

혁신학교가 중학생 학업성취도에 미치는 효과 분석 : 경향점수매칭과 분위회귀분석 활용

최 예 슬(서울대학교, 연구원)
이 소 라(한국교육과정평가원, 부연구위원)*

요 약

본 연구는 경기교육종단연구 자료를 활용하여 혁신학교가 중학교 학업성취도에 미치는 영향을 분석하였다. 경향점수매칭(Propensity Score Matching) 기법을 활용하여 혁신학교에 진학한 학생들과 유사한 특성을 가진 일반학교 학생들을 비교 집단으로 삼아 선택편의(selection bias)를 보정하였다. 이후 경향점수를 도출하는데 활용한 변수들을 회귀분석의 통제변수로 재투입하는 이중강건추정(doubly robust estimation)으로 엄밀한 인과효과를 분석하고자 하였다. 또한 분위회귀분석(Quantile Regression)을 활용하여 학생들의 학업성취 수준별 집단에 따라 혁신학교가 차별적 영향을 미치는지 분석하였다. 연구결과 첫째, 혁신학교와 일반학교의 학업성취는 통계적으로 유의미한 차이가 없는 것으로 나타났다. 둘째, 혁신학교는 일반학교에 비해 국어교과 하위 25%에 속하는 학생들에게 유의미한 학업성취 향상의 효과가 있는 것으로 나타났다. 이러한 연구결과를 바탕으로 다음과 같은 시사점을 제언하였다. 첫째, 국어과 수업에서 혁신학교의 학습자중심 수업방식을 적극적으로 활용할 필요가 있다. 둘째, 혁신학교 효과를 엄밀하게 추정하기 위해 준실험설계를 활용해야 한다. 셋째, 혁신학교 정책의 효과적 시행을 위하여 정책 효과가 미치는 대상과 영역의 특성을 보다 세부적으로 고려해야 한다.

[주제어] : 혁신학교, 학교효과, 학업성취도, 경향점수매칭, 분위회귀분석

I. 서 론

지난 10여 년간 급격하게 전국으로 확산된 혁신학교는 민주적 학교 운영과 학습자중심교육을 통해 지식전달, 입시위주의 공교육을 정상화하고자 하는 학교개혁 모델이다. 혁신학교의 양적 확산과 함께, 혁신학교에서 중점적으로 시도되고 있는 학습자중심수업

* 교신저자(slee486uw@gmail.com)
■ 접수일(2020.06.29), 심사일(2020.07.06), 게재확정일(2020.07.30)

의 방향이 기초학력을 저하시킨다는 비판이 꾸준히 제기되어 왔다.

혁신학교가 전통적 관점에서의 인지적 성장보다 전인적 성장과 시민성, 창의성 등과 같은 새로운 관점에서의 학생역량에 초점(박상현, 김보은, 정주영, 손동빈, 2016)을 두고 있다 할지라도, 공교육 체제 내의 학교가 기초학력을 보장하는 것은 매우 중요한 과제이다. 또한 국가중심의 교육제도가 운영되는 한국의 현실상 국가공통교육과정을 충실히 소화하지 않고 체계화된 교육과정에서 벗어난 수업방식이 시도되는 것에 대한 학부모의 불안(성기선, 2014)은 필연적이다. 따라서 혁신학교의 인지적 성과를 증명하여 사회적 불신을 완화시키기 위한 노력은 혁신학교를 공교육의 혁신모델로 안착시키려는 개혁과정에서 반드시 수행해야하는 책무일 것이다.

학력저하논란에 대응하여 연구자들은 혁신학교와 일반학교 재학생의 학업성취도에 유의한 차이가 없다고 밝힘으로써 적어도 학력저하가 사실이 아니라는 주장을 뒷받침하거나(박상현 외, 2016; 배종현, 김주후, 2016; 윤혜원, 2020; Min, Jung, & Kim, 2017), 혁신학교가 장기적으로는 일반학교보다 높은 성장률을 보인다고 분석하고 있다(김민규, 박세진, 2019; 배종현, 김주후, 2016; 백병부, 박미희, 2014; 서민희, 전경희, 2018; 박세진, 이해니, 이승호, 2019). 나아가, 사회경제적 배경에 따른 교육격차를 감소(백병부, 박미희, 2015; 원순아, 엄기형, 2018)시키거나 기초학력 미달 학생 비율을 감소(배종현, 김주후, 2016)시킨다는 분석도 제시되었다.

그러나 이러한 선행연구들 중 다수는 선택편의(selection bias)에 의한 혁신학교와 일반학교 학생들의 이질적 특성을 통제하지 못하거나, 이러한 선택편의를 고려했다 하더라도(윤혜원, 2020; Min et al., 2017) 혁신학교 진학 전 시점의 자료를 통제하지 못하였다는 제한점을 가지고 있다. 또한 혁신학교의 학습자중심수업 운영이 기초학습이 부족한 학생들에게 부정적인 영향을 미칠 것이라는 우려에도 불구하고, 학생들의 학업능력 수준에 따른 혁신학교의 차별적인 효과를 구체적으로 주목한 연구가 그동안 수행되지 않았다. 따라서 혁신학교의 효과에 대한 보다 정확한 이해를 위해서는 선택편의를 통제하고 학생들의 학업수준 집단별로 미치는 영향에 대한 엄밀하고 체계적 검증이 요구된다.

따라서 본 연구는 혁신학교가 중학교 학업성취도에 미치는 영향을 분석하고자, 경향점수매칭(Propensity Score Matching) 기법을 활용하여 혁신학교에 진학한 학생들과 유사한 특성을 가진 일반학교 학생들을 비교 집단으로 삼아 선택편의를 보정하였다. 이후 경향점수를 도출하는데 활용한 변수들을 회귀분석의 통제변수로 재투입하는 이중강건추정(doubly robust estimation)으로 엄밀한 인과효과를 분석하고자 하였다. 다음으로 분위회귀분석을 활용하여 학생들의 학업성취 수준에 따른 집단별로 혁신학교가 차별적 영향을 미치는지를 분석하였다. 이를 통해 기존의 연구에서 주목하지 못하였던 정책대상에 따른 이질적 효과를 확인함으로써 학력저하를 둘러싼 혁신학교 효과에 대한 학술적 증

거를 확장하고 향후 정책 운영방향에 대한 실질적인 시사점을 제공하는데 기여하고자 하였다. 본 연구는 연구대상으로 특히 중학교 급간에 주목하였다. 고등학교와 달리 주소지 기반의 근거리 배정을 원칙으로 하는 중학교 급간에서 ‘좋은’ 학군에 대한 학부모의 관심이 높으며, 초등학교에 비해 대학입시에 근접한 중학교 단계부터 혁신학교를 꺼리는 경향이 강해지기 때문이다. 이러한 상황에서 중학교 단계의 혁신학교가 학업성취도에 미치는 영향을 보다 세부적으로 분석하는 것은, 현 제도 내에서 중학교 배정에 대한 학부모의 불안을 감소시키고 정책 결정에 대한 시사점을 제공하는데 큰 의의를 가진다.

본 연구의 구체적인 연구문제는 다음과 같다. 첫째, 혁신학교 재학 여부에 따라 중학생의 학업성취도에 차이가 있는가? 둘째, 혁신학교 재학 여부가 중학생의 학업성취 수준에 따라 차별적 효과를 가지는가?

II. 선행연구

1. 혁신학교와 학력저하담론

혁신학교는 “민주적 학교운영 체제를 기반으로 윤리적 생활공동체와 전문적 학습공동체를 형성하고 창의적 교육과정을 운영하여 학생들이 삶의 역량을 기르도록 하는 자율학교(경기도교육청, 2020; p.2)”로 정의되는 학교 개혁 모델이다. 2009년 경기도 교육청의 13개의 시범학교에서 출발한 이후, 2020년 현재 전국에 걸쳐 약 2천여 개의 학교로 확산되어 운영되고 있다. ‘민주적 학교운영체제’, ‘윤리적 생활공동체’, ‘전문적 학습공동체’ 및 ‘창의적 교육과정’이라는 네 가지의 기본원리를 토대로, 기존의 획일화된 지식습득과 암기식, 입시 위주의 교육 환경을 개선하고 학교공동체 차원의 참여적인 학교운영과 역량 및 창의성을 기를 수 있는 학습자중심의 교육방식을 추구한다. 그동안 교육현장에서 공교육 혁신을 추구하는 다양한 교육개혁의 움직임이 있어왔으나, 혁신학교는 현장 교사들의 자발적인 움직임으로부터 ‘상향식(bottom-up)’ 개혁을 통해 정부 정책으로 확대되어 10년 이상 지속된 사례라는 점에서 주목할 만하다.

