

연령이 학업성적에 미치는 효과는 언제까지 지속되는가?: 한국의 경험

남 기 곤¹⁾

요 약

본 논문은 한국직업능력개발원에서 조사한 KEEP 자료를 이용하여, 2004년 중학교 3학년이었던 학생들이 2007년 말 고등학교 졸업 단계에 도달했을 때 연령에 따른 학업성적의 효과가 어떻게 변화하고 있는지를 분석하였다. 또한 이 자료에 대한 분석 결과를 뒷받침하기 위해 TIMSS 1999, 2003, 2007년 자료와 KELS 자료, KYPS 자료를 보충적으로 분석하였다.

분석 결과 중학교 단계에서는 연령이 1개월 많아짐에 따라 학업성적이 0.02 표준편차 정도 증가하는 경향이 유의미하게 나타나지만, 고등학교 졸업단계에서는 연령의 효과가 존재하지 않는 것으로 나타났다. 본 논문에서는 고등학교 단계를 경유하면서 연령이 학업성적에 미치는 유의미한 효과가 사라지는 원인과 관련해서는 다음 두 가지 가능성을 제시하였다. 첫째, 학업이 집중적으로 이루어지는 고등학교 단계에서 상대적으로 학업성적이 뒤쳐졌던 연령이 적은 학생들이 노력을 더욱 배가함으로써 성적 부진을 만회했을 가능성. 둘째, 연령이 많을수록 조기 성숙함에 따라 낮은 수준의 탈선 확률이 더 높았을 가능성. 이를 뒷받침하는 증거로 연령이 많을수록 사교육비가 줄어들고 있고 학습시간의 격차가 사라지고 있다는 분석 결과와, 연령이 많을수록 흡연과 음주 그리고 이성친구 교제 확률이 높아지고 있다는 분석 결과를 제시하였다

I. 서론

어느 나라에서나 하나의 기준 시점 (single cutoff date)에 의거하여 아동의 공식 교육기관에의 입학 여부가 결정된다. 예를 들어 미국의 많은 주에서는 9월 1일 시점에서 5세 이상의 연령이 되는 아동들이 유치원(kindergarten)에 입학할 수 있다. 학교의 개학이 3월에 이루어지는 한국에서는 3월 1일의 시점에서 만 6세 이상의 연령에 도달한 아동들이 초등학교에 입학할 수 있다. 따라서 동일한 연도에 초등학교에 입학한 아동이라도 월 단위의 실질 연령 상에 차이가 존재하게 된다. 한국의 경우 3월생은 가장 연령이 많은 반면 2월생은 가장 연령이 적어, 이들 간에는 평균적으로 11개월의 연령 차이가 발생한다.

최근의 경제학 분야 연구들은 이와 같은 실질 연령의 차이가 초등학교 시절의 학업 성적에 영향을 미칠 뿐만 아니라, 중학교 단계의 (lower secondary education) 학업성취도에도 유의한 영향을 미치고 있음을 밝히고 있다. 다양한 시기의 여러 국가들에 대한 실증분석들은 대체로

1) 한밭대학교 경제학과 교수

일관된 결과를 보여주고 있다. (Bedard and Dhuey, 2006; Elder and Lubotsky, 2008; McEwan and Shapiro, 2008; Strøm, 2004) 그러나 이와 같은 연령의 효과가 고등학교 단계나 (upper secondary education) 대학 이후의 단계 혹은 노동시장에 진입한 이후에까지 지속적으로 나타나는지에 대해서는 엇갈린 분석 결과가 제시되고 있다. 연령이 많은 학생은 연령이 적은 학생에 비해 고등학교 단계에도 학업성적이 높으며, 고등학교를 졸업할 확률이 높고, 대학에 입학할 가능성이 높다는 분석 결과들이 있다. (Bedard and Dhuey, 2006; Fredriksson and Öckert, 2005; Kawaguchi, 2011) 반면 연령의 효과가 고등학교 이후 단계에는 축소되거나 사라지며, 노동시장에 진입한 이후에는 별다른 차이를 나타내지 못한다는 연구들도 존재한다. (Cascio and Schanzenbach, 2007; Dobkin and Ferreira, 2010; Puhani and Weber, 2007)

연령이 학업성적에 유의한 영향을 미치는지 여부도 중요하지만, 이러한 효과가 어느 단계까지 지속되는지 또한 교육 정책의 측면에서 매우 중요한 문제이다. 만약 연령이 중등교육 단계를 넘어 고등교육 단계 그리고 노동시장의 성과에도 영향을 미친다면, 이러한 차이를 해소하기 위한 대폭적인 제도 개선이 필요할 수 있다. 자신의 능력이나 노력이 아닌 생년월과 school cutoff date라는 우연적인 그리고 제도적인 요인 때문에 생애 소득이 달라질 수 있다는 사실에 대해서는 사람들이 동의하기 어려울 수 있기 때문이다. 하지만 시간이 지남에 따라 연령의 효과가 점차 감소하고 사라지는 것이라면, 현재의 교육체계를 유지하면서 각 교육 단계에서 연령이 적은 학생들이 겪는 어려움을 완화시키는 보조적인 방식을 찾아 나가는 것이 합리적일 수 있다. 또한 어떠한 요인 때문에 연령의 효과가 감소되는지를 파악할 수 있다면, 이러한 요인을 강화시키는 방향으로 교육 정책을 수행할 필요가 있을 것이다.

본 논문은 한국의 경우 연령이 학업성적에 미치는 효과가 중학교 시기에서 고등학교 시기로 넘어가면서 어떻게 변화하는지는 추적하고자 한다. 이 주제와 관련된 국내 논문으로는 김태훈(2011)이 주목된다. 이 논문은 할당된 상대 연령을 도구변수로 이용하여 월 단위 연령이 대학 진학 확률에 어떠한 영향을 미쳤는지를 분석하고 있다. 주된 분석 결과는 연령이 높을수록 대학 진학 자체는 유의한 차이를 보이지 않지만, 4년제 대학 진학 확률은 유의하게 높아진다는 것이다. (대략 1달에 0.4%) 그러나 이 연구는 ①1970-1987년이라는 18개 출생년도의 사람 6,489명에 대한 분석이라는 점에서 표본 수가 적고 구성이 이질적이라는 점, ② 생일을음력으로 응답하는 사람에 대한 양력 환산이 철저하게 이루어지지 않았다는 점, ③ 아마도 이러한 영향으로 인해 조기 입학자나 취학 유예자의 비율이 공식 통계에서 확인되는 일반적인 추세보다 높다는 점, ④ 취학유예 여부에 대한 도구변수로 1-2월 출생 여부를 사용하는 등 분석 방식 상에 논란이 있을 수 있다는 점 등의 한계가 있다고 판단된다. 그 외 국내 연구로는 초등학교 입학연령이 초등학교 시기 학업 성적에 미치는 영향을 분석하고 있는 서우석(1991)과 정은희(2003), 그리고 영재 교육을 받는 학생들을 대상으로 이들의 출생월 분포를 분석하고 있는 이순주·박찬웅(2008) 등이 있다. 이들 국내의 교육학 관련 연구들은 평균치를 분석하는 수준에 머무르고 있다는 한계를 가지고 있다. 이 주제에 있어서 한국의 경험은 다음과 같은 측면에서 분석 상의 장점이 있다.

첫째, 한국은 일본과 유사하게 “clean” education system을 가지고 있다. 모든 학생은 특별한 사유가 없는 한 정해진 시간에 초등학교에 입학하여야 하며, 상급 학년으로 진급할 때 낙제하는 경우가 거의 없다. 일반적으로 다른 나라들의 경우, 같은 학년 내에 연령이 많은 학생일수록 능력(ability)이 낮은 학생들의 비율이 높다. 이들 중에는 스스로의 선택에 의해 학교에 늦게 입학하거나 낙제를 하는 경우가 많기 때문이다. 이로 인해 연령이 학업성적에 미치는 인과적 효과를 (능력이 동일한 학생이라도 연령이 많을수록 학업성적이 높아지는 효과) 분석하기 어렵다. 반면 한국의 경우 한 학년 내에 동일한 연령대의 학생이 집중되어 있기 때문에, 학업성적의 분석에서 연령만의 효과를 분리해서 살펴보는 것이 보다 용이하다.

둘째, 많은 나라에서 중학교 학생들을 대상으로는 공통된 학업 시험 결과가 존재하지만, 고등학교 학생에 대한 성적 자료는 존재하지 않는 경우가 대부분이다. 이로 인해 고등학교 단계의 연령 효과를 분석하는 기존 연구들에서는 고등학교 졸업 여부, 대학입학 자격시험을 보았는지 여부, 대학에 입학했는지 여부 등과 같은 변수들을 분석하곤 하였다. 이들 지표도 학업 성과를 나타내는 주요한 변수인 것은 사실이다. 하지만 연령이 학업 성적이라는 연속변수에는 유의미한 영향을 미치더라도 그 정도가 크지 않을 경우, 혹은 연령의 효과가 성적 계층에 따라 차별적일 경우, 이러한 더미변수에 미치는 영향은 유의미하지 않은 것으로 나타날 수도 있다. 반면 한국의 경우 고등학교를 졸업하고 대학에 입학하기 위해서는 모든 학생들이 수학능력시험이라는 국가에서 주관하는 공통된 시험을 보아야 한다. 매년 한차례에 시행되는 이 시험은 대학 입학에 결정적으로 중요한 지표이기 때문에, 학생들은 최선의 노력을 다해 시험을 본다. 최근 들어 이 시험 성적이 설문조사 자료에 공개되곤 하는데, 이는 고등학교 졸업단계 학생들의 학업 성취도를 정확하게 측정할 수 있는 중요한 자료이다.

셋째, 연령이 학업성적에 미치는 효과에 대한 시계열 추세를 분석하는 지금까지의 기존 연구들은 대부분 횡단면 자료를 이용하여 왔다. 한국의 경우 최근 학생들을 대상으로 한 패널조사들이 활발히 진행되고 있기 때문에, 이를 이용하면 동일한 학생이 중학교 단계에서 고등학교 단계로 이행하면서 학업성적에 어떠한 변화를 나타내고 있는지, 그리고 이러한 변화를 야기하는 요인들이 무엇인지 등을 추적하는 것이 가능하다.

본 논문은 한국직업능력개발원에서 조사한KEEP(Korean Education and Employment Panel) 자료를 이용하여, 2004년 중학교 3학년이었던 학생들이 2007년 말 고등학교 졸업 단계에 도달했을 때 연령에 따른 학업성적의 효과가 어떻게 변화하고 있는지를 분석한다. 또한 이 자료에 대한 분석 결과를 뒷받침하기 위해IEA(International Association for the Evaluation of Educational Achievement)에서 조사하는TIMSS(Trends in International Mathematics and Science Study) 1999, 2003, 2007년 자료와 한국교육개발원에서 조사하는 KELS(Korean Education Longitudinal Study) 자료, 한국청소년개발원에서 조사하는 KYPS(Korea Youth Panel Survey) 자료를 보충적으로 분석한다. 모든 자료는 two-stage stratified random sampling 방식으로 추출된 자료이다. 첫 번째 단계에서 학교가 선정되고, 두 번째 단계에서 해당 학교에 할당된 수만큼의 학생이 선정되었다.

본 논문의 분석 결과 중학교 단계에서는 연령이 1개월 많아짐에 따라 학업성적이 0.02 표준편차 정도 증가하는 경향이 유의미하게 나타나지만, 고등학교 졸업단계에서는 연령의 효과가 존재하지 않는 것으로 나타나고 있다. 이러한 현상은 본 논문에서 분석한 모든 자료에서 공통적으로 확인되는 사실이다. 고등학교 단계를 경유하면서 연령이 학업성적에 미치는 유의미한 효과가 사라지는 원인과 관련해서는 다음 두 가지 가능성을 고려할 수 있다. 첫째, 학업이 집중적으로 이루어지는 고등학교 단계에서 상대적으로 학업성적이 뒤쳐졌던 연령이 적은 학생들이 노력을 더욱 배가함으로써 성적 부진을 만회했을 가능성. 둘째, 연령이 많을수록 조기 성숙함에 따라 낮은 수준의 탈선 확률이 더 높았을 가능성. 이를 뒷받침하는 증거로 연령이 많을수록 사교육비가 줄어들고 있고 학습시간의 격차가 사라지고 있다는 분석 결과와, 연령이 많을수록 흡연과 음주 그리고 이성친구 교제 확률이 높아지고 있다는 분석 결과를 제시할 수 있다.

논문의 구성은 다음과 같다. 2절에서는 한국의 교육제도에 대해 설명하고, 3절에서는 실증분석 방식에 대해 설명한다. 다음 4절에서는 KEEP 자료에 대한 분석 결과를 제시하고, 5절에서는 다른 자료에 대한 분석 결과를 살펴본다. 마지막으로 6절에서는 이상의 분석 내용을 정리한다.

