Economics of Education Review 79 (2020) 102063

Contents lists available at ScienceDirect

Economics of Education Review

journal homepage: www.elsevier.com/locate/econedurev

Social networks and college performance: Evidence from dining data

Darius D. Martin, Adam C. Wright *, John M. Krieg

Department of Economics, Western Washington University, 516 High Street, Bellingham, WA 98225, USA

ARTICLE INFO

JEL classification:
I21
I23

Keywords:
social interactions
friendships
heterogeneous peer effects
higher education

ABSTRACT

We investigate the effect of friends in class on academic performance in college using unique data on dining card swipes at a medium-sized public university. We define friendships by academic quarter as repeated meetings among students in the same dining hall. To identify the impact of having a friend in class, we employ models with student- and class-level fixed effects and find having a friend in class has large and positive effects on grades. Our analysis of heterogeneous friend effects reveals that the positive friend effect exists across all types of friend characteristics, suggesting the unconditional importance of social connections.

1. Introduction

Peer interactions have long been argued to play a significant role in academic achievement. It has been established that peers influence a wide range of outcomes including test scores, major choice, and academic grades as well as outcomes and behaviors that indirectly impact human capital accumulation like health and risk-taking behaviors.[1] Due to data limitations, most studies proxy for social interaction using a peer group defined at an aggregated level such as a neighborhood, school, grade, class, or residence hall. Designating peer groups in this manner, however, does not incorporate information on who students interact with nor does it account for the strength of those interactions. Recent work has begun to shed light on the importance of strong social ties, or friendships, for academic outcomes though much remains unexplained.[2] In particular, the effect of friends during post-secondary education, a time in which new friend groups are being formed and important decisions regarding human capital accumulation are being made, has yet to be explored. Understanding the impact of social networks on academic performance is especially policy relevant given the shift of many higher education institutions towards online courses, which has been hastened by COVID-19, where it is likely more difficult to make or leverage social connections.

In this paper, we analyze the effect of social networks on academic performance at the beginning of college. Our contribution is to assess the impact of having peers with strong social ties, or "friends", in the same college courses while carefully addressing the nature of friendship formation and course selection. In the economics of education literature, friendship data is quite limited – most studies rely on the National Longitudinal Study of Adolescent Youth (Add Health) survey in which high school respondents are able to self-nominate up to five friends of each gender.[3] Using this data to identify social networks has three shortcomings: 1) for a given respondent, nominated friendships may entail varying degrees of social interaction that are unobserved to the researcher; 2) the definition of friendship may vary across respondents resulting in friendships of different strengths being measured by the researcher as similar; and 3) respondents may not mutually identify each other as friends leaving researchers to speculate on the nature of relationships between these pairs.

We circumvent many of these problems by using a revealed-preference identifier of friendships. Specifically, we leverage data from university dining halls to identify pairs of students who regularly dine together. Since the vast majority of freshmen frequently use the university dining halls, we can construct quite detailed information about each freshman's dining-related social network. Moreover, if people dine more often with better friends, this data enables us to distinguish between stronger and weaker friendships. We validate our

* Corresponding author at: MS-9074, 516 High Street, Bellingham, WA 98225.
E-mail addresses: darius.martin@wwu.edu (D.D. Martin), adam.wright@wwu.edu (A.C. Wright), john.krieg@wwu.edu (J.M. Krieg).

[1] See Sacerdote (2011), Sacerdote (2014), and Epple and Romano (2011) for a review of this literature.

[2] See for example Pattacchini, et al. (2017), Lavy and Sand (2019), and Lam (2012).

[3] The main draw of using Add Health data is that it contains a nationally representative sample of high school students who are surveyed on a number of typical demographic and outcome variables as well as friendship questions. See for example Fletcher et al. (2020), Bramoullé et al. (2009), Lin (2010), Pattacchini, et al. (2011), and Pattacchini, et al. (2017).

https://doi.org/10.1016/j.econedurev.2020.102063
Received 17 April 2020; Received in revised form 6 August 2020; Accepted 22 October 2020
Available online 2 November 2020

friendship measure by showing that assigned friendship dyads are significantly more likely to exhibit homophily than random pairings of dining hall users. Compared to studies that rely on friendship nomination, like the Add Health survey, measuring friendships in this way provides the distinct advantage that friendships and friendship-intensity are revealed rather than simply stated.

To simplify later discussion, it is important to understand that the term "friendship" as used in this paper represents a symmetric, binary relation between pairs of students who are observed meeting in a dining hall with sufficient frequency. Of course, this differs from the broader social definition of friendship commonly used. We select designating criteria such that if two students are described as friends in this paper, then the social relationship as more broadly understood almost certainly holds between them. That is, under strict enough social interaction criteria, our friendship-determination method is likely a sufficient condition for identifying friendships. Note, however, that it is not a necessary condition for the social relationship; two students may share a strong social bond despite not being identified as friends using our method. We discuss the implications of this later in the paper.

Combining dining data with transcript-level data on student performance, we examine the academic outcomes of a friendship dyad enrolling in the same course within a quarter. Though we concentrate on friends who attend a course together in the same classroom, we expand our analysis to include friends across sections of the same course taught by the same instructor and those taught by a different instructor.[4] Identifying the causal effect of having friends in class on academic performance is a difficult empirical problem because friends and classes are not randomly assigned; students with friends may be fundamentally different from those without friends or students may select into easier (or harder) classes based upon the presence of a friend in that class. To mitigate these concerns, we use a within-student and within-class approach that compares a student's performance relative to classmates in classes taken with friends to their relative performance in classes without friends present.

While this approach has the benefit of controlling for student- and class-level unobservables, it only partially addresses other potential problems, namely correlated effects, non-random selection, and simultaneity. For instance, an unobserved factor that varies within students and across classes, such as subject-specific enthusiasm, may be correlated with having friends in a class and academic achievement (correlated effects). Friends might jointly select into easier classes together (non-random selection) and students may be more likely to make friends in classes they find particularly easy (simultaneity). We address the issues of correlated effects and non-random selection by controlling for a student's stated interest in a subject (indicated on their application for admission), implied subject interest given by the number of credits completed in each academic department, and examining the effect of friends across instructors in the same course, where the results should be qualitatively similar if shared enthusiasm is driving an observed friend effect. Further, we employ the estimator used by Oster (2017) to show that even considerable bias due to selection on unobservables does little to diminish our measured friend effect. The issue of simultaneity is addressed by leveraging the timing of friendships: friendships formed in prior quarters, which cannot be affected by current grades, are used in an alternative specification in place of contemporaneous friendships.

Our main results demonstrate a large and robust effect of social networks on college performance. Using many alternative criteria in designating friendships, we estimate a roughly 0.12 grade point increase (equivalent to 0.12 standard deviations) when students have a friend in the same class, an effect size equal to an increase of a student's SAT test score of about 100 points. In terms of movements in the letter grade distribution, having a friend in class increases the probability of receiving an A grade by about 10% and decreases the probability a D or F grade by roughly 25%.

Additionally, we are able to test how these academic improvements vary by student, friend, and class characteristics. First, we find that the effect of friends increases with the "strength" of the friendship measured by dining hall meetings within an academic quarter. Second, although we find many interesting patterns of heterogeneity based on the interaction of student and friend characteristics, the positive friend effect persists across all subgroups analyzed. In other words, social connections are beneficial regardless of student or friend characteristics. Lastly, friends appear to have a greater positive impact in large classes (i.e., when enrollment is 45 students or greater) but are no more effective at improving grades in hard classes (i.e., classes where the average GPA of students without friends in class is less than 2.9).

This research is closely related to a growing body of evidence identifying the importance of quantity and quality of friends in determining educational achievement and attainment. In general, relatively few studies distinguish between the broader category of peers and the narrower one of friends. The majority of these studies use the aforementioned friendship nomination procedure among high school students in the Add Health data and generally find that friends with better observable characteristics improve own academic outcomes.[5] Lavy and Sand (2019) exploit conditional random assignment of students to classes within Tel Aviv middle schools to examine the impact of having pre-determined friends in class. They find the number of friends in class has either a positive or negative effect on academic performance depending on friends' characteristics and the nature of the friendship. Mora and Oreopoulous (2011) show that self-identified friends of high school students affect intentions to dropout only when the friends are reciprocal. Using a survey in Chinese junior high schools, Lam (2012) is able to distinguish between friends, study mates, emotional supporters, and seatmates and finds that positive personality traits of friends, but not their cognitive traits, positively affect math scores.

None of this literature examines college-aged individuals, where the nature and impact of friendships may differ from those in primary or secondary school. A number of studies have investigated the impact of peers at the college level, but to our knowledge, none has directly assessed the impact of friends. Many studies use the random assignment of students to dormitories to estimate peer effects on academic and social outcomes, and generally find mixed results.[6] However, looking at college roommates may not provide a comprehensive view of students' peer interactions. Carrell et al. (2009) rely on the random assignment of students from the U.S. Air Force Academy to squadrons, which arguably captures a more complete group for peer interactions, and find that better peers positively influence GPA. Findings from the Air Force however may not be generalizable to other university settings. Fischer (2017) leverages random assignment of students to classes at a large public university and finds that women with higher-ability classmates

[4] Here and throughout the paper, we use the terms "class" or "classroom" to describe a section of a course with a specific instructor during a quarter. For example, "ECON 206: Principles of Microeconomics, Section 1" in Fall 2016. Thus, a class is a course-by-section-by-quarter.

[5] Using a difference-in-differences approach, Fletcher et al. (2020) find that more friends with college-educated mothers raises high school GPA for girls but not boys. Exploiting the different waves of Add Health and using indirect friends as an instrument for own friends, Patacchini et al. (2017) show that educational choices of long-lived friendships (lasting more than one year) have a positive effect on own educational attainment. Hill (2015) also uses an instrumental variables approach to demonstrate that a student's share of opposite gender friends negatively affects high school GPA. Bramoullé et al. (2009) and Lin (2010) show that both friends' average grades and average characteristics affect students' educational attainment.

[6] Zimmerman (2003) and Sacerdote (2001) estimate positive peer effects on GPA and take-up of social groups such as fraternities and sororities. Stinebrickner and Stinebrickner (2006) find relatively small and positive peer effects on grades for women but not men. Foster (2006) does not find statistically significant peer effects.

are less likely to graduate with a STEM degree, but cannot identify peer interactions.

