Quantitative Economics 8 (2017), 239–275 1759-7331/20170239

Who is sitting next to you? Peer effects inside the classroom

Sok Chul Hong
Department of Economics, Seoul National University

Jungmin Lee
Department of Economics, Seoul National University and IZA

We examine college students' interaction within classrooms and estimate peer effects on their academic performance. We exploit a unique seating rule at a university in South Korea, known as the fixed-seat system. We propose a novel identification strategy based on students' repeated interaction. Our findings show that a student's performance in a class is significantly influenced by his or her next-seat neighbor's ability. The effect is heterogeneous, varying by student and class characteristics. Also quantile regressions reveal that peer effects are significant among below-average students and among those at the top end.

Keywords. Peer effect, classroom seating rule, academic performance, college students, IV quantile regression, personality.

JEL classification. C31, I21.

1. Introduction

Identification of peer effects is hindered by many econometric problems, such as selection into peer groups, simultaneity bias known as the reflection problem à la Manski (1993), spillover via unobservables, and fuzzy definition of peers. There have been a variety of attempts to overcome such problems. Some exploit random variation in peer composition arising from natural experiments or unexpected shocks (Hoxby (2000), Angrist and Lang (2004), Vigdor and Nechyba (2007), Ding and Lehrer (2007)) or random assignment of peers (Sacerdote (2001), Zimmerman (2003), Foster (2006), Stinebrickner and Stinebrickner (2006), Kang (2007), Carrell, Fullerton, and West (2009), Duflo, Du-

Sok Chul Hong: sokchul.hong@snu.ac.kr
Jungmin Lee: jmlee90@snu.ac.kr
We would like to thank Joshua Angrist, Stacey Chen, Gigi Foster, Kamhon Kan, Daiji Kawaguchi, Jinyoung Kim, Victor Lavy, Chulhee Lee, Myungjae Lee, Simon Sokbae Lee, Hye Lim Son, Petra Todd, and Hongliang Zhang as well as three anonymous referees and seminar participants at Seoul National University, Sogang University, Korea University, the University of Hong Kong, and the Asian Conference for Applied Microeconomics and Econometrics. We also thank David Autor and Sari Pekkala Kerr for sharing their Matlab code for the Chernozhukov–Hansen IVQR estimator. We thank Soohyun Lee at the Office of Academic Administration at Sogang University for providing the data. Lee's work was supported by the National Research Foundation of Korea Grant funded by the Korean Government (NRF-2013S1A5A8024492). Hong's work was supported by the National Research Foundation of Korea Grant funded by the Korean Government (NRF-2011-327-B00079).

DOI: 10.3982/QE434

pas, and Kremer (2011), Sojourner (2013)).[1] In the absence of such randomness, others have tried to eliminate confounding factors by controlling for a rich set of fixed effects, such as school or school-by-grade fixed effects (Hanushek, Kain, Markman, and Rivkin (2003), Betts and Zau (2004), Arcidiacono and Nicholson (2005)) or individual student and/or teacher fixed effects (Betts and Zau (2004), Burke and Sass (2013)).

In this paper, we would like to advance the literature in two major directions. First, we take a close look inside classrooms and examine students' interaction with their neighboring classmates. To our best knowledge, our study is the first to estimate peer effects within classroom subgroups. Our novel data set from a private university in South Korea allows us to do this. Sogang University has long implemented a peculiar classroom seating policy—the so-called fixed-seat system (FSS). Under the FSS, for each course, once students' seats are assigned, they are required to sit in the same seat throughout the semester. We collect data on students' assigned seats and their final examination scores. The data allow us to identify not only a student's exact seat location for each course but also identify those sitting around the student.

Second, we propose a novel identification strategy of exploiting one's repeated interaction with the same peers. Our identification strategy is applicable to many cases for peers. In schools, students study together with their classmates for months, and dormitory roommates meet every day. In workplaces, co-workers interact in various ways. Our identification strategy is based on the idea that peer effects would change over time as peers interact more. In particular, it is likely that peer effects among classmates increase over time. Classmates may start studying in a group or exchanging information about course material or assignments a few weeks after the start of the semester with peers with whom they were not previously acquainted.[2] Students might also be negatively affected by "bad apples" around them, and such effects may accumulate over time.

