Measuring Spillover Effects of High Impact Tutoring onto Peers of Low Performing Students:

Research Prospectus

Eli Mogel

June 2020

Problem Statement

Every year billions of dollars flow into educational interventions that attempt to curb inequalities both locally and nationwide. Federal funds spent on over half of AmeriCorps projects, work-study programs, and Department of Education programs such as ESEA & IDEA go to fund public school educational interventions.[1] One of the most proven and utilized interventions is tutoring. On a per dollar basis, research has pointed to tutoring as one of the most effective interventions for K-12 students as compared to other well-known educational programs (Harris, 2009). Moreover, tutoring has been found to target low socioeconomic students more effectively than other educational interventions (Dietrichenson et al 2017). This makes sense intuitively, as tutoring provides personalized lessons that are adaptable to each student and require little to no fixed costs that other popular interventions typically do.[2] This research has translated into a push from policy makers for a wider adoption of tutoring programs across schools, especially "high impact" tutoring which focuses on small group sessions that occur at least three times a week for over fifty cumulative hours.

While subsequent policy and research have confirmed high impact tutoring's impressive treatment effects, the true effect of such interventions are potentially still unknown. Previous research has often overlooked the positive spillovers onto peers or friends of the tutored student. One could imagine that tutored students return to their classes and provide positive academic effects to their classmates. Research however either ignores the improved learning of peers or uses the tutored student's peers as controls, leading to a biased experiment, understating the true effects of high impact tutoring. Therefore, this experiment plans to recover the true effect of tutoring by studying the educational outcomes of peers and friends of the tutored students. Specifically:

What effects does tutoring low performing students have on peers of the tutored students?

Answering this question carries importance for two different reasons. First, from a policy perspective, understanding what spillover effect tutoring has to classmates can provide a better understanding of the true impacts tutoring and how to best distribute tutoring across a school district, leading to a more efficient use of funds to fund educational programs. Second, from a research perspective, previous tutoring experiments have utilized designs unsuited to measure spillovers primarily due to low power or due to the fact many experiments have control and treatment students integrated into the same schools or classrooms. This study's design focuses on minimizing interference to measure true impacts. Moreover, peer effect literature has mainly relied on observational data to determine how to optimize classroom compositions for learning – tracking – or to understand why achievement differs

---

[1] https://studentsupportaccelerator.com/funding-tutoring-programs

[2] For instance, universal pre-k and smaller classrooms programs typically require capital outlay and the hiring of multiple credentialed staff. Computer assisted learning can be done cheaply but can only aide one student at a time.

with specific classroom compositions. Instead, this study fills gaps in the literature by using a novel experimental approach to try to understand the impact of treating already formed peer groups.

Theory and Mechanisms

Peer effects, in the context of this experiment, can be thought of as an externality in which peers' backgrounds, current behavior, or outcomes affect an outcome of another individual. Therefore, if we think of a simple educational production function of student i we could imagine that:

$$Y_{ict} = f(X_{ict}, \alpha_{ict}, P_{ict})$$

where X_{ict} are the characters of student i in classroom c in year t. α_{ict} would be the classroom fixed effects, such as inputs or teacher experience. Most importantly, P_{ict} would be the peers of student i. These peers could function under two possible frameworks, the bad apple or shining light (Hoxby et al. 2006, Lazear 2001). Where peers provide a template for student i to follow, so that good peers provide good habits to adopt, leading to improved academic outcomes for student i. Bad peers, however, could do the opposite: disrupt learning and provide poor academic habits that would lead to diminished academic output from student i. Furthermore, as student i interacts more with certain peers, student i might inherit more of the peer's habits. This could imply that close friends of student i will impact their outcome more than distant peers.