This study makes three novel contributions to this literature. First, we use a revealed-preference methodology to identify friendships. This allows us to differentiate between stronger and weaker friendships based on how often students coordinate their meals. Coordination entails some cost to students as compared to costlessly nominating friends in a survey setting. Second, we apply the analysis of friendships to a university setting, where students are likely forming new social groups and making important decisions regarding their human capital formation. Third, we estimate the college friend effect along quantity and quality dimensions, which provides rich insights into the nature of the estimated effect.

Several mechanisms may generate our observed effects of friends on grades, including joint production, social pressure, or mutual insurance among friendship pairs (Lavy and Sand, 2019). The first case, joint production, occurs when having more friends in a group improves the group's academic performance. This could happen directly if learning production increases with friends (i.e., group productivity increases with friends), or indirectly if group members increase effort due to the marginal product of effort increasing with friends. The second potential explanation is social pressure, which assumes the cost of students' effort depends on the presence of friends in class. In this case, students will increase their effort if friends reduce the marginal cost of effort, perhaps by creating incentives to compete or generating enthusiasm. Lastly, friends might be able to insure each other against negative productivity shocks. For example, friends may share notes when one of them misses class due to illness. Although we cannot provide a definitive answer as to the mechanism underlying our estimates, we rule out mechanisms that involve substituting effort between friend and non-friend classes and find some evidence consistent with the joint production and social pressure explanations.

2. Data

2.1. Description of the university and student outcomes

We use administrative data from Western Washington University (WWU), a regional, comprehensive university located in Bellingham, Washington with fall undergraduate enrollment of approximately 15,000 students. In the fall of 2017, WWU enrolled 3,078 freshmen from 37 states and 8 countries. WWU is annually ranked among the top five regional, comprehensive universities in the U.S. News and World Report rankings.

WWU operates three quarters during the academic year. In each quarter, instruction lasts for ten weeks, with an eleventh week set aside for final exams. It also operates an optional summer quarter with about one-fourth of its regular enrollment. Because both housing and dining options change significantly in the summer, we exclude these quarters from our analysis. WWU takes pride in offering small instructional experiences and often splits courses into several sections taught by individual instructors. For example, during a quarter there are often five or six stand-alone sections of Introduction to Microeconomics taught by two or three instructors. Across the university, the average section size is 15.8 students and the median section enrolls 11 students.

Using unique student identification (ID) numbers, we merge data from the university dining halls with administrative records to observe each student's background prior to enrolling in WWU (high school GPA, high school attended, SAT scores) as well as students' academic records while at WWU (the courses attempted and grades earned in those courses). We have complete data from five academic years, starting in 2013-2014 and ending in 2017-2018. Although we have academic records for every undergraduate enrolled at the university, we restrict our analysis to freshmen, since freshmen at WWU overwhelmingly live on campus and use the university dining halls.[7] After their freshmen year, students frequently move off campus and are much less likely to eat in the dining halls. Because we cannot identify friendships among these students, we do not include advanced undergraduates in our analytic sample.

Table 1
Descriptive Statistics

	Mean	Standard deviation	Number of observations
Panel A: Student-Level Data			
fraction who dine every quarter	0.86	0.35	14,281
fraction who never dine	0.09	0.29	14,281
fraction who take at least one class with a friend	0.53	0.50	14,281
fraction male	0.42	0.49	14,280
fraction white	0.73	0.44	14,281
fraction black	0.04	0.19	14,281
fraction Hispanic	0.08	0.27	14,281
fraction first generation	0.31	0.46	14,281
SAT score	1124	148	14,250
Panel B: Student by Quarter-Level Data			
fraction who dine	0.89	0.32	40,869
unlimited meal plan	0.26	0.44	40,869
125 meals	0.30	0.46	40,869
100 meals	0.18	0.38	40,869
75 meals	0.15	0.36	40,869
attempted credits	14.67	1.74	40,869
Panel C: Student by Course-Level Data			
grade	2.91	1.01	127,929
prior department credits	1.05	2.41	128,222
class in interest	0.10	0.30	128,222
friend in class	0.14	0.34	128,222

Notes: The data are from the academic years 2013-14 through 2017-18, excluding summer quarters. "Class in interest" is a binary indicator for whether the course is in the academic department of the student's program interest.

Panel A of Table 1 presents descriptive statistics of 14,281 freshmen who make up the analytic sample employed in this paper. Importantly, 86% of the sample are observed to dine on campus each quarter during their freshmen year while only 9% of freshmen are never observed in the dining halls. The average SAT score of the analytic sample is 1124, and about one-third of freshmen are first-generation college students meaning neither of their parents graduated from a four-year college. Finally, the majority of students are white and almost three-fifths are women.

WWU's campus housing is served by three dining facilities. Students may eat at any dining hall of their choice and dining halls remain open throughout the day. There are four different meal plans available for purchase offering, respectively, 80, 100, 125, or an unlimited number of meals per quarter. Meal plans are purchased quarterly with many students changing their plan from one quarter to the next. Panel B of Table 1 shows that the 125-meal plan is the most common choice for

[7] WWU uses the term freshmen to refer to students enrolling at WWU with no intervening educational experience between high school and WWU. This includes students who may have participated in a dual enrollment program while enrolled in high school but excludes students who attended a different degree-granting institution after high school graduation.

Table 2
Distribution of Student Dining Hall Entries per Quarter

Percentile of Dining Entries	Total	Meal Plan Unlimited	125 meals	100 meals	75 meals
10th	42	69	52	39	21
25th	65	101	75	60	39
Median	90	135	96	80	57
75th	117	168	113	92	68
90th	157	197	120	97	73
Number of student-quarters	36,209	10,662	12,296	7,174	6,077

Notes: Each cell gives the number of dining hall entries by percentile and meal plan per quarter.

freshmen, with the unlimited meal plan being the next most popular. Food is provided on an all-you-can-eat basis in the dining hall. Most students who live on campus are required to purchase a meal plan. [8]

Our empirical goal is to understand how the presence of friends in a class impacts a student's grades. WWU assigns letter grades (grade points) of A (4), B (3), C (2), D (1), and F (0). In addition, grades can be assigned a plus or a minus. Grade points are adjusted downwards by three-tenths of a point if a minus is assigned and upwards by the same amount on all letter grades except an A which cannot receive a plus and F which cannot receive a minus. Panel C of Table 1 shows that the average freshmen grade is 2.91 with a standard deviation of 1.01. Although we will express our results in terms of grade points, since the standard deviation is nearly one, we could state our results equivalently in terms of standard deviations. This facilitates comparison of our findings with the wider literature on academic performance. Panel C also highlights the other control variables used in our empirical specifications. Students are asked to list academic areas of interest on their application for admission (*class in interest*). Ten percent of classes taken by freshmen are in one of those areas of interest. In order to control for knowledge of a subject, we sum the prior credits earned in the department in which the student's current course is offered (*prior department credits*).[9] On average, students have taken 1.05 prior credits in a given class department. Finally, using the definition of friendship developed in the next section, 14% of students have at least one friend in a class. In Panel A, we show that 53% of freshmen take at least one class with a friend during their freshmen year.

2.2. Description of friendship measures

Students obtain access to WWU's dining halls after an attendant swipes the student's university identification card through a card reader. The card reader records the student's unique ID number, the location, and the date and time (measured to the second) of each swipe. We refer to any student observed entering the dining hall on one or more occasions in a quarter as a "diner" in that quarter.

Table 2 displays information concerning the frequency of dining hall use per quarter among all freshmen diners. The median freshman visits the dining hall 90 times per quarter. With 77 days in an 11-week quarter, this implies that most students eat in the dining hall at least once per day. Roughly three-fourths of students eat at least 60 meals in the dining hall per quarter, or about five meals per week. 26 percent of students purchase unlimited access to the dining hall, and as we would expect, these students eat at the dining hall much more often than those on a fixed-number meal plan.

[8] Meal plans are not required for students who live in one of two buildings where housing includes kitchens. These buildings contain 544 out of 4040 total beds on campus and are usually populated by upper classman rather than freshmen.

[9] Classes at WWU are assigned one credit for each hour per week of instruction, and for every two hours per week of laboratory or rehearsal.

Table 3
Number of Friendships Observed among Freshmen using Various Designating Criteria, Averaged across All Quarters

Meetings	Time window (seconds) 5	10	30	60
1	44,993	145,484	514,019	918,145
2	8,521	19,058	91,677	253,517
3	5,850	10,560	26,745	83,225
4	4,632	8,479	14,145	34,188
5	3,821	7,281	10,563	18,450
6	3,219	6,402	8,934	12,542
7	2,767	5,748	7,900	9,860
8	2,408	5,217	7,156	8,380
9	2,099	4,757	6,551	7,429
10	1,854	4,370	6,059	6,727
15	1,027	3,000	4,336	4,707
20	613	2,144	3,226	3,494
25	375	1,557	2,453	2,657
30	236	1,145	1,882	2,054
35	148	855	1,440	1,578
40	94	635	1,108	1,219
60	19	191	364	415
75	6	77	163	192
100	1	19	49	65

Notes: All averages of observed friendships are rounded to the nearest whole number.

Table 4
Distribution of Observed Friendships among Freshmen

No. of Friends	Fall Count	(%)	Winter Count	(%)	Spring Count	(%)
0	2823	20	3614	27	4200	32
1	2432	17	2417	18	2555	20
2	2320	16	2145	16	1993	15
3	2079	15	1762	13	1577	12
4	1520	11	1217	9	1074	8
5	1080	8	880	6	624	5
6	720	5	546	4	412	3
7	456	3	386	3	228	2
8	319	2	251	2	207	2
9	187	1	172	1	88	1
10 or more	265	2	209	2	111	1
No. of Students	14,201		13,599		13,069	

Notes: Friendships are student pairs entering the dining hall within 30 seconds of each other on ten or more occasions in a quarter.

We exploit the dining hall data to identify members of each student's social network. We infer that a social relationship exists between any two students who frequently enter the dining hall together. Accordingly, to designate two students as friends, we specify a time window, and define a "meeting" as any occasion when students enter the dining hall within that window. We specify a meeting threshold and consider two students to be friends in a quarter when they meet at least that threshold number of times.

