creative commons

COMMONS DEED

저작자표시-비영리-변경금지 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.

다음과 같은 조건을 따라야 합니다:

저작자표시. 귀하는 원저작자를 표시하여야 합니다.

비영리. 귀하는 이 저작물을 영리 목적으로 이용할 수 없습니다.

변경금지. 귀하는 이 저작물을 개작, 변형 또는 가공할 수 없습니다.

- 귀하는, 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

저작권법에 따른 이용자의 권리는 위의 내용에 의하여 영향을 받지 않습니다.

이것은 이용허락규약(Legal Code)을 이해하기 쉽게 요약한 것입니다.

Disclaimer

dCollection

교육학석사학위논문

국가수준의 종단자료를 활용한 방과후학교의 연속 처치효과 분석

2013년 2월

서울대학교 대학원

교육학과 교육학 전공

이 준 호

국가수준의 종단자료를 활용한 방과후학교의 연속 처치효과 분석

지도교수 박 현 정

이 논문을 교육학석사 학위논문으로 제출함
2012년 10월

서울대학교 대학원
교육학과 교육학전공
이 준 호

이준호의 교육학석사 학위논문을 인준함
2012년 12월

위 원 장 _____ (인)

부위원장 _____ (인)

위 원 _____ (인)

국문초록

이 연구의 목적은 우리나라 고등학교 1·2학년 학생 전체를 대상으로 하는 국가수준 학업성취도 평가의 두 시점 전수 종단자료를 활용하여 방과후학교 참여의 연속적 처치(sequential treatment)에 대한 모집단 인과효과를 분석하는 것이다. 본 연구는 크게 두 단계로 구성되어 있다. 1단계에서는 방과후학교 참여 선택과정의 다층적 속성에 주목하여 방과후학교 참여 여부를 결정하는 학생 및 학교 수준 요인들을 분석한다. 2단계에서는 1단계에서 얻은 다층경향점수의 역확률 가중치를 적용하여 종단적 선택편의를 교정한 후 방과후학교 참여의 횡단적·종단적 처치효과를 산출한다. 연구문제를 제시하면 다음과 같다.

연구문제 1. 방과후학교 참여 여부 결정 요인

1-1. 고1·고2 학생들의 방과후학교 참여 선택에 개인 요인보다 학교 요인이 더욱 큰 영향을 미치는가?

1-2. 고1·고2 시기 방과후학교 참여 여부에 영향을 미치는 구체적인 개인 및 학교 수준 요인은 무엇인가?

1-3. 고2 시기 방과후학교 참여 여부에 고1 시기 시간-의존적 혼동요인이 영향을 미치는 '종단적 선택편의'가 존재하는가?

연구문제 2. 방과후학교 연속적 참여의 처치효과

2-1. 고1·고2 시기 횡단적 처치효과 (single-year treatment effect): 고1·고2 시기에 방과후학교 참여한 것은 그 해 학업성취도에 각각 어떠한 영향을 미치는가?

2-2. 고1 처치 지연 효과 (lagged treatment effect): 고1 시기 방과후학교 참여의 고2 시기 학업성취도에 대한 영향은 어떠한가?

2-3. 고1·고2 연속적 참여의 배가 효과 (amplifying effect): 고1·고2

시기 방과후학교에 연속적으로 참여할 경우 방과후학교의 학업성취에 대한 효과가 더욱 증폭되는가?

주요 연구 결과를 요약하면 다음과 같다.

첫째, 방과후학교 참여 여부에 대한 개인 수준 요인의 영향보다 학교 수준 요인의 영향이 더욱 큰 것으로 나타났다. 집단 내 상관계수(ICC)가 국어·수학·영어 과목 모두에서 50%를 초과하여 방과후학교 참여 여부에 관한 전체 변산 중 절반 이상을 학교 수준 요인이 설명하였다. 과목별 MOR도 7.0~8.0 전후로 산출되어 학생 개인은 동일해도 학교가 달라짐으로써 방과후학교에 참여할 확률 자체가 평균적으로 7배 정도 차이가 났다. 이러한 결과는 단위 학교 방과후학교 운영 방침에 따라 학생들의 방과후학교 참여가 집단적으로 변한다는 사실을 보여준다.

둘째, 방과후학교 참여 여부를 예측하는 요인을 분석한 결과 지역이나 전문계고 여부 등 기본적인 학교 정보만으로도 방과후학교 참여 확률의 학교 간 차이를 52%~55% 정도 설명할 수 있는 것으로 나타났다. 그러한 학교 배경이 통제되었을 때에 '학교 교사를 믿고 성실하게 학교 공부를 수행하는 학생'일수록, 그리고 그러한 학생들이 많은 학교일수록 방과후학교 참여 확률이 증가하였다. 본 연구의 가장 특징적인 발견은 학교 수준 방과후학교 참여 확률에 대해 교장 개인의 특성이 독립적인 영향력을 갖는다는 점이었다. 교장이 여자이거나 공모제로 부임한 경우, 행정업무 및 외부기관 회의 비율이 높은 경우 그 학교의 방과후학교 참여율이 감소하였다. 이는 방과후학교 운영에 관한 학교 수준 운영 방침에 학교장의 성향이 유의하게 반영되고 있다는 점을 보여준다.

셋째, 고2 시기 방과후학교 참여 여부에 고1 시기 시간-의존적 혼동요인(time-varying confounder)이 영향을 미쳐 종단적 선택 편의가 존재하는 것으로 나타났다. 고2 시기 방과후학교 참여에 가장 분명한 영향을 미치는 것은 고1 시기 시간-의존적 혼동변수인 '방과후학교 계속 참여 희망에 대한 학교 평균'이었다. 학생들이 고1 시기에 방과후학교 계속 참여하기를 희망하는 정도의 학교 평균이 높은 학교일수록 고2 시기 방과

후학교에 참여할 학교 수준 로그-승산 평균이 증가하였다.

넷째, 방과후학교 연속적 참여의 처치 효과를 분석한 결과 고1·고2 시기 횡단적 처치효과와 고1·고2 연속적 참여의 배가효과가 존재하는 것으로 나타났다. 횡단적 처치효과의 경우, 고1 시기 방과후학교 참여는 고1 시기의 학업성취도를 100점 만점에 약 4~6점 정도 높였고, 고2 시기의 방과후학교 참여는 고2 시기 학업성취도를 약 6~9점 정도 향상시켰다. 종단적 처치효과의 경우, 고1 시기 방과후학교 참여의 고2 시기 성적에 대한 처치 지연 효과(lagged treatment effect)는 통계적으로 0으로 나타났다. 그러나 고1-고2 시기 영어 방과후학교에 연속적으로 참여하면 고1 시기 방과후학교 참여의 효과와 고2 시기 방과후학교 참여 효과의 단순합을 넘어서는 약 3점 정도의 추가적인 성적 상승 효과, 즉 배가효과(amplifying effect)가 존재하는 것으로 분석되었다. 고2 시기에만 영어 방과후학교에 참여한 학생은 100점 만점에 6점 정도 성적이 향상되지만, 고1 시기부터 참여해 온 학생은 100점 만점에 9점(6+3) 정도 성적이 향상된다는 본 연구의 결과는 교육 프로그램에 대한 연속적인 참여를 유도하는 학교 정책의 중요성을 환기시키고 있다.

주요어 : 연속적 처치효과, 다층경향점수(multilevel propensity score), 역확률 가중치(inverse probability of treatment weight), Marginal Structural Modeling

학 번 : 2009-23375

목 차

제 1 장 서론 ····· 1
제 1 절 연구의 필요성 및 목적 ····· 1
제 2 절 연구 문제 ····· 4

제 2 장 이론적 배경 및 선행연구 검토 ····· 5
제 1 절 방과후학교 참여 결정 요인의 다층적 속성 ··· 7
제 2 절 방과후학교 효과 추정의 종단적 확장 ····· 16

제 3 장 연구방법 ····· 20
제 1 절 분석 자료 ····· 20
제 2 절 변수 구성 및 기술통계 ····· 28
제 3 절 분석 방법 ····· 40

제 4 장 연구결과 ····· 53
제 1 절 방과후학교 참여 선택과정 ····· 53
제 2 절 방과후학교 연속적 참여의 처치효과 ····· 65

제 5 장 결론 ····· 68
제 1 절 요약 ····· 68
제 2 절 논의 ····· 70

참고문헌 ····· 72
Abstract ····· 77

표 목 차

<표 2-1> 방과후학교 효과 추정 관련 국내 주요 선행연구 ····· 6
<표 3-1> 국가수준 학업성취도 평가 제도의 변화 ····· 21
<표 3-2> 2009·2010 각 년도 전수자료의 처치·결과변수 결측치 비율 ····· 25
<표 3-3> 2009·2010 연계 실패 결측 사례들의 규모 및 학업성취도 차이 ····· 27
<표 3-4> 2009·2010 횡단 전수 자료에서 연도별 방과후학교 참여 ····· 29
<표 3-5> 2009-2010 종단 전수 자료에서 처치시퀀스별 방과후학교 참여 ····· 30
<표 3-6> 국가수준 학업성취도 평가의 척도점수 체계 ····· 32
<표 3-7> 교과별 학업성취도 척도점수 백분위 기술통계 ····· 33
<표 3-8> 개인 수준 예측 변수의 기술통계 ····· 36
<표 3-9> 학교 수준 예측 변수의 기술통계 ····· 39
<표 3-10> 시간 및 방과후학교 참여 변수 코딩에 따른 처치효과 정의 ····· 52
<표 4-1> 방과후학교 참여 여부의 과목별 학교 수준 군집화 정도 ····· 53
<표 4-2> 학교수준 랜덤 효과 u_{0j}의 잔차 분산과 변수군의 설명량 ····· 58
<표 4-3> 방과후학교 참여에 대한 학교 수준 변수 고정효과 ····· 59
<표 4-4> 방과후학교 참여에 대한 개인 변수의 학교 내-학교 간 고정효과 ·· 62
<표 4-5> 고2 시기 참여에 대한 고1 처치·결과·시간 의존변수들의 영향 ··· 63
<표 4-6> 고1-고2 방과후학교 연속적 참여의 처치효과 추정 결과 ····· 67

그 림 목 차

[그림 2-1] 교차적 다층 설계와 배속적 다층 설계 ····· 7
[그림 2-2] 방과후학교 참여확률과 SES_{ij}간의 가설적 관계 ····· 13
[그림 2-3] 고1 · 고2 시기 연속적 처치에 따른 잠재적 결과값 및 처치효과 ·· 17
[그림 2-4] 고2 시기 방과후학교 참여 결정 요인의 종단적 구조 ····· 18
[그림 3-1] 국가수준 학업성취도 전수 평가의 유사-종단 설계 ····· 22
[그림 3-2] 2009년 학교별 방과후학교 참여율의 분포 ····· 31
[그림 3-3] 2009년 학교별 수학 방과후학교 참여율의 지역별 분포 ····· 37
[그림 3-4] 다층경향점수와 경향점수의 커널밀도함수 비교 ····· 47
[그림 4-1] 2009년 학교 목적유형 · 성별에 따른 방과후학교 참여율 ····· 55
[그림 4-2] 2009년 중식지원학생비율에 따른 방과후학교 참여율 ····· 56
[그림 4-3] 2009년 교장 외부기관회의 비율에 따른 방과후학교 참여율 ····· 57

제 1 장 서론

제 1 절 연구의 필요성 및 목적

방과후학교가 국가 핵심 교육정책으로 제도화된 2006년 이후 현재까지 방과후학교는 일선 학교 현장에 충분히 뿌리를 내렸다. 2012년 9월 교육과학기술부에서 발표한 방과후학교 운영 현황 자료에 따르면 전국 초·중·고 11,361개 학교의 99.9%가 방과후학교를 운영하고 있다(교육과학기술부, 2012). 교육과학기술부와 각 시·도 교육청이 2011년부터 연평균 약 60% 수준의 급격한 방과후학교 자유수강권 예산 증액을 계획·실행하고 있는 사실(교육과학기술부, 2011)은 방과후학교가 정부의 핵심적인 초·중등정책 사업이라는 점을 보여준다.

이러한 방과후학교 정책의 중요성을 반영하여 학계에서도 그 정책의 효과를 실증적으로 분석하려는 작업을 지속해왔다. 방과후학교의 효과를 대규모 서베이 자료를 활용해 실증하려는 노력은 2008년을 전후로 본격화하였는데, 이는 방과후학교 변수를 비중있게 조사한 한국교육종단연구(KELS) 2차년도 자료가 2008년 연구자 일반에 공개되는 등 가용한 자료들이 점차로 증가한 맥락과 관련이 있다. 그러나 같은 연구 주제를 유사한 모형을 적용해 분석하였으나 분석 자료가 달라 다소 상이한 결과를 산출한 경우나(채재은·임천순·우명숙, 2009과 배상훈·김성식·양수경, 2010) 같은 연구 주제를 동일한 자료를 대상으로 분석하였으나 분석 모형이 달라 반대되는 결과를 산출하는 경우가 있었다(박현정·변종임·조순옥, 2009와 변수용·김경근, 2009). 그 결과 방과후학교의 효과성에 대해서는 학계에 아직 충분한 합의가 이루어지지 않은 상황이다.

방과후학교 효과를 분석한 현재까지의 연구 대부분이 가진 한계점은 그것들이 대개 1개년도 횡단적 효과 추정치만을 산출했다는 데에 있다. 그러나 고등학교 1학년 시기 방과후학교 프로그램 참여의 효과가 잠재 혹은 누적되어 있다가 2학년에 이르러 나타날 수도 있고(처치 지연 효

과), 고등학교 1학년과 2학년에 연속적으로 방과후학교에 참여하는 것이 단순히 한 개 학년에만 참여한 학생들이 얻은 효과의 단순합을 넘어설 수도 있다(배가 효과). 교육 프로그램이 연속적으로 구성되어 있고 교사가 학생들의 이전 프로그램 수혜 경험이나 성취도를 반영하여 교육과정을 설계하는 흔한 교육적 상황을 감안하면, 그러한 연속적 처치에 근거한 종단적 처치효과가 교육 현실에서 실제로 존재할 가능성이 클 것으로 예상된다.

따라서 본 연구는 국가수준 학업성취도 평가의 고1-고2 개별 학생 연계 자료를 활용하여 방과후학교의 연속적 처치(sequential treatment) 상황이 학생들의 학업성취도에 미치는 효과를 분석하는 것을 목적으로 한다. 2009년부터 실시된 국가 수준 학업성취도 평가 자료는 다른 정책 변수들에 비해 방과후학교 관련 변수를 상당히 비중 있게 조사해왔다. 여기에 더해 2009-2010에 걸친 평가 시행 초기의 정책 변경으로 인해 2009년 전체 고등학교 1학년 학생을 2010년 고등학교 2학년 시기에 반복측정한 전수 조사 결과를 사용할 수 있게 되었다. 이러한 '두 시점 전수 종단자료'는 방과후학교 연속적 참여의 효과 추정치를 표본 불일치(sample discrepancy, Cudeck & Henly, 1991)없이 산출하는 데에 있어 다시는 없을 귀중한 연구 기회를 제공한다. 이 자료를 활용하여 본 연구는 "고등학교 1학년 시기 방과후학교 참여가 고등학교 2학년 시기 학업성취도에 미치는 처치 지연 효과(lagged treatment effect)는 어느 정도인가?" 혹은 "고1-고2 시기 방과후학교의 연속적 참여가 갖는 배가효과(amplifying effect)가 존재하는가?"와 같은 연구 질문에 대답할 수 있게 되었다.

방과후학교의 연속적 처치 효과를 추정하는 본 연구는 다음과 같은 세 가지 측면에서 유용하게 활용될 것으로 기대된다.

첫째, 2008년 이후 지금까지 대규모 조사 자료를 활용한 방과후학교 정책 효과 평가 연구들이 보였던 핵심적인 단점들을 보완하여 신뢰로운 모집단 효과 추정치를 얻을 수 있다. 우리가 표본 조사 자료를 분석해서 얻은 모든 효과 추정치는 ① 변수 측정 오차, ② 모집단-표본 간 불일

치, ③ 추정 모형 불일치로 인한 편의를 일정 정도 포함하고 있다. 그러나 국가수준 학업성취도 평가 자료는 ②의 모집단-표본 간 불일치가 완전히 해소된 전수 자료이며 그에 더해 종단 자료로 인한 풍부한 정보 활용과 인과효과 모형 추정으로 ③의 추정 모형 불일치로 인한 편의도 최소화할 수 있다. 따라서 이 연구에서 얻은 방과후학교 참여의 학업성취에 대한 효과추정치는 이전 연구들과 확실히 차별화되는 '모집단 인과효과 추정치'로 방과후학교 관련 정책 결정에 중요하게 활용될 수 있다.

둘째, 연속적인 참여를 유도하는 학교 정책의 중요성에 대한 관심을 환기할 수 있다. 단순히 1개년도만 참여한 학생보다 2개년 연속으로 방과후학교에 참여한 학생의 학업성취도가 배가(amplifying)되는 효과가 크다면, 개별 학교에서는 방과후학교에 기존에 참여했던 학생들이 계속적으로 참여할 수 있도록 독려하는 방침을 고려해야할 것이다. 특히 방과후학교 자유수강권 지원금과 같은 금전적 인센티브가 고등학교 1·2학년에 걸쳐 연속적으로 참여하는 저소득층 학생들에게 더 많이 제공될 수 있도록 하는 정책의 근거로 활용될 수 있다.

셋째, 교육학 연구자들에게 횡단적 처치효과 뿐만 아니라 종단적 처치효과를 추정하는 방법의 중요성을 강조하고 이를 분석하는 새로운 방법을 소개하는 연구로서 의미를 가진다. 학교 정책의 효과를 연구하는 과정에서 1학년 시기 정책의 효과가 시간적으로 지체되어 2학년 시기에 나타날 수도 있고, 방과후학교 만족도와 같은 시간-의존적 혼동변수(time-varying confoudner)가 처치효과를 왜곡하는 요소가 될 수 있다는 점을 연구자들이 인식한다면, 교육학의 각 분야에서 여태껏 드러나지 못했던 교육적 처치효과의 종단적 측면들이 새롭게 발견될 수 있을 것이다. 본 연구에서 적용한 '역확률 가중치를 활용한 Marginal Structural Modeling'은 그러한 종단적 처치효과를 추정하는 실천적인 방법론으로서 교육학 연구자들에게 유용하게 활용될 수 있다.

제 2 절 연구문제

본 연구는 크게 두 가지 단계로 구성되어 있다. 1단계에서는 방과후학교 참여 여부를 결정하는 요인들을 분석한다. 2단계에서는 1단계에서 얻은 정보를 활용하여 선택 편의를 교정한 후 방과후학교 참여의 처치효과 분석한다. 각 단계에서 제기되는 연구 문제를 정리하면 다음과 같다.

1단계- 방과후학교 참여 여부 결정 요인 분석

① 방과후학교 참여의 학교 수준 군집화(clustering) 정도: 고1-고2 시기 학생들의 방과후학교 참여 선택은 개인 요인보다 학교 요인이 더욱 큰 영향을 미치는가?

② 방과후학교 참여 예측 요인: 고1-고2 시기 방과후학교 참여 여부에 영향을 미치는 개인 및 학교 수준 요인은 무엇인가?

③ 종단적 선택 편의 존재 여부: 고2 시기 방과후학교 참여 여부에 고1 시기 방과후학교 참여 여부・고1 시기 학업성취도・고1 시기 시간-의존적 혼동요인(time-varying confounder)이 영향을 미치는가?

2단계- 방과후학교 연속적 참여의 처치효과 분석

① 고1 횡단적 처치효과: 고1 시기 방과후학교 참여의 고1 시기 성적에 대한 영향은 어떠한가?

② 고2 횡단적 처치효과: 고2 시기 방과후학교 참여의 고2 시기 성적에 대한 영향은 어떠한가?

③ 고1 처치 지연 효과: 고1 시기 방과후학교 참여의 고2 시기 학업성취도에 대한 영향은 어떠한가? (lagged treatment effect)

④ 고1-고2 연속적 참여의 배가 효과: 고1-고2 시기 방과후학교 참여의 연속적 처치효과는 고1 및 고2 시기 처치효과들(②+③)의 단순합보다 큰가? (amplifying effect)

제 2 장 이론적 배경 및 선행연구 검토

<표 2-1>은 방과후학교 효과를 분석한 주요 선행연구들을 정리한 것이다. 선행연구들은 방과후학교가 영향을 미치는 학생의 결과변수로 학업성취도, 사교육비, 학교생활(교우관계, 교사-학생관계, 문제행동), 정의적 특성(내재적 동기, 자기효능감, 자아개념) 등을 설정하고 있다. 그 중에서도 방과후학교가 학업성취도와 사교육비 지출에 미치는 영향에 대한 분석이 주종을 이루고 있다. 이는 2006년 당시 교육인적자원부에서 제시한 방과후학교의 4가지 정책 목표 ① 학교교육의 기능 보완, ② 사교육비 경감, ③ 교육복지 실현, ④ 학교의 지역 사회화 중에서 첫 번째 학교교육의 기능 보완과 두 번째 사교육비 경감 목표에 비중을 두고 정책을 평가해왔다는 점을 보여준다.

본 연구에서 관심을 갖는 것은 특히 방과후학교 참여가 학생들의 학업성취도에 미치는 영향이다. 대부분의 선행연구들은 방과후학교 참여가 학업성취도에 긍정적인 영향을 미친다고 보고한다(박현정, 변종임, 조순옥, 2009; 채재은, 임천순, 우명숙, 2009; 배상훈, 김성식, 양수경 2010; 우명숙, 이수정, 2010; 김인숙, 변종임, 조순옥, 2010). 그러나 변수용, 황여정과 김경근(2011)은 오히려 중학생의 방과후학교 참여가 국어·영어·수학 학업성취도에 유의하게 부정적인 영향을 미친다는 전혀 상반된 결과를 제시하기도 한다.

이하에서는 방과후학교의 학업성취도에 대한 기존의 연구들에서 대부분 안고 있는 두 가지 핵심적인 문제, 즉 학교 수준 군집화의 문제와 종단적 처치의 문제를 차례대로 검토한다. 첫째는 방과후학교 참여에 대한 선택과정이 본질적으로 다층적이라는 점을 고려하지 않는 문제다. 둘째는 방과후학교의 효과를 횡단적으로만 추정했을 때 포착할 수 없는 성격의 종단적 처치효과가 존재하는 문제다.

