# Peer Effects From Exits to Private Schools: Evidence from Project STAR

**Sanjay Satish**
**ECON495S**
**09/30/21**

*Professor Michelle Connolly, Faculty Advisor*
*Professor Robert Garlick, Faculty Advisor*

---

*Honors Thesis submitted in partial fulfillment of the requirements for Graduation with Distinction in Economics in Trinity College of Duke University.*

Duke University
Durham, North Carolina
2021

## Introduction

Imagine yourself as an elementary school teacher trying to manage an unruly group of students. Do you seat them across the classroom from each other, hoping that once separated they stop their disruptive behavior? Or does this choice mean that they'll simply find someone else to talk to, resulting in further disruption? Answering these questions may prove instrumental for the success of your students. There exists extensive empirical literature examining the effect of exogenous policy shifts on student achievement; however, most of this research has generated inconclusive results. Much of the focus today is centered around the estimation of peer effects: the unobserved, peer-to-peer spillovers from students both inside and outside the classroom. Studying these mechanisms through which student outcomes may be maximized or diminished has significant consequences for policymakers.

It is well documented in the literature that student-to-student peer effects do in fact exist and that such effects are, in part, influenced significantly by peer quality. It is also well known that both private schools provide a channel through which richer, high-ability students are able to leave the public school system. Many private schools have entrance exam requirements which deliberately select for high-achieving students. Given the existence of such rationing, it is perhaps the case that the act of exiting a school may have unobserved effects on peers in the classroom. **That is, if the distribution of students that exit public schools for private schools consists primarily of high-achieving students, perhaps there exist negative or reduced positive peer effects from such withdrawal. My paper aims to analyze the effects of this attrition, especially in relation to the varying degree of these effects based on student ability.**

Publicly-provided education can be described as a public good. That is, within the economic context it satisfies two conditions: non-excludability and non-rivalry. Individuals cannot be excluded from the public school system (non-excludability) and one student's consumption of education does not reduce the availability of the good to others (non-rivalry). Private schools create rivalrous conditions by introducing rationing through two distinct channels: tuition rationing and cream-skimming. Tuition rationing is the practice through which private schools select for students with high-income backgrounds by charging for seats – disallowing lower-income families from consumption without some form of subsidization. Cream-skimming is the practice through which private schools select for high-ability students by introducing entrance exam cutoffs as a requirement for admission. The practice can be seen as “skimming” the highest performing students from the public school system into private schools. Although, in some cases, private schools may “cross-subsidize” (i.e. high-income, low-ability students’ tuition is used to provide scholarships for low-income high-ability students) it remains that the plurality of students that enter private-schools who engage in cream-skimming are likely to be of high-ability.

Utilizing evidence from Tennessee’s Student/Teacher Achievement Ratio Project (Project STAR) – an experiment in which elementary school students and teachers were randomly assigned to different class sizes – my paper aims to estimate the peer effects of student attrition on academic achievement between the 1st and the 3rd grade. My analysis draws upon subsequent data on graduation and dropout rates of the students involved in STAR as well as the identification of students that left the public school system.

## Literature Review

Since the publication of the Coleman Report in 1966 there has been suspicion that student unobservables have a significant impact on education production.[1] The classic education production function predicts that test scores can be represented as a function of class size and confounds.[2] Subsequent research has shown that such a relationship is likely incorrect, especially with the lack of an identified relationship between class size and achievement.[3] Some, as first described by Lazear (2001), theorize that disruption in the classroom presents a tradeoff between learning and class size – a higher proportion of "better behaved students" leads to a larger optimal class size.[4] That is, disruptive students may have a greater negative impact on smaller class-sizes than larger ones. Such negative spillovers from disruption can perhaps be best understood through the lens of peer effects. The channel through which one's own production results in spillovers or externalities on peers achievement is measured by peer effects. Subsequent literature has broadened this model of production to focus both on exogenous inputs to production (e.g. class size, curriculum, etc.) as well as inputs that cannot be controlled by policymakers (e.g. parental income, student's attentiveness, peer effects, etc.).[5]

