Does Taking a Break Make Students (un)Productive?*

Steven Bednar† and Kathryn Rouse‡

May 2019

Abstract

Cognitive fatigue has been shown to cause academic performance to deteriorate over the course of the school day for students of all ages. We rely on random placement of physical education classes to test the effect of a break from classroom learning on academic achievement for adolescent students. We find support for a simple model that suggests a break can help students rebound from cognitive fatigue later in the day, but removing students who are actively engaged from the classroom can disrupt the flow of learning earlier in the day and decrease achievement.

---

*We thank seminar participants at the Association for Education Finance and Policy meetings, the South Carolina Applied Micro Day meeting, UNC Greensboro, and Yale University for helpful comments and suggestions.

†Corresponding Author. Department of Economics, Elon University, 2075 Campus Box, Elon, NC 27244, USA. E-mail: sbednar@elon.edu. Phone: +1 (336) 278-5935. Fax: +1 (336) 278-5952. This paper uses confidential data from Los Angeles Unified School District. The data can be obtained by filing a request directly with LAUSD at https://achieve.lausd.net/research. The authors are willing to assist.

‡Department of Economics, Elon University, 2075 Campus Box, Elon, NC 27244, USA. E-mail: krouse@elon.edu. Phone: +1 (336) 278-5933. Fax: +1 (336) 278-5952.
I. Introduction

Cognitive fatigue has been shown to negatively impact learning over the course of the school day for students of all ages (Pellegrini, Huberty and Jones 1995, Pope 2016, Shapiro and Williams 2018). Given constraints organizations like schools inherently face, however, it is not possible for all tasks to be moved to the beginning of the day. Utilizing breaks to mitigate cognitive fatigue is unexplored in the economics literature. We address how the timing of a break from classroom learning in the form of physical education (PE) may be used to boost student productivity and consider that a break may have differential impacts throughout the day.

A simple theoretical model illustrates two competing effects of a PE induced break on learning. First, PE serves a rebounding mechanism. The classroom break combined with the mental benefits of exercise can potentially combat cognitive fatigue. If fatigue increases over the course of the day, then there is more room to rebound later in the day.

Second, PE disrupts the flow of classroom learning. If the ability to learn depends on learning in the previous period, then this disruption has a negative effect on subsequent learning. That is, students might perform worse following a break. The theory predicts that the disruption effect dominates early in the day as students have just settled into learning and do not need much of a rebound to help focus. In this case, the break hinders learning. When PE occurs later in the day, the rebounding effect dominates as the capacity to learn has diminished more at that point, so the break leads to increased productivity.

\footnote{The medical literature suggests that short bouts of physical activity can increase attention-to-task (Budde et al. 2008, Mahar 2011).}
learning.

Using data from 2003-2012 on student schedules and learning outcomes from Los Angeles Unified School District (LAUSD) for middle school students in the sixth through eighth grades, we provide evidence that the timing of breaks matters differentially throughout the day as the theory suggests for an end of year math test. Limiting our sample to students with a standard six period day, we focus on students who have their high stakes testing class in periods two through five, so that there is variation in whether PE occurs before or after their math class. Students in a third period math class perform worse on the math test when the flow of learning is disrupted by having PE in second period. Students in a fifth period math class perform better on the test if they have PE immediately prior to math. This break completely mitigates the negative effect of having math in the afternoon compared to the morning (Pope 2016).

Several heterogenous effects emerge when we look at subgroups in the data. First, the disruption effect is larger and has greater statistical significance for high-achieving students. Csikszentmihalyi (1990) describes flow as the experience of being completely immersed in a task, and that it takes skills and a challenge to achieve flow. Therefore, it is possible that higher-ability students are more likely to achieve this state and thus have more to lose from the break. Second, the magnitude of the effects decreases as student age. Just as we hypothesized that students get used to being in the classroom and learning throughout the day, students may become better able to learn the more exposure they have had to being in school. Further, Peña and Duckworth (2018) show that perseverance of effort despite struggle, or grit, increases with age, so students may
learn how to deal with disruptions and cognitive fatigue over time. We provide some evidence that student behavior improves in afternoon classes following PE, which could be the mechanism through which PE enhances achievement, as found in Jarrett et al. (1998).

We also find important spillover effects. Students perform worse in third period and better in fifth period when a greater percentage of their peers have PE before math. If students are more disruptive then their classmates will be impacted. Teachers are also given a break from classroom teaching in the form of a preparation period, so there might be spillover effects from teacher fatigue to student learning. Unlike students who spend their time away from the classroom outside in a less cognitively challenging environment, teachers use the time to prepare lessons or get caught up on grading. We find that this disrupts their teaching flow early in the day but does not let them refresh for their afternoon classes.

Our results imply that schools can increase student learning by more strategically managing the timing of PE relative to math. We recognize that school schedulers are faced with many constraints which may limit their ability to move around a break from the classroom that is built into students’ schedules. However, in our sample 12% of students in a fifth period math class have PE to end the day and 21% of students in a third period math class have PE immediately prior to it, the times that are most harmful to student learning. This implies that there is some scope to rearrange schedules to benefit students in these math periods.

The results also provide a possible mechanism for why many studies on the amount of
time spent in PE on academic achievement find insignificant results (e.g. Dills, Morgan and Rotthoff 2011, Cawley, Frisvold and Meyerhoefer 2013, Bednar and Rouse 2018). Based on the timing of PE, some students benefit from and other students are hindered by the break, so on average these effects cancel out.

II. Literature Review

Our paper adds to a growing literature that looks at time of day effects and educational achievement. Using a panel of student-level data for sixth through eleventh grades, Pope (2016) finds that having math in the first two periods rather than the last two periods of the day is associated with a 0.02 standard deviation increase on an end of year math test score, an effect similar to raising teacher quality by a quarter of a standard deviation. Pope (2016) does not find similar results for an end of year English test. It is not possible for every middle school student to have math at the beginning of the day, so some schools have responded by introducing a rotating schedule, where classes occur in a different order throughout the week. The idea is that every child gets a chance to learn each subject at the time that is best for them, but it also means that each child learns each subject at the time that is worst for them.

Other studies have separated cognitive fatigue from time of day effects. For example, Shapiro and Williams (2018) exploit the fact that first year students in the US Air Force Academy are randomly scheduled into their core classes. They find that students perform better if their first class is in the afternoon rather than in the morning. However, if the students have a class earlier in the day then this decreases the benefit of having an
afternoon class due to cognitive fatigue.

Sievertsen, Gino and Piovesan (2016) incorporate breaks into their analysis of cognitive fatigue of Danish students taking end of year exams. Students take tests on computers, so the tests cannot all be administered at the same time. They exploit the randomization of test time and show that performance decreases throughout the day, except when students are coming off a nutrition break. Having a quick break and a snack right before the test affects performance on test day. Pope and Fillmore (2015) show that students who take two AP tests perform better on the second test when there are more days between the two tests. Having time to refresh or study before taking a test is important, but these papers do not inform us about breaks and knowledge accumulation over the school year.

Other papers suggest that students tend to lose focus if they are not offered breaks (Jarrett et al. 1998, Pellegrini and Bjorklund 1997, Mahar 2011). Pellegrini et al. (1995) exploit variation in the number of minutes elementary school children were in school before their first break and find that inattention was greater for all students before the break compared to after the break and inattention before the break was greater when the first break occurred later. Our paper adds to this literature by using variation in the timing of breaks from classroom learning caused by PE, and considers that breaks may have differential impacts throughout the day.

Our paper also relates to the literature on the academic effects of physical education. A report by the Institute of Medicine cites a large body of literature which finds moderate to vigorous physical activity increases child brain functioning (Kohl III and Cook,
eds 2013). It suggests both child attention spans and cognitive processing speeds are improved with increased physical activity. Rasberry et al. (2011) reviews the literature on the association between PE and achievement and finds near unanimous support that the relationship is positive or unrelated at worst. Studies that have utilized the ECLS-K find no statistical impact of additional time in PE on achievement (Dills et al. 2011, Cawley et al. 2013). Using the updated nationally representative sample, the ECLS-K:2011, which occurred during a time of increased accountability, Bednar and Rouse (2018) find similar, negligible effects of time spent in PE on academic achievement.