Table 3 illustrates the prevalence of friendships using alternate designating criteria. Data in the table are the averages across quarters, rounded to the nearest whole number. For example, in the average quarter, there are 6,059 pairs of freshmen who entered the dining hall within 30 seconds of each other on ten or more occasions. We observe a small number of pairs who meet extremely often – for instance, there are 163 pairs of friends who enter the dining hall within 30 seconds of each other at least 75 times – about once per day – each quarter.

To explore the role of friends on academic outcomes, we focus on a time window of 30 seconds and meeting threshold of 10 meetings. That is, our default criteria are that two students must enter the same dining hall within 30 seconds of each other at least 10 times within a quarter to be designated as friends in that quarter. As we show in Section 5, our results are generally not sensitive to these specifications. Using these criteria, Table 4 shows the frequency of observed friends in our five-year

Table 5
Characteristics of Friendships and of All Pairs of Freshmen, Weighted Averages across Academic Years

	Fall		Winter		Spring	
	friends	all pairs	friends	all pairs	friends	all pairs
both male	33.6%	17.6%	36.4%	17.9%	36.4%	18.0%
both female	43.3%	33.7%	40.8%	33.2%	40.9%	33.1%
both minority	8.8%	7.4%	8.9%	7.4%	9.3%	7.5%
both white	56.4%	53.0%	57.2%	53.1%	57.8%	52.9%
both black or both Hispanic	1.4%	0.8%	1.4%	0.8%	1.4%	0.8%
both first generation	9.7%	8.9%	9.3%	8.7%	9.2%	8.5%
neither first generation	54.0%	49.3%	53.8%	49.8%	54.6%	50.3%
same high school	19.1%	0.5%	15.4%	0.5%	15.2%	0.5%
same interest	7.3%	4.6%	7.0%	4.6%	7.3%	4.6%
observations	7,373	6,692,518	6,117	5,847,163	4,814	5,100,806

Notes: Friendships are student pairs entering the dining hall within 30 seconds of each other on ten or more occasions in a quarter. "All pairs" represent the outcomes if students were paired randomly from the WWU freshmen student population.

Table 6
Distribution of Classes Taken with Friends

Panel A: Distribution of Classes Taken with Friends

	Fall		Winter		Spring	
No. of Classes	Count	(%)	Count	(%)	Count	(%)
0	10,269	72	9,158	67	9,213	70
1	2,743	19	2,993	22	2,524	19
2	959	7	1,075	8	948	7
3	213	1	309	2	327	3
4 or more	17	0	64	0	57	0
No. of Students	14,201		13,599		13,069	

Panel B: Distribution of Friends in Classes

	Fall		Winter		Spring	
No. of Friends	Count	(%)	Count	(%)	Count	(%)
0	37,783	88	36,805	85	36,306	87
1	4,347	10	4,929	11	4,515	11
2	783	2	1,004	2	865	2
3	166	0	291	1	176	0
4 or more	72	0	104	0	72	0
No. of student-class pairs	43,151		43,133		41,930	

Notes: Friendships are student pairs entering the dining hall within 30 seconds of each other on ten or more occasions in a quarter. A class is a section of a course with a specific instructor during a quarter, or a course-by-section-by-quarter.

sample. Table 4 includes all freshmen and designates those without a meal plan as having zero friends, a possible source of measurement error to which we return later. Table 4 shows that 80% of fall quarter freshmen have at least one friend. This number falls slightly as the academic year progresses, most likely because students tend to use the dining hall less frequently later in the year.

Prior research on friendships emphasizes homophily, or the tendency for friendships to form between people who are similar to each other in some respect.[10] For instance, homophily among friends has been shown along racial and ethnic dimensions, gender, age, behavior, and occupational interest.[11] To validate our method for designating friendships, Table 5 shows that pairs of freshmen diners identified as friends are exceptionally similar across a range of observable characteristics relative to random pairings of students. Friends are particularly more likely to be of the same gender, to have attended the same high school, and to have reported an interest in majoring in the same department upon enrollment (*same interest*). Table 5 also indicates that friends are more likely to have a similar demographic and socioeconomic background, using first-generation status as a proxy for the latter.[12] This, along with the improbability of repeatedly entering a dining hall with the same stranger, provides evidence that pairs of students we identify as friends indeed have a social relationship.

Recall that our research focus is to assess the effect of friends in class on academic performance. In Table 1, we showed that 53% of freshmen enroll in at least one class with a friend over the academic year. Table 6 gives more detailed information about the frequency of classes taken with friends and the number of friends in class. Panel A of Table 6 shows the distribution of classes taken with at least one friend. Here, each observation is a unique student. During the average fall quarter, 72% of freshmen have no observed friends in any of their classes. About one-in-five freshmen take one class with a friend and a small number of students have friends in multiple classes. As the academic year progresses, students are observed taking more courses with friends—about one-third of students have at least one class with a friend during the winter and spring quarters. Panel B of Table 6 shows the distribution of friends in class, where each observation is a student-by-class. Among students with a friend in class, 80% have one friend in class and 15% have two friends in class. This density of observations with one friend in class and the exceedingly small percentage of students with more than two friends motivates the empirical specification discussed in the next section.

3. Empirical model and identification strategy

3.1. Empirical model

To estimate the effect of having a friend in a class, we start with a basic econometric model of student-course grade performance, Y_{ikst}, with student-level controls of the form:

$$Y_{ikst} = \alpha + \beta Friend_{ikst} + \mathbf{X}'_{ikst}\gamma + \varepsilon_{ikst} \quad (1)$$

where students are indexed by i, courses by k, course-sections (i.e., instructor-by-meeting time) by s, and quarter by t.[13] The variable $Friend_{ikst}$ is an indicator variable that takes the value one if student i has a friend in course k, section s during quarter t. The parameter of interest is βwhich measures the academic performance of students with a friend in class relative to the performance of students when a friend is not present. We choose a simple binary representation for friendship as our preferred parametric form as few students have more than one friend in class (as shown in Table 6, Panel B), but we do allow the effect of friends on performance to vary by the number of friends within a class in some specifications. We do this in two ways: 1) additional binary variables indicating the number of friends in class; and 2) with a linear measure of

[10] See for example Block and Grund (2014).

[11] See Goodreau, et al. (2009), Smith-Lovin and McPherson (1993), Fischer (1977), Knecht, et al. (2010), and Kalmijn (1998) for evidence of homophily across each of these dimensions, respectively.

[12] Recall that first-generation students are defined as those whose parents did not complete a university degree.

[13] Quarters included are fall, winter, and spring from fall 2013 to spring 2018.

the number of friends in class. While we prefer the binary variable approach because it can allow for non-linearities in the relationship between friendships and grades, we also show the linear approach as it eases discussion when examining heterogeneity in the results. The vector X_{ikst} contains student-level characteristics designed to capture effort (attempted number of credits within the quarter), ability (SAT score), and interest in the class (prior credits within the academic department at the beginning of the quarter and an indicator for stated area of interest).[14] Finally, ε_{ikst} is a stochastic error term adjusted for two-way clustering at both the student and class level.[15]

To arrive at our preferred econometric specification, we include student and class-section fixed effects, α_i and δ_{kst} respectively, which necessitates dropping student SAT score which is collinear with the student fixed effects. Our preferred model can thus be represented as:

$$Y_{ic} = \alpha_i + \beta Friend_{ic} + \mathbf{X}_{ic}^{'}\gamma + \delta_c + \varepsilon_{ic} \quad (2)$$

where the combination of indices *k, s*, and *t* have been replaced by the single class-section index *c*.[16] This within-student and within-class model identifies β by comparing a student's performance (relative to his or her classmates) in classes taken with one or more friends to the student's own relative performance in classes without friends.

3.2. Threats to validity

In this section, we address three potential sources of bias in estimating β from Eq. (2): non-random selection, simultaneity, and measurement error.

3.2.1. Non-random selection

The inclusion of the fixed effects in Eq. (2) help address many threats to obtaining unbiased estimates of β using ordinary least squares (OLS). Employing classroom fixed effects implicitly controls for course and instructor fixed effects, which, in turn, controls for the possibility that students and their friends select into relatively easy courses or instructors. Additionally, classroom fixed effects have the advantage of standardizing grades across classes as students' grades are modeled relative to the average grade within each class. We include student fixed effects to allow a student's performance in classes taken without friends to serve as a counterfactual for grades in classes taken with friends present. This ensures that our results are not driven by time-invariant unobserved characteristics of those students who have friends or choose to take classes with friends. For example, if better students are more likely to take classes with friends, omitting the student-level fixed effect would lead us to mistakenly conclude that friends are associated with higher grades—a positive bias of β in our specification.[17] To mitigate concerns about bias induced by time-varying student characteristics, we also report results from a specification that include a student-by-quarter fixed effect. While this identification strategy addresses issues of quarter-specific shocks (such as attachment to school) to both students' friendships and grades, it has the limitation of only identifying the friend effect for students who have classes both with and without friends during the same quarter.

The estimated friend effect from Eq. (2) may remain biased if some other unobserved factor affects both the likelihood of having a friend in class and a student's relative performance. For instance, if interest in the class is not well captured by prior department credits or stated interest, then friends may tend to be present in those classes that a student would do well in regardless. If this is driving the measured friend effect, then it follows that we should also observe higher grades when friends enroll in the same course but have a different instructor. We test for friend effects across both sections and instructors of a course in Section 5.2 (Table 12).

Another possible source of bias would be if friends altered the likelihood of course completion. For instance, if friends encouraged each other to remain in a course they otherwise would drop, perhaps because they expect a low grade, it would likely downward bias the estimated friend effect. Of course, the opposite could happen whereby individuals encourage struggling friends to withdraw from a class resulting in fewer observed "bad grades" given to those with friends in class, thus biasing the friend effect upwards. In Section 4.1 (Table 9), we provide evidence that friends reduce the probability of withdrawing from a course, which likely downward biases our estimates of the friend effect.