<표 2-1> 방과후학교 효과 추정 관련 국내 주요 선행연구

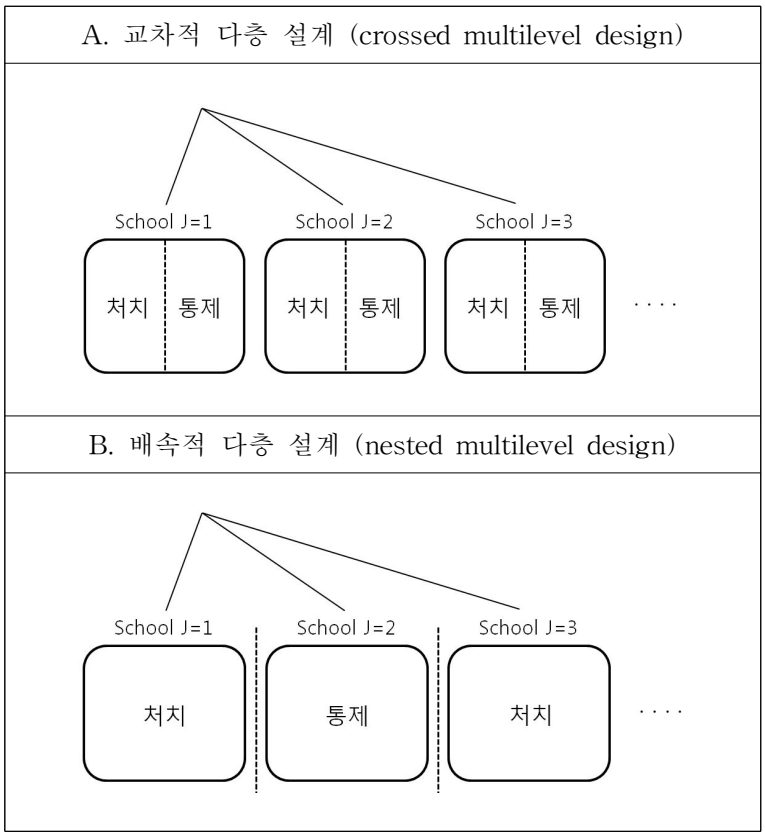
연구자 (연도)	결과변수	분석자료 및 분석대상	분석방법
박소영(2008)	· 사교육비	· 한국교육개발원 사교육 실태조사 자료(2006) · 고2 대상	HLM 배속모형
박현정, 변종임, 조순옥(2009)	· 수업이해도, 사교육비 · 내재적 동기,자기효능감 · 교우관계,교사학생 관계, 문제행동	· 한국교육종단연구 (2006-2008년 자료) · 중2, 중3, 고1 대상	HLM 성장모형
변수용, 김경근(2009)	· 사교육 참여 및 사교육비	· 한국교육종단연구 (2006-2007년 자료) · 중3 대상	OLS, logit, Tobit, Heckit, 경향점수매칭 · 지연 종속변수[1]
채재은, 임천순, 우명숙(2009)	· 학업성취도 · 사교육비	· 한국청소년정책연구원 설문조사 자료(2007년) · 고2 대상	HLM 배속모형
배상훈, 김성식, 양수경(2010)	· 학업성취도 · 사교육비	· 통계청 사교육비 조사자료(2008년) · 초중학생,일반고교생 전체 학년 대상	OLS, Multinomial logit
우명숙, 이수정(2010)	· 학업성취도(영어) · 사교육비(영어)	· 한국교육종단연구 (2005-2007년 자료) · 중3 대상	HLM 배속모형 · 지연 종속변수[1]
김인숙, 변종임, 조순옥(2010)	· 학업성취도(국어) · 사교육 시간(국어) · 교우관계, 교사학생관계, 문제행동	· 2007년 국가수준 학업 성취도평가 표집자료 · 고1 대상	HLM 배속모형
변수용, 황여정, 김경근(2011)	· 학업성취도(국 · 영 · 수)	· 한국교육종단연구 (2005-2007년 자료) · 중3 대상	경향점수 매칭 · 지연 종속변수[1]
김혜숙(2012)	· 학업성취도(국 · 영 · 수) · 사교육비(국 · 영 · 수) · 자기주도적 학습태도, 자아개념 · 교우관계, 교사학생관계	· 서울교육종단연구 (2010-2011년 자료) · 서울지역 초5, 중2, 고2	경향점수 가중회귀분석
이준호, 박현정(2012)	· 학업성취도(국 · 영 · 수)	· 2009년 국가수준 학업 성취도평가 전수자료 · 고1 대상	다층경향점수 층화, HLM 배속 모형
김성식(2012)	· 학업성취도(국 · 영 · 수)	· 2011년 국가수준 학업 성취도평가 전수자료 · 중3, 고2 대상	HLM 배속모형 · 지연 종속변수[1]

1) 지연 종속변수(lagged-dependent variable)를 독립변수로 추가. 예컨대 고2 성적을 예측하는 변수로 고1 성적을 투입하는 것. 그러나 종속변수를 종단적으로 모형화한 것은 아님.

제 1 절 방과후학교 참여 결정 요인의 다층적 속성

1. 교차적 다층 설계 (crossed multilevel design)

처치·통제 집단에 피험자들을 랜덤 할당하였을 때 두 집단 간 결과 변수의 평균 차이로 정의되는 처치의 인과효과는 학교에 학생이 배속되어 있는 다층 구조 상황에서는 더욱 복잡한 이슈를 만들어낸다. 다층 구조 상황에서의 인과효과 추론을 위한 실험 설계 구조는 기본적으로 교차적 다층 설계와 배속적 다층 설계로 구분된다(Seltzer, 2004; Kim, 2006). [그림 2-1]에 두 가지 설계 구조가 도해되어 있다.



[그림 2-1] 교차적 다층 설계와 배속적 다층 설계

교차적 다층 설계(crossed multilevel design) 하에서는 개인이 처치할당의 기본 단위가 되므로 각각의 군집 안에 처치집단과 통제집단이 동시에 존재한다. 반면 배속적 다층 설계(nested multilevel design)는 군집 전체가 처치할당의 기본 단위가 되기 때문에 학교 1이 처치집단에 할당되었다면 학교 1에 속한 모든 학생이 처치집단이 되고, 통제집단에 할당되었다면 학생들 모두 통제집단이 된다. 이러한 특성 때문에 배속적 다층 설계를 군집-랜덤할당 실험(cluster-randomized experiment)이라고도 한다(Murnane & Willett, 2011). 배속적 다층 설계에서는 특정한 교육 프로그램을 실행한 학교와 그렇지 않은 학교 간의 학업성취도 평균 차이를 통해 교육 프로그램 효과를 추정한다.

방과후학교 참여가 학업성취도에 미치는 효과를 추정하기 위한 연구 설계는 배속적 다층 설계가 아니라 전형적인 교차적 다층 설계라고 볼 수 있다. 방과후학교 참여의 기본 단위는 학생 개인이므로 한 학교 내에서 방과후학교에 참여하는 처치집단과 참여하지 않는 통제집단이 동시에 형성된다. 방과후학교 참여의 처치효과는 개별 학교 내에서 참여-미참여 집단 간 학업성취도 평균 비교를 통해 추정이 될 것이다. 따라서 이론적으로 각각의 학교마다 방과후학교 참여의 처치효과가 개별적으로 정의된다. 교차적 다층 설계 구조 하에서 SUTVA(Stable Unit Treatment Value Assumption)[1]의 위배를 생각하면 그러한 학교 내 효과(within-school effect)를 방과후학교 참여의 효과로 정의하는 것이 타당하다고

1) SUTVA는 두 가지 기본 가정으로 구성되어 있다. 첫째는 처치집단에 할당된 개체들은 동일한 처치(constant treatment)를 받는다는 가정이다. 이 가정에 따르면 방과후학교에 참여하는 학생들은 동일한 주제(예컨대 수학 교과)를 동일한 방법과 동일한 강도로 교육받는다. 둘째는 처치와 관련한 개체들 간의 간섭(interference)이 없어야 한다는 가정이다. 예컨대 방과후학교에 참여하지 않는 학생 A의 학업성취도는 방과후학교에 참여하는 그의 절친한 친구 학생 B에 의해 영향을 받지 않는다는 것이다. 학교-학생 2수준 다층구조 하에서 이 가정은 쉽게 위배된다. 방과후학교 프로그램은 학교에서 어떤 교사가 가르치는지 혹은 어떤 방식과 강도로 운영되는지에 따라 상당히 다르고, 학교 내에서 학생들끼리 서로 활발한 사회적 상호작용이 일어나 처치-통제 집단 간 경계가 모호해지는 경우도 있다.

볼 수 있다.

문제는 방과후학교 참여라는 처치 상황에 학생들이 랜덤하게 할당되지 않았다는 데에 있다. 방과후학교 참여 과정에 학생 개인의 자율적인 선택이 개입되는 부분이 존재하므로 방과후학교 효과 추정에서 선택 편의(selection bias)의 교정이 핵심적인 이슈가 된다. 만약 '학교 교사를 믿고 성실하게 학교 공부를 수행하는 학생'이 방과후학교에 참여할 확률이 높다면 방과후학교 참여 집단의 높은 성취도는 방과후학교 프로그램 참여 그 자체의 효과가 아니라 참여하는 학생들이 체계적으로 성실하고 공부 잘하는 학생들로 구성되어 있기 때문일 가능성이 크다.

교차적 다층 설계 상황에서는 개인 요인의 영향 뿐만 아니라 단위 학교의 제반 특성 차이가 학교 간 평균적인 방과후학교 참여 선택의 차이를 야기할 수도 있다(Hong & Raudenbush, 2006; Kim, 2006). 단위 학교의 방과후학교 관련 운영 방침(school policy), 학생·학부모·교사의 구성적 특징(compositional characteristics)이 그러한 원인이 된다. 예컨대 수도권의 A 고등학교는 학교장의 강력한 학력 향상 방침으로 특별히 예외적인 경우가 아니면 방과후학교 교과 보충수업에 학생이 전원 참여할 것을 결정했다고 가정하자. 한편 인근의 B 고등학교는 학생들의 자기 결정권을 존중하는 교원노조 가입 교사들이 60%를 차지하고 있어 방과후학교 참여 결정을 전적으로 학생들의 개인 결정에 맡겼다고 가정하자. A 고등학교와 B 고등학교에 다니는 두 학생이 개인 배경 변수들에서 정확히 동일하더라도 이러한 학교 수준 특성에 따라 방과후학교 참여 확률은 현저히 달라질 것이다.

특히 고등학교에서는 방과후학교 참여 여부의 학교 수준 군집화(clustering) 정도가 개인 수준의 영향력을 압도할 가능성이 높다. 고등학교의 경우 수학·영어·국어 교과 방과후학교는 사실상 학교 교사가 진행하는 방과 후 보충수업의 의미를 가지는 경우가 많기 때문이다(한국일보, 2009; 경향신문, 2012). 단위 학교에서 학교장이나 교사들이 전체 학생들의 학업 성적을 높이려는 노력으로 가장 손쉽게 선택하고 조작할 수 있는 학교 방침 중 하나가 방과후학교 프로그램에 학생들을 참여하도록

독려 혹은 강제하는 것이다.

기존의 연구들에서는 이러한 학교 수준 특성의 영향력이 잘 반영되지 못했다. 경향점수 분석을 활용한 연구들(변수용, 김경근, 2009; 변수용, 황여정, 김경근, 2011; 김혜숙, 2012)에서 그러한 문제점이 특히 두드러진다. 이 연구들은 로지스틱 회귀모형으로 방과후학교 참여 확률을 구하는 과정에서 개인 배경 변수만을 투입하고 학교 수준 변수는 지역 요인만을 고려한다. 사실상 개인 특성들이 동일하면 방과후학교 참여 확률이 동일하다고 가정하는 것이다. 그러나 다층 로지스틱 모형에서 추정되는 학교 수준 랜덤 효과는 관찰되지 않은 학교 수준 변수들의 종합적인 효과를 의미하므로(Rabe-Hesketh & Skrondal, 2012), 단층 로짓 모형을 활용한 경향점수 산출 과정에서 이 랜덤 효과를 추정하지 않는 것은 사실상의 누락 변수 편의(omitted variable bias)를 의미한다. 그 편의의 정도는 방과후학교 참여 선택에 대한 학교 수준 요인의 영향력이 커질수록 더욱 심각해질 것이다.

이러한 문제는 경향점수 추정에 개인 수준 요인과 학교 수준 요인을 종합적으로 고려하는 다층경향점수(multilevel propensity score, Hong & Raudenbush, 2006; Kim & Seltzer, 2007) 추정의 필요성으로 연결된다. 국내 방과후학교 연구에서는 이준호와 박현정(2012)의 연구에서 다층경향점수를 활용한 사례가 있으나 개인 수준 요인과 학교 수준 요인이 각각 어느 정도의 영향을 미치는지 분석하지 못한 한계가 있다. 따라서 방과후학교 선택에 영향을 미치는 개인 및 학교 수준의 영향력을 분해하고 적절히 계량화해볼 필요가 있다.

2. 다층적 영향 요인 분석에서 중심점 교정의 문제

방과후학교 효과 추정시에 연관 요인들이 공변량으로 투입되는 경우는 많지만 방과후학교 참여 여부 자체에 대한 영향 요인을 본격적으로 분석한 연구는 매우 적다. 국내에서는 김경근과 황여정(2009)이 한국교육종단연구의 중3 학생 표본을 대상으로 방과후학교 참여 결정 요인을

분석한 것이 유일하다. 그렇지 않은 경우에는 방과후학교 참여에 대한 경향점수 추정을 위해 평면적인 로지스틱 회귀분석을 시행한 결과를 부가적으로 제시한 것이 발견되는 정도다(변수용, 김경근, 2009; 변수용, 황여정, 김경근, 2011; 김혜숙, 2012).

그러나 방과후학교 참여 여부에 대한 선택과정이 다층적 속성을 갖는다는 점을 간과한 결과, 선행연구들에서 제시한 방과후학교 참여 결정요인에 대한 추정치는 심각한 결함을 가지고 있다. 그 문제는 방법적으로는 중심점 교정(centering)의 문제로 요약된다. 여기에서는 학생 및 학교 수준 변수들에 있어서 선행연구들의 추정치가 어떻게 잘못되었는지를 살피고 새로운 변수와 분석 방법의 필요성을 제기하고자 한다.

기존 방과후학교 연구들은 가정의 사회경제적 배경이나 학업성적이 좋은 학생들일수록 방과후학교에 참여할 확률이 낮고, 반대로 가정의 사회경제적 배경과 학업성적이 좋지 않은 학생들일수록 방과후학교에 참여할 확률이 높다고 보고한다(김경근, 황여정, 2009; 김혜숙, 2012). 특히 방과후학교 참여 학생들이 평균적으로 소득수준이 낮다는 김경근과 황여정(2009)의 연구는 방과후학교 효과 관련 문헌에서도 광범하게 인용되며 재생산되고 있다. 그러나 김경근과 황여정(2009)의 연구는 전체 평균 중심점 교정(grand-mean centering)을 하였기 때문에 소득과 같은 학생 수준 변수의 회귀계수를 의미있게 해석해서는 안 된다. 다음의 두 가지 모형을 살펴보자.

$$P(T_{ij}=1) = \beta_{0j} + \beta_{1j}(SES_{ij} - \overline{SES_{..}}) + u_{0j} + r_{ij} \quad \cdots \text{ 모형(1)}$$

$$P(T_{ij}=1) = \beta_{0j} + \beta_{1j}(SES_{ij} - \overline{SES_{.j}}) + \beta_{2j}\overline{SES_{.j}} + u_{0j} + r_{ij} \quad \cdots \text{ 모형(2)}$$

모형(1)과 모형(2) 모두 j번째 학교 i번째 학생이 방과후학교에 참여($T_{ij}=1$)할 확률을 그 학생의 사회경제적 배경(SES_{ij})으로 예측하는 모형이다. β_{1j}가 우리가 얻고자 하는 두 변수의 연관성에 대한 추정치다. u_{0j}는 학교 수준 오차항, 혹은 랜덤효과(random-effect)이고 r_{ij}는 학생 수준

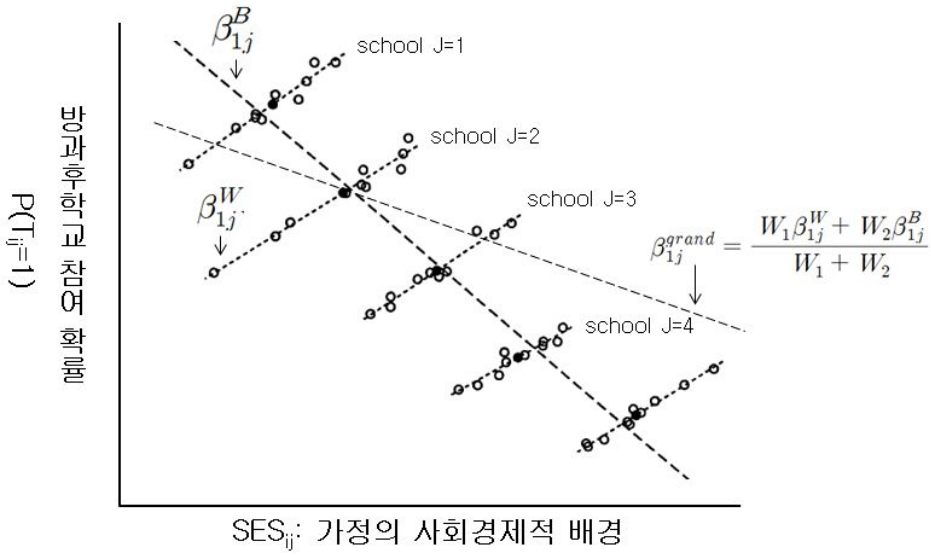
오차항이다. 모형(1)은 개인의 SES_{ij} 값에서 전체 학생 평균 $\overline{SES_{..}}$을 빼준 예측값을 투입한 전체 평균 중심점 교정(grand-mean centering) 모형이다. 모형(2)는 개인의 SES_{ij} 값에서 그 학생이 속한 학교 j의 평균적인 $\overline{SES_{.j}}$를 차분해준 변수를 투입한 집단 평균 중심점 교정(group-mean centering) 모형이다.

김경근과 황여정(2009)의 연구는 기본적으로 모형(1)을 적용하였다. 이러한 방식으로 추정한 β_{1j}은 Raudenbush와 Bryk(2002)에 따르면 "해석 불가능한 혼합물(an uninterpretable blend)"이다. 전체 평균 중심점 교정을 통해 산출된 β_{1j}은, 같은 학교 내에서 SES_{ij}가 1단위만큼 다른 두 학생의 방과후학교 참여 확률의 차이를 의미하는 학교 내 효과(within-school effect, β_{1j}^{W})도 아니고, 학교 평균 $\overline{SES_{.j}}$가 1단위만큼 차이나는 두 학교의 방과후학교의 평균적 참여 확률의 차이를 의미하는 학교 간 효과(between-school effect, β_{1j}^{B})도 아니다. 그 뿐 아니라 학교 단위 군집성을 무시하고 평면적인 OLS 회귀분석으로 추정한 전체 효과(total effect, β_{1j}^{T})와도 그 의미가 다르다. 전체 평균 중심점 교정을 통해 산출되는 1수준 효과 추정치 β_{1j}는 본질적으로 β_{1j}^{W}과 β_{1j}^{B} 두 가지 요소를 가중적으로 혼합하여 산출된 의미가 불분명한 추정치[2]다(Raudenbush & Bryk, 2002).

이를 좀더 개념적으로 살펴보면 왜 선행연구들에서 가정의 사회경제적 배경과 학업성적이 높을수록 방과후학교 참여 확률이 감소한다고 잘못된 결론을 내리고 있는지 이해할 수 있다. 방과후학교 참여 확률과 가정의 사회경제적 배경의 가설적 관계를 나타낸 [그림 2-2]를 살펴보자.

---

2) 예측변수를 전체 평균 중심점 교정(grand-mean centering)을 했을 때 1수준 회귀계수는 $\beta_{1j}^{grand} = (W_1\beta_{1j}^{W} + W_2\beta_{1j}^{B})/(W_1 + W_2)$와 같이 표현된다(Raudenbush & Bryk, 2002). W_1과 W_2는 각각 β_{1j}^{W}과 β_{1j}^{B}의 추정의 정확도(precision)를 의미하며, W_1은 자료의 총 사례수(Jn)에, W_2는 2수준 학교 수 J에 비례한다.



[그림 2-2] 방과후학교 참여확률과 SES_{ij} 간의 가설적 관계

학교 간 평균을 연결하는 선인 β_{1j}^{B}이 보여주듯이 가정의 사회경제적 수준이 평균적으로 높은 학교일수록 그 학교의 평균적인 방과후학교 참여 확률은 낮아질 것이다. 서울 지역이 대표적인 예다. 서울은 다른 지역보다 학교 평균 사회경제적 배경이 높은 학교들이 밀집해있고 학교 주변의 사교육 기관 공급도 다른 지역보다 현저히 많아 방과후학교 참여 확률이 다른 지역보다 평균적으로 낮다. 지역 및 학교 수준 효과를 모두 통제하면 β_{1j}^{W}와 같이 단위 학교 내에서 방과후학교 참여확률과 SES_{ij}의 관계가 일관되게 양(+)의 상관관계를 나타낸다고 가정하여도, 김경근과 황여정(2009)이 사용한 한국교육종단연구 자료처럼 학교당 표집크기가 작은 경우에는 β_{1j}^{W}의 추정의 정확도인 W_1이 작아져 전체 평균 중심점 교정한 예측변수의 회귀계수 추정치가 β_{1j}^{B}에 가까운 음(-)의 상관관계를 보일 수 있다. 김혜숙(2012)의 연구와 같이 학교 단위 군집성을 무시한 평면적인 로지스틱 회귀분석 또한 [그림 2-2]와 같은 자료 구조 하에서는 음의 값을 갖는 β_{1j}^{T}를 산출할 것이다.

이러한 문제는 2수준 오차항 u_{0j}와 1수준 예측변수 SES_{ij}가 상관을 갖

는 학교 수준 내생성(school-level endogeneity)의 문제로 요약할 수 있다(Rabe-Hesketh & Skrondal, 2012). 앞서 살펴본 바와 같이 방과후학교에 참여할 학교 수준의 관찰되지 않은 특성을 요약하는 랜덤효과 u_{0j}는 SES_{ij}가 증가할수록 감소한다($Corr(u_{0j}, SES_{ij}) < 0$). 달리 말하면, SES_{ij}가 높은 서울의 학교가 갖는 평균적인 방과후학교 참여 성향(u_{0j})은 SES_{ij}가 낮은 읍・면지역의 학교가 갖는 평균적인 방과후학교 참여 성향($u_{0j'}$)보다 낮다. 이러한 학교 수준의 내생성 문제를 해결하지 않으면 '서울 지역의 잘 사는 학생과 읍면 지역의 못 사는 학생 간의 방과후학교 참여 확률을 비교'하여 잘 살수록 방과후학교에 참여할 확률이 낮다고 결론내리는 오류를 범하게 된다. 이러한 오류를 극복하기 위해서는 '서울 지역의 동일한 학교에 다니지만 소득 수준이 다른 두 학생의 방과후학교 참여 확률을 비교'하는 학교 내 효과 추정치를 구해야 한다.

이를 위해서는 집단 평균 중심점 교정(group-mean centering)을 하거나 2수준에 1수준 변수의 학교별 평균 $\overline{SES_{.j}}$를 투입하여 2수준 랜덤효과 u_{0j}와 1수준 예측변수들의 상관을 통제해야 한다. 모형(2)는 그 두 가지 처방 모두를 적용한 모형이다. 모형(2)를 반영하여 방과후학교 참여 확률에 대한 가정의 사회경제적 배경의 학교 내 효과(within- school effect, β_{1j}^{W})를 구하면 [그림 2-2]에서 제시된 바와 같이 정적인 관계를 보일 것으로 예상된다. 실제로 가정 배경이 좋지 않은 방과후학교 자유수강권 필수지원대상 저소득층 학생들의 방과후학교 출석률이 저조하다는 사실이 일선 현장으로부터 꾸준히 보고되어 왔다(서울특별시교육청, 2010). 해외의 연구 결과에서도 가정 형편이 좋지 않은 학생들이 형제자매를 돌보거나 직접 돈을 벌어야하는 경우 방과후 프로그램 참여가 저해되는 것으로 나타난다(Lauver & Little, 2005). 2009년 고등학교 1학년 학생의 전수자료를 활용하여 학교 내 효과 추정치를 구한 이준호와 박현정(2012)의 연구도 결손 가정의 학생일 경우 방과후학교에 참여할 승산이 참여하지 않을 승산에 비해 70% 수준인 것으로 보고하고 있다. 이것은 가정의 사회경제적 배경이 좋지 않은 학생들이 방과후학교에 더 많이

참여한다는 기존 연구들의 주장과 정반대되는 이야기다.

학생의 학업성취도가 방과후학교 참여에 미치는 영향 또한 가정의 사회경제적 배경과 유사한 맥락에서 이해할 수 있다. 일반적으로 방과후학교라는 학교 수준 정책에 참여할 수 있는 학생 개인의 자율권이 주어졌을 때, 성적이 좋은 학생이 같은 학교의 성적이 좋지 않은 학생보다 방과후학교에 참여할 확률이 더 높다. 이준호와 박현정(2012)의 연구에서는 학생 개인 수준 변수들의 학교 내 효과 경향성을 '학교 교사를 믿고 성실하게 학교 공부를 수행하는 학생'일수록 학교에서 제공하는 교과 방과후학교 프로그램에도 참여할 확률이 높아지는 것으로 요약한다. 학교교육에 대한 신뢰 및 학교 공부 태도가 고등학생들의 교과 방과후학교 프로그램 참여의 중요한 예측변수라는 것이다.