This research, which largely relies on large-scale nationally representative data, faces constraints that limit the analyses to estimation of the average effect of PE time on achievement. While this is an important question to address, average effects mask whether the timing of PE matters. PE may have a positive impact in the afternoon by providing a rebound for slowed student learning, but may be disruptive in the morning and negatively impact learning. The net average effect estimated in the most recent literature is consistent with this possibility, but the policy implications are quite different. A finding that PE does not matter may lead policymakers to discount its importance when, in fact, it may be possible to strategically place PE in a student's schedule to boost their performance. The data available to us in this study includes detailed course schedules. This attribute allows us to tackle this important possibility, which is unaddressed in the current literature.
III. Theoretical Framework

We consider a simple model of student learning, where the ability to learn and thus achievement depends solely on the amount of time that the student has spent in active classroom learning. Let

\[ A(t) = A[e_i(t), f(t)] \]  

(1)

describe the ability to learn the subject taken at time t. Embedded in the ability to learn are two competing effects. \( e_i(t) \) describes the process of becoming engaged in learning. This is what Csikszentmihalyi (1990) describes as being in the flow, or in the zone. If students who are more engaged in the learning process have a higher propensity to learn, then \( e_i > 0 \). It is also important to consider how engagement changes throughout the day. In first period there is a transition away from before school activities, which are less likely to be academic in nature. As the day progresses, students become more accustomed to being in the classroom, and learning becomes easier. We assume that \( e_0 > 0 \) and \( e_0 < 0 \), so that the greatest gains in the engagement effect occur early in the day. \( i = \{H, L\} \) indexes high-ability and low-ability students, who may have different levels of engagement. We assume that \( e_H(t) > e_L(t) \). According to Csikszentmihalyi (1990), getting into the flow requires both high skills and a challenge, so this last assumption says that it is the higher-ability students who are most likely to become engaged quickly in the learning process.

\( f(t) \) describes how students become fatigued as the day goes on, which reduces the capacity to learn, so \( \frac{\delta A}{\delta f} < 0 \). If fatigue intensifies as the day goes on then \( f'(t) > 0 \) and
Assuming engagement and fatigue are additively separable, the marginal ability to learn as a function of time is thus

\[
\frac{dA(t)}{dt} = \frac{\partial A(t)}{\partial e(t)} \cdot \frac{de(t)}{dt} + \frac{\partial A(t)}{\partial f(t)} \cdot \frac{df(t)}{dt}.
\]

Given the assumptions above, \( \frac{dA(t)}{dt} \) is a downward sloping function, that is initially positive but at some point becomes negative. This produces a concave ability to learn function over the course of the day, as the positive impact of engagement increases quickly for low values of \( t \) and the negative impact of fatigue is low in the morning but increases rapidly in the afternoon. This means that students will be more suited to learning at the beginning of the day than at the end of the day, consistent with the results of Pope (2016).

The implications arising from the theoretical model are illustrated graphically in Figure 1 and Figure 2. The dashed line shows the ability to learn throughout the day for someone who has PE in the final, sixth period, so all of their academic learning occurs during the first five periods of the day. The solid line represent the ability to learn when there is a break caused by PE, and we will examine several cases in turn.

A. PE Before Versus After a Morning Math Class

In the simplest setting, PE means that the student is not engaged in classroom learning, so time is essentially stopped. Two people in the same period can have different values
of $t$ and thus be at different points of the ability to learn function. If ability to learn increases early in the day, then someone with a math class in third period will be able to learn more without having had a break from PE than someone who has had the break. That is, $t$ will be larger for someone without the break and thus $A(t)$ will be bigger. We call this the disruption effect, where $A(x) < A(y)$ if $x < y$. It is easier to enter and remain in the state of flow without a disruption.

Figure 1 illustrates this point by comparing the ability to learn throughout the day for students in a third period math class, depending on whether PE occurs before or after math. The solid line represents learning for a student who has PE in second period, immediately before their math class. The dashed line represents learning for a student who has PE in sixth period and thus never has a break from academic learning throughout the day. As we are interested in learning during math class, it is important to focus on the ability to learn during the third period of the day. The student who has PE during second period is removed from the classroom, which in this extreme example halts the effect of time at school on the ability to learn. The student who continues classroom learning gains the extra benefit from staying in the classroom, getting into the flow, and thus has a higher ability to learn during third period. Our assumptions above that $e_H(t) > e_L(t)$ implies that it is the high-ability students who have more to lose from being removed from the academic learning setting when they are in the flow.
B. PE Before Versus After an Afternoon Math Class

On the other hand, if learning ability is decreasing later in the day due to cognitive fatigue, then someone who has a fifth period math class will learn more if they have a smaller value of \( t \) as a result of the PE break. That is, \( t \) will be larger for someone without the break and thus \( A(t) \) will be smaller. We call this the rebounding effect, where \( A(x) > A(y) \) if \( x < y \).\(^2\)

Figure 2 illustrates this point by comparing the ability to learn throughout the day for students who have math in fifth period, depending on whether PE occurs before or after math. The solid line represents learning for a student who has PE in fourth period, immediately before their math class. The dashed line represents learning for a student who has PE in sixth period and once again never has a break from academic learning throughout the day. In this case we are interested in learning ability during fifth period. The positive effect that getting used to being in the classroom has on the ability to learn becomes swamped by cumulative cognitive fatigue by the afternoon. A student who has PE in fourth period is thus able to rebound from a high level of fatigue and reduce its damaging effects on learning.\(^3\)

\(^2\)Pellegrini et al. (1995) use the term *rebound* in their discussion of elementary school students returning to social activities during recess that they were deprived of during classroom instruction.

\(^3\)It is also possible that PE changes the slope of the ability to learn curve and that students do not simply pick back off where they ended when they return from their PE break. The medical literature cited earlier suggests that there are cognitive benefits to exercise. In either case, we should still expect to see the disruption effect in the morning and the rebounding effect in the afternoon.
C. PE Several Periods Before Math

So far we have used Figure 1 and Figure 2 to compare learning in a given period when a break occurs immediately before that period to when the break occurs at the very end of the day. Turning back to Figure 1, we can examine the ability to learn in fifth period for someone who has PE in second period compared to someone who has PE in sixth period. While the student who had an early break faced negative consequences earlier in the day, this break allows them to delay the ensuing cognitive fatigue, and their ability to learn is better at the end of the day. We exploit this insight when we turn to the data to examine whether there is a difference in learning outcomes when a break caused by PE occurs immediately prior to math or in other periods before math.

D. Math in the Middle of the Day

If there is some degree of heterogeneity among students for when the rebounding effect outweighs the disruption effect, then in a given period some students would benefit from a break and some would be harmed. In this case, the competing effects may cancel out on average so that empirically it looks like the break makes no difference. There have been several studies that exploit variation in the number of minutes allocated to PE per week on academic achievement that find small, insignificant effects (e.g. Dills and Hernández-Julián 2008, Bednar and Rouse 2018). The same argument can be made to learning in the middle of the day. Depending on the timing of PE some students benefit while others are harmed, so on average there is no effect.
IV. Data

Our data comes from Los Angeles Unified School District (LAUSD) for the years 2003-2012. We have demographic characteristics, complete school schedules, and test scores for all students in this district. California education code sections 51220-51229 stipulates that for 400 minutes every 10 days “physical education, with emphasis given to physical activities that are conducive to health and to vigor of body and mind” is required for students in grades 7-12. Students in grade 6 are required to have no less than 200 minutes of PE every 10 days. In LAUSD, students in grades 6-8 are grouped together in middle schools with the same bell schedule, so students have the same amount of time dedicated to PE.

We want to avoid selection concerns that make physical education (PE) period not random with respect the testing classes, so we focus on middle school students in grades 6-8. There is a process for exempting students from the PE regulation, particularly in high school. For example, high school students who play sports typically have their practice during the last period of the day, and this is what counts as PE. In addition, high school students are only required to take two years of math to graduate.

Schools implement the state law by having a dedicated period PE. Since we have information on complete school schedules, we can identify the timing of PE relative to testing classes. We create a variable that indicates if PE occurs immediately prior to the testing class and another variable that groups all other periods before the testing class together. Our omitted category is therefore pe periods after the testing class.

The California Standards Test (CST) is a high-stakes multiple choice test given to all
students towards the end of the academic year. We have scores for Math and English, which are tested every year. We turn the raw score into a Z-score at the test name year level so we interpret all of results in terms of standard deviations.

We only keep students at schools with a traditional six period day plus a homeroom that typically occurs at the beginning or end of the day. The school day begins around 8am and students have first and second periods before a short nutrition break. They have third and fourth periods before a thirty minute lunch break. After lunch they have fifth and sixth periods and the day ends around 3pm.

We drop students from charter schools and schools with more than six periods, which often signifies the use of block scheduling, where PE is not offered every day and its timing relative to testing classes changes throughout the week. Students in middle school in this school district typically have year-long classes in math, English, science, social science, PE, and an elective, with the same teacher both semesters and in the same order. We drop students who change schools mid-year and students whose schedule changes from one semester to the next.