II. 한국의 교육제도

한국에서 초등학교 입학일은 3월 1일이며, 이 날이 school cutoff date의 기준 시점이다. 2009년부터 한국에서 school cutoff date는 1월 1일로 변경되었다. 그 이전까지는 계속 3월 1일이었다. 이 시점까지 만 6세에 도달한, 즉 태어나서 6번째의 생일이 경과된 아동만이 초등학교에 입학할 수 있는 자격이 부여된다. 물론 만 6세가 지났어도 아동의 성숙 정도가 더디거나 특별한 사유가 있는 경우, 부모의 신청에 의해 입학이 유예되는 것이 가능하다. 하지만 이 비율은 매우 낮았다. 본 논문의 주된 자료인 KEEP 중학생 자료에서의 조사대상인 1989년생(1990년 2월까지)은 1996년도에 초등학교에 입학하게 되는데, 당시 공식통계를 보면 1996년 초등학교 입학대상자 650,743명 중 ‘적령아동(children at the right age)’이 아닌 ‘유예 및 과령아(overage children)’는 8,050명으로 1.2%에 불과하였다. 또한 만 6세 이전의 아동이 초등학교에 입학하는 비율도 매우 낮았다. 사실 1995년 이전까지는 이와 같은 조기 입학이 법적으로 불가능한 상황이었고, 1995년 교육개혁으로 1996년부터 초등학교 조기입학이 가능해졌지만 초기 시행 과정에서 이를 이용하는 경우는 매우 드물었다. 만5세 아동에 대한 조기입학이 허용되기 시작한 첫 해인 1996년의 통계를 보면, 전체 취학자 수는 636,160명이고 조기입학자 수는 5,661명으로 0.8% 수준이었다. 이후에도 조기입학자 수는 1997년 5,790명, 1998년 7,923명, 1999년 8,851명에 머무른 것으로 나타나고 있다. (매일경제신문, 1999.10.4)

한국의 교육기간은 초등학교가 6년, 중학교 3년, 고등학교 3년이다. 고등학교를 졸업한 학생은 2년 과정의 전문대학에 입학하거나 4년 과정의 일반대학에 진학할 수 있다. 중학교까지는

의무교육이며 고등학교 진학부터는 선택이지만, 대부분의 한국 학생들은 고등학교까지 교육을 받는다. 한국에서 초등학교에 대한 의무교육은 1954년부터 실시되었다. 중학교의 경우 도서벽지는 1986년부터, 읍면지역은 1992년부터, 그리고 도시지역은 2002년부터 의무교육이 실시되었다. 하지만 중학교 진학률은 이미 1980년대부터 거의 100%에 가까운 수준을 유지해 왔다. 2002년 당해 연도에 초등학교를 졸업한 학생 중 바로 그 해 중학교에 입학한 학생의 비율은 99.9%였고, 2005년 당해 연도에 중학교 졸업자 중 바로 고등학교에 입학한 학생의 비율은 99.7%였다. 대학진학률은 이보다는 낮지만 역시 높은 수준이다. 2008년 당해 연도에 고등학교를 졸업한 학생 중, 그 해 전문대학이나 일반대학에 진학한 학생의 비율은 83.8%였다. 또한 학년이 올라갈 때 자격고사를 보는 것이 아니기 때문에, 특별한 개인적 사정이 없는 한 유급하거나 이로 인해 복학하는 경우는 매우 드물다. 1996년 초등학교 통계, 2002년 중학교 통계, 2005년 고등학교 통계를 분석해보면, 이 기간 동안 재취학(재입학, 편입, 복학) 학생의 비율은 매년 0.2-0.4% 수준에 머물고 있다.

이처럼 한국의 경우 생년월을 기준으로 초등학교 입학 허용 여부가 엄격히 통제될 뿐만 아니라, 초등학교에서 중학교 그리고 고등학교로의 진학이나 그 내에서 학년이 올라갈 때에도 탈락하는 학생이 매우 적기 때문에, 한 학년 내에 동일한 연령대의 학생이 집중되어 있는 clean education system 성격을 잘 보여주고 있다.

한국에서는 중등교육 단계까지 교육이 평준화되어 있는 경향이 강하다. 초등학교와 중학교는 본인의 학업 성적이나 능력과는 상관없이 거주지에 가까운 학교에 배정을 받으며, 모든 학생들은 동일한 커리큘럼에 의해 교육을 받는다. 고등학교는 학문적인 교육에 중점을 두는 일반계 고등학교와 직업교육에 중점을 두는 실업계(전문계) 고등학교, 예체능 고등학교, 그리고 과학고와 외국어고와 같은 특수목적 고등학교로 분리된다. 2005년 고등학교 입학자 중 일반계 고등학교에 진학한 학생의 비율은 69.2%, 실업계 고등학교에 진학한 학생의 비율은 28.3%였고, 예체능 고등학교에 1.1%, 특수목적 고등학교에 1.4%의 학생이 진학하였다. 일반계 고등학교 진학자는 대도시의 경우 평준화 제도(high school equalization policy)에 적용을 받아 거주 지역의 학군 내 특정학교에 무작위 강제 배정을 받는 반면(평준화지역), 중소도시나 시골 지역의 경우 각 학교마다 학생들을 선발하고 있다.(비평준화지역) 예체능 고등학교와 특수목적 고등학교 역시 자체 기준에 따라 입학생을 선발한다.

한국에서도 고등학교 단계에서는 학생의 재능이나 학업 성적에 따라 진학하는 고등학교의 유형이 달라지는 차별적인 교육시스템이 사용되고 있지만, 50%에 가까운 학생들이 평준화제도에 의해 거주지역 학군 내 특정학교에 강제배정되고 있다는 특성을 가지고 있다. 여기서 50%란 전체 고등학생 중 평준화지역 일반계 고등학교 재학생의 비율을 의미한다. 한국의 교육제도, 특히 고교평준화제도의 역사 및 자세한 특성에 대해서는 Chung(1998), Nam and Sung(2009), 남기곤(2012) 참조. 물론 매우 적은 수를 선발하는 특수목적 고등학교나 비평준화지역의 자립형사립 고등학교에 입학하려는 경쟁이 중학교 단계에서 상위 성적 학생들을 중심으로 이루어지고 있는 것은 사실이다. 하지만 대부분의 중학교 학생들은 거주 지역 근처의 고등학교에 입학하는 것 자체에 큰 문제가 없고, 평준화제도에 의해 어느 고등학교나 큰 차별성이 없기 때문에, 중학교 단계의

학업 경쟁은 상대적으로 약한 상황이다. 한국의 경우 교육열이 워낙 높기 때문에 초등학교나 중학교 단계에서도 학생들의 공부에 대한 압박감은 대단히 크다. 여기서 학업경쟁이 약하다는 표현은 고등학교 단계와 비교할 때 상대적으로 약하다는 것을 의미한다. 그리고 중학교 단계에서 하위 성적권 학생들은 자신의 희망과는 달리 일반계 고등학교에 배정되지 못하고 실업계 고등학교 배정될 위험이 있기 때문에, 이에 대한 부담도 존재한다.

어느 유형의 고등학교에 다니든 대학에 입학하기 위해서는 국가에서 주관하는 수능능력시험을 보아야 한다. 매년 한차례씩 시행되는 이 시험에서는 언어, 수학, 영어 과목이 중요하며, 그 외 사회 과목들이나 과학 과목들 그리고 다른 외국어 과목들 중에서 일부 과목을 선택하여 시험을 보게 된다. 각 대학마다 입시 전형 방식은 차이가 있지만 고등학교 단계의 내신성적이나 간단한 면접고사와 더불어 이 수능능력시험 성적은 대학 입학 여부에 결정적인 영향을 미친다. 따라서 고등학교 학생들은 이 시험에 대한 준비를 위해 매우 치열하게 학습을 한다. 2004년 통계청에서 조사한 생활시간조사에 따르면 하루 평균 학습시간이 초등학생과 중학생의 경우 각각 6시간 14분과 7시간 19분인 반면, 고등학생은 8시간 52분이었다.

이와 같은 한국의 교육 시스템 상의 특성은 대부분의 나라에서 확인되어지는 '연령이 학업 성적에 미치는 효과'에도 영향을 미쳤을 가능성이 있다. 중학교 단계까지 평준화된 교육이 실시되고 있고 학업에 대한 본격적인 경쟁이 고등학교 단계에서 이루어진다는 사실은, '연령이 학업성적에 미치는 효과'가 존재한다고 하더라도 확대되고 강화되는데 한계가 있을 수 있다는 가능성, 그리고 고등학교 단계에서 모든 학생들이 학업에 집중할 경우 그 효과가 약화될 수 있다는 가능성을 시사해준다.

III. 분석 모델

초등학교 입학 초기 단계에서 발생하는 아동들의 실질 연령 상의 차이가 학업 성적의 격차로 연결되는 경향이 강하다는 것은 일반적으로 잘 알려진 사실이다. (Stipek, 2002) 문제는 이러한 추세가 학업 초기 단계에 국한되는 현상인지, 아니면 이후 학업 단계에서도 지속적으로 나타나는 것인지 여부이다. 이 문제와 관련하여 TIMSS 1995년과 1999년의 자료를 이용하여 OECD 국가의 학생들을 대상으로, 월 단위의 연령의 차이가 학업성적에 어떠한 인과관계를 미치고 있는지를 분석하고 있는 Bedard and Dhuey(2004, 2006)의 연구 결과는 흥미롭다. Bedard and Dhuey(2004)에서는 한국과 터키가 분석 대상에 포함되어 있지만, 이를 발전시켜 학술지에 게재한 Bedard and Dhuey(2006)에서는 이 두 국가에 관한 분석 결과를 제외하고 있다. 그 이유는 특정 월 출생자의 비율이 비합리적으로(unreasonably) 높아(한국의 경우 1월과 2월, 터키의 경우 1월), 출생월이 무작위적이라는 가정이 의심스럽기 때문이었다. 한국에 대해서는 그 이유에 대해 다음과 같이 설명하고 있다. “한국의 동료들에 따르면 이 나라에서는 상대적으로 연령이 높은 자녀(3-6월생)를 가진 부모들이 자녀를 학교에 일찍 입학시키기 위해, 출생증명서

상의 출생일자를 조정하도록 관리들에게 뇌물을 제공하는(to bribe officials) 것은 흔한(not uncommon) 현상이라고 한다.” Bedard and Dhuey(2004)와 Bedard and Dhuey(2006)에서는 회귀분석 시 통제변수의 차이가 있기 때문에, 최종 추정 결과도 다소 차이가 있지만 분석 결과의 전체적인 추세는 동일하다. 이 연구에서 주목하고 있는 사실은 관측된 연령(observed age)이 학업성적에 미치는 영향을 단순한 OLS 방식으로 추정하게 되면 하향편의(downward bias)가 발생할 수밖에 없다는 점이다. 관측된 연령이 많은 그룹에는 학교를 정상적인 경우보다 늦게 입학하였거나 낙제한 학생들이 다수 포함되어 있는데, 이들은 원래부터 ability가 낮은 학생들의 비율이 높을 수 있기 때문이다.

이러한 문제를 해결하기 위해 이들은 관측된 연령에 대한 도구변수(instrument variable)로 ‘할당된 상대 연령(assigned relative age)’이라는 변수를 사용할 것을 제안하고 있다. 이는 학생들의 출생연도는 고려하지 않고 단지 출생월만을 고려하여 계산한 연령인데, 초등학교 입학 시점보다 출생월이 몇 개월 더 앞서 있는지를 나타내준다. (birth month relative to the school cutoff date) 한국의 경우 2월생은 0의 값을 가지는 반면, 1월생은 1, 12월생은 2, 이러한 방식으로 3월생은 11의 값을 가지게 된다. 자녀의 출생월이 무작위적(random)이라고 한다면, 할당된 상대 연령이 가정 환경이나 학생의 능력과 관련이 있다고 보기 어렵다. 그럼에도 불구하고 할당된 상대 연령에 따라 학생들의 성적이 차별적이라고 한다면, 이는 할당된 상대 연령에 따라 실제 관측된 연령이 달라지기 때문에 나타나는 효과라고 볼 수 있다. 이러한 의미에서 할당된 상대 연령은 실제 관측된 연령의 유의미한 도구변수가 될 수 있다. 서론 부분에서 지적했던 2000년대 중반 이후대부분의 기존 연구들은 이러한 도구변수를 사용하여 분석하고 있다.