Peer effects and the intensity of their effects based on friendship leads to the question, what occurs if peer (or friend) j in student's i class receives high impact tutoring? Figure 1 shows a DAG of what we could expect. What is well known from previous tutoring literature is which mechanisms improve peer j's academic achievement (listed under input). The aspect important to this experiment and the part overlooked in other RCTS are the mechanisms in which peer j would improve student i's achievement.
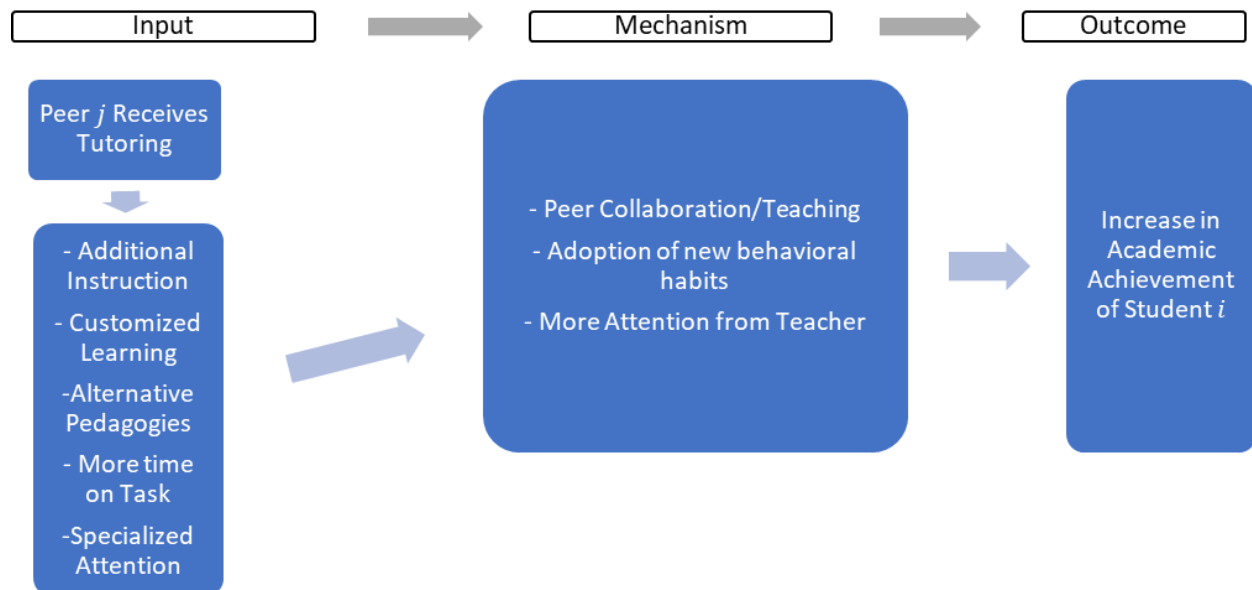

Figure 1: DAG of Peer Effects from Tutoring

2

Literature Review

Conducting an experiment focused on the spillover effects from tutored students onto nontutored peers relies on two streams of literature: tutoring and peer effects literature.

A recent meta-analysis from a team of J-PAL North American researchers, led by Philip Oreopoulos analyzed over 90 tutoring RCTs to find a pooled significant effect of 0.37 standard deviations (SD) across all programs. Stronger effects of tutoring are found during primary education, with an outcome of 0.41SD. Moreover, smaller sessions had strongest treatment effect at 0.46SD. However, groups as large as 3 still provided impressive results: 0.38SD. More importantly, tutoring proves most effective when the intervention is carried out during school, when students are ready to learn; research on tutoring outside of school hours led to insignificant results. One feature of tutoring literature is its relatively well documented large-scale experiments, with over 15 studies being conducted on samples of over 400 students. For instance, Gersten et al. (2015) led a 1000 student, four state, 76 school RCT on math tutoring and found a 0.34SD treatment effect. Parker et al. (2019) ran a multi-grade (4th to 8th) math RCT in Minnesota and found strong pooled treatment effects of 0.24SD. These strong effects, especially in the large-scale setting play well into the context of this experiment which requires sampling from many schools to detect a treatment effect. Unfortunately, there are no RCTs that attempt to measure spillover effects of tutoring, so evidence as to any spillovers is best found in the peer effects literature.

Peer effects literature focused on education has utilized several research designs and theoretical underpinnings to understand how and why peers influence each other in the classroom. The earlier literature on the subject largely used microdata of students coupled with student, classroom, teacher, or school fixed effects to find noisy but consistently significant results (Hanushek 2003, Betta et al. 2004, Sass 2008). More notable works have utilized quasi-experiment designs, exploiting randomized classroom assignment in primary schools to find more precise estimates between 0.27SD to 0.5SD (Kang 2007, Wang 2010). Within the peer effects literature there has been consistent evidence for heterogenous peer effects, where weaker students receive stronger treatment effects due to the presence of higher achieving peers (Kang 2007, Sass 2008, Imberman et al 2012). For instance, Kang (2007) found the lowest quarter of students in class received treatment effects 74% larger than pooled estimates, with a 0.47SD treatment effect for the bottom quarter of students.