Figure 3 shows the average math Z score by class period. In general, scores decrease as the day progresses, indicating the presence of cognitive fatigue effects. However, there is a clear jump up for students who have math during fifth period, which occurs right after a thirty minute lunch break. While the average score is still below that in the first or second period, this is the first indication that an afternoon break can combat

---

4We drop students for whom the elective is in the PE, math, or English departments. A student who takes traditional PE as well as dance or drill team as an elective will have multiple classes in the PE department, although this makes up a small number of observations. Some students take classes called Math Exploration, Adventures in Math, or Math Skills in addition to their traditional math class. Other students have two periods in the English department, if they take journalism, for example.
cognitive fatigue. There is also a shorter break before third period, but in this case the downward trend in performance continues, just as the theory above suggests.

Table 1 shows the summary statistics broken down by whether PE occurs before or after a student’s math class. This district has about 73% Hispanic, 10% Black, and 10% Asian students. About 76% participate in the National Student Lunch Program (NSLP).\(^5\) Around 40% of students have parents with a high school degree or below as their highest level of educational attainment. Importantly, the summary statistics provide evidence that whether a student has PE before or after their math class is not driven by demographic characteristics. The CST math Z-score is about 0.02 standard deviations lower for students who have PE before their math class than those who have it afterwards. However, having PE before math implies that the math class must occur later in the day, and in particular not in the first period, which Pope (2016) and Figure 3 show is when students perform their best. The rest of the table shows that there is balance in demographic characteristics based on the placement of PE relative to the math class. The p-values from t-tests of equal means suggest there are some significant differences in a statistical sense, however the magnitudes are economically insignificant.

Table 2 shows the fraction of students in each PE period (rows) for a given math period (columns). Again, we focus on math periods two through five since there is variation in having PE before or after the math class. Given that there are five periods in which PE can occur other than the math period, we would expect each number to be 0.20 if PE were evenly distributed throughout the day. There is no striking trend that

\(^5\)This means they receive free or reduced price lunch.
suggests that PE is consistently placed in a particular period relative to the math class.

To further check that schools do not place PE strategically to advantage particular students, we run two linear probability regressions where the dependent variable is an indicator for having PE in any period before math in Table 3 and an indicator for having PE immediately prior to math in Table 4 on the demographic variables discussed above. We also include grade fixed effects and a school by course by year fixed effect so that the students taking a specific course, such as algebra, in the same school, year, and period are compared. The first column in each table is the same, because there is only one period before second. Neither table indicates that schools place PE strategically for specific students throughout the day. Most of the coefficients are not statistically significant, and those that are statistically significant are not economically significant. For example, Table 4 indicates that females are about one percentage point less likely than male students to have PE immediately before a fifth period math class.

V. Empirical Analysis

A. Overview of Estimation Strategy

Our data provide every student’s entire class schedule. This allows us to determine when PE occurs relative to high stakes testing periods. We exploit randomness in the relative placement of PE in a student’s schedule. For math classes that occur in periods two through five, some students will have PE before the testing course and others will have it afterwards. As long as this placement is not systemically related to expected
achievement, our results will identify the effect of the timing of PE relative to the high stakes testing class on test performance.

To motivate our empirical strategy, Figure 4 shows how students in each math period fare depending on whether they have PE before or after math. The point estimates and 95% confidence intervals come from t-tests of equal means. Figure 4 provides evidence backing the theory in Section III. The raw data indicate that students perform equally well in second period math no matter if PE is held in first period or later in the day. In third period math, however, students perform worse by about 0.05 standard deviations if PE occurs before math rather than after math. In fourth period math classes students still perform worse if PE occurs before math, but the magnitude is much smaller. Finally, students perform much better in fifth period math on average when PE occurs during any period before math rather than when they end the day with PE.

**B. Empirical Specification**

We estimate the following specification, for student $i$ in course $c$ in school $s$ in year $t$:

$$Z_{icst} = \beta_1 \text{Before}_{icst} + \beta_2 \text{Prior}_{icst} + \eta X_{it} + \lambda_{cst} + \varepsilon_{icst}. \quad (3)$$

$\beta_1$ and $\beta_2$ are our main coefficients of interest, as they measure the impact on the math test $Z$ score of having PE immediately prior to math or in any other period before the math period relative to after it. We distinguish immediately prior from any other period that occurs before math to determine whether it matters where the break from
classroom learning is placed. Our theory suggests that the timing of PE will have a differential effect throughout the day. We therefore estimate our specification separately for periods two through five, where there is variation in whether PE is before or after the class.\footnote{Running separate regression for each period also controls for the fact that students perform better earlier in the day as Pope (2016) finds.} X is a vector of demographic controls including race/ethnicity, gender, a set of indicators for parent’s highest level of education, free and reduced price lunch status (NSLP), English language learner status (ELL), cumulative GPA, and a set of grade indicators. We control for grade level because students in the same grade may take different math classes. $\lambda_{cst}$ is a course by school by year fixed effect, so we compare test score outcomes based on the timing of PE, for all students who take the same math class in the same period in the same school in the same year. The median number of sections of a course in a period is three for all schools. Ideally we would include a classroom fixed effect, but there were not enough observations per classroom to provide adequate variation across all possible PE periods. All standard errors are clustered at the classroom level to account for unobserved shocks to students in the same class.

We look for heterogeneous impacts by splitting up the sample in several ways. In Section III. we argued that high achieving student are more likely to enter the flow and thus have more to lose by having a break from classroom learning early in the day (Csikszentmihalyi 1990). We therefore estimate the specification separately for students below and above median GPA in their school and grade level in that year. Peña and Duckworth (2018) provide evidence that students have more grit, or perseverance of effort despite struggle, as they age. This means that students may learn how to deal
with cognitive fatigue or may re-enter the state of flow more quickly after a cognitive break. We estimate the model separately by grade level to examine the role of grit. In Appendix A we also split the sample by ELL status, NSLP status, gender, and parental education. Although our theory makes no direct predictions for these subgroups, our results may help policy makers who want to take several factors into consideration when setting course schedules.

Finally, we examine spillover effects in two ways. We calculate the percent of students in a given class that have PE before math. If the disruption or rebounding effects affect classroom behavior, then having a larger percentage of classmates behaving well or poorly means that a student can indirectly be impacted by course schedules of other students. Teachers are also given one period throughout the day with no students in order to prepare. Teachers cannot all take this prep period at the same time, so there is variation across students for when the prep period occurs relative to their math period. Teachers may get disrupted from the flow of classroom teaching when they have their prep period. However, while the prep period does cause a break from classroom teaching, it is less clear whether it combats cognitive fatigue, as the teachers spend the time grading or getting their lesson plans together for the following day.\footnote{Shapiro and Williams (2018) find that students benefit from having instructors who have taught the given class earlier in the day. We concentrate on breaks, partly because teachers are responsible for a variety of classes taught in middle school.}
C. Results

Our theory in Section III. predicts that having PE before math could decrease math test scores for morning math classes and increase test scores for afternoon math classes. The results are shown in Table 5. We only display the coefficient on our main variables of interest, those on Before, which indicates that PE occurred in any period before math other than that which precedes it, and on Prior, which is the period immediately before math. The pattern of coefficients is consistent with the theory developed in Section III. and raw data presented in Figure 4. For students who have math in second period, it does not make a statistical difference if they have PE before or after class. Having PE in first period would delay the start of classroom learning, but it cannot have a disruptive effect. In third period, however, having PE before math class can have the disruptive effect and we find a negative and statistically significant coefficient that implies that having PE immediately prior to math is associated with a 0.03 standard deviation decrease in the math test score. Starting the day with PE does not impact third period math test scores. For students in a fourth period math class, we find no effect of having PE before math. The disruption effect and the rebounding effect may cancel out based on the fraction of students who need a break at this point in the day. By fifth period, however, the rebounding effect clearly dominates as having PE prior to math increases the math test score by almost 0.04 standard deviations. This is enough to offset the negative impact of having math later in the day compared to earlier in the day\(^8\).

The coefficients on the demographic variables are relatively consistent across period

\(^8\)Pope (2016) finds that having math in the afternoon decreases the math CST score by about 0.021 standard deviations.
regressions. Females perform about 0.2 standard deviations worse than male students on average, free and reduced priced lunch is associated with worse performance, Hispanic, Black and Native American students have lower performance compared to white students while Asian students have a better performance than white students, performance increases with the level of parental education, and cumulative GPA has a large and positive correlation with math Z-scores.

1. Heterogeneous Effects

We next split the sample in several ways to examine heterogenous effects. Our theoretical model predicts that high-ability students have more to lose from the disruption effect, as they are more likely to be in the state of flow. Splitting the sample by median GPA at the school-year level shows that it is the higher achieving students who are most harmed by the disruption effect in third period. The results are shown in Table 6. Having PE in second period is associated with a 0.04 standard deviation decrease in test scores for high-achieving students who have math in third period. The coefficients on before and prior are negative, but not significant and are much smaller in magnitude for the lower achieving students. Only the lower ability students experience the positive rebounding effect in fifth period, with a larger magnitude of having PE immediately prior to the math class. High-ability students have positive coefficients but they are not statistically significant.

