

Section I

연구논문

학생 잠재력인가? 부모 경제력인가?⁽¹⁾

김세직 · 류근관 · 손석준

본 논문은 출신 학교나 지역에 따라 학생들 간에 크게 차이가 나고 있는 대학 입학확률(예, 서울대 입학확률)이 학생들의 타고난 잠재력(또는 그를 반영한 진짜 인적 자본) 차이를 주로 반영하는 것인지, 아니면 학부모 경제력(또는 그를 반영한 치장법) 차이를 주로 반영하는 것인지를 통계 자료를 이용하여 평가하는 조그마한 실험적 시도이다. 이를 위하여, 본 논문은 서울시 구별 소득 통계자료, 소득과 잠재력의 상관관계, 부모 잠재력과 자녀 잠재력의 상관관계 등을 이용하여, 몇 가지 제한적 가정하에 서울시 구별 학생들의 잠재력(혹은 진짜 인적자본) 분포를 추정하고 이에 따른 대학 입학확률을 추정한다. 추정 결과에 따르면, 잠재력(혹은 진짜 인적자본)에 따라 대입확률이 결정되는 경우에, 유전적 요소로 인해 소득수준이 높은 지역(구)의 학생들이 낮은 지역의 학생들보다 대입확률이 높게 나온다. 즉 소득수준이 높은 지역에는 잠재력이 높은 학생들이 확률적으로 더 많기 때문에, 대입에서 잠재력 만에 의해 학생들이 평가되는 경우에도 소득수준이 높은 지역의 학생들의 대입 합격확률이 더 높게 얻어진다. 그러나, 본 논문의 추정 결과 타고난 잠재력 차이만으로 설명할 수 있는 서울대 입학확률의 구별 차이는 그 최대 배율 차이가 2배를 넘지 않는다. 이에 비해, 실제 관찰된 서울대 입학확률의 구별 차이는 최대 10~20배까지로 나타난다. 이는 첫째, 타고난 잠재력 차이만으로는 소득수준이 높은 구와 낮은 구 학생들의 서울대 입학확률 차이를 설명할 수 없다는 점, 둘째, 서울시내 구별 서울대 합격확률 차이의 8~9할 이상이 부모 경제력 차이에 따른 치장법 차이 등으로 설명될 수도 있다는 점 등을 시사한다. 동시에 현재의 대학입시 시스템이 잠재력이 높고 진짜 실력이 우수한 인재를 가려내는데 성공적이지 못할 수도 있음도 시사한다. 본 논문은 방법론상 한계점을 갖고 있는 작은 실험적 시도에 지나지 않는다. 향후 새로운 데이터와 방법론에 입각한 다양한 후속연구들이 이루어져 이 문제에 대한 보다 정밀하고 입체적인 이해가 요구된다.

주제어: 학생 잠재력, 부모 경제력, 교육제도, 경제성장, 효율성

(1) 이 논문에 계산된 수치들은 아직 잠정적(preliminary) 수치이니 인용 시 저자 허락을 요합니다. 본 논문에 대해 훌륭한 코멘트를 해 주신 서울대학교 교육종합연구원 교육포럼 세미나 및 경제학부 수요세미나 참석자 여러분들께 깊은 감사를 드립니다.

1. 머리말

본 논문은 현재의 우리나라 교육시스템하에서 출신학교에 따라 학생들간에 크게 벌어지고 있는 대학 입학확률(예, 서울대 입학확률)의 격차가 유전적 요소와 확률적 요소 등 두 가지 요소에 의해 결정된 학생들의 타고난 잠재력(또는 그를 반영한 진짜 인적자본)의 차이를 주로 반영하는 것인지, 아니면 학부모 경제력(또는 그를 반영한 치장법) 차이 등을 주로 반영하는 것인지를 서울시의 구별 통계 자료를 이용하여 평가하는 조그마한 실험적 연구이다.

교육제도, 특히 대학입시 제도가 어떠한가에 따라, 교육은 경제성장의 원동력이 될 수도, 반대로 경제성장의 걸림돌이 될 수도 있다. 입시제도가 학교/지역/계층에 관계 없이 우수한 인재를 제대로 가려내어 그들에게 보다 많은 자원과 재량권이 배분되도록 작동될 경우, 자본주의 경제체제의 생산성이 향상되고 자원배분의 효율성이 증대되어 경제성장이 촉진될 수 있다. 그러나 그렇지 못할 경우 교육과 입시제도는 오히려 경제 성장률을 하락시키는 요인으로 작용할 수도 있다. 이런 점에서, 현재 입시제도에 따른 학생들의 대학 입학확률이 학생들의 타고난 잠재력을 제대로 반영하는지를 평가하는 것은 우리나라 입시제도 자체의 평가를 넘어 경제의 효율성과 성장잠재력을 평가하는데 중요하다.

한국 경제는 과거 나름 효율적인 입시 및 교육제도에 힘입어, 1960년대 이후 40년간에 걸쳐 평균 8%대의 초고속성장을 구가할 수 있었다. 그러나 1990년대 중·후반 이후 한국 경제의 장기성장률은 지속적으로 추락하여 이제 2% 중반까지 추락하였고 지금의 추세라면 향후 추가적인 하락 가능성도 배제할 수 없는데, 이는 한국의 입시 및 교육 시스템이 오히려 성장의 걸림돌로 작용하고 있다는 시그널일 수 있다. 김세직(2014)은 데이터를 통해 한국 교육시스템의 효율성이 급격히 하락하고 있음을 보이고, 먼저 그 원인으로 창조형 인적자본 축적이 필수적인 시대에 아직도 모방형 인적자원 축적을 크게 벗어나지 못한 교육시스템을 들고 있다.⁽²⁾ 그와 함께, 교육의 인적자원 배분 시스템이 제대로 작동하지 못하는 문제점을 또 하나의 중요한 원인으로 지적하고 있다.

(2) 경제성장을 위한 창조형 인적자본과 교육의 역할에 관해 김세직, 류근관, 김진영, 박지형 외(2012), 김세직, 정운찬(2007) 참조. 한국경제의 고도성장과정에서의 교육의 역할과 교육제도의 기여도에 대해서는 Kim(1997) 참조.

장래 근로자가 될 학생의 ‘진짜 인적자본’은 무엇보다도 본인이 공부와 배움에 기울인 노력(시간)과 본인의 타고난 잠재력 등 두 가지 요소에 의해 결정될 것이다. 대학은 입시를 통해 지원자 학생의 인적자본을 제대로 측정하고자 노력하지만 이에 대한 정확한 평가가 어려워 부득이 수능성적, 스펙, 출신고등학교 생활기록부 등의 간접지표를 이용하게 된다. 그런데 이러한 간접지표는 ‘겉보기 인적자본’으로 학생들의 진짜 실력, 즉 진짜 인적자본과 차이가 날 수 있다. 특히 학생 자신의 진짜 인적자본 이상으로 겉보기 인적자본을 늘려주는 방법인 치장법(예, 사교육, 선행학습, 특수고 진학 등)에 의해서도 영향을 받게 되는데, 이 치장 능력은 부모 경제력이 높을수록 증가할 가능성이 높다. 이에 따라 부모 경제력 차이로 인해 치장법에 불리한 지원자가 본인은 진짜 우수함에도 입시에서 상대적으로 불이익을 당할 가능성을 배제하기 어렵다. 그렇다면 과연 이러한 가능성이 어느 정도 심각한 수준일까?

부모 경제력이 학생의 대학입시에 미치는 영향력과 관련하여 김세직(2014)이 제시한 몇 가지 한국의 데이터에 따르면, 실제로 학생들간 부모 경제력 차이가 서울대 진학 확률의 격차와 매우 밀접한 상관관계를 보이고 있다. 예컨대 2014년 기준 서울시내 구별 아파트 가격과 구별 일반고 학생의 서울대 합격확률간에 강한 선형 관계가 존재한다. 또한, 서울시내 구별 학원 개수로 근사한 사교육비 지출과 구별 일반고 학생의 서울대 합격확률 간에도 강한 상관관계가 존재한다. 더해서, 부모 경제력이 앞선 학생들이 그렇지 못한 학생들에 비해 이른바 명문대 합격확률이 훨씬 높게 나타난다. 2014년 기준 서울시내 구별로 일반고 학생의 서울대 합격확률을 비교해 보면, 강북구, 금천구, 구로구는 0.1~0.2%인데 비해, 소득수준이 높은 강남구는 2%로 위 3구의 10~20배에 이른다. 나아가 일반고와 보다 높은 부모의 경제력이 요구되는 특수고를 비교해보면, 서울시 일반고의 서울대 합격확률은 100명당 0.6명인데 비해, 서울시 외고와 과학고의 경우 그 확률은 100명당 10명과 41명으로 각각 서울시 일반고의 15배와 65배에 이른다.⁽³⁾

물론 이러한 지역/계층/학교간 학생의 대입확률의 커다란 차이는 학생들간 치장법 차이가 아니라 진짜 능력의 차이를 반영한 것일 수도 있다. 예컨대 부유한 지역 학생들의 대입확률이 훨씬 높은 것이 이 지역 학생들이 부모님들의 우수한 능력을 유전적

(3) 물론 일반고 학생과 특수고 학생을 단순 비교할 수는 없다. 김희삼(2015)은 부모 경제력이 학생들의 교육성파에 큰 영향을 미치는 현재의 교육시스템이 최근 우리나라 사회의 이동성 약화 현상을 가져왔음을 다양한 데이터를 이용하여 보여주고 있다.