이상의 논의와 같이, 방법적 측면에서 다층 로지스틱 모형에 투입된 예측변수의 중심점 교정(centering) 방식에 따라 그 의미가 전혀 다른 효과 추정치를 산출할 수 있다는 점, 이론적으로 기존 연구들에 상반되는 연구결과가 존재한다는 점에 유의하여 방과후학교 참여 여부 결정 요인을 분석해볼 필요가 있다. 이와 관련된 실제적인 중심점 교정 방식은 분석 방법 절에서 설명한다.

제 2 절 방과후학교 효과 추정의 종단적 확장

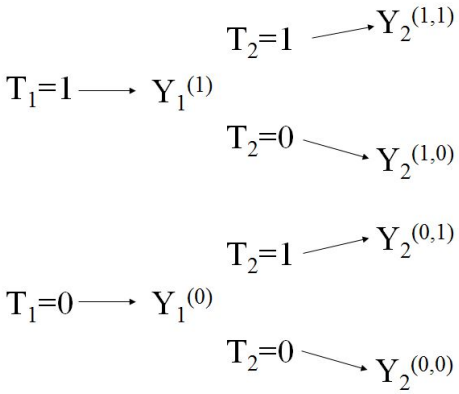
교육 프로그램은 많은 경우 연속적으로 구성되어 있다. 예컨대 영어 과목의 경우 영문법의 이론을 가르치는 프로그램 A를 고등학교 1학년 시기에 들었다는 가정 하에 수능 영문법 문제 집중 풀이반 프로그램 B를 고등학교 2학년 시기에 구성할 수 있다. 방과후학교 프로그램은 이와 같이 의도적인 연속적 프로그램으로 구성된 경우가 많으며 이는 단발성으로 그치는 고립된 프로그램을 여러 개 운영하는 것보다 교육적으로 더욱 바람직하다. 따라서 방과후학교 참여의 효과를 추정할 때 이러한 프로그램의 연속성을 감안하여 종단적으로 분석을 확장해볼 필요가 있다.

[그림 2-3]은 방과후학교의 2개년 연속적 처치 상황을 개념적으로 나타낸 것이다. T_1은 고1 시기 방과후학교 참여 여부, T_2는 고2 시기 방과후학교 참여 여부를 나타낸다. Y_1은 고1 시기 학업성취도, Y_2는 고2 시기 학업성취도의 잠재적 결과값(potential outcome)이다. T_1의 총 두 가지 처치 상황 $T_1=0$, $T_1=1$에 따라 총 2개의 잠재적 결과값 $Y_1^{(0)}$, $Y_1^{(1)}$이 정의된다. T_1과 T_2의 연속적 처치는 총 4가지 처치 시퀀스(sequence)의 조합 (0, 0), (0, 1), (1, 0), (1, 1)이 생기기 때문에 잠재적 결과값도 $Y_1^{(0,0)}$, $Y_1^{(0,1)}$, $Y_1^{(1,0)}$, $Y_1^{(1,1)}$ 4개가 정의된다. 각각의 처치효과는 이 잠재적 결과값들의 차이로 정의된다.

횡단자료를 통해 구할 수 있는 처치효과는 고1 시기 횡단적 처치효과 δ와 고2 시기 횡단적 처치효과 Δ_2다. 몇몇 연구들(변수용, 김경근, 2009; 우명숙, 이수정, 2010; 변수용, 황여정, 김경근, 2011; 김성식, 2012)에서 종단 자료의 장점을 이용하여 전년도 학업성취도와 같은 지연 종속변수(lagged dependent variable)를 차년도 학업성취도를 예측하는 독립변수로 투입한 시도가 관찰된다. 그러나 이러한 지연 종속변수의 추가는 결국 횡단적 효과 추정치 δ나 Δ_2를 좀 더 엄밀하게 구하기 위한 사전-처치 통제변수로서만 기능한다. 따라서 종단적 혹은 연속적 처치의 효과를 모형화한 것이라고 보기는 어렵다.

그러나 처치 상황을 2개년으로 확장하면 고1 시기 처치의 효과가 고2

시기에 나타나는 '처치 지연 효과(lagged treatment effect)' Δ_1이나, 2개년 연속으로 참여했을 경우 Δ_1과 Δ_2의 단순합에 더해 추가적으로 존재하는 '배가 효과(amplifying effect)' Δ^*를 정의할 수 있게 된다. 이는 방과후학교 참여 여부와 학업성취도에 대한 2개년 이상 종단 데이터에서만 정의될 수 있는 처치효과다.



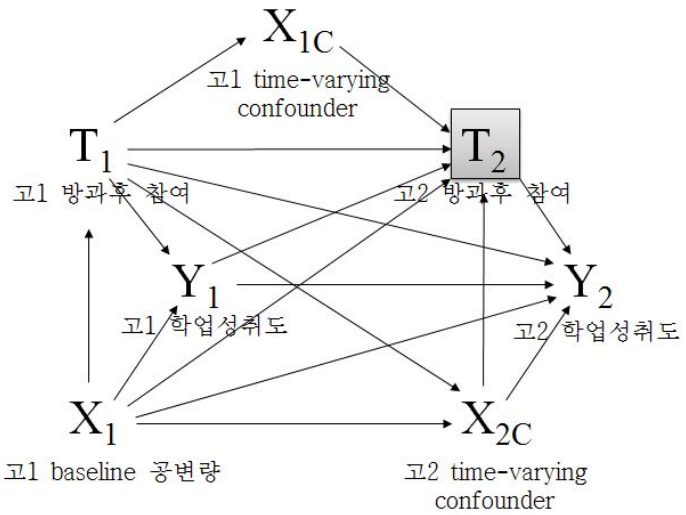
[그림 2-3] 고1·고2 시기 연속적 처치에 따른 잠재적 결과값 및 처치효과

① 고1 방과후학교 참여의 고1 학업성취도에 대한 처치효과 : 고1 시기 참여-미참여 학생 간 기댓값 차이	$E(Y_1^{(1)} - Y_1^{(0)}) = \delta$
① 고1 방과후학교 참여의 고2 학업성취도에 대한 처치효과: 고2 미참여 학생 중 고1 참여-미참여 학생들 간 기댓값 차이	$E(Y_2^{(1,0)} - Y_2^{(0,0)}) = \Delta_1$
① 고2 방과후학교 참여의 고2 학업성취도에 대한 처치효과 : 고1 미참여 학생 중 고2 참여-미참여 학생들 간 기댓값 차이	$E(Y_2^{(0,1)} - Y_2^{(0,0)}) = \Delta_2$
① 고1-고2 연속적 참여의 처치효과: 고1-고2 연속참여 학생과 연속 미참여 학생들 간 기댓값 차이	$E(Y_2^{(1,1)} - Y_2^{(0,0)}) = \Delta_1 + \Delta_2 + \Delta^*$

<표 2-1>의 선행연구들에서 방과후학교 효과를 종단적으로 확장한 연구는 박현정, 변종임과 조순옥(2009)의 연구가 유일하다. 그러나 이 연구는 인과추론의 측면에서는 상당히 취약하다. 해당 연구가 추정치의 인

과성을 주장하려면 연속적 비혼동성(sequential unconfoundedness, Raudenbush, Hong & Rowan, 2002)이라는 매우 강한 가정을 충족해야 한다. [그림 2-4]의 맥락에서 보면 고2 시기 방과후학교 참여 여부 T_2가 가정배경과 같은 사전-처치 변수들은 물론 고1 시기 방과후학교 참여 여부 T_1이나 고1 시기 학업성취 Y_1과도 무관하다고 보는 것이다. 이 가정을 충족하려면 고1 시기에 방과후학교 참여-미참여 집단 간 랜덤 할당을 하고, 다시 이 학생들을 모아 고2 시기에 참여-미참여 집단 간 랜덤 할당을 하는 경우에 가능하다. 그래서 연속적 비혼동성 가정을 연속적 랜덤할당 가정(sequential randomization assumption, Robins, 1987)이라고도 한다.

그러한 상황과는 다르게 현실에서는 방과후학교 참여 여부에 대한 연속적인 자기 선택이 발생한다. 이를 그림으로 나타내면 다음과 같다.



[그림 2-4] 고2 시기 방과후학교 참여 결정 요인의 종단적 구조

[그림 2-4]는 고2 시기 방과후학교 참여 선택과정 모형의 종단적 구조를 보여준다. 고1 시기 방과후학교에 참여한 학생은 고2 시기에도 계속 참여할 가능성이 높다($T_1 \rightarrow T_2$). 고1 시기의 좋은 성적은 학생이 고2 시

기에도 방과후학교에 참여해 열심히 공부해야겠다는 생각을 심어줄 수 있다($Y_1 \rightarrow T_2$). 또한 고1 시기 방과후학교 참여는 '방과후학교 참여 만족도'에 영향을 미치고 이는 다시 고2 시기 방과후학교 참여에 영향을 미칠 것이다($T_1 \rightarrow X_{1c} \rightarrow T_2$). 고2 시기 '수학 과목에 대한 흥미 수준'은 고1 시기 수학 방과후학교 참여로 인해 증대되어 고2 시기 수학 방과후학교 참여에 영향을 미칠 수 있다($T_1 \rightarrow X_{2c} \rightarrow T_2$). 이렇게 고1 시기 방과후학교 참여 여부(T_1)에 영향을 받아 형성된 동시에 고2 시기 방과후학교 참여 여부(T_2)에 영향을 미치는 요인들 X_{1c}, X_{2c}를 시간-의존적 혼동요인(time-dependent confounder)이라고 한다(Robins, Hernan, & Brumback, 2000).

이처럼 고2 시기 방과후학교 참여 여부 T_2에 영향을 미치는 종단적 요인들이 매우 다양하기 때문에 이로 인한 선택 효과가 적절히 배제되어야 T_2의 효과를 인과적으로 해석할 수 있다. 앞서 설명한 T_2의 연속적 비혼동성 가정은 관찰된 고1 시기 배경변수 X_1, 고1 시기 방과후학교 참여 T_1, 고1 시기 학업성취도 Y_1, 그리고 시간-의존적 혼동요인 X_{1c}와 X_{2c}를 통제했을 때 고2 시기 방과후학교 참여 T_2가 관찰되지 않은 다른 요인들과 완전히 독립이라는 것을 의미한다. 시간의존적 혼동요인을 적절히 통제하지 않으면 고2 시기 학업성취도의 차이는 고2 시기 방과후학교 참여 여부에 의한 차이가 아니라 혼동요인들에 의한 크고 작은 선택편의를 반영하는 것이라고 볼 수 있다. 그러한 선택편의는 시간이 경과함에 따라 누적적으로 증가할 것이므로 박현정, 변종임과 조순옥(2009)의 연구에서 나타난 방과후학교 참여의 시간에 따른 상호작용 효과는 시간-의존적 변수에 의한 동적 선택 편의(dynamic selection bias)를 반영하는 것일 가능성이 크다.

제 3 장 연구방법

제 1 절 분석자료

1. 2009-2010 국가수준 학업성취도평가 전수 자료

국가수준 학업성취도 평가(National Assessment of Educational Achievement, NAEA)는 우리나라 초·중등학교 학생들의 교과별 학업성취도를 측정하기 위해 한국교육과정평가원에서 주관하는 전국 단위 시험이다. <표 3-1>에 정리된 바와 같이 이 제도는 1998년 한국교육과정평가원의 설립과 동시에 기본 계획이 수립된 이후 2000년부터 2008년까지 평가 대상 학년의 1~5%를 표집하여 평가하는 방식으로 시행되어 왔다. 그러나 2008년 이명박 정부 출범 이후 표집 평가와 전수 평가를 병행하는 과도기를 거쳐 2009년부터는 해당 학년의 모든 학생들이 평가 대상이 되는 전수 평가 체제로 변경되어 2012년 현재에 이르고 있다.

본 연구는 국가수준 학업성취도평가 고등학생 자료 중 2009, 2010년 각 년도 전수 자료와 2009-2010 개별 학생 연계 자료를 분석한다. 2009년 전수 자료는 2,214개 학교 643,106명의 학생 데이터, 2010년 전수 자료는 2,200개 학교 631,362명의 학생 데이터를 포함하고 있다. 2009-2010년 학생 수준 연계 자료는 2,189개 학교 565,543명의 학생 자료를 포함하는 '두 시점 전수 종단 자료'다. 60만 명이 넘는 고등학교 1학년 전체 학생 모집단이 두 시점(고1·고2)에 걸쳐 학생 개인 수준에서 연계된 매우 보기 드문 귀중한 자료다.

평가 대상 학년의 모든 학생을 조사한 전수 자료를 이용하는 것이 가능하게 된 것은 학교교육의 책무성(accountability) 강화라는 작금의 정책적 흐름과 관련이 있다. 미국의 낙오 학생 방지법 (No Child Left Behind, NCLB)과 같은 학교 책무성 강화 정책은 1) 모든 학교·모든 학생을 대상으로 하는 대단위 전수 평가, 2) 평가 결과에 근거한 개별 학

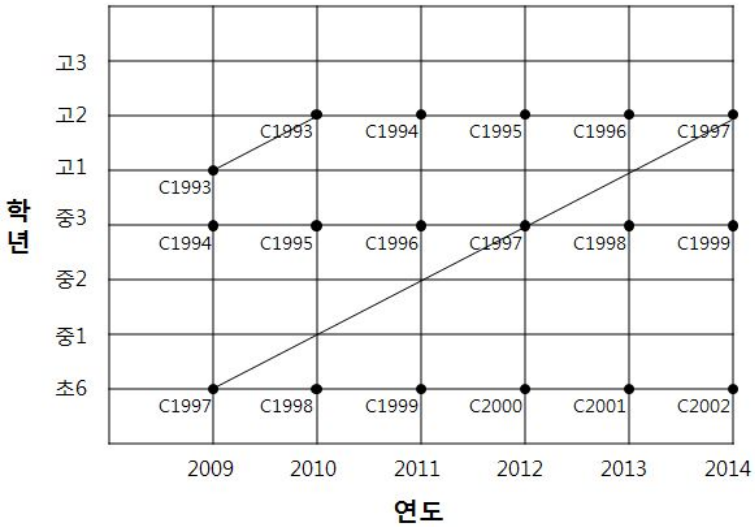
<표 3-1> 국가수준 학업성취도 평가 제도의 변화

구분	연도	개요	대상	영역	시기
표집 평가 시기	1998	• 한국교육과정평가원 설립 • 학업성취도평가 기본계획 수립			
	2000 -2002	• 학년별 1% 표집 • 예비검사 및 본검사로 검사틀 확정 • 교육부 요청에 따라 2002년 고등학교 평가대상 전환 (고2→고1)	초6 중3 고1 고2	국어/수학/ 영어/사회/ 과학	각 년도 6월 하순경
	2003 -2007	• 학년별 약 3~5% 정도 표집 • 검사동등화 방법 개발 및 적용 • 집중교과/기본교과 이원화 체제 운영	초6 중3 고1	국어/수학/ 영어/사회/ 과학	각 년도 10월 중순경
과도기	2008	• 평가원 주관 표집평가와 시도교육청 주관 전수평가의 병행 • 2008년 전수평가 자료 미공개	초6 중3 고1	국어/수학/ 영어/사회/ 과학	2008년 10월 14-15일
전수 평가 시기	2009	• 전수평가 체제로의 전면적 전환 • 국가수준에서 전수자료 축적 시작 • 기초학력미달학생 밀집학교 지원 정책	초6 중3 고1	국어/수학/ 영어/사회/ 과학	2009년 10월 13-14일
	2010	• 고등학교 평가대상 전환 (고1→고2) • 평가 시기 전환 (10월→7월) • 평가 영역 조정 (고2:국어/수학/영어) • 2010년 11월 학교정보공시 시작	초6 중3 고2	국어/수학/ 영어/사회/ 과학	2010년 7월 13-14일
	2011	• 평가 영역 조정 (초6:국어/수학/영어) • 학교 수준 등급별 향상도 산출·공시 • 향상도에 따른 행·재정적 책무성 정책 연계	초6 중3 고2	국어/수학/ 영어/사회/ 과학	2011년 7월 11-12일

교의 학업성취 수준 산출, 3) 평가 결과의 학교 수준 정보 공개, 4) 평가 결과와 행·재정적 지원의 연계라는 네 가지 요소를 통해 작동한다. 표집 학교 선정을 통한 평가는 학력 수준의 전체적인 변화 추이를 보는 데에는 효율적일 수 있으나 개별 학교마다 기초학력 미달 비율을 산출할 수가 없기 때문에 위와 같은 책무성 관리 시스템을 구축하는 기반으로 활용될 수는 없다. 이에 따라 이명박 정부는 집권 직후인 2008년 표집 평가였던 국가수준 학업성취도 평가를 전수 평가로 확장하고 『교육관련 기관의 정보공개에 관한 특례법』을 제정하여 2010년 11월부터 학교별

평가 결과를 공개하도록 하였다. 또한 2009년부터 축적되기 시작한 전수 자료를 연구자들에게 공개하여 교육과정·교육정책·교수학습 개선 등 교육의 질 관리에 유용한 정보를 생산할 수 있도록 하였다.

이러한 전수 자료가 두 시점에 걸친 종단 자료로 구축이 된 것은 2009~2010년 제도 시행 초기의 갑작스런 정책 변경에 의해 가능했다. 기본적으로 국가수준 학업성취도 평가는 같은 학생을 추적하여 매년 반복적으로 측정하는 것이 아니라 매해 특정 학년이 된 학생 집단을 측정하는 유사-종단 설계(quasi-longitudinal design)에 기초한다(Willms & Raudenbush, 1989; 최길찬, 2010). [그림 3-1]은 그러한 국가수준 학업성취도 평가의 코호트 간 설계 특성을 나타낸 것이다.



[그림 3-1] 국가수준 학업성취도 전수 평가의 유사-종단 설계

[그림 3-1]을 보면 2009년에는 초등학교 6학년·중학교 3학년·고등학교 1학년 학생들이 평가 대상 집단이었다. 2009년 초등학교 6학년 학생들은 출생 연도가 1997년인 학생들의 집단(C1997), 중학교 3학년 학생들은 출생 연도 1994년인 집단(C1994), 고등학교 1학년 학생들은 출생 연도 1993년인 집단(C1993)이다. 국가수준 학업성취도 평가는 이 출생

코호트의 학생들을 매년 추적 조사하는 것이 아니라 해당 연도에 초6·중3·고1이 된 학생들을 조사하므로 매년 새로운 출생 코호트가 진입하여 평가 대상이 되는 구조다. 예컨대 2009년에는 출생 연도 1997년인 학생 집단(C1997)이 초등학교 6학년이 되어 시험을 보지만, 이듬해인 2010년에는 출생 연도 1998년인 학생 집단(C1998)이 초등학교 6학년이 되어 시험을 보게 된다[3].

그런데 2010년부터는 고등학교 1학년이 아니라 고등학교 2학년을 평가하는 것으로 제도가 변경되었다. 이에 따라 2009년 고1이었던 출생 연도 1993년인 학생 집단(C1993)이 2010년 고2 시기에도 시험을 보게 되어 장기적인 평가 설계 구조에서는 전혀 의도하지 않았던 이례적인 2개년 연속 전수 종단 자료가 구축이 되었다. 국가수준 학업성취도평가의 유사-종단 설계 구조를 생각하면 이러한 고1·고2 2개년 연속 전수 종단 자료는 2009년 이전의 표집 평가 시기에도 없었고 앞으로도 없을 것으로 예상된다. 왜냐하면 [그림 3-1]에서 보듯이 중3 특정한 출생 코호트의 전수 성적 자료가 고2 시기의 자료와 연계되기 위해서는 그 출생 코호트 학생들이 고2가 될 때까지 최소한 2년이 경과해야 하기 때문이다.

방과후학교 참여의 학업성취도에 대한 연속적 처치 효과를 추정하려는 본 연구에서 이러한 2개년 연속 전수 종단 자료는 종단적 처치효과 정의의 명확성, 외적 타당성(external validity), 그리고 내적 타당성(internal validity)을 확보하는 데에 큰 장점을 지닌다. 첫째로, 두 시점 간 처치의 시간 간격이 2년을 넘지 않는다는 점은 처치지연효과나 배가효과와 같은 종단적 처치효과의 의미를 명확히 해준다는 이점이 있다. 만약 두 시점 연속 처치가 고1과 고3 시기로 구성되었다면, 고3 시기에 포착되는 처치 지연 효과가 고1 시기 방과후학교 참여의 효과인지, 아니면 정보가 누락된 고2 시기 참여의 효과인지 알 수가 없을 것이다. 둘째로, 분석 자료가 원칙적으로 평가 대상 학년의 모든 학생을 평가하는 전

3) 이와 같은 유사-종단 설계는 학생 개인의 학업성취 성장이 아니라 학교의 평균적인 수행의 성장을 추론하는 것에 초점을 맞춘 설계다(최길찬, 2010). 이는 국가수준 학업성취도 평가가 학생 개인보다는 학교 수준 학력 향상도에 근거한 책무성 평가 자료로서의 성격이 더욱 크다는 점을 보여준다.

수 자료이므로 연구 결과의 일반화 가능성을 의미하는 외적 타당성에 대한 문제는 제기되지 않는다. 셋째로 지연 종속변수(lagged-dependent variable)를 비롯한 종단적 정보 활용과 더불어 방과후학교 참여와 관련 있는 풍부한 개인 및 학교 수준 배경변수를 고려하면, 관찰되지 않은 혼동요인으로 인한 대안적 설명 가능성을 최소화할 수 있다. 따라서 방과후학교의 연속적 처치효과에 대한 인과적 질문을 추구하는 본 연구의 내적 타당성의 문제도 효과적으로 해결할 수 있을 것으로 기대된다.

2. 결측 자료 처리

가. 2009·2010 각 년도 전수자료의 결측 처리

본 연구는 전수자료를 사용하지만 다양한 이유로 발생하는 결측 자료(missing data)로 인해 온전한 모집단 전체를 분석하기는 어렵다. 다만 결측 사례들이 무시하지 못할 만큼 큰 비율인지, 일정한 특성을 공유하는 집단이 체계적으로 배제가 되었는지를 살펴보고 적절한 처치를 한다면 편의되지 않은 모집단 효과 추정치를 얻을 수 있다.

분석 자료에서 결측을 고려해야할 변수는 크게 처치변수(treatment)인 교과별 방과후학교 참여 여부, 결과변수(outcome)인 교과별 학업성취도, 그리고 방과후학교 참여 여부를 예측하는 학생 및 학교 수준 공변량(covariate)으로 구분할 수 있다. <표 3-2>는 그 중에서 처치변수와 결과변수의 결측 사례와 비율을 교과별로 계산하여 제시한 것이다. 2009년의 경우 전체 약 64만 명의 사례들 중에서 처치변수인 방과후학교 참여 여부에 응답하지 않은 학생들은 전체 64만 명 학생의 1.72%인 약 1만 명 수준이다. 결과변수인 학업성취도 변수의 2009년 결측 비율은 약 2천 명~3천 명 수준으로 전체의 0.3%~0.4%에 지나지 않는다. 이에 따라 처치변수와 결과변수가 모두 관측된 완전 사례(complete-case)는 2009년의 경우 98% 수준에 이르고 있다. 2010년은 2009년보다 처치-결과 변수의 결측 사례 비율이 더욱 작아서 전체의 약 99.4%가 처치-결과 변수를

모두 포함한 완전 사례로 나타난다. 이러한 결측치의 비율을 고려하면 전체 분석에서 결측 사례로 인한 데이터의 왜곡이 가지는 영향력은 미미할 것으로 판단이 된다. 따라서 처치변수와 결과변수를 모두 포함한 완전 사례만을 분석에 포함하였다.