Next, we split the sample by grade in Table 7. Our simple theoretical model in

\[\text{These are not shown to preserve space, but are available upon request.} \]
Section III. indicates that a student’s ability to learn depends on how engaged they are in the learning process. It could be that as students age it becomes easier to get engaged, or in the flow (Csikszentmihalyi 1990). Alternatively, there is evidence that students exhibit more grit as they age, which suggests they are better able to deal with disruption and fatigue (Peña and Duckworth 2018). Table 7 shows that the magnitude of the effects decreases as students age. The disruption effect in third period is roughly 0.1 standard deviations and the rebounding effect in fifth period is up to 0.2 standard deviations for sixth graders. For seventh graders, the disruption effect in third period and the fifth period rebounding effect are the same sign but the magnitudes are 90% smaller and lose statistical significance. The magnitudes drop even further for eighth graders and retain no statistical significance. This group sees the positive impact of having PE to start the day on second period math scores. As students age the circadian rhythm changes and it is not surprising that essentially starting the academic day fifty minutes later leads to improved learning and test scores for this group (Carrell, Maghakian and West 2011, Edwards 2012, Heissel and Norris 2018).

We show the other heterogenous effects in Appendix A. Table A1 splits the sample by ELL status. Students are labeled as English Only (EO), Initial Fluent English Proficient (IFEP), Reclassified Fluent English Proficient (RFEP), or Limiting English Proficient (LEP) which is commonly called English Language Learner (ELL) and used throughout the paper. Many ELL children take ESL instead of English and are dropped from the sample due to the restriction that students must have one English class. In the analysis sample, about 15% of students are designated as ELL. As a non-native English speaker
is still learning the language, cognitive fatigue is likely to play a large role in a student’s achievement. We find that ELL students who are given a break right before fifth period math see a 0.08 standard deviation increase in their math Z score compared to ELL students who have PE at the end of the day. There is no disruption effect for these students. For English proficient students, the rebounding effect is still positive and nearly significant at the 10% level, and the disruption effect is significant at the 5% level. An English proficient student who has PE in second period and math in third period scores 0.03 standard deviations lower than a similar student who has PE after math.

Students who participate in the NLSP or free and reduced price lunch yield some interesting results shown in Table A2. Students who receive free or reduced price lunch have a similar pattern as before, where there is a negative impact of having PE before a third period math class and a positive impact before a fifth period math class. However, for students without a free or reduced price lunch there is a positive impact of having PE in first period on both second and third period math scores, increasing the score by about 0.04 standard deviations. This amounts to starting the classroom learning later in the day. There is no disruption effect in third period and only Before is significant for fifth period. As shown in Table 1, three quarters of the students in this district receive free or reduced priced lunch, so it is possible that the smaller sample size for non-NSLP students affects statistical significance.

Next, we estimate the specification for females and males separately in Table A3. The general pattern is consistent, although coefficients are larger in magnitude for females
and are statistically significant at the 5% level while the coefficients for males are not statistically significant. On average, females score lower on the end of year math test, so it is possible there is more room to improve.

Finally we consider students whose parents have at most a high school degree separately from those parents who have some college or more in Table A4. About 40% of students have parents with a high school degree or less in this district. The signs of the coefficients follow the same pattern as before, but the only coefficient that is statistically significant is the disruption effect for students with less educated parents.

D. Potential Mechanism

A student receives a cooperation grade of either unsatisfactory, satisfactory, or excellent for each class. Table 8 shows the results from a linear probability model where the left hand side variable is an indicator for receiving an unsatisfactory grade for cooperation in either semester in the math class. This indicates that the student displays disruptive behavior issues that likely affects the entire class. About 15% of students in the sample receive an unsatisfactory grade in at least one semester of math. All of the control variables from the previous regressions are included and standard errors are clustered at the classroom level. The results indicate that having PE before 5th period, either immediately prior or anytime earlier in the day, results in a lower probability of receiving an unsatisfactory cooperation grade by about 1.5 percentage points, or roughly 10%. This limited evidence suggests that PE allows students to focus and channel their energy into their class work, overcoming the effects of fatigue.
E. Spillover Effects

If the mechanism through which student performance changes is through attention and focus, then it is possible that there are spillover effects on classmates. In the top panel of Table 9 we add to our basic specification a variable for each student that represents the fraction of their peers who have PE before the given math class. This table shows that while the main effects of having PE before math persist, peer effects also play a role. A student does worse in a third period math class and better in a fifth period math class when more of their classmates have had PE before the math class. For example, if an additional 10% of other students in a fifth period math class have PE before math, a given student should expect their math score to increase by 0.01 standard deviations.

Teachers are also given a break from teaching during the day in the form of a preparation period. Unlike students, teachers do not get to go outside and have a cognitive break. Instead, they prepare their lesson plans, meet with other teachers or administrative staff, grade, etc. In the bottom panel of Table 9 we investigate how the timing of the teacher prep period relative to a student’s math period affects their learning. An early prep period might disrupt teachers who are in the teaching mindset. Later in the day, a prep period might allow a teacher to refresh, especially if they have taught the same material several times. On the other hand, the prep period is not a real cognitive break, so it is possible that cognitive fatigue and energy levels continue to worsen. We find a negative effect for students in third period math when their teacher has a second period prep period. The teachers who are disrupted early in the day seem to pass the negative effects on to the class. However, we do not find teacher rebounding effects later.
in the day.

F. English CST Scores

CST English exam scores allow us to test whether breaks matter in the same way for different classes. Using the same data as we do, Pope (2016) does not find time of day effects for English CST scores and some of his English falsification tests fail. He notes that many studies in the education literature find larger effects for math than English from education interventions. Having PE before a fifth period English class decreases student performance, contrary to what we found for math. This effect can be attributed to students proficient in English, those who participate in the NSLP, and seventh graders. When we split the sample by GPA or parental education, neither subgroup has coefficients significant at the 5% level.

VI. Conclusion

A growing body of literature has examined different school interventions that aim to increase human capital, many of which are costly to implement (Jacob and Rockoff 2011). In this paper, we build on a newer literature on how the time of day affects the ability to learn to suggest an alternative intervention that does not utilize any additional resources. We explore whether the timing of PE, which removes students from the classroom, relative to testing classes affects a student’s ability to learn and thus their performance on an end of year high-stakes test. We develop a simple theoretical model that predicts
that PE will be disruptive early in the day as it causes a break from classroom learning when students have settled in to the learning process. Later in the day, however, PE can be used to combat fatigue.

Data from math test scores provides evidence for this differential effect of PE throughout the school day. This suggests that policy makers can increase achievement through manipulation of school schedules, by placing PE late in the day for students who have high-stakes testing classes early in the day, and ensuring that student with high-stakes classes late in the day have PE before their testing period. Table 2 shows that 12% of students who have math in fifth period have PE in sixth period, which we argue as the most harmful for success in math. 21% of students in a third period math class have PE immediately prior to it when we suggest it should be later, which provides scope to make a meaningful difference through a rearrangement of schedules.

We do not find similar results of PE induced breaks on end of year English test Z-scores. When coefficients are significant they are often of the opposite sign of those for the math test scores. Future research should explore why differences in the type of learning that occurs in different subjects could cause breaks to have different effects.
References


Jacob, Brian A and Jonah E Rockoff, Organizing Schools to Improve Student Achievement: Start Times, Grade Configurations, and Teacher Assignments, Brookings Institution, Hamilton Project Washington, DC, 2011.