Although there exists a vast literature on peer effects, papers analyzing the peer effects of attrition are somewhat limited, especially given the lack of longitudinal data of student-tracking across both the private and public school systems. One notable study is that of Dills (2005) which utilized the entry of a magnet school in Virginia to estimate the effect of the exit of high-ability students on the proportion of students scoring in the top quartile on a nationally-standardized exam. Dills finds that the loss of high-achieving students to the magnet school lowers the achievement of their lower-scoring

---

[1] Dickinson, E. E. (2016)

[2] Lazear, E. P. (2001) p.777
[3] Hanushek (2008)
[4] Lazear, E. P. (2001) p.779
[5] Hanushek (2008)

counterparts in public schools.[6] Key limitations of Dills' paper, though, include an inability to track individual students' test scores and a lack of demographic controls. Randomized experiments offer the best data to track peer effects given the elimination of selection bias due to peer sorting.

Within the context of Project STAR, papers such as Boozer (2001) focus on estimating the difference in peer-effects due to variation in class-size. Researchers have found that, especially for small classes, much of the effect of class-size on achievement is actually captured by the increased intensity of spillovers from peers in these settings.[7] Given missing data concerns in Project STAR, Sojurner (2011) develops a model for estimating peer effects with the caveat of missing peer data. Sojourner utilizes pre-assignment measures of student ability in conjunction with random assignment to estimate effects, finding that, on average, the effects are positive.[8] Key limitations of Project STAR include differential attrition and a lack of data on where students who left the experiment attended school. Rohlfs & Zilora (2014) exploit whether or not a student had public school testing records after the conclusion of Project STAR as a proxy for identifying if a student remained in the public school system. Such developments should assist in the estimation of peer effects of attrition and reduce the limitations the data impose.

[6] Dills (2005)
[7] Boozer (2001)
[8] Sojurner (2011)

## Data

Project STAR was class-size randomization experiment across grades K-3 in Tennessee between 1985-1989. 11,600 students across 79 schools voluntarily participated in the study. Students and teachers were randomized into one of three class types at the start of the 1985 school year: small classes (13-17 students), regular-sized classes (22-25 students), and regular-sized classes with a teacher's aide (22-25 students in addition to a full-time teacher's aide). After being initially assigned to a class type, students were kept in that class for the duration of the experiment (i.e. 4 years, or between K-3rd grade). Among participating schools, each was required to have at least one of the three class types and randomization occurred within each school. Study administrators collected a variety of data on students from observable characteristics (e.g. race, age, gender, etc.) on both students, teachers, and schools as well as a variety of standardized test scores, administered at the end of each school year. Students who entered a participating school in the middle of the experiment were also randomly assigned to a class type upon entry. There are some limitations to the design, mainly that at the beginning of 1st grade, students in the regular and regular-aide classes were re-randomized across these types. Secondly, approximately 10 percent of students moved across small and regular-sized classes between grades.[9] Lastly, there exists evidence of differential attrition from the experiment by class type. As I work through my project, I may decide to exploit this to determine the severity of the peer effects from attrition based on a student's class type. It is important to note that identification of students who left the experiment, as noted by Rohlfs & Zilora, can only be separated into students who either changed class types, switched to another public school, or left he public school system entirely. Thus, I am unable to accurately identify among the students that left the public school system, which left to a private school Further information about the experiment is well described in Krueger (1999).

[9] Krueger (1999) p. 499

## Works Cited


Boozer, Michael A. and Cacciola, Stephen (2001). Inside the 'Black Box' of Project Star: Estimation of Peer Effects Using Experimental Data.

Dickinson, E. E. (2016). Coleman report set the standard for the study of Public Education. *The Hub*. https://hub.jhu.edu/magazine/2016/winter/coleman-report-public-education/.

Dills, Angela. (2005). Does cream-skimming curdle the milk? A study of peer effects. *Economics of Education Review*. 24. 19-28. 10.1016/j.econedurev.2005.01.002.

Hanushek, Eric A (2008). Education Production Functions. *The New Palgrave Dictionary of Economics*.

