

고등학교 시기의 사교육 경험이 대학생의 사교육 참여 및 사교육비 지출에 미치는 영향 분석

김동일(金東一)*

김영식(金永植)**

김경선(金京嬋)***

논문 요약

본 연구는 고등학교 시기의 사교육 참여 경험이 대학생이 되고 난 이후의 사교육 참여 및 사교육비 지출에 미치는 영향을 살펴보기 위해 수행되었다. 이를 위해 고등학교 3학년 시기의 사교육 참여 집단과 미참여 집단 간에 존재하는 관찰 가능한 특성의 차이를 통제하기 위하여 산출한 성향점수(Propensity Score)를 개인 수준의 통제변수로 활용하여 고등학교 시기의 사교육 참여 경험이 대학생의 사교육 참여 및 사교육비 지출에 미치는 영향을 탐색하였다. 또한, 대학생의 사교육 참여 및 사교육비 지출에 영향을 미치는 대학생 개인 수준 변수의 영향력뿐만 아니라, 개인이 소속된 대학교의 영향력을 함께 고려하기 위하여 위계적 선형모형을 활용하여 연구를 수행하였다.

분석 결과 고3 시기의 사교육경험, 대학입학 이후의 경과 연수, 고 3 시기 사교육 참여 여부에 대한 성향점수가 대학생의 사교육 참여에 정적으로 유의미한 영향력을 미치고 있음을 확인할 수 있다. 또한 2수준 변수인 대학의 소재지, 대학 순위도 대학생의 사교육 참여와 인과관계를 맺고 있음을 확인할 수 있었다. 이와 달리 대학생의 사교육비 지출에 영향을 미친 변수로는 대학입학 이후의 경과 연수와 고 3 시기의 사교육 참여 경험 등의 개인 수준 변수임이 확인되었다. 이러한 연구결과는 대학생의 사교육 참여 및 사교육비 지출에 있어 고교시절의 사교육 참여 경험이 유의미한 영향을 미치고 있음을 보여준다.

주요어 : 사교육, 성향점수(Propensity Score), 위계선형모형(HLM)

* 서울대학교 교육학과 교수

** 서울대학교 교육학과 박사과정

*** 교신저자, 서울대학교 교육학과 박사과정

I. 서론

우리나라 중, 고등학생들의 과다한 사교육비 지출은 학부모의 부담 증가, 교육투자의 양극화로 인한 사회계층간 불평등 심화, 공교육 무력화 등의 개인적·사회적 폐단을 초래한다는 우려 하에 오랜 동안 이슈화되어왔고, 이에 대한 해결책 제시를 위한 국가적 차원의 다양한 정책적 노력이 시도되었다. 그러나 상대적으로 대학생들의 사교육비 지출은 그 규모 면에서 학부모들에게 엄청난 부담이 되고 있음에도 불구하고 그동안 사회적, 국가적 관심사에서 다소 제외되어왔다. 2010년 취업포털 잡코리아가 국내 4년제 대학에 재(휴)학 중인 2~4학년생 812명을 대상으로 조사한 결과에 따르면 현재 취업을 위해 사교육을 받고 있다고 응답한 학생이 전체의 61.5%에 달했으며, 학년별로는 4학년이 가장 높은 68.4%를 차지했고 3학년이 53.0%, 2학년이 50.9% 등의 순을 보여 고학년이 될수록 사교육을 받는 비중이 높은 것으로 나타났다. 그리고 사교육을 받는 학생의 비율 또한 2003년 53.0%, 2005년 56.8%, 2009년 61.9%로 해를 거듭할수록 꾸준히 증가하는 추세를 보이고 있으며, 그 비용도 점차 늘어나고 있는 것으로 나타났다(잡코리아, 2010).

대학생들의 사교육 참여 및 사교육비 지출이 더 좋은 직장을 구하기 위한 스펙 마련에 초점이 맞추어져 있다고 본다면(정지선, 김훈호, 2008), 이와 같은 고등교육에서의 사교육 문제는 사교육비 지불 능력에 따른 취업경쟁력 불평등 및 이에 따른 사회양극화 현상을 심화시킬 우려가 있어 이에 대한 정책적 대안의 마련이 시급한 현안이라고 볼 수 있다.

이처럼 고등교육 단계에서의 사교육 문제가 심각한 사회문제로 대두되고 있음에도 불구하고 이에 대한 사회적, 학술적 차원에서의 논의는 부족한 실정이다. 현재까지 대학생의 사교육비에 관한 연구는 크게 사교육 실태 및 사교육 참여율과 지출 규모에 관한 연구가 중심을 이루고 있다(김지하, 박지은, 2008; 민혜리, 2003; 천세영, 이석열, 이선호, 2004; 정지선, 김훈호, 2009; 남수경, 2006). 이 연구들은 학술적 논의가 거의 없는 대학생 사교육의 실태 파악 및 대학생 사교육에 영향을 미치는 변인들을 분석하는데 중요한 근거를 제시해 주고 있음이 분명하나 연구의 방법과 범위에 있어서 여러 가지 제한점이 있다.

가령 각 개인은 학교나 지역사회에 내재하고 있음에도 불구하고 이러한 특성을 반영하여 자료를 분석한 연구가 거의 없다. 또한 고등학교 사교육 경험과 대학교 사교육 경험을 연속선상에서 분석한 연구가 전무한 실정이다. 고등학교 시기에 사교육을 많이 받은 학생일수록 자기주도적 학습능력을 상실할 가능성이 높으며, 이에 따라 대학에서도 습관처럼 사교육에 의존할 수 있다는 가능성(이준호, 2011)이 제기될 수 있음에도 기존의 연구들은 이러한 문제에 대해서는 많은 관심을 기울이지 못해왔다. 이에 고등학교에서 대학교로 이어지는 사교육의 연속적인 영향력에 대한 분석 결과는 우리나라 사교육비 문제에 대한 새로운 이해를 가능하게 하는 단초를 제공할 수 있다는 측면에서 의의를 가질 수 있을 것으로 기대된다.

따라서 본 연구에서는 한국직업능력개발원이 제공하는 한국교육고용 패널 중년 자료를 활용하여 대학생의 사교육 참여와 이들이 지출하는 사교육비 규모에 영향을 미치는 변인을 밝히고자 하였으며, 특히 고교 시절의 사교육 참여 경험이 대학 진학 이후의 사교육 수요 행위에 미치는 영향력을 살피는데 연구의 주안점을 두었다.

대학생들의 경우 그들이 속해 있는 각 대학의 특성에 따라 사교육 참여행위에 있어 차이를 보일 수 있다. 이에 본 연구에서는 이러한 자료의 위계적 속성을 고려하여 고등학교 시기의 사교육 참여 경험이 이후 대학 단계에서의 사교육 참여 행위에 미치는 영향력을 살피기 위해 위계적 선형모형(HLM)을 활용하는 한편, 이러한 영향력의 인과 효과 추정을 위해 고등학교 사교육 참여 행위와 관련된 관찰가능한 특성의 통제를 통한 집단 간의 동질성 확보 수단인 성향점수(Propensity Score)를 통제변수로 활용하여 분석을 실시하였다. 그리고 이러한 분석 결과를 통해 대학생의 사교육 수요에 대한 이해를 넓히는 한편, 대학의 교육과정 및 사교육 관련 정책의 수립 과정에 유용한 정책적 시사점을 제공하는데 연구의 목적을 두었다. 이러한 연구 목적을 달성하기 위한 구체적인 연구 문제는 다음과 같다.

- 1) 고등학교 시기의 사교육 경험이 대학생의 사교육 참여에 어떠한 영향을 미치는가?
- 2) 고등학교 시기의 사교육 경험이 대학생의 사교육비 지출에 어떠한 영향을 미치는가?

II. 선행연구 분석

1. 대학생의 사교육비 개념

경비의 운영형태를 기준으로 사교육비의 개념에 접근할 경우, 사교육비는 일반적으로 공공회계 절차를 따르지 않고 학생 학부모 등 개인이 자유의사에 따라 지출하는 교육비로 정의된다(천세영, 2001). 이러한 사교육비는 광의로 접근하느냐 협의로 접근하느냐에 따라 그 범위가 달라진다. 즉 광의로 접근할 경우 대학생의 사교육비는 입학금과 수업료를 제외한 정규 교육과정에 참여하기 위해 개인적으로 지출해야 하는 모든 비용으로 정의할 수 있으며 교재 및 부교재 구입비, 학용품비, 수업 준비물비, 학원비, 식비, 교통비, 단체활동비, 하숙이나 자취 등 주거비 등을 포함한다. 반면에 협의에서 보면 사교육비는 학교 외의 교육에 참여하기 위해서 추가로 발생하는 비용이며 학원수강료, 교재구입비, 식비나 교통비, 단체활동비 등을 포함한다(남수경, 2006).

