

노동정책연구
2024. 제24권 제1호 pp.107~132
한국노동연구원
<http://doi.org/10.22914/jlp.2024.24.1.005>

연구논문

교육 및 노동시장 성과에 대한 환경의 영향력 : KLIPS의 쌍둥이 자료를 이용한 분석

민인식*

본 연구는 개인의 사회경제적 성과에 유전적/환경적 요인이 미치는 정도를 분석하는 것을 목적으로 한다. 특히 유전적 유사성을 통제한 상황에서 나이 차이로 인해 발생하는 공유된 환경의 차이가 자녀 간 사회경제적 성과 유사성에 미치는 영향을 살펴보는 것이 본 연구의 주요한 목적이다. 지난 25년 동안 구축된 한국노동패널에서 쌍둥이 데이터를 구축하여 쌍둥이와 비쌍둥이 형제자매간 자녀 상관계수를 추정한다. 연구결과는 다음과 같다. 첫째, 동갑인 쌍둥이 그룹을 처리군, 나이 차이가 1~5세인 자녀를 비교군으로 설정한 결과, 쌍둥이 그룹에서 교육 성과의 자녀 상관계수가 0.51로 가장 높다. 반면 나이 차이=5세이면 ICC=0.31로 가장 낮다. 둘째, 임금으로 측정된 노동시장 성과 역시 쌍둥이 그룹에서 ICC가 0.81로 가장 크다. 나이 차이=1세이면 0.45이고 나이 차이=5세이면 0.38까지 감소한다. 본 연구에서는 KLIPS 데이터에서 쌍둥이 데이터를 구축하여 이를 활용한 새로운 연구주제를 제안하고, 개인의 사회경제적 성과에 미치는 공유된 환경의 중요성을 시사하고 있다. 초기 어린 시절 공유된 환경을 최대한 제공하고 이를 통해 개인 간 취업, 임금, 사회적 지위 격차를 줄이고자 하는 정책에 대한 근거로 기여할 것으로 기대한다.

핵심용어 : 유전적 요인, 공유된 환경, 쌍둥이, KLIPS, 사회경제적 성과

논문접수일 : 2023년 9월 22일, 심사외뢰일 : 2023년 10월 16일, 심사완료일 : 2023년 11월 29일

* 경희대학교 경제학과 교수(imin@khu.ac.kr)

I. 서론

유전학뿐 아니라 사회과학에서도 개인의 사회경제적 성과가 유전적 요인과 환경적 요인에 의해 어떻게 영향을 받는지는 지속적인 연구주제로 자리매김하고 있다. 초기 연구는 환경적 요인이 개인의 교육 및 노동시장 성과에 결정적 영향을 미친다고 주장하였다. Coleman et al.(1966)은 미국의 교육기회 평등에 관한 연구에서 학교의 자원보다는 학생의 가정 및 사회적 배경이 교육성과에 더 큰 영향을 미친다고 강조한다. Bowles and Gintis(1976)는 학교 성적보다는 학생의 사회적 배경과 가정 환경이 그들의 경제적 및 교육적 성과에 더 큰 영향을 미친다고 주장한다.

행동유전학이 부상하면서 유전적 요인이 인간의 행동과 성취에 영향을 미친다는 주장이 흐름을 주도하게 된다. Plomin and DeFries(1980)는 지능과 유전적 요인의 관계를 탐색하고자 하였다. 쌍둥이와 입양 아동 연구를 통해 지능의 변이가 유전적 요인에 기인한다는 증거를 제시한다. Scarr and Weinberg(1978)는 입양된 아동을 대상으로 한 연구에서 입양 부모의 사회경제적 지위와 아동 지능 사이에는 큰 연관성이 없으며 생물학적 부모의 지능과 입양 아동의 지능 사이에는 높은 연관성이 있다는 결과를 제시하였다. 최근 연구에서는 복잡한 인간 행동과 특성의 발현을 단순히 하나의 영향만으로 설명하기보다는 환경과 유전적 요인의 상호작용이 개인의 인지발달과 사회경제적 지위에 중요하게 영향을 미친다고 주장한다(Caspi et al., 2010; Tucker-Drob and Harden, 2011).

Bouchard and McGue(2003)는 인간행동(인지능력)의 개별적 차이를 이해하기 위해 다음과 같이 가법적(additive) 분산 요소(variance component) 관계를 정의한다. P 는 개인행동 또는 인지능력, G 는 유전적 요인(genetic effect), C 는 공유된 환경 요인(shared environmental effect) 그리고 E 는 공유되지 않은 환경 요인(non-shared environmental effect)을 의미한다. 환경적 요인을 공유된 환경과 공유되지 않은 개별적 환경으로 구분한다. 공유된 환경은 같은 부모 밑에서 자라면서 공유하는 부모 훈육과 가정 배경이고 공유되지 않은 환경은 동료 효

과(peer effect)나 예상치 못한 우연한 기회(accidents) 등 자녀 개인에게 적용되는 이질적인(heterogeneous) 환경 효과라고 이해할 수 있다.¹⁾

$$\text{var}(P) = \text{var}(G) + \text{var}(C) + \text{var}(E) \quad (1)$$

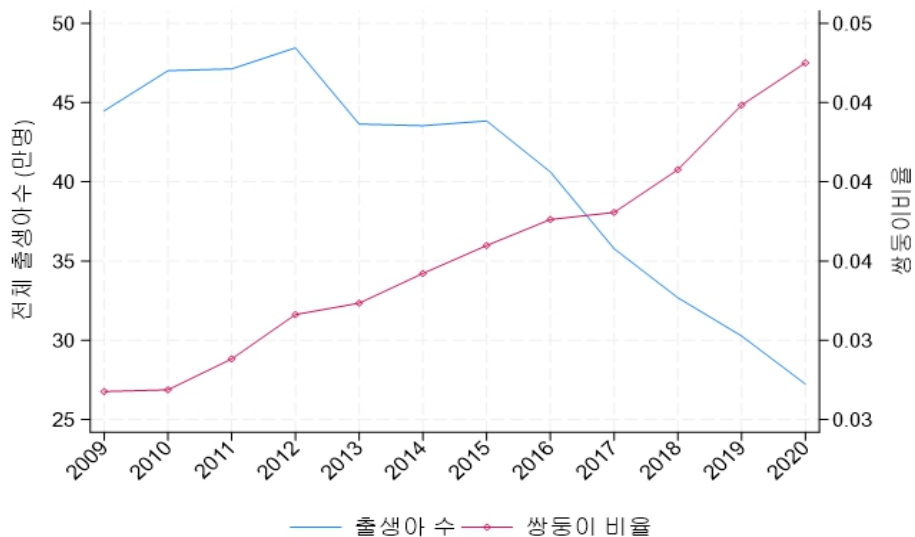
본 연구에서는 개인의 사회경제적 성과 차이가 식 (1)과 같은 3가지 요소에 의해 영향을 받는다고 가정한다. 다만 분석대상을 같은 부모하에서 자란 자녀로 한정하여 유전적 요인 G 는 통제하고자 한다. 다만 유전적 요인을 통제된 상태에서 환경의 영향력을 분석하기 위해 본 연구에서는 쌍둥이 그룹(처리군)과 쌍둥이가 아닌 형제자매(비교군)를 비교하고자 한다. 쌍둥이는 일란성과 이란성 쌍둥이 유형이 있지만 일란성 비율이 현저히 낮다는 점과 본 연구의 실증분석에서 사용하는 한국노동패널(Korea Labor and Income Panel Survey : KLIPS)에서 일란성과 이란성 구별이 불가능하다는 점을 고려하여 모두 이란성 쌍둥이(Dizygotic Twin : DZ)로 간주하기로 한다. 이란성 쌍둥이와 쌍둥이가 아닌 형제자매(non-twin siblings)의 유전적 유사성은 모두 50%로 알려져 있다²⁾. 두 그룹 간 유전적 유사성이 동일하다면 처리군과 비교군의 유일한 차이는 나이 차이(age difference)이다.³⁾

Koepfen-Schomerus et al.(2003)은 이란성 쌍둥이의 인지적 유사성이 쌍둥이가 아닌 형제자매보다 더 크다고 주장한다. 어린 시절 공유된 환경 추정치를 이러한 결론에 대한 근거로 제시한다. 영국과 웨일스 쌍둥이 데이터를 이용한 연구에 따르면 쌍둥이에 대한 공유된 환경의 추정치가 비쌍둥이 형제자매에 비해 두 배 이상 크다. 쌍둥이에 대한 특별한 공유된 환경 효과를 설명할 많은 가능성이 있지만 해당 연구에서는 형제자매가 공유하지 않았던 인지능력과 관련된 경험들이 쌍둥이에게는 동일한 나이 때문에 공유된다고 주장한다.