으로 물려받아 훨씬 우수한 능력을 타고 났기 때문일 가능성을 배제할 수 없다. 고소득 지역 학생들의 부모님들은 타고난 능력이 뛰어나서 소득이 높게 되었고(그래서 고소득 지역에 살 게 되었고), 그 자녀들은 그 부모님들의 뛰어난 능력을 물려받아 역시 타고난 잠재력이 뛰어날 수 있다. 관측된 대입확률의 차이가 진짜 능력의 차이를 제대로 반영하고 있다면, 지역/계층/학교 간 대입확률에 커다란 차이가 존재한다는 것이 오히려 현재의 교육제도가 효율적으로 자원배분 기능을 담당하고 있음을 의미할 수도 있다.

관건은 부모소득과 그 자녀의 대학 합격확률 간 관측된 아주 강한 관계 중 얼마만큼이 유전적 영향에 따른 타고난 잠재력 차이로 설명될 수 있느냐는 것이다. 유전적 영향에 따른 진짜 능력의 차이가 과연 부모 경제력과 자녀 합격확률 간 관측된 관계의 대부분 또는 상당부분을 설명할 수 있을 정도로 클 것인가?

본 논문은 지역간 대입확률 격차, 보다 구체적으로 서울시 구별 일반고 출신학생들의 서울대 입학 확률의 격차가 유전적 요소에 영향 받아 결정된 학생들의 타고난 잠재력(혹은 진짜 인적자본) 차이를 주로 반영하는 것인지, 아니면 부모 경제력 차이에 따른 치장능력 차이를 주로 반영하는 것인지를 추정해보는 실험적 분석이다. 데이터에서 관측되고 있는 서울시 구별 일반고 서울대 합격확률의 차이 중 과연 얼마가 구별 학생들의 잠재력 또는 진짜 인적자본으로 설명되고 얼마가 부모 경제력에 따른 치장법에 의해 설명되느냐는 것이다.

이러한 분석을 위해, 본 논문은 서울시 구별 학부모 소득 차이로 인해 발생하는 구별 학생들의 타고난 잠재력(또는 그를 반영한 진짜 인적자본) 차이를 소득과 지능의 상관관계, 부모지능과 자녀지능의 상관관계를 바탕으로 연역적으로 추론해보고자 한다. 서울시 구별 학생들의 진짜 인적자본이나 잠재력 혹은 ‘지능’에 대한 직접적인 자료가 존재하지 않기에, 본 논문은 이를 다음과 같은 방법으로 간접적으로 추정한다. 우선, 서울특별시가 제공하는 서울시 각 구의 소득분포에 대한 구간별 자료를 이용하여 각 구의 연속적인 소득분포를 추정한다. 다음으로, 추정된 각 구의 소득분포에 기존 연구들에서 얻어진 소득-지능의 상관관계 수치를 결합하여 각 구의 학부모 지능 분포를 추정한다. 마지막으로 이렇게 구한 학부모지능분포와 기존 논문들에서 연구된 부모와 자식 지능의 상관관계를 이용하여 각 구의 학생지능(잠재력) 분포를 도출한다. 진짜 인적자본은 학생의 타고난 능력과 공부에 들인 노력(시간)에 의해 결정되며, 타고난 능력이 동일한 학생들은 똑같은 노력(시간)을 대학 입시 공부에 투자한다

고 가정한다. 이러한 가정 하에서는, 앞에서 추정한 각 구의 학생지능(잠재력) 분포가 구별로 학생들의 진짜 인적자본 분포가 된다.

구별로 추정된 학생 잠재력 혹은 진짜 인적자본 분포를 서울시내 전체 구에 걸쳐 종합하여 예컨대 서울시내 상위 0.5% 안에 드는 학생들(또는 서울대 합격생들)의 구별 분포를 구한다. 이렇게 구한 각 구별 진짜 인적자본 기준 서울대 입학생 분포와 실제 서울대 입학생의 구별 분포를 비교하여 그 차이를 구한다. 좀 더 구체적으로는 서울시내 구별로 해당 구의 일반고 학생 중 진짜 인적자본 기준 서울대 합격생의 비율을 이제껏 설명한 모형을 이용하여 구하고 이를 실제의 구별 서울대 합격생 비율과 비교한다. 그 차이가 매우 작다면, 실제 관찰되는 구별 서울대 합격확률의 차이는 유전에 따른 타고난 능력 차이로 주로 설명될 수 있다. 그러나 그 차이가 매우 크다면, 실제 지역별 서울대 합격확률 차이를 결정하는 주 요인은 진짜 인적자본 차이가 아니라 부모 경제력에 따른 치장법 등의 차이라고 볼 수 있을 것이다.

본 논문의 이러한 실험적 추정에 따르면, 잠재력(혹은 진짜 인적자본)에 따라 대입 확률이 결정되는 경우에, 유전적 요소로 인해 소득수준이 높은 지역(구)의 학생들이 낮은 지역의 학생들보다 대입 확률이 더 높게 나왔다. 즉 소득수준이 높은 지역에는 잠재력이 높은 학생들이 확률적으로 더 많기 때문에, 대입에서 잠재력 만에 의해 학생들이 평가되는 경우에도 이들 지역의 대입 합격확률은 소득수준이 낮은 지역의 대입 합격확률보다 높게 계산된다.

그러나 모형에 따라 진짜 인적자본 기준 서울대 합격확률의 구별 차이를 구해 본 결과 그 차이는 최대 2배가 채 나지 않았다. 구체적으로 부모 소득 수준이 높은 강남구 일반고의 진짜 인적자본 기준 서울대 합격확률은 0.84%로 얻어져 부모 소득이 상대적으로 낮은 강북구 일반고 0.5%의 1.7배에 불과하였다. 반면 실제 관측된 자료에서 서울대 합격률의 구별 차이는 강남구 일반고가 2.1%로 강북구 일반고 0.1%의 20 배씩이나 된다. 모형과 실제 자료 간의 이러한 괴리는 서울시 합격확률 최대구와 최소구 간 차이의 9할 이상이 타고난 능력 차이가 아니라 부모 경제력 차이에 따른 치장법 차이 등으로 설명될 수도 있음을 시사한다. 그렇다면 현재의 대학입시 시스템은 잠재력이 높고 진짜 실력이 우수한 인재를 가려내는데 그다지 성공적이지 못할 수도 있음을 시사한다.

본 논문의 이러한 실험적 추정은 연구를 위해 필요한 데이터들의 부족을 해결하기 위해 몇 가지 제한적 가정들을 하였다. 이러한 데이터의 제한과 방법론 상의 한계로

인해 본 논문의 추정 결과는 새로운 데이터와 방법론을 이용하는 경우 그 결과가 달라질 가능성을 배제할 수 없다. 향후 새로운 데이터와 방법론에 입각한 다양한 후속 연구들이 이루어져 이 문제에 대한 보다 정밀하고 입체적인 분석이 이루어지기를 기대한다.

2. 진짜 인적자본 분포 기본 모형

이 장에서는 학생들의 진짜 인적자본 분포 및 이에 따른 각 구별 가상 합격확률을 시험적으로 추정하기 위한 기본적 모형을 제시한다. 이 기본적 모형은 분석을 위해 몇 가지 가정들을 한다.

먼저, Lucas(1988)에 따라, 학생들의 진짜 인적자본은 학생들의 타고난 능력(잠재력)과 학생들이 인적자본을 늘리기 위해 투입한 노력(시간) 두 가지에 의해서 결정된다고 가정한다. 그리고 공부에 투입하는 시간(노력)은 부모 경제력과 관계없다고 가정한다. 따라서 타고난 능력이 동일하다면, 부유하지 않은 가정에서 태어난 학생들도 부유한 가정에서 태어난 학생들과 동일하게 열심히 공부한다고 가정한다. 그에 따라, 학생들의 진짜 인적자본은 기본적으로 학생들의 타고난 능력(잠재력)에 의해서 결정된다.

둘째, 타고난 능력(잠재력)은 지능으로 측정되며, 학생들의 지능과 학부모 지능은 유전적 요소와 확률적 요소 두 가지 요소에 의해 상관관계가 결정된다고 가정한다. 학생들의 지능과 학부모 지능은 지능에 관계하는 유전자의 공통점으로 인하여 지능의 상관관계가 발생한다. 그러나, 학생과 부모의 지능은 둘에 공통적인 유전적 요소 외에 각각에 랜덤(random)으로 독립적으로 작용하는 요소에 의해서도 영향 받는다. 유전적 요소 때문에, 부모와 자녀 지능이 동일해지는 경향을 보일 수 있으나, 랜덤 요소 때문에 부모의 지능과 자녀의 지능이 완전히 동일해지지 않으며 차이가 날 수 있게 된다. 이 두 요소 중 어느 쪽의 영향력이 크냐에 따라 학생 지능과 부모 지능 사이의 상관관계가 결정된다. 지능이 유전적 요소에 의해서만 결정될 때, 학생 지능과 부모 지능 사이의 상관관계를 나타내는 상관계수는 1이 되어 완전상관을 나타낸다. 이에 비해, 지능이 랜덤 요소에 의해서만 결정될 때, 상관계수는 0이 되어 둘 사이에 아무 상관관계도 존재하지 않게 된다.

