

노동정책연구
2013. 제13권 제1호 pp.1~32
한국노동연구원

연구논문

자녀의 교육성과에 가구배경이 미치는 영향 : 형제 · 자매 · 남매의 학교정보를 이용한 분석*

양정승**

본 연구는 구체적인 가구배경 변수들을 사용하여 자녀의 교육성과에 가구배경과 이웃 등이 미치는 영향을 추정할 때 발생하는 생략변수에 의한 편향의 문제를 최소화하기 위해, 형제 · 자매 · 남매의 대학 정보를 이용하여 기대수능점수를 계산한 후 이를 가구배경과 이웃효과와의 대리변수로 이용하여 그 효과의 크기를 추정하고 하계값을 제시하였다. 추정 결과는 가구배경과 이웃효과가 자녀 교육성과의 20.2% 이상을 설명하는 것으로 나타났으며 평균화지역만으로 한정할 경우 19.2% 이상을 설명하는 것으로 나타났다. 고졸 형제 · 자매 · 남매의 기대수능점수를 계산하여 표본선택 문제를 일부 수정할 경우 23% 이상을, 평균화지역의 경우 20.6% 이상을 설명하는 것으로 나타났다.

핵심용어 : 가구배경, 수능성적, 이웃효과, 학교효과, 교육 기회의 균등

논문접수일 : 2012년 11월 29일, 심사의뢰일 : 2012년 12월 6일, 심사완료일 : 2013년 1월 31일

* 이 글은 양정승(2012a)의 박사학위 논문 제2장을 수정 및 보완한 것이다. 세심한 논문지도를 해주신 이철희 교수님과 많은 조언을 해주신 김대일 교수님, 그리고 심사과정에서 유익한 논평을 해주신 이철인 교수님, 정진화 교수님, 황수경 박사님께 이 자리를 빌어 다시 한 번 감사의 인사를 전하고 싶다.

** 한국직업능력개발원 전문연구원(jseyang@krivet.re.kr)

I. 머리말 : 문제 제기

교육은 개인의 인적자본(human capital)을 증가시켜 소득증가에 기여하게 되는데, Mincer(1974)의 임금방정식을 이용하여 추정된 교육의 임금 기여도 크기가 실질적임이 그동안 많은 실증연구들을 통하여 입증되어 왔다. 또한 교육은 인적자본 축적을 통한 소득의 증가뿐만 아니라 결혼에 있어서 자신과 잘 부합하는 더 나은 배우자를 찾는 데 도움을 주고 건강 유지와 자녀 양육에 있어서 보다 합리적인 결정을 내릴 수 있도록 해주는 역할을 한다. 더욱이 교육 자체를 하나의 소비재로 간주할 수 있어 교육의 소비는 개인의 생애효용을 증가시키고, 또한 교육을 통하여 개인은 인내심을 좀 더 기르고 위험을 기피하는 성향을 가지게 되는 등의 비금전적인(nonpecuniary) 효과를 가진다.¹⁾

이러한 효과들로 인하여 상대적으로 높은 교육 수준을 가지고 있다는 것은 높은 소득을 얻고 다양한 측면에서 보다 가치가 높은 삶을 향유하는 데 있어 다른 어떤 특성들보다도 중요한 요소 중의 하나라고 할 수 있다. 따라서 자녀에 애정을 가지고 있는 이타적인 부모들(altruistic parents)은 그들의 사회경제적 지위를 자녀 세대에 이전시키거나 자녀의 생애효용수준을 증가시킬 주요한 수단으로 자녀의 교육에 투자를 하고자 하는 유인(incentives)을 가지게 된다. 한편 자본시장이 불완전하다면 가구의 경제적 수준이 자녀에 대한 교육투자의 수준을 결정하는 데 중요한 제약조건으로 작용하게 되고, 또한 가구소득 수준과 부모의 학력 수준 등에 따라 시간에 대한 선호율과 교육투자에 대한 선호도, 교육 관련 정보에 대한 접근성의 차이 등이 발생하게 되면, 부모의 투자 행위가 얼마나 효율적인가의 여부는 그 사회의 세대 간 사회경제적 지위의 이동성을 좌우하는 중요한 요소가 된다. 일반적으로 자녀가 높은 교육적 성과를 성취하는 데 가구의 배경이 중요할수록 그만큼 그 사회는 상대적으로 세대 간 이동성이 낮은 경직적인 사회라 할 수 있다.

1) 교육의 비금전적인 효과에 관한 최근 연구들의 리뷰(review)는 Oreopoulous and Salvanes (2011)를 참조할 것.

가구배경은 일반적으로 가구소득이나 부모의 학력 등을 지칭하는데, 넓게 보아 부모의 양육(nurturing) 태도와 유전적으로 결정되는 특성들(genetic inheritance)도 가구배경에 포함된다고 할 수 있다. 부모로부터 물려받은 선천적인 능력은 능력 차이의 원인이 개인이 통제할 수 없는 가구에서 비롯된 것이므로 광의의 가구배경에 포함할 수 있을 것이다. 가령 IQ가 높은 아버지가 IQ가 높은 자녀를 갖는 경향이 있다면 아버지의 높은 IQ에서 비롯되는 부분은 가구적인 요소라 할 수 있다. 부모의 양육태도가 자녀의 교육성과에 영향을 준다면 부모의 양육태도를 자녀가 선택할 수는 없기 때문에 이 또한 가구로부터 비롯되는 비개인적인 요소라는 측면에서 가구배경으로 간주할 수 있다.

거주지역의 환경 또한 자녀의 교육성과에 영향을 줄 수 있다. 빈곤 지역에 거주하는 것이 자녀의 교육성과에 미치는 영향에 관한 연구는 이미 미국 사회의 오랜 연구 주제였으며(Mayer and Jencks, 1989; Jencks and Mayer, 1990; Oreopoulos, 2003), 우리나라의 경우에도 최근 평준화 교육의 재검토 문제와 맞물려 평준화가 오히려 지역적 교육환경의 격리를 조장한다는 주장이 대두됨에 따라 중요한 문제로 인식되고 있다(오호영, 2006). 그러나 실제로 거주지역에 의해 발생하는 이웃효과(neighborhood effect)를 추정하는 것은 쉽지 않은데, 가구배경과 이웃효과 간에는 상관관계가 존재하기 때문이다. 우리나라의 소위 ‘강남학군’ 문제도 마찬가지이지만 거주지역의 결정에는 가구소득 수준이 영향을 주고 다시 거주지역의 이웃효과 혹은 네트워크 효과가 가구소득 증가에 추가적인 향상을 가져오는 메커니즘이 존재하고 있기 때문이다. 그럼에도 불구하고 일정한 가정하에서 거주지역에 따른 이웃효과를 측정하는 방법이 다양한 형태로 시도되어 왔다(Aaronson, 1998; Solon et al., 2000; Katz et al., 2001; Dietz, 2002; Oreopoulos, 2003; Page and Solon, 2003a, 2003b; Haurin et al., 2003; Jacob, 2004). 이 외에도 동료효과(peer effect), 학교효과(school effect) 등이 개인의 교육성과에 영향을 주는 요소로서 연구되어 왔는데, 가구배경, 거주지역과 더불어 이들은 모두 개인적으로 통제할 수 있는 요소들이 아닌 비개인적인 요소라는 측면에서 모두 공통점을 공유하고 있으며, 이러한 측면에서 광의의 가구배경 변수 혹은 비개인적 특성 변수라고 간주할 수 있다.

이러한 비개인적 요소들과 교육성과, 세대 간의 사회경제적 지위의 이동성

간의 관계를 종합해 보면, 유전적 요소, 가족의 사회경제적 배경, 부모의 양육 태도, 거주지역 및 이웃의 특성이 자녀의 교육성과에 미치는 영향이 크면 클수록 세대 간 이동성이 낮고 사회경제적 지위의 대물림이 쉬운 사회라 할 수 있다. 왜냐하면 사회경제적 지위에서 순수한 개인적 요소가 차지하는 기여도가 그만큼 더 낮기 때문이다.

교육적 성취도의 증가는 인적자본 축적을 통하여 소득력의 상승에 크게 기여하고 비금전적 측면에서의 부가적인 효과를 가지므로 개인의 교육적 성과 중 어느 정도가 가구관련 요소와 이웃으로부터 기인하는지를 밝히는 것은 개인의 소득 중 개인적 노력이 아닌 비개인적 특성의 기여도를 추정하는 데 도움을 줌, 재분배 정책을 수립하는 정책적 판단의 참고가 될 수 있을 것이다. 즉 비개인적 특성의 기여도가 크면 클수록 이를 줄이기 위한 재분배 정책은 정당성을 가지고 있다고 할 수 있을 것이다.²⁾

전통적으로 우리나라의 경우 다른 국가들에 비해 자녀교육에 대한 높은 수요와 투자가 있어 왔다. 가령 『교육통계연보』에 따르면 2010년도 기준으로 고등학생의 79%가 대학에 입학하여 교육을 받고 있다. 이러한 높은 대학교육열의 일부분은 빠른 경제성장과 더불어 그동안 교육수준에 대한 보수(return)가 빠르게 상승하면서, 높은 소득을 얻을 수 있는 기회를 자식에게 제공하고자 하는 부모들의 열망에서 기인한다고 할 수 있을 것이다. 또한 대학이 위계적으로 서열화되어 있어, 소위 명문대학을 졸업하였는지의 여부가 생애소득에 큰 영향을 주는 것으로 인식되어 명문대학 입학여부가 중시되면서 이러한 부모들의 자녀교육에 대한 태도가 사교육비 지출 등의 경쟁적 과잉투자에 반영되고 있다고도 할 수 있다. 통계청이 발행한 『2010년 사교육비 보고서』에 따르면, 2010년 현재 일반계 고등학생 1인당 월평균사교육비는 26만 5천 원이고 일반계 고등학생의 사교육 참여율은 61.1%이다. 지역별로는 서울지역의 경우 일반계 고등학생의 1인당 월평균사교육비는 42만 원이고 사교육 참여율은 70.6%에 이르고 있는 반면 읍면지역의 경우 일반계 고등학생의 1인당 월평균사교육비는 11만

2) 본 연구에서 비개인적 특성은 개인의 능력 이외의 특성을 의미하는 것이 아니라 개인의 능력 중 유전적으로 결정되는 부분도 포함하는 넓은 의미로 사용하고 있음에 유의하여야 한다. 즉 타고 태어난 선천적 능력은 개인적 특성이 아니라 비개인적 특성으로 간주하였 다. 왜냐하면 개인이 자신이 타고 태어날 능력을 선택할 수는 없기 때문이다.