While our preferred econometric specification and robustness checks address many threats to internal validity, we cannot directly control for student-by-class unobserved factors in ε_{ic} that are correlated with the friend-in-class indicator. For example, if friends select into a class because they know the instructor's teaching style will be particularly beneficial to them, we may mistake a friend effect for what is really selection on unobservables. As a final approach to dealing with these issues, we employ an estimator designed by Oster (2017) who extends the work of Altonji (2005, 2008) on identifying the extent of bias caused by selection on unobservables. Define W_{ic} as all of the unobserved variables that are jointly correlated with $Friend_{ic}$ and Y_{ic}. Further, define δ as the magnitude of sorting on W_{ic} relative to the sorting on all the observed variables in Eq. (2). Oster (2017) derives that, under some restrictive assumptions,[18] the adjusted value of β^* in Eq. (2)—that is the coefficient on $Friend_{ic}$ that we would have estimated had we been able to control for W_{ic}—is a function of the slope estimate β^0 and the R-squared (R^0) of a null model regressing Y_{ic} on only $Friend_{ic}$; the observed coefficient of *Friend* (β) and R-squared ($\tilde{R}$) of the model in Eq. (2); and the maximum possible R-squared from a model predicting Y_{ic}, R_{max}.[19] Specifically, Oster shows:

$$\beta^* = \beta - \delta\left(\frac{R_{max} - \tilde{R}}{\tilde{R} - R^0}\right)(\beta^0 - \beta) \quad (3)$$

In Eq. (3), δ represents the relative importance of non-random sorting on unobservables. A value of $\delta = 1$ represents the case where the influence of unobservables on Y_{ic} is equal to that of the included explanatory variables and fixed effects. In Section 5.2, we solve Eq. (3) for the δ that causes β^* to equal zero, a solution which demonstrates how much unobserved, non-random selection is required for friendships to have no impact on academic outcomes. To provide an idea of how sensitive our results are to non-random selection, we also solve for β^* when $\delta = 1$, a situation that Oster suggests may be an appropriate upper bound on the extent of bias caused by omitting the unobserved selection variables.

[14] Students may list up to two academic areas of interest on their application for admission to the university. These areas of interest generally align with academic departments. When a student enrolls in a course in an academic department that aligns with a stated area of interest, we assign the variable *class in interest* a value of one, zero otherwise.

[15] Clustering at the student, instructor, class, or quarter levels produce very similar standard errors.

[16] This empirical specification is nearly identical to Fairlie et al. (2014) who examine the role of teacher-student race match on community college outcomes.

[17] Since friendship is a function of dining hall use, the student fixed effect also controls for the possibility that students who use the dining hall more often, and thus are observed having more friends, are fundamentally different from students who use the dining hall less often.

[18] The most restrictive of these assumptions is that the relative contribution of each component of W_{ic} must be the same as their contribution to $Friend_{ic}$.

[19] For all estimations of Eq. (3), we set $R_{max} = 1$ which assumes that all of GPA is explained by observables and unobservables. This assumption is the most conservative we can make and will result in larger deviations of β from β^*.

Table 7
Estimated Effects of Having a Friend in Class on Grades

	(1)	(2)	(3)	(4)	(5)
friend	0.161***	0.149***	0.144***	0.116***	0.111***
	(0.007)	(0.009)	(0.012)	(0.007)	(0.008)
attempted credits	0.047***	0.038***	-0.006***	-0.008***	
	(0.002)	(0.002)	(0.002)	(0.002)	
prior department credits	-0.009***	0.013***	-0.027***	0.004***	0.007***
	(0.001)	(0.002)	(0.002)	(0.001)	(0.002)
class in interest	0.225***	0.063***	0.151***	0.090***	0.080***
	(0.009)	(0.010)	(0.012)	(0.009)	(0.009)
SAT (100s)	0.146***	0.142***			
	(0.002)	(0.004)			
student fixed effect			✓	✓	
student-quarter fixed effect					✓
class fixed effect		✓		✓	✓
Observations	127,673	125,671	127,867	125,835	124,997
R^2	0.063	0.318	0.469	0.670	0.780

Notes: Each column represents a separate regression. The dependent variable is the student's course grade points. Standard errors (in parentheses) are corrected for two-way clustering at the student and class level.
*p<0.1, **p<0.05, ***p<0.01

3.2.2. Simultaneity

If strong (or weak) class performance fosters friendships among classmates, then OLS estimates of the friend effect will be biased. We address this problem by observing the timing of friendship formation. Specifically, we estimate the friendship effect for friendship dyads that form prior to enrolling in a course together in Section 5.2 (Table 12, and Appendix Tables 2 and 3). Assuming current course grades do not impact prior friendship development, this approach eliminates simultaneity concerns in these estimates.

3.2.3. Measurement error

Recall that friendships are determined by two criteria from the dining data: a time window to determine whether a meeting occurred, and a meeting threshold to determine whether a pair of students are designated as friends. This raises two potential sources of measurement error. First, 9% of freshmen do not use the dining hall and thus are assigned zero friends. As demonstrated in Appendix Table 1, our results are essentially unchanged when we drop these students from our analysis. Second, if the criteria for identifying friendships are either too restrictive or too permissive, a downward bias will be introduced into any estimate of a positive friend effect. To see this, it is convenient to describe taking a class with a friend as a "treatment". Criteria that are too restrictive will filter out friendships, incorrectly placing treated observations into the control group, causing the estimated differences between the control and treated groups to be smaller. Criteria that are too permissive will incorrectly identify some observations as treated and assign friendships to pairs who just happen to dine at similar times and locations. Again, this will lead to an estimated difference between the treatment and control group that is too small because non-friends are captured in the treatment group. While there is no obvious "right" criteria for designating friendships, we show in Section 5.1 that our results are robust to a wide range of criteria.

4. Results

4.1. Main results

We estimate Eqs. (1) and (2) by ordinary least squares, adjusting the standard errors for two-way clustering at the student and class level. Table 7 presents the regression results for the entire sample of freshmen.

Table 8
Estimated Friend Effects using Alternate Measures of the Presence of Friends in Class

	(1)	(2)	(3)	(4)
friend	0.116***	0.102***	0.102***	
	(0.007)	(0.008)	(0.008)	
2 or more friends		0.079***	0.081***	
		(0.015)	(0.016)	
3 or more friends			-0.010	
			(0.028)	
number of friends				0.080***
				(0.005)
Observations	125,835	125,835	125,835	124,997
R^2	0.670	0.670	0.670	0.780

Notes: Each column represents a separate regression. The dependent variable is the student's course grade points. All regression models contain student and class fixed effects as well as student-quarter level controls for attempted credits and prior department credits and a student-course level binary indicator for whether the course is in the academic department of the student's program interest. Standard errors (in parentheses) are corrected for two-way clustering at the student and class level.
*p<0.1, **p<0.05, ***p<0.01

The first column of Table 7 gives estimates of the coefficients in Eq. (1), and show that having at least one friend in class raises grades by 0.161 grade points (roughly 0.161 standard deviations), or about one-half of the difference between a B and B+. For context, this equates to increasing SAT scores by about 110 points (100 × 0.161/0.146). The second column of Table 7 adds class fixed effects and the third column includes student fixed effects (and removes the collinear SAT scores).[20] Our results show that the addition of each fixed effect only slightly attenuates the estimated friend effect. We simultaneously include both student and class fixed effects in our preferred model from Eq. (2) in column 4 and find that the presence of a friend raises grades by 0.116 grade points. The final column of Table 7 replaces student fixed effects with student-by-quarter fixed effects and drops the collinear variable *attempted credits* (since they do not vary within quarter).[21] The result is similar to our preferred specification: friends positively impact grades by slightly more than one-tenth of a grade point. The similarity between this coefficient and that in our preferred specification suggests that there is little concern about any endogenous selection of friends based upon unobserved student characteristics that vary from quarter-to-quarter.

The other estimated coefficients from Table 7 are consistent with our expectations. Students perform better in classes that are within their interests. The coefficient on *attempted credits* is negative in specifications that include a student fixed effect, and positive otherwise, suggesting that while better students enroll in more credits, individual students get lower grades in quarters when they attempt more credits than their typical enrollment. The coefficient on *prior department credits* is positive in specifications that include the class fixed effect. As students take more classes in a department, they gain department experience and subject-specific knowledge and perform better relative to other students in the class.

Our preferred empirical specification with a single binary indicator for at least one friend in class is motivated by the fact that among students with a friend in class, over 80% have exactly one friend. We do, however, allow the friend effect to vary by the number of friends in class

[20] The sample size changes slightly from column to column for two reasons. First, a small fraction of students are exempt from the requirement to report SAT scores and are dropped from Eqs. (1) and (2). Second, singleton observations that are collinear with the fixed effects are dropped. Estimates from a consistent sample across specifications are nearly identical to those presented in Table 7 and are available upon request.

[21] Standard errors in this specification are corrected for two-way clustering at the class and student-quarter levels.

Table 9
Estimated Effects of Having a Friend in Class on the Grade Distribution and Withdrawals

	Outcome A grade	B grade	C grade	D or F grade	Withdraw
friend	0.036***	0.004	-0.015***	-0.024***	-0.011***
	(0.004)	(0.005)	(0.004)	(0.002)	(0.001)
Sample mean of dependent variable	0.345	0.372	0.190	0.092	0.025
Observations	126,135	126,135	126,135	126,135	129,343
R^2	0.553	0.286	0.331	0.435	0.237

Notes: Each column represents a separate regression. The dependent variable is a binary indicator for each grade category in the first four columns and an indicator for course withdrawal in the last column. All regression models contain student and class fixed effects as well as student-quarter level controls for attempted credits and prior department credits and a student-course level binary indicator for whether the course is in the academic department of the student's program interest. Standard errors (in parentheses) are corrected for two-way clustering at the student and class level.
*p<0.1, **p<0.05, ***p<0.01

To understand how the estimated impact of having a friend in class affects the distribution of grades, we estimate linear probability models with different letter grades as the outcome using our preferred specification in Eq. (2). Table 9 presents the results of four regressions estimating the probability of an A grade in first column, B grade in the second column, C grade in the third column, and D or F grade in the fourth column. The results indicate that friends in class increase the probability of A grades by 3.6 percentage points and decrease the probability of C grades and D/F grades by 1.6 percentage points and 2.4 percentage points, respectively. Relative to the sample means for grade category, these represent relatively large shifts at the top and bottom of the grade distribution: A-grades increase by 10.4 percent (0.036/0.345) and D/F grades decrease by 26.1 percent (0.024/0.092).