본 논문에서는 기본적으로 이러한 도구변수 방식을 이용하는데, KEEP 자료의 경우 생일에 대한 정보까지 제공되고 있어서 관측된 연령과 할당된 상대 연령 변수를 모두 일단위로까지 확인하여 사용한다. 관측된 연령의 경우 기준시점은 설문조사 시기인 2004년 11월 30일을 사용하였고, 할당된 상대연령은 적령연령 층이 초등학교에 입학한 1996년 3월 1일을 기준 시점으로 사용하였다. 실제 분석에서는 두 변수 모두 30으로 나눈 값을 사용하여, 월 단위 효과를 추정하였다. 생일에 대한 정보를 이용하는 경우와 생월 정보만을 이용하는 경우 분석 결과는 거의 차이가 없었다. KEEP 이외의 자료는 생일에 대한 정보가 없기 때문에, 생월 정보만을 이용하여 분석하였다. 이 외에 본 논문에서는 다음과 같은 두 가지 방식의 분석을 보충적으로 사용한다. 우선 clean education system을 가지고 있는 한국의 경우 한 학년에 대부분 적령 학생으로 구성되어 있다는 점을 감안하여, 이들에 대해서만 연령이 학업성적에 미치는 효과를 OLS 방식으로 추정한다. (예를 들어 KEEP 중학생 자료의 경우 1989년 3월생부터 1990년 2월생까지를 분석) 다음으로 1990년대 중반까지 한국의 부모들은 자녀의 출생월이 1-2월인 경우를 선호하였고, 특히 학력이 높은 부모의 경우 이러한 경향이 보다 강하다는 Nam(2011)의 분석 결과를 감안하여, 적령 학생 중 1-2월생인 표본을 제외하고 (즉 KEEP 중학생 자료의 경우 1989년 3월부터 1989년 12월까지 학생만을 대상으로) 분석을 실시한다.

분석에 사용된 모델은 다음과 같다.

$$Y_{is} = \alpha + \beta \text{ObsAge}_{is} + \gamma X_{is} + \theta_s + \epsilon_{is}$$

종속변수 Y_{is} 는 s 학교에 재학 중인 i 학생의 학업성적에 대한 Z score인데, 경우에 따라서는 인문계 고등학교를 진학했는지 여부나 사교육비 등과 같은 다양한 변수들을 사용한다. ObsAge 는 학생의 관측된 월단위 연령을 나타내며, 이 변수의 계수값 β 가 유의미한지가 분석의 주된 초점이다. X 는 통제변수들의 벡터인데, 여기에는 남학생인지 여부, 아버지와 어머니 학력이 각각 전문대학 이상인지 여부, 첫째 자녀인지 여부, 친부와 친모가 존재하는지 여부, 월평균 가구소득, 자기 집을 보유하고 있는지 여부 변수를 사용한다. 고등학생에 대한 분석에서는 일반계 고등학교에 재학하고 있는지 여부에 대한 더미변수를 추가로 통제한다. 각 변수에 무응답이 존재할 경우 해당 값을 0으로 처리하고, 이 변수가 무응답인지 여부를 나타내는 더미변수를 추가하여 모든 관측치가 분석에 이용되도록 하였다. 그리고 모든 분석에서는 학생이 다니는 학교 s 에 대한 고정효과 θ_s 를 통제하였다. 이는 KEEP 중학생 자료의 성적이 전교 석차이기 때문에 학교를 통제해야 할 필요성이 있기 때문이기도 하며, 동일한 학교에 다니는 학생들은 거주지역이나 생활수준이 유사하고 비슷한 분위기의 교육환경에서 공부한다는 점을 감안하여, 이러한 공통적 요인을 통제하고도 연령의 효과가 존재하는지를 파악하기 위해서이기도 하다. 무응답에 대한 처리 방식이나 학교에 대한 고정효과 통제방식을 다르게 하더라도, 본 논문의 분석 결과의 주요 내용은 변화하지 않음을 밝혀둔다.

핵심적인 변수인 ObsAge 의 효과를 분석하기 위해 할당된 상대 연령을 나타내는 RelAge 변수를 도구변수로 사용하거나, 분석대상을 적령 학생만으로 혹은 그 중에서도 1-2월 출생자는 제외한 표본만을 대상으로 OLS 분석을 실시한다.

IV. KEEP 자료에 대한 분석 결과

(1) 자료에 대한 설명

이 절에서는 한국직업능력개발원이 2004년에 실시한 KEEP(한국교육고용패널조사) 자료를 분석한다. 이 자료는 2004년 현재 중학교 3학년 2,000명과 일반계 고등학교 3학년 2,000명 그리고 실업계 고등학교 3학년 2,000명을 대상으로 학교생활과 가정생활, 학습활동과 진로 등을 조사한 것이다. 학생에 대한 조사와 병행하여 학부모와 담임선생님에 대해서도 함께 조사를 진행하였다. 그 뒤 매년 동일 학생에 대해 추적조사를 실시하여, 현재 7차년도 조사까지 완료되었다. 여기서는 KEEP 중학생 자료 중 중학교 3학년 시기인 1차년도(2004년 실시) 자료부터, 고등학교 3학년 시기인 4차년도(2007년 실시) 자료까지를 이용한다.

우선 이 자료에서 조사한 주민등록번호를 이용하여 학생의 생년월일을 확인하였다. 이 과정에서 6명의 표본이 탈락하여, 1차년도 자료 중 분석에 사용된 표본수는 총 1,994명이다. 다음 <표 1>은 생년월별로 표본이 어떠한 분포를 보이는지 정리한 것이다. 이를 보면 예상했던 대로 이 학년의 적령 학생층인 1989년 3월생부터 1990년 2월생까지의 학생 수가 1,994명 중 1,927명으로 96.6%의 압도적인 비율을 나타내고 있다. (가중치를 적용할 경우 이 비율은 96.3%임) 또 한가지 주목할 수 있는 특징은 다른 월보다 1-2월 출생자의 비율이 매우 높다는 점이다. 이는 Nam(2011)에서 밝혀진 바와 같이 한국에서 신혼부부의 결혼시기가 3-5월의 봄철에 집중되어 있기 때문이기도 하지만, 일부 부모들이 자녀의 초등학교 입학 연도를 한 해 앞당기기 위해 출생월을 조정하는 경향이 있기 때문이기도 하다. <부록 그림 1>에서는 KEEP 중학생 자료에서 학생의 가중치를 이용하여 각 생년월별로 하루 평균 출생자수를 구하여, 이를 당시 출생아 수에 대한 공식통계자료와 비교하고 있다. 두 자료가 거의 유사한 추세를 분포를 보이고 있음을 확인할 수 있다. 만약 상대적으로 연령이 적은 1-2월생 중에 부모가 출생월을 조정할 정도로 자녀에 대한 관심이 높은 경우가 보다 많이 존재한다면, 이는 연령이 학업성적에 미치는 효과를 과소추정하도록 만들 가능성이 있다. 물론 그렇지 않은 경우나 반대의 경우도 가능할 수 있다. 1-2월에 출생한 학생과 부모가 어떠한 특수성을 가지고 있으며, 이로 인해 학업 성적이 어떠한 영향을 받는지 등에 대해서는 좀더 체계적인 분석이 필요한 상황이다. 이러한 점을 감안하여 여기서는 1-2월 출생자를 제외한 분석 결과도 함께 살펴보기로 한다.

1차년도 조사에서 1,994명이었던 유효 표본수가 2차년도에는 1,769명, 3차년도에는 1,746명, 4차년도에는 1,691명으로 감소하였다. 4차년도까지의 표본 유지율은 84.8%이다. 1차년도 조사에서는 담임 선생님에게 해당 학생의 전교 석차 백분율을 기록하도록 요청하였다. 학생들의 성적이 정규분포를 한다고 가정하면, 학생들의 성적 백분율을 이용하여 평균이 0이고 표준편차가 1인 표준정규분포의 Z score를 역산할 수 있다. 성적 백분율이 100%인 경우, Z score 계산을 위해 99.9의 값을 부여하였다. 물론 학교 간에 차별성이 있기 때문에 해당 학교의 전체 석차가 학생의 객관적인 성적을 나타내주는 것은 아니다. 하지만 본 논문의 분석에서는 개별 학교의 고정효과를 통제하고 있기 때문에, 한 학교 내에서 연령이 높을수록 학업 석차가 높아지는 효과가 나타나지는지를 점검하는 의미로 이 지표를 사용하는 것이 가능하다.

이들이 고등학교에 진학하여 정상적인 과정을 밟을 경우, 대학에 진학하고자 하는 학생은 고등학교 3학년 시기인 2007년 말에 수학능력시험을 보게 된다. KEEP 자료에는 이들의 시험 성적 결과를 시험 시행 기관을 통해 제공받아 제시하고 있다. 시험과목은 크게 언어, 수리, 외국어, 탐구, 제2외국어 영역으로 나뉜다. 이 중 탐구 영역과 제2외국어 영역은 해당 영역 내 어떠한 과목을 선택하는지가 학생마다 다르기 때문에, 언어와 수리, 외국어에 대한 성적만을 분석한다. 수학능력시험 성적에 대한 자료는 통상 각 과목별로 표준점수와 백분위점수 및 등급과 같은 다양한 형태로 제공되어 왔지만, 이 시기에는 정부의 방침에 따라 각 과목별로 해당 학생의 성적이 9개의 등급 중에 어떠한 등급에 속하는지만 발표되었다. 예를 들어 백분위점수가 1-4%에 속할 경우 1등급, 5-11%에 속할 경우 2등급 식으로 등급이 부여되는 것이다. 여기서는 각 등급의 중앙값을 그 학생의 성적 백분율이라 가정하고, 앞의 전체 석차를 변환시

키는 것과 동일한 방식으로 Z score를 산출하였다.

<부표 1>에는 학생의 생년월별로 1차년도 중학교 3학년 성적과 4차년도 수학능력시험 성적 자료의 Z score의 평균값을 정리하고 있다. 우선 중학교 성적은 이를 도표화한 <부록 그림 2>에서도 확인할 수 있듯이 뒤의 생년월일수록, 즉 연령이 적어질수록 평균 성적도 낮아지는 경향을 보여준다. 하지만 수학능력시험 성적의 경우에는 어떠한 과목도 일관된 추세를 보여주고 있지 못하다. 물론 학생마다 개인의 특성이나 가정환경 그리고 소속 학교의 특성이 다르기 때문에, 이러한 평균 수준의 분석은 정확하지 않을 수 있다. 따라서 관측가능한 여러 요인들을 통제한 상태에서 연령이 학업성적에 미치는 효과가 어떠한 추세를 보이는지를 확인하기 위해 회귀분석 방식을 사용한다. 이 때 통제하는 변수에 대한 설명과 평균값에 관한 정보는 <부표 2>에 정리하였다. <부표 1>의 마지막 열에서는 생년월별로 부모의학력이 전문대학 이상인 비율을 보여주고 있는데, 생년월별로 일관된추세 변화는 발견되지 않는다. 단 1-2월 출생자인 경우 부모의 학력이 평균 수준을 상회하고 있다는 사실은 확인된다.

(2) 학업성적에 대한 분석

다음 <표 2>는 KEEP 중학생 자료를 이용하여 연령이 학업성적에 미치는 효과에 대한 분석 결과를 정리한 것이다. 표의(1)열은 ‘할당된 상대 연령’을 도구변수로 사용했을 때의 분석 결과이며, (2)열은 이 학년의 적령 학생층인 1989년 3월생부터 1990년 2월생까지만을 분석대상으로 할 경우, 그리고 (3)열은 여기서 1-2월생은 제외하고 1989년 3월생부터 1989년 12월생까지만을 분석대상으로 할 경우의 분석 결과이다. 이 표에는 회귀분석 결과 중 연령 변수(ObsAge)만의 분석 결과를 제시하고 있다. 전체적인 분석 결과는 <부표 3>과 <부표 4>에 제시되어 있다. 분석 시기와 대상 그리고 방법에 따라 각 변수들의 분석 결과는 다소 상이하지만, 대체로 여학생일수록 부모의 학력이 높고 가정환경이 양호할수록 학업성적이 높아지는 경향은 어느 모델에서나 확인되어지고 있다.

우선 중학교 3학년 성적을 보면 모든 모델에서 연령 변수는 대략 0.02의 계수값을 가지며 성적에 유의한 영향을 미치는 것으로 나타났다. IV 분석결과를 기준으로 할 경우 연령이 1개월 증가함에 따라 학업성적이 0.025 표준편차만큼 증가하는 것으로 나타나고 있는데, 이는 3월생이 11개월 어린 다음해 2월생에 비해 0.275 표준편차 가량 성적이 높아진다는 것을 의미한다. 이는 성적 백분율로 분석할 경우 7 percentile 가량 순위가 높아지는 효과이며, 통상 사립학교의 성적 향상 효과가 0.1-0.2 표준편차 수준이라는 점을 감안하더라도 무시하기 어려운 수치이다. 사립학교의 성적 향상 효과에 대한 기존연구로는 Alexander and Pallas(1983), Coleman and Hoffer(1987) 참조. Bedard and Dhuey(2006) 등 기존 연구들에서 제시된 다른 나라의 결과들과 거의 유사한 값이기도 하다.