To preview our identification strategy, suppose that a student's (A) performance is influenced by his or her peer's (B) ability, that is, there exist ability peer effects, and that the magnitude of influence changes as they repeatedly interact.[3] This implies that not only A's own ability but also B's ability should be included with time-varying coefficients in the educational production function. Our estimation strategy is simply to remove unobservable ability using the data structure where we repeatedly observe students' academic outcomes. As we later show, we can still partially estimate the effect of the peer's ability despite that the unobservable ability is removed, because the effect is time-variant.[4] Specifically, we identify a lower bound of the true ability peer effect by regressing B's lagged outcome on A's final outcome, because B's lagged outcome depends on both A's and B's abilities. The ability peer effect is the effect of B's ability on A's out-

[1]Epple and Romano (2011) and Sacerdote (2011) review the literature on educational peer effects.

[2]We conducted student surveys for 667 students in seven additional classes after collecting the data used in this study. The surveys reveal that about 30% of students studied together with the paired student sitting next to them and that 60% exchanged information about examinations and assignments.

[3]Because ability is only partially observable by the econometrician, the peer effects we examine in this paper use unobservable characteristics.

[4]We would like to note up front that our identification strategy cannot estimate spillovers through time-varying unobservable characteristics, such as motivation.

come, so the effect of B's lagged outcome should be a lower bound of the true effect as long as A's ability is positively correlated with A's own outcome.

Our identification approach is closely related to the recent strand of literature focusing on unobservable peer effects. We share the argument of Fruehwirth (2013, 2014) that it is not evident in the existing literature why peer academic outcomes (e.g., test scores) are directly included in the educational production function. Instead, she proposed that peer effects work via peers' unobservable ability. A similar idea is presented by Arcidiacono, Foster, Goodpaster, and Kinsler (2012), who also proposed a structural production function where peer effects work through their unobservable characteristics. Our approach is similar to these studies in that we also model unobservable peer effects instead of endogenous peer effects. Our approach is different from theirs in that we exploit the feature that the effect of unobservable ability varies over time. This also contrasts with previous panel data approaches that utilize changes in peers' characteristics, for example, unexpectedly more high-achieving students over year and grade. Hanushek et al. (2003) noted that the primary source of variation for identification is the variation in peer characteristics within school and grade owing to student mobility into or out of the school. Our approach does not require peer characteristics to change over time (in our setting, they are fixed because of the fixed-seat system) but that repeated outcomes are observed. In contrast, previous panel data approaches controlling for fixed effects assumed that the effect of unobservable ability is time-invariant (so it can be removed by using fixed effects).

To summarize our findings, we find that a student's course performance is significantly affected by his or her neighboring students, particularly by the person who is seated right next to the student.[5] There is little effect from geographically distant classmates. This implies that peer effects should be underestimated when we define peer groups broadly, because irrelevant peers are included. We also find that peer effects are heterogeneous, depending on student and class characteristics. For example, we find that peer effects exist among male students but not significantly among female students. Peer effects do not exist among economics majors. Furthermore, quantile regressions reveal that peer effects are heterogeneous over the distribution of test scores. Peer effects are significant among below-average students as well as those at the top end. Last, we provide the first bit of evidence that personality may play an intermediary role for peer effects to arise.

The remainder of the paper proceeds as follows. In the next section, we present a brief introduction to the FSS. In Section 3, we describe the data and present the summary statistics of primary variables. In Section 4, we explain our empirical strategy. Section 5 presents the results. Section 6 concludes. Replication files are available in a supplementary file on the journal website, http://qeconomics.org/supp/434/code_and_data.zip.

[5]We present a few classroom photos in Figure A.1. Pair students are physically close and are seated right next to each other.

2. Institutional background: Fixed-seat system

Sogang University is one of the leading research and liberal arts universities in South Korea. The university was established in 1960 by the Society of Jesus to provide an education based on Catholic belief; it was inspired by the Jesuit educational philosophy. Although the university is a Jesuit school, it has accepted about 1650 new students regardless of their religion or religious preference; during the past five decades, the entering students have generally been ranked within the top 1% in terms of the national university-entering test score.

Sogang University has also been distinguished by a course-failure policy—the so-called FA grade system. The policy, which was implemented when the university was established in 1960, is intended to encourage or enforce students' class attendance. The acronym FA is a letter grade that appears on transcripts, which means a course failure due to excessive absences. According to the school's regulations on the FA system, a 3-hour-a-week course (i.e., a three-credit course) allows for up to 6-hours of absences (three late attendances are counted as a 1.5-hour absence). Due to the fact that most courses are taught twice every week, this means that a FA grade is given to a student who has been absent in more than four lectures throughout the semester.

For the FA system to work effectively, student attendance should be taken during every lecture. Apparently, taking attendance in every class is very time-consuming, particularly when the class size is large. This has led to the traditional use of the fixed-seat system at Sogang University. Under the fixed-seat system, students are assigned to seats on a "first-come–first-served" basis on the first day of each course[6] and they are required to have the same seat throughout the semester. Teaching assistants (TA), who are graduate students, make a seating chart at the beginning of the semester. Thus, they can simply take student attendance in every lecture just by checking whether each seat on the chart is taken.