Experiments studying peer effects are not well utilized, however some do point to second order effects that positively influence peers of the treated individual. For example, Kremer et al. (2009) found that incentivizing young female students in class induced male students with a bump in test scores as well, possibly due to more teacher interaction. Carrell et al. (2013) randomly formed study pods for incoming Air Force Academy students to study peer effects, but found the weakest achieving peers in random study pods did markedly worse than control peers. This negative effect was due to endogenous formation of peer groups after random study pods were formed, where peers would segregate into subgroups based on academic achievement within the study pod. Moreover, psychology literature speaks to Carrell's results with evidence to the fact students form academic relationships with similarly achieving peers, and that these relationships can influence subsequent academic and behavioral outcomes in schools (Gremmen et al 2017, Shin 2018). All of this indicates that peer effect research ought to study improving already formed peer groups, rather than studying students randomly or "optimally" assigned into peer groups.

Intervention

  To study spillover effects of tutoring onto nontutored peers and friends, I plan to partner with the National Student Support Accelerator to provide math tutoring to students in grades 3 to 5 for one full school year. With that I can study improvements in math achievement on state exams for nontutored peers before and after tutoring is administered across schools.

  The National Student Support Accelerator is an academic accelerator that "aims to catalyze high impact tutoring research, and aid school districts in implementing tutoring programs nationwide."[3] As such, partnering with the Accelerator provides the experiment with a list of potential school districts willing to partake in a tutoring intervention that already have formed deals with tutoring providers in their state. Additionally, the Accelerator works closely with nonprofits such as The Bill and Melinda Gates Foundation, Walton Family Foundation, and others to fund tutoring research and programs.

  The intervention focuses on math tutoring for grades 3 to 5 for several reasons. First, previous research on tutoring has shown math tutoring provides a tighter distribution of results relative to ESL or English interventions. Second, many of the school districts working with the Accelerator only have implemented ESL or English tutoring, leaving less of a two-sided compliance issue with a math-based intervention.[4] Third, state administered math testing typically begin at the end of second grade, therefore administrative data from school districts will contain baseline math achievement starting from grade 3. Lastly, primary education interventions show the strongest treatment effects, thus focusing just on elementary schools allows for easier detectable treatment effects.

Design

Data

Data for this study will come from two sources. First, administrative data from sample school districts can provide baseline and endline math test scores on their respective state exam, allowing me to measure value added math achievement. Administrative data will also contain demographic information on sample students, classroom assignment for current year, and additional behavior information such as days absent, suspensions, etc. An additional benefit of administrative data is its low cost to use, as only data wrangling and cleaning is necessary to make use of the information.

  Second, surveys on student friendships in each sample classroom will be administered as to elicit which nontutored students are friends of the tutored student. This will allow for measurements of how spillover effect potentially change due to social proximity to the tutored students. This data will be administered during the school year and will ask students to list their three "best" friends in the classroom.

Sample

The sample in this experiment are students in the lowest 25% of achievement within each classroom of schools partaking in the experiment, I refer to these students as eligible students. The lowest 25% of students are measured on baseline math achievement as compared to other students in the classroom.

---

[3] https://studentsupportaccelerator.com/about
[4] If I were to roll out an English tutoring intervention I could face low uptake, as many schools do not need English tutoring. Additionally, if I provide English tutoring it will be difficult to have control schools to drop their already implemented high impact English tutoring program.

These eligible students must also be in the baseline data, as to measure their baseline math achievement, and must be in regular classes, i.e., not a specialized class.[5]

This sample set is eligible to receive tutoring and is the main sample that I measure spillovers on. It is assumed that low performing students are friends with other low performing students, so when some low performing students get tutored, the other nontutored low performing students will receive the brunt of the spillover effects.

Treatment

Treatment will be composed of 3-to-1 tutoring of eligible students in the same grade and classroom with a paraprofessional tutor. Paraprofessional tutors are full-time or part-time tutors that have some formal training in the subject matter and in teaching. Sessions will occur three days a week, for fifty minutes, for a total for twenty in-school weeks. In sum, students will receive 50 hours of tutoring. This treatment regimen is largely taken from Gersten et al. (2015) which provided 76 schools with paraprofessional math tutoring. Because this study is concerned with spillovers, not tutoring itself, I chose to follow an already vetted intervention that can stimulate large treatment effects. As reference, Gersten's study found a 0.34SD treatment effect on math achievement after 17 weeks.