반면 고등학교 3학년 말에 보는 수학능력시험 성적 결과는 이와는 전혀 다른 모습을 보여준다. 어느 과목이든 그리고 어떠한 모델이든 연령이 학업성적에 유의한 플러스 경향을 보이는 경우는 나타나지 않고 있다. 언어의 경우는 오히려 마이너스의 부호를 보이고 있으며, (3) 모

델에서는 미약하지만 유의한 마이너스 효과가 나타나고 있다. 수리와 외국어의 경우 모든 모델에서 계수값은 통계적으로 유의하지 않았다.

중학교 3학년까지는 연령이 높을수록 성적이 높아지는 경향이 뚜렷하지만, 대학 입학을 앞둔 고등학교 3학년 말에는 그러한 효과가 나타나지 않는다. 왜 그럴까? 한가지 유력한 설명은 수학능력시험이란 모든 고등학생들이 반드시 보아야 하는 의무적인 시험이 아니라, 대학에 진학하고자 하는 학생만이 선택적으로 보는 시험이라는 특성 때문에 나타나는 결과일 수 있다. 즉 중학교 단계에서 연령이 낮아 학업성적이 좋지 않았던 학생일수록 인문계 고등학교가 아닌 실업계 고등학교에 진학하거나 아예 고등학교에 진학하지 않았을 가능성, 혹은 고등학교 유형은 같더라도 연령이 낮을수록 대학 진학을 스스로 포기하여 수학능력시험을 보지 않았을 가능성이 있다. 그렇다면 고등학교 단계에도 연령이 낮은 학생일수록 학업성적이 낮아지는 효과가 존재함에도 불구하고, 이들이 수학능력시험을 보지 않는 경향이 강하기 때문에, 연령이 수학능력시험 성적에 미치는 효과가 유의미하지 않은 것으로 나타날 수도 있을 것이다.

이러한 가능성을 점검하기 위해서 몇 가지 추가적인 분석 작업을 실시해서 그 결과를 정리한 것이 다음 <표 3>이다. 우선 첫 번째 행에는 연령이 많을수록 인문계 고등학교에 진학할 확률이 높은지를 분석한 결과가 제시되어 있다. 그리고 두 번째 행에는 수학능력시험 중 언어 시험을 보았는지 여부를 종속변수로 했을 때, 연령 변수가 어떠한 계수값을 나타내는지 보여주고 있다. 언어 시험을 보았는지 여부를 기준으로 삼은 이유는, 수학능력시험 중 언어 시험에 응시한 학생 수가 가장 많았기 때문이다. 수리 시험이나 외국어 시험 응시 여부를 기준으로 하더라도 분석 결과는 거의 변하지 않는다. 인문계 고등학교에 진학했는지 여부나 언어 시험을 보았는지 여부는 모두 더미변수이므로, OLS (Linear Probability Model) 방식의 분석에는 추정상의 문제점이 존재할 수 있다. 하지만 Probit 방식으로 분석한 뒤 한계효과를 구해보더라도 여기서의 분석 결과와 거의 차이가 없다. 이 두 결과에서 모든 계수값은 통계적으로 유의하지 않은 것으로 나타난다. 즉 연령이 많을수록 인문계 고등학교에 진학할 확률이 높거나 수학능력시험을 볼 확률이 높아지는 것은 아니라고 판단할 수 있다. 마지막 세 번째 행에서는 수학능력시험을 보는 학생들이 일반 학생들과는 다른 특성을 지닌 계층인지를 확인하기 위해, 언어 시험을 본 학생들만을 대상으로 중학교 3학년 시험 성적에 대한 앞의 <표 2>에서와 같은 분석을 다시 실시하여 그 결과를 정리하였다. 이를 보면 수학능력시험을 본 계층들의 경우에도 중학교 3학년 학업성적에는 연령이, 절대값은 다소 줄어들지만, 대체로 유의한 영향을 미치는 것으로 나타나고 있다. 이러한 분석 결과들에 기초할 때 수학능력시험 성적의 경우 연령의 효과가 유의하지 않은 것으로 나타나는 이유가, 이 시험을 보는지 여부에 대한 선택편의(selection bias) 때문이라 보기는 어렵다고 판단할 수 있다.

(3) 학습 및 생활방식에 대한 분석

그렇다면 왜 고등학교 단계를 경유하면서 연령이 학업성적에 미치는 효과가 사라진 것일까? 물론 여기에는 여러 가지 복합적 요인들이 작용을 하고 있겠지만, 본 논문에서는 이 자료에

대한 추가적인 분석 결과를 기초로 다음과 같은 두 가지 가능성을 제시한다.

첫째, 학습 능력도 누적적인 과정으로 발전하기 때문에, 과거에 성적이 좋은 학생일수록 현재에도 높은 성적을 유지할 가능성이 크다. 하지만 본인의 노력에 따라 과거의 낮은 성적을 극복하는 경우 또한 많다. 대기만성형 학생들은 어디서나 쉽게 관찰된다. 특히 대부분의 학생들이 학업에 보다 집중하고 보다 많은 학습시간을 투여하는 고등학교 시기가 되면, 과거에 성적이 부진했던 학생들이 이를 극복하기 위해 보다 많은 노력을 투입하고, 이를 통해 성적을 만회하는 경향이 나타날 수 있다.

<표 4>에 정리된 몇 가지 지표는 이러한 가능성을 확인시켜준다. 표에서는 지면 관계 상 전체 학생을 대상으로 한 IV 분석 결과만을 제시한다. 표본을 적령 학생층이나 1-2월생을 제외한 적령 학생층으로 한정하여 분석하더라도 거의 유사한 결과가 나타난다. 월평균 사교육비에 대한 분석을 Tobit 모델로 분석하거나, 흡연 경험, 음주 경험, 이성친구 교제 여부에 대한 분석을 Probit 모델로 분석하여도 분석 결과가 달라지지 않는다. 먼저 월평균 사교육비를 종속 변수로 한 회귀분석 결과를 보면, 고등학교 단계에서 연령이 많을수록 사교육비가 적어지는 경향이 뚜렷이 관찰된다. 고등학교 2학년의 경우 연령이 1개월 많을수록 사교육비는 8,900원 가량 적은 것으로 나타나는데, 이는 3월생이 다음해 2월생에 비해 월평균 10만원가량 사교육비 투자가 적다는 것을 의미한다. (고3은 7만원의 차이) 한국과 같이 사교육 비중이 높고 중요한 의미를 가진 교육 환경에서, 연령이 적을수록 사교육비가 유의하게 많았다는 사실은 연령이 학업성적에 미치는 효과를 완화시키는 요인으로 작용하였을 가능성이 있다. 한국의 사교육 문제에 대해서는 다음의 논문을 참조. 남기곤(2008), Baker et al.(2001).

같은 표의 다음 행은 학교나 학원에 가는 것을 제외한 개인적인 학습시간이 연령에 따라 어떠한 차이를 보이는지 분석한 결과이다. 중학교 3학년의 경우 개인 학습시간에 대한 자료는 범주형으로 조사되어 있다. 여기서는 일주일간 학습시간이 5시간 이상인지 여부를 종속변수로 삼았다. 고등학교 1학년 이후의 자료에서는 일주일간 학습시간에 대한 연속변수를 종속변수로 사용하였다. 이를 보면 중학교 3학년과 고등학교 1학년 시기에는 연령이 높을수록 학습시간이 더 긴 것으로 나타나고 있다. 이는 연령이 많은 학생들이 학업성적이 보다 좋아서 공부를 더 열심히 하도록 동기부여가 되었기 때문일 수도 있고, 혹은 연령이 많을수록 사교육에 대한 투자가 적기 때문에 개인적인 학습시간이 많아지는 경향을 보이는 것일 수도 있다. 어쨌든 이러한 경향은 고등학교 2학년과 3학년에는 사라져, 연령이 학습시간에 유의한 영향을 미치지 않는 것으로 나타나고 있다. 절대적인 학습시간이 많아지는 고등학교 단계에서 연령이 적은 학생도 유사한 수준의 학습시간을 투자한다는 사실은, 이들이 과거와는 달리 연령이 많은 학생에 비해 학업성적이 뒤쳐지지 않도록 유도하였을 가능성이 있다.

둘째, 연령이 많아 조기 성숙하다는 사실이 학업성적 향상에 반드시 유리한 것만은 아닐 수도 있다. 오히려 연령이 많을수록 어른들의 세계에 보다 빨리 접하거나 낮은 수준의 탈선 확률이 높아짐으로써, 학습에 대한 집중력이 약화되는 단점도 있을 수 있다. 기존의 연구들 중 학업성적 이외의 측면에서 연령의 효과를 분석하고 있는 연구로는 Dhuey and Lipscomb(2008)과 Fredriksson and Öckert(2005)를 들 수 있다. 전자는 연령이 많은 학생일수록 클럽의 리더가 될

확률이 높다는 사실을 밝히고 있으며, 후자는 연령이 많은 학생일수록 스포츠에 참여하는 경향이 강하다는 사실을 보여주고 있다.

같은 <표 4>를 보면 중학교 3학년에서 고등학교 1학년까지 연령이 많을수록 키가 크고 몸무게가 더 많은 경향이 나타나고 있다. 그러나 이러한 신체적 조기 성숙 경향은 고등학교 2학년 이후에는 사라지고 있다. 학생들의 학습에 방해가 될 수 있는 생활방식으로 흡연과 음주 그리고 이성친구 교제 경험 여부를 상정하고, 이들 변수를 종속변수로 하는 회귀분석을 실시하여 연령 변수의 결과를 정리하였다. 표를 보면 연령이 많을수록 흡연을 경험할 확률이 대체로 높으며(중3, 고2, 고3), 음주 경험 확률이 높아지는 경우도 있고(고2), 이성친구와 교제할 가능성이 높은 경우도 발견된다.(중3) 물론 학년에 따라 그리고 항목에 따라 약간씩의 차이는 있지만, 연령이 많을수록 흡연과 음주 그리고 이성친구 교제 확률이 높아지는 대체적인 경향을 확인할 수 있다. 이들 항목들은 적어도 학업에 대한 집중도를 약화시킬 수 있다는 점에서, 연령이 학업성적에 미치는 효과를 완화시키는 방향으로 작용하고 있을 가능성이 있다고 판단된다. 중학교 3학년 성적을 종속변수로 삼든 혹은 고등학교 3학년 말 수학능력시험 성적을 종속변수로 삼든, 흡연 여부나 음주 여부 그리고 이성친구 교제 여부를 독립변수에 추가할 경우 이들 변수들은 모두 유의한 마이너스의 계수값을 나타내고 있다.

V. 그 외 자료에 대한 분석 결과

(1) TIMSS 자료와 KELS 자료

IEA에서 실시하는 TIMSS는 1995년 이후 매 4년마다 수학과 과학 과목에 대해 독자적인 평가를 실시하는 국제적인 성취도 조사이다. 2007년 조사에서는 59개 국가와 8개의 benchmarking participants가 참가하였는데, 조사 대상 학생 수는 433,785명에 이르고 있다. 한국은 조사 첫 해인 1995년부터 이 조사에 참가하고 있다. 처음 실시된 1995년 조사에서는 3-4학년 학생과 7-8학년 학생을 조사했지만, 1999년 두 번째 조사부터는 한 학년씩에 대해서만 조사를 실시하고 있다. 여기서는 1999년, 2003년, 2007년 한국의 8학년 학생들에 (한국의 학제로는 중학교 3학년에 해당) 대한 학업성적 결과가 학생들의 연령과 어떠한 관련성을 맺고 있는지를 앞 절과 동일한 방식으로 분석한다. TIMSS 자료의 경우 학생과 부모에 대한 설문 조사내용이 다양하지 못하기 때문에, 여기서는 부득이 학생의 성, 아버지와 어머니의 학력에 관한 변수 그리고 학생의 소속 학교에 대한 고정효과만을 통제하였다. 그 외 분석 방식 등은 앞 절의 KEEP 자료에 대한 분석과 동일하다.

이 자료에 대한 분석 결과를 정리한 것이 <표 5>이다. 1999년 자료의 경우 적령 학생층은 1984년 3월생부터 1985년 2월생까지이고, 2003년 자료는 1988년 3월생부터 1989년 2월생까지, 2007년 자료는 1992년 3월생부터 1993년 2월생까지이다. 표를 보면 어느 시기의 자료이거나, 어

떠한 과목이거나, 그리고 어떠한 분석방식을 사용하거나, 연령이 학업성적에 미치는 효과는 대략 0.02 표준편차내외의 일정한 값을 나타내고 있다. 수학 보다는 과학 과목에서 연령의 효과가 다소 강하게 나타나고 있으며, 1999년 자료보다는 2003년 자료 그리고 2007년 자료에서 계수값이 약간씩 커지는 경향을 보이고 있다.

