 가구원의 평균 촌수에 대한 기초통계량

가구원 수(개)	표본 수(개)	평균	표준편차	최소	최대
33,999	315,887	0.75	0.52	0	8

주: KLIPS는 15세 이상 가구원에 대해서만 가족관계 만족도 항목을 조사함.

식 (1)의 패널 선형회귀모형 추정결과는 <표 10>에서 정리한다. 독립변수는 x_{it} =평균 촌수 변수로 선택하였다. 종속변수가 순서를 가진 범주형이라는 점을 고려하여 추정결과를 패널 순서형 로짓모형과 비교한다.

추정결과에 따르면, 평균 촌수가 높을수록 가족관계 만족도는 낮아진다. 평균 촌수가 높다는 것은 거리가 먼 친인척과 동거하는 경우라고 해석할 수 있다. 즉, 상대적으로 거리가 있는 친인척이 동거하는 가구일수록 가족관계 만족도가 낮아진다고 이해할 수 있다. 통제변수 중 균등화 가구소득은 양(+)으로 유의하며, 소득이 높을수록 해당 가구원의 가족관계 만족도는 증가하는 것으로 나타났다. 패널 순서형 로짓모형의 추정결과는 패널 선형모형의 결과와 유사하였다. 평균 촌수가 증가할수록 가족관계 만족도가 큰 값을 가질 확률, 즉 $Pr(Y=5)$ 는 감소하였다.

〈표 10〉 패널모형 추정결과

	패널 선형회귀모형	패널 순서형 로짓모형
평균 존수 x_{it}	-0.027(0.007)***	-0.298(0.021)***
가구원 나이	-0.006(0.000)***	-0.019(0.0004)***
거주지역: 기타 광역시	-0.0002(0.011)	0.160(0.018)***
거주지역: 기타 시도	0.014(0.009)	0.166(0.016)***
가구주 여부	0.059(0.006)***	0.128(0.015)***
배우자 여부	0.176(0.007)***	0.643(0.022)***
균등화 가구소득	0.111(0.002)***	0.672(0.009)***
가구원 수	-0.017(0.002)***	-0.030(0.007)***
n : 가구원 수	33,798	33,798
N : 전체 표본 수	310,214	310,214

주: 1) *, **, ***는 각각 10%, 5%, 1% 유의수준에서 유의함.

2) 거주지역의 기준그룹은 수도권 거주임.

3) 패널 순서형 로짓모형은 개인 이질성을 확률효과로 가정함.

V. 결론

국내에서 가족관계를 논의한 연구의 대부분은 부모와 자녀, 배우자와의 관계에 초점을 맞추었으므로(김영범·박준식, 2004; 김유경 외, 2013; 김태현·박숙자, 1992; 윤현숙·유희정, 2006) 다른 가족 구성원 간의 관계는 매우 제한적으로 다루어졌다. 가족관계는 다양한 구성원과의 상호작용을 통해 변화할 수 있으나, 관련 데이터의 부재로 사실상 가족관계의 역동을 통시적으로 고찰하기 어려웠다. 이에 본 연구에서는 KLIPS 자료를 활용하여 가족관계력을 살펴볼 수 있는 FRM 데이터를 구축하였으며, 가구구성의 시계열적 변화와 가족관계망-가족관계 만족도의 관계를 실증적으로 분석하였다.

본 논문의 주요 내용은 다음과 같이 요약할 수 있다. 첫째, FRM 자료에서 동일한 가구에 속한 구성원뿐만 아니라 분가한 가구원 간 매칭을 시도하였다. 그 결과, 동거하는 가족에 대해서는 4촌 이내 친인척을 매칭하였으며, 비동거 가족은 손자녀와 친조부모, 외조부모의 관계를 연결하였다. 둘째, 형제자매 혹

은 조부모-손자녀로 구성된 가구 비율의 변화를 분석한 결과, 20년간 지속적으로 감소하는 추세를 보였다. 형제자매가 포함된 가구는 1998년 58.6%에서 2021년 27.5%로 감소하였으며, 조부모와 손자녀로 구성된 가구는 1999년에 12.3%로 소폭 증가하였으나, 이후 꾸준히 감소하여 2021년에는 3.4%에 그쳤다. 셋째, 가족관계망이 가족관계 만족도에 미치는 영향에 대한 분석결과를 살펴보면, 평균 촌수가 높을수록 가족관계 만족도는 낮아졌다. 본 연구에서는 가족관계망을 촌수로 측정하므로 이러한 결과는 혈연관계가 먼 사촌과 동거하는 경우 가족관계 만족도가 감소한다는 것을 보여준다.