Research Design

The experiment follows a randomized saturation design with three arms. The first arm are classrooms where 50% of eligible students are tutored. The second arm are classrooms where all but one eligible student is tutored. The last arm serves as a pure control where no student is tutored. This saturated design allows me to study general spillover effects – by comparing arm one and two to control. It also allows me to study intensity of treatment changes spillover effects by comparing arm one to arm two.

In a practical sense this design serves two purposes. First, this experiment relies on a lot of precision to detect reliable outcomes, therefore a saturation approach provides enough power to find such effects. Second, understanding how spillovers change because of treatment saturation provides insights into how to efficiently allocate tutoring across a school, district, or state. If spillovers are nonexistent, then tutoring could be better administered to many in the school or class. If spillovers are strong, then tutoring a few students across many schools or classes would be more efficient.
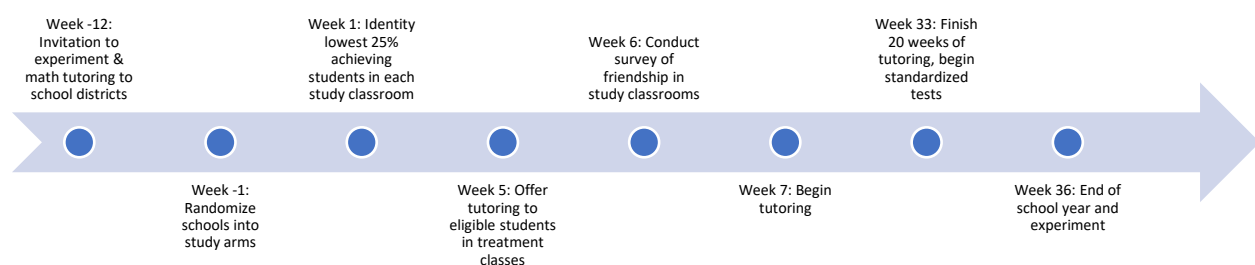
Randomization & Rollout

Figure 2: Rollout timeline

---

[5] Such as special education, accelerated or gifted learning, etc.

The experiment is conducted across one full school year; however, rollout begins during the previous school year. First, I offer the math tutoring intervention to all school districts and schools within the district. Schools that accept must agree to administer the survey regarding in-class friendships. Of schools that accept, each will be blocked into triplets on school district and average baseline math achievement for grades 3 to 5. Of the triplets, each will be randomized into a study arm. This is done to control for block effects, as each district will have heterogenous characteristics and baseline achievement.[6]

Next, during the first week of school I identify the lowest achieving students in each sample classroom, so all treatment and control schools will have identifiable sample students. At the start of week 5, I randomize which 3rd, 4th, and 5th grade classrooms in treatment schools will receive tutoring. I select only one classroom per grade in treatment schools to reduce interference. If I treated multiple classrooms in the same grade, I would not be able to control for cross-class spillovers within a treatment school. Next, I send out offer forms to parents of students eligible for tutoring. Parents that return the forms by week 6 will be randomly drawn to have their children receive tutoring. By offering treatment first, then randomizing conditional on accepting treatment, I try to minimize noncompliance issues before the treatment period begins.

At the end of week 6, each sample classroom will receive friendship forms where all students list their three best friends in class. Then, at the beginning of week 7, tutoring begins and occurs during non-core classes (P.E., art, etc.). Tutoring occurs for 20 active school weeks so that tutored students receive 50 cumulative hours of personalized tutoring. After 20 active weeks, the school year will be in the final weeks where state standardized testing occurs, and end line math results will be recorded. By week 36, schools typically end, and analysis of the experiment will begin. Figure 2 above visualizes the ideal rollout of the experiment.