-
- 1) 공유된 환경의 구체적인 예를 들면 부모의 사회경제적 배경(지위), 자녀 양육환경 및 태도 그리고 자녀 간 유대관계 등이 있다.
 - 2) 어떤 유전자에 대해, 한 형제자매(쌍둥이 또는 non-twin siblings)가 부모로부터 특정 대립유전자(allele)를 물려받으면, 다른 형제자매가 같은 부모로부터 같은 대립유전자를 물려받을 확률은 50%이고, 다른 대립유전자를 물려받을 확률도 50%이다. 이 확률은 어머니와 아버지의 대립유전자 모두에 적용되므로, 전체적으로 두 형제자매가 같은 유전자에 대해 같은 두 대립유전자를 물려받을 확률은 25%이다. 부모로부터 다른 대립유전자를 물려받을 확률도 25%이며, 한 부모로부터는 같은 대립유전자를, 다른 부모로부터는 다른 대립유전자를 물려받을 확률도 50%이다. 쌍둥이 또는 비쌍둥이 형제자매의 예상되는 유전적 유사성은 약 50%이다.
 - 3) 자녀 개인별 이질적 환경에 해당하는 E 는 무작위적인 확률오차(random error)로 가정한다.

유전적 유사성을 통제한 상황에서 나이 차이로 인해 발생하는 공유된 환경(C)의 차이가 자녀 간 사회경제적 성과 유사성에 미치는 영향을 살펴보는 것이 본 연구의 주요한 목적이다. [그림 1]에서는 2009년 이후 우리나라 출생아 수와 그 유형에 대해 보여주고 있다. 출생아 수는 2012년 48만 명을 정점으로 꾸준히 감소한다. 그러나 쌍둥이 출산 비율은 꾸준히 증가하여 2020년에는 전체 출생아 수의 4.75%까지 증가하는 추세이다.

[그림 1] 전체 출생아 수와 쌍둥이 비율



자료 : www.kosis.kr의 출생아 수 자료를 이용하여 저자 작성.

1998년(1차 조사) 시작되어 2022년(25차 조사)까지 수행된 장기간 패널조사인 KLIPS의 장점은 분석에 필요한 쌍둥이 표본을 충분히 찾을 수 있다는 것이다. 서로 동갑인 쌍둥이의 아동·청소년기 공유된 환경이 교육과 노동시장 성과 유사성에 미치는 영향을 자녀 상관계수(sibling-correlation) 추정을 통해 이해하고자 한다.

논문의 구성은 다음과 같다. 제II장에서는 유전적·환경적 요인과 자녀 상관계수 관계를 분석한 선행연구를 정리한다. 제III장에서는 KLIPS 1~25차 데이터에서 구축한 쌍둥이 데이터에 관해 설명하고 쌍둥이 그룹의 주요 변수에 대한 기초통계량을 정리한다. 제IV장에서는 처리군과 비교군에 속한 자녀 상관계수를 추정하기

위한 실증분석 모형과 그 결과를 제시한다. 마지막으로 제V장에서는 실증분석 결과를 요약하고 연구가 가지는 함의를 제시한다.

II. 연구배경 및 선행연구 검토

1. 유전적·환경적 요인과 자녀 상관계수

자녀 간 인지능력이나 사회경제적 성과의 유사성(similarity) 또는 비유사성(dissimilarity)에 관한 연구는 주로 교육수준과 임금 변수의 자녀 상관계수를 추정하는 실증적 방법을 사용한다. Conley and Glauber(2007)는 미국의 PSID (Panel Survey of Income Dynamics) 데이터를 이용하여 형제간 노동시장 및 사회경제적 결과에서 인종 차이를 분석하였다. 연구결과 흑인 형제는 백인보다 더 유사해지는 경향이 나타난다. 이러한 결과는 가구 배경이 사회경제적 성과에 미치는 영향이 흑인이 백인 형제에 비해 더 약하다는 것을 시사한다. 이러한 동적 변화는 가구 자원이 부족하여 노동시장 초기 단계에서 자신을 뒷받침해 주기 어려운 흑인에게서 더욱 두드러진다. 본 연구의 쌍둥이 연구와 Conley and Glauber(2007)는 다음과 같은 점에서 서로 연결되어 있다. 즉, 가구 배경과 같은 환경적 요인이 개인의 사회경제적 성과에 미치는 영향을 분석하고, 가구 배경은 형제자매 간 유사성에 중요한 역할을 하지만 그러한 환경적 영향은 인종에 따라 다르게 나타날 수 있다는 것을 제시하고 있다.

Mazumder and Levine(2003)은 미국 NLSY(National Longitudinal Surveys of Youth)의 2개 코호트(NLSY66과 NLSY79) 표본에 기초하여 형제간 소득 상관관계가 0.26에서 0.45로 증가하였음을 추정하였다. 이는 자녀 간 공유되는 가구 및 지역사회 영향이 사회경제적 성과를 결정하는 데 더 중요해지고 있음을 나타낸다. 이들의 연구는 본 연구와 유사하게 형제자매 간 경제적 성과에 대한 상관관계를 추정하고 있고 유전적 요인과 더불어 가구 및 지역사회 환경이 경제적 성과에 미치는 영향에 대한 추가적인 이해를 제공하고 있다.

자녀 간 소득뿐 아니라 교육수준 역시 자녀 상관관계의 상당 부분을 가구 배경

에 의해 설명할 수 있다는 논의가 있었다. Hauser and Mossel(1985)은 518명의 위스콘신 지역 고등학교 졸업생과 그 형제의 직업과 학력을 조사한 자료를 이용한다. 가구 구성원은 교육수준 분산(변동성)의 50%와 직업적 지위 분산의 1/3 이상을 차지하지만, 가구배경을 통제하지 않더라도 교육수준이 직업적 지위에 미치는 영향이 과대평가(upward bias)되지 않을 것이라고 논의하였다.

Cawley et al.(2023)은 자녀 간 교육수준에 구체적인 유형의 동료효과(peer effect)를 검증한다. 구체적으로 자녀 중 한 명의 교육수준이 다른 자녀의 교육성과에 대한 다항 유전 점수(Polygenic Score : PGS)와 어느 정도 관련되었는지 추정한다⁴⁾. 미국의 NSLY 청소년기~성년 기간에 나타난 유전 데이터를 사용하여 추정된 모형에 따르면 형제-자매간 이러한 유전적 양육(genetic nurture)에 대한 근거를 제시하고 있다. 특히, 형제나 자매의 PGS가 높으면 해당 개인이 대학 학위를 가질 확률이 7.1%에서 7.8% 증가한다. 자기보다 어린 형제가 있는 경우, 성별이 다른 남매 그리고 교육수준이 평균 이하인 PGS를 가진 개인의 경우에 이러한 효과는 더 강하게 나타나는 결과를 얻었다. Cawley et al.(2023)의 연구는 유전적 요인(여기서는 PGS)이 개인의 교육성과에 미치는 영향을 분석하고 있다는 점에서 본 연구와 연결 지점이 존재한다. 특히 이들의 연구는 동료 효과와 유전적 양육을 통해 같은 가정에서 성장한 형제자매 간 상호작용에 대해 논의하고 있다.