Figure 1: PE Before Versus After 3rd Period Math

The horizontal axis measures time throughout the day at school broken up by class period. In this graph, we focus on learning in third period. The solid line represents a student who has PE during second period and the dashed line represents a student who has PE during sixth period.
Figure 2: PE Before Versus After 3rd Period Math

The horizontal axis measures time throughout the day at school broken up by class period. In this graph, we focus on learning in fifth period. The solid line represents a student who has PE during fourth period and the dashed line represents a student who has PE during sixth period.
Students in the data set have six periods with a short nutrition break after second period and a longer lunch break after fourth period.
Figure 4: Average Math Z Score Difference: PE Before Versus After Math Class

Estimates and 95% confidence intervals come from a t-test of equal means of math Z scores for students who have PE before math and after math. A negative point estimate means students score better if PE is after math. A positive point estimate means students score better if PE is before math.
Table 1: Summary Statistics

<table>
<thead>
<tr>
<th></th>
<th>PE Before</th>
<th>PE After</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Math Z-score</td>
<td>-0.011</td>
<td>0.014</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>(0.092)</td>
<td>(0.094)</td>
<td></td>
</tr>
<tr>
<td>Free Lunch</td>
<td>0.759</td>
<td>0.762</td>
<td>0.111</td>
</tr>
<tr>
<td></td>
<td>(0.428)</td>
<td>(0.426)</td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>0.513</td>
<td>0.510</td>
<td>0.163</td>
</tr>
<tr>
<td></td>
<td>(0.500)</td>
<td>(0.500)</td>
<td></td>
</tr>
<tr>
<td>No HS</td>
<td>0.231</td>
<td>0.231</td>
<td>0.982</td>
</tr>
<tr>
<td></td>
<td>(0.421)</td>
<td>(0.421)</td>
<td></td>
</tr>
<tr>
<td>HS</td>
<td>0.181</td>
<td>0.179</td>
<td>0.126</td>
</tr>
<tr>
<td></td>
<td>(0.385)</td>
<td>(0.383)</td>
<td></td>
</tr>
<tr>
<td>Some Coll</td>
<td>0.115</td>
<td>0.116</td>
<td>0.504</td>
</tr>
<tr>
<td></td>
<td>(0.319)</td>
<td>(0.320)</td>
<td></td>
</tr>
<tr>
<td>College</td>
<td>0.086</td>
<td>0.089</td>
<td>0.019</td>
</tr>
<tr>
<td></td>
<td>(0.281)</td>
<td>(0.285)</td>
<td></td>
</tr>
<tr>
<td>Grad School</td>
<td>0.035</td>
<td>0.037</td>
<td>0.011</td>
</tr>
<tr>
<td></td>
<td>(0.183)</td>
<td>(0.188)</td>
<td></td>
</tr>
<tr>
<td>Hispanic</td>
<td>0.729</td>
<td>0.725</td>
<td>0.028</td>
</tr>
<tr>
<td></td>
<td>(0.445)</td>
<td>(0.447)</td>
<td></td>
</tr>
<tr>
<td>Black</td>
<td>0.097</td>
<td>0.096</td>
<td>0.028</td>
</tr>
<tr>
<td></td>
<td>(0.296)</td>
<td>(0.294)</td>
<td></td>
</tr>
<tr>
<td>Asian</td>
<td>0.092</td>
<td>0.094</td>
<td>0.041</td>
</tr>
<tr>
<td></td>
<td>(0.288)</td>
<td>(0.292)</td>
<td></td>
</tr>
<tr>
<td>Native American</td>
<td>0.003</td>
<td>0.003</td>
<td>0.997</td>
</tr>
<tr>
<td></td>
<td>(0.054)</td>
<td>(0.054)</td>
<td></td>
</tr>
<tr>
<td>Cumulative GPA</td>
<td>2.708</td>
<td>2.721</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td>(0.889)</td>
<td>(0.898)</td>
<td></td>
</tr>
<tr>
<td>ELL Status</td>
<td>0.122</td>
<td>0.121</td>
<td>0.375</td>
</tr>
<tr>
<td></td>
<td>(0.327)</td>
<td>(0.326)</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>126,110</td>
<td>105,512</td>
<td></td>
</tr>
</tbody>
</table>

p-value from test of equal means is in parenthesis.
Table 2: PE Course Period Relative to Math Period

<table>
<thead>
<tr>
<th>PE Period</th>
<th>2nd Period</th>
<th>3rd Period</th>
<th>4th Period</th>
<th>5th Period</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st Period</td>
<td>0.228</td>
<td>0.191</td>
<td>0.228</td>
<td>0.226</td>
</tr>
<tr>
<td>2nd Period</td>
<td>0.206</td>
<td>0.206</td>
<td>0.224</td>
<td>0.217</td>
</tr>
<tr>
<td>3rd Period</td>
<td>0.230</td>
<td>0.223</td>
<td>0.223</td>
<td>0.220</td>
</tr>
<tr>
<td>4th Period</td>
<td>0.226</td>
<td>0.160</td>
<td>0.216</td>
<td></td>
</tr>
<tr>
<td>5th Period</td>
<td>0.148</td>
<td>0.220</td>
<td>0.159</td>
<td></td>
</tr>
<tr>
<td>6th Period</td>
<td>0.169</td>
<td>0.224</td>
<td>0.166</td>
<td>0.120</td>
</tr>
</tbody>
</table>

Each number represents the fraction of students for a given math period (columns) that have PE in each period (rows). Columns therefore sum to one. If PE were evenly distributed throughout the day each number would be 0.20.
Table 3: Do Student Characteristics Affect Having PE Before Math?

<table>
<thead>
<tr>
<th></th>
<th>2nd Period</th>
<th>3rd Period</th>
<th>4th Period</th>
<th>5th Period</th>
</tr>
</thead>
<tbody>
<tr>
<td>Free Lunch</td>
<td>-0.003</td>
<td>0.001</td>
<td>-0.004</td>
<td>0.004</td>
</tr>
<tr>
<td></td>
<td>(0.004)</td>
<td>(0.004)</td>
<td>(0.004)</td>
<td>(0.003)</td>
</tr>
<tr>
<td>Female</td>
<td>0.006*</td>
<td>0.002</td>
<td>0.005</td>
<td>-0.002</td>
</tr>
<tr>
<td></td>
<td>(0.003)</td>
<td>(0.003)</td>
<td>(0.003)</td>
<td>(0.002)</td>
</tr>
<tr>
<td>Less than HS</td>
<td>0.000</td>
<td>0.003</td>
<td>0.000</td>
<td>-0.003</td>
</tr>
<tr>
<td></td>
<td>(0.004)</td>
<td>(0.004)</td>
<td>(0.004)</td>
<td>(0.003)</td>
</tr>
<tr>
<td>HS Grad</td>
<td>-0.000</td>
<td>0.000</td>
<td>-0.001</td>
<td>0.002</td>
</tr>
<tr>
<td></td>
<td>(0.004)</td>
<td>(0.004)</td>
<td>(0.004)</td>
<td>(0.003)</td>
</tr>
<tr>
<td>Some College</td>
<td>0.002</td>
<td>0.003</td>
<td>0.003</td>
<td>-0.008**</td>
</tr>
<tr>
<td></td>
<td>(0.005)</td>
<td>(0.005)</td>
<td>(0.005)</td>
<td>(0.004)</td>
</tr>
<tr>
<td>College Grad</td>
<td>-0.008</td>
<td>-0.001</td>
<td>0.008</td>
<td>-0.004</td>
</tr>
<tr>
<td></td>
<td>(0.006)</td>
<td>(0.005)</td>
<td>(0.006)</td>
<td>(0.004)</td>
</tr>
<tr>
<td>College Plus</td>
<td>0.013</td>
<td>-0.005</td>
<td>0.019**</td>
<td>-0.002</td>
</tr>
<tr>
<td></td>
<td>(0.009)</td>
<td>(0.008)</td>
<td>(0.009)</td>
<td>(0.007)</td>
</tr>
<tr>
<td>Hispanic</td>
<td>0.006</td>
<td>-0.014**</td>
<td>0.013**</td>
<td>-0.001</td>
</tr>
<tr>
<td></td>
<td>(0.007)</td>
<td>(0.006)</td>
<td>(0.007)</td>
<td>(0.005)</td>
</tr>
<tr>
<td>Black</td>
<td>0.012</td>
<td>-0.006</td>
<td>0.007</td>
<td>-0.004</td>
</tr>
<tr>
<td></td>
<td>(0.009)</td>
<td>(0.007)</td>
<td>(0.008)</td>
<td>(0.006)</td>
</tr>
<tr>
<td>Asian</td>
<td>0.007</td>
<td>-0.015**</td>
<td>0.015*</td>
<td>0.007</td>
</tr>
<tr>
<td></td>
<td>(0.008)</td>
<td>(0.007)</td>
<td>(0.008)</td>
<td>(0.006)</td>
</tr>
<tr>
<td>Native</td>
<td>0.001</td>
<td>0.024</td>
<td>-0.021</td>
<td>-0.023</td>
</tr>
<tr>
<td></td>
<td>(0.027)</td>
<td>(0.027)</td>
<td>(0.030)</td>
<td>(0.021)</td>
</tr>
<tr>
<td>ELL</td>
<td>0.004</td>
<td>0.007</td>
<td>0.014*</td>
<td>-0.023***</td>
</tr>
<tr>
<td></td>
<td>(0.008)</td>
<td>(0.008)</td>
<td>(0.008)</td>
<td>(0.006)</td>
</tr>
<tr>
<td>GPA</td>
<td>0.000</td>
<td>0.002</td>
<td>-0.005</td>
<td>-0.003</td>
</tr>
<tr>
<td></td>
<td>(0.003)</td>
<td>(0.003)</td>
<td>(0.003)</td>
<td>(0.002)</td>
</tr>
<tr>
<td>adj. $R^2$</td>
<td>0.359</td>
<td>0.483</td>
<td>0.468</td>
<td>0.361</td>
</tr>
<tr>
<td>N</td>
<td>52542</td>
<td>68204</td>
<td>51147</td>
<td>59729</td>
</tr>
<tr>
<td>Mean</td>
<td>0.228</td>
<td>0.397</td>
<td>0.675</td>
<td>0.880</td>
</tr>
</tbody>
</table>

*** p<0.01, ** p<0.05, * p<0.10. The dependent variable is an indicator for having PE any period before the given math period. Regressions also include grade dummies and a course by school by year fixed effect. Standard errors are clustered at the classroom level.
Table 4: Do Student Characteristics Affect Having PE Immediately Before Math?