<표 3-2> 2009·2010 각 년도 전수자료의 처치·결과변수 결측치 비율

	전체	처치변수 결측: 방과후학교 참여	결과변수 결측: 학업성취도 점수	처치-결과변수 complete-case
2009 국어	643,106 (100.0%)	11,060 (1.72%)	2,939 (0.46%)	630,151 (97.99%)
2009 수학	643,106 (100.0%)	11,060 (1.72%)	2,566 (0.40%)	630,668 (98.07%)
2009 영어	643,106 (100.0%)	11,060 (1.72%)	1,937 (0.30%)	631,569 (98.21%)
2010 국어	631,362 (100.0%)	2,870 (0.45%)	1,302 (0.21%)	627,252 (99.35%)
2010 수학	631,362 (100.0%)	2,870 (0.45%)	1,243 (0.20%)	627,410 (99.37%)
2010 영어	631,362 (100.0%)	2,870 (0.45%)	585 (0.09%)	628,331 (99.52%)

처치-결과 변수의 결측과는 달리 공변량의 결측은 본 연구에서 다소 독특한 결측 처리가 가능하다. 그 이유는 2009·2010년 각각의 전수 자료가 방과후학교 참여 선택에 대한 예측모형에서 경향점수를 산출하는 데에 사용되기 때문이다. 통상적으로 단일 대체 방법(single imputation)은 대체값을 실제로 관측된 값처럼 이용하기 때문에 추정치의 표준오차를 과소 추정한다. 그러나 회귀계수에 대한 통계적 추론이 아니라 정확한 예측값을 얻는 것이 목적이라면 단일 대체 방법도 문제가 없다. Haviland et al.(2008)의 연구에서는 경향점수 추정을 위해 공변량의 결측값들을 단일 대체하고 대신 각 공변량의 결측에 대한 지표 더미 변수(missing data indicator)를 함께 투입하여 경향점수를 추정하는 방법의 효과성에 대해 논의하고 있다. 본 연구의 공변량 결측 처리에 있어서도 Haviland et al.(2008)의 연구에서 소개한 “단일대체+결측 더미 변수 추가” 방법을 적용하였다. 이때 단일 대체 방법으로는 EM 알고리즘

(Expection-Maximization algorithm) 방법을 사용하였다.

나. 2009-2010 학생 수준 연계 자료의 결측처리

2009·2010년 각 개년도의 횡단적 전수자료와는 달리 2009-2010 학생 수준 연계 자료는 무시할 수 없는 정도의 결측치 비율을 보이고 있다. 학생 수준 연계자료는 2,189개 학교 565,543명의 학생 자료만을 포함하는데, 이는 2009년 전체 학생의 약 12.1%인 77,563명, 2010년 전체 학생의 약 10.4%인 65,819명의 학생이 매칭 실패로 인해 누락이 된 것이다.

누락된 약 10%~12%의 사례가 과연 완전히 랜덤하게 연계에 실패한 것인지, 아니면 일정한 특성을 보유한 집단이 체계적으로 연계에 실패했는지 분석한 결과를 개략적으로 제시하면 <표 3-3>과 같다. 전반적으로 2010년의 연계 실패 사례들은 연계 성공 사례들과 학업성취도 평균과 표준편차가 거의 유사하다. 2010년의 경우 결측 사례가 특별히 체계적인 자료의 왜곡을 가져오는 것 같지는 않다. 따라서 2010년 연계 실패 사례의 경우 완전 임의 결측 가정(Missing Completely at Random, MCAR)을 충족한다고 볼 수 있다.

그러나 2009년 연계 실패 사례들은 성공 집단에 비해 성적 백분위가 평균적으로 4-5점 정도 낮으며 표준편차도 1-2점 정도 높다. 공부를 못하는 학생들이 체계적으로 탈락한 표본 선택(sample selection)이 일어났다고 볼 수 있다. 이 경우 완전 임의 결측 가정(MCAR)을 바탕으로 통계적 추론을 하면 편의된 추정치를 얻게 된다(Little & Rubin, 2002). 따라서 방과후학교 참여의 연속적 처치효과 분석 표본에서는 이 점을 반영하여 임의 결측 가정(Missing at Random, MAR)을 바탕으로 결측 자료를 처리할 필요가 있다. 임의 결측 가정(MAR)은 자료 안의 관찰된 공변량이 주어지면 결측치의 조건부 분포가 랜덤하다는 가정이다. 예컨대 같은 학업성취도를 보이는 학생들이 종단 자료에서 누락이 될 확률은 서로 같다고 보는 것이다.

임의 결측 가정(MAR)을 분석에 반영하기 위해 본 연구에서는 무응답

확률 가중치(non-response weights) 방법을 적용하였다(Hong & Raudenbush, 2008). 무응답 확률 가중치는 연계 성공=1의 값을 갖는 더미 변수를 결과변수로, 자료 내의 관찰된 공변량들을 예측변수로 하여 '연계 성공을 할 확률'을 추정하고 이에 대한 역수를 취하는 것으로 경향점수 역확률 가중치 방법과 원리가 똑같다. 이 가중치는 연계 실패할 확률과 비례하고 연계 성공할 확률과 반비례하기 때문에 표본으로부터 배제될 순간적 확률을 나타내는 Heckman 2단계 표본 선택 모형의 위험률(hazard rate) 혹은 IMR(Inverse Mill's ratio)과 개념적으로 유사하다. 본 연구에서는 무응답 확률 가중치를 방과후 선택과정 모형에서 산출된 경향점수 가중치와 곱하는 방식을 통해 방과후학교 연속적 처치효과 모형에서 표본 선택 편의를 보정하였다. 구체적인 적용 방법은 제 3절 분석 방법에서 설명한다.

<표 3-3> 2009-2010 연계 실패 결측 사례들의 규모 및 학업성취도 차이

	전체	연계 성공	연계 실패
2009년 사례수 (%)	643,106 (100.0%)	565,543 (87.9%)	77,563 (12.1%)
국어 척도점수 백분위 평균(SD)	51.55 (28.97)	52.07 (28.75)	47.69 (30.30)
수학 척도점수 백분위 평균(SD)	51.67 (28.44)	52.23 (28.33)	47.52 (28.90)
영어 척도점수 백분위 평균(SD)	51.45 (28.64)	51.90 (28.41)	48.08 (30.05)
2010년 사례수 (%)	631,362 (100.0%)	565,543 (89.6%)	65,819 (10.4%)
국어 척도점수 백분위 평균(SD)	50.61 (28.85)	50.57 (28.86)	50.96 (28.77)
수학 척도점수 백분위 평균(SD)	51.30 (28.30)	51.31 (28.30)	51.26 (28.29)
영어 척도점수 백분위 평균(SD)	50.55 (28.77)	50.50 (28.77)	51.05 (28.77)

제 2 절 변수 구성 및 기술통계

1. 처치변수: 교과별 방과후학교 참여 여부

국가수준 학업성취도 평가는 학생 성적 자료와 더불어 학생 설문 자료와 학교장 설문 자료를 포함한다. 학생 설문 자료는 크게 학생의 개인 및 가정사항·방과 후 활동·학습방법 및 태도·학교생활 영역을 조사한다. 방과후학교 참여 여부는 방과 후 활동 영역에서 "학교에서 운영하는 '방과후학교' 프로그램 중 현재 참여하고 있는 것에 모두 표시하십시오"라는 문항으로 측정되고 있다. 해당 문항의 선택지는 ① 참여하지 않음, ② 국어 교과 관련 프로그램, ③ 사회 교과 관련 프로그램, ④ 수학 교과 관련 프로그램, ⑤ 과학 교과 관련 프로그램, ⑥ 영어 교과 관련 프로그램, ⑦ 음악·미술·체육 관련 프로그램, ⑧ 기타 프로그램으로 구성되어 있다. 본 연구에서 사용하는 국어·수학·영어 방과후학교 참여 여부에 대한 0, 1로 코딩된 이분 변수는 기본적으로 위의 과목 선택지에 대한 응답 여부에 따라 제작된다.[4)]

여기서 분명히 할 것은 방과후학교 참여 여부 변수가 가지는 실제적 의미가 2009년과 2010년이 서로 다르다는 점이다. 그 이유는 방과후학교 참여에 대한 측정 문항이 '현재' 참여하고 있는 방과후학교 프로그램을

---

4) 본 연구에서 주로 사용되는 처치변수는 국어·수학·영어 교과별 방과후학교 참여 여부지만 '전체' 방과후학교 참여 여부 변수도 제작하여 기술통계로 제시하였다. 그러나 한국교육과정평가원에서 제공하는 데이터 파일을 통해 '전체' 방과후학교 참여 여부 변수를 제작하는 데에는 한 가지 난점이 있다. 그것은 방과후학교 참여는 확실히 1로 코딩이 되어 있지만 '미참여'와 '결측'이 모두 빈칸으로 처리되어 있어 양자를 구별할 수 없다는 것이다. 이 연구에서는 오류 유형을 방과후학교에 참여한다고 응답했으나 과목별 참여에 하나도 응답하지 않은 경우(오류 유형 1), 방과후학교에 참여하지 않는다고 응답했으나 과목별 참여에 하나 이상 응답한 경우(오류 유형 2)로 구분하고 방과후학교 만족도 응답 변수와의 논리적 관계에 따라 빈칸을 미참여와 결측으로 구별하였다. 향후 평가원의 데이터 클리닝 과정에서 이 부분이 정확히 개선될 필요가 있다.

묻고 있기 때문이다. 2009년에 국가수준 학업성취도 평가 시험과 설문조사를 시행한 시기는 <표 3-1>에 나타나 있듯이 2009년 10월 13-14일로 통상적으로 2학기가 시작된 후 1개월이 경과한 시점이다. 반면 2010년의 평가 시행 시기는 2010년 7월 13-14일로 1학기를 모두 마치고 여름방학을 앞두고 있는 시점이다. 방과후학교 프로그램의 개설과 운영이 대부분의 학교에서 학기마다 갱신이 되는 점을 고려하면, 2009년 방과후학교 프로그램 참여의 처치효과는 2학기 프로그램에 참여한 1~2개월 간의 효과로 정의가 된다. 그에 반해 2010년 방과후학교 프로그램 참여의 처치효과는 2010년 3월부터 7월까지 약 4~5개월 간 1학기 방과후학교 프로그램에 참여한 효과다. 2009년과 2010년 처치의 지속 기간(the duration of treatment)이 다르기 때문에 처치효과의 강도도 2009년보다는 2010년이 더욱 강할 것으로 예상할 수 있다.

<표 3-4>는 2009·2010년 각 년도 횡단 전수자료에서 방과후학교 참여 여부를 나타내는 0·1로 코딩된 처치변수들의 평균과 표준편차를 정리한 것이다. 이 값들은 연도별로 고등학생들의 1·2학년 시기 교과별 방과후학교 참여율을 나타낸다. 2009년 방과후학교 참여율은 전체 평균 75%이고 국어·수학·영어는 각각 54%, 57%, 56%다. 전체 방과후학교 참여율이 주요 교과별 방과후학교 참여율보다 높은 이유는 사회·과학 등 다른 과목의 방과후학교 프로그램 참여까지 고려한 수치이기 때문이다. 2010년 방과후학교 참여율은 전반적으로 참여율이 약간 하락한 정도지만 패턴이나 그 수준은 비슷하다.

<표 3-4> 2009·2010 횡단 전수 자료에서 연도별 방과후학교 참여

연도	결과변수 및 코딩값	N	전체 평균(SD)	국어 평균(SD)	수학 평균(SD)	영어 평균(SD)
2009년	1=방과후학교 참여	630,151	.75 (.44)	.54 (.50)	.57 (.50)	.56 (.50)
2010년	1=방과후학교 참여	627,252	.72 (.45)	.53 (.50)	.55 (.50)	.54 (.50)

<표 3-5>는 학생 개인 수준에서 연계된 2009-2010 종단 전수 자료에서 방과후학교의 연속적 참여 유형을 나누고 기술통계를 정리한 것이다. 방과후학교 참여를 1, 미참여를 0이라고 하면 고1-고2 시기 연속적 처치 시퀀스는 (0, 0), (1, 0), (0, 1), (1, 1) 네 가지 유형으로 분류된다. 전반적인 패턴은 전체・국어・수학・영어 모두 고1-고2 시기에 모두 방과후학교에 참여하는 (1, 1) 처치 시퀀스에 속하는 학생 비율이 가장 높고, 고1-고2 시기에 모두 방과후학교에 참여하지 않는 (0, 0) 처치 시퀀스에 속하는 학생 비율이 가장 낮다. 고1・고2 중 한 개 학년만 방과후학교에 참여하는 학생들 중에서는 고1에만 참여하는 학생 비율이 고2에만 참여하는 학생 비율보다 전반적으로 약간 높다.

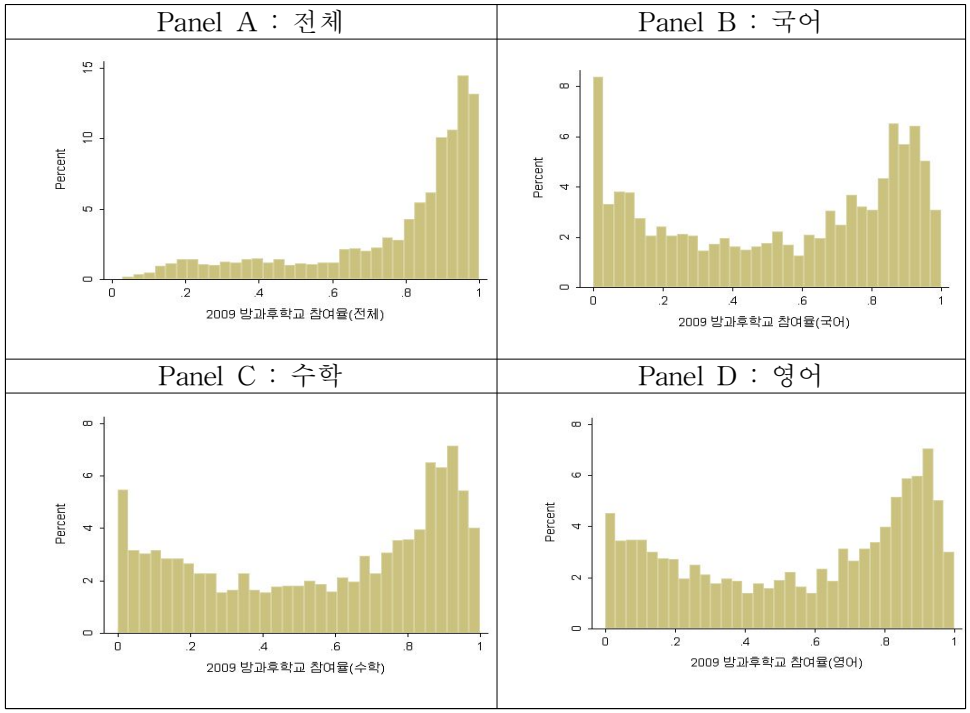
<표 3-5> 2009-2010 종단 전수 자료에서 처치 시퀀스별 방과후학교 참여

처치 시퀀스 (고1, 고2)	N	전체 평균(SD)	국어 평균(SD)	수학 평균(SD)	영어 평균(SD)
① (0, 0): 고1 미참여・고2 미참여	557,151	.07 (.25)	.21 (.41)	.19 (.39)	.20 (.40)
② (1, 0): 고1 참여・고2 미참여	557,151	.21 (.41)	.26 (.44)	.26 (.44)	.26 (.44)
③ (0, 1): 고1 미참여・고2 참여	557,151	.18 (.38)	.24 (.43)	.23 (.42)	.23 (.42)
④ (1, 1): 고1 참여・고2 참여	557,151	.54 (.50)	.29 (.45)	.32 (.47)	.31 (.46)

[그림 3-2]는 방과후학교 참여가 학교 수준에서는 어떠한 분포를 보이는지 기술적으로 나타낸 것이다[5]. 방과후학교 참여 여부 더미 변수(0・1)를 학교 수준으로 총계화(aggregate)하면 해당 학교의 방과후학교 참여율을 의미하게 된다. 이러한 참여율의 학교 수준 분포는 전체 참여율의 경우 부적 편포를 보이나 국어・수학・영어의 교과별 방과후학교 참여율은 0%와 100% 양 극단에 치우친 U자형 분포에 가까운 것으로 나타난다. 만약 학교 수준 군집성이 전무한 개인 수준의 랜덤한 선택 상황

5) 2010년 학교별 방과후학교 참여율의 분포는 2009년과 거의 동일한 분포를 보여 2009년 결과만을 제시하였다.

만을 가정한다면 방과후학교 참여율의 학교 수준 분포는 정규분포에 가까울 것이다. 그러나 방과후학교 참여율이 일률적으로 아주 높거나 아주 낮은 패턴이 국어·수학·영어 전과목에 걸쳐 나타나고 있는 것은 개인의 랜덤한 선택보다는 '방과후학교 강제 참여'와 같은 학교 수준 정책(school policy)의 영향력이 더욱 클 것이라는 추측을 가능하게 한다.



[그림 3-2] 2009년 학교별 방과후학교 참여율의 분포

2. 결과변수: 교과별 학업성취도

방과후학교 참여의 효과를 보기 위한 결과변수로는 국가수준 학업성취도 평가의 국어·수학·영어 학업성취도 점수를 사용한다. 기본적으로 국가수준 학업성취도 평가는 학생들의 동일 학교급 내 평가 점수의 연도간 비교가 가능하도록 척도점수를 산출하고 있다. 척도점수는 문항점수

들의 합산 점수인 원점수를 기초로 하여 만들어진 것으로 원점수를 아크사인(arcsine) 변환한 후 이 값을 다시 특정한 평균과 표준편차를 갖도록 선형 변환한 것이다.

<표 3-6>을 보면 국가수준 학업성취도 평가의 연도별 척도점수 체계가 크게 2003-2008 표집 평가 시기와 2009-2011년 전수 평가 시기로 구분된다는 점을 알 수 있다. 2003-2008년 표집 평가 시기의 척도 점수는 상급 학교의 점수 평균을 높게 설정하였다. 그러나 이 점수들은 같은 학교급의 연도별 성취도 점수만 비교 가능하도록 횡단적 동등화(horizontal equating) 한 점수이기 때문에 다른 학교급 간의 점수 차이가 학생 능력의 절대적인 성장을 의미하는 수직적 동등화(vertical equating) 점수로 해석할 수는 없다(송미영, 남민우, 강태훈, 김창일, 2011). 표집평가와 동일하게 횡단적 동등화 방식을 기반으로 하는 전수평가 체계 또한 학교급별 평균 차이를 두는 것이 아무런 실익이 없으므로 2009-2011년 전수평가 시기에는 초·중·고 모두 평균 200, 표준편차 30을 갖는 척도 점수 체계로 통일되었다.

<표 3-6> 국가수준 학업성취도 평가의 척도점수 체계

연도	학년	평균	표준편차	최솟값	최댓값
2003-2008년	초등학교 6학년	160	8.5	130	190
	중학교 3학년	260	8.5	230	290
	고등학교 1학년	360	8.5	330	390
2009년	초등학교 6학년	200	30	100	300
	중학교 3학년	200	30	100	300
	고등학교 1학년	360	8.5	330	390
2010-2011년	초등학교 6학년	200	30	100	300
	중학교 3학년	200	30	100	300
	고등학교 2학년	200	30	100	300

문제는 표집 평가에서 전면적인 전수 평가로 이행한 2009년에 오직 고등학교 1학년 학업성취도 척도점수만 표집평가 시기 점수 체계에 머물

러 있었다는 점이다. 그래서 2009년 고1 시기 척도점수와 2010년 고2 시기 척도 점수는 같은 학교급의 동일한 학생들을 평가한 결과임에도 불구하고 점수의 직접 비교가 불가능하다. 학년 간 학업성취도 점수에 대한 측정학적인 연계가 되어 있지 않기 때문에 '절대적'인 수준에서 학생 개인별 성장을 계량화하고 종단 모형을 적용하는 데에 근본적인 한계가 있다.

따라서 본 연구에서는 고1·고2 시기의 척도점수들을 각 년도에서의 '상대적' 순위를 나타내는 백분위 점수로 변환하여 동일하게 100점 만점의 척도를 갖도록 하였다. 2009년과 2010년 개별 자료는 전수 자료이기 때문에 개별 학생의 척도점수가 갖는 누적 백분위(cumulative percent)는 전체 모집단 내에서의 상대적 위치를 정확하게 나타낸다. 2009·2010 각 년도에서 척도점수에 대응하는 누적 백분위 점수를 계산한 결과는 <표 3-7>과 같다. 결측치를 제거하여 약간의 평균 차이가 존재하지만 대체로 100점 만점 점수에서 평균 50-51, 표준편차 28-29에 이르는 균등한 분포를 보이고 있다.

<표 3-7> 교과별 학업성취도 척도점수 백분위 기술통계

연도	변수명	N	평균	표준편차	최솟값	최댓값
2009년	국어 척도점수 백분위	632,046	51.55	28.97	0	100
	수학 척도점수 백분위	632,046	51.67	28.44	0	100
	영어 척도점수 백분위	632,046	51.45	28.64	0	100
2010년	국어 척도점수 백분위	627,208	50.61	28.85	0	100
	수학 척도점수 백분위	627,208	51.30	28.30	0	100
	영어 척도점수 백분위	627,208	50.55	28.77	0	100

이러한 백분위 점수를 활용한 종단 모형은 DePascale(2006)의 학업성장 개념 분류에 따르면 '상대적 순위 변화에 기초한 학업성장(growth relative to others)' 모형에 속한다. 만약 어떤 학생이 고1 시기에는 전체 모집단에서 백분위 50에 해당하는 성취를 보이고 고2 시기에는 똑같은

모집단에서 백분위 60에 해당하는 성취를 보였다면 이 학생의 상대적 순위 변화에 근거하여 학업성취도가 증가하였다고 판정할 수 있다. 이러한 접근은 어떤 학생이 고1-고2 간 학업성취도가 절대적인 의미에서 향상되었다고 하더라도 상대적 순위를 나타내는 백분위 점수가 고1·고2 양 년도에서 동일하다면 학업성장이 이루어지지 않았다고 간주하는 특징이 있다. 본 연구에서 추정하는 방과후학교 참여의 종단적 처치효과는 이러한 학업성장 점수의 상대적 의미에 유의하여 해석해야 한다.

3. 학생 및 학교 수준 공변량

가. 학생 수준 공변량

학생 수준 공변량은 2009·2010년 국가수준 학업성취도 평가 학생 설문 자료를 바탕으로 구성되며, 크게 1) 개인 배경 변수, 2) 방과 후 여가 시간(학습활동), 3) 방과 후 여가 시간(비학습활동), 4) 학교 교사에 대한 신뢰 및 학교 적응, 5) 학생의 학습태도 변수군으로 분류된다.

개인 배경 변수는 여학생 여부·편모 가정 여부·편부 가정 여부·조손 가정 여부로 구성되어 있다. 세 종류의 결손가정 더미 변수는 학생 설문 중 '누구와 함께 살고 있습니까' 문항에서 부·모·조부·조모의 보호자 보기에 대한 응답을 바탕으로 생성하였다. 조손 가정은 조부모만이 유일한 보호자인 경우, 편모 가정은 '어머니'에만 응답한 경우, 편부 가정은 '아버지'에만 응답한 경우다. <표 3-8>의 기술통계를 보면 여학생은 전체의 47%를 차지한다. 편모 가정 학생은 전체의 약 8%, 편부 가정 학생은 약 4%, 그리고 조손 가정 학생은 약 1% 정도로 나타나고 있다.

이준호와 박현정(2012)의 연구는 방과후학교 선택 과정을 학생의 방과 후 시간 활용이라는 시간 자원 배분(time resource allocation)의 맥락에서 다루고 있다. 방과 후 학생들의 시간은 한정되어 있고 학생들이 어떤 활동에 시간을 더 투자할 것인지에 관한 선택 과정에 놓여있다는 점에서 접근하는 것이다. 학생들은 방과후 사교육·학교 숙제·TV 시청·PC 게

임·인터넷·친구들과 보내기 등 다양한 활동을 하고 있다. 이러한 활동이 방과후학교 참여와 어떤 관련을 보이고 있는지 탐색하기 위해 각 활동에 대해 0=하지 않음, 1=1시간 미만, 2=1시간 이상 2시간 미만, 3=2시간 이상 3시간 미만, 4=3시간 이상으로 코딩값이 실제 시간과 유사한 의미를 갖도록 구성하여 모형에 투입하였다. 기술통계를 보면 방과 후 평균적인 여가시간 활용은 친구들과 보내는 시간, 텔레비전 시청, 인터넷, 컴퓨터 게임, 학원 강의 및 과외, 학교 숙제, EBS 수강 순으로 많다. 평균적으로 고등학생들은 방과 후 학습활동보다 비학습활동에 시간을 더욱 많이 할애하고 있는 것으로 나타난다.

