Early Intervention in College Classes and Improved Student Outcomes

John Gordanier,* William Hauk* and Chandini Sankaran**

Abstract

This research investigates the effectiveness of an early academic intervention in Principles of Economics courses at a large public university. After the end of the fourth week of classes, students who fell below a 70% threshold on a performance measure, or had an attendance rate below 75%, were referred to the for additional academic support. A referral consisted of students being given optional assistance in course specific skills through tutoring, as well as training in general skills like time management and study skills. Using a regression discontinuity framework at the referral threshold, we find that the performance intervention improved student scores on common questions on the final exam by 6.5 to 7.5 percentage points for students at or near the performance threshold. The gains are particularly large for students who entered college with below average math placement scores. These results indicate that low-cost light-touch interventions may significantly affect student academic performance.

Key

*University of South Carolina: jgordanier@moore.sc.edu, hauk@moore.sc.edu.
**Boston College: Chandini.Sankaran@bc.edu

intervention programs, then implications for state allocation on school financing arise (Breneman, Abraham Jr. and Hoxby, 1998).

The paper will proceed as follows: the second section describes the background literature, the third describes the data, the fourth defines our empirical approach, the fifth presents our main results, the sixth extends our analysis and presents results from falsification tests, and then we conclude.

2. Background

Developmental services and development education in colleges and universities are designed to help students who are underprepared for college classes, or do not possess the necessary academic skills to succeed in college. Development services include student support services such as mentoring, tutoring, advising, skills workshops, and early warning academic monitoring. These services have increasingly been used in higher education to help students achieve success in college. While a few studies have attempted to determine the effectiveness of development services on student learning and retention, evidence on the effectiveness of these programs remains scarce with mixed results[3].

Interventions

Interventions within the semester generally fall into two categories: providing students information about their standing in the course and providing students with nudges, or requiring

[3] Development education can also consist of placement of students into remedial courses. See Calcagno and Long (2008) and Bettiner and Long (2009) for evidence on the effectiveness of remedial courses.

them to use, additional support services[4]. Two articles, Chen and Okediji (2014) and Smith, et. al. (2018) have found that simply reminding students of their standing in the course can improve student outcomes. The former employs a regression discontinuity design similar to this study and finds a gain of 13 percentage points on the final exam grades. The latter uses a randomized trial where students are reminded of their current grade on their homework assignments. This intervention improved homework performance by four percentage points. A third article, Dobkin, Gil, and Marion (2010), uses an intervention on the attendance dimension for students who score below a threshold and finds that an intervention to increase student attendance improves performance.

Nudges and Support Services

In a systematic review of the existing literature on nudges in education, Damgaard and Nielsen

nudging students to use services there are a pair of articles that suggest that these increase student use of services, but do not necessarily improve performance. Pugatch and Wilson (2018) employ a randomized experiment advertising peer tutoring services to students via a postcard and find an increase in attendance of about 7%. However, they find no change on performance.

increase in grades, but this estimate was not statistically significant.[6] Another study in this literature by Angrist, Lang, and Kouropoulos (2009) evaluates the effectiveness of academic support series by devising an experiment that offers some students access to a center that coaches first-year students in study skills, offers other students financial incentives for good grades, and some a combination of the two. They find that female students had a high take-up rate for the study skills tutoring and tended to do better at the end of their first year. They attribute this result to the effectiveness of the intervention.

Heterogeneous Impacts

Past studies on the impact of peer support on minority students have shown that minorities tend to benefit from peer support. An investigation of the effectiveness of a tracking program in a large urban school district on high achievers found significant effects that are concentrated among black and Hispanic participants, with minorities gaining 0.5 standard deviation units in fourth-grade reading and math scores with persistent gains through sixth grade (Card and Guiliano 2016). Furthermore, peer support matters in the retention, persistence, and success among college students of color pursuing degrees in STEM fields (Palmer, Maramba and Dancy 2011). This result is likely due to participants being abf4322n(f f1 iw(n)] 024 571.06 Tm0 g0 G[(stud)-1ue)4d)-

score on the college math placement exam.[12] Students in the principles-level classes are largely

Since we

only have demographic information for students enrolled in the Business School, we limit our primary analysis to students in the Business School[13]. However, our results are robust to the inclusion of Arts and Sciences students in an analysis without covariates. Further, for reasons we explain in the following section, we also restrict our analysis to students who were *not* referred based on the attendance measure. This procedure leaves us with 640 students in the dataset used for our primary analysis. Table 1 reports summary statistics for the whole sample and the sample of students near the discontinuity.