<table>
<thead>
<tr>
<th></th>
<th>2nd Period</th>
<th>3rd Period</th>
<th>4th Period</th>
<th>5th Period</th>
</tr>
</thead>
<tbody>
<tr>
<td>Free Lunch</td>
<td>-0.003</td>
<td>-0.002</td>
<td>0.002</td>
<td>0.004</td>
</tr>
<tr>
<td></td>
<td>(0.004)</td>
<td>(0.004)</td>
<td>(0.004)</td>
<td>(0.004)</td>
</tr>
<tr>
<td>Female</td>
<td>0.006*</td>
<td>-0.005*</td>
<td>0.007**</td>
<td>-0.008**</td>
</tr>
<tr>
<td></td>
<td>(0.003)</td>
<td>(0.003)</td>
<td>(0.003)</td>
<td>(0.003)</td>
</tr>
<tr>
<td>Less than HS</td>
<td>0.000</td>
<td>0.008**</td>
<td>0.003</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>(0.004)</td>
<td>(0.004)</td>
<td>(0.004)</td>
<td>(0.004)</td>
</tr>
<tr>
<td>HS Grad</td>
<td>-0.000</td>
<td>-0.001</td>
<td>-0.002</td>
<td>0.003</td>
</tr>
<tr>
<td></td>
<td>(0.004)</td>
<td>(0.004)</td>
<td>(0.004)</td>
<td>(0.004)</td>
</tr>
<tr>
<td>Some College</td>
<td>0.002</td>
<td>0.003</td>
<td>-0.003</td>
<td>-0.001</td>
</tr>
<tr>
<td></td>
<td>(0.005)</td>
<td>(0.004)</td>
<td>(0.005)</td>
<td>(0.005)</td>
</tr>
<tr>
<td>College Grad</td>
<td>-0.008</td>
<td>-0.001</td>
<td>-0.012*</td>
<td>0.009</td>
</tr>
<tr>
<td></td>
<td>(0.006)</td>
<td>(0.005)</td>
<td>(0.006)</td>
<td>(0.006)</td>
</tr>
<tr>
<td>College Plus</td>
<td>0.013</td>
<td>-0.004</td>
<td>0.018*</td>
<td>0.005</td>
</tr>
<tr>
<td></td>
<td>(0.009)</td>
<td>(0.007)</td>
<td>(0.009)</td>
<td>(0.009)</td>
</tr>
<tr>
<td>Hispanic</td>
<td>0.006</td>
<td>-0.004</td>
<td>0.016**</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>(0.007)</td>
<td>(0.006)</td>
<td>(0.007)</td>
<td>(0.007)</td>
</tr>
<tr>
<td>Black</td>
<td>0.012</td>
<td>0.003</td>
<td>0.013</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td>(0.009)</td>
<td>(0.007)</td>
<td>(0.009)</td>
<td>(0.009)</td>
</tr>
<tr>
<td>Asian</td>
<td>0.007</td>
<td>0.004</td>
<td>0.017**</td>
<td>0.005</td>
</tr>
<tr>
<td></td>
<td>(0.008)</td>
<td>(0.007)</td>
<td>(0.008)</td>
<td>(0.008)</td>
</tr>
<tr>
<td>Native</td>
<td>0.001</td>
<td>0.027</td>
<td>0.022</td>
<td>-0.052*</td>
</tr>
<tr>
<td></td>
<td>(0.027)</td>
<td>(0.028)</td>
<td>(0.027)</td>
<td>(0.029)</td>
</tr>
<tr>
<td>ELL</td>
<td>0.004</td>
<td>-0.001</td>
<td>-0.007</td>
<td>-0.007</td>
</tr>
<tr>
<td></td>
<td>(0.008)</td>
<td>(0.008)</td>
<td>(0.008)</td>
<td>(0.008)</td>
</tr>
<tr>
<td>GPA</td>
<td>0.000</td>
<td>0.001</td>
<td>-0.005</td>
<td>-0.001</td>
</tr>
<tr>
<td></td>
<td>(0.003)</td>
<td>(0.003)</td>
<td>(0.003)</td>
<td>(0.003)</td>
</tr>
<tr>
<td>adj. $R^2$</td>
<td>0.359</td>
<td>0.307</td>
<td>0.347</td>
<td>0.251</td>
</tr>
<tr>
<td>N</td>
<td>52542</td>
<td>68204</td>
<td>51147</td>
<td>59729</td>
</tr>
<tr>
<td>Mean</td>
<td>0.228</td>
<td>0.206</td>
<td>0.223</td>
<td>0.216</td>
</tr>
</tbody>
</table>

*** p<0.01, ** p<0.05, * p<0.10. The dependent variable is an indicator for having PE immediately before the given math period. Regressions also include grade dummies and a course by school by year fixed effect. Standard errors are clustered at the classroom level.
Table 5: Math Test Scores by Period Relative to PE Period

<table>
<thead>
<tr>
<th></th>
<th>2nd Period</th>
<th>3rd Period</th>
<th>4th Period</th>
<th>5th Period</th>
</tr>
</thead>
<tbody>
<tr>
<td>Before</td>
<td>-0.014</td>
<td>-0.014</td>
<td>0.016</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.013)</td>
<td>(0.014)</td>
<td>(0.016)</td>
<td></td>
</tr>
<tr>
<td>Prior</td>
<td>0.013</td>
<td>-0.030**</td>
<td>-0.004</td>
<td>0.041**</td>
</tr>
<tr>
<td></td>
<td>(0.013)</td>
<td>(0.014)</td>
<td>(0.016)</td>
<td>(0.018)</td>
</tr>
<tr>
<td>adj. $R^2$</td>
<td>0.519</td>
<td>0.506</td>
<td>0.527</td>
<td>0.522</td>
</tr>
<tr>
<td>N</td>
<td>52542</td>
<td>68204</td>
<td>51147</td>
<td>59729</td>
</tr>
</tbody>
</table>

*** p<0.01, ** p<0.05, * p<0.10. Column heading indicates the Math class period. Prior indicates that PE occurs in the period before the math class. Before indicates PE occurred before math but not immediately prior to it. All regressions include school by course by year fixed effects and student demographic characteristics, including gender, race/ethnicity, parent’s education, free or reduced price lunch status, ell status, cumulative gpa, and grade level. Standard errors are clustered at the classroom level.
<table>
<thead>
<tr>
<th></th>
<th>2nd Period</th>
<th>3rd Period</th>
<th>4th Period</th>
<th>5th Period</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Panel A: Below Median GPA</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Before</td>
<td>-0.018</td>
<td>0.013</td>
<td>0.021</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.015)</td>
<td>(0.016)</td>
<td>(0.019)</td>
<td></td>
</tr>
<tr>
<td>Prior</td>
<td>-0.004</td>
<td>-0.021</td>
<td>0.017</td>
<td>0.051**</td>
</tr>
<tr>
<td></td>
<td>(0.015)</td>
<td>(0.015)</td>
<td>(0.018)</td>
<td>(0.021)</td>
</tr>
<tr>
<td>adj. $R^2$</td>
<td>0.361</td>
<td>0.371</td>
<td>0.388</td>
<td>0.385</td>
</tr>
<tr>
<td>N</td>
<td>25622</td>
<td>33629</td>
<td>25949</td>
<td>29567</td>
</tr>
<tr>
<td>Mean</td>
<td>-0.393</td>
<td>-0.382</td>
<td>-0.408</td>
<td>-0.392</td>
</tr>
</tbody>
</table>

| **Panel B: Above Median GPA** |            |            |            |            |
| Before                   | -0.000     | -0.017     | 0.031      |            |
|                          | (0.018)    | (0.019)    | (0.021)    |            |
| Prior                    | 0.024      | -0.039**   | -0.016     | 0.037      |
|                          | (0.017)    | (0.018)    | (0.022)    | (0.024)    |
| adj. $R^2$               | 0.517      | 0.501      | 0.527      | 0.518      |
| N                        | 26920      | 34575      | 25198      | 30162      |
| Mean                     | 0.400      | 0.386      | 0.366      | 0.393      |