천세영 외(2004)는 우리나라 대학생들은 생활 범위가 매우 넓고, 교육비 지출에 영향을 미치는 요인이 다양하므로 사교육비의 범위도 매우 포괄적으로 접근할 필요가 있다고 지적하였다.

이 연구에서는 사교육비를 '개인의 자유 계획에 의해 집행되는 경비'라고 정의하였으며, 교재비, 교통비, 단체 활동비 등 대학에 재학하기 위해 필요한 경비와 학원비, 문화생활비 등과 같이 대학 외의 공간에서 지출하는 교육관련 비용을 포괄하였다. 대학생 사교육비의 개념을 포괄적으로 접근한 다른 연구로 이정미(2010)의 연구를 들 수 있는데 이 연구에서 사교육비는 일반 사교육비와 과외 사교육비를 모두 포함하는 개념으로 '등록금 외에 대학교육 및 대학 이외의 교육활동에 참여하기 위해 개인이 부가적으로 지출하는 모든 비용'이라고 정의하였다(이정미, 2010).

이와는 달리 한국교육고용패널에서 제공하는 데이터를 바탕으로 분석을 실시한 김지하, 박지은(2008)과 정지선, 김훈호(2009)의 연구에서는 사교육비의 범위를 대학 외의 교육활동 중 '과외 수업과 학원 수강' 등에 소요되는 비용만으로 제한하였다. 또한 민혜리(2003)는 어학학습, 실무학습, 각종 시험준비 비용 등의 '과외 학습'으로 사교육비를 한정하여 취업준비 사교육비의 유형과 비용에 대한 연구를 실시하였다.

이처럼 사교육비에 대한 정의는 연구의 대상과 목적에 따라 다양하게 정의 되고 그 범위가 제한되어 왔다. 본 연구에서는 선행연구를 바탕으로 사교육비를 '대학 외의 교육활동에 수반되는 비용'으로 정의하며, '과외수업 및 학원 수강'으로 대상을 한정하여 연구를 진행하고자 한다.

2. 대학생 사교육비 지출에 영향을 미치는 요인

지금까지 수행된 대학생 사교육비 관련 연구들은 사교육비의 지출 규모와 사교육비 참여 비율에 대한 집단 간 차이를 조사한 연구가 대부분이다. 민혜리(2002)와 남수경(2006)은 대학생 사교육비 지출과 가장 밀접한 관련이 있는 취업 사교육비에 관한 조사 연구를 실시하였다. 민혜리(2002)는 전국 32개 4년제 대학교의 학생 2500여명을 대상을 하여 취업과외비용을 조사하였다. 이에 따르면 설문조사에 참여한 학생들의 48%는 과외를 받은 경험이 있으며, 과외비 지출은 개인특성, 가정배경에 따라 다르게 나타났다. 즉 남학생인 경우 학부모의 소득과 학력이 높은 가정의 학생일수록 사교육비를 많이 지출했다. 사교육비 항목을 유형별로 구분하여 분석하였을 경우, 재학 중인 학교의 특성에 따라 과외비 지출 양상은 차이를 보였다. 즉 어학학습비용의 경우 수도권 대학의 학생일수록, 자연계열보다는 상경법정계열 및 사회계열인 경우 과외비 지출액이 높게 나타났다. 이와는 반대로 실무학습과외비의 경우 서열이 낮은 대학에 재학할수록, 자연계열 학생일수록 사교육비 지출이 많았다. 마지막으로 각종 시험 준비 비용의 경우 수도권 학생들이 비수도권 학생들보다 높은 과외비를 부담하며 상경법정계열이 타 계열에 비해 사교육비를 많이 지불하였다.

남수경(2006)은 국립대학교 사범대학 학생 626명을 대상으로 사범대학 학생의 사교육비 지출에 대하여 조사연구를 실시하였다. 그 결과 부모의 학력, 성별, 임용고사 준비 시작시기, 사교육

참여 이유 등은 사교육비 지출 규모에 통계적으로 유의미한 영향을 주지 않는 것으로 나타났다. 그리고 부모의 월평균 소득이 높을수록, 6개 대학 중 광역시에 소재한 2개의 학교에 재학 중인 학생이 다른 집단에 비해 사교육비 지출 규모가 큰 것으로 나타났다. 위의 연구는 대학생 사교육비 지출에 대한 집단 간 차이를 분석함으로써 대학생 사교육비에 관련한 논의를 풍부하게 하고 있다는 점에서 의의가 있으나 대상이 사범대학생에 한정되어 있어 대학생 사교육비의 전체적인 특성을 파악하는데 한계가 있다.

천세영 외(2004)와 이정미(2010)는 사교육비의 개념을 일반사교육비와 과외사교육비로 확대하여 연구를 실시하였다. 천세영 외(2004)는 7개 대학 350명을 대상으로 사교육비 지출에 대한 집단 간 차이를 분석하였다. 그 결과 본 연구에서 관심을 가지는 '학원 수강 및 과외비 지출'에 대해 집단 간 유의미한 차이를 보이는 변수는 아르바이트여부와 월평균 가계소득 뿐이었다. 기타 성별, 학년 등의 개인적 배경이나 부모의 학력, 직업, 집의 거주 형태 등과 대학관련 배경인 대학 소재지는 과외비 지출에 있어 유의미한 집단 간 차이를 나타내지 않았다. 이정미(2010)는 대학의 전공계열, 소재지, 규모를 층화변수로 설정하여 유층 비율 군집 표집을 활용하여 전공계열에 따른 사교육 참여 양상 및 사교육비의 차이를 분석하였다. 연구 결과 우리나라 대학생 1인당 연간 사교육비가 약 339만원이며, 예체능계열에 재학 중인 학생이 사교육비 지출규모가 가장 크고 교육계열에 재학 중인 학생이 가장 작은 것으로 나타났다. 사교육 참여는 성별, 학년, 성적, 부모 학력과 소득, 대학 만족도에 따라 유의한 차이가 있는 것으로 나타났으며, 사교육비 지출 규모는 성별은 의학계열에서, 학년은 인문·의학계열을 제외한 모든 계열에서, 부모의 소득은 자연계열을 제외한 모든 계열에서 유의한 차이를 보였다. 천세영 외(2004)와 이정미(2010)의 연구는 개인변수, 가정배경 변수 외에 학교관련 변수를 포함시켜 배경변수별 하위 집단 간 차이를 분석했다는 점에서 대학생 사교육비 관련 연구의 폭을 확장시키는데 기여했지만, 이러한 분석 결과는 배경변수와 사교육 참여 또는 사교육비 지출 간의 인과관계를 설명하지는 못한다는 한계를 가진다.

이에 비하여 김지하, 박지은(2008)과 정지선, 김훈호(2009)의 연구는 각 하위 변수들이 사교육 참여율과 사교육비 지출 규모에 미치는 영향을 밝힌 실증적인 연구라는 점에서 의의가 있다. 김지하, 박지은(2008)은 한국직업능력개발원의 교육고용패널 1,2차년도와 고등학교 3학년 코호트 자료를 분석(최종 1,345명 포함)하였다. 연구결과에 따르면 개인특성변수로서 수능백분위, 강의 태도, 편입계획이, 가정배경변수로서 부모학력이 사교육 참여여부와 월평균 사교육비 지출 규모에 유의미한 변수로 나타났다. 그리고 성별, 고등학교 과외 경험은 사교육 참여여부, 월평균 사교육비, 월평균 사교육 시간에 유의미한 변수로 밝혀졌다. 이에 반해 대학서열, 전공계열, 지역 특성관련 변수들은 선행연구 결과와는 달리 대학생 사교육 수요에 영향을 주지 못하는 것으로 나타났다. 한편, 정지선, 김훈호(2009)는 분석과정에서 사교육에 참여하지 않은 학생의 표본을 임의로 제외할 경우 발생할 수 있는 표본 선택 편향(sample selection bias)의 가능성을 실증하고

사교육 참여비율과 사교육비 규모에 영향을 주는 변수들을 분석하였다. 연구 결과에 따르면 학교 성적이 높을수록, 4년제 대학에 재학할수록, 그리고 예체능계열에 비해 인문·사회·교육계열이 사교육 참여 비율이 높은 것으로 나타났다. 또한 사교육 지출 규모에 통계적으로 유의미한 영향을 미친 변수는 개인과 가정배경변수인 수강태도와 월평균 가계 소득, 고3 시절 사교육경험으로 나타났다. 이 두 연구에서는 다른 연구에서는 포함하지 않았던 고교시절 사교육경험을 변수로 포함시켰는데, 고교시절 사교육경험은 대체적으로 사교육 참여 비율과 사교육 지출 규모에 유의미한 영향을 미치는 변수로 나타났다. 이 두 연구는 대학생 사교육비 지출에 영향을 미치는 요인을 실증적으로 밝히고자 했다는 점에서 큰 의의를 지닌다. 그러나 두 연구 모두 한 해 동안의 사교육 행위 특성만을 분석한 연구이기 때문에 시간의 변화에 따른 사교육 수요 및 참여 특성의 변화 양상을 드러내지는 못하였다는 제한점이 있다.