KELS 조사는 한국교육개발원에서 실시하는 전국 단위의 표본 조사이다. 2005년 150개 중학교 1학년 학생 6,908명을 대상으로 조사를 실시한 뒤, 매년 이들에 대해 추적 조사를 실시하고 있다. 중학교 단계인 1-3차년도 조사에서는 학생들에게 국어, 수학, 영어 세 과목에 대해 독자적인 시험을 보게 하고, 그 결과를 자료에 제공하고 있다.

이 자료에 대해서도 마찬가지로 방식으로 연령이 학업성적에 미치는 효과를 분석하고, 그 결과를 정리한 것이 다음 <표 6>이다. 이 자료의 조사 시기 적령 학생층은 1992년 3월생부터 1993년 2월생까지이다. 표를 보면 이 자료에 대한 분석 결과에서도 연령이 학업 성적에 미치는 효과는 대략 0.02 표준편차의 값을 보여주고 있다. 어느 시기나 어느 과목에서든 거의 유사한 값을 보여준다. 학년이 높아질수록 국어 과목은 연령의 효과가 감소되는 추세를 보이지만, 수학과 영어는 거의 비슷한 값을 보여주고 있다.

앞 절의 KEEP 중학생 자료에서 중학교 3학년 성적은 해당 학교 내 전체 석차를 이용하여 분석한 결과였다. 물론 학교에 대한 고정효과를 통제한 상태에서 분석을 하였지만, 학교마다 학생들의 성적 수준이 다르다는 점은 이 분석의 한계였다. 그러나 지금 이 절에서 TIMSS와 KELS 자료와 같이 모든 학생들에게 동일한 시험을 보게 한 결과를 가지고 분석하더라도, 연령이 학업성적에 미치는 효과는 0.02 표준편차라는 거의 유사한 값을 보여주고 있다. 이러한 사실은 앞 절의 분석 결과가 강건할(robust) 수 있음을 보여주는 증거로 해석될 수 있다.

(2) KEEP 고등학생 자료

KEEP 자료에는 중학교 3학년에 대한 자료뿐만 아니라 고등학교 3학년에 대한 자료도 존재한다. 그렇다면 2004년 당시 고등학교 3학년의 경우, 연령이 학교 내의 성적이나 수학능력시험 성적에 미치는 효과는 어떻게 나타나고 있을까? 이에 대한 분석 결과를 정리한 것이 <표 7>이다. 고등학교 자료도 중학교 자료와 조사 내용이 거의 동일한데, 2004년 당시에는 수학능력시험 성적이 등급 외에도 표준점수와 백분위 점수 지표로도 함께 제공되고 있다는 장점이 있다. 중학생 자료의 경우 수학능력시험에 대한 분석은 9개 등급의 중앙값을 이용했다는 한계가 있지만, 고등학생에 대한 자료는 이 문제가 해결되어 있다. 여기서는 수학능력시험의 경우 각 과목의 표준점수를 Z score로 변환하여 종속변수로 사용하였고, 학교 내 성적은 중학생 자료에서와 마찬가지로 학교 전체 석차를 Z score 변환하여 종속변수로 사용하였다.

표를 보면 우선 학교 내 전체 석차의 경우 연령의 효과는 유의하지 않은 것으로 나타나고 있다. 수학능력시험 성적의 경우에도 연령은 플러스의 유의한 값을 가지지 못하였다. 언어와 외국어의 경우 연령이 높을수록 오히려 성적은 낮아지는 추세를 보이고 있고, 수학의 경우에는 연령의 효과가 유의하지 않은 것으로 나타나고 있다. 이러한 사실은 연령이 학업성적에 미

치는 효과가 고등학교 말에는 존재하지 않는다는 앞 절의 분석 결과를 다시 한번 지지해준다.

(3) KYPS 자료

KYPS는 한국청소년개발원에서 조사하는 자료로, 2003년 중학교 2학년인 학생들을 대상으로 조사를 실시한 이후 이들이 고등학교를 졸업한 1년 후 시점까지 6개년간 추적조사를 실시한 자료이다. 1차년도 조사에서 3,449명의 학생과 부모에 대해 조사를 실시한 이후, 고등학교 3학년 단계인 5차년도 조사에서는 2,967명의 학생이 조사되었다. 패널 유지율은 86.0%이다. 이 조사의 적령 학생층은 1989년 3월생부터 1990년 2월생까지이다. 이 조사에서는 학업과 진로, 생활환경과 관련된 다양한 설문도 조사되고 있다. 단 학업 성적의 경우 본인이 반 석차나 학교 전체 석차, 그리고 수학능력시험 등급을 스스로 기록하도록 하고 있어서 측정오차(measurement error)가 존재할 가능성이 있다는 점은 이 자료에 대한 분석 상의 한계이다. KYPS 자료는 학생의 소속 학교에 대한 아이디를 제공하고 있지 않다. 하지만 학교의 거주 지역(구 단위)에 대한 정보가 제공되기 때문에, 여기서의 분석에서는 이를 고정효과로 통제하였다. 1차년도의 경우 학교의 거주지역 아이디 수는 총 100개로, 전체 표집 학교수인 104개와 거의 비슷한 수준이다.

우선 연령이 학업성적에 미치는 효과를 정리한 <표 8>을 보자. 학업성적 자료가 가지는 한계에도 불구하고, 연령 변수의 계수값은 앞에서의 분석 결과와 거의 유사한 값을 나타내고 있다. 중학교 단계에서는 연령이 학업성적에 미치는 효과는 0.02 표준편차 수준의 유의한 값을 나타내고 있다. 반면 고등학교 졸업 단계에서 본 수학능력시험 성적에는 연령이 유의한 영향을 미치지 못하는 것으로 나타나고 있다. 연령이 학업성적에 미치는 유의한 플러스의 효과가 중학교 단계에는 존재하지만 고등학교 졸업단계에는 나타나지 않는다는 앞의 분석 결과를 다시 지지해준다.

다음 <표 9>에서는 KEEP 중학생 자료에서 분석했던 생활방식과 관련된 변수들을 이 자료를 통해 다시 분석하였다. 표를 보면 연령이 많을수록 흡연 경험에는 별다른 영향이 없지만, 음주 경험이 많아지고, 이성친구 교제 가능성이 높아진다는 사실은 뚜렷이 확인된다. 이성 친구교제 변수는 KEEP 자료에서는 이성친구와 교제한 경험이 있는지 여부를 나타내는 변수인데 반해, KYPS 자료에서는 평일이나 휴일에 여가를 주로 이성 친구와 보내는지 여부를 나타내는 변수이다. 이는 설문 내용 상의차이 때문에 기인한 것이다. 음주 경험의 경우 중2, 고1, 고2 자료에서, 이성친구 교제의 경우 중2부터 고2까지 유의한 플러스의 값을 보여주고 있다. 이러한 사실 역시 연령이 많을수록 오히려 학업에 집중하지 못하도록 하는 반작용 요인 또한 함께 존재한다는 앞의 분석 결과를 다시 확인시켜준다.

VI. 맺음말

다른 나라에서와 마찬가지로 한국에서도 중학교 단계까지는 연령이 학업성적에 유의한 영향을 미친다. 하지만 고등학교를 졸업하는 단계에서는 연령의 효과가 존재하지 않는 것으로 나타났다. 본 논문에서 분석하고 있는 KEEP 자료, TIMSS 자료, KELS 자료, KYPS 자료 중 어느 자료를 이용하든, 그리고 Bedard and Dhuey(2006)에서 제안한 도구변수를 이용하든, 아니면 clean education system을 가지고 있는 한국의 특성을 감안하여 해당 학년의 적령 학생층만을 대상으로 하여 분석하든, 거의 유사한 결과를 확인할 수 있었다. 이는 단순히 고등학교 단계에 연령이 적어 학업성적이 부진한 학생들이 보다 빈번하게 고등학교 성적 자료의 표본에서 탈락하기 때문에 나타나는 현상이라고 보기 어려웠다. 오히려 연령이 적은 학생들이 고등학교 단계에 보다 학업에 전념하고 낮은 수준의 탈선을 경험하지 않음으로써, 중학교 단계까지 부진했던 학업성적을 만회하는 경향이 있기 때문인 것으로 판단된다. 스포츠 분야와는 달리 학업 분야에서는 나이가 들어서도 어린 시절의 부진이 충분히 만회될 수 있으며, 또한 연령의 효과가 항상 긍정적인 효과만을 미치는 것은 아닐 수 있음을 한국의 경험은 시사해준다. 스포츠 분야에서 연령의 효과에 대해서는 캐나다 아이스하키 부문에서의 연령 효과를 분석하고 있는 Allen and Barnsley(1993) 등 많은 선행 연구들이 있다.

물론 이러한 현상들이 모든 나라에서 나타나고 있는지는 불분명하다. 한국의 경우에는 중학교 단계나 고등학교 단계에 까지도 학업성적에 따른 학생들의 분리 교육 시스템이 덜 발전되어 있고, 고등학교 단계에 들어서서 학업 경쟁이 치열해지기 때문에, 중학교 단계까지 부진했던 학생이라도 고등학교에 진입한 이후 충분히 만회할 수 있는 기회가 제공될 수 있다. 고등학교 단계에 연령의 효과가 사라지는 한국 자료에 대한 분석 결과는 이와 같은 한국 교육 시스템 상의 특성에 기인하는 것일 수 있다. 따라서 이와는 다른 교육 시스템을 가지고 있는 나라들에서는, 한국과는 다른 양상으로 연령의 효과가 나타날 수 있을 것이다. 중학교 단계를 넘어 고등학교 그리고 대학 입학 이후의 단계에도, 초등학교 입학 단계의 연령의 차이가 유의한 성과의 차이를 유발하는 것인지에 대한 국제적 분석 결과가 축적될 필요가 있다.

❖ 참고문헌 ❖

- 김태훈, 「초등학교 취학나이가 대학진학에 미치는 영향」, 『노동경제논집』, 제34권, 제1호, 2011, 1-32.
- 남기곤, 「사교육시간과 학업성적과의 관련성: PISA 자료를 이용한 국제비교 분석」, 『한국경제학보』, 제15권, 제1호, 2008, 55-90.
- 남기곤, 「명문고 졸업이 임금에 미치는 효과: 실제 존재하는가?」, mimeo, 2012.
- 서우석, 「초등학교 입학연령에 따른 학업성적에 관한 의사 종단적 조사」, 중앙대학교 대학원 석사학위 논문, 1991.
- 이순주·박찬웅, 「출생월에 따른 학교입학연령과 과학영재교육원 영재선발」, 『영재교육연구』, 제18권, 제2호, 2008, 239-264.
- 정은희, 「입학연령에 따른 초등학생의 학업성취도에 관한 연구」, 상지대학교 대학원 석사학위 논문, 2003
- Alexander, K. L., and A. M. Pallas, "Private Schools and Public Policy: New Evidence on Cognitive Achievement in Public and Private Schools," *Sociology of Education*, Vol. 56, No. 4, 1983, 170-182.
- Allen, Jeremiah, and Roger Barnsley, "Streams and Tiers: The Interaction of Ability, Maturity, and Training in Systems with Age-Dependent Recursive Selection," *Journal of Human Resources*, Vol. 28, No. 3, 1993, 649-659.
- Baker, David P., Motoko Akiba, Gerald K. LeTendre, and Alexander W. Wiseman, "Worldwide Shadow Education: Outside-School Learning, Institutional Quality of Schooling, and Cross-National Mathematics Achievement," *Educational Evaluation and Policy Analysis*, Vol. 23, No. 1, 2001, 1-17.
- Bedard, Kelly, and Elizabeth Dhuey, "The Persistence of Early Childhood Maturity: International Evidence of Long-Run Age Effects," Working Paper, 2004. <http://www.economics.uci.edu/docs/colloqpapers/f04/Bedard.pdf>
- Bedard, Kelly, and Elizabeth Dhuey, "The Persistence of Early Childhood Maturity: International Evidence of Long-Run Age Effects," *The Quarterly Journal of Economics*, Vol. 121, No. 4, 2006, 1437-1472.
- Cascio, Elizabeth, and Diane Whitmore Schanzenbach, "First in the Class? Age and the Education Production Function," NBER Working Paper No. 13663, 2007.
- Chung, Bong Gun 1998, "A Study of the High School Leveling Policy in the Republic of Korea: Its Genesis, Implementation and Reforms, 1974-1995," Ed. D. dissertation, University of Hawaii at Manoa.
- Coleman, J., and T. Hoffer, *Public and Private Schools: The Impact of Communities*, New York: Basic Books, 1987.
- Dhuey, Elizabeth, and Stephen Lipscomb, "What Makes a Leader? Relative Age and High School Leadership," *Economics of Education Review*, Vol. 27, No. 2, 2008, 173-183.
- Dobkin, Carlos, and Fernando Ferreira, "Do School Entry Laws Affect Educational Attainment and Labor Market Outcomes?" *Economics of Education Review*, Vol. 29, No. 1, 2010, 40-54.
- Elder, Todd E., and Darren H. Lubotsky, "Kindergarten Entrance Age and Children's Achievement: Impacts of State Policies, Family Background, and Peers," Working Paper, 2008. <http://www.econ>.