3천 원이고 사교육 참여율은 38.1%에 불과하였다. 이러한 사교육비와 사교육 참여율의 격차가 교육적 성과에 영향을 주는 정도는 결국 가족적 배경이 교육적 성과에 영향을 주는 크기가 크면 클수록 마찬가지로 커질 것이다. 극단적으로 부모의 금전적 투자가 자녀의 교육성과에 매우 중요하다면, 이러한 사교육 기회의 격차는 교육적 성과의 차이를 통해 장차 소득격차로 나타나 부의 대물림 현상으로 귀결될 것이다.

따라서 가구배경과 이웃효과가 자녀의 교육성과에 미치는 영향을 추정해 보는 것은 중요하다고 할 수 있는데, 우리나라의 경우 자녀의 교육성과에 대한 기존의 선행연구들은 주로 관심의 초점이 되는 구체적인 가구배경 변수들의 효과를 알아보는 데 치중되어 왔다. 부모의 학력이나 직업, 부모의 교육연수, 가구원수, 교육비 지출 등의 구체적인 가구배경 변수들을 통제하여 이들이 자녀의 교육에 미치는 영향을 살펴보는 것이 주요 연구 주제였다(고제이 · 이우진, 2011; 안종범 · 전승훈, 2008; 김미란, 2009; 방하남 · 김기현, 2002, 2003 등). 한편 오호영(2006)은 소득계층에 따른 지역적 격리 문제와 학교효과 문제를 다루었으며, 이주호 · 홍성창(2001)과 이주호 · 김선웅(2002)은 평준화정책과 학교 교육의 질, 사교육 지출 등의 연관성 문제를 다루어 가구배경 이외의 문제를 명시적으로 고려하였다. 한편 Kang(2007)과 강창희 · 채창균(2010)은 학급 내에서 발생하는 동료효과를 연구하였다. 본 연구와 연구 목적이 가장 유사하다고 할 수 있는 고제이 · 이우진(2011)은 아버지의 학력 이외의 다른 요소들에 의한 불평등은 모두 아들세대의 노력의 차이에 의한 것이라는 가정하에 아들세대의 학력 불평등 중 16~59% 정도가 아버지의 학력이라는 환경의 불평등에 의한 부분으로 계산하였다.

교육성과에 대한 가구배경의 영향을 추정하는 선행 연구들은 구체적인 가구배경변수들의 효과를 설명한다는 측면에서는 유용한 연구들이라고 할 수 있지만, 모든 가구배경 변수들이 관측 가능한 것이 아니기 때문에 변수 누락에 의한 편향의 발생이라는 고전적인 문제를 가지고 있다. 가령 가구소득을 통제하였으나 거주지역을 통제하지 않은 전형적인 경우를 생각해 보자. 우리가 흔히 생각하듯이 소위 ‘강남학군’이 가구소득 수준으로는 설명할 수 없는 부가적인 이득을 자녀의 교육성과에 가져온다고 한다면, 이러한 강남학군효과를 통제하지 않

았을 경우 이 효과의 일부는 가구소득의 효과로 추정되어 과대추정될 것이다. 가구소득 수준과 강남거주 여부 간에는 양(positive)의 상관관계가 존재하기 때문이다. 또한 이러한 생략변수의 존재는 자녀들의 교육성과 격차에서 가구배경의 차이가 기인하는 부분을 과소추정하는 결과를 가져오게 된다. 관측할 수 없는 가구배경 변수들은 대부분의 경우 항상 존재하므로 관측되지 않은 가구배경 변수들이 많을수록 그만큼 가구배경의 설명력은 감소하게 될 것이다.

본 연구는 자녀의 교육성과에 대한 가구배경의 중요성을 평가하는 데 있어서, 구체적인 가구배경 변수들을 사용할 때 피할 수 없는, 생략변수 문제에 의한 과소추정을 최소화하는 데 주된 목적을 두었다. 형제·자매·남매 데이터를 사용하여 분석하는 서양의 선행 연구들이 주로 사용하는 기본적인 가정에 기초하여 가구배경의 중요성을 설명하고자 시도하였다. 이를 위해 형제·자매·남매의 대학정보를 이용하여 그 효과를 추정하는 시도를 하였다. 연구의 대상 표본으로는 「한국교육고용패널」(Korean Education & Employment Panel)을 사용하였으며, 응답자의 형제·자매·남매 정보와 그들이 다니고 있는 대학정보를 이용하여 형제·자매·남매의 예상수능점수를 생성한 후, 이를 가구배경 및 이웃효과를 나타내는 설명변수로 사용하였다. 형제·자매·남매의 경우 유전적 능력, 가족의 사회경제적 배경(background)과 환경(family environment), 이웃환경의 상당부분을 공유하고 있기 때문에 구체적인 가족배경 관련 변수들을 설명변수로 사용하여 회귀분석하는 것보다는 가족배경에 대한 대리변수로 형제·자매·남매의 교육성과를 사용하는 것이 보다 많은 가족배경 및 이웃에 관한 정보를 제공한다고 할 수 있을 것이다.

본 연구의 추정 결과는 가족배경과 이웃효과가 자녀의 교육성과 분포(variation)의 20.2% 이상을 설명하는 것으로 나타났으며, 평준화지역만으로 한정할 경우 19.2% 이상을 설명하는 것으로 나타났고, 표본선택 문제를 수정할 경우 23%까지의 하계값을 얻을 수 있었다.

본 연구의 기여점은 다음과 같다.

첫째, 자녀의 교육성과에 대한 가구배경과 이웃효과의 대리변수로 형제·자매·남매의 수능점수를 사용한 첫 연구이다.

둘째, 자녀의 수능성적에 가구배경과 이웃효과가 미치는 영향의 크기를 측정

하는 데 있어서 구체적인 가구배경 변수들을 사용하여 분석한 선행연구들이 가지는 문제점을 줄이고 가능한 한 모든 가구배경과 이웃효과의 크기를 추정하고자 시도하였으며 그 값의 하계(lower bound)를 제시하였다.

셋째, 우리나라 수능성적의 경우 대학입학에서 차지하는 중요도가 높아 교육연수보다 자녀의 교육성공을 평가하는 보다 바람직한 측도(measure)라고 할 수 있는데, 교육연수가 아닌 직접적인 교육성취도 측도에 대한 형제 · 자매 · 남매 간의 상관관계를 추정한 최초의 연구라 할 수 있다.

마지막으로, 우리나라 최초로 형제 · 자매 · 남매 데이터를 사용한 연구이다. 서양의 경우 쌍둥이들을 대상으로 한 연구들을 비롯하여 형제 · 자매 · 남매 데이터를 사용한 연구들이 활발히 이루어지고 있는 반면에, 우리나라에서는 이를 활용한 연구가 전무한 실정이다. 형제 · 자매 · 남매 데이터는 선천적인 능력과 환경의 중요성을 연구하는 데 있어서 가장 기본이 되는 데이터라고 할 수 있는데, 이를 활용한 연구의 활성화는 진화론과 유전학이라는 현대생물학의 주요 성과가 각종 사회과학으로 흘러들어가고 있는 학문적 추세를 우리나라 학계에 서도 본격화하기 위해 중요한 첫걸음이라고 할 수 있다.

II. 추정모형

본 연구는 형제 · 자매 · 남매 데이터를 사용한 분석이나 이웃효과를 추정하는 연구들에서 가장 기본적으로 사용하는 다음과 같은 기초적인 모형을 사용한다. 교육성공(educational outcome)가 가구배경(family background), 이웃(neighborhood) 혹은 학교의 특성, 개인 고유의 능력(individual ability)의 세 가지 구성부분으로 이루어져 있다고 가정하면, 지역 h 에 거주하는 가구 f 의 특정 자녀 i 의 교육성공은 다음과 같은 식으로 나타낼 수 있다(Aaronson, 1998; Solon et al., 2000; Oreopoulos, 2003; Page and Solon, 2003a, 2003b).

$$S_{ifh} = F_f + H_h + A_i \quad (1)$$

위 식에서 S_{ifh} 는 지역 h 에 거주하는 가구 f 의 자녀 i 의 교육성공이고, F_f 는

자녀의 교육성과에 영향을 주는 모든 가구적인 요소들을 의미한다. 따라서 F_f 는 부모로부터 물려받은 유전적인 학업 능력 등의 선천적 요소와 부모의 양육(nurturing) 태도 등의 후천적인 요소들을 모두 포함하고 있다. H_h 는 자녀의 교육성과에 영향을 주는 이웃의 특성, 학교의 질, 동료효과(peer effect), 지역의 네트워크 효과(network effect)를 모두 포함하는 것으로 정의한다. 한편 A_i 는 자녀 개인에게 고유한 능력으로 유전적으로 물려받는 선천적 능력이나 혹은 자라나는 가족환경, 거주지역 환경, 학교 등과는 상관관계가 없는 순수한 개인적인 능력을 의미한다. 따라서 $Cov(F_f, H_h) \neq 0$ 이지만, $Cov(F_f, A_i) = 0$ 이고 $Cov(H_h, A_i) = 0$ 인 것으로 가정한다.

동일한 방식으로 지역 h 에 거주하는 가구 f 의 또 다른 자녀 j 의 교육성과는 다음 식과 같이 나타낼 수 있다.

$$S_{jfh} = F_f + H_h + A_j \quad (2)$$

자녀 i 와 j 는 동일한 F_f 와 H_h 를 공유하고, A_i 와 A_j 는 개인에게 고유한 능력을 의미하므로 $Cov(A_i, A_j) = 0$ 이다. 물론 동일한 가족의 자녀들이라 하더라도 일란성 쌍둥이(monozygotic twins)가 아니라면 모든 유전적 특성을 공유하지는 않는다. 가령 선별적 결혼성향이 없고 완전한 무작위 결합(random mating)에 의해 결혼이 이루어진다면, 이란성 쌍둥이(dizygotic twins) 혹은 자녀들 간의 유전적 상관관계는 0.5이다(Bowls and Gintis, 2002). 따라서 유전적 특성의 절반만을 자녀들 간에 공유한다. 그러나 자녀들 간 유전적 상관관계는 선별적 결혼성향이 강화될수록 높아지고 현실적으로 선별적 결혼성향은 존재하므로 분석의 편의상 유전적 특성은 모두 공유하는 것으로 가정하였다. 따라서 공유하지 않는 유전적 특성을 고려하면 가구적 요소가 미치는 영향은 본 연구의 추정치보다 클 것이다. 가구배경은 이러한 선천적인 유전적 능력뿐만 아니라 부모의 사회적 지위, 소득수준, 양육(nurturing)의 질(quality)과 같은 요소들을 모두 포함하는 광의의 개념으로 정의하였다. 따라서 형제·자매·남매 간에 서로 영향을 주는 내생성(endogeneity)을 발생시키는 효과는 모두 정의상 가구배경에 포함되는 것으로 간주한다. 가령 형이 공부를 열심히 하는 것이 동생 또한 공부를 열심히 하게 유도하는 효과를 가진다면, 형제 간에 서로 공부를 열심히

하게끔 이끄는 서로 공유하는 보다 근본적인 요소(factor)들이 있다고 가정하는 것이다. 위의 모형을 형제 간의 교육성과 구조방정식으로 바라본다면, 일종의 축약식(reduced form) 형태로 간주할 수 있다.