고등학생의 경우 방과후학교 강사의 약 98%가 현직 교사기 때문에 학교 교사에 대한 신뢰는 방과후학교 참여에 주요한 예측 요인이다. 교육과학기술부에서 발표한 2009년 방과후학교 현황 분석(09. 10월 기준) 자료에 따르면 교과 방과후학교 프로그램 강사는 일반계고의 경우 97.7%, 그 외의 고등학교의 경우 95.2%가 현직 교사였다. 외부 강사는 각각 2.3%, 4.8%에 지나지 않았다. 따라서 학생 설문 중 학교 교사의 배려 정도('우리 학교 선생님들은 내 의견을 존중한다')와 학교 교사의 열정 정도('우리 학교 선생님들은 열심히 가르친다')를 묻는 문항을 변수로 투입하였다.

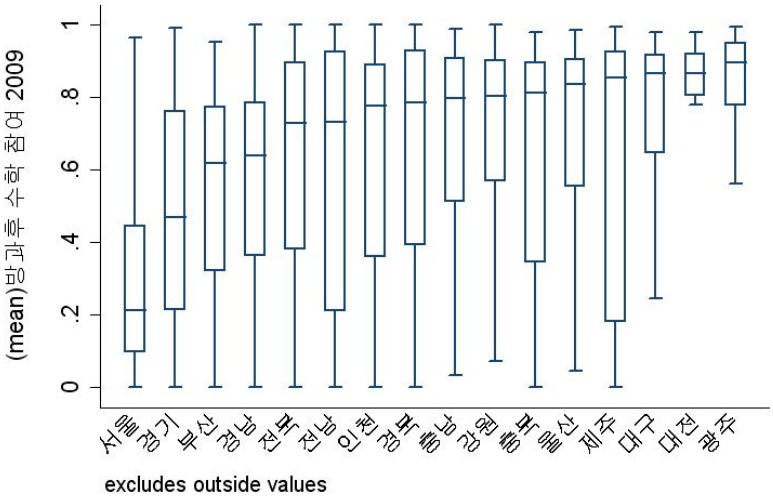
또한 학교에 대한 적응 정도가 방과후에도 계속 학교에 남아있으려는 성향에 영향을 미친다고 보아 학교 생활 즐거움('나는 학교에 가는 것이 즐겁다')을 묻는 문항을 변수로 투입하였다. 동일한 맥락에서 예습·복습·수업 등에서 학교 공부에 성실성을 보이는 학생은 학교에서 제공하는 방과후학교 수업에도 성실하게 참여할 가능성이 크기 때문에 4점 리커트 척도로 구성된 문항을 변수로 투입하였다.

<표 3-8> 개인 수준 예측 변수의 기술통계

구분	변수명	고등학교 1학년 (2009)			고등학교 2학년 (2010)		
		사례수	평균(SD)	범위	사례수	평균(SD)	범위
개인 및 가정배경	여학생	632,046	0.47 (0.50)	0-1	627,208	0.47 (0.50)	0-1
	편모 가정	632,046	0.08 (0.27)	0-1	627,208	0.09 (0.28)	0-1
	편부 가정	632,046	0.04 (0.20)	0-1	627,208	0.04 (0.21)	0-1
	조손 가정	632,046	0.01 (0.10)	0-1	627,208	0.01 (0.10)	0-1
방과후 여가시간 (학습활동)	수업 후 학교숙제 시간	632,046	1.24 (0.80)	0-4	627,208	1.17 (0.80)	0-4
	학원 및 과외 시간	632,046	1.28 (1.44)	0-4	627,208	1.16 (1.38)	0-4
	EBS 수강 시간	632,046	0.73 (1.00)	0-4	627,208	0.82 (1.01)	0-4
방과후 여가시간 (비학습활동)	텔레비전 시청 시간	632,046	1.79 (1.13)	0-4	627,208	1.70 (1.09)	0-4
	컴퓨터 게임 시간	632,046	1.33 (1.25)	0-4	627,208	1.32 (1.22)	0-4
	인터넷 시간	632,046	1.38 (0.96)	0-4	627,208	1.42 (0.94)	0-4
	친구들과 보내는 시간	632,046	2.16 (1.31)	0-4	627,208	2.17 (1.27)	0-4
교사 신뢰 및 학교 적응	학교 교사 배려	632,046	2.77 (0.63)	1-4	627,208	2.79 (0.64)	1-4
	학교 교사 열정	632,046	2.97 (0.56)	1-4	627,208	2.95 (0.55)	1-4
	학교 즐거움	632,046	2.70 (0.76)	1-4	627,208	2.70 (0.75)	1-4
학생의 학습 태도	예습	632,046	1.90 (0.64)	1-4	627,208	1.95 (0.65)	1-4
	복습	632,046	2.29 (0.75)	1-4	627,208	2.31 (0.74)	1-4
	수업 시간 집중	632,046	3.12 (0.65)	1-4	627,208	3.15 (0.63)	1-4

가. 학교 수준 공변량

학교 수준 변수는 크게 지역 더미 변수군, 기본 학교 배경, 학생·학부모·교사 구성, 교장 특성, 방과후학교 관련 특성 변수로 구성되어 있다. 기술통계는 <표 3-9>에 제시하였다. 지역 더미 변수군은 서울 지역을 준거집단으로 부산, 인천, 대구, 대전, 울산, 광주, 경기, 충북, 충남, 전북, 전남, 경북, 경남, 강원, 제주 15개 더미변수를 투입하였다. 15개 더미변수를 모두 투입한 이유는 [그림 3-3]에서 보듯이 학교별 방과후학교 참여율이 지역별로 상당한 이질성을 보였기 때문에 지역 요소를 충분히 통제하는 것이 필요하다고 보았기 때문이다. 따라서 다른 학교 수준 예측변수들의 회귀계수는 모두 같은 광역시도 내에서의 비교를 통해 산출되는 값들이 된다.



[그림 3-3] 2009년 학교별 수학 방과후학교 참여율의 지역별 분포

기본 학교 배경 변수는 학교 성별·목적 유형·설립 유형·학교 규모로 구성하였다. 학교 성별은 남녀공학을 준거집단으로 남학교 더미와 여학교 더미를, 목적 유형은 일반계고를 준거집단으로 전문계고와 종합고

더미를 추가하였다. 설립 유형은 국공립 학교를 준거집단으로 사립 더미 변수를 투입하였다. 학교 규모는 총 학생수의 자연로그값을 사용하였다.

학생·학부모·교사 구성 변수는 학생의 경우 기초생활수급자 비율·중식지원 학생 비율·학생들의 태도에 대한 교장 인지(학교 시설을 아끼려는 마음, 학교생활을 잘 하려는 의지, 학업성취를 높이고자 하는 의지, 선생님에 대한 존경심, 학교에 대한 자부심) 문항 평균을 사용하였다. 학부모 변수는 학부모 태도에 대한 교장의 인지(학업성취도 향상을 위한 학부모의 요구, 자녀의 학업성취에 대한 학부모의 지원, 학교 교육활동에 대한 학부모의 참여) 문항 평균을 사용하였다. 교사 변수는 석사학위 이상 교사 비율과 초임 교사 비율을 투입하였다.

교장 특성 관련 변수는 교장 배경변수로서 교장의 성별(여자 교장=1)과 연령, 교장이 공모제를 통해 부임하였는지 여부를 변수로 구성하였다. 이와 더불어 행정업무·외부 기관 회의에 교장이 전체 업무 중 얼마 정도의 비율을 할애하는지에 대한 응답치를 변수로 투입하였다. 교장 관련 변수를 비중있게 구성한 이유는 학교장의 결정권한이 큰 학교 단위 방과후학교 운영방침이 방과후학교 참여율에 상당한 영향을 미칠 것이라는 본 연구의 가설 때문이다.

마지막으로 학교 수준에 투입되는 시간-의존적 혼동요인(time-varying confounder)으로서 방과후학교가 '교과내용 이해하는 데에 도움이 된다' 정도에 대한 학교 평균, '방과후학교에 계속 참여 하고 싶다' 학교 평균, 그리고 각 년도 1학기 교과 방과후학교 개설 강좌수를 구성하였다. 이 변수들은 고1 시기 학교 수준 방과후학교 참여율에 영향을 받으면서도 고2 시기 학교 수준 방과후학교 참여율에 영향을 미치는 요인들이다. 특기할만한 것은 방과후학교에 대한 두 개의 평가 문항이 4점 만점의 리커트 척도로 구성되어 있음에도 불구하고 고1 시기 방과후학교 재참여 의사의 학교 평균은 평균적으로 2.06, 고2 시기 방과후학교가 교과 내용에 도움이 되었는지에 대한 학교 평균은 평균적으로 1.02에 지나지 않는다는 점이다. 이는 일선 학교 현장에서 방과후학교에 대한 만족도가 그리 높지 않다는 점을 보여주고 있다.

<표 3-9> 학교 수준 예측 변수의 기술통계

변수명		고등학교 1학년 (2009)			고등학교 2학년 (2010)		
		사례수	평균(SD)	범위	사례수	평균(SD)	범위
지역	서울=1	2,204	0.14 (0.35)	0-1	2,185	0.14 (0.35)	0-1
기본 학교 배경	남학교=1	2,204	0.18 (0.39)	0-1	2,185	0.18 (0.39)	0-1
	여학교=1	2,204	0.20 (0.40)	0-1	2,185	0.20 (0.40)	0-1
	전문계고=1	2,204	0.24 (0.43)	0-1	2,185	0.25 (0.43)	0-1
	종합고=1	2,204	0.09 (0.28)	0-1	2,185	0.09 (0.29)	0-1
	사립=1	2,204	0.42 (0.49)	0-1	2,185	0.42 (0.49)	0-1
	로그 총학생수	2,204	6.49 (0.90)	2.3-7.7	2,185	6.51 (0.88)	2.3-7.7
학생 학부모 교사 구성	기초생활수급자 비율	2,204	8.77 (8.20)	0-100	2,185	9.04 (8.62)	0-91.5
	중식지원학생 비율	2,204	18.75 (18.70)	0-100	2,185	23.34 (22.00)	0-148
	학생들의 태도 평균	2,204	2.91 (0.55)	1-4	2,185	2.99 (0.59)	1-4
	학부모의 태도 평균	2,204	2.68 (0.64)	1-4	2,185	2.77 (0.65)	1-4
	석사 이상 교사 비율	2,204	40.33 (15.50)	0-100	2,185	40.34 (15.64)	0-100
	초임교사 비율	2,204	5.32 (8.10)	0-71.4	2,185	4.61 (7.36)	0-95
교장 특성	여자 교장=1	2,204	0.06 (0.24)	0-1	2,185	0.07 (0.25)	0-1
	교장 연령	2,204	58.18 (4.30)	38-87	2,185	58.36 (4.01)	40-75
	교장공모제=1	2,204	0.10 (0.31)	0-1	2,185	0.14 (0.35)	0-1
	교장 행정업무 비율	2,204	31.00 (13.7)	2-80	2,185	32.81 (13.95)	5-90
	교장 외부회의 비율	2,204	10.94 (5.73)	0-50	2,185	10.91 (5.72)	0-50
방과후학교 관련 특성	교과내용 이해 도움	2,204	2.46 (0.27)	1.6-3.4	2,185	1.02 (0.08)	1.0-2.8
	방과후 재참여 의사	2,204	2.06 (0.34)	1.0-3.5		N/A	
	교과방과후 강좌수	2,204	36.30 (43.30)	0-527	2,185	51.84 (64.77)	0-539

제 3 절 분석 방법

이 절에서 설명하는 분석방법은 연구문제에 따라 1단계 방과후학교 참여 여부 선택 모형과 2단계 방과후학교 연속적 참여의 처치효과 분석 모형으로 구분된다.

1. 1단계 : 방과후학교 참여 선택 모형

가. 학교 수준 군집화 정도 추정: ICC와 MOR

결과변수인 수학·영어·국어 과목별 방과후학교 참여 변수는 방과후학교에 참여할 경우 1, 참여하지 않을 경우 0의 값을 갖는 이분(binary) 반응 변수이므로 기본적으로 다음과 같이 학생 i, 학교 j에 대한 로지스틱 연결함수를 적용한다.

$$\eta_{ij} = \log\left(\frac{p_{ij}}{1-p_{ij}}\right),$$

p_{ij}는 j번째 학교 i번째 학생이 해당 과목의 방과후학교에 참여할 확률을 나타낸다. η_{ij}는 해당 확률을 로그 승산(log-odds)으로 변환한 값이다. 이 η_{ij}값으로 다층 로지스틱 모형의 unconditional model을 적용하여 학교 수준 랜덤 효과를 구한다.

Level-1 학생 수준 $\eta_{ij} = \beta_{0j}$

Level-2 학교 수준 $\beta_{0j} = \gamma_{00} + u_{0j}, \quad u_{0j} \sim N(0, \tau_{00})$

j번째 학교의 랜덤 효과(random-effect)인 u_{0j}는 방과후학교 참여 로그 승산에 대한 학교 수준 영향력의 총합으로 볼 수 있다(Rabe-Hesketh

& Skrondal, 2012). 이 랜덤 효과의 분산인 τ_{00}로부터 다음과 같은 집단 내 상관계수(intra-class correlation, ICC)를 구할 수 있다.

$$ICC = \frac{\tau_{00}}{\tau_{00} + \pi^2/3}$$

로지스틱 함수의 분산은 위의 식에서 나타난 바와 같이 $\pi^2/3$로 약 3.29 정도의 상수(constant)로 고정되어 있다. 1수준 분산으로 고정된 $\pi^2/3$와 데이터에서 추정된 2수준 랜덤 효과의 분산을 더하여 총 분산을 정의하고 이 중에서 2수준 랜덤 효과의 분산이 어느 정도 비율을 차지하고 있는지 계산함으로써 ICC를 산출한다.

ICC와 더불어 다층 로지스틱 함수의 군집성을 계량화하는 또 하나의 척도로 Larsen과 Merlo(2005)가 제안한 MOR(Median Odds Ratio)을 사용할 수 있다. MOR의 원리를 간단히 요약하면 다음과 같다. 먼저 공변량이 모두 동일하지만 랜덤 효과가 다른 두 학교 j와 j'의 랜덤 효과 추정치 u_{0j}와 $u_{0j'}$을 반복적으로 표집하여 두 추정치의 차이를 승산비(odds ratio)로 나타낸다. 각각의 랜덤효과 추정치 u_{0j}와 $u_{0j'}$는 정규분포를 다르므로 반복적 표집을 통해서 구해진 다수의 랜덤 효과의 승산비 또한 누적 정규분포로부터 얻을 수 있다. 이 분포의 중간값을 취해 전형적인(typical) 학교 간 랜덤 효과 차이를 요약한 것이 바로 MOR 지수다. 만약 MOR이 1이라면 우리가 가진 표본에서 학교들 간의 랜덤 효과 간 차이가 없다고 볼 수 있다. 그러나 MOR이 2라면 학교 간 랜덤 효과의 평균적인 차이가 2배 정도라는 것을 의미한다.

이러한 원리를 반영하여 MOR 추정식을 유도하면 다음과 같다. 먼저 중간값(50 백분위)과 1보다 큰 다른 백분위 상수 a는 다음과 같은 누적분포 함수(cumulative distribution function)로부터 얻어질 수 있다.

$$P\{\exp(|u_{0j} - u_{0j'}|) \le a\} = P\left\{\frac{|u_{0j} - u_{0j'}|}{\sqrt{2\tau_{00}}}\right\} = 2\Phi\left\{\frac{\ln(a)}{\sqrt{2\tau_{00}}}\right\} - 1$$

만약 누적확률(cumulative probability)이 1/2, 즉 중간값인 50백분위로 설정되어 있다면 백분위 상수 a가 바로 MOR이 된다.

$$2\Phi\left\{\frac{\ln(MOR)}{\sqrt{2\tau_{00}}}\right\}-1=\frac{1}{2}$$

이 방정식을 MOR에 대해 풀면,

$$MOR=\exp\sqrt{2\tau_{00}}\,\Phi^{-1}(3/4)$$

을 얻는다. 이 식에서 나타난 바와 같이 MOR은 랜덤효과 u_{0j}의 분산 추정치(τ_{00})와 누적정규분포 함수, 그리고 상수 3/4으로 추정된다.

나. 고1·고2 시기 방과후학교 참여 예측 모형

고1 시기 방과후학교 참여를 예측하는 다층 로지스틱 모형으로는 다음과 같은 2수준 확률 절편(random-intercepts) 모형을 적용한다. 표기의 편의를 위해 여러 개의 공변량들을 하나의 변수로 나타내었다. 이 확률 절편 모형에서 투입되는 변수들은 이해를 돕기 위해 [그림 2-4]에서 사용한 문자로 표시되어 있다.

Level-1 학생 $\eta_{ij}=\beta_{0j}+\beta_{1j}X_1$

Level-2 학교 $\beta_{0j}=\gamma_{00}+\gamma_{01}\overline{X_1}+\gamma_{02}X_{1j}+u_{0j},\quad u_{0j}\sim N(0,\tau_{00})$

X_1과 X_{1j}은 고1 시기 baseline 공변량들을 나타낸다. 학생 수준에 투입되는 X_1은 개인 배경·방과후 여가 시간 활용·학교 교사에 대한 인식·학생의 수업 태도 변수들이고, 학교 수준에 포함되는 X_{1j}는 지역 더미·기본 학교 배경·학생, 학부모, 교사 구성·교장 특성 등이다. 앞서

이론적 배경 장에서 살펴보았듯이 고1 시기 개인 수준 예측변수 X_1는 모두 집단 평균 중심점 교정(group-mean centering)을 하고 모든 X_1의 집단 평균 $\overline{X_1}$를 2수준에 투입해야 한다. 이렇게 하는 이유는 세 가지다.

첫째는 1수준 예측변수 X_1의 학교 내 효과(within-school effect)를 추정하기 위해서다. 집단 평균 중심점 교정을 하면 1수준 X_1의 회귀계수 β_{1j}는, 예컨대 학교 j내에서 그 학교 교사에 대한 신뢰도가 1단위 차이 나는 학생 i와 i'간의 방과후학교 참여 로그 승산(log-odds)의 차이를 나타내게 된다. 집단 평균 $\overline{X_1}$을 추가하게 되면 학교 내 효과 추정치가 최소한 1수준 예측변수 X_1와 학교 수준 랜덤 효과 u_{0j}와의 상관으로 인해 편향될 가능성은 없다(Gelman & Hill, 2007; Rabe-Hesketh & Skrondal, 2012).

둘째는 2수준에 투입된 1수준 예측변수들의 집단 평균 $\overline{X_1}$의 회귀계수 γ_{01}가 학교 간 효과(between-school effect)를 나타내도록 하기 위해서다. 모든 1수준 변수를 집단 평균 중심점 교정하였기 때문에 β_{0j}는 방과후학교 참여 로그 승산(log-odds)의 j번째 학교 평균을 나타낸다. 따라서 학생들의 수업 태도 학교 평균 $\overline{X_1}$의 회귀계수는 수업 태도 학교 평균이 1단위 차이 나는 학교 j와 j'간 방과후학교 참여 로그 승산 학교 평균의 차이를 의미하게 된다.

셋째는 2수준 예측변수 X_{1j}의 회귀계수가 갖는 편의를 줄이기 위해서다. 만약 1수준에 집단 평균 중심점 교정된 학생의 수업태도 변수만 투입하고 이들의 수업태도의 학교 평균을 2수준에 투입하지 않는다면, 수업태도 학교 평균으로 표상되는 해당 학교의 학습 분위기가 학교 수준 랜덤 효과 u_{0j}에 비관측 이질성(unobserved heterogeneity)으로 남아있게 된다. 이 랜덤 효과 u_{0j}와 전문계고 더미변수와 같은 학교 수준 예측변수는 상관을 가질 가능성이 크다. 따라서 전문계고 더미변수의 회귀계수 추정치는 수업태도 학교 평균과의 상관에 비례하는 정도의 편의를 갖게 된다.

이러한 원리에 따라 고2 시기 방과후학교 참여를 예측하는 다층 로지

스틱 모형 또한 학생 수준의 모든 변수를 집단 평균 중심점 교정을 하고 그 변수들의 집단 평균 $\overline{X_1}$를 2수준에 투입하는 다음과 같은 2수준 확률 절편(random-intercepts) 모형을 적용한다.

Level-1 학생 $\eta_{ij} = \beta_{0j} + \beta_{1j}X_1 + \beta_{2j}T_1 + \beta_{3j}Y_1 + \beta_{4j}X_{2C}$

Level-2 학교 $\beta_{0j} = \gamma_{00} + \gamma_{01}\overline{X_1} + \gamma_{02}X_{1j} + \gamma_{03}X_{1C} + u_{0j}, \quad u_{0j} \sim N(0, \tau_{00})$

위 모형은 고1 시기 방과후학교 참여 여부 예측모형과 다른 것들은 모두 동일하다. 다만 고2 시기 종단적인 다층경향점수 추정식에는 고1 시기 방과후학교 참여 여부 T_1, 고1 시기 학업성취 점수 Y_1, 방과후학교 만족도 학교 평균 등 고1 시기 학교 수준 시간-의존적 혼동변수 X_{1C}, 학습 태도·학교생활 등 고2 시기 개인 수준 시간-의존적 혼동변수 X_{2C}가 추가적으로 투입된 것이 차이점이다.

방과후학교 만족도 변수를 개인 수준이 아니라 학교 수준의 시간 의존적 변수로 투입한 이유는 방과후학교에 참여하지 않은 학생의 경우 방과후학교의 교과 내용 이해 도움 정도나 방과후학교 재참여 의사에 대한 응답이 결측 처리되기 때문이다. 이 값들을 제외하거나 특정값으로 대체하는 것은 다른 변수들의 회귀계수의 의미를 심각하게 왜곡한다. 이러한 상황에서 시간-의존적 혼동변수의 성격을 보존하면서도 분석에 포함할 수 있는 방안은 학교 수준으로 총계화(aggregate)하는 것이다. 고1 시기 방과후학교 참여율은 방과후학교 재참여 의사의 학교 평균에 영향을 미치고, 이는 다시 고2 시기 방과후학교 참여율에 영향을 미칠 것이다. 이렇게 학교 수준 시간-의존적 혼동변수를 고려할 수 있는 것은 다층경향점수 추정만을 통해 가능하다.

다. 다층경향점수 및 안정화 역확률 가중치의 추정

위와 같이 설정된 다층 로지스틱 회귀분석의 확률 절편 모형을 통해 먼저 다층경향점수(multilevel propensity score)를 추정한다. 다층경향점

수 추정의 핵심은 방과후학교 참여 확률에 대한 특정 학교 j의 비관측 특성들을 요약하는 값인 랜덤 효과(random-effect) u_{0j}에 특정한 값을 부여하는 것이다. 본 연구에서는 u_{0j}를 경험적 베이즈 예측(empirical Bayes prediction) 방식으로 추정하고 다른 공변량들을 통해 산출된 고정효과(fixed effects) 예측값과 결합하여 j학교의 i번째 학생이 방과후학교에 참여할 로그-승산(log-odds)을 산출하였다. 이 로그 승산값[4)]이 바로 다층경향점수 추정치다.