wisc.edu/workshop/ELApril4.pdf

- Fredriksson, Peter, and Björn Öckert, “Is Early Learning Really More Productive? The Effect of School Starting Age on School and Labor Market Performance,” IZA Discussion Paper, No. 1659, 2005.
- Kawaguchi, D., “Actual Age at School Entry, Educational Outcomes, and Earnings,” *Journal of the Japanese and International Economies*, Vol. 25, No. 2, 2011, 64 - 80.
- McEwan, Patrick J., and Joseph S. Shapiro, “The Benefits of Delayed Primary School Enrollment: Discontinuity Estimates Using Exact Birth Dates,” *The Journal of Human Resources*, Vol. 43, No. 1, 2008, 1-29.
- Nam, Kigon, “Why Were There So Many February Births in Korea? Are They Successful?,” mimeo, 2011.
- Nam, Kigon, and Kisun Sung, “Are Private Schools More Effective than Public Schools?: Experience from a Natural Experiment in Korea,” *Korean Journal of Labor Economics*, Vol. 32, No. 3, 2009, 91-121.
- Puhani, Patrick A., and Andrea M. Weber, “Does the Early Bird Catch the Worm?: Instrumental Variable Estimates of Early Educational Effects of Age of School Entry in Germany,” *Empirical Economics*, Vol. 32, No. 2-3, 2007, 359-386.
- Stipek, Deborah, “At What Age Should Children Enter Kindergarten? A Question for Policy Makers and Parents,” *Social Policy Report*, Vol. 16, No. 2, 2002, 1-19.
- Strøm, Bjarne, “Student Achievement and Birthday Effects,” Norwegian University of Science and Technology Working Paper, 2004. <http://www.hks.harvard.edu/pepg/PDF/events/Munich/PEPG-04-24Strom.pdf>

〈표 1〉 생년월별 표본 수의 분포: KEEP 중학생 자료

| | 1987 | 1988 | 1989 | 1990 | Total |
|-------|------|------|-------|------|-------|
| Jan | 0 | 0 | 16 | 181 | 197 |
| Feb | 0 | 0 | 20 | 202 | 222 |
| Mar | 0 | 0 | 140 | 18 | 158 |
| Apr | 0 | 0 | 157 | 6 | 163 |
| May | 0 | 0 | 151 | 2 | 153 |
| Jun | 0 | 0 | 130 | 0 | 130 |
| Jul | 0 | 0 | 140 | 0 | 140 |
| Aug | 0 | 1 | 155 | 0 | 156 |
| Sep | 0 | 0 | 166 | 0 | 166 |
| Oct | 0 | 1 | 162 | 0 | 163 |
| Nov | 0 | 1 | 171 | 0 | 172 |
| Dec | 1 | 1 | 172 | 0 | 174 |
| Total | 1 | 4 | 1,580 | 409 | 1,994 |

〈표 2〉 KEEP 중학생 자료: 성적에 대한 분석

| Subject | | 〈분석대상〉 〈분석방식〉 Variables | 전체 IV (1) | 1989. Mar. ~1990. Feb. OLS (2) | 1989. Mar. ~ 1989. Dec. OLS (3) |
|---------|---------|-------------------------------|---------------------|---|--|
| 중3 성적 | | ObservedAge | 0.025*** (0.008) | 0.022*** (0.007) | 0.016* (0.010) |
| | | N. of Observations | 1,773 | 1,716 | 1,370 |
| | | R-squared | 0.1940 | 0.2029 | 0.2054 |
| 수능 성적 | Korean | ObservedAge | -0.002 (0.009) | -0.003 (0.008) | -0.015^ (0.010) |
| | | N. of Observations | 1,309 | 1,263 | 1,013 |
| | | R-squared | 0.2238 | 0.2365 | 0.2664 |
| | Math | ObservedAge | 0.001 (0.009) | -0.000 (0.008) | -0.009 (0.011) |
| | | N. of Observations | 1,229 | 1,186 | 956 |
| | | R-squared | 0.1769 | 0.1829 | 0.2114 |
| | English | ObservedAge | 0.005 (0.009) | 0.002 (0.008) | -0.008 (0.011) |
| | | N. of Observations | 1,298 | 1,252 | 1,003 |
| | | R-squared | 0.2631 | 0.2730 | 0.2860 |

Notes: 모든 모델에는 가중치를 적용함. 괄호 안은 robust standard errors. ^는 15%, *는 10%, **는 5%, ***는 1% 수준에서 통계적으로 유의함을 의미. 모델 (1)은 '할당된 상대 연령' 을 도구변수로 사용함. 모든 모델은 〈부표 2〉에 제시한 변수들과 소속 학교에 대한 고정효과를 통제함. 수능능력시험 성적에 대한 분석에서는 인문계 고등학교에 다니고 있는지 여부를 나타내는 더미변수도 추가로 통제함. 독립 변수에 missing 값이 있을 경우 0으로 대체한 후, 각 변수에 대해 missing 값을 가지는지에 여부를 나타내는 더미변수를 통제함.

〈표 3〉 KEEP 중학생 자료: 성적 분석과 관련된 점검

| Dependent Variable | 〈분석대상〉 〈분석방식〉 Variables | 전체 | 1989. Mar. ~1990. Feb. | 1989. Mar. ~ 1989. Dec. |
|---------------------------------------|-------------------------------|--------------------|---------------------------|----------------------------|
| | | IV (1) | OLS (2) | OLS (3) |
| 인문계 고등학교 진학 여부에 대한 더미변수 (고1 시점) | ObservedAge | 0.002 (0.005) | -0.001 (0.004) | 0.001 (0.005) |
| | N. of Observations | 1,769 | 1,712 | 1,371 |
| | R-squared | 0.2338 | 0.2375 | 0.2365 |
| 수능 Korean 성적이 있는지 여부에 대한 더미변수 | ObservedAge | 0.001 (0.003) | 0.001 (0.003) | -0.001 (0.004) |
| | N. of Observations | 1,994 | 1,927 | 1,544 |
| | R-squared | 0.3380 | 0.3317 | 0.3385 |
| 중3 성적(수능 Korean 성적 있는 사람만) | ObservedAge | 0.019** (0.009) | 0.018** (0.008) | 0.015^ (0.010) |
| | N. of Observations | 1,170 | 1,130 | 904 |
| | R-squared | 0.1702 | 0.1750 | 0.1887 |

Notes: 모든 모델에는 가중치를 적용함. 괄호 안은 robust standard errors. ^는 15%, *는 10%, **는 5%, ***는 1% 수준에서 통계적으로 유의함을 의미. 모델 (1)은 '할당된 상대 연령' 을 도구변수로 사용함. 모든 모델은 〈부표 2〉에 제시한 변수들과 소속 학교에 대한 고정효과를 통제함. 수학능력시험 성적에 대한 분석에서는 인문계 고등학교에 다니고 있는지 여부를 나타내는 더미변수도 추가로 통제함. 독립 변수에 missing 값이 있을 경우 0으로 대체한 후, 각 변수에 대해 missing 값을 가지는지에 여부를 나타내는 더미변수를 통제함.

〈표 4〉 KEEP 중학생 자료: 학습 및 생활방식 분석

| Dependent Variable | 〈분석대상〉 Variables | 중 3 | 고 1 | 고 2 | 고 3 |
|--------------------------|---------------------|---------------------|-------------------|----------------------|--------------------|
| | | (1) | (2) | (3) | (4) |
| 월평균 사교육비 (지난 학기) | ObservedAge | -0.119 (0.162) | -0.343 (0.252) | -0.890*** (0.327) | -0.603* (0.341) |
| | N. of Observations | 1,938 | 1,501 | 1,530 | 1,391 |
| | R-squared | 0.3987 | 0.4660 | 0.3890 | 0.4638 |
| 일주일간 학습시간 (학교나 학원 제외) | ObservedAge | 0.006* (0.003) | 0.106* (0.064) | 0.078 (0.081) | 0.022 (0.117) |
| | N. of Observations | 1,994 | 1,749 | 1,730 | 1,673 |
| | R-squared | 0.1313 | 0.2806 | 0.2805 | 0.3246 |
| 신장 | ObservedAge | 0.176*** (0.049) | 0.086 (0.072) | -0.003 (0.050) | -0.003 (0.051) |
| | N. of Observations | 1,994 | 1,760 | 1,743 | 1,687 |
| | R-squared | 0.4380 | 0.6243 | 0.6495 | 0.6576 |
| 몸무게 | ObservedAge | 0.151* (0.086) | 0.203* (0.109) | 0.110 (0.090) | 0.041 (0.095) |
| | N. of Observations | 1,991 | 1,749 | 1,738 | 1,677 |
| | R-squared | 0.2114 | 0.3808 | 0.3947 | 0.4272 |
| 흡연 경험 | ObservedAge | 0.002^ (0.001) | 0.002 (0.004) | 0.005* (0.003) | 0.006** (0.003) |
| | N. of Observations | 1,994 | 1,761 | 1,746 | 1,691 |
| | R-squared | 0.0621 | 0.2347 | 0.1409 | 0.1570 |
| 음주 경험 | ObservedAge | 0.003 (0.003) | -0.001 (0.006) | 0.007^ (0.005) | 0.001 (0.005) |
| | N. of Observations | 1,994 | 1,760 | 1,746 | 1,690 |
| | R-squared | 0.0693 | 0.2165 | 0.1299 | 0.2085 |

| Dependent Variable | <분석대상> Variables | 중 3 (1) | 고 1 (2) | 고 2 (3) | 고 3 (4) |
|--------------------|---------------------|-------------------|------------------|-------------------|-------------------|
| 이성친구 교제 | ObservedAge | 0.005* (0.003) | 0.006 (0.005) | -0.000 (0.003) | -0.001 (0.004) |
| | N. of Observations | 1,994 | 1,765 | 1,746 | 1,677 |
| | R-squared | 0.0930 | 0.1983 | 0.1228 | 0.1245 |

Notes: 전체 표본을 대상으로 한 IV 모델 분석 결과임. 통제변수 등 모든 설명은 <표 3>과 동일함. 고등학생에 대한 분석에서는 인문계 고등학교에 다니고 있는지 여부를 나타내는 더미변수도 추가로 통제함.

<표 5> TIMSS자료: 성적에 대한 분석 (중 3)

| 과목 | <자료> | TIMSS 1999 | | | TIMSS 2003 | | | TIMSS 2007 | | |
|------|-------------------------------|---------------------|---------------------------------------|---------------------------------------|---------------------|---------------------------------------|---------------------------------------|---------------------|---------------------------------------|---------------------------------------|
| | <분석대상> <분석방식> Variables | 전체 IV (1) | 1984. Mar. ~ 1985. Feb. OLS (2) | 1984. Mar. ~ 1984. Dec. OLS (3) | 전체 IV (4) | 1988. Mar. ~ 1989. Feb. OLS (5) | 1988. Mar. ~ 1988. Dec. OLS (6) | 전체 IV (7) | 1992. Mar. ~ 1993. Feb. OLS (8) | 1992. Mar. ~ 1992. Dec. OLS (9) |
| MATH | ObservedAge | 0.012*** (0.004) | 0.009** (0.004) | 0.011** (0.005) | 0.019*** (0.004) | 0.016*** (0.004) | 0.017*** (0.005) | 0.021*** (0.005) | 0.013*** (0.004) | 0.016*** (0.006) |
| | N. of Observations | 6,114 | 5,885 | 4,591 | 5,309 | 5,185 | 4,170 | 4,240 | 4,032 | 3,384 |
| | R-squared | 0.1286 | 0.1366 | 0.1400 | 0.1393 | 0.1462 | 0.1506 | 0.1570 | 0.1742 | 0.1822 |
| SCI | ObservedAge | 0.020*** (0.004) | 0.017*** (0.004) | 0.022*** (0.005) | 0.020*** (0.004) | 0.017*** (0.004) | 0.018*** (0.005) | 0.029*** (0.005) | 0.022*** (0.004) | 0.025*** (0.006) |
| | N. of Observations | 6,114 | 5,885 | 4,591 | 5,309 | 5,185 | 4,170 | 4,238 | 4,030 | 3,382 |
| | R-squared | 0.1243 | 0.1348 | 0.1417 | 0.1200 | 0.1222 | 0.1296 | 0.1417 | 0.1549 | 0.1655 |

Notes: 모든 모델에는 가중치를 적용함. 괄호 안은 robust standard errors. ^는 15%, *는 10%, **는 5%, ***는 1% 수준에서 통계적으로 유의함을 의미. 모델 (1), (4), (7)은 '할당된 상대 연령' 을 도구변수로 사용함. 모든 모델은 학생의 성, 아버지의 학력이 전문대학 이상인지 여부, 어머니의 학력이 전문대학 이상인지 여부와 소속 학교에 대한 고정효과를 통제함. 독립 변수에 missing 값이 있을 경우 0으로 대체한 후, 각 변수에 대해 missing 값을 가지는지에 여부를 나타내는 더미변수를 추가로 통제함.