이웃효과(neighborhood)는 거주지역의 환경뿐만 아니라 거주지역의 네트워크효과(network effect), 학교의 시설, 교사의 질과 같은 학교효과(school effect), 동급생과의 경쟁 혹은 도움이 주는 동료효과(peer effect) 등을 포괄하는 것으로 정의하였다. 지역 연고에 따른 가정환경이나 양육, 주변환경 등의 후천적 요소도 한 가구 내에서 성별과 출생순서에 따라 다른 것이 일반적이라고 할 수 있지만, 분석의 편의를 위해 이러한 차이가 크지 않다는 가정하에서 동일한 요소를 공유하는 것으로 가정하였다. 또한 본 연구에서는 순수한 네트워크효과와 동료효과, 교사효과(teacher effect), 학교효과 등을 분리하여 추정하는 시도를 하지 않았는데 분해(decomposition)가 쉽지 않은 작업일 뿐만 아니라, 본 연구의 목적은 단지 교육성과에서 어느 정도 비중을 개인적 요소로부터 비롯되지 않는 요인들이 설명할 수 있는가를 파악하는 것이고 구체적인 분해에 그 목적이 있지는 않기 때문이다.

이러한 관계에 의해 자녀들의 교육성과가 담겨 있는 표본에 대하여 다음과 같은 회귀식이 분석될 수 있다.

$$S_{ifh} = \beta_0 + \beta_1 S_{jfh} + \epsilon_i \quad (3)$$

위 식에서 $Cov(S_{jfh}, \epsilon_i) = 0$ 이다.

추가적으로 $\sigma_{A_i}^2 = \sigma_{A_j}^2 = \sigma_A^2$ 를 가정하면, 식 (3)의 계수 β_1 은 회귀분석하였을 때 다음과 같은 관계를 가지게 된다.

$$\beta_1 = \frac{Cov(S_{jfh}, S_{ifh})}{Var(S_{jfh})} = \frac{Var(F_f + H_h)}{Var(F_f + H_h) + \sigma_A^2} \quad (4)$$

식 (4)에서 계수 β_1 은 결국 자녀들 교육성과의 전체 분산(variance) 중에서 가구배경과 이웃효과에 의해 설명되는 부분의 비중이 된다.

자녀들의 실제 교육성과를 모두 관찰할 수 있는 것이 아니라 한쪽만 실제 교육성과 자료를 관측 가능하고, 다른 한쪽은 오직 교육성과의 기대치만이 이용

가능하다면, 대안적으로 회귀식 (3)에서 설명변수를 바꾸어 다음과 같은 두 가지 회귀식이 분석 가능하다.

$$S_{ifh} = \beta_0 + \beta_1 \widehat{S}_{jfh} + \epsilon_i \quad (5)$$

$$S_{ifh} = \beta_0 + \beta_1 \widehat{S}_{jfh} + \beta_2 X_{fi} + \beta_3 N_h + \epsilon_i \quad (6)$$

식 (5)와 (6)에서 \widehat{S}_{jfh} 는 자녀 j 의 교육성과 기대치이고 X_{fi} 는 가구배경 변수들의 벡터, N_h 는 거주지역 더미변수들의 벡터이다. 이 식을 회귀분석하게 되면, β_2 는 \widehat{S}_{jfh} 에 의해 설명되지 않는 가구배경들의 효과 중 일부를 보여주고, β_3 는 \widehat{S}_{jfh} 에 의해 설명되지 않는 모든 이웃효과와 자녀들 간의 이웃효과의 차이를 보여주게 되어 식 (5)에서 실제 교육성고가 아닌 기대교육성고를 설명변수로 사용함에 의해 누락되는 효과들의 일부를 측정하게 된다.

Ⅲ. 자료설명

자녀들의 교육성과에 가구배경과 이웃이 주는 효과를 추정하기 위해 사용한 자료는 「한국교육고용패널」(Korean Education & Employment Panel; 이하 KEEP)이다. KEEP은 2004년부터 시작되어 당시에 중학교 3학년인 학생 2,000명, 일반계 고등학교 3학년 학생 2,000명, 실업계 고등학교 3학년 학생 2,000명을 대상으로 하여 매년 추적조사가 이루어지고 있는 패널조사이다. KEEP은 학생들에 관한 정보뿐만 아니라 가구소득, 부모의 교육수준, 가구 소비지출, 가구원들의 대학교명과 전공 등의 가구관련 정보들과 담임교사와 학교행정가로부터 수집한 학교정보들을 제공하고 있다. KEEP 응답자들 중에서 본 연구의 분석대상이 되는 표본은 일반계 고등학교 3학년 학생 2,000명이다. 중학교 3학년 코호트(cohort)를 분석대상에서 제외한 이유는 그들이 2005년도에 입학한 고등학교의 수가 많아 한 학교당 단지 몇 명의 학생들만이 해당되어 본 연구의 분석에 적합하지 않기 때문이다. 실업계 고등학교 3학년생의 경우 대학입학을 목적

으로 하고 있지 않아 본 연구의 주제와는 관련이 적고, 실제 고등학생 비율에 비해 실업계 고등학교 학생들이 과다하게 표집되는 문제를 안고 있으며, 지역에 의해 학생을 배정받은 것이 아니라 학생들을 선발하기 때문에 지역적으로 동일하지 않아 이웃을 충분히 통제하기 어렵기 때문이다. 본 연구에서는 이웃 효과를 순수한 이웃효과와 동료효과, 교사효과, 그리고 학교효과를 모두 포괄하는 광의로 사용할 것이다. 이러한 개별적 효과를 분해하는 것은 본 연구의 목적을 벗어나기 때문이다.³⁾ 학교효과를 자녀들이 공유한다고 간주하는 이유는 평준화정책에 의해 거주지역에 따라 분류된 학군 내에서 학교가 배정되므로 자녀들은 비록 동일한 학교는 아니라 하더라도 비슷한 교육환경을 가지고 비슷한 교육서비스를 제공하는 학교에 다닌다고 간주할 수 있기 때문이다.

교육성공을 위한 변수는 대학수학능력시험(이하 수능) 언어영역, 수리영역, 외국어영역 표준화 점수의 합계이다. 이 세 가지 영역은 대부분의 학생들이 공통적으로 응시했기 때문에 비교가 상대적으로 쉬우나 과학/사회/직업탐구 영역은 응답자들 간에 시험을 친 과목과 과목수가 일치하지 않아 동등한 비교가 어렵기 때문에 제외하였다. 표준화 점수는 각 영역이 0점부터 200점까지의 점수로 되어 있고 연도마다 난이도를 고려하여 동등하게 비교할 수 있도록 표준화되어 있는 점수이므로 형제 · 자매 · 남매의 대학과 전공정보를 가지고 수능점수를 예측하기 위해 2005년 수능점수만이 아니라 그 이후의 수능점수도 이용할 수 있다는 장점이 있다. 가령 우리나라의 경우 명문대학 입학여부가 차후의 소득과 사회경제적 지위에 중요하다는 인식이 팽배해 있어 적지 않은 학생들이 한 번의 시험으로 만족할 만한 성적을 거두지 못하는 경우 다시 시험을 치고 있다. 따라서 이들의 추가적인 시험정보를 이용할 수 있다는 점에서 표준화 점수의 장점이 있다.

먼저 KEEP 응답자들이 응시한 수능점수와 직후 입학한 대학교명과 학과정보 데이터를 생성한 후 수능점수를 대학교명과 학과정보에 의한 전공계열 더미 변수들에 회귀분석함으로써 각 학교별 · 전공계열별 기대수능점수를 예측하였다.⁴⁾ 이렇게 계산한 학교별 · 전공계열별 기대수능점수를 응답자의 형제 · 자

3) 더욱이 그러한 효과들을 분해하기 어려운 경우가 많다. 이에 관해서는 Manski(1993)를 참조하라.

4) 대학별 학과별 수학능력시험을 예측하기 위해 사용한 데이터의 관측치 수는 1,805개이고

매·남매들이 다니고 있는 대학정보와 학과정보를 이용하여 그들의 기대수능 점수와 연결하였다. KEEP은 응답자들의 수학능력점수만을 제공하고 형제·자매·남매들의 점수는 제공하지 않기 때문에 그들의 교육성과를 이용하기 위해 기대수능점수를 이용했다. 기대수능점수를 계산하는 데 이용한 데이터는 2005년부터 2007년까지의 KEEP 일반계 고등학생 코호트의 수능시험점수와 직후 그들이 다닌 대학과 학과정보이다.⁵⁾

가구배경 변수들은 가구월평균소득, 월평균사교육비, 아버지와 어머니의 교육수준, 그리고 어머니의 보육효과를 통제하기 위한 어머니의 전문직 종사여부, 어머니의 비전문직 종사여부, 아버지와 어머니의 연령이다. 가구소득은 임시소득에 의해 하향편의(attenuation bias)가 발생하는 문제를 줄이기 위해 2004년부터 2006년까지의 3개년도 평균값을 사용하였고, 교육수준 변수로는 교육연수(schooling years)와 학력수준(education level) 더미변수 두 가지를 모두 사용하였다. 금액변수들은 2005년도 기준의 소비자물가지수로 실질화하여 사용하였다.