선행 연구를 분석한 결과 우리나라 대학생 사교육비 관련 연구는 다음과 같은 몇 가지 한계점을 가지는 것으로 나타났다. 첫째, 대학생 사교육비와 관련한 연구들은 개인 및 가정 배경 특성 변수와 학교배경 특성 변수를 동시에 투입하여 각 변수들과 사교육비의 관련성을 분석하였다. 이는 각 개인이 학교에 내재하고 있다는 특성을 무시한 결과이기 때문에 결과 해석의 정확성을 담보하기가 어렵다. 따라서 이러한 특성을 반영한 다층 분석 연구가 필요하다. 둘째, 대부분의 대학 사교육비 관련 연구들이 한 해 동안의 사교육비 지출에 대한 연구에 머물고 있으며, 시간 변화에 따른 사교육비 지출 및 참여 특성의 변화 양상을 밝힌 연구가 전무하다. 학년이 올라갈수록 대학 사교육비 지출이 증가하는 추세를 보이는지, 또한 이러한 경향이 고등학교의 사교육 경험과 어떤 관계가 있는지를 분석해 보는 것은 각 가정의 사교육비 부담의 무게를 총체적으로 진단할 수 있고 이에 따른 정부의 사교육비 경감 대책 마련에 유의한 정보를 제공할 수 있을 것이다. 셋째, 사교육비에 영향을 미치는 변수로 모든 연구가 대학의 소재지와 계열을 포함하고 있지만 대학교의 위계 서열은 김지하, 박지은(2008)의 연구에서만 포함되었다. 우리나라 사교육의 주요한 목적 중 하나가 졸업 후 좋은 직업을 가지고 질 높은 삶을 영위하는 것이란 점을 고려한다면, 소위 높은 서열의 대학교에 재학 중인 학생들의 경우 이러한 지위 획득을 위한 경쟁이 더욱 치열할 것으로 추론해볼 수 있다. 이에 대학의 서열과 대학 사교육의 관계를 실증한 결과는 이러한 추론에 대한 유의미한 시사점을 제공해 줄 수 있을 것이다. 넷째, 기존 연구들은 대학생의 사교육 참여 행위와 관련된 변수들 간의 상관관계에 주로 초점을 맞추어 왔다. 예외적으로 정지선 외(2009)는 Heckit 모형을, 김지하 외(2008)는 도구변수를 활용하여 대학생 사교육에 대한 인과관계 추론(causal inference)을 시도하였으나 해당 연구 방법의 한계로 인해 엄밀한 인과 효과 추정이 이루어진 것으로 보기는 어렵다. 즉 Heckit 모형은 처치변수와 통제변수 간에 존재하는 내생성 문제 해소에는 적합하지 않으며, 도구변수 추정방법은 그 가정을 충족하는 적합한 변수를 찾기가 현실적으로 어려운 것이 사실이며 김지하 외(2008) 연구에서도 이러한 경향이 나

타남을 확인할 수 있다. 따라서 대학생 사교육에 대한 인과관계 분석을 위해서는 좀 더 엄밀한 추정 방법이 요구된다.

본 연구에서는 이와 같은 선행연구 분석 결과와 한계점을 반영하여 대학생 사교육과 관련된 개인 및 가정배경 변수로 가구소득, 대학 입학 후 경과 연수, 고 3시절의 사교육 경험 여부를 포함하였으며, 대학수준 독립변수로서 대학서열과 대학 소재지를 포함하였다. 그리고 대학생들의 사교육 참여 행위에 대해 관찰 가능한 통제변수들의 영향력을 통제하기 위하여 고 3시절 사교육 참여에 대한 성향점수(propensity score)를 통제변수로 활용하였으며, 각 개인이 학교에 내재되어 있다는 특성을 고려하여 다층모형 분석을 실시하였다.

3. 성향 점수(Propensity Score)

본 연구에서는 분석 대상 학생들의 고교 시절 사교육 참여에 대한 성향점수(Propensity Score)를 도출하여 이를 학생 수준의 통제변수로 활용함으로써 고교 시절의 사교육 참여 경험이 대학생의 사교육 참여 및 사교육비 지출에 미치는 인과적인 영향력을 파악하고자 하였다.

성향점수(Propensity Score)는 처치 혹은 통제집단에 속할 참여자들의 관찰 가능한 특성들이 주어졌을 때, 이들이 처치 집단에 참여할 조건부 확률(conditional probability of assignment to treatment)을 나타내며, 이를 식으로 나타내면 아래와 같다(Rosenbaum & Rubin, 1983)

$$\text{Propensity Score} = P(X) = \Pr(T=1|X)$$

Rosenbaum & Rubin(1983)은 성향점수가 주어졌을 때, 처치 집단에의 참여 여부는 관찰된 공변수(covariate)과는 조건부 독립이며 $(X \perp T) | P(X)$, 공변수 벡터에 대하여 강한 무관성의 가정(strongly ignorable treatment assignment assumption)¹⁾이 충족될 경우 이러한 가정이 성향점수에 대해서도 동일하게 성립함을 보였다.

$$(Y_1, Y_0) \perp T | P(X), \quad 0 < \Pr(T = 1 | P(X)) < 1$$

1) 처치집단에의 참여에 대한 강한 무관성의 가정(strongly ignorable treatment assignment assumption)은 처치 혹은 통제집단에의 참여자의 특성에 대한 공변수 벡터(vector of covariate) v 가 주어졌을 때, 처치 집단에의 참여 여부는 종속변수들과 독립적이고, 처치집단과 통제집단 구성원들의 처치 집단 참여 확률의 분포는 공통의 영역에 속한다는 가정이 만족될 때 충족되는 것으로 알려져 있다(이석원, 2003).

$$(Y_1, Y_0) \perp T | v, \quad 0 < \Pr(T = 1 | v) < 1$$

이와 같은 정리는 산출된 성향점수가 동일한 처치 집단과 통제 집단의 구성원 간의 종속변수 값을 비교할 경우 양 집단 간의 관찰 가능한 특성들이 동일한 분포를 가지게 되므로, 이에 따라 무작위 실험평가(randomized experiment)와 같이 선택편의(selection bias)가 없는 인과효과를 산출할 수 있음을 의미한다(이석원, 1983). 즉, 동일한 성향점수를 가진 학생들은 그들의 특성을 나타내는 공변인에 있어 동질성이 확보되므로, 이들을 대상으로 처치 효과를 관찰할 경우 인과효과를 산출할 수 있다는 장점이 있다.

이러한 성향점수를 활용한 연구들은 주로 처치집단과 비교집단 간 구성원의 짝짓기(matching)을 통하여, 두 집단 간 평균의 차이를 살핌으로써 정책 및 처치의 효과를 살피는 연구가 주종을 이루었다(Rosenbaum & Rubin, 1983; Dehejia & Wahba, 2002; Frangakis & Rubin, 2002; 이석원, 2003). 그러나 최근 들어서 이러한 성향 점수 자체를 회귀식 상의 통제변수로 활용함으로써 처치 전 집단 간의 동질성을 꾀하는 연구들이 시도되고 있다(Hong & Raudenbush, 2005).