<표 6> KELS자료: 성적에 대한 분석

| 시험시기 | <Subject> | KOR | | | MATH | | | ENG | | |
|------|-------------------------------|---------------------|---------------------------------------|---------------------------------------|---------------------|---------------------------------------|---------------------------------------|---------------------|---------------------------------------|---------------------------------------|
| | <분석대상> <분석방식> Variables | 전체 IV (1) | 1992. Mar. ~ 1993. Feb. OLS (2) | 1992. Mar. ~ 1992. Dec. OLS (3) | 전체 IV (4) | 1992. Mar. ~ 1993. Feb. OLS (5) | 1992. Mar. ~ 1992. Dec. OLS (6) | 전체 IV (7) | 1992. Mar. ~ 1993. Feb. OLS (8) | 1992. Mar. ~ 1992. Dec. OLS (9) |
| 중 1 | ObservedAge | 0.023*** (0.004) | 0.020*** (0.003) | 0.019*** (0.004) | 0.019*** (0.004) | 0.017*** (0.003) | 0.015*** (0.004) | 0.016*** (0.004) | 0.014*** (0.003) | 0.011*** (0.004) |
| | N. of Observations | 6,721 | 6,431 | 5,388 | 6,721 | 6,431 | 5,390 | 6,789 | 6,497 | 5,443 |
| | R-squared | 0.2257 | 0.2294 | 0.2267 | 0.2217 | 0.2262 | 0.2313 | 0.3174 | 0.3199 | 0.3163 |
| 중 2 | ObservedAge | 0.019*** (0.004) | 0.016*** (0.003) | 0.013*** (0.004) | 0.019*** (0.004) | 0.016*** (0.003) | 0.019*** (0.004) | 0.017*** (0.004) | 0.014*** (0.003) | 0.015*** (0.004) |
| | N. of Observations | 6,385 | 6,109 | 5,112 | 6,404 | 6,129 | 5,128 | 6,397 | 6,121 | 5,123 |
| | R-squared | 0.2362 | 0.2398 | 0.2477 | 0.2314 | 0.2345 | 0.2469 | 0.3008 | 0.3061 | 0.3130 |
| 중 3 | ObservedAge | 0.010*** (0.004) | 0.008** (0.003) | 0.002 (0.004) | 0.019*** (0.004) | 0.015*** (0.003) | 0.014*** (0.004) | 0.020*** (0.004) | 0.016*** (0.003) | 0.015*** (0.004) |
| | N. of Observations | 6,227 | 5,969 | 5,005 | 6,506 | 6,228 | 5,217 | 6,232 | 5,974 | 5,007 |
| | R-squared | 0.2365 | 0.2395 | 0.2399 | 0.2238 | 0.2270 | 0.2330 | 0.2661 | 0.2728 | 0.2757 |

Notes: 괄호 안은 robust standard errors. ^는 15%, *는 10%, **는 5%, ***는 1% 수준에서 통계적으로 유의함을 의미. 모델 (1), (4), (7)은 '할당된 상대 연령' 을 도구변수로 사용함. 모든 모델은 <부표 2>에 제시한 변수들과 소속 학교에 대한 고정효과를 통제함. 독립 변수에 missing 값이 있을 경우 0으로 대체한 후, 각 변수에 대해 missing 값을 가지는지에 여부를 나타내는 더미변수를 추가로 통제함.

〈표 7〉 KEEP 고등학생 자료: 성적에 대한 분석

| Subject | | <분석대상> <분석방식> Variables | 전체 IV (1) | 1986. Mar. ~1987. Feb. OLS (2) | 1986. Mar. ~ 1986. Dec. OLS (3) |
|---------|---------|-------------------------------|--------------------------------|---|--|
| 고3 성적 | | ObservedAge | -0.003 (0.006) | -0.003 (0.006) | -0.009 (0.008) |
| | | N. of Observations | 3,011 | 2,927 | 2,306 |
| | | R-squared | 0.0990 | 0.1028 | 0.1243 |
| 수능 성적 | Korean | ObservedAge | -0.009 [^] (0.006) | -0.010* (0.005) | -0.017** (0.007) |
| | | N. of Observations | 2,182 | 2,132 | 1,676 |
| | | R-squared | 0.4259 | 0.4232 | 0.4372 |
| | Math | ObservedAge | -0.007 (0.007) | -0.007 (0.006) | -0.008 (0.009) |
| | | N. of Observations | 1,941 | 1,902 | 1,483 |
| | | R-squared | 0.2613 | 0.2667 | 0.2817 |
| | English | ObservedAge | -0.010* (0.006) | -0.011* (0.005) | -0.019** (0.008) |
| | | N. of Observations | 2,170 | 2,121 | 1,665 |
| | | R-squared | 0.4117 | 0.4068 | 0.4169 |

Notes: 모든 모델에는 가중치를 적용함. 괄호 안은 robust standard errors. [^]는 15%, *는 10%, **는 5%, ***는 1% 수준에서 통계적으로 유의함을 의미. 모델 (1)은 '할당된 상대 연령' 을 도구변수로 사용함. 모든 모델은 <부표 2>에 제시한 변수들과 소속 학교에 대한 고정효과를 통제함. 인문계 고등학교에 다니고 있는지 여부를 나타내는 더미변수도 추가로 통제함. 독립변수에 missing 값이 있을 경우 0으로 대체한 후, 각 변수에 대해 missing 값을 가지는지에 여부를 나타내는 더미변수를 통제함.

〈표 8〉 KYPS자료: 성적에 대한 분석

| Subject | | <분석대상> <분석방식> Variables | 전체 IV (1) | 1989. Mar. ~1990. Feb. OLS (2) | 1989. Mar. ~ 1989. Dec. OLS (3) |
|---------------------|--|-------------------------------|---------------------|---|--|
| 중2 성적 (반 석차 기준) | | ObservedAge | 0.030*** (0.006) | 0.025*** (0.005) | 0.026*** (0.007) |
| | | N. of Observations | 3,293 | 3,170 | 2,507 |
| | | R-squared | 0.1486 | 0.1515 | 0.1526 |
| 중2 성적 (전교 석차 기준) | | ObservedAge | 0.025*** (0.006) | 0.021*** (0.005) | 0.024*** (0.007) |
| | | N. of Observations | 3,050 | 2,928 | 2,318 |
| | | R-squared | 0.1574 | 0.1591 | 0.1676 |
| 고등학교 졸업 당시 수능성적 | | ObservedAge | -0.000 (0.006) | -0.001 (0.005) | -0.003 (0.007) |
| | | N. of Observations | 1,942 | 1,878 | 1,484 |
| | | R-squared | 0.2502 | 0.2513 | 0.2646 |

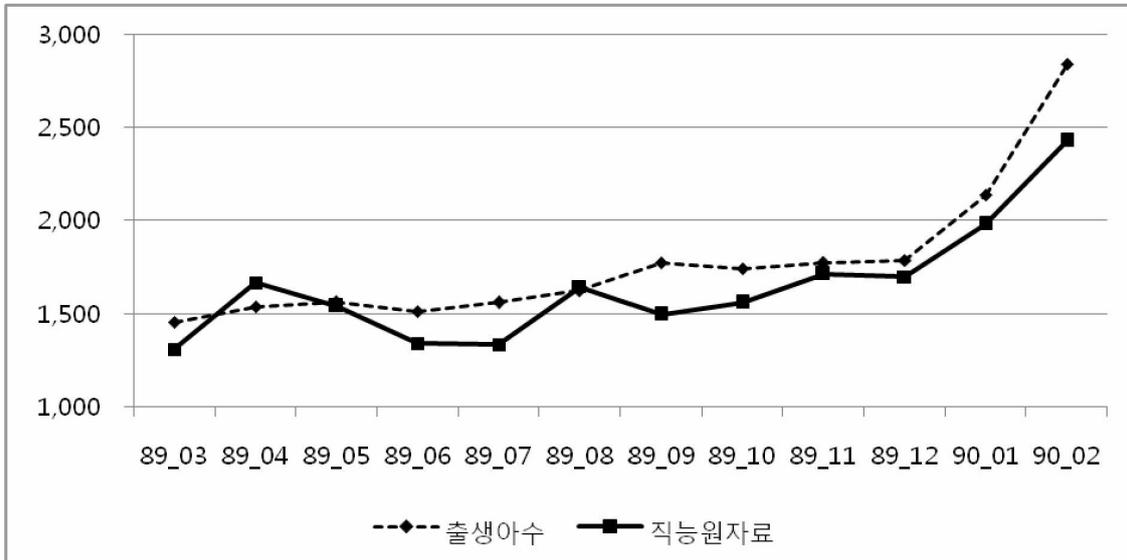
Notes: 모든 모델에는 가중치를 적용함. 괄호 안은 robust standard errors. [^]는 15%, *는 10%, **는 5%, ***는 1% 수준에서 통계적으로 유의함을 의미. 모델 (1)은 '할당된 상대 연령' 을 도구변수로 사용함. 모든 모델은 <부표 2>에 제시한 변수들과 소속 학교의 거주지역(구 단위)에 대한 고정효과를 통제함. 수학적능력시험 성적에 대한 분석에서는 인문계 고등학교에 다니고 있는지 여부를 나타내는 더미변수도 추가로 통제함. 독립 변수에 missing 값이 있을 경우 0으로 대체한 후, 각 변수에 대해 missing 값을 가지는지에 여부를 나타내는 더미변수를 통제함.

〈표 9〉 KYPs자료: 생활방식 분석

| Dependent Variable | <분석대상> Variables | 중 2 (1) | 중 3 (2) | 고 1 (3) | 고 2 (4) | 고 3 (5) |
|--------------------|---------------------|---------------------|---------------------|---------------------|--------------------|-------------------|
| 흡연 경험 | ObservedAge | 0.003 (0.002) | 0.002 (0.002) | 0.001 (0.002) | -0.001 (0.002) | -0.002 (0.002) |
| | N. of Observations | 3,449 | 3,188 | 3,125 | 3,121 | 2,958 |
| | R-squared | 0.0811 | 0.0600 | 0.1165 | 0.1546 | 0.1562 |
| 음주 경험 | ObservedAge | 0.009*** (0.003) | 0.001 (0.003) | 0.008*** (0.003) | 0.005^ (0.003) | 0.002 (0.003) |
| | N. of Observations | 3,449 | 3,188 | 3,125 | 3,121 | 2,958 |
| | R-squared | 0.0851 | 0.0698 | 0.0969 | 0.1046 | 0.1021 |
| 이성친구 교제 | ObservedAge | 0.004** (0.002) | 0.010*** (0.002) | 0.004* (0.002) | 0.005** (0.002) | 0.003 (0.002) |
| | N. of Observations | 3,447 | 3,188 | 3,123 | 3,120 | 2,958 |
| | R-squared | 0.0437 | 0.0550 | 0.0772 | 0.0852 | 0.0747 |