[Table 1]

Of these students, 18% fell below the performance cutoff at four weeks. The average performance score was 82%, while the average on the common final questions was 78%. Females make up 40% of the population, while just over 3% is African American. The average math placement score is 15.4 out of 26, which is slightly higher than the cutoff for a recommended placement in calculus rather than college algebra (a score of 14).

4. Empirical Approach

The difficulty in evaluating the effect of the intervention on student outcomes is that the students who receive the treatment are also likely to do worse in the course. To identify the causal relationship, we adopt a sharp regression discontinuity design. Specifically, we take advantage of

[12] Students took one of two possible placement exams prior to beginning their freshman year: a college algebra or calculus exam. The math score we use is the college algebra score. If a student only took the calculuk o22(p)-5(r)-emmatexau(m)19(.)-,

the intervention assignment rule whereby students who fall below a 70% on the performance measure receive the treatment. This procedure allows us to compare the students who were just above the threshold to the students who fall just below the threshold. This approach is complicated by the fact that students were also treated if their attendance rate was less than 75% for the first four weeks of class. For students who were referred for attendance, there should be no discontinuity at the performance threshold as they were treated regardless. To get around this issue, we restrict our analysis only to students who were *not* referred for attendance. Instead, we use the students who were referred for attendance as part of a falsification test. We choose to focus on the performance threshold instead of the attendance threshold because of the lack of continuity in measuring attendance after a fixed number of classes, i.e. students have either missed zero classes, one class, etc. Further, a student who missed 75% of a course that met on a Tuesday and Thursday would have missed two classes, whereas a student enrolled in a Monday, Wednesday, Friday course would have missed three. Finally, a very small number of students were referred for attendance who were *not* also below the performance threshold. Thus, our results should be interpreted as the effect of the treatment on performance, conditional on the student attending more than 75% of classes in the first four weeks.

Empirically, we estimate the following model:

where is the percent of common questions on the final that student *i* answered correctly, is a dummy variable where a 1 indicates the student received the treatment, and are polynomials that are functions

referral (and the referral cutoff (0.70), which is estimated for right and left sides of the cutoff

respectively, is a vector of individual characteristics, is an instructor fixed effect and is a residual term. In this set-up, our treatment variable of interest is

We estimate the model in two ways: first we estimate the model with 2nd, 3rd and 4th degree polynomials using the entire sample. Second, we estimate a local linear model with robust bandwidth selection (Calonico, Cattaneo and Farrell, 2017).[14]

5. Results

Before presenting the results of our estimation, in order to provide some *prima facie* evidence that the treatment has some effect, we first graph the unconditional relationship between the score on the common final questions against the four-week performance measure for students with a score within one standard deviation of the referral cutoff at four weeks.[15] The dashed line represents the threshold for treatment and the solid lines the linear fit on either side of the cutoff.

We can see a monotonic increase in performance on the final up until the cutoff for the threshold, at which point the average score drops before increasing again beyond the cutoff. This is consistent with what we would expect if the treatment has a positive effect.

[Table 2]

Table 2 presents our main results. In panel A, we report the effect of the treatment variable where we estimate global polynomials of the 2nd, 3rd and 4th order. The treatment variable is largely consistent across specifications, but the standard errors rise with the order of the polynomial. The magnitude is between 6.7 and 10 percentage points or 1.3-2 additional questions correct out of 20. Since we are controlling for performance at four weeks, the coefficients on the covariates should be interpreted as the effect relative to the predicted score on the common final given their performance at four weeks. Because some of our other regressors are almost certainly good predictors of performance at four weeks, their direct effect on the final performance of students will be attenuated in our regression specification. Nevertheless, scores on the college placement math test are positively correlated with higher than otherwise predicted performance, where a standard deviation increase in math score suggests a 1.2 percentage point higher than otherwise predicted score on the common final questions. Females do slightly worse than predicted based on their performance at four weeks.