본 연구의 다층경향점수 추정에서 핵심적인 부분을 차지하는 것은 학교별 랜덤효과 u_{0j}에 값을 부여하는 경험적 베이즈 예측 방법이다. 이 방식의 원리를 간단히 설명하면 다음과 같다. 기본적으로 랜덤효과 u_{0j}를 미지의 상수로 놓고, 특정한 학교 j에서 학교 수준 공변량 X_j와 u_{0j}가 함께 주어졌을 때 해당 학교 학생들 자료가 관찰될 확률을 최대로 하는 다음과 같은 우도 함수를 세워볼 수 있다.

$$Likelihood\,(y_{1j}, y_{2j}, \cdots, y_{nj,j} | X_j, u_{0j})$$

이 우도를 최대로 하는 통상적인 최대우도추정(maximum likelihood estimation)과 경험적 베이즈 방식의 차이점은, 경험적 베이즈 방식의 경우 랜덤효과의 정규확률밀도함수 $u_{0j} \sim N(0, \tau_{00})$를 사전확률분포(prior distribution) 정보로 활용하는 데에 있다. 그러면 다음과 같은 랜덤효과에 대한 사후확률분포(posterior distribution)를 얻을 수 있다.

$$Posterior(u_{0j} | y_{1j,} \cdots, y_{n_j j,} X_j)$$

$$\propto Prior(u_{0j}) \times Likelihood\,(y_{1j}, y_{2j}, \cdots, y_{n_j j} | X_j, u_{0j})$$

---

4) Rosenbaum and Rubin(1985)은 산출된 로그 승산을 확률값으로 변환한 것보다 로그 승산 값 그 자체가 경향점수 분석에 더욱 바람직한 분포 특성을 보인다고 제안한 바 있다.

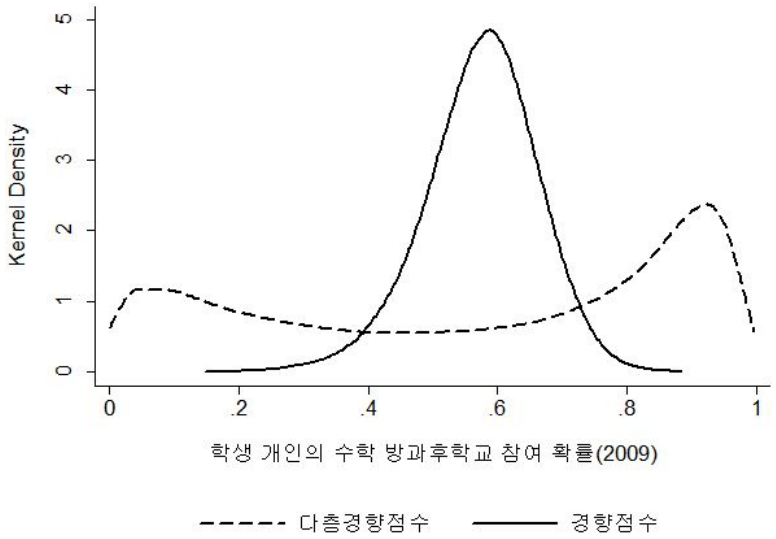
우측항의 사전분포와 우도함수의 결합은 사후분포와 동일한 것이 아니라 단지 비례할 뿐이다. 따라서 수치적분(numerical integration)을 통해 얻어지는 정규화 상수(normalizing constant)로 우측항을 나누어주어야 사후확률분포를 얻을 수 있다. 수치적분식은 다음과 같다.

$$\widetilde{u_{0j}} = \int u_{0j}\, Posterior(u_{0j}|y_{1j,} \cdots, y_{n_j j,} X_j) du_{0j}$$

위의 식으로 계산한 u_{0j}의 실제적인 의미는 학생 개인이 방과후학교 참여할 확률에 대한 학교 수준 특성의 부가분(addition)이다. 예를 들어 개인 수준 특성이 정확히 동일한 학생 A, B 둘이 동일한 개인 수준 방과후학교 참여 확률 0.5를 갖는다고 가정해보자. 이것이 통상적인 경향점수 추정치다. 그러나 학생 A, B가 다른 학교에 다닌다면 학교 수준 특성의 방과후학교 참여에 대한 효과 부가분인 u_{0j}의 값이 다르게 더해진다. 학생 A가 다니는 학교는 방과후학교 참여에 대한 자율적 선택 방침으로 방과후학교 참여 확률이 전체적으로 낮아서 음(-)의 로그-승산값을 갖는 u_{0j}가 더해지고, 학생 B가 다니는 학교는 강제 참여 방침으로 인해 학교 수준 방과후학교 참여 확률이 높아 양(+)의 로그-승산값을 갖는 u_{0j}가 더해진다. 그렇다면 학생 A의 방과후학교 참여 확률은 0.2, 학생 B의 방과후학교 참여 확률은 0.8 정도로 현저히 달라질 수 있다.

학생 개인 수준만을 고려한 경향점수 추정치와 학교 수준 랜덤효과 u_{0j}를 함께 고려한 다층경향점수 추정치의 커널밀도함수를 비교한 [그림 3-4]가 이러한 측면을 잘 보여준다. 그림을 보면 다층경향점수 추정치의 분포와 경향점수 추정치의 분포는 확연하게 다르다는 점을 알 수 있다. 학생 개인 수준의 예측변수들만 고려한 경향점수 추정치는 평균 0.57, 표준편차 0.09인 정규분포에 가깝게 나타난다. 이 경향점수에 따르면 우리나라 학생들은 방과후학교에 참여할 확률이 0.57 정도되는 학생들이 가장 많고 방과후학교에 참여할 확률이 0이거나 1에 가까운 학생들은 매우 적은 종모양의 대칭 분포를 이룬다. 그러나 학교 수준 랜덤효과 u_{0j}를 고

려한 다층경향점수 추정치의 분포는 방과후학교 참여 확률이 0과 1 양극단에 가까운 학생들이 많은 U자형 양봉분포를 보인다. 이것은 [그림 3-2]의 학교 수준 방과후학교 참여율의 분포와 거의 일치하는 모습이다. 다층경향점수 추정치의 분포에 따르면 우리나라 고등학생들은 방과후학교에 참여할 확률이 아주 높거나 아주 낮은 학생들로 구성되어 있다. 이 역시 방과후학교 참여는 학생 개인의 랜덤한 선택보다는 '방과후학교 강제 참여'와 같은 학교 수준 정책(school policy)의 영향력이 더욱 크다는 점을 보여주고 있다.



[그림 3-4] 다층경향점수와 경향점수의 커널밀도함수 비교

이렇게 산출된 고1·고2 시기 다층경향점수를 활용하여 다음과 같은 안정화 역확률 가중치(stabilized inverse probability of treatement weight, SIPTW)를 구할 수 있다(Van der Wal & Geskus, 2011).

$$sw_{tij} = \frac{1}{P(T_1 = 1|X_1)} \times I_{고2} \cdot \frac{P(T_2 = 1|T_1, X_1)}{P(T_2 = 1|T_1, Y_1, X_1, X_{1c}, X_{2c})}$$

위의 가중치 산출식에서 t는 시점, i는 개인, j는 학교를 나타낸다. 기본적으로 종단 연구에서 역확률 가중치는 해당 시점까지의 가중치 곱으로 표현된다. 예컨대 j학교 i번째 학생이 고2 시기 방과후학교 참여에 대한 역확률 가중치는 그 학생의 고1 시기 방과후학교 참여에 대한 역확률 가중치와 고2 시기 가중치를 함께 곱해준 값이다. 위의 식에서 고2 시기 가중치 앞에 고2 시기면 1, 아니면 0의 값을 갖는 지표함수 $I_{고2}$가 붙어 있지만 고1 시기 가중치 앞에는 지표함수 없이 고1 시기든 고2 시기든 변함없이 포함되어 있다. 한편 위의 식이 '안정화' 가중치인 이유는 시간-의존적 혼동 변수(X_{1c},X_{2c})와 고1 시기 학업성취도(Y_1)를 제외한 뒤 방과후학교에 참여할 로그-승산값을 추정하여 고2 시기 가중치의 분자 부분에 추가했기 때문이다. 이 부분은 분자와 분모 사이의 극단적인 차이를 방지하고 분포를 좁게 만들어주는 이점이 있다(Cole & Hernan, 2008).

똑같은 원리로 2009-2010년 연계 실패 자료의 표본 선택성을 보정하기 위한 무응답 가중치(non-response weight)를 다음과 같이 구성할 수 있다(Hong & Raudenbush, 2008).

$$nrw_{tij} = \frac{1}{P(a_{1ij}=1|X_1)} \times I_{고2} \cdot \frac{P(a_{2ij}=1|T_1, X_1)}{P(a_{2ij}=1|T_1, Y_1, X_1, X_{1c}, X_{2c})}$$

$a_{tij}=1$은 j번째 학교의 i번째 학생이 $t(t=1,2)$ 시점에 관측치 Y_{tij}를 가지는 것을 의미한다. 반대로 $a_{tij}=0$은 해당 시점에 Y_{tij}가 결측인 것을 의미한다. 무응답 가중치(nrw_{tij})는 개념적으로 t시점에 관측될 확률에 반비례하는 역확률 가중치다. 관측될 확률이 높은 사례에 낮은 가중치를 주고 낮은 사례에 높은 가중치를 주어서 표본 선택 편의를 보정하겠다는 것이다. 이러한 무응답 가중치를 SIPTW와 같이 종단적 안정화 추정치로 확장하면, MCAR 가정을 충족하는 것으로 보이는 2010년 사례에 혹시 남아있을지도 모르는 결측치 분포 상의 이질성을 더욱 섬세하게 통제할 수 있다.

최종적으로 적용하는 가중치 ϕ_{tij}는 SIPTW인 sw_{tij}와 무응답 가중치인 nrw_{tij}를 곱한 것으로 $\phi_{tij} = sw_{tij} \times nrw_{tij}$로 표현된다. ϕ_{tij}가 각 사례에 적용되면 선택과정에 관여하는 공변량들이 모두 적절히 투입되었다는 가정 하에 분석표본을 선택편의 없는 연속적 랜덤할당 표본에 근사시킨다. 이때 sw_{tij}가 큰 값을 가져도 nrw_{tij}이 작은 값이면 ϕ_{tij}가 작은 값을 가질 수 있다. 반대로 sw_{tij}가 작은 값을 가져도 nrw_{tij} 큰 값이면 해당 가중치 ϕ_{tij}가 적용된 사례는 표본에서 큰 비중을 차지하도록 재구조화될 수 있다.

2. 2단계 : 방과후학교 참여의 연속적 처치효과 추정 모형

고1 시기 사전-처치 혼동요인 X_1과 고1·고2 시기 시간-의존적 혼동요인 X_{1c}, X_{2c}가 충분히 관찰되었다고 가정하면 위에서 추정한 역확률 가중치 ϕ_{tij}를 적용하여 재구조화한 표본은 연속적 비혼동성(sequential uncofoundedness) 가정을 충족하는 연속적 랜덤 할당 표본에 근사한다(Robins, Greenland, & Hu, 1999). 이렇게 역확률 가중치로 재구조화된 종단 자료를 바탕으로 처치변수와 결과변수 간 관계를 모형화한 것을 Marginal Structural Model(Robins, Hernan, & Brumback, 2000)이라고 하며 생명통계학(biostatistics)의 전통에서 주로 사용된다.

역확률 가중치를 적용하여 혼동요인들을 적절히 통제했다는 가정 하에 본 연구가 해결해야 할 또 하나의 문제는 처치변수인 방과후학교 참여 여부(T_1, T_2)와 결과변수인 고1·고2 시기 학업성취도(Y_1, Y_2) 간의 관계를 어떻게 모형화할 것인가 하는 점이다. 주어진 데이터를 다층 성장 모형의 관점에서 보면 1수준 2시점 반복 측정치, 2수준 학생, 3수준 학교인 3수준 다층모형으로 볼 수 있다. 그런데 1수준에서 추정해야할 모수는 절편·기울기·잔차항 3개인데 우리가 가진 정보는 2개 시점 자료뿐이다. 따라서 절편·기울기·잔차항 3개의 모수치 중 하나를 고정시켜서 모형의 자유도가 최소한 0이 되도록 해야 한다(최길찬, 2010).

만약 문항반응이론을 통한 시점 간 척도 동등화가 이루어지고 각 개인의 성취도에 대한 표준오차가 추정되었다면 두 시점 자료와 측정의 표준오차를 이용한 3수준-2시점 잠재변수 성장모형(Bryk, Thum, Easton, & Luppescu, 1998; Choi, Seltzer, & Herman, 2004)을 추정하는 것이 가능하다. 이 모형은 1수준 잔차항의 분산을 특정한 상수 $SE(Y_{tij})^2$로 고정시키는 방식이다. 그러나 현재 국가수준 학업성취도 평가의 측정 체계는 문항반응이론에 근거하여 개별 학생들의 학업성취도 측정치의 불확실성을 계량화지 않기 때문에 $SE(Y_{tij})^2$ 값을 구할 수 없다.

따라서 본 연구에서는 1수준 잔차항을 특정 상수 $SE(Y_{tij})^2$가 아닌 0으로 고정시켰다. 이는 기존의 3수준 다층모형을 사실상 학생 수준-학교 수준으로 구성된 2수준 다층 모형으로 적용하겠다는 것이다. [표 3-10]에 본 연구에서 적용한 2수준 다층모형이 제시되어 있다.

기본적으로 [표 3-10]의 모형은 두 시점 반복측정치를 학생 수준에서 병합(pooling)한 것이다. 전형적인 데이터 구조는 좌측 3개 열의 $time$, T_1, T_2의 코딩을 보면 이해할 수 있다. 시간($time$) 변수는 고1=0, 고2=1인 더미변수로 학생 수준에 투입되며 방과후학교 참여에 관한 고1・고2 시기 처치변수 두 개(T_1, T_2)는 1수준에 시간불변 공변량(time-constant covariate)으로 추가되어 있다. 우리가 구하고자 하는 처치효과들은 시간($time$)변수와 고1・고2 시기 방과후학교 참여 여부 변수들(T_1, T_2) 간의 삼원 상호작용(three-way interaction)을 통해 정의된다.

고1 시기 방과후학교 참여의 처치 지연 효과(lagged treatment effect)는 '④ 고1만 방과후학교에 참여한 학생의 고2 시기 성적 기댓값'에서 '② 고1-고2 연속 미참여 학생의 고2 시기 성적 기댓값'을 뺀 것으로 정의된다. $time=1$인 고2 성적 데이터에서 처치 시퀀스 (1, 0)인 학생과 (0, 0)인 학생을 비교하는 것이다. 처치 지연 효과의 회귀계수는 $\beta_1+\beta_5$로 Stata의 lincom 명령어를 통해 두 회귀계수의 선형결합에 대한 유의성 검정을 할 수 있다. 두 시점이 병합된 데이터이므로 이때 로버스트 표준오차(robust standard error)를 추정한다.

고1・고2 시기의 연속적 참여가 갖는 배가 효과(amplifying effect)는

우선 처치 시퀀스 (1, 1)인 2년 연속 참여한 학생과 처치 시퀀스 (0, 0)인 2년 연속 참여하지 않은 학생의 고2 시기 성적 격차를 정의한다. '⑧ 고1-고2 연속 참여 학생의 고2 시기 성적 기댓값'으로부터 '② 고1-고2 연속 미참여 학생의 고2 시기 성적 기댓값'을 빼서 구한다. 이 격차에서 고2 성적에 대한 고1 시기 방과후학교의 효과를 의미하는 '고1 시기 처치 지연 효과'와 고2 성적에 대한 고2 시기 방과후학교의 효과를 의미하는 '고2 시기 횡단적 처치효과'를 추가적으로 빼면 '연속적 참여가 갖는 배가효과'를 추정할 수 있다. 이 배가효과를 회귀계수로 나타내면 $\beta_4 + \beta_7$로 정의되는데 모형에서 보면 고1·고2 시기 방과후학교 참여의 상호작용항인 $T_1 \cdot T_2$가 갖는 시간에 따른 효과로 볼 수 있다.

<표 3-10> 시간 및 방과후학교 참여 변수 코딩에 따른 처치효과 정의

$$Y_{ij} = \beta_0 + \beta_1 T_1 + \beta_2 T_2 + \beta_3 time + \beta_4 T_1 \cdot T_2$$
$$+ \beta_5 T_1 \cdot time + \beta_6 T_2 \cdot time + \beta_7 T_1 \cdot T_2 \cdot time + u_{0j} + r_{ij}$$

연번	$time$	T_1	T_2	회귀계수	의미
①	0	0	0	β_0	고1-고2 연속 미참여 학생의 고1 시기 성적 기댓값
②	1	0	0	$\beta_0 + \beta_3$	고1-고2 연속 미참여 학생의 고2 시기 성적 기댓값
③	0	1	0	$\beta_0 + \beta_1$	고1만 참여 학생의 고1 시기 성적 기댓값
④	1	1	0	$\beta_0 + \beta_1 + \beta_3 + \beta_5$	고1만 참여 학생의 고2 시기 성적 기댓값
⑤	0	0	1	$\beta_0 + \beta_2$	고2만 참여 학생의 고1 시기 성적 기댓값
⑥	1	0	1	$\beta_0 + \beta_2 + \beta_3 + \beta_6$	고2만 참여 학생의 고2 시기 성적 기댓값
⑦	0	1	1	$\beta_0 + \beta_1 + \beta_2 + \beta_4$	고1-고2 연속 참여 학생의 고1 시기 성적 기댓값
⑧	1	1	1	$\beta_0 + \beta_1 + \beta_2 + \beta_3 + \beta_4 + \beta_5 + \beta_6 + \beta_7$	고1-고2 연속 참여 학생의 고2 시기 성적 기댓값

처치효과	정의	회귀계수
1. **고1 횡단적 처치효과**: 고1 시기 방과후학교 참여의 고1 시기 성적에 대한 효과	③-①	β_1
2. **고2 횡단적 처치효과**: 고2 시기 방과후학교 참여의 고2 시기 성적에 대한 효과	⑥-②	$\beta_2 + \beta_6$
3. **고1 처치 지연 효과**: 고1 시기 방과후학교 참여의 고2 시기 성적에 대한 효과	④-②	$\beta_1 + \beta_5$
4. 고1-고2 연속참여의 고2 시기 학업성취도에 대한 효과	⑧-②	$\beta_1 + \beta_2 + \beta_4 + \beta_5 + \beta_6 + \beta_7$
5. **고1-고2 연속참여의 배가 효과**	4-(2+3)	$\beta_4 + \beta_7$

제 4 장 연구결과

제 1 절 방과후학교 참여 선택과정

1. 방과후학교 참여 여부의 학교 수준 군집화 정도

<표 4-1>은 전체 및 국어·수학·영어 과목별로 연도별 ICC와 MOR 값을 구하여 정리한 것이다. 이로부터 학생 개인의 방과후학교 참여 여부가 얼마나 학교 수준 요인에 큰 영향을 받는지 알 수 있다. 2009년의 경우 과목별 ICC는 국어 59.2%, 수학 56.5%, 영어 54.2%이고 2010년에는 국어 58.0%, 수학 56.7%, 영어 55.8%로 모두 50%를 초과하고 있다. 통상적으로 학업성취 점수의 ICC가 20% 초반대인 것을 감안하면 그 2~3배의 군집성이 방과후학교 참여 여부에서 나타나고 있는 것이다.

<표 4-1> 방과후학교 참여 여부의 과목별 학교 수준 군집화 정도

연도	군집화 척도	학교 N	전체	국어	수학	영어
2009년	ICC	2,204	44.5%	59.2%	56.5%	54.2%
	MOR	2,204	4.70	8.02	7.17	6.56
2010년	ICC	2,185	41.2%	58.0%	56.7%	55.8%
	MOR	2,185	4.26	7.63	7.25	6.97

과목별 MOR 또한 2009년과 2010년 각각 국어 8.02, 7.63, 수학 7.17, 7.25, 영어 6.56, 6.97로 7.0-8.0 전후의 매우 높은 수치를 보이고 있다. 다른 개인 수준 변수가 정확히 동일해도 학교가 달라지면 학생 i와 i'간의 방과후학교 참여 승산(odds)이 7-8배 차이가 난다는 것이 위에서 구한 MOR값들이 전달하는 메시지다. 달리 말하면, 어떤 학생이 학교 j에서 학교 j'으로 옮겨갔을 때 그 학생 개인은 변한 것이 없지만 학교가 달라짐으로써 방과후학교에 참여할 확률 자체가 평균적으로 7배 정도 차

이가 날 것이라고 예상할 수 있다. ICC가 국어·수학·영어 각 과목에서 50%를 넘어 개인 수준 요인에 따른 변산을 압도한다는 점, 7.0-8.0 정도 되는 MOR 추정값은 <표 4-4>에 나타난 모든 학생 수준 예측변수들의 OR(odds ratio)값보다 크다는 점에서 방과후학교 참여 여부는 개인 수준 요인보다는 학교 수준 요인에 의해 결정되는 부분이 더욱 크다고 볼 수 있다.

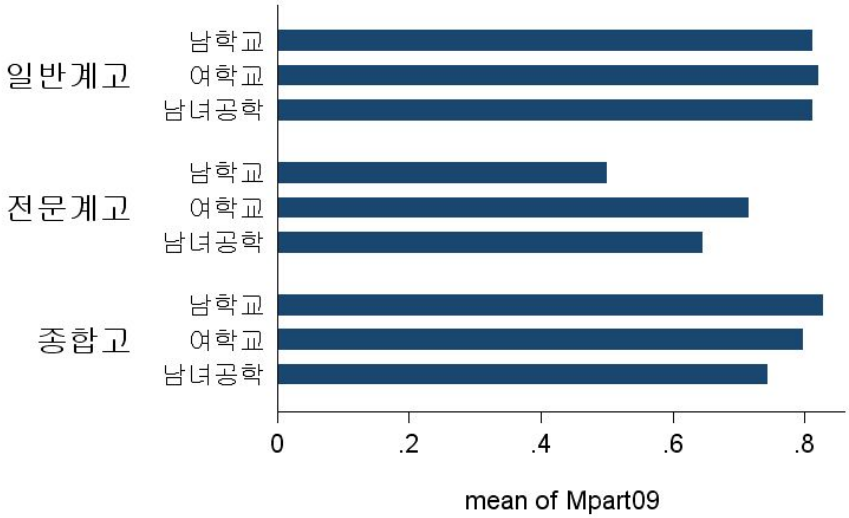
2. 고1 시기 방과후학교 참여 선택 예측 요인

학교 수준 요인 중 어떤 것이 방과후학교 참여에 대한 큰 연관성을 보이는지 알아보기 위해 <표 4-2>에 학교 수준 랜덤 효과(random-effect), <표 4-3>에 학교 수준 고정 효과(fixed-effect)를 산출해 정리하였다.

<표 4-2>는 확률 절편 모형에서 정의되는 각 과목별 학교 수준 랜덤 효과 u_{0j}의 분산 τ_{00}를 단계적으로 제시한 것이다. 학교 수준 변수군을 차례로 투입하면서 해당 변수군의 τ_{00}에 대한 '설명된 분산 비율(proportion of variance explained)'을 계산하였다. 먼저 기저 모형은 학생 수준 예측변수만을 투입하고 학교 수준 변수는 아무 것도 투입하지 않은 모형이다. 이때 학교 수준 랜덤 효과의 잔차 분산 τ_{00}는 국어 4.06, 수학 4.18, 영어 3.82으로 로지스틱 회귀모형의 1수준 분산 상수 3.29보다 모두 크게 나타나 학교 수준 요인의 영향력이 개인 요인의 영향력보다 크다는 점을 다시 한 번 확인할 수 있다. 여기에 서울을 준거집단으로 한 15개의 지역 더미 변수군을 투입하면 지역 요인으로 설명된 만큼 잔차 분산이 줄어든다. 그 비율을 계산하면 국어 18.0%, 수학 15.8%, 영어 18.3%다. 지역 요인이 방과후학교 참여의 학교 간 차이를 약 16~18% 정도 설명한다는 의미다.