그러나 혁신학교가 강조하는 토론 및 체험 중심의 학습자중심교육으로 인해 상급학년 진학에 필요한 학습결손을 유발하여 기초학력을 저하시킨다는 비판이 오랫동안 제기되었다. 2016년도 국가수준 학업성취도 평가에서 혁신고교가 전체 고교 평균 4.5%보다 2배 높은 11.9%의 학생들이 기초학력 미달평가를 받는 등, 전국 평균에 비해 혁신학교 학생들의 기초학력 미달 비율이 높게 나타났다(조선일보, 2017.10.12.). 경기도교육청에 설문조사에 따르면, 교사, 학부모, 학생들은 혁신학교 교육에 만족한다는 대답이 78.2%,

76.3%, 83.5%에 이를 정도로 높은 만족도를 나타내지만, 혁신학교에 다니면 일반학교에 비해 기초학력이 떨어질 수 있다는 불안 역시 가지고 있는 것으로 나타난다(연합뉴스, 2019.10.21.). 학부모들 사이에서 혁신학교에 다니면서 기초학습을 위한 사교육 의존도가 오히려 높아졌다는 불만이 제기되기도 하였다(나민주, 박소영, 2013). 이러한 학력저하를 우려한 학부모들이 초등학교보다 입시와 직결되는 중·고등학교 급간에서 혁신학교를 기피하고(서울신문, 2019.1.8), 서울 일부 지역에서는 혁신학교 지정에 반대하는 학부모들이 집회를 여는 등 사회적 갈등이 일어나기도 하였다(동아일보, 2019.6.30.).

혁신학교를 옹호하는 입장에서는, 전통적 영역의 학업성취 효과를 논하는 것이 혁신학교의 본질을 왜곡하며 학습자중심의 수업방식이 새로운 시대의 학력을 반영하기 때문에 장기적으로 보았을 때 진정한 학력 향상에 도움을 준다고 주장한다(박상현 외, 2016). 또한 진보적 성향의 교육감들과 교사들에 의해 주도적으로 확산된 배경에 의해, 혁신학교의 학력저하 담론을 진보진영의 교육개혁에 대한 보수진영의 억압이나 진영논리에 따른 이데올로기적 논쟁과 같은 정치적 대립으로 바라보는 시선들도 존재한다(백병부, 성열관, 양성관, 2019; 안혜정, 2019).

그러나 교육현장에서 국가교육과정과는 달리 체계화되지 않은 혁신적인 교육방식과 부족한 상급학교 진학을 위한 준비 등에 대한 학부모의 불안(성기선, 2014)은 교육 수요자에게 타당한 문제제기로 보인다. 실제로 혁신학교 운영의 질적 사례를 살펴보면, 교사들 사이에서도 정규교과과정을 망라하지 않는 혁신학교의 수업 운영에 대해 우려하는 목소리가 있으며(정혜진, 조영하, 2020), 학부모와 학생들이 사교육을 통해 학업을 보충하려고 하는 경향도 나타나고 있다(김수경, 2011; 유경훈, 2014). 이처럼 혁신학교를 둘러싼 학력저하담론은 단순히 정치적인 것으로 치부되기에는 매우 현실적인 교육현장의 목소리에서 비롯되었음을 알 수 있다.

2. 혁신학교가 학업성취도에 미치는 영향에 관한 선행연구

앞에서 살펴본 혁신학교의 학력저하담론에 대응하여 혁신학교가 인지적 성취에 미치는 영향을 분석하려는 연구들이 수행되었다. 혁신학교의 학업성취 효과를 분석한 대표적인 선행연구를 종합하여 정리하자면 <표 1>과 같다.

<표 1> 혁신학교가 학업성취에 미치는 영향에 대한 선행연구 정리

저자	데이터	급간	종속변수	분석방법	주요결론
백병부, 박미희(2015)	GEPS (1~2차)	초중고	국어, 영어, 수학,	2수준 HLM	혁신학교 재학생은 일반학교보다 가정의 사회경제적 배경의 영향을 적게 받음
서민희, 전경희(2018)	GEPS (2~5차)	초,중	국어, 영어, 수학	다집단 잠재성장모형	초등 혁신학교에 재학한 학생들의 학업성취 성장률이 높음
김민규, 박세진(2019)	GEPS (1~6차)	중	국영수 평균	3수준 HLM	혁신학교에 재학한 학생들의 학업성취 성장률이 높음
원순아, 엄기형(2018)	GEPS (1~4차)	중	국어, 영어, 수학, 국영수 평균	하이브리드모형	혁신학교는 중1 학업성취에 유의미한 영향 없으나, 부의 학력의 영향을 유의미하게 감소시킴
박상현 외(2016)	SELS (3차, 5차)	중,고	국어, 영어, 수학	PSM DID	중학교, 고등학교 급간에서 혁신학교-일반학교 간의 차이 없음
배중현, 김주후(2016)	학교 알리미 (경기도)	중	국어, 영어 교과 (국가수준 학업성취도 평가)	반복측정 공분산 분석	학교유형별 차이 없으나, 국어 기초학력 미달 및 영어 기초학력 성취수준 비율이 혁신학교에서 가파른 감소세를 보임
안영은, 박세진(2019)	GEPS (4~6차)	중	국영수 평균	잠재성장/ 성장혼합모형	초기치는 차이 없으나, 혁신학교에 재학한 학생들의 학업성취 성장률이 높음
박세진, 이해니, 이승호(2019)	SELS (4~6차)	고	국어, 영어, 수학	반복측정 공분산분석	특목고, 자사고에 비해 학업성취도는 낮으나, 높은 성장으로 인해 최종적인 학업성취도 격차는 감소
윤혜원(2020)	GEPS (1, 3차)	고	국어, 영어, 수학	PSM OLS	혁신학교-일반학교 간의 성취도 차이 없음
Min et al. (2017)	GEPS (1~3차)	고	국어, 영어, 수학	PSM 2수준 HLM	혁신학교-일반학교 간의 성취도 차이 없음

GEPS 경기교육종단연구; SELS 서울교육종단연구; PSM(Propensity Score Matching) 경향점수매칭; DID (Difference in Difference) 이중차분법; HLM(Hierarchical Linear Model) 위계적 선형분석

기존 연구들은 혁신학교가 학업성취에 부정적인 영향이 없고, 일반학교와의 차이가 없다는 것을 밝힘으로써, 앞서 논의된 혁신학교 재학이 학력을 저하시킨다는 우려를 반박하고자 하였다. 박상현 외(2016)는 경향점수매칭으로 중학교 및 고등학교 급간의 일반학교와 혁신학교를 매칭한 뒤 이중차분법으로 선택편의를 통제하였는데, 학교 유형 간 학업성취의 유의미한 차이는 발견되지 않았다. 배중현, 김주후(2016)는 학교알리미 자료

를 활용하여 경기도교육청 소속의 혁신학교 30개와 동일한 소재지에 위치하고 학교 규모가 비슷한 일반학교 30개를 1대 1 대응 기법으로 비교하였다. 국가수준의 학업성취도 평가의 국어, 영어 교과와 기초학력 및 기초학력에 미달하는 성취수준의 학생들의 비율을 살펴보았을 때, 학교유형별 차이가 유의미하지 않았다. 그러나 국어교과에서 기초학력 미달 성취수준 학생들과 영어교과에서 기초학력 성취수준 학생들의 비율이 시간이 지남에 따라 혁신학교가 일반학교보다 가파르게 감소한 것을 확인하였다.

윤혜원(2020)과 Min et al. (2017)은 GEPS 자료를 활용하여 경향점수매칭을 통해 혁신학교와 일반학교 간의 이질성을 통제한 후, 혁신학교가 고등학교 학업성취에 미치는 영향을 분석하였다. 두 연구 역시 중학교와 동일하게 고등학교 급간에서 혁신학교와 일반학교의 학업성취도 면에서 차이를 발견하지 못하였다. 그러나 GEPS 자료에서 고등학교 진학 전 학생들의 개인 및 가정 특성에 관련한 자료를 제공하지 않고 있어, 이를 통제하지 못했다는 한계를 가진다. 이처럼 혁신학교와 일반학교의 학업성취도의 차이가 없다고 밝힌 연구들은 중고등학교 급간에서 학력저하 논란이 사실이 아니라는 주장을 뒷받침하는 동시에, 혁신학교 운영이 지속될수록 특정 교과의 기초학력 획득에 일부 긍정적인 영향을 줄 수 있다(배종현, 김주후, 2016)고 제안한다.

반면, 혁신학교가 학력에 부정적인 영향을 미치지 않을 뿐만 아니라 장기적으로 보았을 때 오히려 학업성취도 변화율에 긍정적인 영향을 미친다고 분석한 연구들도 있다(김민규, 박세진, 2019, 배종현, 김주후, 2016; 백병부, 박미희, 2014). 서민희, 전경희(2018)가 다집단 잠재성장모형을 통해 분석한 결과, 혁신 초등학교에 재학한 학생들과 일반학교 학생들의 초기 학업성취도의 통계적인 차이는 없었으나, 학년이 올라갈수록 혁신학교 학생들이 높은 성장변화율을 보여 장기적으로는 초등단계에 혁신학교에 재학한 학생들의 인지적 성취가 높은 것으로 나타났다. 그러나 이 연구는 가정환경을 통제하지 않은 상태에서 학업성취의 기술적인 변화추이만을 분석하였으며, 초등학교와 중학교 급간을 구분하지 않았다는 한계를 지니고 있다.