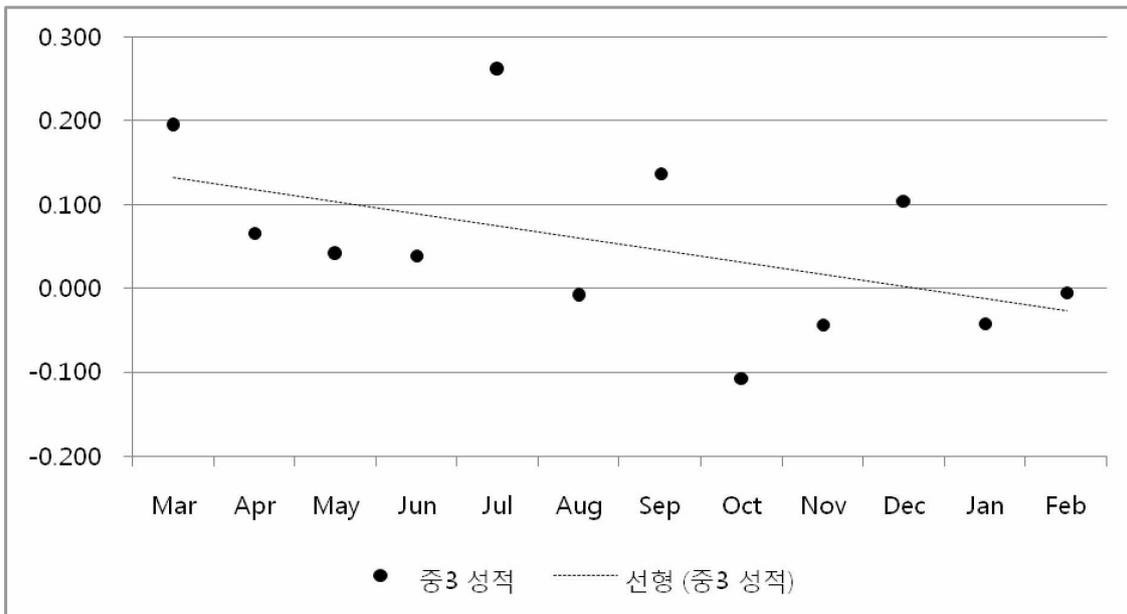
Notes: 모든 모델에는 가중치를 적용함. 괄호 안은 robust standard errors. ^는 15%, *는 10%, **는 5%, ***는 1% 수준에서 통계적으로 유의함을 의미. 모델 (1)은 '할당된 상대 연령' 을 도구변수로 사용함. 모든 모델은 <부표 2>에 제시한 변수들과 소속 학교의 거주지역(구 단위)에 대한 고정효과를 통제함. 고등학생에 대한 분석에서는 인문계 고등학교에 다니고 있는지 여부를 나타내는 더미변수도 추가로 통제함. 독립 변수에 missing 값이 있을 경우 0으로 대체한 후, 각 변수에 대해 missing 값을 가지는지에 여부를 나타내는 더미변수를 통제함.

부 록



Notes: 출생아 수는 통계청의 인구동태연보 자료임. 직능원 자료는 KEEP 중학생 자료에서 가중치를 이용하여 추산한 것임. 각 월의 총 출생아 수를 해당 월의 일수로 나눈 값임.

[부록 그림 1] 생년월별 출생아 수와 표본 수의 비교



Notes: 각 생년월별 평균 성적(Z score). <부표 1>의 자료를 도표화한 것임.

[부록 그림 2] 생년월별 평균 성적 (KEEP 중학생 자료, 중3성적)

〈부표 1〉 생년월별 성적 및 부모 학력 분포

| Birth Month | 중3 성적 | 수능 성적 | | | 전문대이상 학력 | |
|-------------|--------|--------|--------|---------|----------|-------|
| | | Korean | Math | English | 아버지 | 어머니 |
| Mar | 0.194 | -0.184 | -0.121 | -0.147 | 0.272 | 0.133 |
| Apr | 0.066 | -0.271 | -0.358 | -0.189 | 0.270 | 0.160 |
| May | 0.041 | -0.373 | -0.252 | -0.459 | 0.268 | 0.190 |
| Jun | 0.039 | -0.181 | 0.058 | -0.095 | 0.231 | 0.154 |
| Jul | 0.261 | -0.113 | -0.114 | -0.123 | 0.300 | 0.143 |
| Aug | -0.009 | -0.058 | -0.035 | -0.065 | 0.314 | 0.179 |
| Sep | 0.136 | -0.081 | -0.113 | -0.132 | 0.277 | 0.120 |
| Oct | -0.107 | -0.116 | -0.145 | -0.211 | 0.356 | 0.166 |
| Nov | -0.044 | -0.146 | -0.253 | -0.130 | 0.320 | 0.157 |
| Dec | 0.104 | -0.115 | -0.133 | -0.203 | 0.259 | 0.121 |
| Jan | -0.042 | -0.118 | -0.080 | -0.138 | 0.355 | 0.178 |
| Feb | -0.005 | -0.127 | -0.160 | -0.142 | 0.306 | 0.189 |
| Total | 0.044 | -0.153 | -0.144 | -0.167 | 0.296 | 0.158 |

〈부표 2〉 변수의 정의 및 평균값

| Variables | Definition | Mean | Std. Dev. |
|--------------|------------------------|-------|-----------|
| ObservedAge | 연령(일/30) | 185.4 | 3.8 |
| Male | 남학생=1 | 0.497 | 0.500 |
| DadCollege | 아버지 학력이 전문대학 이상=1 | 0.296 | 0.457 |
| MomCollege | 어머니 학력이 전문대학 이상=1 | 0.158 | 0.365 |
| FirstChild | 첫째 자녀=1 | 0.508 | 0.500 |
| ParentsExist | 친부와 친모인 경우=1 | 0.868 | 0.338 |
| Earnings | 지난 1년간 월평균 가구소득(백만원단위) | 2.915 | 1.956 |
| OwnHouse | 자가인 경우=1 | 0.657 | 0.475 |

〈부표 3〉 KEEP 중학생자료: 중3성적에 대한 분석에 대한 전체 결과

| <분석대상> | 전체 | 1989. Mar. ~ 1990. Feb. | 1989. Mar. ~ 1989. Dec. |
|--------------|----------------------|-------------------------|-------------------------|
| <분석방식> | IV | OLS | OLS |
| Variables | (1) | (2) | (3) |
| ObservedAge | 0.025*** (0.008) | 0.022*** (0.007) | 0.016* (0.010) |
| Male | -0.229*** (0.063) | -0.245*** (0.064) | -0.227*** (0.073) |
| DadCollege | 0.262*** (0.067) | 0.283*** (0.070) | 0.207** (0.083) |
| MomCollege | 0.309*** (0.082) | 0.318*** (0.086) | 0.325*** (0.103) |
| FirstChild | 0.150*** (0.049) | 0.139*** (0.050) | 0.157*** (0.060) |
| ParentsExist | 0.418*** (0.151) | 0.437*** (0.160) | 0.502*** (0.171) |

| <분석대상> | 전체 | 1989. Mar. ~1990. Feb. | 1989. Mar. ~ 1989. Dec. |
|--------------------|----------------------|------------------------|-------------------------|
| <분석방식> | IV | OLS | OLS |
| Variables | (1) | (2) | (3) |
| Earnings | 0.056*** (0.014) | 0.055*** (0.014) | 0.054*** (0.016) |
| OwnHouse | 0.193*** (0.053) | 0.186*** (0.055) | 0.179*** (0.063) |
| Constant | -5.401*** (1.502) | -4.882*** (1.325) | -3.766** (1.823) |
| N. of Observations | 1,773 | 1,716 | 1,370 |
| R-squared | 0.1940 | 0.2029 | 0.2054 |

Notes: 모든 모델에는 가중치를 적용함. 괄호 안은 robust standard errors. ^는 15%, *는 10%, **는 5%, ***는 1% 수준에서 통계적으로 유의함을 의미. 모델 (1)은 '할당된 상대 연령' 을 도구변수로 사용함. 모든 모델은 소속 학교에 대한 고정효과를 통제함. 독립 변수에 missing 값이 있을 경우 0으로 대체한 후, 각 변수에 대해 missing 값을 가지는지에 여부를 나타내는 더미변수를 통제함.

<부표 4> KEEP 중학생자료: 수능성적에 대한 분석에 대한 전체 결과

| <Subject> | Korean | | | Math | | | English | | |
|--------------------|---------------------|------------------------|------------------------|---------------------|------------------------|------------------------|---------------------|------------------------|------------------------|
| | 전체 | 1989. Mar. ~1990. Feb. | 1989. Mar. ~1989. Dec. | 전체 | 1989. Mar. ~1990. Feb. | 1989. Mar. ~1989. Dec. | 전체 | 1989. Mar. ~1990. Feb. | 1989. Mar. ~1989. Dec. |
| <분석방식> | IV | OLS | OLS | IV | OLS | OLS | IV | OLS | OLS |
| Variables | (1) | (2) | (3) | (1) | (2) | (3) | (1) | (2) | (3) |
| ObservedAge | -0.002 (0.009) | -0.003 (0.008) | -0.015^ (0.010) | 0.001 (0.009) | -0.000 (0.008) | -0.009 (0.011) | 0.005 (0.009) | 0.002 (0.008) | -0.008 (0.011) |
| Male | -0.146** (0.070) | -0.135* (0.072) | -0.123^ (0.081) | 0.018 (0.076) | 0.021 (0.076) | -0.029 (0.089) | -0.149** (0.068) | -0.145** (0.069) | -0.138* (0.079) |
| GeneralHigh | 0.576*** (0.083) | 0.584*** (0.084) | 0.641*** (0.089) | 0.491*** (0.084) | 0.503*** (0.084) | 0.533*** (0.100) | 0.535*** (0.078) | 0.558*** (0.078) | 0.654*** (0.086) |
| DadCollege | 0.213*** (0.074) | 0.215*** (0.077) | 0.081 (0.086) | 0.178** (0.075) | 0.182** (0.077) | 0.129 (0.091) | 0.306*** (0.070) | 0.306*** (0.073) | 0.242*** (0.085) |
| MomCollege | 0.304*** (0.090) | 0.323*** (0.093) | 0.355*** (0.104) | 0.016 (0.097) | 0.016 (0.098) | -0.006 (0.114) | 0.315*** (0.088) | 0.340*** (0.091) | 0.287*** (0.107) |
| FirstChild | 0.093* (0.056) | 0.070 (0.057) | 0.085 (0.064) | 0.103* (0.059) | 0.100* (0.060) | 0.089 (0.069) | 0.093* (0.055) | 0.080 (0.057) | 0.071 (0.066) |
| ParentsExist | 0.180^ (0.124) | 0.248** (0.124) | 0.222* (0.134) | 0.147 (0.172) | 0.151 (0.174) | 0.144 (0.187) | 0.241** (0.110) | 0.284** (0.113) | 0.293** (0.130) |
| Earnings | -0.006 (0.010) | -0.008 (0.010) | -0.010 (0.010) | 0.022** (0.009) | 0.019** (0.009) | 0.015^ (0.010) | 0.005 (0.008) | 0.003 (0.008) | 0.006 (0.009) |
| OwnHouse | 0.146** (0.067) | 0.143** (0.069) | 0.144* (0.078) | 0.122^ (0.079) | 0.108 (0.079) | 0.085 (0.089) | 0.163** (0.066) | 0.138** (0.068) | 0.130^ (0.081) |
| Constant | -0.535 (1.751) | -0.371 (1.515) | 1.736 (1.916) | -0.627 (1.766) | -0.405 (1.527) | 1.186 (2.126) | -1.949 (1.738) | -1.603 (1.490) | 0.172 (1.986) |
| N. of Observations | 1,309 | 1,263 | 1,013 | 1,229 | 1,186 | 956 | 1,298 | 1,252 | 1,003 |
| R-squared | 0.2238 | 0.2365 | 0.2664 | 0.1769 | 0.1829 | 0.2114 | 0.2631 | 0.2730 | 0.2860 |

Notes: 모든 모델에는 가중치를 적용함. 괄호 안은 robust standard errors. ^는 15%, *는 10%, **는 5%, ***는 1% 수준에서 통계적으로 유의함을 의미. 모델 (1)은 '할당된 상대 연령' 을 도구변수로 사용함. 모든 모델은 소속 학교에 대한 고정효과를 통제함. 인문계 고등학교에 다니고 있는지 여부를 나타내는 더미변수도 추가로 통제함. 독립 변수에 missing 값이 있을 경우 0으로 대체한 후, 각 변수에 대해 missing 값을 가지는지에 여부를 나타내는 더미변수를 통제함.

❖ Abstract ❖

Until When Does the Effect of Age on Academic Achievements Persist: Experience of South Korea

Kigon Nam

This paper analyze the effect of age on test scores using KEEP(Korean Education and Employment Panel) middle school student data. Also TIMSS(Trends in International Mathematics and Science Study) 1999, 2003, 2007 data, KELS(Korean Education Longitudinal Study) data, and KYPS(Korea Youth Panel Survey) data are analyzed to confirm the results.

In the stage of middle school, the test score tends to increase around 0.02 standard deviation as a student is one month older. In the stage of high school, however, the effect of age is not statistically significant. This paper proposes two possibilities as the reason of this change. First, younger student with low test score makes more efforts in high school period when the competition among students is very severe. Second, older student tends to contact with an adult world in earlier times. As the evidence of the former possibility, this paper shows the empirical results on private tutoring cost and studying time. And the results on the experience of smoking, alcoholic drinking, and intersexual friendship are shown as the evidence of the latter possibility.

Keywords: age effect, testscore, instrument variable, clean education system