셋째, 학부모 지능과 학부모 소득 사이에 상관관계가 존재한다고 상정한다. 학부모

지능이 높을수록 학부모 소득이 높아질 가능성이 커지리라 예상된다. 학부모의 소득은 자신의 타고난 지능(또는 진짜 인적자본)외에도 랜덤 요소에 의해서도 영향 받아 결정된다고 가정한다. 학부모의 소득이 두 요소 중, 지능(또는 진짜 인적자본)에 의해 크게 영향 받을수록 소득과 지능의 상관관계는 높아질 것이고, 소득이 랜덤 요소에 의해 크게 영향을 받을수록, 소득과 지능의 상관관계는 낮아질 것이다.

넷째, 잠재력(지능)에 영향을 미치는 랜덤 요소와 소득에 영향을 미치는 랜덤 요소들은 서로 독립이며, 유전적 요소와도 독립이라고 가정한다. 따라서, 부모의 소득은 자녀의 타고난 잠재력(그에 따른 진짜 인적자본)에 유전적 요소를 통해서만 영향을 미친다.

다섯째, 거주 지역과 학부모 지능 혹은 자녀 지능에 영향을 미치는 랜덤 요소는 관련이 없다고 상정한다. 즉, 어느 지역에 사느냐에 따라 사람의 지능이 달라지지 않는다는 것이다. 따라서, 거주 지역에 따라 학부모 소득의 분포는 달라지지만, 거주 지역이 학부모 지능 혹은 자녀 지능에 추가적인 영향을 주지는 않는다. 달리 표현하면, 거주 지역은 오로지 학부모 소득만을 통해서 학부모 지능 혹은 자녀 지능에 관계된다.

이상의 학생 지능과 학부모 지능 사이의 상관관계에 대한 가정과 학부모 지능과 학부모 소득과의 상관관계에 대한 가정 등으로부터 우리는 학생 지능과 학부모 소득 사이의 상관관계를 도출할 수 있다. 특히 위와 같은 가정하에서, 학부모 소득과 학생 자녀 지능 사이의 상관관계를 상관계수로 표현하면, 학생 지능과 학부모 지능 사이의 상관계수와 학부모 지능과 학부모 소득과의 상관계수의 곱으로 표현될 수 있다.

이러한 기본적 분석 모형은 수식을 이용하여 추가적인 가정과 함께 다양한 형태의 수식적 모형으로 표현될 수 있는데, 본 논문에서는 아래와 같은 수식 모형을 시험적으로 이용한다.⁽⁴⁾ 이 수식적 모형에서 우리는 서울시 구 i 에서 임의로 추출된 학부모의 로그소득을 확률변수 I_i 로 표현하고,⁽⁵⁾ 이러한 서울시 각 구의 학부모 로그소득 분포가 서울시 학부모 로그소득 분포와 전국 학부모 로그소득 분포와 다음과 같은 관계를 갖는 것으로 상정한다. 서울시 전체의 학부모 로그소득을 I_s , 모집단(전국)의 학부모 로그소득은 I 로 표현하고, 학부모 소득은 모든 지역에서 로그노멀 분포를 따른다

(4) 본 논문 모형과 세부 가정들을 약간 달리하는 수식모형의 예로 김세직·류근관·손석준(2015) 참조.

(5) 여기서 모형 설정의 편의상 어머니는 고려하지 않고 아버지만 계산한다고 가정한다. 기존의 지능 관련 문헌에 따르면 이 가정은 계산 결과에 큰 영향을 주지 않는다.

고 가정한다. 따라서 로그소득은 정규분포를 따르게 된다. 이 때, 임의의 소득 구간 A에 대해, I_i 와 I_S 의 관계는 다음과 같은 식으로 나타낼 수 있다.

$$(2.1) \quad \Pr(I_S \in A) = \sum_{i=1}^{25} \Pr(\text{구} = i) \Pr(I_i \in A)$$

즉, 서울시 구별 학부모 소득의 가중확률밀도를 쌓아 올리면 서울시 전체 학부모 소득의 확률밀도가 된다.⁽⁶⁾

서울시 전체 학부모 로그소득의 확률분포는 평균을 μ_{Z_S} , 표준편차 1로 표준화한다. 평균을 0이 아닌 μ_{Z_S} 로 표준화하는 것은 서울시의 학부모들이 전국(전체모집단) 평균보다 더욱 부유할 가능성을 포착하기 위해서이다.⁽⁷⁾ 이에 따라 표준화된 서울시 각 구의 학부모 로그소득분포 확률변수는

$$(2.2) \quad Z_i = \frac{I_i - \mu_I}{\sigma_I} + \mu_{Z_S}$$

가 된다. 여기에서 μ_I, σ_I 는 전국 학부모 로그소득 평균 및 표준편차를 나타내고, 위의 가정에 따라 σ_I 는 서울시 학부모 로그소득 표준편차와 동일하다.⁽⁸⁾

위와 같이 설정한 서울시 각 구의 학부모 로그소득 분포에 소득-지능 상관계수를

(6) 일반적으로 정규분포의 밀도의 가중 평균은 정확히 정규분포의 밀도가 되지는 않는다. 하지만 자료에서 보여지는 로그소득이 정규분포를 따른다고 보는 것이 이미 근사라는 측면에서, 위 식과 같은 근사가 성립한다고 보아도 무리는 없다.

(7) 이에 따라 본 논문의 분석에서 서울시 부모지능 분포 평균이 전국 평균보다 더욱 높은 계산 결과가 나오고, 따라서 아이들 지능 평균도 전국 평균보다 높게 계산된다. 그러므로 이를 바탕으로 추정된 서울시의 서울대 입학확률 (혹은 전국 상위 0.5%일 확률) 또한 전국 평균보다 높아지게 된다. 그리고 본 논문은 이렇게 추정된 지능을 바탕으로 계산한 서울시의 ‘서울대 진짜 인적자본기준 입학확률’이 실제 서울시의 서울대 입학확률과 같게끔 μ_{Z_S} 를 설정한다. 향후 전국 학부모 월평균 소득에 대한 서베이 등이 이루어져 그 데이터가 공개된다면, 그 자료를 이용하여 μ_{Z_S} 를 보다 직접적으로 계산할 수도 있을 것이다.

(8) 마찬가지로 방법으로 표준화한 서울시 로그소득분포 확률변수는 Z_S 라고 하자. 이렇게 표준화하면

$$(2.3) \quad \Pr(Z_S \in A) = \sum_{i=1}^{25} \Pr(\text{구} = i) \Pr(Z_i \in A)$$

가 성립하여 위의 식이 그대로 성립하게 된다.

결합하면 각 구 학부모들의 지능(또는 진짜 인적자본)의 분포가 추정될 수 있다. 이제, X , X_S 와 X_i 를 각각 전체모집단(전국), 서울시, 그리고 구 i 의 학부모 지능분포를 나타내는 확률변수라고 하자. 그리고 위에서 본 대로 Z , Z_S 와 Z_i 는 각각 전체모집단, 서울시, 그리고 구 i 의 표준화된 로그소득을 나타내는 확률변수이다. Z 와 마찬가지로 X 또한 $N(0, 1)$ 을 따른다고 가정한다.

우리는 전체모집단의 표준화된 로그소득과 지능 (Z, X)가 2변량 정규분포를 따른다고 가정한다. 이 경우, 해당 분포는 상관계수 ρ_2 를 통해 완전히 특정(specify)된다. 즉,

$$(2.4) \quad \begin{pmatrix} Z \\ X \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho_2 \\ \rho_2 & 1 \end{pmatrix}\right)$$

를 따르게 된다. 그리고 이변량 정규분포의 성질을 이용하면, $X|Z = z$ 의 분포는 $N(\rho_2 z, 1 - \rho_2^2)$ 가 된다. 다르게 표현하면,

$$(2.5) \quad X = \rho_2 Z + \xi,$$

$$(2.6) \quad \xi \sim N(0, 1 - \rho_2^2), \quad Z \perp \xi$$

가 된다. 이는 학부모의 소득은 학부모의 타고난 능력(또는 이에 따른 진짜 인적자본)과 랜덤 요소 두 가지 요소에 의해 결정되며 두 가지 요소는 서로 독립임을 의미한다. 그리고 학부모 지능과 학부모 소득의 상관관계는 학부모의 능력 요소가 커질수록 높아지고, 반대로 랜덤요소가 커질수록 낮아짐을 의미한다.

이 식을 이용하면, 각 구 i 의 학부모 지능분포 X_i 를 도출할 수 있다. 어느 구에 거주할지의 선택이 소득으로 조정된 후에 부모지능분포와는 관련이 없다는 가정 하에서, 위의 식을 사용하면

$$(2.7) \quad \begin{aligned} X_i &= \rho_2 Z_i + \xi, \\ \xi &\sim N(0, 1 - \rho_2^2), \quad Z_i \perp \xi \end{aligned}$$

으로 나타낼 수 있다. 실제 데이터로부터 구한 각 구 i 의 표준화된 로그소득 Z_i 들의

분포를 이용하여 각 구 i 의 학부모 지능 X_i 의 분포를 위 관계식을 통해 구할 수 있게 된다.

이렇게 구한 학부모들의 지능 분포에 이제 부모지능-자식지능 상관관계를 결합하면 학생들의 진짜 인적자본 분포를 구할 수 있다. 학생들의 지능(잠재력)을 Y 라 하고, (X, Y) 가 X 와 Y 각각의 평균은 0, 분산은 1인 2변량 정규분포를 따른다고 가정한다. 이 경우 상관계수가 ρ_1 일 때 (X, Y) 의 분포는 상관계수 ρ_1 에 의해 완전히 특정(specify)된다.