이러한 제한점을 극복하기 위해, 김민규, 박세진(2019)은 동일한 GEPS 자료로 가정환경을 통제한 상태에서 일반학교와 혁신학교에 재학 중인 학생들의 초등학교(1~3차)와 중학교(4~6차) 국어, 영어, 수학 점수의 평균값의 변화추이를 분석하였다. 초등학교 단계에서는 초기치와 성장률에 있어 혁신학교와 일반학교의 유의미한 차이는 없었으나, 중학교의 경우 혁신학교 재학이 학생들의 성장률에 긍정적인 영향을 미치는 것으로 나타났다.

같은 GEPS 자료를 활용한 안영은, 박세진(2019) 역시 혁신학교 재학생은 중학교 학업성취의 초기치에 차이가 없으나 높은 성장률을 보이는 것을 확인하였다. 특히 이 연구는 성장혼합모형을 사용하여 혁신 중학교와 일반 중학교 학생들을 학업성취 성장 유형

에 따라 상위, 중위상승, 중위하락, 하락이라는 네 개의 집단으로 분류하여, 혁신학교가 일반학교에 비해 상위, 하락집단의 학업성취 성장과 하락집단의 감소폭에 긍정적인 영향을 미치고 있음을 보여주었다. 그러나 이 연구는 두 학교유형에 입학한 서로 다른 이질적인 학생집단의 선택편의를 통제하지 않았으며, 성장혼합모형으로 분류한 네 개의 성취도 집단유형의 비율이 두 학교유형에서 다름에도 불구하고 초기치와 변화율을 단순 비교하였다. 따라서 혁신학교가 일반학교에 비해 여러 성취도 집단의 유형별로 모두 긍정적인 영향을 미쳤다고 결론내리기에는 한계가 있으며 엄밀한 탐색이 요구된다.

박세준, 이해니, 이승호(2019)는 고등학교 급간에서 혁신고 학생들은 특목고, 자사고 학생들에 비해 낮은 학업성취도를 보였으나, 변화추이를 살펴보았을 때 혁신학교 학생들의 높은 성장으로 인해 학교유형간 성취도 격차가 감소한 것을 밝혔다. 이처럼 학생들의 장기적인 성장률을 분석한 연구들은 혁신학교가 초·중·고등학교 급간에서 학업성취의 변화율에 긍정적 효과가 있다는 공통적인 결과를 보여준다.

나아가 혁신학교의 재학여부가 가정배경에 따른 교육격차를 감소시킨다는 주장을 뒷받침하는 연구들도 수행되었다. 백병부, 박미희(2014)는 중학교와 고등학교 단계에서 혁신학교가 일반학교에 비해 사회경제적 가정배경의 영향력에 따른 학업성취 격차를 감소시키는데 기여하는 것을 보여주었다. 이와 비슷하게 원순아, 엄기형(2018)은 혁신학교 재학여부가 중1 학업성취에 유의미한 영향을 미치는 것은 아니지만, 부의 학력이 낮은 학생의 성적을 향상시켜 학력 차이를 줄인다는 것을 확인하였다. 또한 혁신학교의 학습자중심수업이 국어과 학업성취에 긍정적인 영향이 있음을 발견하였다.

종합했을 때, 선행연구들은 혁신학교의 학업성취가 일반학교에 비해 높은 것은 아니지만 성장률에 긍정적인 영향을 미칠 수도 있음을 시사하고, 사회경제적 배경에 따른 학력 격차를 감소시키거나 특정과목에 차별적 효과가 가질 가능성을 보여주었다. 그러나 기존의 연구들은 다음과 같은 한계를 가지고 있다. 첫째, 선행연구들 중 다수는 선택편의를 일으키는 혁신학교와 일반학교 재학생 집단의 이질성을 제거하지 못하였다. 혁신학교는 열악한 지역에 우선적으로 선정되었기 때문에 일반학교와 비교했을 때 재학생의 사회경제적 배경이 낮고(박상현 외, 2016; 백병부, 박미희, 2014) 학업성취도는 일반학교보다 낮은 경향성(김민규, 박세진, 2019; 서민희, 전경희, 2018)이 있다. 그러나 대다수의 연구들은 혁신학교에 입학하는 학생들의 개인 및 가정 배경 차이에 따른 실험집단과 통제집단의 이질성을 고려하지 못하였으므로 혁신학교의 인과적 효과를 주장하기 어렵다. 경향점수매칭을 통해 선택편의를 통제하고자 한 연구들 역시 혁신학교 진학 전 시점의 자료를 활용하지 못한(윤혜원, 2020; Min et al., 2017) 제한점을 가지고 있다.

둘째, 혁신학교의 학습자중심수업 운영 방식이 기초학습이 부족한 하위권 학생들에게 부정적인 영향을 미칠 것이라는 현장의 불안이 높은 가운데, 학생들의 학업수준에 따른

혁신학교의 차별적인 효과를 구체적으로 주목하지 못하였다. 대다수의 연구(백병부, 박미희, 2015; 원순아, 엄기형, 2018)들이 가정배경에 따른 영향력에 초점을 맞추고 있으며, 기초학력 미달 학생들의 비율을 감소시킬 가능성을 제시한 연구(배종현, 김주후, 2016)가 있었으나 학생 개인을 추적한 것이 아닌 학교단위의 횡단적인 비교에 그쳤다는 한계를 가지고 있었다.

따라서 본 연구에서는 혁신학교 재학생과 일반학교 재학생의 이질성을 통제하고 엄밀한 인과효과 추정을 위해, 경향점수매칭 기법과 회귀분석을 결합한 이중강건추정을 통하여 혁신학교가 중학생 학업성취도에 미치는 영향을 분석하고자 하였다. 더 나아가 분위회귀분석을 활용하여 기존연구에서 주목하지 않았던 학생들의 학력 수준별 혁신학교의 보다 세부적인 효과를 확인하였다.

Ⅲ. 연구 방법

1. 분석 대상과 자료

본 연구에서는 혁신학교가 중학생의 학업성취에 미치는 효과를 분석하기 위해 경기교육종단연구(Gyeonggi Education Panel Study, 이하 GEPS)의 1, 3, 4, 6차년도 자료를 활용하였다. 초등학교 4학년부터 중학교 3학년까지 추적한 초등학교 패널의 경우 혁신중학교 진학 이전의 정보와 혁신중학교 재학 정보를 알 수 있다. 1차년도 자료에서는 부모의 학력, 3차년도 자료에서는 초등학교 학업성취도, 혁신 초등학교 재학 여부 등 경향점수 산출을 위해 사용된 변인들과 학업효능감, 수업태도, 수업집중시간, 교사만족도, 학교만족도 등을 추출하였다. 4차년도 자료에서는 혁신 중학교 입학에 대한 정보를 이용하였고, 6차년도 자료에서는 중학교 3학년의 학업성취도를 추출하였다.

결측치를 제외하고 최종 분석에 포함된 대상은 총 201개 중학교 학생 2192명으로, 혁신중학교 46개교의 학생 417명, 일반중학교 155개교의 학생 1775명이다.

2. 변수 선정

본 연구는 경향점수매칭을 활용하여 처치집단에 대한 선택편의를 제거한 후, 다중회귀분석(multiple regression analysis)을 통해 학업성취도에 대한 혁신학교의 순수 처치효과를 산출하였다. 다중회귀분석에 사용된 종속변수와 통제변수의 종류, 처리 방법 및 각 변수에 대한 기초통계치는 <표 2>와 같다.