Rowe et al.(2003)은 유전 가능성은 부모 교육수준이 우수한 환경에서 더 높아진다고 주장한다. 이를 위해 부모의 교육수준이 어떻게 언어 IQ 변화에 대한 유전적 및 환경적 조절효과를 갖는지 분석하였다. 전체 표본(일란성, 이란성, 동일 부모 형제자매, 같은 가정의 사촌, 생물학적으로 관련 없는 형제)에서 유전 가능성에 대한 분산 추정치는 공유환경에 대한 추정치보다 크게 나타난다. 그러나 유전 가능성과 공유된 환경 역시 모두 부모 교육수준에 의해 조절된다. 구체적으로 교육수준이 높을수록 유전 가능성이 높고 공유된 환경 분산 추정치가 작아진다. 반면 교육수준이 낮은 가구에서는 유전 가능성이 감소하고 공유된 환경 영향이 더 크게 작용한다. 이 연구는 부모의 교육수준이라는 특정 환경적 요인이 형제자매 간 유사성에 미치는 유전적 및 환경적 기여를 분석한다. 이를 통해 유전과 환경의 상호작용을 이해할 수 있는 중요한 논의를 포함한다.

4) 교육수준 PGS는 교육수준과 관련된 다양한 유전자 변이를 기반으로 하며, 이 점수가 높을수록 교육수준이 높을 것으로 예상된다.

Dong et al.(2023)은 부모의 교육수준과 청소년기 동료(peer)의 교육환경이 더 높을 때 언어 지능의 유전 가능성이 증가하고 공유된 환경 효과는 감소한다는 결과를 얻었다. 동료의 조절 효과는 성인기까지 지속되며, 이러한 효과는 인지 발달 과정에서 발생한 새로운 유전-환경 상호작용에 부분적으로 기인하기 때문으로 예상된다. 지능의 유전 가능성이 어떻게 부모의 교육수준과 같은 환경적 요인에 의해 영향을 받는지 분석하였다는 점에서 본 연구와 연결 지점이 존재한다.

유혜림(2022)은 KLIPS 데이터를 이용하여 자녀 간 사회경제적 성과의 유사성을 추정했다는 점에서 본 연구와 관련이 깊다. 이 연구에서는 자녀 상관관계를 통해 부모 배경효과를 보는 것이 목적이다. 교육수준 전체 분산의 약 36%를 가구 요인이 설명한다. 임금의 경우 교육에 비해 가구 배경효과가 상대적으로 작다. 형제보다 자매간 임금의 자녀 상관계수가 더 높게 나타나는 경향이 있다는 것을 발견하였다.

본 연구는 Haveman and Wolfe(1995)의 Wisconsin 모형과 밀접한 관련이 있다. Wisconsin 모형은 부모의 사회경제적 상태가 자녀의 교육 및 노동시장 성과에 미치는 영향을 중점적으로 다루고 있다. 본 연구는 이러한 맥락에서 한 걸음 더 나아가 유전적 및 환경적 요인의 상호작용이 어떻게 자녀의 사회경제적 성과에 영향을 미치는지를 구체적으로 탐구하고자 한다. 특히, 공유된 환경의 영향력에 초점을 맞추고 있으며 이는 Wisconsin 모형에서 강조하는 가족의 사회경제적 배경이 자녀의 성과에 미치는 영향과 맥락적으로 연결된다. 본 연구는 가구 내의 공유된 환경이 어떻게 자녀의 교육 및 노동시장 성과에 영향을 미치는지에 대한 더 깊은 이해를 제공함으로써 교육 정책의 설계에 중요한 시사점을 제공할 수 있을 것으로 기대할 수 있다.

2. 쌍둥이 관련 연구

쌍둥이 연구는 일란성(Monozygotic twins : MZ)과 이란성(DZ) 쌍둥이를 비교 대상으로 삼아 유전적 동일성이 쌍둥이 성과 차이에 미치는 영향을 분석하는 것이 대표적인 연구 방향이다. 개인의 심리적 특성에 영향을 미치는 유전과 환경 요인을 이해하기 위해 유전적으로 같은 일란성 쌍둥이와 50% 동일한 이란성 쌍둥이의 상관계수를 비교한다(허윤미, 2005). MZ와 DZ 간 환경 요인이 같다고 가정할 때

MZ의 상관계수가 DZ의 상관계수보다 높다면 그 차이는 유전적 요인이 작용한다고 예상할 수 있다. 개인의 심리적 특성이 유전적 영향에 의해서만 결정되고 공유된 환경의 영향이 없다면 DZ의 상관계수는 MZ의 상관계수의 절반이 된다. 만약 추정된 DZ의 상관계수가 MZ의 상관계수 절반보다 크다면 이는 공유된 환경이 그 특성에 영향을 미친다고 해석할 수 있다.

Loehlin and Nichols(1976)는 고등학교에 재학 중인 쌍둥이 학생에게 자아검사를 실시하여 일반 자아개념(general self-concept)에 미치는 유전적 영향이 약 40% 정도 되는 것을 확인하였다. McGuire et al.(1999)은 720쌍의 쌍둥이 형제 자매에게 자아개념 척도를 적용한 결과 유전의 영향이 최대 61%에 이른다는 것을 발견하였다. 공유된 환경의 영향은 의미있게 나타나지 않았고 유전 이외는 비공유 환경과 측정오차에 의해 설명되었다.

다른 자녀와 달리 쌍둥이의 특징이 동갑이라는 점에 초점을 맞추어 “공유된 환경”이 쌍둥이의 인지능력에 미치는 연구도 존재한다. Plomin and Daniels (2011)는 개인행동 특성을 결정하는 데 유전보다 환경적 영향이 중요하다고 주장하며, 이를 한 가구 내에서 자란 입양 아이에 관한 연구를 통해 보여주고 있다. 동일한 환경을 공유하지만 유사한 심리적 특성이 나타나지 않는 것은 개별적으로 독특한 비공유 환경이 개인의 심리적 특성에 기여하는 것을 의미한다. Matteson et al.(2013)은 쌍둥이와 입양 데이터를 동시에 모델링하여 청소년 성격에 대한 공유된 환경의 영향을 추정하였다. 다차원 성격 질문지(Multi-dimensional Personality Questionnaire : MPQ)에서 흡수(Absorption) 분산의 15%, 소외감(Alienation) 분산의 10%, 피해 회피(Harm Avoidance) 분산의 14% 그리고 전통주의(Traditionalism) 분산의 26%에서 중요한 공유환경 영향을 발견하였다.

허윤미(2005)는 서울 쌍둥이 가족 연구에 등록된 초등학생 226쌍의 일관성 쌍둥이와 168쌍의 이관성 쌍둥이를 대상으로 자아개념 검사를 실시하였다. 유전의 영향은 12~35%이고 공유된 환경의 영향은 14~38% 정도 요인으로 작용하는 것으로 나타났다. 특히 행동, 외모, 신체적 특성과 불안감 영역의 자아개념은 유전 영향력이 약하고 공유/비공유 환경이 주된 역할을 한다는 결과를 얻었다.

쌍둥이에 대한 실증분석은 표적 집단을 선정하고 설문조사 및 검사를 진행하는 방식이 대부분이다. 본 연구에서는 이러한 제약을 넘어 패널서베이에서 관찰 가능

한 쌍둥이를 식별하고 이들의 사회경제적 성과를 비교하고자 한다. 기존 쌍둥이 연구가 개인의 심리적·인지능력에 미치는 유전적/환경적 요인의 영향을 분석하는 데 비해 본 연구에서는 장기간 패널조사에서 확인할 수 있는 개인의 사회경제적 성과에 미치는 환경 효과를 분석하고 있다는 점에서 차별화가 존재한다.