Panel B, reports the results from local linear regressions around the cutoff. The robust bandwidth is 11.8 percentage points around the cutoff (Calonico, Cattaneo and Farrell, 2017). We report the results for the robust bandwidth and also for 3/2s and 2/3rds of the robust bandwidth. The estimates are similar to that of the global polynomial. The treatment effect is estimated to be between 5.4 and 7.6 percentage points, although in the smallest bandwidth the standard errors are larger, and the effect is not statistically significant due to the relatively low numbers of

observations. Once we restrict ourselves to students just around the cutoff, the math placement test score, no longer has any further association with the score on the common questions (beyond its association with performance at four weeks). This result is not surprising, as we would expect students near either side of the cutoff to be similar in ability.

Again, females do worse than predicted by 2 to 4 percentage points. This difference by gender is surprising, as one would expect students near the cutoff to be very similar to each other. One possible explanation for the coefficient on female is that the performance score at four weeks includes some non-exam performance measures. If females do systematically better on those

(Damgaard and Nielsen, 2018), we also look at the effect of our treatment on students based on their college math placement score.[16] We group them into students who scored above average on the placement test and those who scored below average. This score provides some measure of the particularly analytical skills necessary for principles-level economics when they began college.

[Table 3]

Table 3 reports the results from the local linear regression for each group. The treatment effect for the students below average is 17.4%, while just 2.1% on the above average students. This suggests that it is the students who are least prepared when entering college that see the largest benefits.

Identification

It is possible that the threshold of 70% in the regression discontinuity design we employ might, in itself, have some effect on student effort. For example, if students who scored just below a 70 decided to study hard, while those just above a 70 did not, then the treatment variable could spuriously attribute this behavior to the intervention. Indeed, 70% was chosen because it is the grade needed to receive a C in the class, which is necessary for all students in the Economics and Business majors. Hence, we may have reason to believe that students just below the cutoff will study harder than students above the cutoff even without an intervention. In this section we attempt to perform two falsification exercises to see if we were to see a similar effect around a

[16] A natural question given previous work is whether there are differences by race. Unfortunately, we do not have enough minority students in our sample to reliably report results by race.

performance score of 70 when there is no change in treatment for students with scores below 70 and those above 70.

In the first exercise, we consider students during this semester who were treated regardless of whether they fell below the performance cutoff: the students who were below the attendance threshold. We perform the same analysis as in the previous section. However, for this group there should be no jump at the performance cutoff of 70.

[Table 4]

Table 4 reports the results of this exercise. While the sample size is limited, there does not appear to be any systematic pattern as the coefficient on the treatment variable changes signs and magnitude across specifications.

In our second exercise, we use a group of students that were never treated: students from a previous semester.[17] We use data from instructors who taught in both the Spring 2016 and Spring 2017 semester. Since there were no referrals or common questions on the final exam in the 2016 semester, we use the score on the final exam as the outcome variable. We construct a

Due to lacking data on the demographic characteristics and math test scores for these students, we perform the analysis without covariates.

[Table 5]

There are a number of limitations to using data from prior semesters in addition to the lack of data on student characteristics. Most notably, the referral initiative required instructors to track attendance for at least four weeks. This change in tracking attendance confounds efforts to perform a difference in difference approach across semesters.

The results are reported in Table 5. Overall the effect is small and statistically insignificant. There does not appear to be any evidence that students just below the 70% threshold did any better than those just above the 70% threshold in previous semesters.

Besides the possibility that the threshold of 70% has some effect independent of the intervention, there is a potential for student manipulation of the treatment. Since information regarding the referral process is provided to students via the syllabus, it is possible that students who were near the performance cutoff tried harder to avoid being referred. Thus, the students right above the cut-off might be worse on average than if assigned randomly. As long as any difference in ability between students just above the threshold and those just below the threshold is explained by the college placement math scores, then this should not weaken our identification. For this to weaken our identification, it must be the case that a) students are aware of the policy b) are willing to study harder to avoid referral and c) are capable of affecting the four-week performance score at the margin. We think this is unlikely as there is little reason to believe that the intervention is particularly costly to students who do not wish to participate in the intervention, as participation was completely voluntary. Furthermore, the authors are skeptical that information provided on the syllabus implies that students are aware of the policy four weeks into the course.[18]

Finally, there is also the possibility of a selection bias in the data arising from students withdrawing from the course. If receiving the referral increased the likelihood of a student near the cutoff of withdrawing from the course and the students who withdrew were on average worse than those that did not (with a similar m80 g0 G[(sk)-10(e)4(pti)-3(clg0 G12.22 251.0.000018243 F3 12 TnBT/F

from the final exam using data from Spring 2017. Then, we proceed with falsification tests by conducting similar analyses using students from a previous semester before the policy was implemented. Finally, we investigate which students benefit the most from the intervention.