The FA and fixed-seat system has been proved very effective at the university. Above all, the system has resulted in high rates of class attendance. According to the annual statistics reported by the university, the ratio of FA grades was only 1.08% in the 2010 fall semester or 473 out of 43,596 grades. The ratio has been consistently low over the past 10 years: 0.91% in the fall semester of 2000 and 0.63% in the fall semester of 2005.

3. Data collection and description

Our sample is constructed from three data sources.[7] First, we collected data on students' course exam scores and their seat location for 36 courses in economics. The courses were held in the spring and fall of 2010 and in the spring of 2011, of which class sizes were more than 80, taught by tenured or tenure-track professors. Second, we also collected data on the basic individual characteristics from the University Registration System. The variables includes sex, age, university entrance year, major, an indicator as to

[6]We will discuss the potential endogeneity problem due to seat selection later in detail.

[7]We use an auxiliary data set, which contains information on the personality of students aggregated by their sex, major, and entering year. Freshmen take the standard personality test, and the results are published in the university journal. We will explain the data set in more detail in Section 5.4.

whether the student is an exchange student, and an indicator as to whether the student is retaking the course.[8] Third, we also obtained transcript records on all the courses taken by each student in our sample since entering the university as well as their letter grades from the Office of Academic Administration. That is, we know the students' academic performance (i.e., grade point average (GPA)) before they took the courses under study. The final sample includes 4155 student-by-class observations. The sample is large relative to the size of the university, which admits about 1600 new students each year.

From each course's seating chart, we identified each student's exact seat location in terms of row and column and defined his or her classroom "peers." Figure 1 is a seating chart as an example to show how we define classroom peers. First, we define "pairs." When two seats in two successive columns on the same row are adjacent, we treat the two students as a pair. Note that there exists a small passage (aisle) between pairs. In some classrooms, those in the first or last column do not have a pair student. We focus on pairs because presumably, if peer effects do exist, the effects are expected to be stronger between pair students. This is in part because of their proximity and in part because classroom pairs possess some special cultural meanings in Korea. Unlike the United States, in Korea, from elementary to high school, students do not move between classrooms. Instead, teachers go around each classroom and teach. Students are assigned to their own classroom, which is fixed during a school year. Their seats rarely change within the different semesters. Thus, one's pair student is considered special and even has a unique nickname, *jjak*, in Korean. It is conceivable that after having spent 12

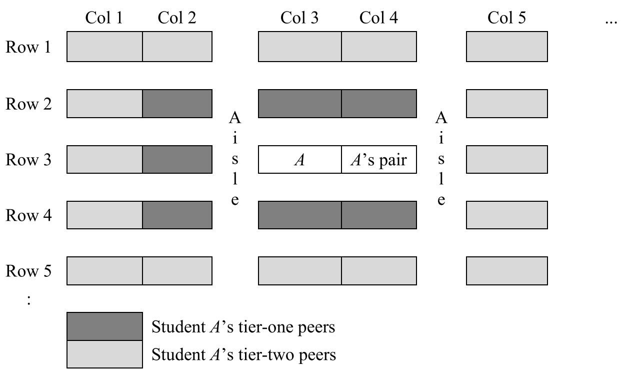

FIGURE 1. Seating chart and definition of pair and peers. *Note*: The figure shows an example of a seating chart. In our samples, the number of columns is between 11 and 14, and the number of rows is between 9 and 15. Student A's pair, which is the key variable in this study, is defined as that person seated next to student A in the adjacent seat of the same row. In addition to the pair, students who surround student A are defined as tier-one peers. The next outer layer of students is defined as tier-two peers.

[8]The university's undergraduate programs consist of seven academic units: humanities, social sciences, natural sciences, engineering, economics, business, and communication. We categorized them into four groups. The university has offered the option of multiple majors. Economics is one of the most popular majors at the university. About 30% of the courses in economics are taught in English so as to meet the demand of foreign-exchange students.

years in this kind of a classroom environment, even university students may have a special feeling about their pair student.

Second, we define neighboring peers. Figure 1 shows that a student is surrounded by up to seven students excluding his or her pair. We name those students *tier-one peers*. Similarly, we define another outer layer of students (up to eight) as *tier-two peers*. It is interesting to examine the role of physical distance in peer effects in the context of education, as it is well known that distance matters in knowledge spillover or technology diffusion (Jaffe, Trajtenberg, and Henderson (1993), Keller (2004)).