이 때 $Y|X = x$ 의 분포는 $N(\rho_1 x, 1 - \rho_1^2)$ 가 된다. 다르게 표현하면 다음과 같다.

$$(2.8) \quad \begin{aligned} Y &= \rho_1 X + \epsilon, \\ \epsilon &\sim N(0, 1 - \rho_1^2), X \perp \epsilon \end{aligned}$$

이는 학생의 잠재력(지능) 혹은 이에 따른 진짜 인적자본은 학부모의 지능(잠재력)의 유전적 요소와 랜덤 요소 두 가지 요소에 의해 결정되며 두 가지 요소는 서로 독립임을 의미한다. 그리고 유전적 요소가 커질수록 학생의 잠재력(진짜 인적자본)과 학부모의 잠재력과의 상관관계는 높아지고, 반대로 랜덤요소가 커질수록 그 상관관계는 낮아짐을 의미한다.

어떤 지능(잠재력)의 아이를 낳는지의 여부가 부모지능 정보를 알고 나면 어떤 구에 사는지 여부에 대해서는 영향을 받지 않는다는 가정 하에서, 구별 학생 진짜 인적자본 Y_i 은 다음을 만족시킨다.

$$(2.9) \quad \begin{aligned} Y_i &= \rho_1 X_i + \epsilon \\ \epsilon &\sim N(0, 1 - \rho_1^2), X_i \perp \epsilon \end{aligned}$$

여기서 우리는 학부모 지능(잠재력)을 결정하는 랜덤요소와 학생의 진짜 인적자본(지능)을 결정하는 랜덤요소(ζ, ϵ)끼리 서로 독립이고, 각각 표준로그소득분포 Z_i 와 독립이라고 가정한다.

이제 X_i 의 분포가 $X_i = \rho_2 Z_i + \zeta$, $Z_i \sim N(\mu_{Z_i}, \sigma_{Z_i}^2)$, $\zeta \sim N(0, 1 - \rho_2^2)$ 을 따른다고 위에서 밝혔으므로 이를 위 식에 대입하면, 서울시 각 구 i 의 학생들의 지능, 잠재력(또는 진짜 인적자본) Y_i 의 분포는 결국 다음과 같이 나타낼 수 있다.

$$(2.10) \quad Y_i = \rho_1 \rho_2 Z_i + \rho_1 \xi + \epsilon$$

$$Z_i \sim N(\mu_{Z_i}, \sigma_{Z_i}^2), \xi \sim N(0, 1 - \rho_2^2), \epsilon \sim N(0, 1 - \rho_1^2)$$

$$\xi \perp Z_i, \quad \epsilon \perp (\xi, Z_i)$$

가 된다.

이 식이 의미하는 바는, 지능을 진짜 인적자본의 대용물(proxy)로 사용했을 때에, 학생들 지능과 학부모 지능 사이의 상관계수 및 학부모 지능과 학부모 소득 과의 상관계수, 그리고 학부모 소득분포를 알면 그로부터 학생의 진짜 인적자본 분포를 도출할 수 있다는 것이다. 특히, 학부모 소득과 학생자녀 진짜 인적자본의 상관계수가 학생 지능과 학부모 지능 사이의 상관계수(ρ_1)와 학부모 지능과 학부모 소득 과의 상관계수(ρ_2)의 곱으로 표현될 수 있음을 알 수 있다. 따라서, 두 상관계수 ρ_1 과 ρ_2 의 크기가 작으면 작을수록 서울시 구별 진짜 인적자본의 분포는 모집단의 분포로 회귀하므로 구별 진짜 인적자본의 차이는 줄어든다.⁽⁹⁾

3. 진짜 인적자본 추정

이번 장에서는 위의 기본 모형을 이용하여 실제 서울시 구별 진짜인적자본 분포를 시험적으로 추정한다. 위의 기본 모형의 추정을 위해서는 학생들 지능과 학부모 지능 사이의 상관계수 및 학부모 지능과 학부모 소득 과의 상관계수, 그리고 학부모 소득 분포에 대한 정보가 필요하다. 이 장에서는 이들 값을 이용하여 서울시 구별 학생들의 진짜 인적자본 분포를 추정한다.

우리는 먼저 서울시 각 구의 학부모 소득 분포를 추정한다. 이를 위해, 서울특별시 행정국 정보공개정책과에서 제공하는 ‘2013년 서울서베이 도시정책지표조사’의 가구조사 자료를 이용한다. 이 자료는 서울시 가구 2만 가구를 대상으로 서베이를 한 것으로, 각 구의 월평균 소득(임대료, 이자수입, 연금 등을 포함 가족구성원의 모든 수입을 합한 소득), 거주하고 있는 자치구, 그리고 초·중·고에 다니는 자녀가 있는지 여부 등에 대한 정보를 제공하고 있다.

(9) 만약 학생들의 구별 ‘지능’ 분포가 아닌 ‘창조성’ 분포가 관심의 대상이라면, 더욱 교란(confounding)하는 요소가 늘어나기에 적절한 가정 하에서 구별 차이는 더욱 줄어든다. 이에 관련한 자세한 논의는 5장을 참조하시오.

우리는 각 구별 소득분포를 구함에 있어서, 초·중·고에 다니는 자녀가 있는 가구 즉 ‘학부모 가구’들로 한정하여 월평균 소득을 구한다. 다만, 이 서베이의 소득 자료는 구간 자료로, 각 가구의 월평균 소득이 0~50만원, 50~100만원, ..., 650~700만원, 그리고 700만원 이상 중 어디에 속하는지만 제시하는 불완전한 소득 자료이다.⁽¹⁰⁾ 즉, 각 구간 내에서 소득이 어떻게 분포하고 있는지는 제시하고 있지 않다. 또한, 그래프를 보면 $[100n, 100n + 50]$ 구간은 과대 리포트되고 $[100n + 50, 100n + 100]$ 구간은 과소 리포트 되는 것 같은 경향성을 확인할 수 있다. 또한 700만원 이상인 소득은 모두 한 구간으로 처리되어 있기에 이 넓은 구간에 소득이 실제로 어떻게 분포하고 있는지를 제시해 주지 않고 있다.

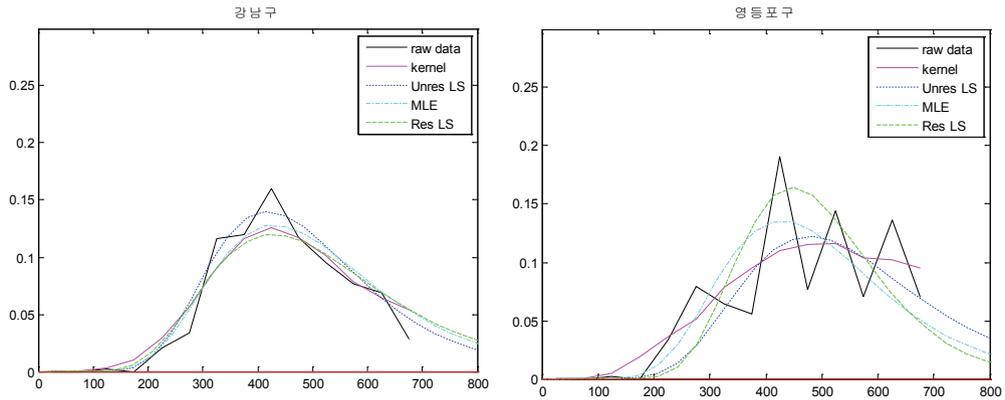
우리는 서울서베이 자료가 갖는 이러한 문제점들을 극복하도록 각 구별 학부모 소득분포를 추정한다. 추정을 위해 우리는 각 구의 소득분포가 로그노멀 분포를 따른다고 가정한다. 그리고 서베이의 소득 데이터를 바탕으로 ‘가장 가까운’ 로그노멀 분포를 찾아 이것이 실제 소득분포라고 가정하고 이를 추정한다.

실제 소득데이터와 ‘가장 가까운’ 로그노멀 분포를 추정하는 방법으로는 기존 연구들에서 많이 쓰이는 방법인 데이터를 바탕으로 작성된 히스토그램과 모형 확률밀도함수사이의 거리의 제곱을 적분한 것을 최소화하는 방법인 무제약최소제곱추정(unrestricted least squares)의 개념을 이용하되, 맨 오른쪽 구간, 즉 700만원 이상인 확률을 실제 데이터와 완전히 맞추도록 제약을 준 상태에서 최소화하는 방법인 제약최소제곱추정(restricted least squares)을 이용하였다.⁽¹¹⁾

이런 방법을 통해 우리는 각 구별 학부모 소득 분포를 추정하였다. <그림 1>은 이렇게 구해진 구별 소득분포의 예로서 강남구와 영등포구의 추정 소득 분포를 나타낸다.

(10) 각각 소득 구간의 중간값을 기준으로 (즉, [100만, 150만]원 구간은 125만원으로) 해당 구간에 위치하는 학부모들의 비율을 히스토그램으로 나타내면 부록의 <그림 4>와 같다. 단, 700만원 이상에 해당하는 사람들의 비율은 임의로 800만원에 표시해 놓았다.