Methodology: Outcome & Model

The primary outcome of interest is value added math achievement of eligible non-tutored students in treatment classrooms, the sample in this case only contains the bottom quarter of students in both treatment and control classrooms:

$$Y_{icgt} = \pi_0 + \pi_1\big(Z_{icgt} * E_{icgt}\big) + \pi_2 E_{icgt} + \pi_3 Z_{icgt} + \pi_4 X_{icg,t-1} + B_c + G_g + \varepsilon_{icgt} \ (1)$$

Y_{icgt} is a standardized end line math test score for student i, in class c, in grade g, in year t. Further, Z_{icgt} is a dummy variable indicating if the student was in a treatment classroom and E_{icgt} is if the student was a nontutored eligible student. π_1 measures the spillover onto nontutored students in treatment classrooms and compares it against eligible students in control classrooms. $X_{icg,t-1}$ measures all baseline characteristics of students, including baseline math test scores. Lastly, B_c are blocking fixed effects for each sample classroom, G_g are grade fixed effects, and ε_{icgt} is the error term.

While model (1) measures general spillover effects onto other non-tutored low achieving students, spillovers might affect all students heterogeneously conditional on explicit friendship; then model (2) measures spillovers onto friends regardless of eligibility of tutoring, the sample contains all students in both control and treatment classrooms:

---

[6] Each district contracts to different tutoring services, have different demographics, and achievement. Controlling for each block will help attenuate this heterogeneity for final analysis.

$$Y_{icgt} = \pi_0 + \pi_1(Z_{icgt} * Fr_{icgt}) + \pi_2(Fr_{icgt} * Num_{icgt}) + \pi_3 Fr_{icgt} + \pi_4 Z_{icgt} + \pi_5 X_{icg,t-1} + B_c + G_g$$
$$+ \varepsilon_{icgt} \ (2)$$

Fr_{icgt} is a dummy variable indicating if student i is present on a friendship form of a tutored student or eligible student in the control group. Num_{icgt} counts how many tutored student's friendships forms student i is present on. Then, π_1 measures if being a friend of any tutored student produces spillover effects. While π_2 measures if popularity or intensity of friendship to tutored students produces differential spillovers. Model 2, unlike model 1, estimates spillovers for both eligible and ineligible students of tutoring if the student is present on a tutored student's friendship form.

Lastly, with a saturation design I can also compare how spillovers differ between treatment arms, I do this with model (3), where the sample is only treatment classrooms and the lowest quarter of students:

$$Y_{icgt} = \pi_0 + \pi_1\{100 - 1\%\}T_{icgt} + \pi_2\{100 - 1\%\}E_{icgt} + \pi_3 X_{icg,t-1} + B_c + G_g + \varepsilon_{icgt} \ (3)$$

$\{100 - 1\%\}T_{icgt}$ is if student i is a tutored student in a "all but one" treatment classroom. In this case, the comparison group are other tutored students in the 50% treated group, so estimates measure how saturation of tutoring affects the tutored student's math achievement. $\{100 - 1\%\}E_{icgt}$ indicates if student i is an eligible nontutored student in a "all but one" treatment classroom. So, comparison measures if saturation of tutoring alters spillover effects onto eligible nontutored students in class.

Secondary outcomes of interest would be on any behavior effects and their spillovers onto nontutored low achieving students. These models would be the same as models above, but the outcome would be changed to days absent, in-school suspensions, or any other behavioral information available in administrative data. Lastly, this experiment opens many avenues in which heterogenous effects could occur. Therefore, I also plan to study outcomes disaggregated by grade level, as spillovers and collaboration between students could occur at different rates in specific grade levels. I also plan to study heterogenous effects by sex of nontutored student and sexes of tutored and nontutored friends. Some evidence points to the fact that students of different sexes collaborate and adopt habits in different ways, leading to heterogenous outcomes.

Power Calculations & Costs

Power Calculations

One issue facing this study is determining an adequate minimal detectable effect. By using previous peer effect point estimates, coupled with tutoring treatment effects I can derive a conservative point estimate that spillovers would produce.

Following the quasi-experimental designs from peer effect literature, treatment effects were on average 0.3SD for a student when their peer's achievement increased by 1.0SD, i.e., an effect 30% the size of their peers. Further, large scale tutoring RCTs found, on average, 0.3SD treatment effects for tutored students. Therefore, if a tutored student's achievement increases by 0.3SD, we could expect that the nontutored student achievement would increase by 30% of the initial treatment effect, so 0.09SD.