본 연구의 함의는 두 가지로 정리할 수 있다. 첫째, 국내에서 처음으로 가족관계력을 분석할 수 있는 FRM 데이터를 구축하였다는 점이다. 기존 2차 자료는 가구주를 중심으로 가족 구성원과의 관계를 조사하였기 때문에 다양한 가족 구성원 간 관계의 역동을 고찰하는 것은 사실상 불가능에 가까웠다. 그러나 본 연구는 가구주 여부와 관계없이 각 가구원을 서로 매칭하여 KLIPS 데이터가 가진 정보를 확장하였다. 예컨대, FRM 데이터에서는 며느리와 시어머니가 매칭되어 있으므로 이들이 각각 응답한 값을 활용하면 관계의 변화를 면밀하게 분석할 수 있다.

둘째, FRM 데이터를 구축하는 과정에서 동거하는 가족뿐만 아니라 비동거하는 가족을 포함하였다는 점이다. 이는 본 연구에서 구축한 FRM 데이터의 장점이자 PSID에서 생성한 FRM 데이터와의 차별점이라 할 수 있다. PSID의 FRM 데이터는 가구 내 구성원 사이의 관계를 정의하지만, 본 연구에서 생성한 FRM 데이터는 이와 더불어 가구 외 구성원 즉, 동거하지 않는 가족과의 관계도 포함한다. 따라서 본 연구에서 구축한 FRM 데이터로는 PSID의 FRM 데이터로 관찰할 수 없는 떨어져 사는 가족과의 관계(가령, 한 번도 함께 산 적 없는 손자녀와 조부모의 관계)를 분석할 수 있다. 이를 활용하면 조부모-손자녀 간 소득이동성 및 교육이동성 등 다세대 간 이동성(multi-generational mobility)을 분석할 수 있을 것이다. 이처럼 본 연구에서 구축한 FRM 데이터는 가구주 여부와 관련 없이 동거 혹은 비동거 가족 구성원 간 관계를 통시적이고 다면적으로 분석하는 데 활용할 수 있으므로 가족관계에 관한 연구의 범위를 확장하는 데 기여할 것으로 기대된다.

참고문헌

- 김미혜 · 신경림(2005). 「한국 노인의 ‘성공적 노후 척도’ 개발에 관한 연구」. 『한국노년학』 25 (2) : 35~52.
- 김영범 · 박준식(2004). 「한국노인의 가족관계망과 삶의 만족도: 서울 지역 노인을 중심으로」. 『한국노년학』 24 (1) : 169~185.
- 김유경 · 진미정 · 송유진 · 김가희(2013). 『가구 · 가족의 변동과 정책적 대응방안 연구』. 한국보건사회연구원.
- 김은경(2002). 「농촌 여성노인과 남성노인의 생활만족도에 영향을 미치는 성인 자녀 관련변인에 관한 연구」. 『가정과삶의질연구』 20 (4) : 27~36.
- 김태현 · 박숙자(1992). 「한국의 비동거 가족 연구」. 『가족학논집』 4 (1992) : 45~74.
- 박경숙(2000). 「한국 노인의 사회적 관계: 가족과 지역사회와의 연계정도」. 『한국사회학』 34 (3) : 621~647.
- 윤현숙 · 유희정(2006). 「가족관계가 성공적 노화에 미치는 영향」. 『한국가족복지학』 18 : 5~31.
- 이서영(2015). 「성별차이에 따른 농촌지역 노인의 가족관계망 특성이 삶의 만족도에 미치는 영향」. 『한국케어매니지먼트 연구』 17 : 181~200.
- 이혜자 · 박경애(2009). 「농촌노인의 가족관계망 유형과 생활만족도」. 『한국노년학』 29 (1) : 291~307.
- Institute for Social Research and Survey Research Center(2021a). “A Panel Study of Income Dynamics : 1968~2019 Relationship Matrix File Documentation (Release 1)”. University of Michigan.
- _____ (2021b). “PSID Family Identification Mapping System User Manual”. University of Michigan.
- Zheng, Z. and H. Chen(2020). “Age Sequences of the Elderly’s Social Network and Its Efficacies on Well-Being : An Urban-Rural Comparison in China”. *BMC Geriatrics* 20 (1) : 1~10.

Abstract

Establishment and Application of FRM(Family Relationship Matrix) Data from KLIPS

Min, Insik · Hong, Kyung-Zoon · Jeon, Yeji

This study utilizes the Korea Labor and Employment Panel Survey(KLIPS) to establish a Korean-version Family Relationship Matrix(FRM) dataset, which identifies and maps familial relationships among all household members, including those within and outside the household. The FRM data enables a diachronic and multifaceted analysis of relationships between family members, regardless of their household status. To demonstrate the potential of the FRM data, this study presents visualizations of changes in household composition over time. Furthermore, the effect of the family relationship network on family members' satisfaction is investigated. Empirical results show that family relationship network degree is positively associated with family satisfaction. FRM data expands the scope of research on family relationships and provides a novel approach to studying relationships between family members.

Keywords : FRM(Family Relationship Matrix), KLIPS(Korea Labor and Income Panel Survey), out-of-household FRM, family satisfaction