(11) 본 논문은 이 방법 외에 아무 제약 없이 최소화를 통해 추정하는 방법인 무제약최소제곱추정(unrestricted least squares) 방법과 최대가능도 추정(maximum likelihood estimation) 방법을 이용해서도 추정해 보았다. 각각의 방법에 따른 추정 결과를 비교해 본 결과, 어느 방법을 사용하든 큰 차이는 없지만 본 장에서 택한 방법이 가장 적절하다고 판단되었다. 각각의 방법으로 추정해본 각 구의 2013년 학부모 가구 평균소득과 2008년 전체 가구 평균소득(서울시 비공식 자료)이 가지는 상관관계를 비교해 본 결과 우리가 택한 방법이 2008년 (전체 가구) 소득을 보다 잘 예측한다. 또한 우리가 택한 방법으로 추정하였을 시, 영등포구의 부자 비율이 과다추정되고 강남구의 부자 비율은 과소추정되는 경향을 상당히 완화시킬 수 있었다.



〈그림 1〉 강남구, 영등포구의 원자료 분포와 추정 분포의 확률밀도⁽¹²⁾

이렇게 각 구의 학부모 소득 분포가 구해진 상태에서, 학생들 지능과 학부모 지능 사이의 상관계수 및 학부모 지능과 학부모 소득과의 상관계수를 이용하여 학생들의 진짜 인적자본의 분포를 구할 수 있다. 우리는 학생들 지능과 학부모 지능 사이의 상관계수 및 학부모 지능과 학부모 소득과의 상관계수의 값에 대해 기존연구들의 추정치를 이용한다. 단, 기존 연구들의 추정치들이 다양함에 비추어, 우리는 매우 보수적인 추정치를 이용한다.

우리는 두 가지 상관계수 중 학생들 잠재력과 학부모 잠재력 사이의 상관계수로 부모 지능과 자녀 지능의 상관관계에 대한 기존 연구들의 추정치를 이용한다. 부모 지능과 자녀 지능의 상관관계에 대해서는 오래 전부터 많은 기존 연구가 있어 왔다. Bouchard(1981)가 기존 연구들을 정리한 바에 따르면, 부모와 자식 지능 간의 상관관계에 관해 기존 연구들에서 보고된 추정치가 36개가 존재한다. 보고된 상관관계 추정치 중 4개는 부모와 그 부모로부터 떨어져 양육된 자식 IQ 사이의 상관관계였으며, 이들 상관관계의 가중 평균은 0.22이었다. 그리고, 보고된 32개의 상관관계 추정치는 부모와 그 슬하에서 자란 자식 IQ 사이의 상관관계에 대한 것이었으며, 이 추정치들의 가중 평균은 0.42였다.

(12) 〈그림 1〉에서 raw data는 서울서베이에서 제공하는 데이터를 히스토그램으로 나타낸 것이며(700만원 이상 구간은 생략), ‘kernel’은 커널함수를 통해 해당 데이터를 부드럽게 한 것, ‘Unres LS’는 무제약최소제곱추정, ‘MLE’는 최대가능도추정, 그리고 ‘Res LS’는 본 장에서 채택한 제약최소제곱추정 방법에 따른 추정 분포를 나타낸다.

우리는 학생들 지능과 학부모 지능 사이의 상관계수로 Bouchard(1981)가 정리한 부모와 그 슬하에서 자란 자녀 지능 사이의 상관관계 가중평균인 0.42를 사용한다. 우리가 부모지능-자녀지능 상관관계로 이용한 0.42는 보수적인 숫자로 같이 생활한 부모-자식의 상관관계여서 유전적 요인에 더해 환경적인 영향까지 포함되어 있다고 볼 수 있다. 유전적 요인만 반영하기 위해, 따로 생활한 부모-자식의 상관관계 가중평균을 이용한다면 이 상관관계는 0.22에 불과하지만, 우리는 보다 보수적인 추정치를 이용하였다.

또 하나 필요한 상관계수인 학부모 지능과 학부모 소득과의 상관계수로 지능과 소득 간의 상관관계에 대한 기존 연구들의 추정치를 이용한다. 지능과 소득의 상관관계에 대해서도 여러 기존 연구가 있다. Zagorsky(2007)가 ‘National Longitudinal Survey of Youth 1997’ 자료를 활용하여 IQ와 33~40세 소득 사이의 상관관계를 추정할 바에 의하면 상관계수가 0.30이었다. Irwing and Lynn(2006)가 8세 때의 IQ와 43세 소득 사이의 관계를 분석한 바에 의하면, 상관계수가 남자의 경우 0.37, 여자의 경우 0.32였다. 한편, Jencks and others (1972)는 25~65세의 미국 백인 남성의 소득과 IQ 사이의 상관계수로 0.31을 보고하였다. 우리는 이러한 위 논문들에서 보고하는 값 중 제일 보수적인 숫자인 0.37을 소득과 IQ 사이의 상관관계로써 사용한다.⁽¹³⁾

우리는 서울시 소득 데이터를 이용해 구한 각 구별 학부모 소득분포와 학생 지능과 학부모 지능 사이의 상관계수에 대한 보수적 추정치 0.42 및 학부모 지능과 학부모 소득과의 상관계수에 대한 보수적 추정치 0.37를 결합하여 서울시 각 구별 학생의 진짜 인적자본 분포를 구하였다.

4. 학생 겉보기 인적자본 분포

입시를 통해 대학들은 학생들의 진짜 인적자본을 평가하려 많은 노력을 기울이지만, 진짜 인적자본은 직접 관찰하기 어렵다. 이에 여러 가지 간접지표들을 보는 데, 이러한 간접지표들은 겉보기 인적자본으로 진짜 인적자본과 차이가 날 수 있다.⁽¹⁴⁾

(13) 본 논문은 분석에 필요한 소득-지능 상관계수의 추정치를 주로 미국의 연구들에 따른 추정치를 이용하되 이에 따른 문제를 극복하기 위해 가장 보수적인 수치들을 이용하였다. 향후 한국에서의 소득과 지능의 상관관계에 대한 연구들이 이루어져, 한국에서의 상관계수가 구해진다면, 이를 이용하여 보다 정확한 추정이 가능해질 수 있을 것이다.

(14) 특히 부모 경제력에 의해 영향 받는 치장법(예, 사교육, 선행학습, 특수고 진학 여부)의 영

대학 입학담당자가 평가 혹은 관찰할 수 있는 각 학생들의 속성을 모두 암묵적으로나마 점수화하고 이 점수를 기준으로 합격 여부를 판가름한다면, 이 점수가 나타내는 것이 바로 겉보기 인적자본 점수이다. 가상적으로, 이 겉보기 인적자본 점수가 전국의 모든 수험생들에게 주어져 있다고 생각해보자. 각 대학은 학생들을 겉보기 인적자본에 따라 평가하고, 그 결과 학생들의 겉보기 인적자본 점수의 전국적 분포가 형성된다. 그리고 대학들은 이 점수의 ‘합격선’을 정하고 이를 넘긴 사람들을 입학시킨다.

전국적으로 학생들의 겉보기 인적자본 점수 S 의 분포는 표준정규분포 $N(0, 1)$ 을 따른다고 가정하자. 서울대의 경우는 전국의 학생들 중 약 0.5%만이 서울대로 입학하고, 따라서 $\Pr(S > 2.575) = 0.005$ 이다. 따라서, 위에서 말한 가정에 의하면 겉보기 인적자본 점수가 2.575를 넘는 학생들만이 서울대에 입학한다.

더해서, 서울시 각 구의 겉보기 인적자본 점수 또한 정규분포를 따른다고 가정한다. 다만, 평균과 표준편차는 구마다 다를 수 있는데, 겉보기 인적자본 점수 분포의 표준편차는 각 구 학생들의 진짜 인적자본 분포의 표준편차와 일치한다고 상정한다. 이러한 가정 하에, 각 구별로 겉보기 인적자본의 평균과 표준편차를 구하기 위해, 겉보기 인적자본 점수 2.575점을 넘기는 사람의 비율이 우리가 실제로 관측한 각 구별 서울 대학합격확률과 일치함을 이용한다. 실제로 대학들은 겉보기 인적자본에 입각하여 학생들을 평가하게 되는데, 따라서 각 구별 겉보기 인적자본 분포에서 합격선을 넘는 학생의 비율은 실제 대학에 합격한 각 구별 학생들의 비율, 즉 각 구별 합격확률과 같아야 한다. 이를 이용하여, 우리는 각 구별로 겉보기 인적자본의 평균과 표준편차를 구하고, 그 결과 서울시 구별 겉보기 인적자본의 분포를 추정하였다.

단, 여기서 각 구에 거주하는 학생들의 실제 합격확률을 계산함에 있어서, 각 구의 일반고 학생만의 합격확률을 이용하였다. 부득이 일반고 자료만 이용한 이유는 데이터의 부재로 인해 특수고(과학고, 외고, 자율고)에 다니는 학생들의 거주 지역에 대한 정보가 없기 때문이었다. 만약 특수고 학생들의 출신 구에 대한 자료를 구할 수 있다면, 이 정보까지 이용하여 서울시 구별 학생들의 대학입학 확률을 계산하는 것이 보다 정확할 것이다. 특수고에 다니는 학생들의 상당수가 강남구 등 소득수준이 높은

항 때문에, 동일한 진짜 인적자본을 갖고 있는 학생 간에도 겉보기 인적자본의 차이가 날 수도 있다. 만약 대학이 학생부종합전형 등의 평가에 있어서 학생들의 잠재력을 평가할 능력이 충분하지 못하다면, 불가피하게 스펙, 특수고 진학 여부 등에서 진짜 인적자본의 시그널을 찾게 될 가능성도 생기며, 이 경우에 부모 경제력에 따른 치장 능력 차이로 인해 오히려 겉보기 인적자본과 진짜 인적자본의 괴리가 더 커질 수도 있다.