<표 2> 회귀분석에 사용한 변수

구분	응답 주체	변수명	사례 수	평균	표준 편차	변수 처리
종속 변수	학생	국어	2831	515.14	93.01	중학교 3학년 국어, 영어, 수학 교과와의 수직적도 점수
		수학	2830	509.41	101.06	
		영어	2831	511.75	98.71	
		평균	2835	511.98	82.77	
처치 변수	학교	중학교 혁신학교 여부	2704	0.18	-	더미변인 (일반=0, 혁신=1)
통제 변수		국어	2819	3.70	0.90	교과효능감 관련 5개 문항의 평균값
		영어	2815	3.50	1.03	
		수학	2809	3.53	1.06	
		평균	2806	3.58	0.87	
	학생	국어	2824	3.85	1.05	교과별 수업 집중 시간 (10분 단위 5단계 척도값)
		영어	2824	3.63	1.18	
		수학	2822	3.60	1.22	
		평균	2822	3.69	0.98	
	수업 태도	국어	2822	3.52	0.86	교과별 수업태도 5개 문항의 평균값
		영어	2821	3.51	1.00	
		수학	2821	3.60	1.03	
		평균	2819	3.54	0.85	
	교사만족도	교사만족도	2826	3.64	0.88	교사만족도 6개 문항의 평균값
		학교만족도	2828	3.37	0.82	학교만족도 8개 문항의 평균값
성별	성별	2840	.51	-	더미변인(여=0, 남=1)	
부모	소득수준(ln)	2759	6.03	0.53	월평균 가구소득의 자연로그 변환값	
	부	2774	3.12	1.13	학력에 따른	
	모	2792	2.85	0.97	6단계 척도값	
학교	초등 혁신학교 재학 경험	2840	.25	-	더미변인 (일반=0, 혁신=1)	
	학교소재지	2667	.88	-	더미변인 (읍면=0, 도시=1)	

연구모형의 종속변수는 학업성취도로, 국어, 영어, 수학 교과점수 및 국영수 평균 점수를 이용하였다. 통제변수로는 혁신 중학교 재학여부를 비롯하여 학업성취도에 영향을 미치는 학생, 부모, 학교 특성관련 변수를 투입하였다. 성별은 여학생을 0, 남학생을 1로

코딩하였고 소득수준은 월평균 가구소득을 자연로그로 변환하여 사용하였다. 부모의 학력은 중졸이하는 1, 고졸은 2, 2~3년제 대졸은 3, 4년제 대졸은 4, 석사 수료 및 졸업은 5, 박사 수료 및 졸업은 6으로 코딩되었고, 부와 모의 학력을 각각 사용하였다. 학업효능감, 수업태도, 교사만족도, 학교만족도는 5점 척도로 된 관련 문항들의 평균값을 분석에 사용하였다¹⁾. 수업집중은 각 과목별로 수업 시간을 10분 간격으로 나누어 5단계 척도로 나타냈다. 학업효능감, 수업 집중, 수업 태도는 종속변수에 따라 각 과목별로 구분하여 사용되었고, 국영수 평균점수 분석에 대해서는 전체 평균값을 사용하였다. 혁신학교 여부는 일반학교를 0, 혁신학교를 1로, 학교소재지는 읍면지역을 0, 도시지역을 1로 코딩하였다.

3. 분석 모형

본 연구는 준실험설계 방식인 경향점수매칭(Propensity Score Matching; PSM) 기법을 활용하여, 혁신 중학교에 진학한 처치집단과 일반 중학교에 진학한 통제집단에 미치는 변인들의 영향과 선택편의를 최소화하고자 하였다. 이후 경향점수를 도출하는데 활용한 변수들 중 일부와 새로운 변수들을 통제변수로 재투입하여 통상최소자승법(Ordinary Least Squares; OLS)으로 다중회귀분석을 실시하고 학교유형에 따른 학업성취도의 효과가 매칭 전후로 어떤 차이가 있는지 살펴보았다. 이러한 경향점수와 회귀분석을 결합하여 이중으로 통제변인을 모형에 투입하여 처치효과를 추정하는 이중강건추정 방식은 각각의 기법만을 활용하는 전통적인 분석에 비해 모형 설정오류(misspecification)에 의한 편향을 줄일 수 있다는 장점을 가진다(Funk et al., 2011). 마지막으로, 학생들의 학력수준에 따른 혁신학교의 차별효과를 분석하기 위하여 분위회귀분석(Quantile Regression Analysis)을 활용하였다. 분석을 위해 STATA 16.0과 R 프로그램을 이용하였다.

가. 경향점수매칭(PSM)

경기도의 경우 중학교 진학은 무시험 추첨 배정의 원칙을 따르지만, 거주 지역 내 희망하는 중학교에 우선 지원 신청을 할 수 있도록 되어있다. 즉, 거주지 내에 혁신 중학교가 존재한다면 혁신학교에 지원한 후 추첨에 따라서 진학여부가 결정되게 되는 것이다. 혁신학교 진학여부가 학부모 및 학생의 선택에 의해서 완전히 좌우되는 것은 아니

1) 교사 및 학교만족도는 학생이 느끼는 교사와의 관계 및 학교에 대한 만족도를 의미함. 구인에 대한 문항들의 신뢰도를 나타내는 Cronbach's α 값은, 학업효능감 5문항은 .94(국어), .95(영어), .95(수학), 수업태도 5문항은 .84(국어), .88(영어), .89(수학), 교사만족도 6문항은 .93, 학교만족도 8문항은 .93임.

지만, 학생과 학부모의 혁신학교 진학 희망 여부에 어느 정도 영향을 받는다는 것이다. 혁신학교 진학은 완벽한 무작위 배정이 아니기 때문에 학부모 및 학생들의 선택편의가 존재하여 관련한 차이를 통제하지 않고 효과를 추정한다면 추정의 타당성이 저해된다. 이처럼, 처치집단과 통제집단을 나누는 방식이 무작위배정이 아니면서 동시에 뚜렷한 배정규칙이 없는 경우를 비동질적 통제집단 설계(non-equivalent control group design) 상황이라 하고, 배정과 관련된 변수에 대한 가정을 가지고 매칭을 할 수 있다(노화준, 2015). 정책 투입 이전 자료를 이용하여 변수들을 통제하고 처치집단과 통제집단을 동질적으로 구성하기 위해 일반화된 선형모형(generalized linear model)을 활용하여 처치집단에 속할 확률을 경향점수로 산출한 뒤, 처치집단과 통제집단에서 각각 비슷한 점수를 얻은 사례들을 매칭하고 이를 가중치로 활용하여 선택편의를 보정할 수 있다(Rosenbaum & Rubin, 1983). 본 연구에서는 로지스틱 회귀분석을 이용하여 응답자의 경향점수를 산출하였으며 그 모형은 식 (1)과 같다.

$$\ln\left(\frac{P_i}{1-P_i}\right) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k \quad (1)$$

나. 통상최소자승 회귀모형(OLS)

혁신 중학교에 진학한 학생들과 유사한 특성을 지녔으나 일반 중학교에 진학한 학생들을 경향점수매칭 방법으로 선정한 후, 식 (2)와 같이 OLS 회귀분석을 적용하여 혁신학교 효과를 분석하였다.

$$Y = \beta_0 + \beta_t T + \sum \beta_k X_k + \epsilon \quad (2)$$

학생의 학업성취도 점수는 종속변수(Y)를, 혁신학교 재학 여부(T)와 그 외 배경 특성(X)은 독립변수를, ϵ 는 오차항을 구성한다.

다. 분위회귀분석(Quantile Regression)

본 연구는 혁신학교 재학여부가 학업성취도에 미치는 분위별 처치효과(quantile treatment effects)를 분석하기 위하여 분위회귀분석을 활용하였다. 일반적인 회귀분석이 처치효과에의 평균적인 경향성만을 드러낸다면, 분위회귀분석은 종속변인의 분위별로 각각의 회귀계수를 추정하여 처치변인이 어떤 수준에서 더 효과적으로 작용하는지 보여준

다. 아래의 식 (3)은 분위회귀모형을 나타낸다(Koenker & Bassett, 1978).

$$Q_{\tau}(y_i) = \beta_{0(\tau)} + \beta_{1(\tau)}x_{i1} + \cdots + \beta_{p(\tau)}x_{ip}, \quad i = 1, 2, \dots, n \quad (3)$$

식 (3)에서 $0 < \tau < 1$ 의 범위를 가지며, τ 번째 분위에 해당하는 회귀계수를 도출한다. 종속변수(y_i)인 학업성취도 하위 25%, 중위 50%, 상위 25% 분위별 학생에 대해 처치변인인 혁신학교 재학 여부가 각각 어떤 효과를 미치는지를 분위회귀계수($\beta_{(\tau)}$) 값을 통해 추정하였다.

IV. 연구 결과

1. 경향점수 매칭 전 OLS 결과

<표 3>은 경향점수 매칭 전, 혁신학교가 학업성취도에 미치는 영향을 OLS 방법으로 분석한 결과이다. 공변인들을 통제했을 때, 국영수 평균과 국어 과목에서 혁신학교가 일반학교에 비해 각각 8.77점, 8.91점이 유의하게 향상하는 것으로 나타났다. 성별에 따른 학업성취도의 차이를 보면, 수학 과목을 제외하고 여학생들의 성취수준이 유의하게 높은 것으로 드러났다. 전과목에 걸쳐 부학력이 높아질수록 학업성취도는 유의하게 증가하고, 모학력은 영어 교과에서 유의한 영향을 미치는 것으로 드러났다. 도시에 있는 학생들의 수학 과목 성취도는 읍면 지역 학생들보다 높았으나 그 외 과목은 차이가 없었다. 교사에 대한 만족도는 학업 성취에 부적인 영향이 있었고, 학생들의 수업 집중 시간과 학업효능감은 학업 성취에 긍정적으로 작용하였다.