학교 성별·설립 유형·목적 유형·학교 규모를 포함하는 기본 학교 배경 변수들은 방과후학교 참여 확률의 학교 간 변산을 가장 많이 설명한다. 기본 학교 배경 변수들은 학교 간 분산을 약 35%~39% 정도 설

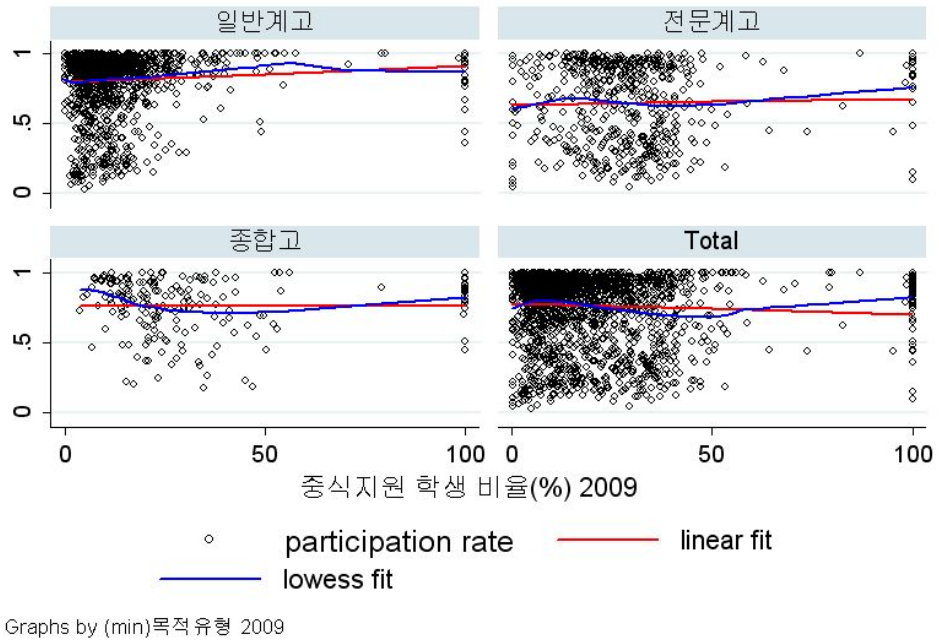
명하고 있다. 이로부터 우리는 지역이나 학교 배경 등 기본적인 정보만으로도 방과후학교 참여의 학교 간 차이를 52%~55% 정도 예측할 수 있다는 점을 알 수 있다. 특히 <표 4-3>의 학교 수준 변수들의 고정효과를 보면 가장 영향력이 큰 기본 학교 배경 변수는 전문계고·종합고 더미 변수다. 전문계고와 종합고인 경우 방과후학교 참여 로그 승산의 학교 평균이 각각 일반계고의 14%, 36% 수준이다. [그림 4-1]이 이를 기술적으로 보여주고 있다.



[그림 4-1] 2009년 학교 목적유형·성별에 따른 방과후학교 참여율

그런데 기초생활수급자 비율, 중식 지원 학생 비율, 학부모 적극성 평균, 석사학위 이상 교사 비율 등으로 구성된 학생·학부모·교사 구성 관련 변수군은 세 과목 모두 설명량이 0%로 방과후학교 참여 로그 승산의 학교 간 분산을 전혀 설명하지 못하는 것으로 나타난다. [그림 4-2]에서 중식지원 학생 비율과 방과후학교 참여율 간의 이변량 관계가 편평하게 나타나는 것을 보면 알 수 있다. 회귀모형에서 학교의 인적 구성 요소의 영향력을 대부분의 1수준 변수 학교 평균 변수들이 설명하는 것이 중요한 원인이다. 1수준 변수의 학교 평균에는 '학교 교사가 열정적으

로 가르친다고 생각하는 정도'의 학교 평균, '학생이 수업 후 복습을 하는 정도'의 학교 평균 등이 포함되어 있다. 이는 해당 학교의 교사와 학생의 구성적 특징을 반영하는 것으로 설명량은 10~14% 내외다.



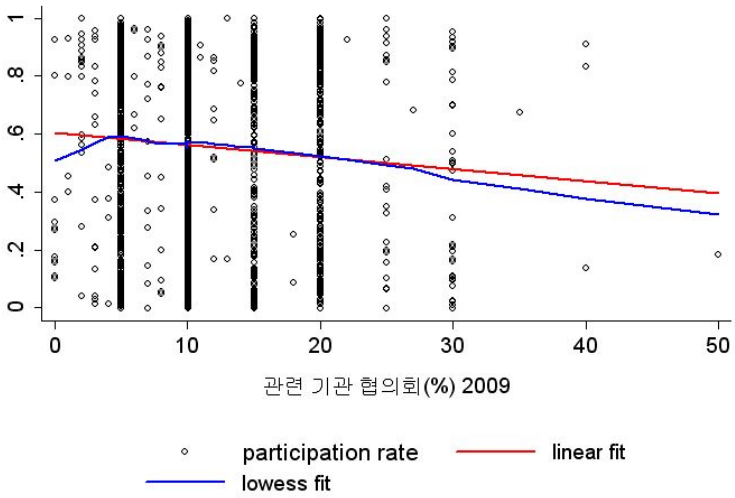
[그림 4-2] 2009년 중식지원학생비율에 따른 방과후학교 참여율

선행연구들과 다른 본 연구의 특징적인 발견은 바로 교장 특성 변수군의 영향력이다. 교장 특성 변수군이 지역이나 기본 학교 배경, 학생·학부모·교사의 인적 구성 등의 요인과 관계없이 독립적으로 설명하는 방과후학교 참여의 학교 간 분산은 국어 3.2%, 수학 2.8%, 영어 2.5%로 산출된다. 방과후학교 운영에 관한 학교 수준 운영 방침에 학교장의 성향이 유의하게 반영되고 있다는 점을 보여주는 결과다.

구체적인 계수의 방향성을 살펴보면 교장이 여자이거나 공모제로 부임한 교장일 경우 국어·수학·영어 모든 과목에서 학교 평균 방과후학교 참여의 로그 승산이 감소한다. 이러한 분석 결과의 '실제적' 의미에 대해

서는 좀 더 면밀한 고찰이 필요할 것이나 이를 학교장에 의한 '방과후학교 강제적 참여' 방침을 시사하는 결과로 해석할 수 있을 것이다. 여자 교장이나 공모제로 부임한 교장이 방과후학교 참여를 학생들 개인의 선택에 자유롭게 맡기는 경향이 더 많고, 반대로 남자 교장이나 공모제로 부임하지 않은 교장이 방과후학교 전원 강제 참여 방침 등을 정하는 경향이 더욱 클 가능성이 있다.

또한 교장의 행정업무 비율이나 외부 기관 회의 비율이 높아지면 해당 학교의 방과후학교 참여 확률이 유의하게 감소하는 것으로 나타난다. 특히 교장이 행정업무에 치중하는 학교일수록 방과후학교 참여 경향성이 낮은데 이는 [그림 4-3]에 경향적으로 나타나 있다. 방과후학교 운영 방침에 적극적으로 관여하는 것이 학생들의 학력향상을 위한 교장의 '교수적 리더십(instructional leadership)'을 의미한다면, 행정업무나 외부기관 회의에 치중하는 교장은 학교 내 교수-학습 활동에 대한 관여가 적은 것으로 해석할 수도 있다.



[그림 4-3] 2009년 교장 외부기관회의 비율에 따른 방과후학교 참여율

<표 4-2> 학교수준 랜덤 효과 u_{0j}의 잔차 분산과 변수군의 설명량

	변수군	τ_{00}	df	χ^2	Δdf	$\Delta\chi^2$	변수군 설명량
	기저 모형	4.06	2,168	174449.5***			
	지역 더미변수군	3.77	2,153	164510.4***	15	9939.1***	18.0%
국어	기본 학교배경	2.47	2,147	132463.6***	6	32046.8***	34.5%
	학생·학부모·교사구성	2.47	2,141	132458.7***	6	4.9	0.0%
	교장 특성	2.39	2,136	130859.2***	5	1599.5***	3.2%
	1수준 변수 학교 평균	2.15	2,119	124213.0***	17	6646.2***	10.0%
	기저 모형	4.18	2,168	172499.8***			
	지역 더미변수군	3.52	2,153	154886.7***	15	17613.1***	15.8%
수학	기본 학교배경	2.15	2,147	126340.7***	6	28546.0***	38.9%
	학생·학부모·교사 구성	2.15	2,141	126045.3***	6	295.4***	0.0%
	교장 특성	2.09	2,136	124523.2***	5	1522.1***	2.8%
	1수준 변수 학교 평균	1.80	2,119	115378.2***	17	9145.0***	13.9%
	기저 모형	3.82	2,168	171885.6***			
	지역 더미변수군	3.12	2,153	156421.5***	15	15464.1***	18.3%
영어	기본 학교배경	2.02	2,147	124023.8***	6	32397.7***	35.3%
	학생·학부모·교사구성	2.02	2,141	123652.5***	6	371.3***	0.0%
	교장 특성	1.97	2,136	122212.5***	5	1440.0***	2.5%
	1수준 변수 학교 평균	1.73	2,119	113727.6***	17	8484.9***	12.2%

***$p \leq .001$, **$.001 < p \leq .01$, *$.01 < p \leq .05$, †$.05 < p \leq .10$

<표 4-3> 방과후학교 참여에 대한 학교 수준 변수 고정효과

	변수명	국어 OR(t)		수학 OR(t)		영어 OR(t)	
기본 학교 배경	남학교=1	1.41 *	(2.24)	0.85	(-1.12)	1.23	(1.47)
	여학교=1	0.78	(-1.38)	1.43 *	(2.20)	1.01	(0.04)
	사립=1	1.07	(0.86)	1.02	(0.23)	1.17 *	(2.15)
	전문계고=1	0.19 ***	(-10.78)	0.14 ***	(-13.71)	0.24 ***	(-10.43)
	종합고=1	0.46 ***	(-5.45)	0.36 ***	(-7.82)	0.44 ***	(-6.47)
	로그 총학생수	1.13 †	(1.65)	1.12	(1.62)	1.04	(0.65)
학생 학부모 교사 구성	기초생활수급자 비율	1.00	(-0.20)	1.00	(0.28)	1.00	(-0.22)
	중식지원학생 비율	1.00	(0.34)	1.00	(-0.33)	1.00	(0.53)
	학생들의 태도 평균	0.93	(-0.86)	0.93	(-0.95)	0.88	(-1.58)
	학부모의 태도 평균	0.97	(-0.40)	0.98	(-0.28)	1.04	(0.54)
	석사 이상 교사 비율	1.00	(-1.03)	1.00	(0.22)	1.00	(-0.45)
	초임교사 비율	1.01	(1.14)	1.00	(-0.05)	1.00	(0.46)
교장 특성	여자 교장=1	0.83 **	(-2.87)	0.70 **	(-2.75)	0.81 *	(-2.56)
	교장 연령	1.01 †	(1.66)	1.01	(1.36)	1.01	(1.40)
	교장공모제=1	0.81 †	(-1.96)	0.76 **	(-2.81)	0.84 *	(-2.31)
	교장 행정업무 비율	0.99 **	(-2.89)	0.99 **	(-2.77)	0.99 **	(-2.85)
	교장 외부회의 비율	0.99 *	(-2.27)	0.99 †	(-1.91)	0.99 †	(-1.85)

***$p \leq .001$, **$.001 < p \leq .01$, *$.01 < p \leq .05$, †$.05 < p \leq .10$

<표 4-4>는 학생 수준 변수의 학교 내 효과(within-school effect)와 학생 수준 변수들을 학교 수준으로 평균 낸 변수들의 학교 간 효과(between-school effect)를 정리한 것이다. 예컨대 학교 숙제 시간의 학교 내 효과는 동일한 학교 내에서 방과 후 학교 숙제 시간이 1단위 차이 나는 학생 간 방과후학교 참여 로그승산의 차이를 의미한다. 한편 학교 숙제 시간의 학교 간 효과는 학교 숙제 시간 학교 평균이 1단위 차이 나는 학교 간 방과후학교 참여 로그 승산의 평균 차이를 의미한다. 계수는 승산비(odds ratio)로 표현되어 있으며 괄호 안의 숫자는 t값이다.

학교 내 효과 계수들이 보여주는 일련의 경향성을 요약하면 이준호와 박현정(2012)의 연구결과와 같이 '학교 교사를 믿고 성실하게 학교 공부를 수행하는 학생'일수록 학교에서 제공하는 교과 방과후학교 프로그램에도 참여할 확률이 높아지는 것으로 나타난다[5]. 방과 후 여가 시간에 TV 시청이나 컴퓨터 게임 같은 비학습활동보다 학교 숙제, EBS 수강같은 학습활동에 치중하는 학생, 학교 교사가 배려해주고 열정적으로 가르친다고 인식하며 학교생활을 즐겁게 하는 학생이 그렇지 않은 학생보다 방과후학교 참여 확률이 더 높다. 그러나 학원 및 과외 시간은 유의한 관련성을 보이지 않고 있다. 전수 자료라서 사례수가 매우 큼에도 불구하고 영 효과(null effect)가 나타난 것은 방과후학교와 학원은 한 재화의 수요가 증가함에 따라 다른 재화의 수요가 감소하는 대체재 관계가 아닐 가능성이 크다는 점을 보여준다.

학교 평균을 예측변수로 하는 학교 간 효과도 일반적으로는 '학생들이 열심히 공부하고 교사도 열정적으로 가르치는 학교'의 방과후학교 참여율이 높다는 것을 보여준다. 특히 복습과 수업시간 집중의 학교 평균이 1단위 증가하면 방과후학교 참여의 학교 평균 로그 승산은 수학 과목의 경우 각각 15배, 5배 정도로 증가한다. 또한 학교 교사가 열정적으로 가

5) 학교 내 효과를 보면 대부분의 계수들이 통계적으로 유의하며 상당히 큰 t값을 보이고 있다. 이는 실제 효과 크기(effect size)가 커서 그렇다기보다는 전수 자료로서 사례 수가 매우 커서 표준오차가 작게 추정되기 때문인 것으로 보인다. 따라서 학교 내 효과는 계수의 방향성과 크기에 초점을 맞춰 해석할 필요가 있다.

르친다고 인식하는 정도의 학교 평균이 1단위 높아지면 방과후학교 참여의 학교 평균 로그 승산이 수학과 영어 과목에서 8배 정도 증가한다. 그러나 '학원 및 과외 시간'의 학교 평균과 '학교 교사 배려에 대한 인식'의 학교 평균이 높아지면 정반대로 학교 수준 방과후학교 참여 확률이 낮아지고 있다. 만약 방과 후 시간을 학교에서 강제 보충수업이나 자율학습에 참여시키지 않고 학생들의 선택에 맡기는 등의 방침을 정하고 있다면, 해당 학교 학생들은 학교 교사가 배려해준다는 인식이 평균적으로 높아짐과 동시에 학원 및 과외 시간이 늘어날 수 있다.

<표 4-4> 방과후학교 참여에 대한 개인 변수의 학교 내-학교 간 고정효과

구분	변수명	Within-school Effect OR (t-value)			Between-school Effect OR (t-value)		
		국어	수학	영어	국어	수학	영어
개인 및 가정배경	여학생	1.37*** (30.12)	1.24*** (20.90)	1.30*** (25.41)	1.47 (1.11)	0.49* (-2.23)	1.18 (0.54)
	편모 가정	0.92*** (-6.00)	0.92*** (-6.30)	0.92*** (-6.51)	0.72 (-0.29)	1.32 (0.26)	0.60 (-0.49)
	편부 가정	0.85*** (-9.30)	0.86*** (-8.67)	0.85*** (-9.60)	1.36 (0.23)	1.81 (0.47)	0.30 (-1.01)
	조손 가정	0.80*** (-6.23)	0.77*** (-7.18)	0.78*** (-7.01)	3.25 (0.83)	2.70 (0.75)	5.12 (1.25)
방과후 여가시간 (학습활동)	수업 후 학교숙제 시간	1.17*** (30.44)	1.19*** (34.79)	1.18*** (33.32)	1.90*** (3.77)	2.77*** (6.51)	2.30*** (5.45)
	학원 및 과외 시간	1.02*** (8.18)	1.00 (-0.35)	1.00 (1.10)	0.68** (-3.11)	0.55*** (-5.27)	0.61*** (-4.39)
	EBS 수강 시간	1.02*** (4.05)	1.02*** (6.34)	1.02*** (5.14)	3.29*** (6.87)	2.34*** (5.35)	2.74*** (6.49)
방과후 여가시간 (비학습활동)	TV 시청 시간	0.96*** (-12.12)	0.95*** (-13.66)	0.96*** (-12.53)	1.63* (2.28)	1.61* (2.42)	1.24 (1.15)
	컴퓨터 게임 시간	0.95*** (-14.34)	0.97*** (-10.39)	0.95*** (-13.72)	0.89 (-0.52)	0.93 (-0.37)	1.09 (0.44)
	인터넷 시간	0.99** (-3.59)	0.98*** (-5.27)	0.99** (-2.86)	0.62 (-1.60)	0.65 (-1.57)	0.72 (-1.27)
	친구들과 보내는 시간	0.99* (-2.48)	0.99* (-2.33)	1.00 (-0.09)	0.70* (-2.36)	0.76† (-1.92)	0.73* (-2.28)
교사 신뢰 및 학교 적응	학교 교사 배려	1.04*** (5.44)	1.05*** (7.52)	1.05*** (7.78)	0.10*** (-4.94)	0.15*** (-4.44)	0.17*** (-4.26)
	학교 교사 열정	1.05*** (7.36)	1.07*** (9.84)	1.06*** (8.88)	11.91*** (6.00)	7.74*** (5.39)	8.02*** (5.62)
	학교 즐거움	1.04*** (6.80)	1.05*** (10.42)	1.04*** (7.81)	0.22*** (-4.96)	0.27*** (-4.70)	0.27*** (-4.72)
학생의 학습 태도	예습	0.95*** (-7.49)	0.96*** (-6.87)	0.94*** (-8.68)	0.28* (-2.41)	0.21** (-3.23)	0.25** (-2.92)
	복습	1.12*** (17.58)	1.15*** (22.79)	1.12*** (18.89)	11.05*** (5.50)	14.76*** (6.69)	12.78*** (6.47)
	수업 시간 집중	1.20*** (28.75)	1.21*** (30.31)	1.19*** (29.00)	3.99** (3.32)	4.96*** (4.17)	4.67*** (4.09)

$^{***}p \leq .001$, $^{**}.001 < p \leq .01$, $^{*}.01 < p \leq .05$, $^{\dagger}.05 < p \leq .10$

3. 고2 시기 참여에 대한 시간-의존적 혼동변수의 영향

고2 시기 방과후학교 참여 예측 모형도 기본적으로는 2009년 고1 시기 방과후학교 참여 예측 모형과 동일하다. 다만 2010년 고2 시기 예측 모형에서는 고1 시기 방과후학교 참여(고1 처치변수), 고1 시기 학업성취도 척도점수 백분위(고1 결과변수), '방과후학교가 교과내용 이해에 도움이 된 정도'와 '방과후학교 계속 참여 희망'에 대한 학교 평균(고1 시기 시간-의존적 혼동변수), 고2 시기 교과 방과후학교 개설 강좌수(고2 시기 시간-의존적 혼동변수)가 추가된다. 고1 시기 baseline 공변량들은 앞서 제시한 결과들과 대동소이한 패턴을 보였고 그 계수들도 매개변수로 기능하는 시간-의존적 혼동변수들이 추가되어 사후-처치 편의(post-treatment bias)를 포함하므로 제시하지 않는다. <표 4-5>에 해당 변수들이 고2 시기 방과후학교 참여에 미치는 영향이 정리되어 있다.

<표 4-5> 고2 시기 참여에 대한 고1 처치·결과·시간 의존변수들의 영향

변수명		국어 OR(t값)		수학 OR(t값)		영어 OR(t값)	
처치	고1 방과후학교 참여	0.98 *	-1.96	0.99	(-1.10)	0.99 †	(-1.69)
결과	고1 척도점수 백분위	1.00	(1.30)	1.00	(0.06)	1.00	(1.27)
시간의존	교과이해 도움 평균	0.97	(-0.47)	0.96	(-0.72)	1.00	(0.03)
시간의존	계속 참여 희망 평균	1.08 **	(2.80)	1.12 **	(2.72)	1.05 *	(2.21)
시간의존	교과방과후 강좌수 (고2)	1.00	(1.10)	1.00	(1.10)	1.00	(0.22)

***$p \leq .001$, **$.001 < p \leq .01$, *$.01 < p \leq .05$, †$.05 < p \leq .10$

분석 결과 고2 시기 방과후학교 참여에 가장 분명한 영향을 미치는 것은 고1 시기 시간-의존적 혼동변수인 '방과후학교 계속 참여 희망에 대한 학교 평균'이다. 학교 수준에서 학생들이 고1 시기에 방과후학교 계속 참여하기를 희망하는 정도의 학교 평균이 높은 학교일수록 자연스럽

게 고2 시기 방과후학교에 참여할 학교 수준 로그-승산 평균을 증가시키는 것으로 보인다. 이 결과는 추정 모형에서 지역, 기본 학교 배경, 교장 특성, 학생·학부모·교사 구성, 개인 수준 변수들의 학교 평균 등을 모두 통제하고도 나타나는 결과이다. 반면 교과내용 이해에 도움이 된 정도의 학교 평균은 다른 학교 수준 변수들이 통제되면 별다른 유의성을 보이지 않았다.

다음으로 고2 시기 방과후학교 참여에 영향을 미치는 고1 시기 변수는 '고1 시기 방과후학교 참여' 변수다. 그런데 수학·영어·국어 과목 추정계수의 통계적 유의성과 관계없이 계수의 방향성만 보면 학교 내에서 고1 시기 방과후학교에 참여한 학생은 그렇지 않은 학생에 비하여 고2 시기에 해당 과목의 방과후학교에 참여할 승산이 체계적으로 낮은 경향을 보인다. 이 결과는 앞서 분석 자료 절에서 살펴본 바와 같이 고1 시기에 공부 못하는 학생들이 체계적으로 탈락한 표본 선택의 효과일 가능성이 있어 이를 결측 확률 가중치(non-response weight)로 교정하고 분석한 것이다. 또한 집단 평균 중심점 교정(group-mean centering)을 하고 1수준 변수들의 학교 평균을 통제한 결과이기 때문에 학교 수준 영향력이 배제된 '학교 내 효과'다. 그럼에도 불구하고 이러한 결과가 나타난 것은 기술통계 절에서 방과후학교 재참여 의사의 고1 시기 학교 평균이 낮은 점에서도 확인할 수 있듯이 일선 학교 현장에서 방과후학교에 대한 만족도가 그리 높지 않을 가능성을 시사한다. 고1 시기에 방과후학교에 참여한 학생들은 개인적인 선택이 가능한 경우 고2 시기에는 되도록 방과후학교에 참여하지 않으려는 경향을 보인다고 해석할 수 있다. 다만 이것의 통계적 유의성이 영어 유의수준 10%, 국어 유의수준 5%에서 나타나고 있기 때문에 해석에 좀 더 신중할 필요가 있다.

제 2 절 방과후학교 연속적 참여의 처치효과

<표 4-6>은 방과후학교 참여 예측모형으로부터 산출한 다층 경향점수의 역확률 가중치를 적용했을 때와 적용하지 않았을 때 각 과목당 방과후학교 연속적 참여의 처치효과를 계산한 것이다. <표 4-6>에 제시된 추정치는 기본적으로 '학교 내 효과'를 의미한다. 처치효과 추정 모형 역시 학생 수준 변수를 집단 평균 중심점 교정을 하고 학교 수준에 학생 수준 변수 학교 평균을 투입하였기 때문에 같은 학교 내에서의 비교를 의미하게 된다. 다층경향점수의 역확률 가중치를 적용하면 개인 수준 공변량들의 방과후학교 참여-미참여 집단 간 평균 차이가 사라지는 것으로 나타난다. 따라서 관찰된 공변량에 의한 선택 편의가 적절히 해소된 것으로 보았다.

전반적으로 가중치를 적용하지 않은 모형은 계수 추정치의 크기가 크고 표준오차가 작다. 반면 가중치를 적용한 모형은 계수 추정치의 크기가 감소하고 표준오차는 반대로 증가한다. 학생 수준 효과는 전수 자료를 사용하는 관계로 대부분의 평균 차이를 통계적으로 유의한 것으로 포착함에도 불구하고 통계적 유의성이 없는 것으로 나타난 것은 실제적으로는 강력한 영 효과(null effect)를 지지하는 결과로 해석할 수 있다.