이에 본 연구에서는 한국교육고용패널의 2004년도 학생 및 가계 데이터를 활용하여 고 3 사교육 참여 경험과 유의미한 상관을 보이는 32개의 변수를 처치전 공변인(covariate)으로 설정한 후, 이 중 15개의 변수를 활용하여 propensity 모형을 설정하였다. 그리하여 앞서 선정한 32개 공변인에 대해 처치집단과 통제집단 간의 평균값 차이를 확인하는 balance 조정을 거쳐, 이를 토대로 성향점수를 산출하였으며, 고교 3학년 시절 사교육에 참여한 학생들과 참여하지 않은 학생들 간의 이질성을 해소하기 위하여 이를 회귀식 상의 통제변수로 활용하였다.

III. 연구방법

1. 분석 자료

본 연구는 고교 시절의 사교육 경험이 대학생의 사교육 참여 및 사교육비 지출에 미치는 영향을 알아보기 위해 한국직업능력개발원이 2004년부터 2008년에 걸쳐 실시한 한국교육고용패널자료(KEEP)의 고등학교 3학년 집단(코호트) 자료를 활용하였다. 한국교육고용패널은 2004년에 고등학교 3학년이던 학생 4,000명을 표집하여 학생의 개인 특성 및 배경 변수와 사교육 수요 관련 변수, 진학 대학 관련 특성 변수 등에 대해 매년 추적조사를 실시하고 있어 기존 연구에 비해 추정치의 신뢰성을 확보할 수 있다는 장점을 지니고 있다(김지하 외, 2008).

본 연구에서는 우선 패널 1차년도인 2004년 인문계 및 실업계 고등학교 3학년 학생 패널 4,000명을 대상으로 프로빗 분석을 실시하여 각 학생이 고등학교 재학 당시 사교육에 참여할 성향점수를 산출하였다. 그리고 이 학생들이 대학에 진학한 이후, 즉 2차년도(2005년), 3차년도(2006년), 4차년도(2007년), 5차년도(2008년)별로 해당 학생들의 사교육 참여 여부 및 사교육비

지출에 있어서의 변화 양상을 살펴보았다. 그리고 필요한 변수가 누락된 경우 이를 분석대상에서 제외하였으며, 이에 최종 분석에 포함된 표본수는 총 6,350명으로 연도별로 살펴보면 2005년 2,327명, 2006년 1,692명, 2007년 974명, 2008년 1,357명이었다. 한편, 사교육비 지출을 종속변수로 한 연구모형의 경우 사교육비 지출 행위는 우선 사교육에 참여하고, 이에 대한 비용을 지불한 경우에만 피회귀변수가 관찰될 수 있으므로, 사교육비 지출에 대한 분석을 시행하기 위해서는 사교육에 참여하되, 사교육비를 지출한 학생들만을 대상으로 분석을 실시해야 한다. 이에 본 연구의 두 번째 연구문제인 사교육비 지출에 대한 분석을 위해 포함된 표본수는 총 816명이다.

2. 변수 선정

본 연구의 종속변수는 대학생들의 사교육 참여 여부와 사교육에 지출한 월평균 사교육비이다. 본 연구에서는 해당 분석 기간 중 외국어공부(영어, 제2외국어)와 전공 관련 교육, 공무원 시험 준비를 위한 사교육 중 하나에라도 참여한 경험이 있을 경우 이를 대학생의 사교육참여 행위로 규정하였다. 이에 사교육 참여 여부 변수의 경우 2004년 당시 고등학교 3학년이었던 학생이 2005년 대학에 입학한 시점을 기준으로 t 년 후 시점의 사교육 참여 여부(참여: 1, 미참여: 0)를 나타내며, 사교육비 변수의 경우 사교육 참여 여부 변수와 마찬가지로 2005년 기준 t 년 후 시점의 월평균 사교육비 지출액(만원 단위)의 자연로그값으로 나타낸다. 이 중, 사교육 참여 여부를 종속변수로 한 모형의 경우 총 표본수는 6,350개, 사교육비 지출액을 종속변수로 한 모형의 경우 표본수는 816개이다.

선행 연구를 통해 대학생의 사교육 참여 및 사교육비 규모에 영향을 미칠 것으로 예상되는 독립변수는 크게 개인수준 변수와 대학수준 변수로 구분하였으며, 개인 수준 변수는 가계소득, 대학 입학후년 경과연수, 고3 사교육 경험 여부로 구성되었다. 이 중 가계소득 변수 및 입학 후 경과연수 변수는 2005년을 기준으로 해당 연도의 값을 투입하였다. 그리고 이와 같이 시간에 따라 변화하는 1수준 개인변수들의 영향력을 통제하고, 고교 3학년 시절의 사교육 경험이 대학생의 사교육 참여 및 사교육비 지출에 미치는 영향력을 추정하기 위해 처치 전 집단 간 차이에 대한 통제변수로서 성향점수를 포함하였다. 성향점수는 2004년 고 3사교육 참여 경험과 유의미한 상관을 보이는 32개의 공변수(covariate)에 대하여 15개의 변수를 포함하는 propensity모형을 설정한 후 처치집단과 통제집단간의 균형(balance)을 맞춘 후 산출하였다.

대학수준 변수는 대학서열과 대학소재지로 구성되었다. 이 중 대학 서열 변수는 중앙일보 교육개발연구소에서 매년 발표하는 대학평가 종합순위를 활용하여 2006년 기준 1위에서 30위에 해당하는 대학을 1, 해당 순위 밖의 대학을 0으로 코딩하여 더미변수화하였다. 대학 소재지 변수는 지방 소재 대학을 참조 집단으로 하여, 서울 소재 대학을 1로 코딩한 더미변수를 활용하였다.

독립변수와 종속변수에 대한 설명 및 기술통계량은 <표 1> 과 같다.

<표 1> 변수 설명 및 기술통계

| 구분 | 변수명 | 설 명 | 평균 | 표준 편차 | 최소값 | 최대값 | N | 비고 |
|----------|------------|--|------|----------|-------|-------|------|----|
| 종속 변수 | 사교육 참여 | 입학년도(2005년) 기준 t년 후 시점의 사교육 참여 여부 | 0.22 | 0.41 | 0.00 | 1.00 | 6350 | |
| | 사교 육비 | 입학년도(2005년) 기준 t년 후 시점의 월 사교육비(만원) 로그값 | 2.62 | 1.14 | -1.68 | 8.01 | 816 | |
| | 가구 소득 | 수능시험연도(2005년) 기준 t년 후 시점의 월평균 가구소득(만원)의 자연로그값 | 5.54 | 0.71 | 0.00 | 10.31 | 6350 | 참여 |
| | | | 5.71 | 0.67 | 0.00 | 10.31 | 816 | 비용 |
| 개인 수준 | 입학후 연수 | 2005년 입학년도 기준(0) 경과 연수 | 1.21 | 1.15 | 0.00 | 3.00 | 6350 | 참여 |
| | | | 1.53 | 1.06 | 0.00 | 3.00 | 816 | 비용 |
| 독립 변수 | 성향 점수 | 고3 사교육 참여에 대한 성향 점수 | 0.55 | 0.26 | 0.00 | 1.00 | 6350 | 참여 |
| | | | 0.71 | 0.20 | 0.04 | 1.00 | 816 | 비용 |
| | 고3교 육경험 | 고3 시절 사교육경험 여부 더미변수 (1: 참여, 0: 불참) | 0.55 | 0.50 | 0.00 | 1.00 | 6350 | 참여 |
| | | | 0.76 | 0.43 | 0.00 | 1.00 | 816 | 비용 |
| 대학 수준 | 대학 서열 | 중앙일보 선정 대학랭킹 더미변수 (1: 1-30위, 0: 31위~) | 0.08 | 0.28 | 0.00 | 1.00 | 350 | |
| | | | 0.22 | 0.42 | 0.00 | 1.00 | 94 | |
| 독립 변수 | 대학 소재지 | 더미변수 : 서울(1), 지방(0) | 0.16 | 0.37 | 0.00 | 1.00 | 350 | |
| | | | 0.30 | 0.46 | 0.00 | 1.00 | 94 | |

3. 분석 모형

전통적인 회귀모형은 균일한 모집단으로부터의 독립적인 무선 표집을 가정하고 있지만, 실제로 대학생들은 자신이 재학 중인 대학교에 내재되어(nested) 있어, 대학교의 영향으로부터 자유로울 수 없는 것이 사실이다. 이에 본 연구에서는 대학생의 사교육 참여 행위에 대한 개인 특성, 그리고 상위 조직으로서 대학교의 특성을 동시에 고려하기 위하여 위계적 선형모형을 사용하였다. 그리고 분석에 사용된 독립변수 중 가구 소득 및 성향점수에 대해서는 연구대상 대학의 실제 평균 사교육참여율 및 사교육비 지출을 반영하기 위해 개별 대학 평균으로의 중심점 교정(group-mean centering)을 실시하였다. 그리고 본 연구의 분석모형 도출을 위해 아래와 같은 과정을 거쳤다.