표본은 2005년도에 수학능력시험에 응시한 1,752명의 일반계 고등학교 학생들이고 그들 중 1,202명은 평준화지역에 거주하고 있었고 나머지 550명은 비평준화지역의 학생들이었다. 먼저 형제·자매·남매는 손아래인 경우 3차년도 조사까지도 대학진학의 수능점수는 되지만 원하는 학교에 진학하지 못해 지속적으로 대입을 준비하고 있는 상태인 경우가 있을 가능성이 높아 순위로 한정하였다. 순위들의 경우 3차년도에는 대부분 최종대학 진학이 종료된 상태로 간주할 수 있어 대학과 학과정보를 가지고 이들의 수능점수를 보다 잘 예측할 수 있다고 판단하였다. 또한 되도록 응답자와 서열이 가까운 형제·자매·남매를 선택하였고 18세 이상으로 제한하였다. 최종적으로 표본 중 651명이 대학을 다니고 있는 형제·자매·남매들의 정보를 확인할 수 있었는데, 이들 중 440명이 평준화지역, 그리고 211명은 비평준화지역의 학생들이었다. 이후에는 편의상

R^2 는 0.626이다.

- 5) KEEP은 응답자 이외의 가구의 입학연도를 제공하지 않기 때문에 입학연도는 고려하지 못하였다. 따라서 학교별 전공별 수능점수 분포가 연도에 따라 달라질 가능성을 통제하지 못하였다. 만약 이러한 분포를 고려할 수 있다면 추정치는 본 연구의 결과보다 상승하게 된다.

응답자 전체를 대상으로 한 표본을 ‘전체표본’으로, 형제 · 자매 · 남매의 대학 정보를 확인할 수 있는 표본은 ‘형제/자매/남매표본’으로 지칭하였다. 이러한 표본들에 의해 확인되는 학교의 수는 65개가 평준화지역으로 분류되었고 35개가 비평준화지역으로 분류되었다. 평준화지역에서는 같은 학교에 다니는 경우 같은 지역으로 볼 수 있기 때문에 본 연구에서는 학교를 하나의 지역으로 간주하였으며 분석의 초점을 평준화지역에 두었다.⁶⁾ 본 연구에서는 이미 언급한 바와 같이 학교효과와 이웃효과 등을 분해하는 시도를 하지 않았기 때문에 학교를 지역으로 간주하여도 평준화지역에서 문제는 발생하지 않는다.

〈표 1〉 평준화지역 형제 · 자매 · 남매표본의 기술 통계표(descriptive statistics)

	평준화지역			비평준화지역			전체		
	N	평균	표준 편차	N	평균	표준 편차	N	평균	표준 편차
응답자의 수능성적	440	292.34	59.31	211	267.45	59.88	651	284.27	60.58
형제/자매/남매의 기대수능성적	440	291.87	56.24	211	274.41	58.93	651	286.21	57.66
형제/자매/남매의 수	437	2.29	0.56	210	2.45	0.72	647	2.34	0.62
출생서열	437	2.08	0.53	210	2.15	0.59	647	2.10	0.55
로그가구소득	436	5.65	0.52	211	5.40	0.44	647	5.57	0.51
로그월평균사교육비 지출	346	3.63	0.91	138	3.08	0.94	484	3.47	0.95
아버지의 연령	419	48.97	3.08	205	48.71	3.48	624	48.89	3.22
아버지의 교육연수	416	13.03	2.74	205	11.57	3.02	621	12.55	2.91
어머니의 연령	429	45.83	2.84	207	45.36	3.15	636	45.68	2.95
어머니의 교육연수	427	11.81	2.46	207	10.62	2.24	634	11.42	2.45

6) 한 지역 내에서도 여러 학교를 배정받는 경우가 있고 일정 지역이 묶여 학군을 형성하는 문제는 본 연구에서는 고려하지 않았다. 또한 이후에는 특별한 언급이 없는 경우 학교효과와 이웃효과는 동일한 것으로 간주하였다.

IV. 결 과

1. 가구배경과 이웃이 자녀의 교육성과에 미치는 영향 추정

자녀의 교육성과 중에서 가구배경에서 기인하는 부분은 $F = \hat{F} + \eta_f$ 로, 이웃효과에서 기인하는 부분은 $H = \hat{H} + \xi_h$ 로 각각 나타낼 수 있다. \hat{F} 은 가구배경에서 기인하는 부분 중 관찰할 수 있는 변수들에 의해 설명되는 부분이고 η_f 는 관찰할 수 없는 부분을 의미한다. 마찬가지로 \hat{H} 는 이웃효과 중 관찰할 수 있는 부분이고 ξ_h 는 관찰할 수 없는 이웃효과를 의미한다. 이때 $Cov(\hat{F}, \eta_f) = 0$ 이고 $Cov(\hat{H}, \xi_h) = 0$ 으로 가정한다. 이러한 가정들하에서 형제·자매·남매의 실제 수능점수를 설명변수로 하고 응답자의 실제 수능점수를 종속변수로 하여 식(3)을 회귀분석하였을 때의 계수값을 의미하는 식(4)는 다음과 같이 표현할 수 있다.

$$\begin{aligned} \beta_1 &= \frac{Var(F+H)}{Var(F+H) + \sigma_A^2} = \frac{Var(\hat{F} + \eta_f + \hat{H} + \xi_h)}{Var(F+H) + \sigma_A^2} \\ &= \frac{Var(\hat{F} + \hat{H}) + Cov(\eta_f, \hat{H}) + Cov(\hat{F}, \xi_h) + Cov(\eta_f, \xi_h)}{Var(F+H) + \sigma_A^2} \end{aligned} \quad (7)$$

한편 형제·자매·남매의 기대수능점수를 설명변수로 하여 동일한 식(3)을 회귀분석하면 그 추정량(estimator)은 다음과 같이 된다.

$$\hat{\beta}_1 = \frac{Var(\hat{F} + \hat{H})}{Var(\hat{F} + \hat{H}) + \sigma_A^2} \quad (8)$$

이때 다음의 부등식(inequality)이 항상 성립하게 된다.

$$\beta_1 \geq \hat{\beta}_1 \frac{Var(\hat{S}_j)}{Var(S_i)} = \frac{Var(\hat{F} + \hat{H})}{Var(F+H) + \sigma_A^2} \quad (9)$$

부등식(9)의 우변은 응답자의 실제 수능점수를 형제·자매·남매의 기대수

〈표 2〉 자녀의 수능성적을 형제 · 자매 · 남매의 기대수능점수에 회귀분석한 결과

	평준화지역	비평준화지역	전체
형제/자매/남매의 기대수능성적	0.2134*** (0.0494)	0.1708*** (0.0693)	0.2234*** (0.0403)
상수항	230.0702*** (14.6699)	220.5635*** (19.4443)	220.3322*** (11.7662)
관측치 수	440	211	651
R^2	0.041	0.028	0.045
$V(S_i)$	3,518.212	3,585.258	3,670.467
$V(\hat{S}_j)$	3,162.485	3,472.466	3,324.632
$Corr(S_i, S_j)$	0.192	0.165	0.202

주 : *** p<0.01, ** p<0.05, * p<0.1.

능점수에 대하여 회귀분석한 결과로부터 항상 계산될 수 있는데, 부등식 (9)는 이렇게 계산된 값이 응답자의 실제 수능점수의 분포에서 가구배경과 이웃효과의 차이에 의해 설명되는 부분의 하계(lower bound)가 됨을 의미한다.

〈표 2〉는 응답자의 실제 수능점수를 그들의 형제 · 자매 · 남매의 기대수능점수에 대하여 평준화지역의 응답자들과 비평준화지역, 전체 응답자들을 대상으로 각각 회귀분석한 결과이다. 이 결과로부터 위에서 도출한 가구배경과 이웃효과의 하계값을 계산하면 평준화지역의 학생들만을 대상으로 할 경우 0.192이고 비평준화지역의 경우 0.165, 전체지역을 대상으로 할 경우 0.202이다. 즉 자녀의 수능성적 중 평준화지역의 경우 최소 19.2%, 비평준화지역의 경우 16.5%, 우리나라 전체적으로는 최소 20.2%가 가구배경과 거주하고 있는 지역 또는 다니고 있는 학교에서 기인하고 있음을 의미한다. 다시 말하면 수능성적의 5분의 1가량은 비개인적인 요소로부터 발생하고 있다고 할 수 있다.7)

7) 비평준화지역의 경우 일반적으로 평준화지역보다 경제적 여건이 부족한 지역에 해당하여 <표 1>에서 확인할 수 있는 것처럼 평준화지역과는 가구의 일반적인 특성들이 다를 뿐 아니라, 거주지역에 따라 학교가 배정되는 것이 아니라 학교가 학생들을 선발하는 시스템이므로 형제 · 자매 · 남매가 학교의 질이 다른 학교에 다닐 가능성이 평준화지역 가구에 비해 높아 상호 간에 공유하는 자원 또한 적을 가능성이 크다. 따라서 <표 2>의 결과를 평준화 혹은 비평준화의 효과로 해석할 수는 없다.

2. 가구배경과 이웃의 대리변수로서의 형제·자매·남매의 기대수능성적의 적정성 검토

형제·자매·남매의 실제 수능성적이 아닌 기대수능성적을 설명변수로 사용하였을 경우 실제 수능성적 중 대학교명과 학과정보에 의해 설명되지 않는 부분에 의해 항상 가구배경과 이웃효과의 기여도가 과소추정됨을 앞에서 보였다. 이 절에서는 가구배경과 이웃의 대리변수로서 형제·자매·남매의 기대수능성적을 사용하는 것의 적절성의 정도를 살펴보고 과소추정문제를 보완하고자 구체적인 가구배경 변수들과 이웃을 나타내는 학교더미들을 통제한 결과를 살펴보고자 한다.

<표 3>은 평준화지역의 응답자들을 대상으로 하여, 형제·자매·남매의 기대수능성적과 더불어 가구소득, 아버지와 어머니의 교육수준, 월사교육비지출, 어머니의 전문직 종사여부와 비전문직 종사여부,⁸⁾ 아버지와 어머니의 연령 등의 관측할 수 있는 가구배경 변수들을 설명변수로 추가하여 수능표준점수 결정식을 회귀분석하였을 경우의 결과들이다. 열 (1)부터 (4)는 형제·자매·남매의 기대수능성적을 통제하였을 경우의 회귀분석 결과들을 정리한 것이고, 열 (5)부터 (8)은 형제·자매·남매의 기대수능성적을 통제하지 않았을 경우의 회귀분석 결과들을 정리한 것이다. 한편 열 (1), (2), (5), (6)은 학교 더미들을 포함하지 않고 회귀분석한 결과이고 나머지 열들은 추가로 학교더미들을 포함하여 이웃효과를 통제하였을 경우의 회귀분석 결과들이다.