Next, parameters for this study are composed of the number of schools, classrooms, students, the ICC between student's achievement, the variation explained by baseline covariates (R^2), and

proportion of classrooms in treatment or control arms. First, using Texas public school data, I assume there are 240 students in grades 3 to 5 in each school. Each class has 25 students, so there are 9 classrooms per school. Further, treatment schools can only provide one classroom per grade to minimize interference and control schools can provide all 9 classes to the study (interference is not an issue for control schools). Because each class has 25 students, there will be 6 eligible students per class. Using Hedges (2007), I predict the ICC to be 0.11 and R^2 to be .5, as I have all demographic and baseline achievement data to use as controls. Lastly, using PowerCalculator I find the optimal treatment/control assignments to be 30% in the half-treated arm, 20% in the "all but one arm," and 50% in the control arm.[7] The only parameter I can manipulate to gain more power would be the number of schools. Table 1 and 2 below displays parameter values and possible school quantities to sample. I need at least 260 schools to comply in the study to reliably detect a spillover effect onto nontutored of 0.9SD.

Table 1: Minimal Detectable Effect Sizes with Different School Quantities		
Schools	**75% Compliance**	**50% Compliance**
350	SNT: 0.089SD; Slope-SNT: 0.19SD	SNT: 0.12SD; Slope-SNT: 0.23SD
400	SNT: 0.08SD; Slope-SNT: 0.18SD	SNT: 0.1SD; Slope-SNT: 0.22SD
450	SNT: 0.07SD; Slope-SNT: 0.17SD	SNT: 0.96SD, Slope-SNT: 0.21SD

SNT: Spillover on nontreated
Slope-SNT: Slope of spillover between two treatment arms

Table 2: Parameters for Study	
Schools:	*Variable*
Classrooms- Treatment: *Classrooms – Control:*	*3* *9*
Eligible Students per Classroom:	*6*
Proportion in 50%-Treatment/All-But-One-Treatment/Control:	*30/20/50*
R^2	*0.5*
ICC	*0.11*

Costs

Costs of this experiment are mostly borne from paying the salary of tutors across all treatment schools. Using the Accelerator's cost calculator, I can estimate all costs of the intervention.[8] First, full-time tutors are paid $20 per hour and can teach three kids per session. Moreover, each student taught carries a fixed cost of 60$ for supplies. Therefore each 20-week, 50-hour, 3-to-1 tutoring session costs $1180.

---

[7] A. Bohren, P. Staples, S. Baird, C. McIntosh, and B. Özler, (2016). Power Calculation Software for Randomized Saturation Experiments, Version 1.0. Available from http://pdel.ucsd.edu/tools/index.html
[8] https://studentsupportaccelerator.com/node/294

This implies that that the half-treated arm will have 78 schools (260 schools * 30%), with 3 groups of students needing tutoring at $1180 per group. This sums to $276,120 dollars for treatment arm one.

Treatment arm two, "all but one," consists of 52 schools (260 schools * 20%) with 5 groups of students needing tutoring. Therefore, treatment sums to $306,500 for arm two. Also, the control schools do not cost any money for this intervention.

Lastly, costs for a research manager and assistant are baked into the experiment so clear communication and data wrangling of administrative data can be done in a timely manner. This adds an additional $245,000 for two years of employment.

In sum, the total cost of the experiment will cost $827,620. However, these costs are extremely variable depending on which school districts and tutoring services are used for the study. We could imagine school districts in coastal California will be more expensive to fund than schools in southern Texas.

Table 3: Costs of Tutoring	
Tutor Salary	20$/Hour
Fixed Cost per Student	60$
Duration of Tutoring:	50 Hours
Tutoring Group Size:	3 Students per Group
Total Cost per Group:	*1180$*

Limitations & Conclusion

Limitations to this study are like many other education experiments but are compounded by the intervention's cross-school and cross-grade comparisons. First, attrition issues from tutored students and nontutored friends pose a problem. For instance, if parents sign their children up to tutoring unwillingly, it could increase unwanted behavior and force parents to switch classrooms, schools, or drop the tutoring program halfway through the year. Additionally, if the single eligible nontutored student leaves the "all but one" treatment classroom, I lose a whole cluster for my analysis. To this end, I plan to track students in the event they switch schools within the same district using administrative data. If the student leaves the district, I plan to use research staff to track down the student and try to cover end line scores in their new school.

Second, compliance in tutoring poses a risk. Schools might already have some other intervention that aims to helps low achieving students in math that is not necessarily tutoring. To this end, I plan to record what other interventions schools have and see if they record student level information on if students partake in those interventions as to control for it. Regardless, because I randomize within district, control and treatment schools should have the same type and compliance into preexisting educational programs. Baseline tests will be run to determine if this could be an issue in analysis as well. I also plan to use research staff to nudge control schools to administer surveys to ensure compliance across all schools.