III. 분석자료

1. KLIPS 쌍둥이 데이터 구축

민인식 외(2023)의 연구에서는 KLIPS 1~24차 데이터를 이용하여 가족관계행렬(Family Relation Matrix : FRM)을 부가 데이터로 생성하고 그 데이터를 활용하는 방법을 소개하고 있다. 기존 KLIPS 데이터에서 가구 구성원 간 관계는 “가구 주와의 관계”에 의해 정의된다. 가구주가 아닌 가구원 간 관계를 새롭게 정의한 데이터가 FRM이다. 본 연구에서는 가구 내 쌍둥이를 식별하기 위해서 FRM의 “가구원 간 관계” 코드를 이용한다. FRM 데이터에서는 [그림 2]와 같이 pid_ego(본인 pid)와 pid_alter(동일 가구 내 다른 가구원 id)가 주어진다. ea_rel 변수에서는 상대 가구원에 대해 본인의 역할을 확인할 수 있다. [그림 2]에서 hhid=44 가구에 속한 pid_ego=4403과 pid_alter=4404 가구원이 서로 형제자매 관계로 나이가 서로 같다는 것을 확인할 수 있다. 따라서 가구주의 자녀이면서 쌍둥이 관계라는 것을 예상할 수 있다. pid_ego=5703과 pid_alter=5704 역시 쌍둥이 관계에 해당한다.⁵⁾

FRM 데이터에서 다음과 같은 조건을 만족하는 개인 id 쌍(pair)을 본 연구의 분석대상에 해당하는 쌍둥이로 선택한다. 조건 1과 조건 2에 해당하는 두 사람의 나이가 가장 어린 조사 시점(wave)에서 가구주의 자녀 또는 손자녀인 경우만 남

5) FRM 데이터에서 pid_ego(본인)를 중심으로 같은 가구 내 다른 가구원(pid_alter)을 매칭시킨 데이터이다. 따라서 가구원이 4명인 가구에서는 총 16가지 매칭이 발생한다 (본인-본인 매칭 포함). 매칭된 가구원 간 관계가 “자녀”, “배우자”, “형제자매”, “부모”의 관계를 정의한 변수를 포함하고 있다. 그 변수가 ea_rel 변수이다. ea_rel 변수값 중에서 “형제자매”에 해당하는 매칭만 남긴 후 그 형제자매 간 나이가 서로 같으면 쌍둥이로 식별할 수 있다.

졌다. 가능한 최근 출생 코호트를 분석하기 위해 생년을 1970년 이후로 한정하고, 겹쌍둥이 또는 세쌍둥이 가구에서는 2명만 남기기 위해 첫째-둘째만 분석대상에 포함한다.⁶⁾

[그림 2] FRM에서 쌍둥이 예시

	hhid	pid_ego	age_ego	pid_alter	age_alter	ea_rel
1	44	4403	3	4404	3	형제자매
2	44	4404	3	4403	3	형제자매
3	57	5703	4	5704	4	형제자매
4	57	5704	4	5703	4	형제자매

<표 1> 분석대상 쌍둥이 식별 조건

조건 1	ea_rel=="형제자매"로 정의되는 관계
조건 2	pid_ego와 pid_alter의 나이가 같은 경우
조건 3	가구주의 자녀 또는 가구주의 손자녀인 경우
조건 4	생년이 1970년 이후인 경우
조건 5	겹쌍둥이 또는 세쌍둥이인 가구에서는 첫째-둘째만 분석대상에 포함
최종 분석대상	167쌍(334명)

분석대상에 포함된 전체 쌍둥이 가구는 167쌍(334명)이며, 1972년생~2022년생 쌍둥이까지 찾을 수 있다. 쌍둥이 성이 다른 경우(이성)는 42쌍으로 전체의 25%, 동성(남성) 쌍둥이는 68쌍으로 전체의 40% 그리고 동성(여성) 쌍둥이는 57쌍으로 전체의 34%를 차지한다.

KLIPS는 패널조사이므로 쌍둥이를 매 조사 시점에서 반복적으로 관찰하여 그들의 변화를 추적하는 장점이 있다. KLIPS의 개인레벨 조사는 15세 이상인 가구원에 대해서만 설문을 답하고 있다. 따라서 쌍둥이 데이터와 개인레벨 데이터를 병합하면 15세 미만 시점의 쌍둥이는 분석대상에서 제외된다. 쌍둥이 조건과 15세 이상 조건을 모두 만족한 개인은 163명이고 1,371개 관측치로 나타난다. 이에 대한 기초통계량은 <표 2>에서 정리한다.

6) 1~25차 데이터에서 조건 1~4를 모두 만족한 겹쌍둥이 가구는 2가구이고 세쌍둥이 가구는 6가구를 찾을 수 있었다.

1~25차에서 반복 조사되는 경우를 모두 고려하면 163명이 1,371개 관측치로 주어진다. 평균 나이는 27세이고 최소 15세 그리고 최고령은 50세 시점까지 관찰된다. sex(성별) 평균 0.5 미만으로 남자보다는 여자가 더 많이 포함되어 있다. 쌍둥이가 취업한 경우는 모두 642개 관측치로 전체의 46%를 차지한다. 평균 임금(자영업자는 월 평균 소득을 계산)은 230만 원임을 확인할 수 있다.

〈표 2〉 쌍둥이에 해당하는 개인의 기초통계량

	obs	평균	표준편차	최솟값	최댓값
age	1,371	27	8.8	15	50
male (남자=1, 여자=0)	1,371	0.44	0.49	0	1
edu (교육수준 6개 범주)	1,371	3.88	1.42	2	6
econstat (1=취업자, 2=실업자, 3=비경활)	1,371	2.03	0.98	1	3
job_type (1=임금근로자, 2=비임금근로자)	642	0.89	0.30	0	1
wage (임금)	617	230	189	0	3,827

개인 : 163명 관측치 : 1,371개

주 : 1) wage(임금)는 월 평균 임금(자영업자는 소득), 단위는 만 원, 2022년 기준 실질임금.

2) edu는 6개 범주로 구성되어 있으며 1=무학, 2=고졸 미만, 3=고졸, 4=대재 및 중퇴, 5=전문대졸, 6=대졸 이상.

쌍둥이 간 교육수준과 임금(소득)의 상관계수를 계산한 결과는 〈표 3〉에서 정리한다. 먼저 교육수준을 1~6 범주 연속형으로 간주하고 피어슨(Pearson) 상관계수를 계산한다. 쌍둥이의 취업상태가 비임금근로자(자영업자)인 경우도 포함하는 경우와 취업상태를 임금근로자로 한정된 표본에서 상관계수를 제시한다.

쌍둥이 자녀의 교육수준은 시간에 따라 반복적으로 관찰되지만, 최종 학력만 고려하기 위해 조사 wave 중에서 교육수준이 가장 높은 값을 선택한다. 개별 쌍둥이의 교육수준을 연속형 변수로 간주하고 상관계수를 계산한다. 임금(소득) 변수

의 쌍둥이 간 상관계수는 임금근로자와 비임금근로자인 개인을 모두 포함하여 계산한 결과와 임금근로자인 쌍둥이만을 포함한 상관계수를 모두 계산한다. <표 3> 결과에 따르면 교육수준의 상관계수는 0.70으로 매우 높다는 것을 확인할 수 있다. 다만 1980년생과 1990년생의 대학 진학률이 상대적으로 높다는 점을 고려하면 쌍둥이 대부분은 대학졸업 이상 학력을 가지고 있을 가능성이 크다. 임금 상관계수는 비임금근로자를 포함하면 상관계수는 0.60으로 상대적으로 낮다. 그러나 임금근로자로 한정하면 상관계수가 0.72로 더 높아지는 것을 확인할 수 있다.

<표 3> 쌍둥이 간 교육수준과 임금의 상관계수

	교육수준	임금(소득) 수준	
		임금&비임금 근로자 모두 포함	임금근로자만
$corr(Y_{j, T1}, Y_{j, T2})$	0.70	0.60	0.72

주: 같은 가구 j 에 속한 쌍둥이 첫째 $T1$ 과 쌍둥이 둘째 $T2$ 의 상관계수.

<표 3>에서 알 수 있듯이 쌍둥이 간의 교육/임금에서 계산된 높은 양의 상관관계가 유전적 유사성을 지닌 다른 자녀(non-twin siblings)에게도 나타나는지 확인할 필요가 있다. 같은 부모 환경 아래에서 성장한 자녀 간 사회경제적 성과 유사성이 쌍둥이에게 더 크게 나타난다면 제II장에서 논의한 바와 같이 공유된 환경의 영향력을 예상할 수 있다.

2. 처리군과 비교군 설정

본 연구에서 공유된 환경의 정도는 두 자녀 간 나이 차이에 의해서 정의한다. 특히 쌍둥이 그룹은 공유된 환경 정도가 가장 높은 그룹으로 가정하고 있다. 쌍둥이는 나이 차이=0이기 때문에 공유된 환경의 정도가 가장 큰 처리군(treatment group)으로 가정한다. 쌍둥이 그룹과 비교를 위해 공유된 환경의 정도가 다른 자녀그룹을 비교군(comparison group)으로 설정한다. 비교군에 속한 쌍둥이가 아닌 자녀그룹은 다음과 같이 5개 자녀그룹으로 설정한다.