*** p<0.01, ** p<0.05, * p<0.10. Column heading indicates the Math class period. Median GPA is calculated at the school-year level. Prior indicates that PE occurs in the period before the math class. Before indicates PE occurred before math but not immediately prior to it. All regressions include school by course by year fixed effects and student demographic characteristics, including gender, race/ethnicity, parent’s education, free or reduced price lunch status, ell status, cumulative GPA, and grade level. Standard errors are clustered at the classroom level.
Table 7: Math Test Scores by Period and Grade Level

<table>
<thead>
<tr>
<th></th>
<th>2nd Period</th>
<th>3rd Period</th>
<th>4th Period</th>
<th>5th Period</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Panel A: 6th Graders</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Before</td>
<td>-0.056</td>
<td>-0.093</td>
<td>0.138*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.043)</td>
<td>(0.058)</td>
<td>(0.072)</td>
<td></td>
</tr>
<tr>
<td>Prior</td>
<td>-0.046</td>
<td>-0.080*</td>
<td>-0.070</td>
<td>0.225***</td>
</tr>
<tr>
<td></td>
<td>(0.069)</td>
<td>(0.044)</td>
<td>(0.081)</td>
<td>(0.080)</td>
</tr>
<tr>
<td>adj. $R^2$</td>
<td>0.551</td>
<td>0.511</td>
<td>0.527</td>
<td>0.544</td>
</tr>
<tr>
<td>N</td>
<td>8346</td>
<td>24792</td>
<td>8877</td>
<td>12234</td>
</tr>
<tr>
<td>Mean</td>
<td>-0.048</td>
<td>0.015</td>
<td>-0.064</td>
<td>-0.013</td>
</tr>
</tbody>
</table>

| **Panel B: 7th Graders** |            |            |            |            |
| Before         | -0.030     | -0.014     | 0.004      |            |
|                | (0.020)    | (0.021)    | (0.024)    |            |
| Prior          | -0.023     | -0.032     | -0.004     | 0.021      |
|                | (0.020)    | (0.021)    | (0.024)    | (0.025)    |
| adj. $R^2$     | 0.549      | 0.532      | 0.553      | 0.552      |
| N              | 21376      | 20591      | 19619      | 22909      |
| Mean           | 0.078      | 0.045      | 0.008      | 0.060      |

| **Panel B: 8th Graders** |            |            |            |            |
| Before         | -0.001     | 0.008      | -0.008     |            |
|                | (0.017)    | (0.019)    | (0.021)    |            |
| Prior          | 0.042**    | -0.007     | 0.017      | 0.006      |
|                | (0.017)    | (0.016)    | (0.022)    | (0.025)    |
| adj. $R^2$     | 0.490      | 0.487      | 0.515      | 0.492      |
| N              | 22820      | 22821      | 22651      | 24586      |
| Mean           | -0.024     | -0.035     | -0.042     | -0.039     |

*** p<0.01, ** p<0.05, * p<0.10. Column heading indicates the Math class period. Prior indicates that PE occurs in the period before the math class. Before indicates PE occurred before math but not immediately prior to it. All regressions include school by course by year fixed effects and student demographic characteristics, including gender, race/ethnicity, parent’s education, free or reduced price lunch status, ell status and cumulative gpa. Standard errors are clustered at the classroom level.
Table 8: Unsatisfactory Cooperation Grade Math Class in Either Semester

<table>
<thead>
<tr>
<th></th>
<th>2nd Period</th>
<th>3rd Period</th>
<th>4th Period</th>
<th>5th Period</th>
</tr>
</thead>
<tbody>
<tr>
<td>Before</td>
<td>0.003</td>
<td>0.000</td>
<td>-0.015**</td>
<td>-0.015**</td>
</tr>
<tr>
<td></td>
<td>(0.005)</td>
<td>(0.006)</td>
<td>(0.007)</td>
<td></td>
</tr>
<tr>
<td>Prior</td>
<td>-0.003</td>
<td>-0.002</td>
<td>-0.009</td>
<td>-0.015**</td>
</tr>
<tr>
<td></td>
<td>(0.005)</td>
<td>(0.005)</td>
<td>(0.006)</td>
<td>(0.008)</td>
</tr>
<tr>
<td>adj. $R^2$</td>
<td>0.236</td>
<td>0.249</td>
<td>0.256</td>
<td>0.247</td>
</tr>
<tr>
<td>N</td>
<td>52542</td>
<td>68204</td>
<td>51147</td>
<td>59729</td>
</tr>
<tr>
<td>Mean</td>
<td>0.147</td>
<td>0.152</td>
<td>0.160</td>
<td>0.164</td>
</tr>
</tbody>
</table>

*** p<0.01, ** p<0.05, * p<0.10. Column heading indicates the Math class period. Students are given a cooperation grade of Excellent, Satisfactory, or Unsatisfactory each semester. The dependent variable is an indicator for receiving unsatisfactory in at least one semester in their math class. Prior indicates that PE occurs in the period before the math class. Before indicates PE occurred before math but not immediately prior to it. All regressions include school by course by year fixed effects and student demographic characteristics, including gender, race/ethnicity, parent’s education, free or reduced price lunch status, ell status, cumulative gpa, and grade level. Standard errors are clustered at the classroom level.
Table 9: Spillover Effects - Math Class

<table>
<thead>
<tr>
<th>Panel A: Percent of Classmates who have PE before Math</th>
<th>2nd Period</th>
<th>3rd Period</th>
<th>4th Period</th>
<th>5th Period</th>
</tr>
</thead>
<tbody>
<tr>
<td>% of Class with PE Before</td>
<td>-0.049</td>
<td>-0.059*</td>
<td>-0.008</td>
<td>0.113**</td>
</tr>
<tr>
<td></td>
<td>(0.039)</td>
<td>(0.031)</td>
<td>(0.040)</td>
<td>(0.045)</td>
</tr>
<tr>
<td>adj. $R^2$</td>
<td>0.519</td>
<td>0.506</td>
<td>0.527</td>
<td>0.522</td>
</tr>
<tr>
<td>N</td>
<td>52518</td>
<td>68183</td>
<td>51129</td>
<td>59719</td>
</tr>
<tr>
<td>Mean</td>
<td>0.014</td>
<td>0.008</td>
<td>-0.026</td>
<td>0.005</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel B: Timing of Teacher Prep Period Relative to Math Period</th>
<th>2nd Period</th>
<th>3rd Period</th>
<th>4th Period</th>
<th>5th Period</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prep Before</td>
<td>-0.043*</td>
<td>0.028</td>
<td>-0.012</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.026)</td>
<td>(0.026)</td>
<td>(0.034)</td>
<td></td>
</tr>
<tr>
<td>Prep Prior</td>
<td>0.020</td>
<td>-0.064**</td>
<td>0.020</td>
<td>-0.022</td>
</tr>
<tr>
<td></td>
<td>(0.027)</td>
<td>(0.026)</td>
<td>(0.031)</td>
<td>(0.038)</td>
</tr>
<tr>
<td>adj. $R^2$</td>
<td>0.520</td>
<td>0.503</td>
<td>0.523</td>
<td>0.509</td>
</tr>
<tr>
<td>N</td>
<td>41219</td>
<td>55644</td>
<td>41070</td>
<td>48092</td>
</tr>
<tr>
<td>Mean</td>
<td>-0.039</td>
<td>-0.035</td>
<td>-0.071</td>
<td>-0.045</td>
</tr>
</tbody>
</table>

*** p<0.01, ** p<0.05, * p<0.10. Column heading indicates the Math class period. In Panel A we include the fraction of students in the classroom who have PE before math calculated before we drop students for reasons listed in Section IV. In Panel B we create dummy variables for whether the teacher’s prep period was immediately prior to the given math class or in any other period before the math class period. All regressions include school by course by year fixed effects and student demographic characteristics, including gender, race/ethnicity, parent’s education, free or reduced price lunch status, ell status, cumulative gpa, grade level, as well as PE before and prior indicators. Standard errors are clustered at the classroom level. Standard errors are clustered at the classroom level.
## A Additional Math Heterogeneity Results