Krueger, A. B. (1999). Experimental Estimates of Education Production Functions. *The Quarterly Journal of Economics*, 114(2), 497–532. http://www.jstor.org/stable/2587015

Lazear, E. P. (2001). Educational Production. *The Quarterly Journal of Economics*, 116(3), 777–803. http://www.jstor.org/stable/2696418

Rohlfs Chris Zilora Melanie (2014). Estimating Parents' Valuations of Class Size Reductions Using Attrition in the Tennessee STAR Experiment. *The B.E. Journal of Economic Analysis Policy, De Gruyter*, vol. 14(3), pages 755-790.

Sacerdote, Bruce (2011). Peer effects in education: how might they work, how big are they and how much do we know thus far? *Handbook of the Economics of Education*, vol. 3, pp. 249–277.

Sojourner, Aaron. (2013). Identification of Peer Effects with Missing Peer Data: Evidence from Project STAR*. The Economic Journal.

**Sanjay Satish** **September 27, 2021**
**Duke University – Economics Honors Thesis**

---

# Research Plan & Empirical Strategy

At a high level, my main empirical strategy can be described by the following procedure:

1. Given data in the controls set for Project STAR or surrounding private schools, utilize this data to find (or construct) some measure of what the "typical private school student" looks like.

2. Create two partitions of the dataset by selecting a proportion of schools at random. I would aim for the following distribution of observations in the partition: 30% of schools would exist in a "training" dataset and the other 70% in a "testing" dataset.

3. Within the training dataset, utilize a k-nearest neighbors algorithm to classify the students that left the public school system who "look-like" the typical private school student as individuals who may have, in fact, actually gone to a private school. K-nearest neighbors is a classification technique through which for a given binary outcome variable (in this case, an indicator for whether or not a student left for a private school) individuals are assigned an outcome given their degree of "closeness" to its k-nearest neighbors. To explain more rigorously:

   Imagine we have a collection of $N$ students who we know left the public school system $x_1, \ldots, x_n$ denoted by the set $X$. For each of these students there exists a $d \times 1$ vector $\vec{v_i} \in \mathbb{R}^D$ where each row of the vector represents the value of each covariate for student $x_i$ (i.e. row one is race, row two is gender, etc.). Imagine that we also have a data on $N$ private school students $y_1, \ldots, y_n$ denoted by the set $Y$ which are represented by the set of vectors $W = \{\vec{w_1}, \ldots, \vec{w_n}\}$ where each vector $w_i$ is also $d \times 1$ with each row representing the value of the same data points that are held in each vector $v_i$. Through taking a measure of "closeness" (e.g. the Euclidean distance between each $\vec{v_i}$ and $\vec{w_i}$) we then look at the $k$ closest distance values and look at whichever group the majority of those k-values belong to (i.e. private school student or not a private school student). We then classify each $x_i$ as belonging to the group which the majority of these values fall under.

4. Given these newly classified students, estimate the lagged-achievement effects of these students leaving by some form of the following regression equation:

$$T_{i,g,c,t,s} = \beta_0 + \beta_1(L_{i,g,c,t-1,s}) + \beta_2(A_{i,g,c,t-1,s}) + \beta_3(PA_{i,g,c,t-1,s}) + \beta_4(X_{i,g,c,t-1,s}) + \alpha_s + \gamma_{t-1} + \epsilon \tag{1}$$

Where:

- $T_{i,g,c,t,s}$ is the test score for student $i$ in grade $g$ in classroom $c$ in school $s$ at time $t$.
- $L_{i,g,c,t-1,s}$ is the proportion of students in student $i$'s class that left for private schools at time $t-1$.
- $A_{i,g,c,t-1,s}$ is student $i$'s ability, measured by their test score in time $t-1$
- $PA_{i,g,c,t-1,s}$ is the peer ability for student $i$. That is, the average test scores of the students that left for private school in $t-1$
- $X_{i,g,c,t-1,s}$ is a vector of controls for each student including observable student, teacher, and school data
- $\alpha_s$ is a school fixed effect
- $\gamma_{t-1}$ is a year fixed effect
- $\epsilon$ is the regression error term

If I am unable to accurately classify students who left Project STAR for private schools, perhaps I may refine my research question to only explore the effect of attrition broadly, or perhaps the effect of attrition across class-types. Evidence that attrition may perhaps have greater effects in smaller classrooms may also result in significant policy consequences.