<표 3>에서 확인할 수 있는 바와 같이 어머니의 비전문직 종사여부와 연령을 제외한 다른 구체적인 가구배경 변수들은 대부분 통계적으로 유의하지 않았다.⁹⁾

주목할 만한 것은 어머니의 교육수준의 경우 그 기여도가 음의 값을 가졌고 유의하지 않아 일반적인 예상과는 다른 결과를 보여준다. 교육연수 대신 어머니의 학력더미를 사용하여도 상황을 바꾸지 않았다.¹⁰⁾

8) 전문직여부는 직업분류가 ‘의회의원, 고위임직원 및 관리자’, ‘전문가’, ‘기술공 및 준전문가’인 경우 전문직으로 분류하고 그렇지 않은 경우 비전문직으로 간주하였다.

9) 사교육비 지출의 경우 각 자녀별 지출액 규모를 분리할 수 없어 본 논문에서는 사교육비 지출의 효과를 자세히 논의할 수 없다는 한계를 가지고 있다.

10) 변수들 간의 상관관계 때문에 어머니의 학력 역할이 과소평가될 가능성을 살펴보기 위해 각 변수들을 개별적으로 통제하면서 회귀분석해 보았지만 일관되게 어머니의 학력은 아버지의 학력보다 그 크기가 작았으며 신뢰도도 낮았다. 이러한 결과가 본 연구의 표본특

〈표 3〉 평준화지역 형제/자매/남매표본을 대상으로 한 수능표준점수 결정식 추정결과

	형제/자매/남매의 기대수능성적 통제				형제/자매/남매의 기대수능성적 통제하지 않음			
	학교 통제 없음		학교 통제		학교 통제 없음		학교 통제	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
형제/자매/남매의 수능성적	0.1538** (0.0622)	0.1622*** (0.0625)	0.1388** (0.0636)	0.1529** (0.0639)				
로그가구소득	5.9244 (9.1101)	7.6439 (9.1550)	12.1067 (9.8678)	13.9353 (9.9421)	9.3482 (9.0767)	11.2525 (9.1313)	13.9477 (9.9046)	16.1975 (9.9905)
로그월사교육비	0.6254 (4.1367)	0.0699 (4.1476)	2.2550 (4.4864)	1.3783 (4.4988)	-0.1423 (4.1582)	-0.6294 (4.1766)	1.5434 (4.5077)	0.6781 (4.5317)
교육연수	1.4213 (1.6542)		0.9355 (1.7312)		2.0210 (1.6495)		1.6098 (1.7160)	
고졸		16.3413 (13.1709)		13.4323 (13.0974)		20.0128 (13.2142)		17.0993 (13.1303)
초대졸		20.8998 (17.4461)		23.4245 (17.6870)		23.6166 (17.5736)		26.7624 (17.7984)
부 대출		20.8467 (15.4388)		14.3016 (16.0156)		26.6486* (15.4155)		20.5770 (15.9487)
모 교육연수	-0.0983 (1.9498)		-1.6967 (2.0483)		-0.2080 (1.9650)		-1.8935 (2.0615)	
고졸		-1.9191 (9.6102)		-5.2780 (10.2245)		-4.6572 (9.6393)		-8.4960 (10.2314)
초대졸		-46.1701** (19.6923)		-43.7322** (20.0534)		-45.4585** (19.8700)		-43.1628** (20.2414)
대출		-2.7500 (15.2337)		-14.5558 (16.3116)		-2.8787 (15.3726)		-16.2455 (16.4502)
어머니의 전문직종사여부	0.0624 (12.2686)	7.4673 (12.6484)	-6.7823 (12.7407)	-0.6673 (12.9707)	-0.5762 (12.3646)	5.4550 (12.7398)	-8.5891 (12.8081)	-3.5036 (13.0385)
어머니의 비전문직종사여부	-13.5214* (7.3936)	-12.5087* (7.4554)	-7.2871 (7.6249)	-6.5758 (7.6996)	-15.3990** (7.4137)	-14.3274* (7.4901)	-9.0518 (7.6382)	-8.4378 (7.7325)
아버지의 연령	-1.5311 (1.4771)	-1.3766 (1.4748)	-0.6836 (1.5165)	-0.5716 (1.5323)	-1.2659 (1.4851)	-1.0999 (1.4843)	-0.2871 (1.5167)	-0.0939 (1.5336)
어머니의 연령	3.1380* (1.6489)	2.9887* (1.6441)	1.2231 (1.6812)	1.2806 (1.7010)	2.9257* (1.6599)	2.7437* (1.6564)	0.9362 (1.6885)	0.8993 (1.7095)
상수항	131.4607* (75.9498)	123.7139 (79.9510)	35.7651 (89.2054)	3.0894 (93.2521)	150.9831** (76.1464)	149.7957* (80.0407)	48.7544 (89.6666)	19.5547 (93.8763)
관측치수	326	326	326	326	326	326	326	326
R ²	0.076	0.099	0.373	0.388	0.058	0.080	0.361	0.374

주: *** p<0.01, ** p<0.05, * p<0.1.

성에서 기인하는 것으로는 보이지 않는다. 어머니의 역할이 중요하리라는 일반적인 기대와는 달리 대부분의 실증분석 결과는 아버지의 교육수준의 역할이 큰 것으로 나타난다. Behrman and Rosenzweig(2002)는 부모의 교육연수와 자녀의 교육연수 간의 관계를 분석한 선행연구들을 검토해 본 결과 아버지의 교육연수의 효과가 어머니의 교육연수 효과보다 자녀에게 미치는 영향이 크다는 점을 지적하였다. 우리나라의 선행연구들도 부모의 동질혼으로 인한 상관관계 때문에 일반적으로 두 변수를 함께 통제하는 것을 피하고 있는데, 부모의 학력을 동시에 고려한 김현주·이병훈(2005)에서 부의 학력은 모의 학력보다 자녀의 학업성취에 미치는 영향이 큰 분석결과를 보였다.

<표 3>에서 형제·자매·남매의 기대수능성적의 경우 0.14~0.16사이의 값을 보였고 모두 통계적으로 유의하였다. <표 2>의 결과값보다 계수값이 낮아진 이유는 추가적인 설명변수의 도입으로 이들의 직접적인 효과가 통제되었기 때문이다. 다시 말하면 <표 2>에서 별다른 추가변수 없이 형제·자매·남매의 기대수능성적만을 통제하여도 이러한 변수들의 효과의 상당 부분을 통제할 수 있었음을 의미한다. 만약 기대수능성적이 아닌 실제 수능성적을 사용하였다면 더욱 효과적인 추정 결과를 얻을 수 있었을 것으로 생각할 수 있다. 또한 형제·자매·남매의 기대수능성적을 포함하였을 때 로그가구소득 등의 구체적인 가구배경 변수들의 효과가 크게 감소하였다는 사실은 형제·자매·남매의 기대수능성적이 구체적인 가구배경들의 상당부분을 설명하고 있어 좋은 대리변수가 될 수 있음을 의미하여 모형의 예상과도 일치하였다.¹¹⁾

한편 학교더미를 통제하였을 때 모형의 설명력이 크게 상승하고 형제·자매·남매의 기대수능성적의 기여도는 별다른 변화를 보이지 않는 것으로 비추어 보아 학교더미에 의해 설명되는 상당부분은 형제·자매·남매 간에 공유하는 순수한 이웃효과보다는 학교의 질, 교사의 질, 동료효과 등에 의해 설명되는 순수한 학교효과일 가능성을 암시하고 있다. 달리 말하면 평준화지역이라 하더라도 거주지역 내의 학교들이 비슷한 크기의 학교효과를 가지고 있다기보다는 그 효과가 학교에 고유한 성질일 것임을 암시하고 있다.

3. 지역적 격리(segregation)에 의한 추가적인 학교효과

서론에서 언급했던 바와 같이 최근 우리나라의 교육문제에 있어서 관심을 끄는 주장 중 하나는 가구소득과 높은 지가에 의해 지역 간 격리(segregation)가 발생하고 그 결과 고소득층이 거주하는 소위 강남지역에 양질의 학교와 각종

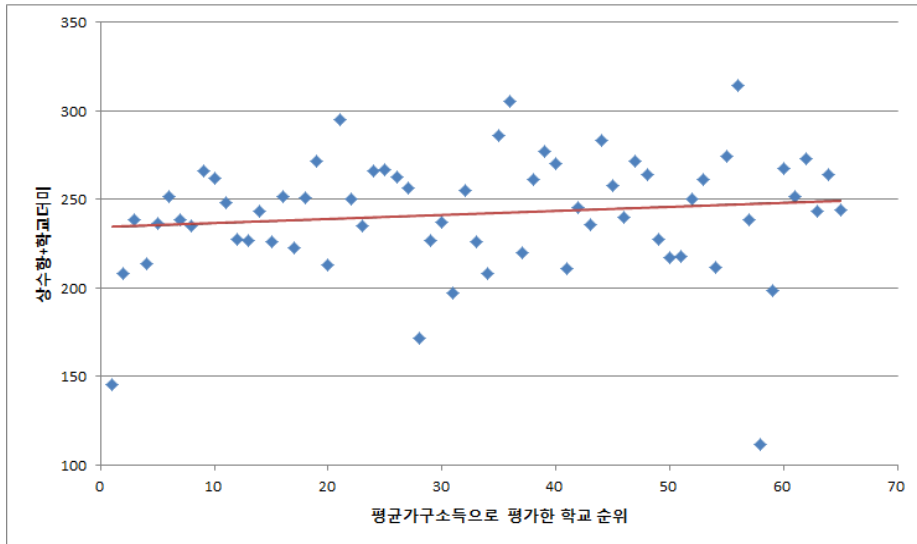
11) <표 3>의 결과는 기대수능성적의 통제로 다른 가구배경 변수들의 설명력이 크게 감소하지 않는 것으로 해석할 여지도 있다. 그러나 <표 3>은 지면의 한계로 모든 변수들을 동시에 통제할 경우들만 보여준 것이고 변수들을 개별적으로 통제하였을 경우에는 <표 3>보다 설명력이 감소하는 정도가 뚜렷했다. 가령 로그가구소득의 경우 유일한 설명변수로 통제할 경우 계수값은 13.2549이고 P값은 0.016이나 기대수능성적을 추가로 통제할 경우 계수값은 8.3636로 낮아지고 P값은 0.131이었다. 반면 기대수능성적은 계수값이 0.2067이고 1% 수준에서도 유의하여 큰 변화가 없었다.