(모형 1) 기초모형

(모형 2) 기초모형 + 학생 수준 변수

(모형 3) 기초모형 + 학생 수준 변수 + 학교 수준 변수

기초모형(Random-Effects Anova Model)에서는 독립변수를 포함하지 않은 상태로, 전체 분산을 대학 내(within) 그리고 대학 간(between) 분산으로 분해하기 위하여 수행하였다. 그리고 이러한 기초모형에 수준별로 독립변수들을 포함하여 분석 모형을 구성하였다. 이를 위하여 모형 2에서는 1)절편은 무선효과, 기울기는 고정효과로 설정하는 경우(Random Intercept Model)와, 2)절편과 기울기를 모두 무선효과로 설정하는 경우(Random Coefficient Model)와, 3) Random Coefficient Model의 결과를 토대로 1수준의 회귀계수를 선택적으로 고정 혹은 무선효과로 설정하여 분석한 모형으로 나누어 이들에 대한 분석을 각각 실시한 후, 이들 세 모형에 대하여 1수준 회귀계수의 분산에 대한 카이제곱(χ^2) 검정, 모형에 대한 로그우도비(Log Likelihood) 차이 검정, 2수준의 임의효과(random effect)에 대한 신뢰도(Reliability) 점검을 통하여 가장 적합한 모형을 선정하여 분석을 실시하였다. 그리고 선정된 모형 2의 1수준 회귀계수에 영향을 미치는 2수준 독립변수들을 투입한 모형 3을 통하여 대학교의 특성이 소속 대학생들의 사교육 참여 및 사교육비 지출에 미치는 영향력을 통제한 후 관심변수의 영향력을 확인하였다.

이를 바탕으로 고교 3학년 시절의 사교육 경험이 대학생의 사교육 참여에 미치는 영향을 알아보기 위해 설정한 연구 문제 1에 대해서는 이분형 종속변수의 분석에 적합한 위계적 일반화선형 모형(HGLM)을 아래와 같이 활용하였다.

1) 모형 1 : 기초모형

$$1\text{수준 방정식} : \eta_{ij} = \beta_{oj}$$

$$2\text{수준 방정식} : \beta_{oj} = \gamma_{oo} + u_{oj} \quad (u_{oj} \sim N(0, \tau))$$

2) 모형 2, 3 : 연구모형

$$1\text{수준 방정식} : \eta_{ij} = \beta_{0j} + \sum_{q=1}^4 \beta_{qj} X_{qij} + r_{ij} \quad (r_{ij} \sim N(0, \sigma^2))$$

$$2\text{수준 방정식} : \beta_{oj} = \gamma_{00} + \sum_{s=1}^2 \gamma_{os} W_{sj} + u_{oj} \quad (\mu_{0j} \sim N(0, \tau_{00})),$$

$$\beta_{qj} = \gamma_{qj} \quad (q = 1, 2)$$

그리고 고교 시절의 사교육 경험이 대학생의 사교육비 지출에 미치는 영향을 알아보기 위해 설정한 연구문제 2에 대해서는 위계적선형모형(HLM)을 활용하였으며, 기초 모형과 연구 모형은 다음과 같다.

1) 모형 1 : 기초모형

$$1\text{수준 방정식} : Y_{ij} = \beta_{0j} + r_{ij} \quad (r_{ij} \sim N(0, \sigma^2))$$

$$2\text{수준 방정식} : \beta_{0j} = \gamma_{00} + u_{0j} \quad (u_{0j} \sim N(0, \tau))$$

2) 모형 2 : 연구모형

$$1\text{수준 방정식} : Y_{ij} = \beta_{0j} + \sum_{q=1}^4 \beta_{qj} X_{qij} + r_{ij} \quad (r_{ij} \sim N(0, \sigma^2))$$

$$2\text{수준 방정식} : \beta_{0j} = \gamma_{00} + \sum_{s=1}^2 \gamma_{0s} W_{sj} + u_{0j} \quad (\mu_{0j} \sim N(0, \tau_{00})),$$

$$\beta_{qj} = \gamma_{qj} \quad (q = 1, 2)$$

IV. 분석 결과

1. 대학생의 「사교육 참여 여부」에 대한 위계적일반화선형모형 분석

<표 2>는 위계적 일반화선형모형(HGLM)을 활용하여 고교 시절의 사교육 경험이 대학생의 사교육 참여 여부에 미치는 영향력의 크기 및 관련 통제변수들의 영향력을 함께 나타낸 것이다.

우선 앞서 제시한 모형 1(기초모형)을 통하여 사교육 참여 여부에 있어 대학 간에 존재하는 변량(variation)의 크기를 확인하였다. 그러나 일반적인 2수준 HLM 모형에서 전체 분산에 대한 조작·집단 간 변량의 비율 크기를 나타내는 집단내 상관계수(ICC: Intraclass Correlation Coefficient)는, 위와 같이 이분변수를 종속변수로 다루는 비선형모형에 대해서는 1수준 분산의 이분산성으로 인해 정확한 정보를 제공하는데 어려움이 있다(Raudenbush & Byrk, 2002). 이에 연구 대상 대학들에 대한 전반적인 사교육 참여 수준을 살펴보기 위해 해당 대학의 평균적인 사교육 참여율을 나타내는 1수준의 절편계수인 β_{0j} 의 95% 신뢰값 구간(plausible values)을 제시하였다. 이에 따르면 β_{0j} 에 대한 대학들의 95% 신뢰값 구간은 $-1.365 \pm 1.96 * \sqrt{0.296} = (-2.431, -0.299)$ 으로 나타났다. 이에 대한 승산값(Odds)의 구간은 (0.088, 0.742), 해당 확률은 (0.081 0.426)의 구간에 있는 것으로 나타나 학교 간 평균 사교육 참여율에 있어 상당한 편차가 있음을 확인할 수 있다.

모형 2는 학생 개인의 사교육 참여에 영향을 미치는 학생 수준 변수들만을 투입한 이후 이러한 1수준 변수들의 회귀계수들을 고정 혹은 임의 효과로 간주할 것인지를 판단하기 위해 설정한 연구모형이다. 우선 이 연구의 관심사인 대학생의 사교육 참여에 고교 시절의 사교육 참여 여부가 미치는 영향력을 파악하기 위해 해당 변수를 투입하였고, 학생 수준의 통제변수로 2005년 대

학 입학 이후의 경과연수, 해당 연도 가구 소득의 자연로그값, 앞서 언급한 고교 시절의 사교육 참여 여부에 대한 학생 개인의 성향 점수(propensity score)를 통제변수로서 추가로 투입하였다. 그리고 이러한 독립변수들이 대학생 개인의 사교육 참여에 미치는 영향력의 크기(1수준 기울기 계수)와 해당 대학의 사교육 평균 참여율을 나타내는 절편계수를 모두 임의효과로 설정한 후, 1수준 회귀계수에 대한 분산 검정, 2수준 임의 효과에 대한 신뢰도 검정을 실시하여 무선절편모형(Random Intercept Model)이 적합함을 확인하였다. 이에 더하여 무선절편모형과 무선계수모형(Random Coefficient Model) 간의 로그우도비 차이 검정을 실시하여 무선절편모형이 모형 2에 적합한 모형임을 재확인하였다.