We find that students who were just below the performance threshold, i.e. those who received an intervention notice, performed 6.5 to 7.5 percentage points better on a set of questions on the final exam than students who were just above the threshold. We also find that the students who are least prepared when entering college, measured through placement scores on the math exam which students take before their first year of classes, benefit the most from the intervention. Students who had above average math placement scores increased their grade by 2.1% whereas students with below average math placement scores experienced a grade increase of 17.4% as a result of the intervention. The College Board shows that high school seniors from lower income families and those with less educated parents tend to score lower on the SAT math scores

References

1. Angrist, Joshua, Daniel Lang, and Phillip Oreopoulos (20 Incentives and Services for College Achievement: Evidence from a Randomized Trial *American Economic Journal: Applied Economics*, 1(1), 136-163.
2. Arco-Tirado, José L. Francisco D. Fernández-Martín and Juan-Miguel Fernández-Balboa, of a peer-
Higher Education, 62 (6): 773-78.
3. Can a summer make a difference? The impact of the American Economic Association Summer Program on *Economics of Education Review*, 53 (2016) 46 71.
4. Bettinger, E. P., & Baker, R. B. (2014). The effects of student coaching: An evaluation of a randomized experiment in student advising. *Educational Evaluation and Policy Analysis*, *36*(1), 3-19.
5.

8. Breneman, David W., Ansley A. Abraham Jr. and Caroline M. Hoxby (1998). *Brookings Papers on Education Policy*, 1: 359-383 Published by: Brookings Institution Press.

9. -based guidance *Educational Evaluation and Policy Analysis*, 35: 3 (298-323).

10. [illegible]

15.

Agricultural & Applied Economics Association, 78 (3): 499-507.

16.

Economics of Education Review, 29: 566 575.

17. *Economics of Education Review*, 64: 313-342.

18. *The Journal of Economic Education*, 32 (2): 99-109

19. Moss, B. G., & Yeaton, W. H. (2015). Failed warnings: Evaluating the impact of academic probation warning letters on student achievement. *Evaluation Review*, *39*(5), 501-524.

20. Munley, Vincent G., Eoghan Garvey, and Michael J. McConnell. 2010. "The Effectiveness of Peer Tutoring on Student Achievement at the University Level." *American Economic Review*, 100(2): 277-82.

21.

23.

Innovations in Education and Teaching International, 46 (4).

24. Pugatch, T., & Wilson, N. (2018). Nudging study habits: A field experiment on peer tutoring in higher education. *Economics of Education Review*, *62*, 151-161.

25.

The Review of Economics and Statistics, 93 (2): 436-454.

26. Sankaran, Chandini, Abdullah Al-Bahrani, and Breyon Williams, 2018

27. Smith, B. O., White, D. R., Kuzyk, P. C., & Tierney, J. E. (2018). Improved grade outcomes with an e- *The Journal of Economic Education*, *49*(1), 1-7.

Table 2

Table 3

Results by Math Placement Test

	Below Average	Above Average
Treatment	0.174***	0.021
	(0.044)	(0.043)
Math Score		

Table 4

Panel A: Global Polynomial Results

Polynomial Degree	2nd	3rd	4th
Treatment	0.088	-0.089	-0.146
	(0.06)	(0.119)	(0.19)
Math Score	0.000	0.000	0.000
	(.002)	(.003)	(.003)
African American	-0.06	-0.08	-0.095
	(0.071)	(0.068)	(0.071)
Female	-0.035	-0.035	-0.034
	(0.036)	(0.035)	(0.035)
Number of Observations	96	96	96

Panel B: Local Linear Results

Bandwidth	Robust	1.5X	0.67X
Treatment	0.004	0.036	-