비교군은 형제자매(가구주의 자녀) 중에서 나이 차이가 1~5세인 경우에 해당하는

다. 나이 차이가 클수록 같은 부모 밑에서 자랐지만 공유된 환경이 상대적으로 적다고 가정한다. KLIPS에서 15세 이상 가구원에 대해서만 개인레벨 변수가 존재하기 때문에 처리군과 마찬가지로 비교군 선정에서도 15세 이상 형제자매로 한정한다. 나이 차이=2세인 형제자매를 KLIPS에서 가장 많이 찾을 수 있다는 것을 아래 결과에서 확인한다.

〈표 4〉 처리군과 비교군의 정의 및 대상 표본 수

	나이 차이	표본 수	비고
처리군 (twins)	0	163명	
비교군 (non-twin siblings)	1	1,242명	1) 쌍둥이가 속한 가구 내 자녀 쌍은 비교군에서 제외 2) 한 가구 비교군에 속한 자녀 쌍이 2쌍 이상이면 첫째-둘째만 분석대상에 포함
	2	3,142명	
	3	2,048명	
	4	1,246명	
	5	792명	

주 : 비교군 역시 15세 이상인 경우로 한정.

IV. 실증분석 : 교육과 노동시장 성과

1. 성과변수의 기초통계량

본 연구의 실증분석에서 사회경제적 성과변수로 분석대상 개인의 교육수준과 취업한 개인의 임금을 측정한다. 쌍둥이 처리군과 마찬가지로 비교군에서도 교육수준은 조사 wave 중에서 최고 학력을 최종 교육수준으로 정의한다. 제Ⅲ장 〈표 3〉에서 언급하였듯이 모든 대출자를 하나의 범주로 묶는 것이 높은 대학진학률을 고려할 때 적절한지 의문이다. 실증분석에서는 개인의 교육성과를 다음과 같이 6개 순서형 값으로 재구성하고자 한다. 새로운 교육수준 변수(edu_new)에서는 고졸 이하=1, 대학중퇴 및 전문대 졸=2, 그리고 4년제 대학 졸업자 중에서 대학 랭

킹에 따라 3~6 그룹에 배정하였다. KLIPS에서는 졸업한 대학의 코드를 제공하고 있다. 해당 변수를 이용하여 4년제 대학졸업자의 대학 랭킹을 4개 그룹으로 배정하였다. 본 연구에서 사용한 대학랭킹은 이지영·고영선(2023)의 부록 <표 1>에서 제시한 대학랭킹을 그대로 활용하였다.

<표 5> 교육성과 변수의 재구성

교육수준 원변수 : edu	새로 구성한 교육수준 변수 : edu_new
1 무학	1 무학, 고졸미만, 고졸
2 고졸 미만	2 대재 및 중퇴, 전문대졸
3 고졸	3 4년제 대학 중에서 하위 40% 가야대 외 48개 대학, 방송통신대학, 사이버대학
4 대재 및 중퇴	4 4년제 대학 중에서 40~60% 건국대 충주캠퍼스 외 28개 대학
5 전문대졸	5 4년제 대학 중에서 60~80% 강원대학교 외 32개 대학
6 대졸 이상	6 4년제 대학 중에서 상위 20% 서울대학교 외 31개 대학, 사관학교

주 : 3~6그룹에 속한 구체적인 대학교 이름은 이지영·고영선(2023)의 Table A.1을 참고.

노동시장 성과변수로는 <표 2>에서 제시한 바와 같이 임금근로자인 경우에는 월 평균 임금(wage)을 선택하고 비임금근로자는 월 평균 소득(income) 변수를 사용한다.

2. 실증분석 모형

자녀 간 교육 및 노동시장 성과변수의 상관계수를 추정하기 위해서 Conley and Glauber(2008)와 유혜림(2022) 연구에서 사용한 멀티레벨 회귀모형(multi-level regression model)을 설정한다.

먼저 교육성과 변수는 1~6의 순서형 범주에 속하지만 본 연구모형에서는 연속형 변수로 간주하고 다음과 같이 two-level 모형을 설정한다. 각 개인에게 교육성과는 최고 학력수준으로 한 번만 주어지는 데이터 구조이다. 식 (2)에서는 종속변수의 분산을 상위레벨(가구 배경)과 하위레벨(개인 특성)의 분산 합으로 형성된 무조건부 모형(unconditional model)을 설정한다.

$$\begin{aligned} edu\ new_{ij} &= \alpha + u_j + e_{ij} \\ \text{where } cov(u_j, e_{ij}) &= 0 \end{aligned} \quad (2)$$

식 (2)에서 교육성과의 분산 $\sigma_{edu\ new}^2 = \sigma_u^2 + \sigma_e^2$ 임을 쉽게 이해할 수 있다. 위 결과를 이용하여 j 가구 내에 속한 쌍둥이 i, k 의 교육성과 상관계수는 다음과 같이 도출할 수 있다.

$$corr(edu\ new_{ij}, edu\ new_{kj}) = \frac{cov(edu\ new_{ij}, edu\ new_{kj})}{var(edu\ new_{ij})} = \frac{\sigma_u^2}{\sigma_u^2 + \sigma_e^2} \quad (3)$$

식 (2)의 two-level 모형을 추정한 가구레벨 오차항 분산 추정치 $\hat{\sigma}_u^2$ 와 개인레벨 오차항 분산 추정치 $\hat{\sigma}_e^2$ 를 이용하여 처리군에 속한 2명의 쌍둥이 간 교육성과 상관계수 추정치를 얻을 수 있다. 식 (2) 모형을 비교군 표본만 이용하여 추정하면 (쌍둥이가 아닌) 자녀 간에도 상관계수 추정치 역시 구할 수 있다.

두 번째 모형은 쌍둥이 간 교육성과 상관계수가 개인 성별(sex)과 출생연도 코호트(birth-year cohort)에 의해서 영향을 받을 수 있다고 가정한 조건부 모형(conditional model)이다. 가령 쌍둥이 간 교육성과 상관계수가 크다면 그 이유 중 하나는 최근 출생 코호트가 비교군에 비해 더 많기 때문이라고 주장할 수 있다. 이러한 주장을 통제하기 위해 다음과 같이 조건부 모형을 설정한다.⁷⁾

$$\begin{aligned} edu\ new_{ij} &= \alpha + \beta_1 male_{ij} + \beta_2 bcohort_{ij} + u_j + e_{ij} \\ \text{where } cov(u_j, e_{ij} \mid male_{ij}, bcohort_{ij}) &= 0 \end{aligned} \quad (4)$$

7) 유혜림(2022) 연구모형에서는 통제변수로 부모특성(소득과 교육수준)을 포함하고 있다. 본 연구에서는 공유된 환경의 상당 부분이 부모특성일 가능성이 크다. 따라서 개별적으로 해당 변수의 효과를 식별하는 대신 공유된 환경에 포함하는 연구모형을 설정한다.

조건부 교육성과의 상관계수는 식 (4)를 추정한 후 다음과 같이 가구레벨 오차항 u_j 분산과 개인레벨 오차항 e_{ij} 분산을 이용하여 유사하게 얻을 수 있다.

$$\text{corr}(\text{educnew}_{ij}, \text{educnew}_{kj} \mid \text{male}_{ij}, \text{bcohort}_{ij}) = \frac{\sigma_u^2}{\sigma_u^2 + \sigma_e^2} \quad (5)$$

교육성과와 달리 노동시장 성과인 임금은 wave별로 변할 수 있다는 점을 고려하여 다음과 같이 three-level 모형을 설정한다. 식 (6)과 식 (7)에서는 무조건부 모형과 조건부 모형을 설정한다. 식 (7)의 조건부 모형에서는 개인 성별, 나이, 출생코호트 변수를 포함한다. 취업한 이후에는 나이가 많아질수록 임금이 높아지는 임금체계를 반영하고자 한다. 식 (6)과 식 (7)에서 μ_{tj} 는 같은 가구 j 에 속한 쌍둥이의 임금이 영향을 미칠 수 있는 시점 효과(가령 경기 호황 또는 불황)를 의미한다.

$$\text{wage}_{itj} = \alpha + u_j + \mu_{tj} + e_{itj} \quad (6)$$

$$\begin{aligned} \text{wage}_{itj} &= \alpha + \beta_1 \text{male}_{ij} + \beta_2 \text{age}_{itj} + \beta_3 \text{bcohort}_{ij} + u_j + \mu_{tj} + e_{itj} \\ \text{where } \text{cov}(u_j, \mu_{tj}) &= \text{cov}(u_j, e_{itj}) = \text{cov}(\mu_{tj}, e_{itj}) = 0 \end{aligned} \quad (7)$$

Three-level 모형에서는 민인식·최필선(2022: 13장)에서 제시된 바와 같이 쌍둥이(i, k) 간 임금 상관계수를 다음과 같이 계산한다.

$$\text{corr}(\text{wage}_{itj}, \text{wage}_{ktj}) = \frac{\sigma_u^2 + \sigma_\mu^2}{\sigma_u^2 + \sigma_\mu^2 + \sigma_e^2} \quad (8)$$

자녀 간 교육성과 상관계수에 해당하는 식 (5)와 노동시장 성과에 해당하는 식 (8)의 결과를 얻기 위해 비교군과 처리군에 속한 각 표본데이터를 이용한다.