<표 2> 사교육 참여 여부에 대한 HGLM 분석 결과

| 변인 | 모형 1 (기초모형) | | 모형 2 (기초모형+학생변수) | | 모형 3 (기초모형+학생변수+학교변수) | |
|---|----------------------|---------------|----------------------|---------------|--------------------------|---------------|
| | 회귀 계수 (S.E) | Odds ratio | 회귀 계수 (S.E) | Odds ratio | 회귀 계수 (S.E) | Odds ratio |
| 학생 수준 변수 | | | | | | |
| 절편, r_{00} | -1.365*** (0.047) | 0.255 | -2.050*** (0.073) | 0.129 | -2.125*** (0.075) | 0.119 |
| 입학후연수, r_{10} | | | 0.261*** (0.028) | 1.298 | 0.258*** (0.028) | 1.294 |
| ln가구소득, r_{20} | | | -0.013 (0.051) | 0.987 | -0.008 (0.052) | 0.992 |
| 고3 사교육경험, r_{30} | | | 0.596*** (0.072) | 1.814 | 0.524*** (0.075) | 1.688 |
| 성향점수, r_{40} | | | 1.387*** (0.187) | 4.003 | 1.468*** (0.192) | 4.425 |
| 학교 수준 변수 | | | | | | |
| 대학소재지 r_{01} | | | | | 0.502*** (0.112) | 1.653 |
| 대학서열 r_{02} | | | | | 0.227* (0.127) | 1.255 |
| 95% plausible values (β_{oj}) | | | | | | |
| | -2.431 | -0.299 | -2.942 | -1.216 | | |

* $P < 0.1$, ** $P < 0.05$, *** $P < 0.01$

모형 2에 따르면 대학입학 이후의 경과연수와 고교 시절의 사교육 참여 여부에 대한 성향점수가 대학생의 사교육비 참여에 정적으로 유의미한 영향력을 미치고 있음을 확인할 수 있다. 즉, 대학입학 이후 1년이 경과할 때마다 학생들의 사교육 참여에 대한 승산(odds)이 1.3배 높은 것으로 나타나 고학년으로 진학할수록 사교육 참여비율이 높아지는 것을 확인할 수 있다. 마찬가지로 고교 시절 사교육 참여에 대한 성향점수가 높은 학생일수록 대학생이 되어서도 사교육에 참

여할 확률이 높은 것으로 나타났다. 이에 반해 가구소득 변수의 경우 대학생의 사교육 참여 비율에 유의미한 영향을 미치지 못하는 것으로 나타나 대체적인 선행연구의 결과와는 상반된 양상을 보였다. 이는 선행연구들에서 미처 통제하지 못한 처치 전 공변인의 이질성을 제거하고 난 이후의 분석을 통해 나타난 결과로서, 대학생의 사교육 참여는 초중등 단계에서의 사교육 참여와 달리 학생이 속한 가계의 소득에 관계없이 미래의 취업 준비 및 습관화된 사교육경험에 따라 이루어질 수 있음을 암시한다고 할 수 있으며, 이러한 분석 결과는 김지하 외(2008)의 연구 결과와도 맥락을 같이 한다. 그러나 이러한 분석 결과에 대해서는 추후 다른 데이터 및 연구모형을 통해 다시 한 번 점검해볼 필요가 있을 것으로 판단된다.

한편, 본 연구의 주된 관심사인 고교 시절의 사교육 참여 경험과 대학생의 사교육 참여간의 인과관계는 모형 2에 따르면 고교 시절 사교육 참여 경험이 있는 대학생이 그렇지 않은 대학생에 비해 사교육 참여에 대한 승산(odds)이 1.8배 높은 것으로 나타나, 두 변수 간에 강한 인과관계가 존재하고 있음을 확인할 수 있었다. 즉, 고교 시절 사교육에 참여한 경험이 있는 학생이 그렇지 않은 학생에 비해 사교육에 참여할 확률이 유의미하게 높게 나타나는 것이다.

모형 3에서는 모형 2의 분석 결과를 토대로, 대학생이 재학 혹은 속해 있는 대학교의 특성이 학생의 사교육 참여에 미치는 영향력을 통제한 후, 고교 사교육 참여가 대학 사교육 참여에 미치는 영향을 살펴보기 위해 해당 대학의 소재지와 전국 대학 순위를 대학 특성 변수로 투입하여 결과를 분석하였다. 분석 결과 서울 소재 대학을 다니는 학생들이 그렇지 않은 학생들에 비해 사교육 참여에 대한 승산이 1.7배, 전국 대학 순위 30위권 대학에 다니고 있는 학생들이 그렇지 않은 학생에 비해 사교육 참여에 대한 승산이 1.3배 높은 것으로 나타나 대학의 특성이 학생의 사교육 참여에 유의미한 영향을 미치고 있음을 확인하였다. 즉, 서울 소재 지역 대학에 재학 중인 학생일수록, 서열이 높은 대학에 재학 중인 학생일수록 사교육 참여 확률이 높아지는 것이다.

그럼에도 불구하고 모형 2에서 대학생의 사교육 참여에 유의미한 영향을 미치는 것으로 나타난 변수들은, 대학 특성 변수들을 통제한 후에도 여전히 사교육 참여에 유의미한 영향을 미치는 것을 확인할 수 있다. 특히, 고교 시절의 사교육 참여경험이 있는 학생들은 대학 특성을 통제하고서도 여전히 그렇지 않은 학생들에 비해 사교육 참여확률이 높은 것으로 나타나 두 변수간의 강한 인과관계가 여전히 존재함을 확인할 수 있었다. 이를 통해, 고교 시절의 사교육 경험이 대학생의 사교육 참여 결정에 정적으로 유의미한 영향을 미치고 있음을 알 수 있다.

2. 대학생의 사교육비 지출에 대한 위계적 분석

<표 3>은 위계적 선형모형(HLM)을 활용하여 고교 시절의 사교육 경험이 대학생의 사교육비 지출에 미치는 영향력의 크기 및 통제변수들의 영향력을 함께 나타낸 것이다.

우선 앞서 제시한 모형 1(기초모형)을 통하여 사교육비 지출에 있어 대학 간에 존재하는 변량 (variation)의 크기를 확인하였다. 이를 위하여 전체 분산에 대한 조직·집단 간 변량의 비율 크기를 나타내는 집단내 상관계수(ICC: Intraclass Correlation Coefficient)를 아래와 같이 산출하였다.

$$\text{집단내상관계수(ICC)} = \frac{0.07}{0.07 + 1.24} = \frac{0.07}{1.31} \approx 0.05$$

이는 대학교 간 차이로 설명할 수 있는 대학생 사교육비의 분산이 전체 분산의 5%정도이며, 나머지 95%는 동일 대학교 내 학생들 간의 차이로 인해 발생함을 의미한다. 비록 전체 분산에 비해 대학 간의 차이로 인한 분산이 상대적으로 낮기는 하지만, 통계적으로 유의미한 수준의 학교 간 차이가 존재한다는 판단 하에 본 연구에서는 이에 기반하여 후속 연구를 진행하였다.

<표 3> 사교육비에 대한 위계적 분석 결과

| 변 인 | 모형 1 (기초모형) | | 모형 2 (기초모형+학생변수) | | 모형 3 (기초모형+학생변수+학교변수) | |
|---|----------------|----------|---------------------|----------|--------------------------|----------|
| | 회귀 계수 | 표준 오차 | 회귀 계수 | 표준 오차 | 회귀 계수 | 표준 오차 |
| 학생 수준 변수 | | | | | | |
| 절편, r_{00} | 2.623*** | (0.047) | 2.075*** | (0.104) | 2.069*** | (0.106) |
| 입학후연수, r_{10} | | | 0.226*** | (0.037) | 0.226*** | (0.037) |
| ln가구소득, r_{20} | | | 0.038 | (0.065) | 0.038 | (0.065) |
| 고3사교육경험, r_{30} | | | 0.264*** | (0.096) | 0.257*** | (0.098) |
| 성향점수, r_{40} | | | 0.110 | (0.261) | 0.114 | (0.261) |
| 학교 수준 변수 | | | | | | |
| 대학소재지 r_{01} | | | | | 0.095 | (0.101) |
| 대학서열 r_{02} | | | | | -0.062 | (0.105) |
| 변 량 | | | | | | |
| 학교수준 | 0.07 | | 0.04 | | 0.04 | |
| 학생수준 | 1.24 | | 1.18 | | 1.18 | |
| 전체 | 1.31 | | 1.22 | | 1.22 | |
| 집단내상관(ICC) | 0.05 | | 0.03 | | 0.03 | |
| 누적설명분산 R^2 (추가된 설명분산) | | | | | | |
| 학교수준 | | | 42.8 (42.8) | | 42.8 (0) | |
| 학생수준 | | | 4.8 (4.8) | | 4.8 (0) | |
| 전체 | | | 6.9 (6.9) | | 6.9 (0) | |