To test whether students' decisions to remain in a course are affected by the presence a friend, which may bias the OLS estimates, we estimate a linear probability model of student course withdrawal in the last column of Table 9. The results show that taking a course with a friend decreases the probability of withdrawing by 1.1 percentage points. Since roughly 2.5 percent of students withdraw from a course, this represents a rather substantial 44 percent decrease. Given that friends appear to help each other complete courses, it is likely that our estimate of the effect of friends on grades understates the true academic impact.

Table 10
Estimated Effects of the Number of Friends in Class on Grades, by Student and Friend Characteristics

		Student Characteristic							
		Panel A: Gender					**Panel B:** Ability			
Friend Characteristic		All	(a) Male	(b) Female	**(a) – (b)**		All	(a) High SAT	(b) Low SAT	**(a) – (b)**
	All	0.080***	0.098***	0.064***	**0.034*** **	All	0.080***	0.070***	0.087***	**– 0.016**
		(0.005)	(0.007)	(0.007)			(0.005)	(0.007)	(0.007)	
	(i) Male	0.089***	0.098***	0.058***	**0.040****	(i) High	0.071***	0.055***	0.094***	**–**
		(0.008)	(0.009)	(0.014)		SAT	(0.008)	(0.010)	(0.011)	**0.039*****
	(ii)	0.072***	0.098***	0.065***	**0.033****	(ii) Low	0.087***	0.093***	0.084***	**0.009**
	Female	(0.007)	(0.013)	(0.098)		SAT	(0.007)	(0.092)	(0.010)	
	(i) – (ii)	**0.017***	**0.000**	**– 0.007**		**(i) – (ii)**	**– 0.016**	**– 0.037****	**0.010**	
		Panel C: Gender and Friendship Strength					**Panel D**: First-Generation Status			
		All	(a) Male	(b) Female	**(a) – (b)**		All	(a) 1st Gen.	(b) Not 1st Gen	**(a) – (b)**
	All	0.080***	0.098***	0.064***	**0.034*** **	All	0.080***	0.096***	0.073***	**0.022****
		(0.005)	(0.007)	(0.007)			(0.005)	(0.007)	(0.006)	
	(i)	0.097***	0.103***	0.090***	**0.013**	(i) 1st Gen.	0.073***	0.088***	0.064***	**0.023**
	Strong	(0.006)	(0.009)	(0.009)			(0.009)	(0.015)	(0.011)	
	(ii)	0.055***	0.090***	0.028***	**0.062*****	(ii) Not 1st	0.083***	0.100***	0.077***	**0.023***
	Weak	(0.008)	(0.012)	(0.010)		Gen.	(0.006)	(0.010)	(0.007)	
	(i) – (ii)	**0.042*****	**0.013**	**0.062*****		**(i) – (ii)**	**– 0.010**	**– 0.013**	**– 0.012**	

Notes: Each panel contains estimated coefficients for number of friends in class interacted with various student and friend characteristics. Estimates are derived from four separate regressions and are grouped accordingly. The dependent variable is the student's course grade points. All regression models contain student and class fixed effects as well as student-quarter level controls for attempted credits and prior department credits and a student-course level binary indicator for whether the course is in the academic department of the student's program interest. Standard errors (in parentheses) are corrected for two-way clustering at the student and class level. "High" and "Low" SAT scores represent scoring above and below the median SAT score of the sample, respectively. "1st Gen." designates first-generation college students. "Strong" indicates friendships with at least 20 dining hall meetings in the academic quarter. The last column and row of each panel is a Wald test for the difference in coefficients by student and friend characteristics.

in Table 8. The first column of Table 8 contains the estimated friend effect from the preferred specification (column 4 of Table 7) for reference. The second column adds an indicator variable for two or more friends in class and the third column adds an indicator for three or more friends in class. The results suggest that there are diminishing returns to additional friends in class: one friend increases grades by 0.102 grade points, a second friend increases grades by an additional 0.081 grade points, and further friends after the second have no impact. To simplify later analyses of heterogeneous treatment effects, we also include a linear measure of the number of friends in class in column 6. The marginal impact of a friend in this specification is 0.080 grade points (or 0.080 standard deviations), which is similar with results presented by Lavy and Sand (2019) who estimate that each reciprocal friend in class raises middle school test scores by 0.098 standard deviations.

4.2. Heterogeneity

Several natural follow-up questions arise from our central finding that friends in class have a strong, positive influence on grades. For instance, how does this effect vary by the characteristics of either the student, friend, or class? Through what mechanism exactly do friends exert their positive influence? Although we are unable provide a definitive answer to the last question, answers to the question about variation in the friend effect by student, friend, and class characteristics provide some insight.

We concentrate our student and friend heterogeneity analysis on four different characteristics: gender, SAT score, strength of friendship, and first-generation status. We designate SAT scores as "low" or "high" based on whether they are below or above the median SAT score, respectively,

among all freshmen in our data. We define strong friendships as those with at least 20 dining hall meetings in the same quarter while enrolled in class together.[22] Lastly, first-generation status identifies students whose parents never obtained a four-year post-secondary degree and serves as a proxy for socioeconomic background.

For each of these characteristics, we conduct analyses based on 1) the student characteristic; 2) the friend characteristic; and 3) the interaction between student and friend characteristics. To facilitate interpretation, we concentrate on the linear measure of friends (number of friends in class) interacted with relevant student and friend traits. Table 10 presents these estimates organized in four panels which each contain estimates from four regressions: the three categories described above and the overall estimate of the marginal effect of a friend from the last column of Table 8 for reference. We also include a Wald test for the difference in coefficients by student and friend characteristics in the last column and row of each panel of Table 10. Importantly, friends have a positive and statistically significant effect on grades regardless of the student or friend characteristic under consideration in Table 10, implying that many subgroups of students benefit from taking classes with friends.

The estimates of heterogenous treatment effects by gender are provided in Panel A of Table 10. Although both men and women benefit from friends, the first row shows that men benefit more: the impact of a friend on men's grades is 0.098 grade points while the impact of a friend on women's grades is 0.064 grade points, and this difference is significant at the 1% level. The first column of Panel A shows that male friends have slightly larger impact (significant at the 10% level) than a female friend does—a male friend raises their friend's grades by 0.089 grade points as opposed to a female friend's effect of 0.072. The interior rows and columns of Panel A show the impact of same- and opposite-gender friends. For men, the impact of a same-gender or opposite-gender is nearly identical. Women appear to benefit slightly more from a female friend than a male friend, though the difference is not statistically significant.

Panel B repeats this exercise with student ability as measured by SAT scores. Similar to the gender analysis, both low- and high-ability students benefit from friends. Some interesting patterns emerge when interacting student and friend traits. First, high-ability students benefit more (0.037 grade points) from a low-ability friend compared to a high-ability friend in class. For low-ability students, the marginal effect of a friend is similar regardless of the friend's ability. These results suggest that the greatest improvements occur for friendships pairs consisting of students with different ability.

We allow the friend effect to vary by strength of the friendship and student gender in Panel C. Note that while "strong" and "weak" are listed as friend characteristics in Table 10, they are actually characteristics of the friendship since all measured friendships are symmetric. Overall, strong friendships exhibit nearly twice the positive impact of weak ones. This result is driven entirely by women who experience a threefold increase in the measured friend effect when a stronger friend is present and provides evidence of differences in the effect of social relationships across genders.

The first row of Panel D shows that first-generation students benefit more (0.022 grade points) than non-first-generation students from having a friend in class. Comparing the second and third rows of Panel D, we generally find that non-first-generation friends have a larger impact on grades than first-generation friends, though this difference is not statistically significant.[23] This suggests that social interaction between students of lower and higher socioeconomic background may benefit disadvantaged students with no academic detriment to more advantaged students.

Table 11
Estimated Effects of Having a Friend in Class on Grades, by Class Characteristics

	(1)	(2)	(3)
friend	0.116*** (0.007)	0.083***(0.010)	0.122*** (0.010)
friend × *(large class)*		0.051*** (0.014)	
friend × *(difficult class)*			-0.012 (0.014)
Observations	125,835	125,835	125,835
R^2	0.670	0.670	0.670

Notes: Each column represents a separate regression. The dependent variable is the student's course grade points. The *large class* indicator takes the value one if the class has at least 45 students and is zero otherwise. The *difficult class* indicator takes the value one if the class grade point average is less than 2.9 and is zero otherwise. All regression models contain student and class fixed effects as well as student-quarter level controls for attempted credits and prior department credits and a student-course level binary indicator for whether the course is in the academic department of the student's program interest. Standard errors (in parentheses) are corrected for two-way clustering at the student and class level. *p<0.1, **p<0.05, ***p<0.01

As a second way of exploring heterogeneity, Table 11 provides estimates of the binary friendship effect from Eq. (2) by two class characteristics: size and difficulty. A class is considered "large" if it enrolls at least 45 students and "difficult" if the class GPA among students without friends is less than 2.9.[24] The results from including an interaction between indicators for each of these groups and the friend indicator from Eq. (2) is given in columns 2 and 3 of Table 11. The estimate of the friend effect from column 4 of Table 7 is provided in the first column for reference. The point estimates indicate that friends have a significantly greater impact in larger classes but no differential impact in more difficult classes, perhaps suggesting a mechanism through which friends improve grades is through mutual instruction (if time with an instructor is harder to come by in larger classes).

Though the estimation output is not presented here, we further investigate mechanisms behind the friend effect by determining whether academic performance decreases in classes taken without friends during the same quarter when a friend is present in at least one class. If so, it may suggest substitution of effort from classes without friends to those with friends. To do this, we create an indicator that takes the value one if a friend is present in another class in the same quarter. We add this friend-in-other-class indicator to Eq. (2) and drop the friend-in-class indicator. This method and the subsequent results suggest that grades do not significantly change in non-friend classes. If effort is zero-sum across classes within a quarter (i.e., additional effort in one class leads to less effort in others), the fact that grades do not decrease in classes without friends suggests that effort is not being substituted across classes and that the observed positive friend effect may therefore be a result of increases in productivity (i.e., joint production). If effort is not zero-sum across classes, then having a friend may increase effort, either because the return to effort has increased (joint production) or because friends decrease the marginal cost of effort (social pressure).