3. 실증분석 결과

〈표 6〉에서는 자녀 간 교육성과 상관계수를 얻기 위해 처리군과 비교군 표본을 이용하여 two-level(가구-자녀) 회귀모형 추정결과를 정리한다. 각 개인(자녀) 쌍(pair) 중에서 두 자녀가 모두 포함된 케이스만 남긴 후 회귀모형을 추정한다. 따

〈표 6〉 교육성과 회귀모형 추정결과

	처리군 :		비교군 :		
	무조건부 모형	조건부 모형	조건부 모형		
	쌍둥이 그룹	쌍둥이 그룹	age_diff =1세	age_diff =3세	age_diff =5세
<i>male</i> (남자=1, 여자=0)		-0.465* (0.240)	-0.156** (0.077)	-0.189*** (0.064)	0.008 (0.108)
<i>bcohort80</i> (80년대생)		-0.014 (0.340)	0.293** (0.125)	0.525*** (0.099)	0.108 (0.157)
<i>bcohort90</i> (90년대생)		0.016 (0.305)	0.108 (0.122)	0.083 (0.095)	0.022 (0.161)
σ_u	0.888	0.850	0.875	0.794	0.800
σ_e	0.829	0.830	1.073	1.150	1.187
$\log L$	-195	-193	-1,847	-2,806	-1,086
가구 obs	64	64	543	813	310
개인 obs	128	128	1,086	1,626	620

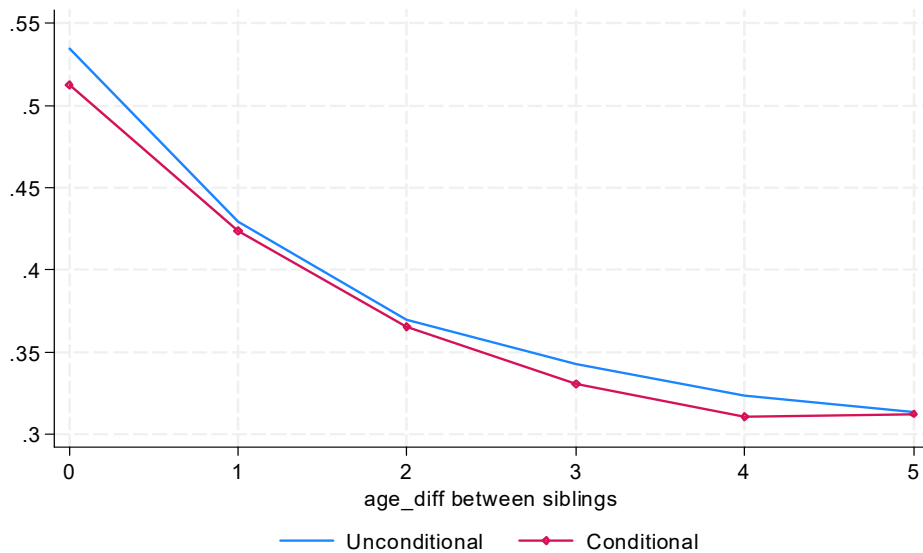
- 주 : 1) ***, **, *는 각각 1%, 5%, 10% 유의수준에서 유의함을 의미.
 2) 괄호 안은 추정계수의 표준오차를 의미.
 3) 출생 코호트 변수의 기준범주는 bcohort70(70년대생).
 4) age_diff=2세와 age_diff=4세 비교그룹에 대해서 추정결과를 얻었지만, 지면 제약상 생략.
 5) 비교군에 대해서도 무조건부 모형 추정결과를 얻었지만 지면제약상 생략.

라서 각 가구에서는 2명의 자녀가 선택되어 분석에 포함된다. 처리군에 속한 쌍둥이 표본에 대해서 무조건부와 조건부 모형 추정결과를 제시한다. 두 모형에서 상위레벨(가구) 오차항 표준편차와 하위레벨(자녀 개인) 오차항 표준편차는 큰 차이는 없는 것으로 얻어진다. 식 (3) 공식을 이용하여 ICC(Intra-Class Correlation)를 얻을 수 있다. ICC=0.51(조건부 모형)~0.53(무조건부 모형)이고 쌍둥이 자녀 간 교육성과 상관관계가 매우 크다는 것을 확인할 수 있다. 반면 비교군에 속한 비쌍둥이 자녀 간 ICC 역시 양의 값으로 추정된다. age_diff(나이 차이)가 커질수록 두 자녀 간 ICC는 점차 감소하는 것을 확인할 수 있다. age_diff=1세(1살 차이 자녀)의 ICC=0.41 정도이고 age_diff=3세(3살 차이 자녀)의 ICC=0.33, 그리고 age_diff=5세(5살 차이 자녀)의 ICC=0.31 정도이다. 같은 부모 밑에서 성장하였더라도 두 자녀의 나이 차이에 따라 공유된 환경의 정도가 다를 수 있고 동갑인 쌍둥이가 서로 공유한 환경의 정도가 가장 크고 이러한 현상은 교육성과의 유사성

으로 확인할 수 있다.

[그림 3]에서는 무조건부 모형과 조건부 모형에서 추정된 자녀 간 교육성과의 상관관계수(ICC)를 시각적으로 보여준다. 자녀 간 나이 차이가 증가할수록 두 자녀의 교육성과 유사성은 점차 낮아지는 것을 확인할 수 있다. age_diff=0세(쌍둥이)인 경우에 가장 유사성이 크고 age_diff=4~5세인 경우는 가장 낮은 수준을 보여준다.

[그림 3] 교육성과의 ICC : 처리군과 비교군



주 : age_diff의 증가에 따른 추세를 파악하기 위해 lowess mean smoothing 결과를 제시.

노동시장 성과인 임금은 2022년 기준 실질임금으로 환산한 후 로그값을 취해 사용한다. <표 7>에서는 식 (6)(무조건부 모형)과 식 (7)(조건부 모형)에 해당하는 three-level 회귀모형을 추정한 결과를 정리한다. 각 레벨에 해당하는 오차항의 표준편차 추정치를 이용하여 자녀 간 상관계수를 계산할 수 있다. 먼저 처리군인 쌍둥이 표본에서는 0.808(조건부)~0.809(무조건부) 정도의 높은 상관관계를 얻는다. 같은 부모 밑에서 자라고 같은 시점(wave)에 취업해 있는 쌍둥이 2명의 임금 상관관계가 매우 높다는 것을 확인할 수 있다. 비교군으로 나이 차이가 있는 자녀 간 상관계수는 age_diff=1세인 경우는 0.45(조건부 모형)로 쌍둥이 그룹에 비해

크게 감소한다. age_diff=3세인 경우는 0.41 그리고 age_diff=5세인 표본에서는 0.38까지 지속적으로 감소한다. 그러나 나이 차이=3~5세 그룹에서는 서로 유의미한 차이가 있다고 판단하기는 어렵다.