모형의 절편인 β_0는 고1・고2 모두 방과후학교에 참여하지 않은 학생의 고1 시기 성적의 기댓값의 의미를 갖는다. 이는 대체로 100점 만점에 36~46점 정도인 것으로 나타난다. 이 학생들과 고1 시기에만 방과후학교에 참여한 학생들의 고1 시기 성적 기댓값과의 차이를 의미하는 β_1은 방과후학교 참여의 '고1 시기 횡단적 처치효과'로 정의된다. 가중치를 적용하지 않은 모형에서 이 값은 국어 12점, 수학 15점, 영어 10점 정도로 계산된다. 그러나 선택편의를 역확률 가중치로 통제하면 국어 6.3점, 수학 5.7점, 영어 3.7점 정도로 계수의 크기가 줄어든다. 이것은 고1 시기 방과후학교 참여는 고1 시기의 학업성취도를 약 4~6점 정도 높이는 효과가 있다는 것을 의미한다.

고2 시기 횡단적 처치효과는 고2 시기 방과후학교 참여 여부 T_2의 계

수 β_2와 이 변수의 시간과의 상호작용 효과 β_6와의 합으로 정의된다. 고2 시기 횡단적 처치효과는 역확률 가중치를 적용하지 않았을 때 국어 14점, 수학 18점, 영어 11점 정도로 산출되나 가중치를 적용하면 그 절반 수준으로 줄어든다. 고2 시기의 방과후학교 참여는 그렇지 않은 학생들에 비해 국어 7.3점, 수학 8.9점, 영어 6.0점 정도의 성적 향상 효과가 있는 것으로 나타나며 이는 고1 횡단적 처치효과보다 더 높은 수준이다.

고1 시기 방과후학교 참여가 고2 시기 성적에 미치는 효과를 의미하는 처치 지연 효과(lagged treatment effect)는 가중치를 적용하지 않은 경우 계수의 방향이 모든 과목에 걸쳐 일관되게 음(-)의 방향으로 나타나고 있어 고1 시기에만 방과후학교에 참여한 것은 고2 시기 성적을 오히려 감소시킨다는 점을 암시하고 있다. 그러나 가중치를 적용한 결과 모든 과목에서 통계적으로 유의하지 않아 실제적으로 그 효과는 0인 것으로 볼 수 있다.

반면 고1-고2 연속적 참여의 배가 효과(amplifying effect)는 영어 과목에서 특히 통계적으로 유의하게 관찰되고 있다. 영어 과목의 경우 고1-고2 시기 방과후학교에 연속적으로 참여하면 고1 시기 방과후학교 참여의 효과와 고2 시기 방과후학교 참여 효과의 단순합을 넘어서는 약 3점 정도의 추가적인 성적 상승 효과가 있는 것으로 나타난다. 달리 표현하면, 단지 고2 시기에만 영어 방과후학교에 참여한 학생은 100점 만점에 6점 정도 성적이 향상되지만 고1 시기부터 참여해 온 학생은 100점 만점에 9점(6+3) 정도 성적이 향상되는 것으로 볼 수 있다. 국어 과목은 오직 유의수준 10%에서만 통계적으로 유의하나 역시 2.3점 정도의 배가 효과를 방과후학교에 연속적으로 참여함으로써 얻을 수 있다. 수학은 고2 시기의 횡단적 처치효과가 약 9점으로 영어(6점), 국어(7.3점)보다 높은 대신 고1-고2 연속적 참여로 인한 배가 효과는 통계적으로 0으로 관찰된다. 영어와 국어 과목과 같은 '언어' 과목은 꾸준히 감을 익히고 공부를 지속해오는 것이 중요한 반면, 상대적으로 수학은 해당 학년에 배운 것을 방과 후에 꾸준히 복습하는 것이 더 학습 효과가 나은 것일 가능성이 있다.

<표 4-6> 고1-고2 방과후학교 연속적 참여의 처치효과 추정 결과

$$Y_{ij} = \beta_0 + \beta_1 T_1 + \beta_2 T_2 + \beta_3 time + \beta_4 T_1 \cdot T_2$$
$$+ \beta_5 T_1 \cdot time + \beta_6 T_2 \cdot time + \beta_7 T_1 \cdot T_2 \cdot time + u_{0j} + r_{ij}$$

회귀계수	가중치 적용 안한 모형			가중치 적용한 모형			처치효과 정의
	국어	수학	영어	국어	수학	영어	
β_0	45.13*** (0.21)	43.39*** (0.16)	45.67*** (0.21)	38.53*** (0.5)	36.02*** (0.42)	43.02*** (0.69)	
β_1	12.12*** (0.27)	15.44*** (0.22)	10.46*** (0.27)	6.29*** (0.775)	5.67*** (0.68)	3.69*** (0.87)	고1 횡단적 처치효과
β_2	-1.51*** (0.27)	0.10 (0.22)	-1.50*** (0.27)	0.61 (0.68)	0.59 (0.56)	0.82 (0.93)	
β_3	-1.89*** (0.3)	-2.20*** (0.24)	-1.57*** (0.30)	5.50*** (0.73)	6.04*** (0.64)	2.38** (0.96)	
β_4	4.03*** (0.33)	0.13 (0.30)	3.94*** (0.33)	8.62*** (1.23)	0.64 (0.92)	5.91*** (1.41)	
β_5	-15.63*** (0.39)	-15.52*** (0.32)	-13.85*** (0.39)	-6.98*** (1.08)	-6.03*** (1.68)	-4.46*** (1.17)	
β_6	15.09*** (0.39)	17.83*** (0.32)	13.03*** (0.39)	6.65*** (0.98)	8.27*** (0.84)	5.21*** (1.26)	
β_7	2.48*** (0.45)	0.90*** (0.42)	3.01*** (0.45)	-6.34*** (1.73)	0.01 (1.30)	-2.99 (1.86)	
$\beta_2 + \beta_6$	13.58*** (0.27)	17.93*** (0.11)	11.53*** (0.27)	7.26*** (0.70)	8.86*** (0.31)	6.03*** (0.84)	고2 횡단적 처치효과
$\beta_1 + \beta_5$	-3.51*** (0.27)	-0.08 (0.11)	-3.39*** (0.27)	-0.69 (0.72)	-0.36 (0.39)	-0.77 (0.73)	고1 처치 지연 효과
$\beta_4 + \beta_7$	6.51*** (0.33)	1.03*** (0.15)	6.95*** (0.3)	2.28† (1.20)	0.65 (0.47)	2.92** (1.23)	고1-2 연속참여 배가효과

***p≤.001, **.001<p≤.01, *.01<p≤.05, †.05<p≤.10

제 5 장 결론

제 1 절 요약

이 연구는 국가수준 학업성취도 평가의 고1-고2 시기 개별 학생 반복 측정 자료를 활용하여 방과후학교의 연속적 처치(sequential treatment) 상황이 학생들의 학업성취도에 미치는 종단적 효과를 분석하였다. 1단계에서는 방과후학교 참여 선택에 영향을 미치는 요인들을 다층적 속성에 주목하여 분석하고 2단계에서는 1단계에서 얻은 정보를 활용하여 역확률 가중치 방법으로 선택 편의를 교정한 후 방과후학교 참여의 연속적 처치 효과를 분석하였다. 연구문제에 대한 대답을 정리하면 다음과 같다.

첫째, 방과후학교 참여 여부에는 개인 수준 요인보다 학교 수준 요인의 영향력이 더욱 큰 것으로 나타났다. 집단 내 상관계수(ICC)는 국어·수학·영어 과목 모두에서 50%를 초과하여 방과후학교 참여 여부 선택에 관한 전체 변산 중 절반 이상을 학교 수준 요인이 설명하였다. 과목별 MOR도 7.0-8.0 전후로 산출되어 학생 개인은 동일해도 학교가 달라짐으로써 방과후학교에 참여할 확률 자체가 평균적으로 7배 정도 차이가 나는 것으로 나타났다.

둘째, 방과후학교 참여 여부를 예측하는 요인을 분석한 결과 지역이나 전문계고 여부 등 기본적인 학교 정보만으로도 방과후학교 참여 확률의 학교 간 차이를 52%~55% 정도 설명하는 것으로 나타났다. 이러한 학교 배경이 통제되었을 때에 기본적으로 '학교 교사를 믿고 성실하게 학교 공부를 수행하는 학생'일수록, 그리고 그러한 학생들이 많은 학교일수록 방과후학교 참여 확률이 증가하였다. 본 연구의 가장 특징적인 발견은 학교 수준 방과후학교 참여 확률에 대해 교장 개인의 특성이 독립적인 영향력을 갖는다는 점이다. 여자 교장 혹은 공모제 교장일수록, 교장의 행정업무 및 외부기관 회의 비율이 높을수록 해당 학교의 방과후학교 참여율이 감소하여 방과후학교 운영에 관한 학교 수준 운영 방침에 학교

장의 성향이 유의하게 반영되고 있다는 점을 보여주었다.

셋째, 고2 시기 방과후학교 참여 여부에 고1 시기 시간-의존적 혼동요인이 영향을 미쳐 종단적 선택 편의가 존재하는 것으로 나타났다. 고2 시기 방과후학교 참여에 가장 분명한 영향을 미치는 것은 고1 시기 시간-의존적 혼동변수인 '방과후학교 계속 참여 희망에 대한 학교 평균'이었다. 학생들이 고1 시기에 방과후학교 계속 참여하기를 희망하는 정도의 학교 평균이 높은 학교일수록 자연스럽게 고2 시기 방과후학교에 참여할 학교 수준 로그-승산 평균이 증가하였다.

넷째, 방과후학교 연속적 참여의 처치 효과를 분석한 결과 고1·고2 시기 횡단적 처치효과와 고1-고2 연속적 참여의 배가효과(amplifying effect)가 존재하는 것으로 나타났다. 고1 시기 방과후학교 참여는 고1 시기의 학업성취도를 100점 만점에 약 4~6점 정도 높이는 횡단적 처치효과가 있었고, 고2 시기의 방과후학교 참여는 고2 시기 학업성취도를 약 6~9점 정도 향상시키는 횡단적 처치효과가 있었다. 고1 시기 방과후학교 참여가 고2 시기 성적에 미치는 효과를 의미하는 처치 지연 효과(lagged treatment effect)는 통계적으로 0으로 나타났지만, 영어 과목의 경우 고1-고2 시기 방과후학교에 연속적으로 참여하면 고1 시기 방과후학교 참여의 효과와 고2 시기 방과후학교 참여 효과의 단순합을 넘어서는 약 3점 정도의 추가적인 성적 상승 효과, 즉 배가효과가 존재하는 것으로 나타났다.

제 2 절 논의

방과후학교 참여에 학교 수준 요인이 상당한 영향을 미칠 것이라는 점은 일면 상식인 것처럼 보이지만 기존 방과후학교 연구들에서는 잘 고려되지 않던 부분이다. 본 연구의 결과는 단위 학교 방과후학교 운영 방침(school policy)에 따라 학생들의 방과후학교 참여가 집단적으로 변한다는 사실을 보여준다. 특히 운영 방침의 결정 권한을 가진 학교장 특성 변수가 가지는 독립적인 영향력은 일선 학교 현장에서 방과후학교가 사실상 학교장의 '전원 강제 참여' 방침에 좌우되는 경우가 많다는 현장의 전언이 전국 전수 데이터를 통해 포착된 것으로 볼 수 있다. 방과후학교 제도가 팽창하여 일정한 궤도에 오른 이명박 정부 시기는 2009년부터 국가수준 학업성취도 평가를 전수로 시행하여 시·도 교육청 성적 순위를 공개하는 등 외부로부터의 학력 향상 압박이 함께 증가한 시기다. 2008년 51%이던 교과 방과후학교 프로그램이 2012년 60.9%로 4년간 10%포인트 가까이 늘어난 것은 학력향상에 대한 압박을 단위 학교에서 교과 중심 방과후학교 운영 및 강제 참여 유도 방식으로 대응했을 가능성을 보여준다. 따라서 방과후학교 연구의 맥락에서 선택과정을 모형화할 때는 다층 경향점수 추정하는 방식 등 반드시 이와 같은 학교 수준 요인을 고려해야 한다.

다음으로 고1-고2 시기 방과후학교 연속 참여의 배가 효과가 존재한다는 본 연구의 결과는 교육 프로그램에 대한 연속적인 참여를 유도하는 학교 정책의 중요성을 환기시키고 있다. 고2 시기에만 영어 방과후학교에 참여한 학생은 100점 만점에 6점 정도 성적이 향상되지만 고1 시기부터 참여해 온 학생은 100점 만점에 9점(6+3) 정도 성적이 향상된다. 이러한 배가 효과를 고려하여 학교 현장에서는 단발성으로 그치는 고립된 프로그램을 여러 개 운영하는 것보다는 학년에 걸쳐 연속적으로 쌓아올릴 수 있는 '시리즈 강의'를 적절히 구성하여 제공할 필요가 있다. 특히 방과후학교 자유수강권 지원금이 고1-고2에 걸쳐 연속적으로 참여하는 저소득층 학생들에게 더 많이 제공될 수 있도록 금전적 인센티브를 설계

하는 것도 좋은 방안이 될 것이다.

마지막으로 이 연구는 교육 연구자들 사이에서 연속적 처치(sequential treatment)에 대한 인과효과 추정 방법의 중요성을 제안하는 의의를 가진다. 특히 방과후학교 만족도와 같은 시간-의존적 혼동변수가 처치효과를 왜곡하는 요소가 될 수 있다는 점을 인식하고 연구 모형에 적절히 반영할 필요가 있다. 본 연구의 결과처럼 '방과후학교 만족도'가 차년도 방과후학교 선택에 영향을 미친다는 사실 그 자체만으로도 정책에서 환류(feedback)의 중요성을 상기시키는 중요한 연구결과가 될 수 있다. 다만 그러한 시간-의존적 혼동변수 중 비관측 요인들(unobserved factors)에 대한 민감도 분석을 수행하지 못했기 때문에 본 연구에서 추정한 처치효과에 크고 작은 선택 편의가 남아있을 수 있다는 점은 한계로 지적되며, 차후 적절한 분석을 통해 연구의 결론을 확증해가는 작업을 지속할 필요가 있다.

참 고 문 헌

교육과학기술부(2012). 2012년 방과후학교 운영 현황 분석. http://moe.go.kr에서 2012. 9. 7. 인출.

교육과학기술부(2011. 10. 5.). 교과부 2012년도 예산 52조 9,426억원 편성. http://www.moe.go.kr에서 2012. 3. 1. 인출.

경향신문(2012. 9. 6.). 방과후학교 '불편한 진실'… 억지 실적 띄우기. http://news.khan.co.kr에서 2012. 9. 6. 인출.

김경근・황여정(2009). 중학생의 방과후학교 참여 결정요인. **교육사회학연구**, 19(2), 31-57.

김성식(2012). 학업성취도에 대한 방과후학교 효과 분석. **제 16회 KICE 교육과정・평가 정책포럼 자료집**, 129-150.

김인숙・변종임・조순옥(2010). 방과후학교 참여의 영향. **열린교육연구**, 18(4), 246-270.

김혜숙(2012). 방과후학교 프로그램이 학업성취도 및 사교육에 미치는 영향. **아시아교육연구**, 13(3), 87-114.

박소영(2008). 방과후 학교와 EBS 수능강의의 사교육비 경감효과. **교육행정학연구**, 26(1), 391-411.

박현정・변종임・조순옥(2009). 방과후학교 효과 분석. **제 3회 한국교육종단연구 학술대회 자료집**, 3-31.

배상훈・김성식・양수경(2010). 방과후학교 참여와 사교육비 지출 및 학업성취 수준과의 관계. **교육행정학연구**, 28(2), 55-79.

변수용・김경근(2009). 중학생의 방과후학교 참여가 사교육 수요에 미치는 영향. **제 3회 한국교육종단연구 학술대회 자료집**, 63-101.

변수용・황여정・김경근(2011). 방과후학교 참여가 학업성취에 미치는 영향. **교육사회학연구**, 21(2), 57-85.

서울특별시교육청(2010. 3. 2.). 2010년 서울 방과후학교 자유수강권 운영계획. 서울: 서울특별시교육청.

송미영・남민우・강태훈・김창일(2011). **학교교육 성과 평가를 위한 학**

업성취지표와 평가 모형 개발: 국가수준 학업성취도 평가의 종단척도 구축. 한국교육과정평가원 연구보고 RRE 2011-6-1.

우명숙・이수정(2010). 중학생의 방과후학교 프로그램 참여가 영어 학업성취도와 사교육비에 미치는 영향분석. **교육재정경제연구**, 19(4), 105-134.

이준호・박현정(2012). 방과후학교 자유수강권 지원금은 교육 기회 및 결과의 평등에 기여하였는가? -서울지역 고등학생의 경우. **교육재정경제연구**, 21(2), 53-85.

채재은・임천순・우명숙(2009). 방과후학교와 수능강의가 사교육비 및 학업성취도에 미치는 효과 분석. **교육재정경제연구**, 18(3), 37-62.

최길찬(2010). 두 시점 유사종단자료를 이용한 학교수행 성장모형. **아시아교육연구**, 11(2), 163-177.

한국일보(2009. 9. 4.). 방과후 학교, 일제고사용 '강제 보충수업' 변질. http://news.hankooki.com에서 2012. 9. 4. 인출.

Barrow, L. & Rouse, C.(2005). *Causality, causality, causality: the view of education inputs and outputs from economics*. Federal Reserve Bank of Chicago Working Paper Series with number WP-05-15.

Bryk, A., Thum, Y. M., Easton, J., & Luppescu, S.(1998). Assessing school academic productivity: The case of Chicago school reform. *Social Psychology of Education, 2*, 103-142.

Choi, K., Seltzer, M., & Herman, J.(2004). Children left behind in AYP and non-AYP schools: Using student progress and the distribution of student gains to validate AYP. *CRESST technical report #637.* Los Angeles:CA.

Cole, R. & Hernan M. A.(2008). Constructing inverse probability weights for marginal structural models. *American Journal of Epidemiology, 168*(6), 656-664.

Cudeck, R. & Henly, S. J.(1991). Model selection in covariance

structure analysis and the "problem" of sample size : A clarification. *Psychological Bulletin, 109*, 512-519.

DePascale, C. A.(2006). *Measuring Growth with the MCAS Tests: A consideration of vertical scales and standards*. Retrieved from http://www.nciea.org/publications.

Gelman, A., & Hill, J.(2007). *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Cambridge University Press.

Haviland, A., Nagin, D. S., Rosenbaum, P. R., and Tremblay, R. E.(2008). Combining group-based trajectory modeling and propensity score matching for causal inference in nonexperimental longitudinal data. *Developmental Psychology, 44*(2), 422-436.

Hong, G. & Raudenbush, S. W.(2006). Evaluating kindergarten retention policy: A case study of casual inference for multi-level observational data. *Journal of American Statistical Association, 101*, 901-910.

Hong, G., & Raudenbush, S. W.(2008). Causal inference for time-varying instructional treatments. *Journal of Educational and Behavioral Statistics, 33*(3), 333-362.

Kim, J.(2006). *Causal Inference in Multilevel Settings: Estimating and Using Propensity Scores when Treatment is Implemented in Nested Settings*. Unpublished Doctoral dissertation, University of California, Los Angeles.

Kim, J., & Seltzer, M.(2007). *Causal inference in multilevel settings in which selection processes vary across schools (No. 708)*. Los Angeles: Center for the Study of Evaluation (CSE).

Larsen, K., & Merlo, J.(2005). Appropriate assessment of neighborhood effects on individual health: Integrating random and fixed effects in multilevel logistic regression. *American Journal of*

Epidemiology, 161, 121-160.

Lauver, S. & Little, P. M. D.(2005). Strategies for attracting and sustaining participation in afterschool program. *The School Administrator*, 2005(May).

Little, R. J. A., & Rubin, D. B.(2002). *Statistical Analysis with Missing Data(2nd ed.)*. Hoboken, NJ:Wiley.

Murnane, R. J. & Willett, J. B. (2011). *Methods matter: Improving causal inference in educational and social science research*. New York: Oxford University Press.

Oakes, J. M.(2004) The (mis)estimation of neighborhood effects: causal inference for a practicable social epidemiology. *Social Science and Medicine, 58,* 1929-1952.

Rabe-Hesketh, S., Skrondal, A.(2012). *Multilevel and Longitudinal Modeling Using Stata(3rd Edition).* Texas: Stata Press.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical Linear Models: Applications and Data Analysis Methods (2nd ed.).* Thousand Oaks.:Sage Publication.

Raudenbush, S. W., Hong, G., & Rowan, B.(2002). *Studying the Causal Effect of Instruction with Application to Primary-School Mathematics*. Paper presented at the Research Seminar : Instructional and Performance Consequences of High Poverty Schooling, Washington, D.C.

Robins, J.(1987). A graphical approach to the identification and estimation of causal parameters in mortality studies with sustained exposure periods. *Journal of Chronic Diseases, 40*(2), 139S-161S.

Robins, J., Greenland, S., & Hu, F.-C.(1999). Estimation of the causal effect of a time-varying exposure on the marginal mean of repeated binary outcome. *Journal of the American Statistical*

Association, 94(447), 687-700.

Robins, J., Hernán, M. & Brumback, B.(2000). Marginal structural models and causal inference in epidemiology. *Epidemiology, 11*(5), 550-560.

Rosenbaum, P. R., & Rubin, D. B.(1985). Constructing a control group using multivariate matched sampling methods that incorporate the propensity score. *The American Statistician, 39*(1), 33-38.

Rubin, D.(2001). Using propensity score to help design observational studies: Application to the tobacco litigation. *Health Services and Outcomes Research Methodology, 2*, 169-188.

Seltzer, M.(2004). The use of hierarchical models in analyzing data from experiments and quasi-experiments conducted in field settings. In Kaplan(ed), *The Handbook of Quantitative Methodology for the Social Science*. Thousand Oaks: Sage Publication.

Van der Wal, W. M., & Geskus, R. B. (2011). ipw: An R Pakage for Inverse Probability Weighting. *Journal of Statistical Software, 43*(13).

Willms, J., & Raudenbush, S. W.(1989). A longitudinal hierarchical linear model for estimating school effects and their stability. *Journal of Educational Measurement, 26*(3), 209-232.

Abstract

Causal Inference for Time-varying After-school Program Participation

Joon-ho Lee

Department of Education

The Graduate School

Seoul National University

The purpose of this study is to analyze the effects of time-varying after-school program participation on repeatedly observed student achievement. This study responds to three methodological challenges: (a) The students' after-school program participation may depend on a host of confounding factors including both school-level and student-level characteristics; (b) Students' learning outcomes under a certain after-school program may depend on the program participation of other students; and (c) time-varying confounding poses special problems of endogeneity. I addressed these challenges by modifying the stable unit treatment value assumption (SUTVA) to identify potential outcomes and causal effects, and by integrating

inverse probability of treatment weighting (IPTW) into multilevel models. Using data from the National Assessment of Educational Achievement (NAEA), I applied these methods to study the impact of the sequential treatment of after-school program in Grades 1 and 2 of Korean high schools. The result can be summarized as follows.

First, the results of the selection process of after-school program participation showed that students with same characteristics might have different chances depending on the schools they are attending. The differences were due to a combination of student and school-level factors including school policy decided by principal, resource availability or school compositional characteristics. This study also confirmed the presence of time-varying confounder, an intent to participate again in the after-school program, which provided a rationale for inverse probability of treatment weighting(IPTW) to remove observed confounding in the multilevel longitudinal context.

Second, there was a significant positive effect of the after-school program participation in Grade 1 on Grade 1 academic achievement, and that in Grade 2 alone on Grade 2 outcome (single year treatment effect). However, there was no evidence of the effect of grade-1 treatment on grade-2 outcome for those not receiving grade 2 treatment (lagged treatment effect). Participating after-school program again in Grade 2 significantly enhanced the effect of English after-school program treatment received in Grade 1 (amplifying effect). This result shows that if the second-year after-school program is built upon deliberately and effectively on the first-year after-school program, the benefit of the sequence may significantly exceed the sum of two average effects of the after-school program participation of each single year in Grade-1 and Grade-2.

keywords : sequential treatment effect, multilevel propensity score, inverse probability of treatment weighting, marginal structural model
Student Number : 2009-23375