사교육기관의 집중으로 소득수준의 차이를 넘어서는 교육기회의 불균등성이 심화된다는 것이다. 학교더미들을 통제하면서 설명변수로 형제 · 자매 · 남매의 기대수능성적과 로그가구소득만을 통제하여 회귀분석할 경우 총 64개의 학교 더미들 중 42개가 1% 유의수준에서 통계적으로 유의했으며, 13개는 5% 수준에서, 4개는 10% 수준에서 통계적으로 유의하였고 5개의 학교더미들만이 기각되었다. 따라서 이러한 결과는 이웃효과가 상당히 존재하고 있을 가능성을 암시하고 있다. 그러나 앞서서도 언급했던 바와 같이 형제 · 자매 · 남매의 기대수능성적을 통제하였음에도 학교더미들을 포함하였을 경우 모형의 설명력이 크게 높아지고 또한 형제 · 자매 · 남매의 기대수능성적은 학교더미의 통제여부에 민감하지 않다는 것은 오히려 학교효과들이 거주지역의 환경과는 크게 관련이 없는 순수한 학교효과일 가능성이 크다는 것을 암시하고 있다.

가구소득과 높은 지가에 의한 지역 간 격리(segregation)의 발생과 이로 인한 교육기회의 불균등성 심화 주장의 타당성을 확인하기 위하여 본 절에서는 학교의 평균가구소득과 이웃효과의 크기 간의 관련성을 살펴보았다. 먼저 응답자의 수능성적을 형제 · 자매 · 남매의 기대수능성적과 학교더미에 대하여 회귀분석을 하였다. 다음은 각 학교의 가구소득의 평균치를 구하여 그 값이 낮은 지역부터 높은 지역으로 오름차순으로 순위를 설정하였다. 이렇게 설정한 학교의 순위를 횡축으로 한 후 종축에는 회귀분석 결과의 상수항과 학교더미의 계수값을 합한 값을 나타내어 [그림 1]을 작성하였다.

[그림 1]은 평균가구소득과 이웃효과 혹은 학교효과 간에 아주 약한 선형관계가 존재함을 보여준다. 평균가구소득이 높은 학교일수록 학교효과 또한 높은 경향을 가지고는 있으나 전체 학교효과 중에서 차지하는 정도는 상대적으로 크지 않은 수준이라고 할 수 있다. 즉 형제 · 자매 · 남매의 기대수능성적에 의해 표현되는 개인적인 가구배경들이 통제되면 부유한 가구들이 많이 거주하는 지역의 학교를 다님에 의해서 얻게 되는 추가적 이익은 전체 학교효과에 비해 그다지 크지 않다는 것이다. 형제 · 자매 · 남매의 기대수능성적이 통제되고 나면 이것에 의해 설명되지 않는 학교효과는 무작위과정에 가까워서 좋은 교사와 좋은 학생들이 부유한 지역에 집중되어 발생하는 지역적 격리현상이 두드러진 것으로는 볼 수 없었다. 학교효과가 가장 컸던 학교는 과학고등학교였다. 과학고

(그림 1) 평균가구소득으로 평가한 학교순위와 학교효과

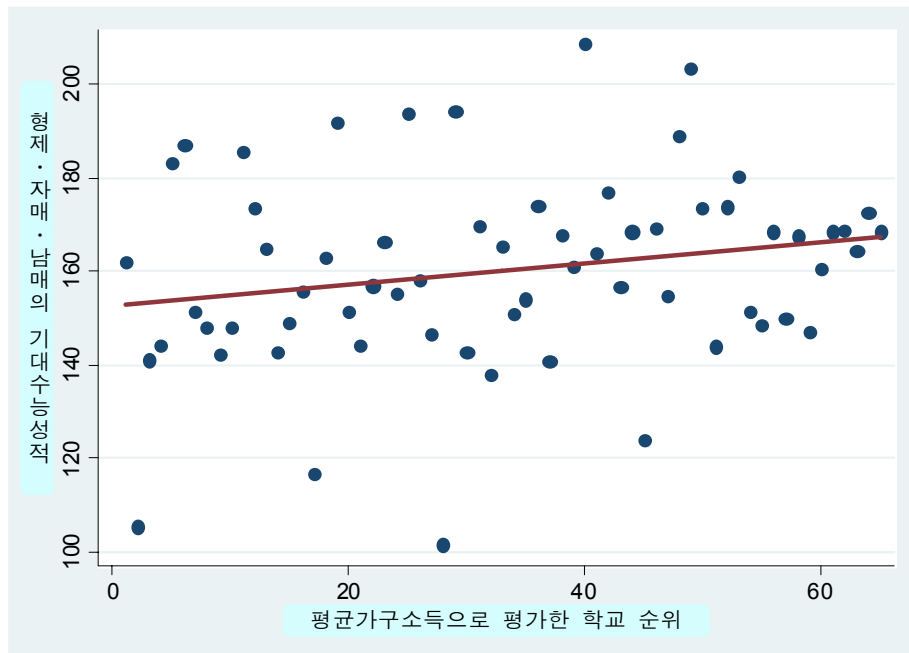


등학교는 비록 평준화지역에 있지만 별도의 시험을 통하여 학생들을 선발하므로 동료효과가 클 것이고 양질의 선생님들에 의한 교사효과 또한 클 것으로 추정해 볼 수 있다. 흥미로운 것은 과학고등학교 학생들 부모의 평균소득이 상대적으로 높은 편에 속한다는 것이다. 그러나 과학고등학교는 단 한 개에 불과하여 신뢰성에는 문제를 가지고 있다. 한편 학교효과가 가장 낮았던 학교는 예술고등학교였다. 그러나 형제·자매·남매의 기대수능성적과 다른 가구배경 변수들에 의해 설명되지 않는 가구배경 변수들의 누락으로 인하여 이러한 누락변수들의 효과의 일부는 상관관계를 가지고 있는 이웃효과에 포함되어 상향편의를 발생시키고 있을 것임에 주의하여야 한다. 또한 상수항은 개인에 고유한 능력의 평균치를 포함하고 있으므로 그 자체가 학교효과를 의미하지 않는다는 것도 주의하여야 할 것이다.

한편, 형제·자매·남매의 기대수능성적이 학교효과와 어떠한 관계를 가지고 있는지 살펴보기 위해 추가적으로 형제·자매·남매의 기대수능성적을 종속변수로 하여, 로그가구소득에 대하여 회귀한 후 상수항과 잔차의 합을 종속으로 하고 평균가구소득으로 평가한 학교순위를 횡축으로 하여 그 관계를 [그림 2]에 나타내었다.¹²⁾ 가구소득을 통제 한 이후에도 동생이 다니는 학교의 평균가구

소득이 높을수록 형제 · 자매 · 남매의 기대수능성적 또한 상승하는 경향이 있는 것으로 보이는데, 가장 평균가구소득이 높은 지역의 경우 낮은 지역에 비해 평균적으로 약 20점가량의 수능성적 향상을 기대할 수 있는 것으로 볼 수 있다. 따라서 평균가구소득이 높은 지역의 경우 평균적으로 소득 이외의 수능점수 향상에 기여하는 좋은 양육환경 등의 가구배경을 가지고 있거나 이웃효과로 인하여 추가적인 수능성적의 향상이 있는 것으로 간주할 수 있을 것이다. 그러나 그 크기는 [그림 1]의 경우와 유사하다는 것을 확인할 수 있으며 따라서 [그림 1]이 암시하는 바와 같이 지역적 격리에 의한 추가적인 이득은 학교 고유의 효과에 비하면 그 크기가 상대적으로 작음을 다시 한번 보여주고 있다고 할 수 있다.

[그림 2] 평균가구소득으로 평가한 학교 순위와 형제 · 자매 · 남매의 기대수능성적



12) 가구배경효과와 이웃효과를 구분하는 작업은 Solon et al.(2000)의 방법을 이용하여 각 효과를 분해(decomposition)해 볼 수 있는데 본 연구의 목적을 벗어나기 때문에 시도하지 않았다.

4. 표본선택문제(Sample Selection Problem)

형제·자매·남매의 기대수능성적을 가구배경과 이웃효과의 대리변수로 사용하는 경우 발생하는 가장 큰 문제는 대학을 다니고 있는 형제·자매·남매가 있는 경우에만 분석대상에 포함됨으로써 발생하는 표본선택문제이다.

먼저, 대학을 다니고 있는 형제·자매·남매가 있는 경우에만 표본에 포함되므로 첫째 자녀인 경우 표본에서 빠지게 되는데, 출생순서가 교육성과에 미치는 영향을 분석한 선행연구들은 대부분 출생순서가 늦은 자녀가 교육성과가 더 좋지 못함을 발견하였다(조현국, 2008; Behrman and Taubman, 1986; Hanushek, 1992; Black et al., 2005; Kantarevic and Mechoulan, 2006). 따라서 본 연구에서 사용한 표본은 평균보다 성적이 좋지 못한 학생들일 가능성이 있다.

둘째, 가구의 자녀수도 가구소득이 동일할 경우 그만큼 분배해야 할 자원이 제약될 것이므로 자녀수가 많은 가구의 자녀일수록 교육성과가 좋지 않을 가능성이 있다. 이러한 경우 형제·자매·남매가 있는 경우에만 표본에 한정하는 것은 마찬가지로 사회 평균보다 성적이 좋지 못한 학생들이 선택될 가능성이 있다. 그러나 Black et al.(2005)은 가구의 크기가 자녀의 교육성과에 그다지 영향을 주지 않는다는 연구결과를 보여주었다.

셋째, 형제·자매·남매가 있다 하더라도 대학에 진학하지 못한 경우 표본에서 누락되는데, 대학입학을 하지 못한 형제·자매·남매를 둔 응답자의 경우 그렇지 않은 경우보다 유전적 혹은 환경적으로 좋지 않은 가구배경을 가지고 있을 가능성이 높다.

마지막으로, 가구배경과 자녀수와의 관계이다. 학력수준이 낮을수록 출산에 있어서의 계획성이 낮다는 것을 감안하면 형제·자매·남매 표본의 가구배경은 사회 평균보다 가난할 가능성이 있다. 반대로 가구소득이 높을수록 출산을 후생증가로 받아들인다는 가설을 받아들이면 오히려 형제/자매/남매표본은 사회 평균보다 부유할 것이다.