Table A1: Math Test Scores by ELL Status

<table>
<thead>
<tr>
<th></th>
<th>2nd Period</th>
<th>3rd Period</th>
<th>4th Period</th>
<th>5th Period</th>
</tr>
</thead>
<tbody>
<tr>
<td>Panel A: English Language Learner</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Before</td>
<td>0.000</td>
<td>-0.001</td>
<td>0.052</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.033)</td>
<td>(0.030)</td>
<td>(0.034)</td>
<td></td>
</tr>
<tr>
<td>Prior</td>
<td>0.023</td>
<td>0.007</td>
<td>-0.025</td>
<td>0.086**</td>
</tr>
<tr>
<td></td>
<td>(0.031)</td>
<td>(0.031)</td>
<td>(0.037)</td>
<td>(0.039)</td>
</tr>
<tr>
<td>adj. $R^2$</td>
<td>0.289</td>
<td>0.304</td>
<td>0.319</td>
<td>0.308</td>
</tr>
<tr>
<td>N</td>
<td>6580</td>
<td>7849</td>
<td>6535</td>
<td>7129</td>
</tr>
<tr>
<td>Mean</td>
<td>-0.726</td>
<td>-0.744</td>
<td>-0.759</td>
<td>-0.764</td>
</tr>
</tbody>
</table>

| Panel B: English Proficient |            |            |            |            |
| Before               | -0.013     | -0.017     | 0.010      |            |
|                      | (0.014)    | (0.015)    | (0.018)    |            |
| Prior                | 0.006      | -0.030**   | -0.004     | 0.032      |
|                      | (0.014)    | (0.014)    | (0.017)    | (0.020)    |
| adj. $R^2$           | 0.504      | 0.489      | 0.513      | 0.505      |
| N                    | 45962      | 60355      | 44612      | 52600      |
| Mean                 | 0.119      | 0.105      | 0.081      | 0.109      |

*** p<0.01, ** p<0.05, * p<0.10. Column heading indicates the Math class period. Prior indicates that PE occurs in the period before the math class. Before indicates PE occurred before math but not immediately prior to it. All regressions include school by grade by year fixed effects and student demographic characteristics, including gender, race, ethnicity, parent’s education and free or reduced price lunch indicator. Standard errors are clustered at the classroom level.
Table A2: Math Test Scores by NSLP Status

<table>
<thead>
<tr>
<th></th>
<th>2nd Period</th>
<th>3rd Period</th>
<th>4th Period</th>
<th>5th Period</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Panel A: Receives Free or Reduced Price Lunch</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Before</td>
<td>-0.031**</td>
<td>-0.007</td>
<td>0.018</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.015)</td>
<td>(0.016)</td>
<td>(0.018)</td>
<td></td>
</tr>
<tr>
<td>Prior</td>
<td>0.004</td>
<td>-0.036**</td>
<td>-0.004</td>
<td>0.055***</td>
</tr>
<tr>
<td></td>
<td>(0.015)</td>
<td>(0.015)</td>
<td>(0.018)</td>
<td>(0.020)</td>
</tr>
<tr>
<td>adj. $R^2$</td>
<td>0.481</td>
<td>0.478</td>
<td>0.493</td>
<td>0.486</td>
</tr>
<tr>
<td>N</td>
<td>39799</td>
<td>52331</td>
<td>38728</td>
<td>45249</td>
</tr>
<tr>
<td>Mean</td>
<td>-0.069</td>
<td>-0.074</td>
<td>-0.109</td>
<td>-0.085</td>
</tr>
<tr>
<td><strong>Panel B: Does not Receive Free or Reduced Price Lunch</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Before</td>
<td>0.041*</td>
<td>-0.025</td>
<td>0.014</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.024)</td>
<td>(0.024)</td>
<td>(0.030)</td>
<td></td>
</tr>
<tr>
<td>Prior</td>
<td>0.044</td>
<td>-0.005</td>
<td>0.005</td>
<td>-0.001</td>
</tr>
<tr>
<td></td>
<td>(0.023)</td>
<td>(0.022)</td>
<td>(0.027)</td>
<td>(0.033)</td>
</tr>
<tr>
<td>adj. $R^2$</td>
<td>0.573</td>
<td>0.541</td>
<td>0.572</td>
<td>0.568</td>
</tr>
<tr>
<td>N</td>
<td>12743</td>
<td>15873</td>
<td>12419</td>
<td>14480</td>
</tr>
<tr>
<td>Mean</td>
<td>0.272</td>
<td>0.276</td>
<td>0.231</td>
<td>0.283</td>
</tr>
</tbody>
</table>

*** p<0.01, ** p<0.05, * p<0.10. Column heading indicates the Math class period. Prior indicates that PE occurs in the period before the math class. Before indicates PE occurred before math but not immediately prior to it. All regressions include school by grade by year fixed effects and student demographic characteristics, including gender, race, ethnicity, parent’s education and free or reduced price lunch indicator. Standard errors are clustered at the classroom level.
Table A3: Math Test Scores by Gender

<table>
<thead>
<tr>
<th></th>
<th>2nd Period</th>
<th>3rd Period</th>
<th>4th Period</th>
<th>5th Period</th>
</tr>
</thead>
<tbody>
<tr>
<td>Panel A: Females</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Before</td>
<td>-0.008</td>
<td>-0.006</td>
<td>0.017</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.015)</td>
<td>(0.017)</td>
<td>(0.019)</td>
<td></td>
</tr>
<tr>
<td>Prior</td>
<td>0.023</td>
<td>-0.039**</td>
<td>0.003</td>
<td>0.050**</td>
</tr>
<tr>
<td></td>
<td>(0.016)</td>
<td>(0.016)</td>
<td>(0.019)</td>
<td>(0.021)</td>
</tr>
<tr>
<td>adj. $R^2$</td>
<td>0.518</td>
<td>0.511</td>
<td>0.536</td>
<td>0.531</td>
</tr>
<tr>
<td>N</td>
<td>26839</td>
<td>35096</td>
<td>26031</td>
<td>30570</td>
</tr>
<tr>
<td>Mean</td>
<td>-0.007</td>
<td>-0.012</td>
<td>-0.050</td>
<td>-0.017</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Panel B: Males</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Before</td>
<td>-0.018</td>
<td>-0.024</td>
<td>0.022</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.017)</td>
<td>(0.018)</td>
<td>(0.021)</td>
<td></td>
</tr>
<tr>
<td>Prior</td>
<td>0.004</td>
<td>-0.014</td>
<td>-0.013</td>
<td>0.039</td>
</tr>
<tr>
<td></td>
<td>(0.017)</td>
<td>(0.017)</td>
<td>(0.021)</td>
<td>(0.024)</td>
</tr>
<tr>
<td>adj. $R^2$</td>
<td>0.524</td>
<td>0.504</td>
<td>0.522</td>
<td>0.518</td>
</tr>
<tr>
<td>N</td>
<td>25703</td>
<td>33108</td>
<td>25116</td>
<td>29159</td>
</tr>
<tr>
<td>Mean</td>
<td>0.035</td>
<td>0.028</td>
<td>-0.002</td>
<td>0.027</td>
</tr>
</tbody>
</table>

*** p<0.01, ** p<0.05, * p<0.10. Column heading indicates the Math class period. Prior indicates that PE occurs in the period before the math class. Before indicates PE occurred before math but not immediately prior to it. All regressions include school by grade by year fixed effects and student demographic characteristics, including gender, race, ethnicity, parent’s education and free or reduced price lunch indicator. Standard errors are clustered at the classroom level.
Table A4: Math Test Scores by Parent Education

<table>
<thead>
<tr>
<th></th>
<th>2nd Period</th>
<th>3rd Period</th>
<th>4th Period</th>
<th>5th Period</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Panel A: HS Grad or Less</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Before</td>
<td>-0.014</td>
<td>-0.011</td>
<td>0.000</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.018)</td>
<td>(0.019)</td>
<td>(0.022)</td>
<td></td>
</tr>
<tr>
<td>Prior</td>
<td>-0.000</td>
<td>-0.044**</td>
<td>-0.012</td>
<td>0.022</td>
</tr>
<tr>
<td></td>
<td>(0.019)</td>
<td>(0.019)</td>
<td>(0.021)</td>
<td>(0.024)</td>
</tr>
<tr>
<td>adj. $R^2$</td>
<td>0.476</td>
<td>0.464</td>
<td>0.486</td>
<td>0.477</td>
</tr>
<tr>
<td>N</td>
<td>21624</td>
<td>27909</td>
<td>21088</td>
<td>24607</td>
</tr>
<tr>
<td>Mean</td>
<td>-0.079</td>
<td>-0.096</td>
<td>-0.128</td>
<td>-0.106</td>
</tr>
</tbody>
</table>

| **Panel B: Some College or More** |            |            |            |            |
| Before           | -0.018     | -0.038     | 0.004      |            |
|                  | (0.024)    | (0.025)    | (0.029)    |            |
| Prior            | 0.034      | -0.020     | 0.013      | 0.036      |
|                  | (0.024)    | (0.024)    | (0.029)    | (0.032)    |
| adj. $R^2$       | 0.546      | 0.514      | 0.551      | 0.543      |
| N                | 12438      | 16541      | 12081      | 14223      |
| Mean             | 0.324      | 0.325      | 0.278      | 0.322      |

*** p<0.01, ** p<0.05, * p<0.10. Column heading indicates the Math class period. Prior indicates that PE occurs in the period before the math class. Before indicates PE occurred before math but not immediately prior to it. All regressions include school by grade by year fixed effects and student demographic characteristics, including gender, race, ethnicity, parent’s education and free or reduced price lunch indicator. Standard errors are clustered at the classroom level.