Note that while we might conceptualize productivity as a measure of learning production, we can only directly observe grades. Consequently, in our context "productivity" is technically the marginal impact of effort on grades, rather than learning. Although we find that friends improve grades, we cannot rule out the possibility that friends lead to worse learning outcomes if, for example, division of work leads to less learning. To investigate this, we would need more information about how

[22] 3,226 out of 6,059 friendship pairs (53%) meet at least 20 times within the quarter.

[23] This result is similar to Lavy and Sand (2019) who find that friends with more educated parents have a greater positive impact on educational attainment than friends whose parents have lower levels of educational attainment.

[24] These group designations were chosen to roughly equalize the number of student-course observations in each group.

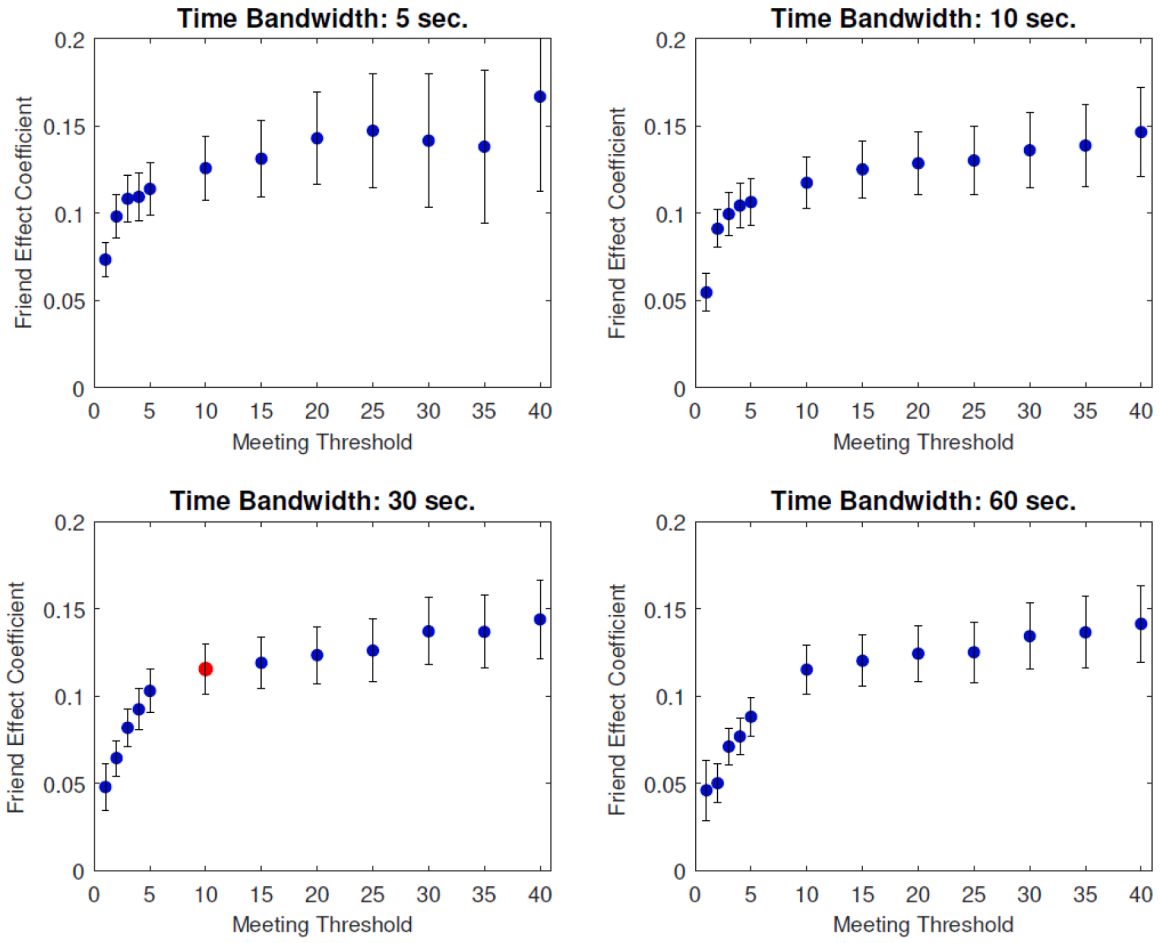

Fig. 1. Estimated Effects of Having a Friend in Class on Grades, by Alternate Friendship Criteria. *Notes:* Each dot represents the point estimate of the friend effect for different friendship criteria (time windows / bandwidths and meeting thresholds) from our preferred specification in column 4 of Table 7. The corresponding bars represent the 95% confidence interval. The red dot represents the default friendship criteria of 10 or more meetings within 30 seconds in a quarter used in the models of Tables 7, 8, 9, 10, 11, and 12.

students interact with their friends in classes as well as measures of learning outcomes. A detailed exploration of the impact of friends on attendance and performance on the various assignments within a class is left for future research.

5. Robustness checks

In this section, we exploit unique features of our data to check the robustness of our results. We first show in Section 5.1 that our findings are generally insensitive to the criteria for assigning friendships. Section 5.2 provides a detailed discussion of possible sources of bias in estimating the friend effect and several tests exploring these. Taken together, these robustness checks mitigate concerns about measurement error, simultaneity and non-random selection.

5.1. *Friendship criteria*

Thus far, we have designated two students as friends if they enter a dining hall within thirty seconds of each other at least ten times during a quarter. Since these criteria are arbitrary, the observed friendship distribution may be measured with error.[25] To test whether our results are generally sensitive to friendship criteria, we plot estimates of the friend coefficient from our preferred specification in Eq. (2) using alternate time windows and meeting thresholds in Fig. 1. Specifically, we illustrate the friend effect using 5-, 10-, 30-, and 60- second time windows, and meeting thresholds ranging from 1 to 40. Fig. 1 shows that we observe a significantly large and positive friend effect using any reasonable criteria for designating students as friends.

Fig. 1 also shows that as the meetings threshold increases, the observed friend effect increases and is simultaneously measured less precisely. If people who meet more often are better friends, this implies that stronger friends have a larger impact on course performance and is consistent with the "strong" versus "weak" friend findings from Table 10. The estimated standard errors increase because fewer observations pass through the higher threshold and we therefore observe fewer friendship dyads from which to estimate the friend effect.

An alternative way to assess social connections among classmates would be to dispense of the meetings threshold entirely. To do this, we construct a new variable that measures the total number of dining hall meetings that each student has with their classmates. We then replace the binary variable *Friend* in Eq. (2) with a fourth-degree polynomial in the total number of dining hall meetings with classmates. Because it is difficult to interpret the coefficients of high-order polynomials, we display the predicted grades (in grade points) in Fig. 2. Predicted grades are calculated at the mean values of the regressors in X_{ic}' in Eq. (2), for zero to 100 meetings with classmates. Note that the graph is increasing and concave indicating that while grades increase when students have more social ties with their classmates, this effect exhibits diminishing marginal returns.

[25] Though as discussed in Section 3.2.3, this measurement error should attenuate our estimates of the positive friend effect.

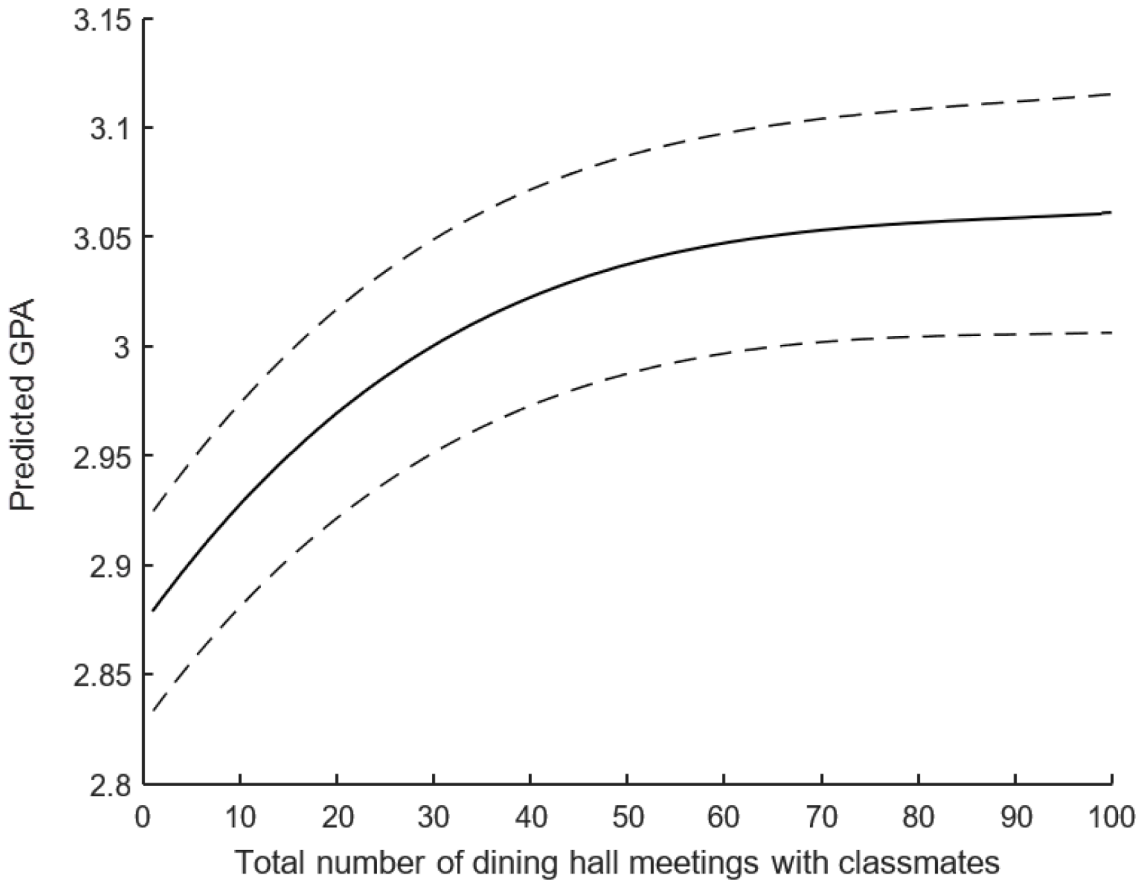

Fig. 2. Estimated Effects of Total Dining Hall Meetings with Classmates on Grades**.** *Notes*: This figure presents predicted grades and the corresponding 95% confidence interval obtained from estimates from the preferred specification in column 4 of Table 7 where the friend indicator has been replaced with a fourth-degree polynomial for the total number of dining hall meetings (using a 30-second time bandwidth) with classmates in a quarter. Predicated grades are calculated at the mean values of the regressors in $X_{ic}^{'}$ in Eq. (2), for zero to 100 meetings with classmates.