<표 4>는 본 연구에서 사용된 평준화지역의 ‘형제·자매·남매표본’과 ‘전체 표본’ 중에서 이 표본에서 누락된 응답자들 간의 본인 수능성적과 주요 가구들의 변수들의 특성을 비교한 것이다. 첫 번째 행은 표본에서 누락된 응답자들의

성적과 가구특성들을, 두 번째 행은 ‘형제 · 자매 · 남매표본’에 포함된 응답자들의 수능점수와 가구특성들의 평균값을 나타내고 있다. 마지막 행은 이렇게 구한 수능점수와 각 가구특성들이 두 그룹 간에 평균값이 같다는 귀무가설을 검정(test)하였을 때 귀무가설하에서 <표 4>와 같은 차이가 발생할 확률을 나타내고 있다. 결과를 보면 수능성적은 표본누락자들에 비해 표본에 포함된 응답자들의 성적이 약 2.8점 정도 높은 것으로 나타난 반면 가구소득과 사교육비지출, 아버지와 어머니의 교육연수 등은 오히려 표본누락자들이 약간 좋은 것으로 나타나고 있다. 어머니의 고용률은 표본누락자의 경우가 5%포인트 높았다. 두 그룹의 평균이 같다는 귀무가설에 대한 검정 결과는 가구소득과, 사교육비지출, 어머니의 교육연수는 두 그룹 간에 차이가 있다고 할 수 있었고, 수능성적과 아버지의 교육연수는 두 그룹 간에 차이가 있다고 볼 수 없었다. 어머니의 고용률은 10% 유의수준에서 기각되지 않지만 경계값 근처에 해당되어 두 그룹 간 차이가 있을 가능성을 배제할 수는 없었다. 따라서 ‘형제 · 자매 · 남매 표본’은 전체평균보다 가구소득과 사교육비지출, 어머니의 교육연수는 약간 낮으며 어머니의 고용률 또한 약간 낮을 가능성이 있다고 할 수 있었고 이러한 측면에서 앞의 결과들에 대한 신뢰성에 문제가 있을 수 있다는 가능성을 배제할 수 없었다.

이질적인 두 집단 중 한 집단만을 대상으로 하였을 때 발생하게 될 편위의 방향을 확정적으로 계산할 수는 없으나 본 연구의 경우에는 <표 2>의 결과들에 근거하여 추정해보면 하향편의가 발생했을 것으로 예상되고 따라서 대상을 확장할 경우 <표 2>에서 보여준 이전의 결과들보다는 더 큰 하계값을 얻을 것으로 여겨진다. <표 2>에서 평준화지역만을 대상으로 한 결과인 19.2%보다 비평준화지역을 대상으로 한 전체집단을 대상으로 추정하였을 때 이보다 1%포인트 높은 20.2%라는 결과를 얻었기 때문이다.

본 연구에서는 표본선택문제에 의한 편의발생문제를 수정하고 그 방향을 살펴보기 위해, 고졸 형제 · 자매 · 남매만을 가진 표본을 추가해 보았다. 이를 위해 고졸의 기대수능성적이 필요한데 이는 다음과 같은 절차로 추정하였다. 먼저 KEEP에서 3차년도까지 대학입학 정보가 없는 응답자들의 2005년도 수능성적의 평균을 학교별로 평균하였다. 평준화지역의 경우 학교는 거주지역에 따

<표 4> 평준화지역 형제·자매·남매표본과 표본 누락 응답자들 간의 가구특성 비교

	수능점수	로그가구 소득	로그월평균 사교육비지출	아버지의 교육연수	어머니의 교육연수	어머니의 고용률
표본누락	289.50	5.70	3.89	13.21	12.14	0.61
형제·자매· 남매표본	292.34	5.64	3.61	13.06	11.81	0.56
전체	290.54	5.68	3.80	13.15	12.02	0.59
P-value	0.43	0.04	0.00	0.36	0.02	0.12

<표 5> 고졸 형제·자매·남매를 포함한 수능표준점수 결정식 추정결과

	평준화지역	비평준화지역	전 체
형제/자매/남매의 기대수능성적	0.2531*** (0.0484)	0.2355*** (0.0651)	0.2768*** (0.0388)
상수항	216.1520*** (14.3022)	199.9552*** (18.0163)	202.3891*** (11.2268)
관측치 수	500	252	752
R^2	0.052	0.050	0.064
$V(S_i)$	3,575.966	3,508.997	3,698.536
$V(\hat{S}_j)$	2,907.748	3,149.574	3,069.727
$Corr(S_i, S_j)$	0.206	0.211	0.230

주: *** p<0.01, ** p<0.05, * p<0.1.

라 배정받으므로, 동일 학군 내의 학교들은 다른 학군의 학교들보다는 서로 공유하는 부분이 클 것이기 때문이다. 물론 앞에서 언급한 바와 같이 학교효과는 공유하지 않을 가능성 또한 높지만 최소한 추정결과를 과대추정하는 결과를 가져오지는 않을 것이기 때문에 표본선택문제로 인한 방향을 살펴보는 데는 문제가 되지 않는다. 이렇게 구한 학교별 고졸자 평균수능점수를 고졸 형제·자매·남매의 기대수능점수로 결합하였다.¹³⁾ <표 5>은 이렇게 고졸 형제·자매·남매까지 표본을 확장하여 <표 2>의 회귀식을 동일하게 추정한 결과이다. 예상한 것처럼 모두 <표 2>의 결과보다는 더 높았다. 평준화지역의 경우 하계

13) 대학을 입학한 형제·자매·남매를 확인할 수 없는 경우에 한정하여 고졸 형제·자매·남매만이 있는 경우 이를 결합하였다.

값은 20.6%로 상승하였으며 전체적으로 23%로 상승하여 <표 2>보다 각각 1.4%포인트, 2.8%포인트 증가하였다. 그러나 여전히 형제 · 자매 · 남매가 없거나 첫째인 경우 회귀식에서 빠지게 되는데, <표 2>와 <표 5>의 결과에 비추어 보면 이들의 포함은 추정치를 낮추지는 않을 것으로 생각되어 평준화지역 20.6%, 비평준화지역 21.1%, 전체지역 23%는 자녀의 수능성적에 가구배경과 이웃효과 등의 비개인적 요소가 미치는 영향의 하계로 간주할 수 있다.¹⁴⁾

V. 맺음말 : 토론과 연구함의

본 연구의 목적은 자녀의 교육성가에 가구배경과 이웃효과가 미치는 영향을 측정하는 것이다. 구체적인 가구배경 변수들을 사용할 경우 관측 가능한 가구배경 변수들의 효과만 측정가능하기 때문에 가구배경과 이웃효과의 제한된 효과만이 측정가능하다는 문제가 있다. 이러한 문제를 최소화하기 위해 본 연구에서는 구체적인 가구배경 변수들을 사용하는 대신 대학을 다니고 있는 형제 · 자매 · 남매의 대학과 학과정보를 이용하였다. 우리나라의 대학은 대학과 학과 별로 서열화되어 있어 이들 정보를 사용하여 기대수능점수를 예측하는 데 좋은 조건을 제공한다. 이렇게 대학과 학과정보를 이용하여 형제 · 자매 · 남매들의 기대수능점수를 예측한 후 이를 가구배경과 이웃효과의 대리변수로 사용하여 가구배경과 이웃효과가 자녀의 수능성적에 미치는 영향을 회귀분석하였는데 주요 결과는 다음과 같다.

첫째, 가구배경과 이웃효과는 수능성적으로 평가한 자녀의 교육성파에서 적어도 20.2%를 설명할 수 있는 것으로 나타났다. 평준화지역에만 한정하였을 경우 자녀의 수능성적의 19.2% 이상을 설명할 수 있는 것으로 나타났다. 또한 고졸 형제 · 자매 · 남매들까지 포함할 경우 23%의 하계값을 얻을 수 있었다.

둘째, 형제 · 자매 · 남매의 기대수능성적은 관측 가능한 구체적인 가구배경

14) 형제 · 자매 · 남매의 졸업 시점을 감안하지 못하였는데, 이로 인한 문제가 어느 정도인지 확인하기 위하여 형제 · 자매 · 남매와의 터울을 독립변수로 추가한 모형을 추정해보았는데 결과에 있어서 실질적인 차이는 없었다.

변수들과 관련성이 높고 구체적인 가구배경 변수들보다 그 효과의 크기가 크고 유의해서 가구배경의 대리변수로 적합하다고 할 수 있었다.

셋째, 이웃효과의 대부분은 지역적 격리에 의한 기회의 불균등 문제보다는 순수한 학교효과일 가능성이 높은 것으로 나타났다.

그러나 본 연구에서 사용한 표본은 관측치의 수가 많지 않을 뿐만 아니라 형제·자매·남매가 없거나 첫째인 경우 분석대상에서 제외되어 표본선택문제가 발생할 가능성이 있으므로 이에 대한 주의가 필요하다. 특히 어머니의 양육(nurturing)역할에 관한 보다 심도 깊은 추가연구가 필요한 것으로 생각된다. 또한 가구배경과 이웃효과를 좀 더 구체적으로 양육효과(nurturing effect), 동료효과(peer effect), 교사효과(teacher effect), 순수한 이웃효과(pure neighborhood effects), 학교자원(school resources)효과 등으로 나누어 살펴보는 연구 또한 귀중한 연구가 될 것이다.

한편 Björklund and Salvanes(2011)는 미국과 서부유럽을 대상으로 한 형제·자매·남매의 상관관계 추정 연구들의 결과들을 정리하여, 통독 이전의 서독을 제외하고는 대부분 40~65% 사이에 있는 것으로 결론 내렸다. 심지어 스웨덴과 노르웨이와 같은 북유럽 복지국가조차 추정치는 0.40~0.50이었다. 이에 따르면 20~25% 정도의 우리나라 추정치는 매우 낮은 것으로 볼 수 있을 것이다. 본 연구의 결과는 우리나라의 경우 타국과 비교하여 상대적으로 자녀의 교육성장에 부모의 사회경제적 배경이 미치는 영향이 작을 것임을 시사하고 있다고 할 수 있다. OECD(2010)에 따르면 학생의 성취도평가에 부모의 사회경제적 배경이 미치는 비율이 우리나라는 OECD 34개 회원국 중 7번째로 낮은 국가로 나타나 본 연구의 시사점과 일치하였다. 이렇게 우리나라에서 자녀의 성적에 사회경제적 배경의 영향력이 낮게 나타나는 원인을 밝히는 것도 앞으로 귀중한 연구가 될 것으로 생각한다. 한 가지 분명한 것은 그동안의 우리나라 교육정책이 형평성의 측면에서는 성공을 거두었을 가능성이 높다는 것이다. 과거 교육정책의 가장 큰 중심축의 하나가 평준화정책이었음을 고려하면, 평준화정책의 선부른 수정은 형평성 측면에서 안 좋은 결과를 가져올 수도 있음에 주목하여야 할 것이다. 자녀의 성적에 대한 사회경제적 배경의 영향력이 낮게 나타나는 이유를 밝히는 연구와는 별도로 평준화정책의 효과에 대한 연구가 보다 활발히

이루어질 필요가 있을 것이다.