〈표 7〉 임금 회귀모형 추정결과

	처리군 :	처리군 :	비교군 :		
	무조건부 모형	조건부 모형	조건부 모형		
	쌍둥이 그룹	쌍둥이 그룹	age_diff =1세	age_diff =3세	age_diff =5세
<i>male</i> (남자=1, 여자=0)		0.295*** (0.106)	0.151*** (0.018)	0.200*** (0.015)	0.347*** (0.023)
<i>age</i>		0.038*** (0.002)	0.036*** (0.001)	0.042*** (0.001)	0.032*** (0.001)
<i>bcohort80</i> (80년대생)		0.105 (0.198)	0.113** (0.045)	0.137*** (0.035)	0.138** (0.055)
<i>bcohort90</i> (90년대생)		0.354*** (0.209)	0.143*** (0.052)	0.195*** (0.044)	0.258*** (0.078)
σ_u	0.476	0.493	0.302	0.288	0.269
σ_μ	0.189	0.000	0.000	0.000	0.000
σ_e	0.249	0.239	0.338	0.347	0.341
logL	-114	-55	-1142	-1,693	-626
Obs	324	324	2,608	3,712	1,436

- 주 : 1) ***, **, *는 각각 1%, 5%, 10% 유의수준에서 유의함을 의미.
 2) 괄호 안은 추정계수의 표준오차를 의미.
 3) 기준범주는 bcohort70(70년대생).
 4) age_diff=2세와 age_diff=4세 비교그룹에 대해서 추정결과를 얻었지만, 지면 제약상 생략.
 5) 비교군에 대해서도 무조건부 모형 추정결과를 얻었지만 지면제약상 생략.

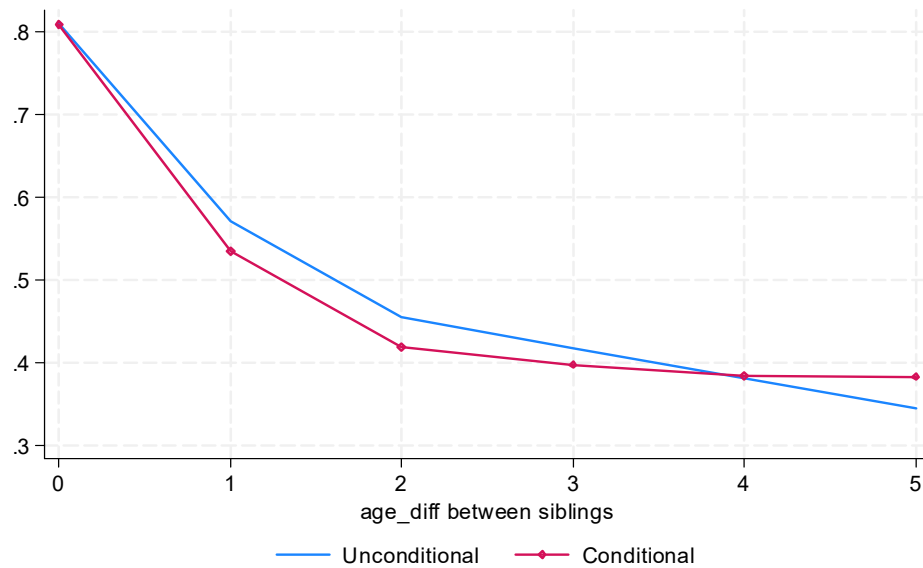
[그림 4]에서는 자녀 간 임금 상관계수의 평활화된 값을 그래프로 제시하고 있다. 쌍둥이 그룹에서는 상관계수가 매우 크지만, 자녀 간 나이 차이가 증가할수록 그러한 양의 상관관계는 계속 감소하고 age_diff=3세 이상인 그룹에서는 상관관계 정도가 큰 변화가 없이 0.4 정도를 유지하고 있다. 다만 3~5세 그룹에서는 자녀 상관계수에서 유의미한 차이가 없는 것으로 판단된다. 이러한 결과를 종합하면

쌍둥이 그룹에서 특이하게 자녀 상관계수가 큰 값으로 추정되고 자녀 나이 차이가 증가할수록 자녀 상관계수가 감소하기는 하지만 3세 이상이면 그 차이가 거의 없다고 해석할 수 있다.

노동시장 성과의 상관계수는 교육성과보다 더 큰 수치로 추정된다. 그 이유 중 하나는 노동시장 성과가 자녀 간 교육성과의 유사성을 포함하고 있기 때문이다.⁸⁾ 쌍둥이의 교육성과가 유사하다는 것을 이미 [그림 3]에서 확인하였고 이러한 교육 수준의 유사성이 임금 유사성을 확대시키는 것으로 판단할 수 있다.

교육수준뿐 아니라 자녀 간 임금 상관관계는 공유된 환경에 의해 영향을 받는다는 것을 [그림 4] 결과에서 확인할 수 있다. 공유된 환경의 정도가 가장 크다고 예상되는 쌍둥이 그룹에서 양의 상관관계가 가장 크게 나타나고 나이 차이가 증가할수록(즉, 공유된 환경 정도가 감소할수록) 자녀 간 상관관계가 줄어든다.

[그림 4] 노동시장 성과의 ICC : 처리군과 비교군



주: age_diff의 증가에 따른 추세를 파악하기 위해 lowess mean smoothing 결과를 제시.

8) <표 3>에서 쌍둥이의 교육 상관계수는 0.7 정도로 높지만 [그림 3]에서 교육 ICC=0.53 정도로 차이가 있는 이유는 <표 3>에서 교육수준은 1.무학~6.대졸 이상 범주형 변수를 사용하여 계산한 결과이다. 반면 [그림 3]에서는 <표 5>에서 제시한 새로운 교육수준 범주형 변수를 사용하고 있다.

V. 결론 및 시사점

본 연구는 개인의 사회경제적 성과에 유전적/환경적 요인이 미치는 정도를 분석하는 것을 목적으로 시작하였다. 유전적 영향력을 통제한 상태에서 성과변수에 관해 같은 가구 내에 속한 자녀 상관계수를 추정한다. 특히 환경 요인을 공유된/비공유된 환경으로 구분한다는 점에서 선행연구와 차별화를 갖는다. 지난 25년 동안 축적된 한국노동패널에서 쌍둥이 데이터를 구축하여 쌍둥이와 쌍둥이가 아닌 형제자매의 자녀 상관계수를 추정하여 공유된 환경이 사회경제적 성과에 미치는 영향을 추정하였다.

연구결과를 요약하면 다음과 같다. 첫째, 민인식 외(2023)에서 제시한 가족관계행렬(FRM)을 활용하면 쌍둥이를 쉽게 식별할 수 있고 1970년생 이후 쌍둥이 167쌍(334명)을 쌍둥이 데이터로 구축하여 쌍둥이의 사회경제적 성과 유사성을 쉽게 측정할 수 있다. 둘째, 쌍둥이 그룹을 처리군, 나이 차이 1~5세인 자녀를 비교군으로 설정하고 2단계(가구-자녀) 회귀모형을 설정한다. 개인의 고유한 특성을 통제하였을 때 교육수준의 급내상관계수(ICC)는 0.51로 가장 높다. 나이 차이가 많은 형제자매일수록 ICC 값이 줄어든다. 나이 차이=5세이면 ICC=0.31이다. 셋째, 임금(자영업자는 소득)으로 측정된 노동시장 성과 역시 동갑이 쌍둥이 자녀그룹에서 ICC=0.81로 가장 크다는 것을 확인한다. 나이 차이=1세인 자녀그룹에서는 0.45 그리고 나이 차이=5세인 자녀그룹에서 임금 유사성은 0.38까지 감소한다. 이러한 실증분석 결과는 나이 차이가 적은 자녀일수록 공유된 환경 요인이 증가하고 이러한 공유환경 요인의 크기가 자녀 간 사회경제적 성과 유사성에 기여하고 있다.⁹⁾ 다만 나이 차이가 일정 수준 이상이면 자녀 간 성과 유사성에는 유의미한 변화가 없다고 판단할 수 있다.

본 연구의 함의는 다음과 같이 두 가지로 정리할 수 있다. 첫째, KLIPS 데이터

9) 본 실증분석에서는 유전적 요인을 통제한 상태에서 환경요인의 효과만 분석하고 공유/비공유된 환경의 차이를 분석하고 있다. 자녀 상관성(유사성)은 유전적 요인이 가장 크다고 알려져 있다. 다만 환경적 요인 중에서는 공유된 환경이 비공유된 환경보다 크다는 것을 설명하는 것이 본 연구의 주요한 목적이다.