<표 3> 경향점수매칭 전의 OLS 분석 결과

독립변수	종속변수			
	국영수 평균	국어	수학	영어
혁신학교	8.765** (3.168)	8.912* (4.151)	7.921 (4.316)	6.955 (4.068)
성별 (남자)	-19.421*** (2.476)	-43.662*** (3.244)	0.199 (3.368)	-27.347*** (3.170)
부학력	8.348*** (1.459)	8.864*** (1.898)	9.582*** (1.979)	12.160*** (1.878)

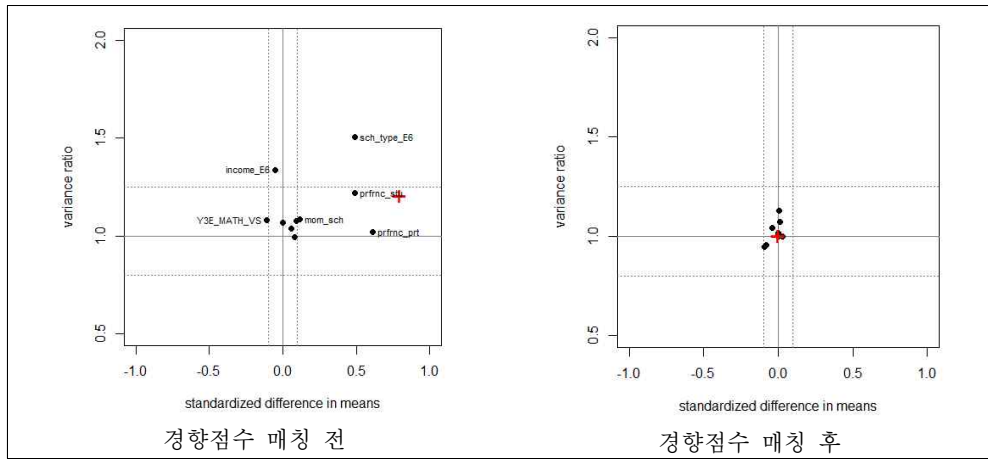
독립변수	종속변수			
	국영수 평균	국어	수학	영어
모학력	1.755 (1.669)	0.605 (2.189)	3.910 (2.272)	4.515* (2.150)
소득	-0.001 (0.003)	-0.006 (0.004)	0.009* (0.005)	-0.004 (0.004)
교과성적(초6)	0.562*** (0.019)	0.384*** (0.023)	0.370*** (0.017)	0.421*** (0.021)
교사만족도	-4.846** (1.641)	-4.894* (2.167)	-3.663 (2.189)	-5.361** (2.092)
학교소재지 (도시)	6.207 (3.755)	5.280 (4.906)	13.646** (5.115)	7.273 (4.822)
학교만족도	2.459 (1.753)	7.466** (2.282)	-2.227 (2.385)	2.408 (2.234)
수업집중시간	10.931*** (1.727)	10.426*** (1.934)	5.379** (1.901)	9.619*** (1.837)
수업태도	0.402 (2.246)	-2.665 (2.541)	4.069 (2.572)	4.011 (2.416)
학업효능감	16.453*** (2.002)	15.416*** (2.277)	22.858*** (2.197)	16.413*** (2.093)
혁신학교 재학경험(초6)	1.733 (2.900)	-0.245 (3.803)	-2.950 (3.958)	7.289 (3.730)
constant	140.287*** (10.161)	251.716*** (13.180)	175.284*** (12.128)	185.534*** (11.747)

*p < .05, **p < .01, ***p < .001 (괄호 안은 표준오차를 나타냄)

2. 경향점수를 이용한 혁신학교가 학업성취도에 미치는 영향 분석

경향점수 매칭을 위해 통제변수들이 혁신 중학교의 재학여부에 영향을 받지 않도록 GEPS의 3차년도(초등학교 6학년) 데이터를 사용하였다. 혁신 중학교에 재학할 확률인 경향점수 산출을 위해 로지스틱(logistic) 회귀분석에서 사용한 변인은 부학력, 모학력, 가구 소득 수준, 부모와 학생의 혁신학교 진학 희망 여부, 초등학교 6학년의 국어, 수학, 영어 성적, 초등학교 때 혁신학교 재학 여부, 성별 등 10개의 변수들이다. 투입된 공변인의 영향력이 처치집단과 통제집단에 대해 잘 통제되었는지 균형(balance)을 확인하기 위해 표준화된 평균 차이(standardized mean difference)와 분산 비(variance ratio) 방법을 사용하였다(Rubin, 2001). 표준화된 평균 차이 값은 각각의 공변인이 처치집단과 통제집단에서 갖는 평균값의 차이에서 통합된(pooled) 표준편차를 나누어준 값으로서, 절대값이 0.1보다 작으면 균형이 잘 이루어진 것으로 볼 수 있다. 처치집단과 통제집단에

서 각각의 공변인이 갖는 분산의 비를 계산한 분산 비는 1에 가까울수록 균형이 잘 이루어진 것이며 0.8에서 1.25 값을 가지면 수용할 만한 범위라고 할 수 있다(Steiner & Cook, 2013).



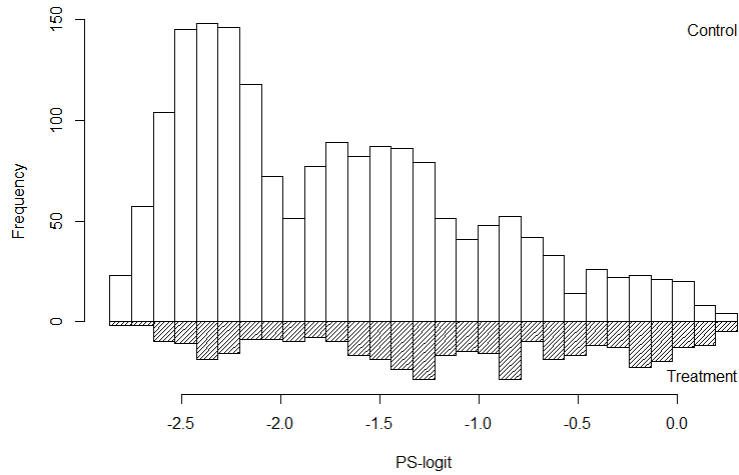
[그림 1] 경향점수 매칭 전후 공변인의 균형

[그림 1]의 좌측 패널은 경향점수 매칭 전 표준화된 평균 차이와 분산 비에 따른 각각의 공변인의 위치를 나타낸 것이다. 대부분의 공변인들이 균형의 기준인 점선 범위 안에 들어가 있으나 대체적으로 흩어진 양상을 보이고 있으며, 처치그룹과 통제그룹 별로 차이가 큰 변인은 초등학교의 혁신학교 재학 여부, 학생의 혁신학교 진학 희망 여부, 부모의 혁신학교 진학 희망 여부, 부모의 소득 수준 등이었다. [그림 1]의 오른쪽 패널은 매칭 후의 위치를 보여주며, 표준화된 평균 차이는 0에 가깝고 분산 비는 1에 가까워져 그룹 별 공변인의 균형이 잘 이루어진 것을 확인할 수 있다.

다음으로, 경향점수의 공통적인 범위에 처치그룹과 통제그룹이 겹쳐서 분포하고 있는지를 살펴보고, 겹치지 않는 범위의 사례들을 제거하고 처치효과 해석에 유의하기 위해 각 집단의 분포 오버랩(overlap)을 확인하였다. [그림 2]에 나타난 것처럼 로짓(logit) 척도로 나타낸 경향 점수의 전 범위에 걸쳐 두 집단의 사례들이 공통적으로 분포하고 있고, 일반학교 학생들의 사례수가 더 많은 것을 확인할 수 있다. [그림 1]과 [그림 2]를 통해 경향점수 매칭이 두 집단에 대한 공변인의 영향력을 통제하여 준실험설계 목적에 맞는 균형이 이루어졌음을 확인하였다.

로지스틱 회귀분석을 통해 산출한 경향점수를 활용하여 역확률 가중치(inverse probability treatment weighting; IPTW)를 계산하였고, 이를 통해 처치집단과 통제집단의 표본 차이를 보정하여 선택편의를 제거한 후 분석이 이루어지도록 하였다(Austin &

Stuart, 2015).