궁극적으로 자녀의 교육성과에 부모의 사회경제적 배경이 미치는 영향을 분석하는 가장 중요한 동인으로서 앞서서도 이야기한 바와 같이 세대 간의 소득 이동성을 고려하지 않을 수 없을 것이다. 교육을 통한 세대 간의 경제적 부의 대물림 가능성은 형평성 측면에서 주요하게 고려하여야 할 사항 중의 하나이다. 양정승(2012b)는 우리나라의 세대 간 소득탄력성이 0.3~0.4일 것으로 추정하여 미국보다는 낮지만 서부유럽국가들보다는 높을 것으로 추정하였는데, 이러한 추정치에도 불구하고 자녀의 교육성과에 부모의 사회경제적 배경이 미치는 영향이 낮다는 본 연구의 결과는 두 가지 시사점을 제공한다. 첫째, 우리나라의 세대 간 소득이동성이 미국보다 높은 이유 중의 하나로 자녀의 교육에 대한 가구배경의 영향력이 작다는 것이 작용하고 있을 가능성이 높다는 것이다. 둘째, 가구배경이 영향력이 자녀에게 미치는 영향력이 서부유럽에 비해 낮음에도 불구하고 소득이동성은 서부유럽보다 낮다는 것은 우리나라의 세대 간 소득이전의 교육을 통하지 않은 다른 경로의존성이 크다는 것이다. 따라서 세대 간 소득이전의 구체적인 메커니즘을 연구하는 것이 필요하다는 것을 제기하고 있다.

그러나 서구의 연구는 교육성과지표로서 시험점수가 아닌 교육연수를 사용하고 있어 측정수단이 다르며, 또한 교육연수는 최종적 교육성과지표이기 때문에 아직 교육을 마치지 않은 세대를 대상으로 한 시험점수를 사용한 본 연구와의 직접비교는 한계를 가지고 있음을 유의하여야 할 것이다. 또한 본 연구의 추정치는 형제 · 자매 · 남매의 경우는 실제 교육성과가 아닌 기대치이기 때문에 앞서서도 지적한 바와 같이 하향편의를 가짐에 유의하여야 한다.

추가적으로 자녀의 수능성적이 과연 교육적 성과를 측정하는 가장 좋은 지표인가 하는 문제가 남아 있다. 가령 명문대 효과가 매우 크다고 하다면 교육성과를 명문대 입학여부 혹은 학교의 서열로 측정하는 것이 보다 바람직할 수도 있을 것이다. 이에 대해서도 추가적인 연구가 진행될 것으로 기대해 본다.

참고문헌

- 강창희·채창균(2010). 「대입수학능력시험 성적 자료를 이용한 학급 내 동료 효과의 분석」. 『노동경제논집』 33 (1): 1~30.
- 고제이·이우진(2011). 「아버지의 학력과 아들의 성취」. 『재정학연구』 4 (2): 47~97.
- 김미란(2009). 「가계의 사교육비 지출 규모와 영향요인 분석」. 『교육재정경제 연구』 18 (2): 1~25.
- 김현주·이병훈(2005). 「부모의 가족배경과 사회문화적 자원이 자녀의 학업성취에 미치는 영향」. 제1회 한국교육고용패널 학술대회자료집.
- 방하남·김기현(2002). 「기회와 불평등: 고등교육 기회에 있어서 사회계층간 불평등의 분석」. 『한국사회학』 36 (4): 193~222.
- _____(2003). 「한국사회의 교육계층화: 연령코호트간 변화와 학력단계별 차이」. 『한국사회학』 37 (4): 31~65.
- 안중범·진승훈(2008). 「교육 및 소득수준의 세대간 이진」. 『재정학연구』 1 (1): 119~142.
- 양정승(2012a). 『세대간 소득이동과 소득분배에 관한 연구』. 서울대학교 박사 학위논문.
- _____(2012b). 「한국의 세대 간 소득이동성 추정」. 『노동경제논집』 35 (2): 79~115.
- 오호영(2006). 「소득계층과 학업 성취도」. 제2회 한국교육고용패널 학술대회 자료집.
- 이주호·김선웅(2002). 「학교 정책과 과외의 경제 분석」. 『한국경제의 분석』 9 (2): 1~51.
- 이주호·홍성창(2001). 「학교 대 과외: 한국 교육의 선택과 형평」. 『경제학 연구』 49 (1): 39~56.
- 조현국(2008). 「출생순서가 교육성과에 미치는 영향」. 제4회 한국교육고용패널

널 학술대회자료집.

- Aaronson, Daniel(1998). "Using Sibling Data to Estimate the Impact of Neighborhoods on Children's Educational Outcomes." *Journal of Human Resources* 33 (4): 915~946.
- Behrman, Jere R. and Paul Taubman(1986). "Birth Order, Schooling and Earnings." *Journal of Labor Economics* 4 (3): S121~45.
- Behrman, Jere R. and Mark R. Rosenzweig(2002). "Does Increasing Women's Schooling Raise the Schooling of the Next Generation?" *The American Economic Review* 92 (1): 323~334.
- Björklund and Salvanes(2011). "Education and Family Background: Mechanisms and Policies." In Eric A. Hanushek, Stephen Machin and Ludger Woessmann(eds). *Handbook of the Economics of Education*. Amsterdam: Elsevier: pp.201~247.
- Black, Sandra E., Paul J. Devereux, and Kjell G. Salvanes(2005). "The More the Merrier? The Effect of Family Composition on Children's Outcomes." *The Quarterly Journal of Economics* 120 (2): 669~700.
- Dietz, Robert D.(2002). "The estimation of neighborhood effects in the social sciences: An interdisciplinary approach." *Social Science Research* 31: 539~575.
- Hanushek, Eric A.(1992). "The Trade-off between Child Quantity and Quality." *Journal of Political Economy* 100 (1): 84~117.
- Haurin, Donald R., Robert D. Dietz, and Bruce A. Weinberg(2003). "The Impact of Neighborhood Homeownership Rates: A Review of the Theoretical and Empirical Literature." *Journal of Housing Research* 12 (2): 119~151.
- Jacob, Brian A.(2004). "Public Housing, Housing Vouchers, and Student Achievement: Evidence from Public Housing Demolitions in Chicago." *The American Economic Review* 94 (1): 232~258
- Jencks, Christopher and Susan E. Mayer(1990). "The Social Consequences of

- Growing Up in a Poor Neighborhood.” In L. Lynn and M. McGeary(eds). *Inner-City Poverty in the United States*. Washington, D. C.: National Academy Press: pp.111~186.
- Kantarevic, Jasmin and S. Mechoulan(2006). “Birth Order, Educational Attainment and Earnings: An Investigation Using the PSID.” *Journal of Human Resources* 41 (4): 755~777.
- Katz, Lawrence F., Jeffrey R. Kling, and Jeffrey B. Liebman(2001). “Moving to Opportunity in Boston: Early Results of a Randomized Mobility Experiment.” *The Quarterly Journal of Economics* 116 (2): 607~654.
- Kang, C.(2007). “Classroom Peer Effects and Academic Achievement: Quasi-randomization evidence from South Korea.” *Journal of Urban Economics* 61 (3): 458~495.
- Manski, Charles F.(1993). “Identification of Endogenous Social Effects: The Reflection Problem.” *Review of Economics Studies* 60 (3): 531~542.
- Mayer, Susan E. and Christopher Jencks(1989). “Growing Up in Poor Neighborhoods: How Much Does It Matter?” *Science* 243 (17): 1441~1445.
- Mincer, Jacob(1974). *Schooling, Experience and Earnings*. Columbia University Press: New York.
- OECD(2010). “PISA 2009 Results: Overcoming Social Background - Equity in Learning Opportunities and Outcomes(Volume II).” <http://dx.doi.org/10.1787/9789264091504-en>.
- Oreopoulos, Philip(2003). “The Long-run Consequences of Living in a Poor Neighborhood.” *The Quarterly Journal of Economics* 118 (4): 1533~1575.
- Oreopoulos, Philip and Kjell G. Salvanes(2011). “Priceless: The Nonpecuniary Benefits of Schooling.” *Journal of Economic Perspectives* 25 (1): 159~184.
- Page, Marianne E. and Gary Solon(2003a). “Correlations between Sisters and Neighboring Girls in Their Subsequent Income as Adults.” *Journal of Applied Econometrics* 18 (5): 545~562.

_____(2003b). “Correlations between Brothers and Neighboring Boys in Their Adult Earnings: The Importance of Being Urban.” *Journal of Labor Economics* 21 (4): 831~855.

Solon, Gary, Marianne E. Page, and Greg J. Duncan(2000). “Correlations Between Neighboring Children in Their Subsequent Educational Attainment.” *The Review of Economics and Statistics* 82 (3): 383~392.

The Effect of Family Background on Children's Educational Outcomes in Korea : Estimation Using Siblings and School Data

Jung-Seung Yang

In this study, I try to overcome the omitted variable bias that always happens at the regression of educational outcome on observable specific family background and neighborhood variables. To deal with this problem, I estimate the effect of family background and neighborhood on child's educational outcomes with the college scholastic ability test score of siblings and provide the lower bound of it. The empirical results indicate that the effect of family background and neighborhood explains at least 20.2 percent of the variation of educational outcomes in Korea and 19.2 percent when restricted to the regions under leveling policy. When the high school graduated siblings are included to reduce the sample selection problems, the results show that the size increase to at least 23% in whole region and 20.6% in the region ruled by leveling policy.

Keywords : family background, college scholastic ability test score, neighborhood effect, School effect, educational equal opportunity