* P < 0.1, ** P < 0.05, *** P < 0.01

모형 2는 앞의 위계적일반화선형모형의 경우와 같이 모형 1에 학생 개인의 사교육 참여에 영향을 미치는 학생 수준 변수들만을 투입한 이후, 이러한 1수준 변수들의 회귀계수들을 고정 혹은 임의 효과로 간주할 것인지를 판단하기 위해 설정한 연구모형이다. 우선 이 연구의 관심사인 대학생의 사교육비 지출에 고교 시절의 사교육 참여 여부가 미치는 영향력을 파악하기 위해 해당 변수를 투입하였고, 학생 수준의 통제변수로 2005년 대학 입학 이후의 경과연수, 해당 연도 가구 소득의 자연로그값, 고교 시절의 사교육 참여 여부에 대한 학생 개인의 성향 점수(propensity score)를 추가로 투입하였다. 그리고 이러한 네 독립변수들이 대학생 개인의 사교육 참여에 미치는 영향력의 크기(1수준 회귀계수)와 해당 대학의 사교육 평균 참여율을 나타내는 절편계수를 모두 임의효과로 설정한 후, 1수준 회귀계수에 대한 분산 검정, 2수준 임의 효과에 대한 신뢰도 검정을 실시하여 절편(β_{0j})과 성향점수에 대한 회귀계수(β_{4j})를 임의효과로 설정하고, 나머지 계수들은 고정 효과로 설정하였다. 그리고 이렇게 도출된 모형(Random Intercept & Slope Model)과 무선절편모형(Random Intercept Model) 간의 로그우도비 차이 검정을 실시한 결과, 두 모형 간에 차이가 없음을 확인하였고, 이에 따라 무선절편모형이 모형 2에 적합한 모형임을 확인하였다²⁾.

모형 2의 분석결과는 앞선 HGLM에 따른 분석결과와는 다른 양상으로 나타난다. 즉, 대학생의 사교육비 지출에 대한 모형 2의 분석 결과에 따르면 대학입학 이후의 경과 연수와 고교 시절의 사교육 참여 여부만이 대학생의 사교육비 참여에 정적으로 유의미한 영향을 미치고 있는 것으로 나타나고 있다. 즉, 대학입학 이후 1년이 경과할 때마다 학생들의 사교육비 지출은 22.6%씩 높아지는 것으로 나타나 고학년으로 진학할수록 대학생들의 사교육비 지출이 늘어나는 경향을 보이고 있음을 알 수 있다. 특히, 고교 시절 사교육 참여 경험이 있는 대학생의 경우 그렇지 않은 대학생에 비해 월평균 사교육비 지출이 26.4% 높은 것으로 나타나 두 변수 간의 강한 인과관계를 확인할 수 있었다. 한편, 앞선 HGLM 분석 모형에서 사교육 참여에 유의미한 영향을 미치는 것으로 나타난 고교 시절의 사교육 참여에 대한 성향점수가 대학생의 사교육비 지출에는 유의미한 영향을 미치지 못함을 알 수 있다. 이러한 분석 결과는 대학생들의 사교육 참여에 대한 의사결정 행위와 사교육비 지출에 대한 의사결정 행위에 있어 차이가 존재할 수 있음을 암시한다. 즉, 대학생들의 사교육 참여에 대한 의사결정은 우선, 사교육 참여 여부를 결정하고, 이후 얼마만큼의 비용을 지불할지를 결정한다는 측면에서 볼 때 이러한 순차적인 의사결정 단계(sequential decision making)에 대한 분석은, 단계별로 그 접근 방식을 달리할 필요가 있을 것으로 보인다. 그럼에도 불구하고, 본 연구의 관심변수인 고 3시절의 사교육 참여 경험은 대학생의 사교육비 지출에 유의미한 영향을 미치고 있음을 알 수 있다.

2) Random Intercept Model : Deviance-2490.82, df-2, Random Intercept & Slope Model Deviance-2490.15, df-4.

한편, 모형 3에서는 모형 2의 분석 결과를 토대로, 대학생이 재학 혹은 속해 있는 대학교의 특성이 대학생의 사교육비 지출에 미치는 영향력을 통제하기 위하여 해당 대학의 소재지와 전국 대학 순위를 대학 특성 변수로 투입하여 결과를 분석하였다. 분석 결과 고교 시절의 사교육 참여 경험이 있는 학생들은 대학 특성을 통제하고서도 그렇지 않은 학생들에 비해 여전히 사교육비 지출이 25.7% 높은 것으로 나타나 두 변수간의 강한 인과관계가 여전히 존재함을 확인할 수 있다. 이러한 분석 결과는 고교 시절의 사교육 경험이 대학생의 사교육비 지출에 정적으로 유의미한 영향을 미치고 있음을 확인할 수 있었다. 그리고 앞선 사교육 참여 여부에 대한 분석결과와 같이 해당 학생의 가계 소득은 사교육비 지출에 유의미한 영향을 미치지 못하고 있음을 확인할 수 있다. 이러한 분석 결과는 대학생의 사교육비 지출에 있어서 가계의 영향력에 대한 재고가 필요함을 다시 한 번 보여준다.

한편 모형 3에 따른 분석 결과 학생들이 속해 있는 대학의 소재지 및 대학 순위는 사교육비 지출에는 거의 아무런 영향을 미치지 못하고 있음을 확인할 수 있었다. 이는 앞선 HGLM의 분석 결과는 다른 결과로, 학생들의 사교육 선택 여부와 달리 사교육비 지출 결정에는 이러한 대학 특성 변수들의 영향력이 거의 없음을 나타내는 것이다. 그 결과 모형 2와 모형 3의 고정효과 계수들의 크기에도 큰 차이가 없음을 확인할 수 있다. 이러한 연구결과는 대학생의 사교육비 지출에 대한 연구는 통상최소자승모형(OLS)으로도 충분히 수행될 수 있음을 암시하고 있다.

IV. 논의

본 연구는 고교 시절의 사교육 참여 경험이 해당 학생이 대학에 진학하고 난 이후의 사교육 참여 및 사교육비 지출에 미치는 효과를 살펴보기 위해 수행되었다. 이를 위해 본 연구는 고3 사교육 참여 집단과 미참여 집단 간에 존재하는 관찰가능한 특성의 차이를 통제하기 위하여 산출한 성향 점수(Propensity Score)를 개인 수준의 통제변수로 활용하여 고등학교 3학년의 사교육 참여 경험이 대학생의 사교육 참여 및 사교육비 지출에 미치는 인과 효과(causal effect)를 확인하였다.

그리고 대학생의 사교육 참여 및 사교육비 지출에 영향을 미치는 대학생 개인 수준 변수 뿐만 아니라, 개인이 내재된 대학교의 영향력을 함께 고려하기 위하여 위계적 선형모형을 활용하여 연구를 수행하였다. 특히, 종속변수가 대학생의 사교육 참여 여부로서 이분변수인 경우 위계적 일반화선형모형(HGLM)을 활용하였으며, 종속변수가 대학생의 사교육비 지출로서 연속변수인 경우에는 위계적 선형모형(HLM)을 활용하여 이를 살펴보았다.

한편, 대학생의 사교육비 지출을 종속변수로 한 모형의 경우 사교육에 참여하지 않는 학생들이 표본에 포함될 경우 과도한 영(0)편의가 발생하여 변수의 영향력 추정에 있어 편의가 발생할

수 있어 사교육에 참여한 학생들만을 대상으로 분석을 수행하였다.

분석 결과, 고3 시절의 사교육 참여 경험은 대학생의 사교육 참여 여부 및 사교육비 지출 모두에 유의미한 영향을 미치고 있음을 확인할 수 있었다. 즉, 고3 시절 사교육 참여경험이 있는 학생들은 그렇지 않은 학생들에 비해 1.7배 가량 사교육 참여에 대한 승산(odds)이 높은 것으로 나타났으며, 사교육비도 월평균 25.7% 더 많이 지출하는 것으로 나타났다. 이는 고3 시절의 사교육 경험이, 이후 해당 학생이 대학교로 진학하더라도, 학생의 사교육 수요에 지속적으로 영향을 미칠 수 있음을 나타낸다.

이러한 분석 결과는 고등학교 시절 경험한 사교육이 해당 학생의 자기학습능력의 저하를 가져오고, 이에 따라 대학에 진학해서도 습관적으로 사교육에 의존하게 된다는 사교육 습관화 가설(이준호, 2011)에 대해 개념적 근거를 제공해준다고 할 수 있다. 물론, 이러한 가설의 타당성 여부에 대해서는 추후 더욱 엄밀한 연구 모형의 설계를 통한 검증이 필요하지만, 본 연구의 결과는 고등학교에서 대학으로 이어지는 사교육의 연결고리에 대한 가설적 단초를 제공했다는 데서 의의를 찾을 수 있을 것으로 보인다.