[그림 2] 혁신학교와 일반학교 학생들의 경향점수에 따른 분포

<표 4> 경향점수매칭 후의 OLS 분석 결과

독립변수	종속변수			
	국영수 평균	국어	수학	영어
혁신학교	3.538 (3.424)	5.301 (4.616)	3.812 (4.672)	1.461 (4.744)
성별 (남자)	-19.743*** (2.733)	-42.872*** (3.630)	-1.218 (3.548)	-26.547*** (3.519)
부학력	8.739*** (1.507)	8.309*** (1.999)	10.094*** (2.008)	13.221*** (1.988)
모학력	1.524 (1.709)	1.032 (2.299)	3.547 (2.348)	4.109 (2.316)
소득	-0.001 (0.004)	-0.005 (0.004)	0.010 (0.006)	-0.003 (0.006)
교과성적(초6)	0.558*** (0.023)	0.393*** (0.025)	0.366*** (0.020)	0.419*** (0.022)
교사만족도	-3.838* (1.852)	-3.910 (2.403)	-2.948 (2.477)	-4.566 (2.380)
학교소재지 (도시)	3.707 (3.609)	2.787 (4.932)	11.136* (5.197)	3.850 (4.801)

독립변수	종속변수			
	국영수 평균	국어	수학	영어
학교만족도	1.786 (1.861)	6.172* (2.512)	-1.815 (2.616)	1.454 (2.432)
수업집중시간	12.022*** (2.060)	11.636*** (2.270)	5.178* (2.110)	11.578*** (2.068)
수업태도	-0.420 (2.852)	-4.733 (2.993)	4.213 (2.890)	3.122 (2.779)
학업효능감	15.612*** (2.505)	15.412*** (2.813)	22.768*** (2.540)	14.873*** (2.338)
혁신학교 재학경험(초6)	1.172 (3.338)	-1.380 (4.376)	-2.584 (4.159)	5.849 (4.180)
constant	144.587*** (10.133)	253.770*** (14.587)	177.012*** (12.319)	189.247*** (11.580)

* $p < .05$, ** $p < .01$, *** $p < .001$ (괄호 안은 표준오차를 나타냄)

<표 4>는 PSM 매칭 후 학업성취도에 대한 처치변인과 그 외 공변인의 영향을 나타낸 결과이다. 국어, 영어, 수학 및 국영수 평균 모두에서 혁신학교의 긍정적 효과는 줄어들었으며, 특히 매칭전 유의미했던 국어 과목과 국영수 평균에서의 효과가 사라졌다. 즉, 매칭 후에도 혁신학교가 일반학교에 비해 학업성취도가 다소 높게 나타나기는 하지만 그 차이는 통계적으로 유의하지 않았다. 성별에 따른 차이는 매칭 전과 마찬가지로 수학 과목을 제외하고 여학생의 성취도가 유의미하게 높았다. 과거 교과성적, 수업 집중시간, 그리고 학업효능감은 매칭 후에도 여전히 학업성취도에 긍정적으로 영향을 미치고 있었다. 매칭 후에도 교사에 대한 만족도가 클수록 학업성적은 낮아지는 경향이 있었으나 각 과목에서 그 영향력의 유의성은 없었다.

3. 학업성취 수준별 혁신학교의 영향 분석

<표 5>는 혁신학교가 학업성취도에 미치는 분위효과를 나타낸다. 경향점수매칭을 통해 본 국어, 수학, 영어 및 국영수 평균 점수에 대한 혁신학교 효과를 교과 성적 하위 25분위(25Q), 중위수(50Q), 상위 25분위(Q75)에서 어떠한 차이가 있는지 살펴보았다. 우선, 혁신학교의 국어 과목 학업성취도에 대한 영향을 보면, 하위 수준의 학생들에서 유의미한 도움이 되는 것으로 나타났다. 즉, 국어 성적이 하위 25%에 해당하는 학생들에게서 혁신학교의 효과가 12.63으로 추정되어, 일반학교에 비해 유의미한 성적 향상이 있었다. 수학과 영어 과목의 하위 25% 학생들에게 혁신학교의 성적향상 효과는 통계적으로 유의미하지 않았다. 중간 수준과 상위 수준의 학생들에서는 전 과목에 걸쳐 혁신학

교의 효과가 통계적으로 유의하지 않았다.

<표 5> 혁신학교가 학업성취도에 미치는 분위효과

		종속변수			
		국영수 평균	국어	수학	영어
혁신학교 효과추정치 (S.E)	Q25	7.433 (4.080)	12.627* (5.342)	15.283 (8.090)	2.929 (4.959)
	Q50	.060 (3.286)	2.982 (6.059)	4.086 (4.830)	-.308 (5.167)
	Q75	-3.280 (3.993)	-.331 (5.187)	1.834 (4.560)	-.148 (4.140)

$p < .1$, $*p < .05$ (괄호 안은 표준오차를 나타냄); 처치변인 외 공변인 생략

V. 논의 및 결론

본 연구는 GEPS 자료를 활용하여 혁신학교가 중학교 학업성취에 미치는 영향을 분석하였다. 혁신학교와 일반학교 진학 학생 간의 이질성을 통제한 경향점수매칭 기법으로 엄밀한 인과성을 추구하고, 분위회귀분석으로 학업성취 수준별 집단에 따른 차별적 효과를 추정하였다. 본 연구의 연구 결과를 논의하자면 다음과 같다.

첫째, 경향점수매칭 기법을 활용하기 전 OLS 분석 결과, 일반학교보다 혁신학교 학생들의 중학교 3학년 국영수 평균 점수가 더 높은 경향성을 나타냈으나, 매칭 후에는 혁신학교와 일반학교의 학업성취는 통계적으로 유의미한 차이가 없는 것으로 나타났다. 즉, 입학 전 두 집단의 이질성과 학업성취에 영향을 미치는 공변인을 통제하였을 때 혁신학교의 긍정적 효과는 사라진다. 다양한 급간에서 혁신학교와 일반학교의 이질성을 통제하지 않은 기존의 연구(김민규, 박세진, 2019; 박세진, 이해니, 이승호, 2019; 서민희, 전경희, 2018; 안영은, 박세진, 2019)에서는 혁신학교의 긍정적인 효과를 발견하였으나, 본 연구와 마찬가지로 이질성을 통제한 준실험설계를 도입한 선행 논문들(박상현, 2016; 윤혜원, 2020; Min et al., 2017)에서 혁신학교 효과를 발견하지 못했다는 점은 혁신학교의 교육 효과 탐구에 세심한 주의가 필요함을 시사한다.

혁신학교가 학업성취의 성장률에 정적 효과를 준다고 밝힌 연구들은 모두 두 집단의 초기 이질성을 통제하지 않는 연구들이었다. 이는 초기 시점에 대한 두 집단의 이질성을 통제하지 못한 선행연구들이 효과에 대한 과대추정을 했을 가능성을 의미한다. 특히, 중학교 급간에서 국영수 평균 학업성취를 가지고 잠재성장모형을 활용한 연구 결과들

(김민규, 박세진, 2019; 안영은, 박세진, 2019)에서 혁신학교 성장률이 일반 중학교 학생들보다 높게 나타난 것은 이처럼 두 집단의 초기 이질성을 고려하지 않고 공변인을 통제하지 않았기 때문이라고 추정할 수 있다. 따라서 본 연구는 혁신학교 효과의 엄밀한 추정을 위해 두 집단을 초기 이질적 특성을 통제하는 준실험연구가 중요하다는 점을 시사한다.

둘째, 본 연구의 분석 결과는 중학교 급간에서 혁신학교가 학업성취에 긍정적인 영향을 미치지 못하는 것만 일반학교에 비해 학력저하가 있는 것은 아니라는 기존의 연구결과를 지지한다(박상현, 2016; 배종현, 김주후, 2016). 이러한 분석 결과를 바탕으로 한 대부분의 연구자들은 두 집단의 차이가 없기 때문에 혁신학교의 학력저하가 근거 없는 주장이라고 결론 내리고 있다. 그러나 통계 결과의 이면에 존재하는 실재를 심층적으로 이해하기 위해서는 혁신학교가 운영되는 현장을 보다 세밀히 살펴볼 필요가 있다.

혁신학교가 학생들의 학력에 긍정적 영향을 미치거나, 반대로 기초학력저하와 관련성이 있을 것이라고 추정되는 핵심요소는 학습자중심의 수업 방식이다. 본 연구에서도 t-검증을 통해 분석한 결과, 수학에는 유의한 차이가 없었지만 국어와 영어교과에서 일반학교에 비해 혁신학교가 학습자중심 수업을 더 많이 실시하는 것으로 나타났다²⁾. 그러나 혁신학교 운영에 대한 질적 연구들을 살펴보면, 혁신학교가 오롯이 학습자중심교육이라는 가치에 부합하는 방향으로 운영되기보다는 학력저하담론과 입시에 대응하기 위해 노력하고 있다는 것을 엿볼 수 있다(김수경, 2011; 유경훈, 2014; 정혜진, 조영하, 2020). 공교육 체제에서 요구하는 기본적인 교육과정 범위를 모두 가르치는 것을 상대적으로 중요하게 생각하지 않고 혁신학교의 가치를 고수하는 교사들도 있지만, 상급학교 진학을 대비해야 한다는 현실과제 앞에서 학력저하를 우려하는 일부 교사들은 이를 보충하고자 애를 쓰기도 하는 것이다. 예를 들어, 혁신학교 철학을 가진 ‘소신 있는’ 교사들이 학습 진도를 많이 나가지 못하는 바람에 상급학년의 동료교사들이 수업결손을 보충하기 위한 부담을 가지고, 학생들은 지난 진도와 관련해서 시험에 나올 문제를 ‘짚어달라’고 하는 등의 현상이 포착된다(정혜진, 조영하, 2020).