구 출신임을 감안하면, 일반고와 특수고 학생을 모두 고려하여 계산한 구별 합격확률은 소득수준이 높은 구에서 상대적으로 더욱 높아질 가능성도 배제할 수 없다.

5. 평가되는 겉보기 인적자본과 진짜 인적자본의 괴리

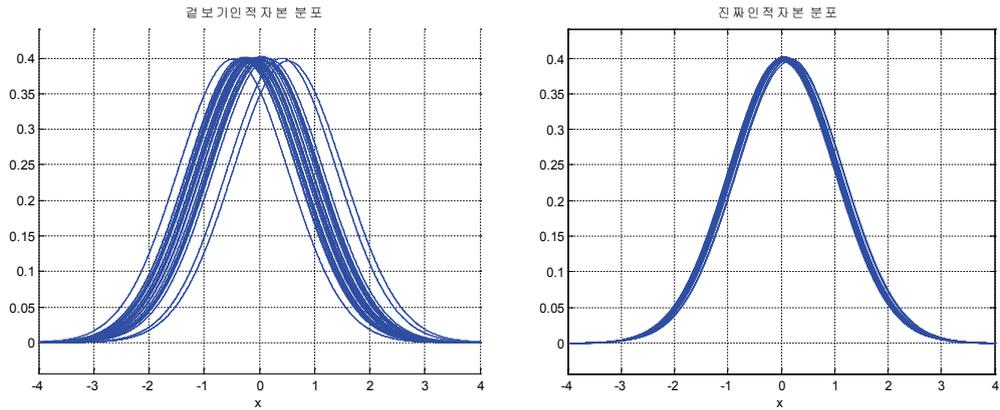
이 장에서는 앞에서 구한 학생들의 겉보기 인적자본 분포와 진짜 인적자본 분포를 비교한다. 특히, 겉보기 인적자본 분포에 따른 서울시 구별 서울대 합격확률의 차이가 학생들의 진짜 인적자본 분포의 차이로 어느 정도 설명이 될 수 있는지 실험적으로 분석한다.

앞에서 구한 각 구별 진짜 인적자본의 분포와 겉보기 인적자본의 분포를 나타낸 것이 <그림 2>이다. 이 그림에서 보여지듯이 겉보기 인적자본의 분포에 비해 진짜 인적자본의 분포는 서울시 구별로 큰 차이가 나지 않는다.

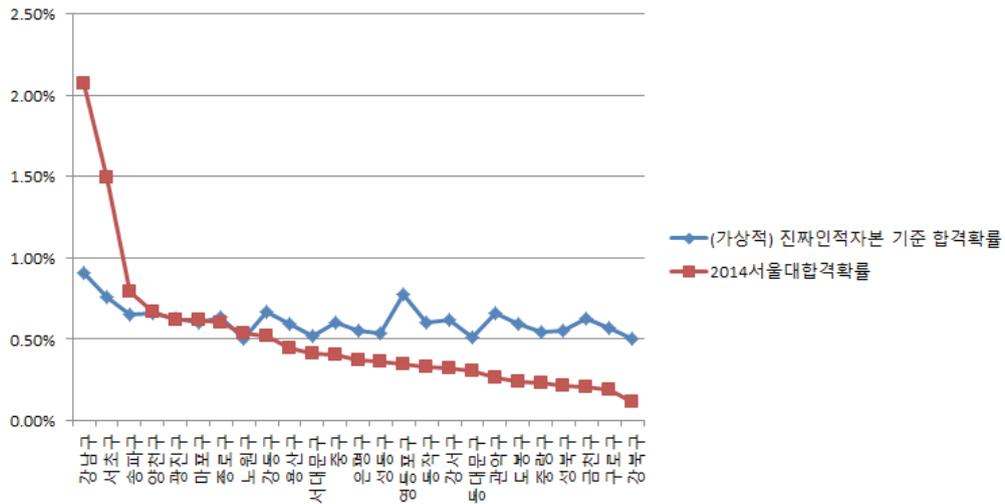
우리는 서울시 각 구별 학생들의 진짜 인적자본 분포에 입각하여 각 구별로 진짜 인적자본 기준 전국 상위 0.5% 안에 드는 학생들의 비율을 구하였다. 이는 대학입시가 진짜 인적자본을 정확히 가려낸다면 이루어질 가상의 각 구별 서울대 합격확률이다. 그리고 이를 2014년 각 구별 겉보기 인적자본 분포에 따른 전국 상위 0.5% 안에 드는 학생들의 비율, 즉 2014년의 각 구별 실제 서울대 합격확률과 비교한다. 2014년의 실제 서울대 합격자 자료는 중앙일보 2014년 2월 21일 기사 ‘수시 맥 찾은 일반고 웃었다’에서 제시한 최초합격자 자료를 사용하였다.⁽¹⁵⁾ 각 학교들의 2014년의 수험생 숫자는 서울교육통계의 연도별 학교현황에서 보고된 2011년의 입학생 숫자로 근사되었다.

이러한 본 논문의 실험적 추정결과는 <그림 3>에 요약되어 있다. 이를 통해, 학생들의 진짜 인적자본 분포로부터 도출되는 가상의 서울대 합격확률의 구별 차이가 대입에서 겉보기 인적자본 평가에 따른 실제 합격확률의 구별 차이보다 훨씬 작다는 것을 알 수 있다. 특히, 진짜 인적자본 분포로 예측하였을 때 최대 합격률을 보인 강남구와 최소 합격률을 보인 강북구의 진짜 인적자본 기준 서울대 합격률은 각각 0.84%, 0.50%로 1.7배 정도 차이에 불과한 반면, 2014년 입시에서 실제 합격확률은 두 지역 출신 일반고 학생만 고려해 보아도 각각 2.07%, 0.11%로 스무 배 정도로 차이가 난

(15) 기사에 따르면 이는 서울대가 국회 교육문화체육관광위원회 박인숙 의원에게 제출한 ‘2014 학년도 서울대 합격자 고교별 현황’ 자료를 입수해 분석한 결과이다.



〈그림 2〉 겉보기 인적자본과 진짜 인적자본 분포



〈그림 3〉 겉보기인적자본 기준 및 진짜인적자본 기준 서울대 합격확률

다. 이는, 실제 강남구와 강북구의 2014년 서울대 입학확률의 90% 이상이 부모 경제력에 따른 치장법 차이로 설명될 수 있음을 시사한다.

진짜 인적자본 기준 서울대 합격확률 구별 차이를, 강남구 대비 구별 차이의 절대값(즉 $|강남구 - 구|$)의 평균을 통해서 구하여 볼 수도 있는데, 이 경우 이 값은 0.22%이다. 이에 비해 2014년 겉보기 인적자본(서울대 입학확률로 표현된) 기준 구별 차이 절대값의 평균은 1.58%이다. 따라서, 2014년 구별 평균 서울대 입학확률 차

이 중 부모경제력에 따른 치장법 차이가 설명하는 부분이 86%에 달함을 알 수 있다.

이상의 실험적 추정 결과는 각 구별로 서울대 입학확률의 실제 차이가 나는 주요 요인이 유전에 따른 각 구별 타고난 지능(잠재력) 또는 진짜 인적자본의 차이 때문 보다는 부모 경제력 차이에 따른 치장법 차이 때문일 가능성이 보다 높음을 시사한다.⁽¹⁶⁾ 동시에 현재의 대학입시 시스템이 잠재력이 높고 진짜 실력이 우수한 인재를 가려내는데 성공적이지 못할 수도 있음도 시사한다.

단, 본 논문의 분석 결과는 입시나 소득, 잠재력 등과 관련된 데이터의 한계 속에서 제한된 방법론과 강한 가정 하에 도출되었기 이를 단정적으로 받아들이기보다 하나의 실험적인 결과로 보아야 한다. 즉, 향후 후속 연구에 의해 보완되어야 할 하나의 예비적인 결과(preliminary results)로 해석할 필요가 있다. 물론, 본 논문의 실험적 분석에 따른 추정치가 예비적인 수치이기는 하지만, 치장법의 역할이 매우 크게 나와 향후 새로운 데이터와 방법론을 이용하는 경우에도 그 구체적 수치는 변할지라도, 치장법이 매우 중요한 역할을 한다는 시사점은 크게 달라지지 않은 가능성도 배제할 수 없다.

6. 창의성에 따른 진짜 인적자본 분포

앞에서 살펴 본 우리 논문 분석의 기본적인 아이디어는 다음과 같다. 일단 학생들이 공부하는데 투자하는 시간은 부모 경제력과 관계없이 동일하다고 가정한다. 이 경우, 학생들의 진짜 인적자본은 학생들의 타고난 잠재력에 의해 결정된다. 학생들의 진짜 인적자본을 결정하는 타고난 잠재력은 유전의 영향도 받는다. 따라서, 부모의 경제력은 치장법 만이 아니라 유전을 통해 타고난 잠재력에도 영향을 미칠 수 있는데, 그 미치는 영향은 학부모 소득과 학생 잠재력 간의 상관관계를 통해 계량화한다.