한편, 사교육 참여 행위와 사교육비 지출 행위에서 나타난 독립변수들 간의 영향력 차이는 사교육과 관련된 의사결정이 단계별로 다르게 나타남을 보여준다. 우선 사교육에 대한 참여 여부를 결정하고, 이후 얼마만큼의 비용을 지불할지에 대한 결정이 이루어진다는 측면에서 볼 때, 이러한 순차적인 의사결정 단계(sequential decision making)로서의 사교육에 대한 분석은 단계별로 그 접근 방식을 달리할 필요가 있을 것으로 보인다. 본 연구 결과에 따르면 사교육에 대한 의사결정 단계별로 그러한 의사결정 행위에 영향을 미치는 변수들 및 그 영향력의 크기가 달라지므로, 대학생의 사교육 행위에 대한 정책 대안 구상 시 사교육 참여 여부에 대한 정책 대안과, 사교육비 지출 행위에 대한 정책 대안은 분리할 필요가 있을 것으로 보인다. 즉, 현재와 같이 획일적인 사교육비 경감 대책보다는 사교육 참여 자체를 줄일 것인지, 이후에 발생할 사교육비 지출 규모를 줄일 것인지에 대한 고민이 선행될 때 정부 당국의 사교육 감소 대책은 효과를 발휘할 수 있을 것으로 기대할 수 있다.

본 연구결과는 사교육 문제가 더 이상 초·중등학교만의 문제가 아님을 암시한다. 과도한 대학 졸업자 배출과 이에 따른 막대한 실업자 양산 문제는 취업을 위한 인적자본 개발이라는 측면에서 대학생 사교육을 지속적으로 유발시켜 사회문제로서 기능할 가능성이 크며, 이미 사회 곳곳에서 이러한 징후가 나타나고 있다. 이러한 현실 속에서 대학 사교육 문제의 원인을 찾기 위한 노력은 이러한 문제를 해결하기 위한 초석으로 기능할 수 있다. 즉, 사교육 문제는 더 이상 초·중등학교와 대학교로 나누어 생각할 문제가 아니라 서로가 긴밀히 연계되어 그 영향력을 지속적으로 미래에 행사하는 긴밀한 인과관계 속에 놓여 있는 것이다. 이러한 인식에 기초할 때 현재 심각하게 제기되고 있는 대학생 사교육 문제의 해결을 위한 논의를 체계적으로 시작할 수 있을 것이다.

한편, 본 연구의 경우 한국교육고용패널의 데이터를 이용함에 따라 사교육 수요 행위에 영향을 미칠 것으로 예상된 투입 변수들이 제한적이며, 이로 인해 사교육 수요의 목적 및 효과에 대한 구체적인 정보를 담고 있지 못하다는 제한점이 있다. 그리고 성향점수(propensity score)를 활용한 연구의 경우 관찰가능한 특성변수들(observable variables)을 활용하여 처치 집단과 통제 집단 간의 동질성을 꾀한다는 장점이 있지만, 처치 집단에의 참여 여부에 관찰가능하지 않은 변수(unobservable variables)들이 영향을 미친다면 이에 따른 집단 간 동질화는 애초에 의도한 수준에 이르기 어렵다는 단점이 존재한다. 이에 추후 사교육 수요에 대한 보다 엄밀한 문항 및 변수 설계가 이루어진 데이터를 바탕으로 분석을 시행한다면 더욱 정밀한 추정 결과를 도출할 수 있을 것으로 기대할 수 있을 것이다.

참고문헌

- 김지하, 박지은(2008). 대학생의 사교육 수요에 관한 탐색적 연구. **교육재정경제연구**, 17(1), 93-121.
- 남수경(2006). 사범대학 학생의 임용고사 관련 사교육비 지출 실태 분석. **교육행정학연구**, 24(4), 337-362.
- 민혜리(2003). 대학에서의 취업준비 과외학습의 유형과 비용에 관한 연구. **교육사회학 연구**, 13(3), 133-149.
- 이석원(2003). Propensity Score Matching 방법에 의한 실업자 직업훈련 사업의 효과성 평가. **한국행정학보**, 37(3), 181-199.
- 이정미(2010). 대학생의 사교육 참여와 사교육비 지출 규모의 계열별 차이 분석. **교육재정경제연구**, 19(2), 65-94.
- 이준호(2011). **대학생의 영어학습 선택 결정 요인 분석**. 제6회 한국교육고용패널 학술대회 자료집, 421-457.
- 정지선, 김훈호(2009). 대학생의 사교육 참여 및 사교육비 지출에 영향을 미치는 변수 분석. **교육재정학연구**, 18(3), 89-122
- 천세영(2001). 교육비 개념 및 분류체계에 대한 연구: 국제비교의 관점. **교육재정경제연구**, 10(2), 221-250.
- 천세영, 이석열, 이선호(2004). 대학생의 사교육비 규모와 차이 분석. **교육재정경제연구**, 13(1), 1-24.
- 잡코리아 홈페이지. 2010. 6. 17일 인용. <http://www.jobkorea.co.kr/>
- Dehejia, R., & Wahba, S. (2002). Propensity Score-Matching Methods for Nonexperimental Causal Studies. *The Review of Economics and Statistics*, 84(1), 151-161.
- Frangakis, E., & Rubin, D. (2002). Principal Stratification in Causal Inference. *Biometrics* 58(1), 21-29.
- Hong, G., & Raudenbush, S. (2005). Effects of Kindergarten Retention Policy on Children's Cognitive Growth in Reading and Mathematics. *Educational Evaluation and Policy Analysis*, 27(3), 205-224.
- Rosenbaum, R., & Rubin, D. (1983). The Central Role of the Propensity Score in Observational Studies for Causal Effects. *Biometrics*, 70(1), 41-55.

Raudenbush, S., & Byrk, A. (2002). *Hierarchical Linear Model: Applications and Data Analysis Method*. CA: Sage Publications.

* 논문접수 2011년 11월 5일 / 1차 심사 2011년 12월 10일 / 게재 승인 2011년 12월 19일

* 김동일 : 서울대학교 교육학과 및 동대학원을 수료하고 미국 미네소타대학교 교육심리학과에서 학습장애전공으로 석사 및 박사학위를 취득하였다. 현재 서울대학교 교육학과 교수, 글로벌인재연구센터 소장으로 재직 중이며, BK21 역량기반 교육혁신 사업단 참여 교수로 있다. 주요저서로는 '학습장애아동의 이해와 교육', '특수아동상담' 등이 있다.

* E-mail : dikimedu@snu.ac.kr

* 김영식 : 서울대학교 국어교육과를 졸업하고, 동 대학원 교육학과에서 석사학위를 취득하였다. 현재 동 대학원 교육학과에서 박사과정에 재학 중이다.

* E-mail : sandfom3@snu.ac.kr

* 김경선 : 대구대학교 특수교육과를 졸업하고, 호주 Macquarie University에서 석사학위를 취득하였다. 서울대학교 대학원 특수교육전공 박사과정에 재학 중이다.

* E-mail : muindok@hanmail.net

Abstract

Analysis of Private Education Experiences in High School Influences the Probability of Private Education and the Costs of Private Education in University

Kim, Dongil*

Kim, Youngsik**

Kim, Kyoungsun***

This study examined the effect of Private Education experiences of third grade in high school on the probability and the costs of the private education in university. Propensity Score was used to control the observable variables existed between the groups which underwent private education and which did not undergo private education. Furthermore, we used HLM to create 2-level hierarchical linear models to consider students are nested within university. The result of this study showed that the private education experience of third grade in high school, year's condition since entrance to university, and Propensity Score influenced positively on the probability of private education in university. In addition, the location and ranking of university were also variables which had casual effect on the probability of private education in university. On the other hand, private education experience of third grade in high school, and the year's condition since entrance to university were the variables that showed significant positive effect on the costs of private education in University.

Key words: private education, propensity score, HLM(Hierarchical Linear Models)

* Professor, Seoul National University

** Ph. D. Candidate, Seoul National University

*** Corresponding author, Ph. D. Candidate, Seoul National University