더욱이 혁신학교 자체적으로도 ‘학력향상’에 초점을 맞추거나, ‘개인별 스펙 쌓기’나 체험식 동아리를 지원하면서 입시에 유리한 입학사정관제에 적극 대비하거나(김수경, 2011), 재학 중인 학생들이 학교에서 따로 공부를 하거나 사교육을 찾는 등 대비하는 모

2) 국어, 영어, 수학 각 과목별로 학생들이 학습자중심의 수업이 이루어진다고 인식하는 정도를 측정된 각 15개의 문항의 평균을 비교하였음. 국어교과에서의 평균값은 일반학교 3.78(sd=.84), 혁신학교 3.83(sd=.76)으로 혁신학교에서 유의미하게 높게 나타남($t = -1.97$, $p = .049$). 영어교과의 평균값은 일반학교 3.64(sd=.89), 혁신학교 3.71(sd=.81)로 혁신학교에서 유의미하게 높게 나타남($t = -2.25$, $p = .025$). 수학교과와 경우, 평균값은 일반학교 3.71(sd=.88), 혁신학교 3.74(sd=.83)로 혁신학교의 평균값이 높았으나 그 차이가 유의하지 않았음($t = -.93$, $p = .35$).

습도 나타난다(유경훈, 2014). 혁신학교가 일반학교에 비해 학업성취 면에서 차이가 없는 결과가 학력저하 우려에 대한 다양한 대응이 존재하기 때문이라고 추측할 수도 있다. 이처럼 혁신학교는 학교 및 수업 운영 원리에 있어 일반학교와는 다른 철학을 가지고 있으나, 실제 운영되고 있는 각 요소들이 혁신학교의 효과와 어떠한 관련성을 맺고 있는지에 관해서는 여전히 베일에 싸인 “블랙박스(black box)”의 상태로 남아있다. 따라서 혁신학교의 핵심 운영원리인 학습자중심 수업 방식을 고수하는 혁신학교가 학력저하와 무관하다고 주장하기 위해서는 실제 운영 현장에 대한 면밀한 분석이 뒷받침되어야만 한다.

셋째, 혁신학교가 학업성취의 수준별 집단에 미치는 영향을 분석한 결과, 국어교과 하위 25%에 속하는 학생들에게 일반학교에 비해 유의미한 학업성취 향상의 효과가 있는 것으로 나타났다. 그러나 수학과 영어 및 국영수 평균 점수에는 효과가 없었다. 수학, 영어 교과의 하위권 학생들의 학업성취에도 혁신학교와 일반학교의 차이가 없으므로, 적어도 혁신학교가 기초 학업수준이 낮은 하위권 학생들에게 부정적인 영향을 미치지 않는다는 것을 보여준다. 오히려 국어과의 하위권 학생들에게는 혁신학교가 긍정적인 영향을 미치고 있었다. 이는 국어교과에서는 기초학력 미달 학생의 비율이 일반학교에 비해 감소했다는 기존의 연구 결과(배중현, 김주후, 2016)와도 부합한다. 이와 유사하게 한 학교의 사례이긴 하나, 혁신 초등학교에서 국어과는 기초학력 부진이 감소한 반면, 수학과 영어교과에서 기초학력 부진이 증가했다는 결과(김수경, 2011) 역시 유사한 맥락을 제공한다.

혁신학교가 타 과목보다 특히 국어과목 하위권 학생들에게 긍정적 효과를 보인 것은 과목의 특성에서 유추할 수 있다. 국어교과에 비해, 수학과 외국어와 같은 영어 과목은 일상생활에서 쉽게 학습의 기회를 얻기 어렵기 때문에, 학교가 아닌 가정의 사회경제적 배경의 영향을 더욱 받는 것으로 알려져 있다(이영자, 차윤경, 1996; 최선헌, 2003; Cooper, Nye, Charlton, Lindsay, & Greathouse, 1996). 실제로 개인 혹은 소집단 프로젝트 방식의 수업이 국어과에서는 사회경제적 배경의 영향력을 감소시키지만, 영어의 경우 영향력을 더 증가시키는 것으로 나타나기도 한다(백병부, 2012).

이러한 특성을 고려했을 때, 수학과 영어는 기초적 개념이나 지식이 부족한 학생들에게 개인의 활동을 중시하는 학습자중심의 교육이 큰 영향력을 발휘하기 어려울 수 있다. 반면, 학생들 사이의 토론과 협업과 같은 상호작용 활동을 강조하는 학습자중심수업의 특성이 상대적으로 타 교과에 비해 국어 하위권 학생들의 성적에 긍정적인 영향을 미쳤을 가능성이 있다. 원순아, 엄기형(2018)의 연구는 국어과에서만 혁신학교의 학습자중심수업이 학업성취에 유의미한 정적 효과를 보이는 것으로 나타나고 있어 이러한 해석을 지지한다. 이처럼 본 연구는 혁신학교의 과목별, 학업성취수준 집단별 차별적 효과

를 분석함으로써 앞에서 제기한 혁신학교 효과와 관련한 “블랙박스”의 세부적인 요소들을 이해할 수 있는 단서를 제공한다는 점에서 큰 의의를 갖는다.

본 연구의 논의를 바탕으로 시사점을 제언하면 다음과 같다. 첫째, 혁신학교의 국어과 수업에서 활용되는 학습자중심수업 방식의 장점을 세밀히 분석하고 더욱 적극적으로 추진할 필요가 있다. 특히 본 연구를 통해 국어교과 하위권 학생들에게 효과가 있음이 확인되었으므로 혁신학교 뿐 아니라 일반학교에서도 학습자중심수업의 활용이 촉진될 가치가 있다. 반면, 혁신학교가 국어, 영어, 수학 교과에서 모두 평균적 효과가 없었고, 학업성취 수준별 집단에서도 수학과 영어 과목에서 모두 일반학교와 차별적인 효과가 없었다는 점에서 혁신학교의 운영방식이 가지는 효과에 대한 현장에서의 보다 면밀한 분석이 요구된다.

둘째, 혁신학교의 정책 효과에 관한 과대추정을 배제하고 보다 엄밀한 추정을 위해서는 본 연구와 같은 실험집단과 처치집단의 이질성을 균형적으로 통제할 수 있는 준실험설계를 활용하는 것이 중요하다. 특히 혁신학교에 입학하는 학생들과 일반학교 학생들의 배경적 특성이 다르기 때문에 초기의 선택편의를 제거할 필요가 있다. 따라서 혁신학교의 정확한 효과를 분석하기 위해 본 연구와 같은 준실험설계 기법을 활용한 연구가 적극적으로 이루어질 필요가 있다.

셋째, 혁신학교 정책을 더욱 효과적으로 시행하기 위해서 정책 효과가 미치는 대상과 영역의 특성을 보다 세부적으로 고려해야 한다. 본 연구는 분위회귀분석 기법을 활용하여 혁신학교가 국어과목의 하위권 학생들에게 차별적 효과가 있음을 발견하였다. 선행연구들을 종합했을 때 이러한 결과는 혁신학교가 초·중·고등학교에 다른 효과를 가질 뿐 아니라, 학생들의 가정배경, 학업능력수준, 그리고 각 과목별로도 다른 영향을 미칠 수 있다는 것을 보여준다. 따라서 대상급간, 과목, 학력수준, 성취 영역 등에서 혁신학교의 운영 요소들과 관련한 정교하고 세분화된 분석이 활발히 이루어져야 한다. 이를 통해서, 혁신학교라는 “블랙박스” 안에서 일어나고 있는 과정을 이해하고 지속적인 수정·보완과 지원을 위한 방향 설정이 가능할 것이다.

Slavin(1989)은 새로운 교육정책이 빠른 속도로 확산되었다가 엄밀한 검증이 이루어지기 전에 이내 관심이 다른 곳으로 옮겨가버리는 교육계의 유행주의(faddism)에 대해 경고하였다. 혁신학교 역시 교육개혁에 대한 현장의 열망에서 출발하여 과학적 검증이 수행되기도 전에 전국으로 급격하게 확산되었다. 그 과정에서 학력저하에 대한 학부모와 교사들의 불안감이 커지는 것이 당연하다. 이러한 교육 현장의 필연적 요구를 혁신학교를 지지하는 입장에서 진보적 학교개혁을 저지하려는 정치적 이념의 공격으로만 바라보는 것은, 오히려 혁신학교 정책의 기반을 약화시키고 향후 지속적인 정착을 위한 걸림돌로 작용할 것이다. 따라서 혁신학교 모델을 정착시키기 위해서는 새롭고 혁신적

인 정책에 대한 높은 만족도의 함정에 빠지는 위험성을 경계하고, 증거기반(evidence-based)의 논의를 활성화해 나가야만 할 것이다.