Lastly, interference can still pose a risk to the study. Either cross-grade spillovers between treatment classroom within the same school or cross-school spillovers within the same school districts could occur. While there are no immediate methods to control for this problem, I expect this to be small in the educational context. Students across grades might interact on the playground or after school,

melding new behavioral habits, but I would not expect for low achieving students (or any student) of different grade levels to be working on schoolwork together during their free time. Controlling for average distance between schools could be an adequate control for cross-school interference if one expects treatment or control students could affect each other outside of their respective schools.

To end, tutoring has continually been proven to target and help students struggling with learning. While many open questions in tutoring remain unanswered, understanding what spillovers could exists can better direct funds into educational interventions that close learning gaps.

Works Cited

Betts, Julian R. "Peer Groups and Academic Achievement: Panel Evidence from Administrative Data." (2004).

Burke, Mary A. and Sass, Tim, Classroom Peer Effects and Student Achievement (June 13, 2008). FRB of Boston Working Paper No. 08-5, Available at SSRN: https://ssrn.com/abstract=1260882 or http://dx.doi.org/10.2139/ssrn.1260882

Carrell, S.E., Sacerdote, B.I. and West, J.E. (2013), From Natural Variation to Optimal Policy? The Importance of Endogenous Peer Group Formation. Econometrica, 81: 855-882. https://doi.org/10.3982/ECTA10168

David C. Parker, Lisa H. Stewart, Ruth A. Kaminski, Susan Thomson, Sandy M. Pulles. (2020) Outcomes of a Vocabulary Intervention Implemented by Community AmeriCorps Members. School Psychology Review 49:3, pages 321-332.

Gersten, Russell. (2016). What We Are Learning About Mathematics Interventions and Conducting Research on Mathematics Interventions. Journal of Research on Educational Effectiveness. 9. 684–688. 10.1080/19345747.2016.1212631.

Gremmen, Mariola & Dijkstra, Jan & Steglich, Christian & Veenstra, René. (2017). First Selection, Then Influence: Developmental Differences in Friendship Dynamics Regarding Academic Achievement. Developmental psychology. 53. 10.1037/dev0000314.

Hanushek, E.A., Kain, J.F., Markman, J.M. and Rivkin, S.G. (2003), Does peer ability affect student achievement? J. Appl. Econ., 18: 527-544. https://doi.org/10.1002/jae.741

Hoxby, Caroline & Weingarth, G.. (2000). Taking race out of the equation: School reassignment and the structure of peer effects. Education. 18.

Imberman, Scott A., Adriana D. Kugler, and Bruce I. Sacerdote. 2012. "Katrina's Children: Evidence on the Structure of Peer Effects from Hurricane Evacuees." American Economic Review, 102 (5): 2048-8

Kang, Changhui. "Classroom peer effects and academic achievement: Quasi-randomization evidence from South Korea." Journal of Urban Economics 61 (2007): 458-495.

Michael Kremer, Edward Miguel, Rebecca Thornton; Incentives to Learn. The Review of Economics and Statistics 2009; 91 (3): 437–456. doi: https://doi.org/10.1162/rest.91.3.437

Nickow, Andre and Oreopoulos, Philip and Quan, Vincent, The Impressive Effects of Tutoring on Prek-12 Learning: A Systematic Review and Meta-Analysis of the Experimental Evidence (July 2020). NBER Working Paper No. w27476, Available at SSRN: https://ssrn.com/abstract=3644077

Lazear, Edward P. "Educational Production." The Quarterly Journal of Economics 116, no. 3 (2001): 777-803. Accessed June 10, 2021. http://www.jstor.org/stable/2696418.

Shin H, Ryan AM. Early adolescent friendships and academic adjustment: examining selection and influence processes with longitudinal social network analysis. Dev Psychol. 2014 Nov;50(11):2462-72. doi: 10.1037/a0037922. Epub 2014 Sep 15. PMID: 25221841.

Wang, Liang Choon. "Peer Effects in the Classroom: Evidence from a Natural Experiment in Malaysia",
Job Market Paper. Accessed June 10, 2021.
http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.603.967&rep=rep1&type=pdf