부유한 학부모와 능력이 뛰어난 학생간의 상관관계는 부모의 타고난 능력과 소득

(16) 이에 대한 한가지 다른 설명 가능성은 동일한 잠재력을 갖고 있더라도 소득이 높은 가정의 자녀들이 공부하고자 하는 열망이나 인센티브가 더욱 커서 소득이 낮은 가정의 자녀들 보다 공부에 더 많은 시간을 투자하여, 그 결과 진짜 인적자본 자체가 더 커진다는 설명이다. 그러나 입시가 진짜 인적자본을 잘 가려낸다면, 공부에 더 많은 노력을 기울일 가능성은 오히려 부유하지 않은 가정의 학생들이 부유한 가정의 학생들보다 더 높을 수도 있다. 공부를 열심히 해 이른바 명문대학에 들어가고 졸업 후 고소득 직업에 매칭될 때의 한계이익이 부유하지 않은 가정의 학생들이 부유한 가정의 학생들보다 더 클 가능성이 높기 때문이다.

간의 상관관계와 부모 지능과 자녀 지능과의 상관관계를 이용하여 구한다. 잠재력(지능)에 영향을 미치는 랜덤 요소와 소득에 영향을 미치는 랜덤 요소들은 서로 독립이며, 유전적 요소와도 독립이다. 따라서, 부모의 소득은 자녀의 타고난 잠재력(그에 따른 진짜 인적자본)에 유전적 요소를 통해서만 영향을 미친다. 이러한 상황에서, 부모 소득과 학생 잠재력 상관계수는 학생 지능과 학부모 지능 사이의 상관계수와 학부모 지능과 학부모 소득과의 상관계수의 곱으로 표현된다.

타고난 능력이 뛰어난 부모가 소득이 높은지에 관해서는 소득-지능 상관관계에 관한 기존의 여러 연구들이 있다. 우리가 찾은 소득-지능 상관계수에 대한 기존 문헌들은 그 추정치로 0.3에서 0.37 사이의 값을 보고하고 있는데 우리는 보수적으로 0.37을 이용한다. 부모의 타고난 능력이 자녀에게 유전되는지에 관해서도 부모지능과 자녀지능의 상관관계에 대한 여러 연구들이 있다. 우리는 부모와 따로 산 자녀와 부모 사이 지능의 상관계수 추정치 대신, 보수적으로 부모와 같이 사는 자녀와 그 부모 지능 간의 상관계수의 가중평균인 0.42를 이용한다.

이 두 상관계수에 보수적 추정치를 이용하여 부유한 부모와 능력이 뛰어난 자식의 상관관계를 구하면

$$(6.1) \quad 0.37 \times 0.42 = 0.16$$

으로 구해질 수 있다. 이를 통해 알 수 있는 것은 부유한 부모와 능력이 뛰어난 자식의 상관관계는 1보다 작은 두 상관관계 0.37과 0.42를 곱하면서 값이 크게 작아져 0.16으로 그다지 크지 않은 값으로 희석(dilution)된다는 것이다. 이에 따라, 각 구별 학부모 소득에서 막대한 차이가 나지 않는 한, 학생들의 타고난 잠재력 혹은 진짜 인적자본이 구별로 커다란 차이가 나지 않게 된다. 이것이 진짜 인적자본으로 예상되는 서울시 강남구와 강북구의 서울대 입학확률이 겉보기 인적자본 차이에 따른 실제 서울대 입학확률에 비해 십 분의 일 수준밖에 안 되는 기본적인 이유이다.

이상의 2장의 기본적 모형에 따른 분석에서 우리는 학생들의 잠재력 분포를 지능 분포와 동일한 것으로 가정하고 지능-소득 상관관계와 부모-자식 지능 상관관계를 이용하여 논의를 전개하였다. 하지만 이러한 가정은 자료의 한계로 인하여 설정하였던 것이며, 실제로는 당연히 학생들의 지능이 곧 아이들의 잠재력 그 자체는 아닐 수 있다.

그리고 기술 변화에 따라 그리고 경제발전 단계에 따라, 요구되는 인적자본의 특성과 이에 따른 잠재력의 성격도 변한다. 특히 우리나라 근로자에게 현재 시대가 요구하고 있는 인적자본은 더 이상 모방형 인적자본이 아니라 창조형 인적자본이다. 따라서, 학생들에게 요구되는 진짜 인적자본의 성격이 창조형 인적자본에 가까워졌음을 감안하면, 유전의 영향력은 더욱 약화될 가능성이 높다.

학부모 세대가 모방형 인적자본이 요구되는 시대였고, 이에 따라 현재 고소득 학부모들은 모방형 인적자본을 축적하는 능력에 비교우위가 있었다면, 이러한 부모들의 자녀들은 유전을 통해 모방형 인적자본을 습득하는 능력을 물려받았을 가능성이 높을 수 있다. 현재 학생들 세대는 부모세대와 질적으로 다른 창조형 인적자본이 근로자로서의 능력, 즉 진짜 인적자본을 구성하는 주 요소가 되었다면, 따라서 소득이 높은 학부모를 둔 학생들의 진짜 인적자본이 높을 가능성은 더욱 줄어들게 된다.

이를 보이기 위해, 우리는 학생들의 잠재력에 있어서 ‘지능’이 아닌 ‘창의성’이 핵심이 된 경우, 진짜 인적자본의 분포를 이론적으로 서울시 각 구별로 구할 수 있다. 앞 장에서의 논의에서와 마찬가지로, 서울시 각 구 i 의 학생들의 창의성을 H_i 로 나타내고, 그 자체는 표준정규분포를 따르며, 창의성과 지능 (H_i, Y_i) 쌍은 2변량 정규분포를 따른다고 가정하자. 그리고 이 두 변수의 상관관계는 ρ_3 를 가진다고 상정하자. 그리고 어느 구에 거주하는지의 여부는, 부모의 소득이 주어지고 나면, 아이들의 창의성과는 상관이 없다고 가정하자. 이 경우, 구별 창의성 분포는

$$(6.2) \quad \begin{aligned} H_i &= \rho_3 Y_i + v \\ v &\sim N(0, 1 - \rho_3^2), \quad Y_i \perp v \end{aligned}$$

로 나타낼 수 있다. 그리고 부모 지능, 아이 지능을 결정하는 오차 항들인 ξ, ϵ 과 v 가 독립이라고 가정하자. 이 가정 하에서, 2장에서와 비슷한 계산을 통해 학생들의 창의성 H_i 의 분포가 다음과 같이 됨을 보일 수 있다.

$$(6.3) \quad H_i \sim N(\rho_1 \rho_2 \rho_3 \mu_{Z_i}, (\rho_1^2 \rho_2^2 \rho_3^2) \sigma_{Z_i}^2 + (1 - \rho_1^2 \rho_2^2 \rho_3^2))$$

여기서, 상관계수 ρ_3 가 추가됨으로써 거주하는 구 i 의 (소득과 관련된) 특성의 중요성이 더욱 떨어지고 거의 원래 H_i 의 분포인 표준정규분포로 회귀하는 모습을 볼 수 있

다. 즉, 부모 세대의 고소득을 위해 요구되는 능력(지능)과 아이 세대에게 요구되는 능력(창의성)의 상관관계가 약해지면 약해질수록 자녀세대 학생들에게 요구되는 진정한 인적자본의 분포는 거주 지역과는 거의 상관이 없어지는 것을 알 수 있다.

이를 실험적 수치로 보이기 위해, 학부모 세대와 자녀세대 간 요구되는 인적자본의 상관관계가 양수이기는 하지만 1보다 작은 값, 예를 들어 0.5라고 하자. 이 경우 부유한 부모와 능력이 뛰어난 자식의 상관관계는 결국

$$(6.4) \quad 0.37 \times 0.42 \times 0.5 = 0.08$$

으로 구해져 0.1도 안 되는 더욱 작은 값이 된다.

그 결과, 학생들의 진짜 인적자본 분포는 부모 경제력과 거의 상관이 없어질 것으로 예측 될 수 있다.⁽¹⁷⁾ 비유를 들자면, 창의적인 용의 씨는 학교, 지역, 계층 관계 없이 골고루 뿌려진다고 가정하는 편이 용의 씨가 높은 확률로 유전된다는 주장에 비해 데이터에 보다 부합될 가능성이 높다. 물론, 이러한 결과는 창의성에 대한 데이터와 기존 선행 연구의 부족이라는 한계 속에서 이루어진 하나의 예비적인 결과로 보아야 한다. 그럼에도 불구하고 이 장의 예비적인 결과들은 학생부종합전형을 중심으로 한 현 대학입시제도 하에서 일부 학교/지역/계층에 선발학생들이 심하게 쏠리는 현상들이 현 교육/입시제도가 창의적 인재 선발이라는 목표를 충분히 달성하지는 못하고 있다는 증거임을 시사할 수도 있다.

6. 맺음말

자본주의 경제의 효율성을 보장하고 경제성장을 촉진하기 위해서는 교육시스템 특히 입시시스템이 학생들의 잠재력을 반영한 진짜 인적자본을 제대로 평가하는 것이 바람직하다. 교육 혹은 대학 입시제도가 학생들의 진짜 인적자본을 적절히 평가하고 있다면 진짜 인적자본 분포에 기반하여 본 논문에서 도출한 것과 같은 대학 합격확률

(17) 소득이 높은 학부모의 인도/관리 하에 어려서부터 선행학습과 사교육 등 치장법에 많은 시간을 투자해 온 학생들은 타고난 자기주도성을 제대로 계발을 못해, 부모 도움 없이 자기 길을 개척해 온 학생들에 비해 창의성의 전제가 되는 자기주도성이 약하고 이에 따라 창의성이 오히려 더 낮을 가능성도 배제할 수 없다.