본 연구는 준실험설계를 통하여 혁신학교가 중학교 급간에서 적어도 학업성취에 부정적인 영향을 미치지 않는 것이며, 오히려 국어 하위권 학생들에게 긍정적인 영향을 미치는 것을 보여주었다. 이러한 결과는 혁신학교가 학력을 떨어뜨린다고 가해지는 일방적 비판이 타당하지 않다는 것을 지지한다. 그러나 한 가지 간과해서 안 되는 것은 혁신학교와 일반학교의 재정지원의 차이다. 경기도 혁신 중학교의 경우 1인당 평균 목적사업비가 일반 중학교의 42% 가량을 더 지원받고 있으며, 최상위 목적사업비를 받는 혁신학교와 최하위 일반학교 양극을 비교하면 약 7.6배의 차이가 난다(베리타스 알파, 2015.4.3.). 일반적으로 학생 1인당 투자비용이 높아질수록 학생의 학업성취는 높아지는 것으로 알려져 있다(Greenwald, Hedges, & Laine, 1996). 혁신학교가 일반학교에 비해 상당히 높은 재정적 투자를 지원받는다라는 것을 고려할 때, 학업성취도의 효과가 없다는 것은 정책의 투자대비 효과에 심각한 의문을 제기할 수 있다. 따라서 후속연구에서는 혁신학교의 효과를 정책적 관점에서 더욱 객관적으로 바라보기 위해서 재정투자 대비 효과를 더욱 엄밀하게 분석할 필요가 있다.

또한 현재 전국으로 확산되어 시행되고 있는 혁신학교라는 교육 개혁모델에 대한 사회적 불신을 완화시키기 위해서는 효과성 연구와 더불어 혁신학교 운영 현장의 “블랙박스”를 면밀하게 분석하는데 많은 노력을 기울여야 한다. 앞에서 논의한 바와 같이 기존의 연구들은 혁신학교가 교사들의 자율적인 논의에 의해 운영되기 때문에, 같은 기본원리를 추구하는 혁신학교 간 혹은 내에도 교육과정이나 수업 운영 방향이 서로 다른 양상으로 나타날 가능성을 제시하고 있다. 뿐만 아니라 본 연구의 결과에서 하위권 학생들이 혁신학교에 더 긍정적인 영향을 받을 수 있음이 나타난 것처럼, 수혜대상인 학생 및 학부모들 역시 혁신학교의 운영과 수업 방식에 서로 다른 영향을 받거나 대응을 하는 모습도 포착되고 있다. 따라서 혁신학교에서 운영되고 있는 교육과정과 수업 운영방식이 어떠한 양상으로 나타나고 있는지, 서로 다른 배경을 가진 학생 및 학부모가 혁신학교라는 교육모델을 구체적으로 어떻게 해석하고 영향을 받고 있는지 등 실행과정을 이해할 수 있는 체계적인 후속연구의 설계가 적극적으로 이루어져야 할 것이다.

참고문헌

- 경기도교육청. (2020). 2020 혁신학교 운영기본계획.
- 김민규, 박세진. (2019). 혁신학교가 학업성취도 향상에 미치는 영향. **교육행정학연구**, 37(2), 149-174.
- 김수경. (2011). 혁신학교 운영의 실태와 성과 분석. **교육문화연구**, 19(4), 145-168
- 나민주, 박소영. (2013). 자율학교 성과분석 연구: 혁신학교모형을 중심으로. 한국교육개발원.
- 노화준. (2015). **정책평가론** (제 5판). 법문사.
- 동아일보. (2019. 6. 30.). 서울형 혁신학교 '넘비(NIMBY)' 논란 "학부모 반발 심한 강남보다 예산 목마른 강북 최고 13배".
<https://shindonga.donga.com/3/all/13/1771338/1> 에서 2020. 08. 11 인출.
- 박상현, 김보은, 정주영, 손동빈. (2016). 서울형혁신학교의 중단적 효과 분석: 새로운 학력 관점의 학생역량을 중심으로. 서울교육정책연구소.
- 배중현, 김주후. (2016). 혁신학교와 일반학교의 학업성취도 비교 분석: 경기도 중학교의 국어와 영어교과를 중심으로. **교육연구논총**, 37(1), 27-56.
- 백병부. (2012). 가정배경에 따른 학업성취 격차와 수행평가의 관계. **교육사회학연구**, 22(3), 111-132.
- 백병부, 박미희. (2015). 혁신학교가 교육격차 감소에 미치는 효과. **교육사회학연구**, 25(1), 105-128.
- 백병부, 성열관, 양성관(2019). 혁신학교의 지속가능성 위협 요인 분석: 혁신학교 확산기 경기도의 경험을 중심으로. **한국교육**, 46(3), 99-128.
- 베리타스 알파. (2015. 4. 3). 도넛은 혁신학교 몰아주기..'44배까지 지원금'.
<http://www.veritas-a.com/news/articleView.html?idxno=39125> 에서 2020. 08. 11 인출.
- 안혜정. (2019). 혁신학교와 학력: 「학교교육 제 4의 길」과 「사토마나부, 학교개혁을 말하다」의 교육개혁과 학력 논의를 중심으로. **교육비평**, 44, 208-240.
- 연합일보. (2019. 10. 21.). 경기혁신학교 10년...'만족도 높지만, 기초학력 저하 우려'.
<https://www.yna.co.kr/view/AKR20191021118500061> 에서 2020. 08. 11 인출.
- 원순아, 엄기형. (2018). 혁신학교의 교육효과 분석: 경기도 혁신학교의 계층 간 학력격차를 중심으로. **교육사회학연구**, 28(4), 57-86.
- 유경훈. (2014). 혁신 고등학교 운영과정의 특징에 관한 문화기술적 사례연구: 양가성(ambivalence)을 중심으로. **교육행정학연구**, 32(4), 229-261.
- 윤혜원. (2020). 혁신학교가 고등학생의 학업성취도에 미치는 영향: 경향점수 매칭 활용. **한국교육학연구**, 26(1), 45-73

- 이영자, 차윤경. (1996). 가정의 사회·경제·문화적 배경과 학교의 영어교육환경이 초등 학교 학생의 영어구사능력에 미치는 영향. **응용언어학**, **12**, 277-295.
- 정혜진, 조영하. (2020). 서울형 혁신학교에서 민주적 학교운영에 의한 혁신은 실현되고 있는가: A중학교 교사들의 경험에 관한 질적사례연구. **교육행정학연구**, **38**(1), 79-108.
- 조선일보. (2017. 10. 12.). 혁신학교, 기초학력 미달 학생 3배 많다.
https://news.chosun.com/site/data/html_dir/2017/10/12/2017101200053.html
 에서 2020. 08. 11 인출.
- 최셋별. (2003). 한국 사회에서의 영어실력에 대한 문화자본론적 고찰. **사회과학연구논총**, **11**, 5-21.
- Austin, P. C., & Stuart, E. A. (2015). Moving towards best practice when using inverse probability of treatment weighting (IPTW) using the propensity score to estimate causal treatment effects in observational studies. *Statistics in Medicine*, *34*(28), 3661-3679.
- Cooper, H., Nye, B., Charlton, K., Lindsay, J., & Greathouse, S. (1996). The effects of summer vacation on achievement test scores: A narrative and meta-analytic review. *Review of Educational Research*, *66*(3), 227-268.
- Funk, M. J., Westreich, D., Wiesen, C., Stürmer, T., Brookhart, M. A., & Davidian, M. (2011). Doubly robust estimation of causal effects. *American Journal of Epidemiology*, *173*(7), 761-767.
- Greenwald, R., Hedges, L. V., & Laine, R. D. (1996). The effect of school resources on student achievement. *Review of Educational Research*, *66*(3), 361-396.
- Koenker, R., & Bassett Jr, G. (1978). Regression quantiles. *Econometrica: Journal of the Econometric Society*, 33-50.
- Min, K. S., Jung, H., & Kim, C. M. (2017). Examining a causal effect of Gyeonggi innovation schools in Korea. *KEDI Journal of Educational Policy*, *14*(2), 3-20.
- Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, *70*(1), 41-55.
- Rubin, D. B. (2001). Using propensity scores to help design observational studies: application to the tobacco litigation. *Health Services and Outcomes Research Methodology*, *2*(3-4), 169-188.
- Steiner, P. M., & Cook, D. (2013). Matching and propensity scores. *The Oxford Handbook of Quantitative Methods*, *1*, 237-259.

ABSTRACT

The Effect of Innovation School on Middle School Students' Academic Achievement : Using Propensity Score Matching and Quantile Regression

Yeseul Choi(Research Fellow, Seoul National University)

Sora Lee(Associate Researcher, Korea Institute for Curriculum and Evaluation)

This study examined the effect of “Innovation(*Hyukshin*) school” on middle school students’ academic achievement. Using Gyeonggi Education Panel Study(GEPS) data, we employed propensity score matching methods to control for potential selection bias and heterogeneity among students from two different school groups, Innovation school and general middle school. We then performed doubly robust estimation that combines the regression model with weighting by propensity score. Furthermore, we utilized quantile regression to examine whether the effects of Innovation school differ across the distribution of academic achievement. Relative to the general school, we did not find the significant effect of the Innovation school on middle school students’ academic achievement. Our findings, however, indicate that Innovation school improves students’ academic achievement in Korean language, especially among students who performed to the bottom 25% of the academic distribution. Based on these findings, the significance of the study and its practical implications for policy and future research are discussed.

[Key words] Innovation School, school effect, academic achievement, propensity score matching, quantile regression