One last concern has to do with the potential undercounting of friendships. Our sample of freshmen introduces this possibility. Specifically, our approach utilizes all freshmen observations, including those who do not live on campus and, hence, do not have dining hall privileges. Because these individuals are never observed entering the dining hall, they are assigned zero friends. However, these individuals likely form unobserved friendships and, in the terminology of treatment and control groups, are incorrectly assigned to the control group. This problem is mitigated somewhat by the fact that 86% of freshmen observations eat in the dining hall each quarter, and so the number of false zeros is likely small. To address this issue, recall that we estimate all regressions from Table 7 in Appendix Table 1, but exclude from the sample all freshmen without a dining hall plan. Using our preferred specification, a friend in class raises GPA by 0.115 grade points, almost identical to that reported in column 4 of Table 7.

Table 12
Estimated Effects of Having a Friend in Class on Grades: Robustness

	(1)	(2)	(3)	(4)
friend	0.116***			0.117***
	(0.007)			(0.007)
past friend		0.107***		
		(0.009)		
strictly past friend			0.067***	
			(0.013)	
friend – same course and instructor, different section				0.112***
				(0.015)
friend – same course, different instructor				0.005
				(0.010)
Observations	125,835	83,318	83,318	125,835
R^2	0.670	0.701	0.700	0.670

Notes: Each column represents a separate regression. The dependent variable is the student's course grade points. All regression models contain student and class fixed effects as well as student-quarter level controls for attempted credits and prior department credits and a student-course level binary indicator for whether the course is in the academic department of the student's program interest. The variable *past friend* indicates the presence of at least one student in class who was a friend in a past quarter and *strictly past friend* identifies the presence of at least one student in class who was a friend in a past quarter but is not in the current quarter. Standard errors (in parentheses) are corrected for two-way clustering at the student and class levels.
*p<0.1, **p<0.05, ***p<0.01

5.2. Endogenous friend formation

We now address potential bias in estimates of the friend effect that result from endogeneity, specifically due to simultaneity and non-random selection of students to classes. As first described in Section 3.2.2, we address simultaneity, or higher grades causing friendships instead of friendships causing higher grades, by leveraging the timing of friendships. To rule out simultaneity, we estimate the effect of past friends on current grades on the assumption that current grades cannot affect past friendship formation. To do this, we introduce two new binary measures of friendship: *past friend* and *strictly past friend*. The variable *past friend* indicates the presence of at least one student in class

who was a friend in a past quarter and *strictly past friend* identifies the presence of at least one student in class who was a friend in a past quarter but is not a friend in the current quarter. Because these measures rely on observing friends in prior quarters, we drop observations from students' first quarter on campus (typically the Fall quarter). We return to our preferred friendship criteria of at least 10 meetings and a 30-second window for identifying these new variables.

The second and third columns of Table 12 presents estimates of the coefficients on the indicators of past friendship, with the estimate of the binary (contemporaneous) friend effect from column 4 of Table 7 provided in the first column for reference.[26] Importantly, friendships formed in the past (*past friend*) have a positive impact on grades (0.107 grade points) and the estimate is nearly as large as the contemporaneous friend effect from the preferred specification (0.116 grade points). This resolves concerns that the friend effect is entirely a product of reverse causality as it is unlikely that current course grades impact previous dining hall choices. Limiting the measure of friendships to those formed strictly in the past reduces the estimated effect size to 0.067 grade points. While this measured effect is significantly smaller than the baseline treatment estimate of 0.116 grade points, it is reasonable to presume that strictly past friends are also weaker friends, and hence produce a smaller treatment effect.

Another potential source of bias is student-by-course level unobservable factors that affect both taking a class with a friend and student performance. Our identification strategy in Eq. (2) is to use student and classroom fixed effects to control for unobserved factors at the student and classroom level, respectively, but this does not preclude a student-by-course unobserved factor, like enthusiasm for the subject, from driving both in-class friendships and grades.[27] This could occur if students befriend and take classes with others who will turn out to perform well in the same course subjects. To address this type of potential bias, we leverage the fact that WWU often offers multiple sections of the same course each quarter with some sections taught by the same instructor. For example, in the fall of 2019, five instructors taught eight sections of Principles of Microeconomics, three of whom taught two sections each. If what is driving our estimated friend effect is not friend interaction but rather that friends happen to do well in the same subjects, then we should observe a positive effect on grades even when friends do not share the same course instructor, where collaboration among friends is likely more difficult than for friends who share the same instructor.

The last column of Table 12 presents estimates of the friend effect across class sections within the same course. We use three indicator variables to denote categories of friends within a course: friends in the same classroom (*Friend*), friends in different sections that share the same instructor and course, and friends with different instructors in the same course. The results indicate that friends influence grades when they share an instructor and the influence is nearly identical whether students are in the same class (0.117 grade points) or not (0.112 grade points). These results make intuitive sense as sections of the same course taught by the same instructor likely cover nearly identical material, allowing for collaboration between friends across sections.

Importantly, there is no estimated friend effect across instructors, suggesting that friends interacting to produce better outcomes, and not student-by-course omitted variables, is what is driving our primary results.

As a final check on bias due to selection on unobservables, we follow the technique outlined by Oster (2017) to estimate the friend effect under varying levels of non-random selection. Using our preferred specification, we first estimate Eq. (3) for cases where the true effect of friends on GPA, β^*, is equal to zero. This occurs when $\delta = 2.48$ (95% CI), indicating that if non-random selection bias has almost two-and-one-half times the magnitude of the observed explanatory variables (including the fixed effects), then friends would have no effect on GPA. Given the importance of the explanatory variables (we estimate an R^2 of 0.67 is our preferred model), we find this an unlikely scenario. A more likely case would be if the unobserved variables had an effect equal to that of the included variables. Using Eq. (3), when $\delta = 1$, the estimated $\beta^* = .092$, which, while less than the estimate from our preferred specification in column 4 of Table 7, is still quite large and of practical significance. Taken as a whole, this approach suggests that bias due to non-random selection and omitted variables is likely to be small.

6. Conclusion

Using a revealed-preference based method of determining friendships, this paper documents that students who take college-level courses with friends earn significantly higher grades than they do when enrolled in courses without friends. This increase in grades is significant both statistically and practically, with the impact on grades being about one-half of the difference between an A- and an A, or equivalently, about the same magnitude as increasing a student's SAT score by about 100 points. This effect occurs in the presence of student, student-quarter, and classroom fixed effects suggesting that these results are not driven by unobservables at any of these levels. This effect increases with the strength of friendship (measured by the number of dining hall meetings) and the number of friends in a course. The friend effect is also present for friendships that were formed prior to course enrollment indicating that these results are not driven by the simultaneous formation of friendships and academic performance. Further, the effect of having a friend in the same course taught by a different professor is statistically no different than zero implying that our measure of friendship does not serve as a proxy for shared enthusiasm for a subject or another student-by-course unobservable factor. Moreover, there are reasons to believe that we have understated the true friendship effect. For instance, we document that enrolling in a class with a friend reduces a student's likelihood of dropping the course or of earning a D or F. Under these circumstances, friendships increase course completion and may therefore reduce the time and costs of earning a degree. Taken as a whole, it appears that friendships play an important role in academic achievement and, through that, in human capital formation.

As pointed out in the introduction, there is a vast literature on peer effects and their role in academic achievement. However, the literature on friendships is more sparse, partly because of its reliance on self-reported friendship data. In addition, the analysis of friendship in the academic literature has been limited to elementary and secondary aged children. Ours is the first to document the role of friendships on academic outcomes among college-aged adults. Another important finding in this study is that we measure a positive friend effect regardless of the student or friend's characteristics, suggesting the unconditional importance of friendships in enhancing academic achievement. As might be expected, there are similarities and differences between the extant literature and our findings. For instance, Lavy and Sand (2019) estimate the marginal effect of a friend in class on middle school test scores to be 0.098 standard deviations, which is comparable to our estimate of 0.080 standard deviations. Lavy and Sand also find that the education level of a friend's parents has a positive impact on academic achievement among middle school students. We estimate qualitatively similar results: friend effects are larger when the friend comes from a family whose parents completed a college education compared to first-generation friends, though the difference is not statistically significant. When examining high school students, Hill (2015) finds that opposite gender friends in high school decrease own GPAs which seemingly stands in contrast to our findings that both same and opposite gender friends in class have a consistent, positive effect on grades. However, this difference could simply be due to measuring the effect of overall friendship network characteristics versus the effect of friend characteristics in the same

[26] Appendix Tables 2 and 3 present full results for all the specifications of Table 7 using these two alternative measures of friendship.

[27] Note that this would only be an issue if enthusiasm is not well captured by our department-level controls for interest in Eq. (2).

Table A1
Estimated Effects of Having a Friend in Class on Grades, Only Dining Hall Users

	(1)	(2)	(3)	(4)	(5)
friend	0.160***	0.143***	0.145***	0.115***	0.110***
	(0.007)	(0.009)	(0.012)	(0.007)	(0.008)
attempted credits	0.047***	0.036***	-0.006***	-0.010***	
	(0.002)	(0.002)	(0.002)	(0.002)	
prior department credits	-0.009***	0.012***	-0.028***	0.003***	0.006***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
class in interest	0.218***	0.073***	0.166***	0.102***	0.091***
	(0.009)	(0.011)	(0.013)	(0.009)	(0.010)
SAT (100s)	0.144***	0.141***			
	(0.002)	(0.004)			
student fixed effect			✓	✓	
student-quarter fixed effect					✓
class fixed effect		✓		✓	✓
Observations	113,944	112,017	113,955	112,011	111,327
R^2	0.062	0.321	0.465	0.670	0.778

Notes: The sample of students is all freshmen observed using the dining hall at least once from Fall 2013 to Spring 2018. Each column represents a separate regression. The dependent variable is the student's course grade points. Standard errors (in parentheses) are corrected for two-way clustering at the student and class level.
*p<0.1, **p<0.05, ***p<0.01

class.