과 실제로 관측된 대학 입학확률간 차이가 크게 나지 않을 것으로 예상할 수 있다. 만약 모형으로부터 도출한 구별 대입 합격확률과 실제의 구별 대입 합격확률간 차이가 많이 나면, 이는 현행 대학입시가 진짜 인적자본을 충분히 잘 가려내지 못하고 있음을 시사할 수도 있다.

본 논문은 현재 교육/입시제도가 진짜 인적자본을 잘 가려내어 자원배분의 효율성 향상에 기여하는지 그렇지 못한지를 평가하기 위해, 서울시 구별 소득 통계 등의 데이터를 이용하여 구별 진짜 인적자본의 분포를 구하고 다시 이에 입각하여 서울시 내 구별로 서울대 합격 확률을 구하고 이를 실제로 관측된 구별 서울대 합격 확률과 비교하는 조그마한 실험적 분석을 하였다. 이러한 실험적 분석 결과는 서울시의 경우 소득수준이 높은 구와 낮은 구 학생들의 대학 합격확률 차이의 8~9할 이상이 타고난 잠재력 차이 보다는 부모 경제력에 따른 치장법 등의 차이로 설명될 수도 있음을 시사하고 있다. 예컨대 부모의 경제력에 따라 학생들의 치장법과 겉보기 인적자본이 달라질 수 있는데 현재 입시제도는 이러한 겉보기 인적자본과 진짜 인적자본을 충분히 구분하지 못하여 결과적으로 부모 경제력 차이에 따라 학생의 대학 입학확률 차이가 지나치게 커질 개연성을 배제하지 못한다. 부모 경제력에 따라 치장능력이 상당히 차이가 나는 상태에서 교육 및 입시제도가 겉보기 인적자본과 진짜 인적자본을 충분히 구분해 내지 못한다면, 이는 막대한 사회적 비효율성을 유발하여 장기적으로 경제 성장에 걸림돌이 될 수도 있기 때문에 향후 이에 대한 보다 면밀한 연구와 검토가 요구된다.

본 논문의 분석 결과는 입시나 소득, 잠재력 등과 관련된 데이터의 한계 속에서 제한된 방법론과 강한 가정 하에 도출되었기에 이를 단정적으로 받아들이기보다 하나의 실험적인 결과로 보아야 한다. 즉, 향후 후속 연구에 의해 보완되어야 할 하나의 예비적인 결과(preliminary results)로 해석할 필요가 있다. 본 논문은 서울시 내 구별 학생들의 입학확률 차이를 계산함에 있어서 일반고 자료만 대상으로 분석하였다. 특수고(과학고, 외고, 자율고) 학생들의 출신 구별 자료를 추가한다면 향후 구별 학생들의 대학입학 확률 차이를 보다 정확히 계산할 수 있을 것이다. 아울러 후속 연구를 통해 데이터의 한계를 극복하기 위한 다양한 방법론적 시도도 이어지기를 기대한다.⁽¹⁸⁾

(18) 김세직, 류근관, 손석준(2015) 참조.

김세직

서울대학교 경제학부 교수

151-746 서울특별시 관악구 관악로 1

전화: (02) 880-4020

팩스: (02) 886-4231

Email: skim@snu.ac.kr

류근관

서울대학교 경제학부 교수

151-746 서울특별시 관악구 관악로 1

전화: (02) 880-6397

팩스: (02) 886-4231

Email: ryu@snu.ac.kr

손석준

예일대 경제학과 박사과정

28 Hillhouse Avenue, New Haven, CT 06511 USA

전화: +1 (475) 209-0029

팩스: +1(203) 436-2626

Email: luisfigokr@gmail.com

참고문헌

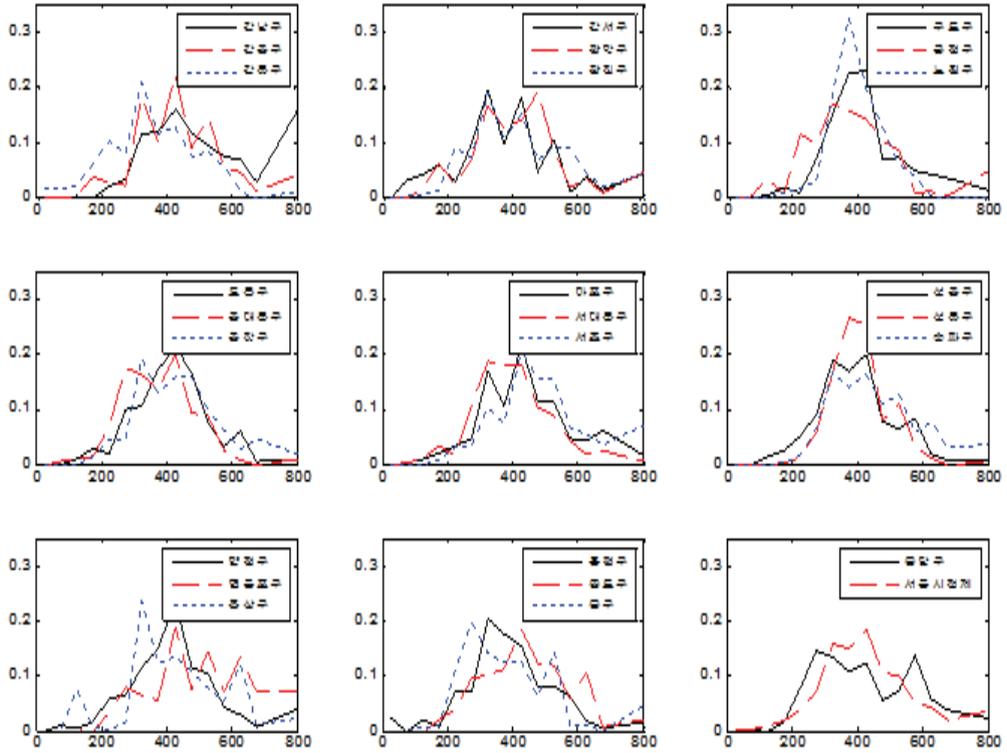
김세직(2014): “경제성장과 교육의 공정경쟁”, 서울대학교 경제연구소 편, 『경제논집』, **53**, **1**, 3-20.

김세직, 정운찬(2007): “미래 성장동력으로서의 창조형 인적자본과 이를 위한 교육개혁”, 『경제논집』, **46**, **4**, 187-214.

김세직, 류근관 외(2012): “성장동력으로서 창조형 인적자본 육성의 필요성”, 『협동연구총서』, 10-05-01.

- 김세직, 류근관, 손석준(2015): “수치적 베이지안 분석방법을 이용한 부모 소득과 자녀 성적 간의 관계: 명문대 합격생의 구별 분포”, 미출간 논문.
- 김희삼(2015): “사회 이동성 복원을 위한 교육정책의 방향”, 『KDI Focus』.
- Bouchard, Thomas J., and Matthew McGue(1981): “Familial Studies of Intelligence: A Review,” *Science* **212**, **4498**, 1055-1059.
- Irwing, Paul, and Richard Lynn(2006): “The Relation Between Childhood IQ and Income in Middle Age.” *The Journal of Social, Political, and Economic Studies*, **31**, 191-196.
- Jencks, Christopher, Marshall Smith, Henry Acland, and Mary Jo Bane(1972): *Inequality: A Reassessment of the Effect of Family and Schooling in America*, New York, Basic Books, Inc.
- Kim, Se-Jik(1997): “Growth Effects of Taxes in an Endogenous Growth Model: To What Extent Do Taxes Affect Economic Growth?” *Journal of Economic Dynamics and Control*, **23**, 125-158.
- Lucas, Robert(1988): “On the Mechanics of Economic Development.” *Journal of Monetary Economics*, **22**, 3-42.
- Zagorsky, Jay L.(2007): “Do You Have to Be Smart to Be Rich? “The Impact of IQ on Wealth, Income and Financial Distress.” *Intelligence*, **35**, 489-501.

〈부록〉



〈그림 4〉 각 구별 학부모 소득 원 자료로 그린 히스토그램

Abstract

Parental Wealth, Children's Ability and Entering Prestigious Colleges

Se-Jik Kim, Keunkwan Ryu, and Suk Joon Son

Depending on parental wealth, there is a marked difference in the probability of entering more prestigious colleges in Korea. A hypothesis to explain this difference could be that the children of rich parents are more likely to be born with higher intellectual ability, which is genetically inherited from their parents who have become rich because of innate abilities. An alternative explanation would be that the wealth of rich parents enables the children to purchase more of effective means to raise the probability of entering top colleges (for example, attending expensive private schools) though the children are not born with higher abilities.

This paper empirically explores whether the difference in the chance of entering top colleges in Korea reflects students' innate abilities (say, intelligence) or other factors that are correlated with parental wealth, such as enhanced signaling. Based on income survey data for Seoul, we estimate the probability distribution of parents' income across 25 districts of Seoul. By using the existing estimates for income-intelligence correlation and parents-children intelligence correlation, we then estimate the probability distribution of students' innate intelligence, which is used to predict the probability of entering top colleges due to students' innate ability for each of 25 districts of Seoul. Finally, we compare between the predicted and the actual difference in the probability across the districts.

We find that the predicted difference in the probability of entering Seoul National University due to students' innate ability across the districts may account for only 10-20% of the actual difference. This suggests that 80-90% of the actual difference may

be explained by the factors that are correlated with parental wealth, other than children's innate ability.

Keywords: Parental Wealth, Children's Innate Ability, Education, Economic Growth