의 새로운 활용 가능성을 제시하고 있다. 기존 쌍둥이 연구는 대상 그룹을 따로 표집하여 설문조사 방식으로 진행한다. 이를 위해 비용과 시간을 투자해야 하는 어려움이 있다. 25년 동안 같은 가구-개인을 조사하고 있는 한국노동패널에서는 자연스럽게 쌍둥이 출산을 관찰할 수 있고 이들의 사회경제적 지위 변화를 추적할 수 있는 장점이 있다. KLIPS의 파생데이터로 쌍둥이 데이터를 구축하고 이를 활용하는 새로운 연구과제를 제안할 수 있는 기여가 있다.

둘째, 유전적 유사성이 동일한 형제자매를 대상으로 선택한 분석이라는 한계를 가지고 있지만, 공유된 환경의 중요성을 시사하고 있다는 점에서 의미가 있다. 형제자매간 나이 차이가 적을수록 초기 어린 시절의 공유된 환경이 더 많을 것으로 예상된다. 이러한 공유된 환경은 청소년기-성인기를 거쳐 사회경제적 성과의 유사성에 영향을 미친다는 것을 확인하였다. 부모 배경에 따른 개인 간 사회경제적 지위의 불평등이 심화되는 최근 사회적 현상 아래에서 이를 완화하기 위해 공유된 환경을 정책적으로 제공할 필요가 있다. 소외계층에 대한 유아 학비 및 보육료 지원 및 시설 제공을 통해 출발점부터 격차를 줄이는 정책이 지속적으로 필요하다. 청소년기에는 공교육의 질적 향상을 통해 환경 격차를 줄일 수 있어야 한다. 이러한 정책적 노력은 교육적 성취를 통한 취업·임금·사회적 지위 격차를 줄이는 데 기여할 수 있을 것으로 기대한다.

참고문헌

- 민인식·홍경준·전예지(2023). 「KLIPS 가족관계행렬 구축과 활용」. 『노동정책연구』 23(2): 85~107.
- 민인식·최필선(2022). 『Stata 고급통계분석 16-17버전』. 지필미디어. 서울.
- 이지영·고영선(2023). 「대학서열과 생애임금격차」. 『경제학연구』 71(2): 155~204.
- 유혜림(2022). 「자녀 간 사회경제적 성과의 유사성 및 차이성에 관한 연구: 자녀 상관관계를 중심으로」. 『노동정책연구』 22(3): 101~125.
- 허윤미(2005). 「아동의 자아개념에 미치는 유전과 환경의 영향: 쌍둥이 연구」. 『한국심리학회지: 사회 및 성격』 19(2): 91~102.

- Bouchard, T. J. and M. McGue(2003). "Genetic and Environmental Influences on Human Psychological Differences". *Journal of Neurobiology* 54 (1) : 4~45.
- Bowels, S. and H. Gintis(1976). *Schooling in Capitalist America: Educational Reform and the Contradiction of Economic Life*. Basic Books.
- Caspi, A., A. R. Hariri, A. Holmes, R. Uher, M. R. C. Psych, and T. E. Moffitt(2010). "Genetic Sensitivity to the Environment : The Case of the Serotonin Transporter Gene and Its Implication for Studying Complex Diseases and Traits". *American Journal of Psychiatry* 167 (5) : 509~527.
- Cawley, J., E. Han, J. Kim, and E. C. Norton(2023). "Sibling Correlation in Educational Attainment : A Test of Genetic Nurture". *NBER Working Paper* No. w27336.
- Coleman, J. S. and Others(1966). *Equality of Educational Opportunity*. U.S. Department of Health, Education, Welfare.
- Conley, D. and R. Glauber(2007). "Family Background, Race, and Labor Market Inequality". *The Annals of the American Academy of Political and Social Science* 609 (1) : 134~152.
- _____(2008). "All in the Family?: Family Composition, Resources, and Sibling Similarity in Socioeconomic Status". *Research in Social Stratification and Mobility* 26 (4) : 267~306.
- Dong, L., E. J. Giangrande, S. R. Womack, K. Yoo, C. R. Beam, K. C. Jacobson, and E. Turkheimer(2023). "A Longitudinal Analysis of Gene X Environment Interaction on Verbal Intelligence across Adolescence and Early Adulthood". *Behavior Genetics* 53 (4) : 311~330.
- Hauser, R. M. and P. A. Mossel(1985). "Fraternal Resemblance in Educational Attainment and Occupation". *American Journal of Sociology* 91 : 650~673.
- Haveman, R. and B. Wolfe(1995). "The Determinants of Children's Attainments :

- A Review of Methods and Findings". *Journal of Economic Literature* 33 (4) : 1829~1878.
- Koepfen-Schomerus, G, F. M. Spinath, and R. Plomin(2003). "Twins and Non-twin Siblings : Different Estimates of Shared Environmental Influence in Early Childhood". *Twin Research and Human Genetics* 6 (2) : 97~105.
- Loehlin, J. C. and R. C. Nichols(1976). *Heredity, Environment and Personality : A Study of 850 Sets of Twins*. Austin, University of Texas Press.
- Matteson, L. K., M. McGue, and W. G. Lacono(2013). "Shared Environmental Influences on Personality : A Combined Twin Adoption Approach". *Behavioral Genetics* 43 (6) : 491~504.
- Mazumder, B. and D. I. Levine(2003). "The Growing Importance of Family and Community : An Analysis of Changes in the Sibling Correlation in Earnings". *The Institute for the Study of Labor (IZA) Discussion Paper* No. 6987.
- McGuire, S., B. Manke, K. J. Saudino, D. Reiss, E. M. Hetherington, and R. Plomin(1999). "Perceived Competence and Self-Worth during Adolescence : A Longitudinal Behavioral Genetic Study". *Child Development* 70 : 1283~1296.
- Plomin, R. and D. Daniels(1987). "Why Are Children in the Same Family So Different from One Another?". *Behavioral and Brain Sciences* 10 (1) : 1~16.
- Plomin, R. and J. C. DeFries(1980). "Genetics and Intelligence : Recent Data". *Intelligence* 4 (1) : 15~24.
- Rowe, D. C., K. C. Jacobson, and E. J. Van den Oord(2003). "Genetic and Environmental Influences on Vocabulary IQ : Parental Education Level as Moderator". *Child Development* 70 (5) : 1151~1162.
- Scarr, S. and R. A. Weinberg(1978). "The Influence of Family Background on Intellectual Attainment". *American Sociological Review* 43 (5) : 674~692.

Tucker-Drob, E. M. and K. P. Harden(2011). "Early Childhood Cognitive Development and Parental Cognitive Stimulation : Evidence for Reciprocal Gene-Environment Transactions". *Development Science* 15 (2) : 250~259.

Abstract

**Environment and Socio-Economic Outcomes : Using Twins
Data from KLIPS****Min, Insik**

This study analyzes how genetic and environmental factors impact an individual's socioeconomic outcomes. By controlling for genetic variability, we estimate the sibling correlation for outcome variables among children within the same household. Distinct from prior research, this study differentiates between shared and non-shared environmental factors. We construct twin data from the Korea Labor and Income Panel (KLIPS), built over the past 25 years, and compare the sibling correlation between twins and non-twin siblings. The findings are as follows. First, treating the twin group as the treatment group and siblings with age differences of 1~5 years as the control group, the sibling correlation for educational outcomes in the twin group is the highest at 0.51. In contrast, when the age difference is 5 years, the ICC is the lowest at 0.31. Second, labor market outcomes, measured by wages, also have the highest ICC in the twin group at 0.81. It steadily decreases to 0.45 when the age difference is 1 year and 0.38 when it is 5 years. Through this study, we propose a new research topic using twin data constructed from KLIPS data, suggesting the importance of shared environments on individual socioeconomic outcomes. Targeted policies emphasizing shared experience during early childhood can help bridge the gaps in employment, wages, and social status among individuals.

Keywords : genetic factors, shared environment, twins, KLIPS, socioeconomic outcomes