One open question in this research has to do with the mechanism by which friendship leads to higher grades. Certainly one can imagine friends aiding and encouraging each other, by serving as insurance for each other in cases of missed course material, or by eliciting more effort from an individual. While we cannot identify the specific mechanism, having a friend in one class does not appear to decrease grades in other classes thus ruling out mechanisms that involve substituting effort or resources between classes.

Given that friends are strongly associated with better grades, it is worth commenting on policies that might encourage friendships. Universities have long created structures and programs to increase social interactions among their students. Indeed, WWU has a number of programs that encourage friendships even before students set foot on campus as well as social events that take place in residence halls throughout the year. Further, faculty can create class environments that encourage social interaction. To the extent that these programs enable friendships to form, they may have indirect academic benefits that outweigh their costs. Of course universities are also expanding programs in which it is difficult to have friendship support—most notably in online courses. Our read of the online course literature suggests that academic outcomes in these courses tend to be lower than face-to-face courses and our research suggests an additional reason for this: less meaningful social interactions might disadvantage students in online courses.

CRediT authorship contribution statement

Darius D. Martin: Conceptualization, Software, Formal analysis, Data curation, Visualization, Writing - original draft, Writing - review & editing. **Adam C. Wright:** Conceptualization, Methodology, Formal analysis, Writing - original draft, Writing - review & editing. **John M. Krieg:** Conceptualization, Software, Formal analysis, Data curation, Supervision, Writing - original draft, Writing - review & editing.

Declarations of interest

none.

Table A2
Estimated Effects of Having a Past Friend in Class on Grades

	(1)	(2)	(3)	(4)	(5)
past friend	0.158***	0.121***	0.169***	0.107***	0.107***
	(0.016)	(0.010)	(0.014)	(0.009)	(0.010)
attempted credits	0.049***	0.042***	-0.008***	-0.008***	
	(0.002)	(0.002)	(0.002)	(0.002)	
prior department credits	-0.013***	0.013***	-0.031***	0.007***	0.007***
	(0.002)	(0.002)	(0.002)	(0.001)	(0.002)
class in interest	0.219***	0.044***	0.153***	0.080***	0.070***
	(0.014)	(0.013)	(0.013)	(0.011)	(0.011)
SAT (100s)	0.137***	0.131***			
	(0.004)	(0.003)			
student fixed effect			✓	✓	
student-quarter fixed effect					✓
class fixed effect		✓		✓	✓
Observations	84,707	83,244	84,815	83,318	82,820
R^2	0.060	0.320	0.512	0.701	0.778

Notes: A *past friend* is a student observed as a friend in one prior quarter. The sample includes all freshmen observed using the dining hall at least once from Fall 2013 to Spring 2018. Each column represents a separate regression. The dependent variable is the student's course grade points. Because we cannot observe prior friendships in the first quarter students are on campus, we drop all student-course observations from the Fall quarter. Standard errors (in parentheses) are corrected for two-way clustering at the student and class levels.
*p<0.1, **p<0.05, ***p<0.01

Table A3
Estimated Effects of Having a Strictly Past Friend in Class on Grades

	(1)	(2)	(3)	(4)	(5)
strictly past friend	0.106***	0.038**	0.131***	0.067***	0.071***
	(0.021)	(0.015)	(0.019)	(0.013)	(0.014)
attempted credits	0.049***	0.042***	-0.008***	-0.008***	
	(0.002)	(0.002)	(0.002)	(0.002)	
prior department credits	-0.014***	0.013***	-0.032***	0.006***	0.007***
	(0.002)	(0.002)	(0.002)	(0.001)	(0.002)
class in interest	0.210***	0.042***	0.148***	0.080***	0.069***
	(0.014)	(0.014)	(0.013)	(0.011)	(0.009)
SAT (100s)	0.136***	0.130***			
	(0.004)	(0.003)			
student fixed effect			✓	✓	
student-quarter fixed effect					✓
class fixed effect		✓		✓	✓
Observations	84,707	83,244	84,815	83,318	82,820
R^2	0.058	0.318	0.510	0.700	0.777

Notes: A *strictly past friend* is a friend in a prior quarter, but not in the current quarter. The sample includes all freshmen observed using the dining hall at least once from Fall 2013 to Spring 2018. Each column represents a separate regression. The dependent variable is the student's course grade points. Because we cannot observe prior friendships in the first quarter students are on campus, we drop all student-course observations from the Fall quarter. Standard errors (in parentheses) are corrected for two-way clustering at the student and class levels.
*p<0.1, **p<0.05, ***p<0.01

Acknowledgements

This paper has greatly benefited from helpful comments from Dick Startz, Dan Goldhaber, Roddy Theobald, Dan Black, Peter Kuhn and participants at the following conferences and seminars: 2020 Association for Education Finance and Policy Conference, Center for Education Data & Research at the University of Washington, and Western Washington University. The authors declare that they have no known competing financial interests or personal relationships that could have

appeared to influence the work reported in this paper. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Appendix

Table A1, Table A2, Table A3

References

Altonji, J. G., Elder, T. E., & Taber, C. R. (2005). Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools. *Journal of Political Economy, 113*(1), 151–184.

Altonji, J. G., Elder, T. E., & Taber, C. R. (2008). Using Selection on Observed Variables to Assess Bias from Unobservables when Evaluating Swan-Ganz Catherization. *American Economic Review, 98*(2), 345–350.

Block, P., & Grund, T. (2014). Multidimensional Homophily in Friendship Networks. *Network Science, 2*(2), 189–212.

Bramoullé, Y., Djebbari, H., & Fortin, B. (2009). Identification of peer effects through social networks. *Journal of Econometrics, 150*(1), 41–55.

Carrell, Scott E., Fullerton, Richard L, & West, James E. (2009). Does Your Cohort Matter? Measuring Peer Effects in College Achievement. *Journal of Labor Economics, 27*(3), 439–464.

Epple, D., & Romano, R. (2011). Peer Effects in Education: A Survey of the Theory and Evidence. In J. Benhabib, A. Bisin, & M. Jackson (Eds.), *In Handbook of Social Economics, 1A* pp. 1053–1163). Amsterdam: North-Holland. edited by.

Fairlie, R. W., Hoffman, F., & Oreopoulos, P. (2014). A community college instructor like me: Race and ethnicity interactions in the classroom. *American Economic Review, 104* (8), 2567–2591. https://doi.org/10.1257/aer.104.8.2567.

Fischer, C. S. (1977). *Networks and Places: Social Relations in the Urban Setting*. New York: Free Press.

Fischer, S. (2017). The Downside of Good Peers: How Classroom Composition Differentially Affects Men's and Women's STEM Persistence. *Labour Economics, 46*, 211–226.

Foster, Gig (2006). It's Not your Peers, and it's Not your Friends: Some Progress Toward Understanding the Educational Peer Effect Mechanism. *Journal of Public Economics, 90*(8), 1455–1475.

Fletcher, Jason M., Ross, Stephen L., & Zhang, Yuxiu (2020). The Consequences of Friendships: Evidence on the Effect of Social Relationships in School on Academic Achievement. *Journal of Urban Economics*, 116.

Goodreau, S. M., Kitts, J. A., & Morris, M. (2009). Birds of a feather, or friend of a friend? Using exponential random graph models to investigate adolescent social networks. *Demography, 46*(1), 103–125.

Hill, A. (2015). "The Girl Next Door: The Effect of Opposite Gender Friends on High School Achievement. *American Economic Journal: Applied Economics, 7*(3), 147–177.

Kalmijn, M. (1998). Intermarriage and homogamy: causes, patterns, trends. *Annual Review of Sociology*, (24), 395–421.

Knecht, A., Snijders, T. A. B., Baerveldt, C., & Steglich, C. E. G. (2010). Friendship and Delinquency: Selection and Influence Processes in Early Adolescence. *Social Development, 19*(3), 494–514.

Lam, C.T. 2012. "Estimating various kinds of peer effects on academic performance", Working Paper, Clemson Department of Economics.

Lavy, V., & Sand, E. (2019). The effect of social networks on students' academic and non-cog- nitive behavioural outcomes: evidence from conditional random assignment of friends in school. *Econ. J., 129*(617), 439–480.

Lin, X. (2010). Identifying peer effects in student academic achievement by spatial autoregressive models with group unobservables. *Journal of Labor Economics, 28*(4), 825–860.

Mora, T., & Oreopoulos, P. (2011). Peer effects on high school aspirations: Evidence from a sample of close and not-so-close friends. *Economics of Education Review, 30*(4), 575–581.

Oster, E. (2017). Unobservable Selection and Coefficient Stability: Theory and Evidence. *Journal of Business and Economic Statistics, 37*(2), 187–204.

Patacchini, E., Rainone, E. and Zenou, Y., 2011. "Dynamic aspects of teenage friendships and educational attainment', Working Paper, CEPR Discussion Papers No. 8223.

Patacchini, E., Rainone, E., & Zenou, Y. (2017). Heterogeneous peer effects in education. *Journal of Economic Behavior and Organization, 134*, 190–227.

Sacerdote, B. (2001). Peer Effects with Random Assignment: Results for Dartmouth Roommates. *The Quarterly Journal of Economics, 116*(2), 681–704.

Sacerdote, B. (2011). Peer effects in education: how might they work, how big are they and how much do we know thus far? In E. A. Hanushek, S. Machin, & L. Woessmann (Eds.), *Handbook of Economics of Education, 3* pp. 249–277) Amsterdam: Elevier Science.

Sacerdote, B. (2014). Experimental and quasi-experimental analysis of peer effects: two steps forward. *Annu. Rev. Econ., 6*(1), 253–272.

Smith-Lovin, L. and J.M. McPherson, 1993. "You are who you know: a network perspective on gender." In England, P. editor, Theory on Gender. Aldine, New York, 1993, pp. 223-2251.

Stinebrickner, Ralph, & Stinebrickner, Todd R. (2006). What Can Be Learned About Peer Effects Using College Roommates? Evidence from New Survey Data and Students from Disadvantaged Backgrounds. *Journal of Public Economics, 90*(8), 1435–1454.

Zimmerman, David J. (2003). "Peer Effects in Academic Outcomes: Evidence from a Natural Experiment. *Review of Economics and Statistics, 85*(1), 9–23.