Table 1 presents the summary statistics of the crucial variables. Because we are interested in students' relative performance within classes, we standardize the test scores within classes and examinations.[9] The average number of registered years is 2.7. There are more male students (59%), and about 54% of the students are economics majors. About 13% of the students are course retakers.

The last two columns in Table 1 present the correlations between students' characteristics and their test scores. We find that students in the front seats, those with an economics major, and those with a higher GPA do better. According to our definition of a pair, 17% of students do not have a pair. We find that these students without pairs do worse in their class on average. Midterm and final examination scores are strongly correlated, and the correlation coefficient is about 0.61.

The statistics in the bottom panel of Table 1 are suggestive of the existence of peer effects. We find that a student's final exam score is significantly correlated with the pair student's midterm and final exam scores.[10] In particular, it is interesting to find that the correlation becomes larger over time. Similarly, the final exam score is significantly correlated with the pair student's GPA, while the correlation with the midterm score is insignificant. The results suggest that peer effects are likely to become larger over time as students interact more.

4. Empirical strategy

4.1 *Identification from repeated interaction*

To derive our estimation equation, we begin with a "structural" educational production function.[11] We assume that peer effects arise due to "something unobservable about the peers, such as their ability, motivation or behavior" (Fruehwirth (2013, 2014)). Our

[9]Testing and grading policies are standardized because all classes are economics courses. All classes in our sample require two examinations, a midterm and a final, which are given during official university-wide examination weeks. Grading policies are similar in the sense that there are some strict rules at the university level. For example, the proportions of As and Bs are strictly regulated.

[10]One might think that correlations between neighboring students' test scores could arise due to cheating. However, this is unlikely because students are reassigned to seats at exams according to their student ID numbers. Also, it is very hard to cheat on exams because typically, two or more teaching assistants proctor in a large class, similar to those in our sample.

[11]We adopt the distinctions made by Fruehwirth (2013, 2014): the statistical model of student achievement with peer effects and the structural educational production function. In the statistical model, the peer effect works through peer achievement (i.e., test scores), whereas in the structural model, it works through unobservable "ability."

Table 1. Summary statistics of own and peer characteristics.

			Correlation Coefficient With Own Scores	
Variables	Mean	S.D.	Own Midterm	Own Final
Panel A: Own Characteristics				
Final score	0.000	0.996	0.611***	1.000
Midterm score	0.010	0.988	1.000	0.611***
Grade (the number of registered years)	2.684	2.206	−0.040***	0.002
Boy	0.589	0.492	0.022	0.028*
Row of seat	6.161	3.522	−0.083***	−0.077***
End columns	0.192	0.394	−0.042***	−0.019
Major: Economics	0.535	0.499	0.014	0.006*
Business	0.159	0.366	−0.002	0.002
Humanity or social science	0.190	0.392	0.005	0.002
Natural science or engineering	0.104	0.305	0.007	0.018
Retaking	0.131	0.338	−0.030*	−0.009
Exchange student	0.013	0.112	−0.090***	−0.090***
No pair	0.173	0.378	−0.074***	−0.076***
Panel B: Peer Characteristics				
Pair (next-seat peer)				
Final score	0.035	0.959	0.074***	0.091***
Midterm score	0.039	0.971	0.082***	0.075***
Tier-one peers				
Avg. final score	0.018	0.466	0.020	0.005
Avg. midterm score	0.022	0.451	0.006	0.021
Grade composition	1.302	1.001	−0.006	−0.001
Gender composition	0.427	0.238	−0.006	−0.047***
Panel C: Average GPA in Economics				
Own GPA	3.129	0.619	0.378***	0.403***
Next-seat pair's GPA	3.153	0.612	0.028	0.069***

Note: The statistics for own characteristics in Panel A are obtained for 4155 students, and the own previous semester GPA reported in Panel C is for 3249 students. Statistics for peer characteristics are calculated for 3437 students in Panel B and 2640 paired students in Panel C. Exam score is normalized using the class mean and standard deviation. The grade composition variable for tier-one peers is measured using the difference between own average and each tier-one peer's average number of registered years. Here, gender composition is the proportion of tier-one peers that are of opposite sex to the central student. The asterisks *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels.

assumption is reasonable in the educational context, because we usually do not directly observe peers' relevant behaviors, and it is ex ante unknown how peers' academic outcomes should be included in the educational production function.[12] Our identification idea is along the lines of previous papers that try to specify peer effects more structurally, such as Altonji, Huang, and Taber (2010), Arcidiacono et al. (2012), and Fruehwirth (2013, 2014). Also, in the sense that our paper is an attempt to tackle endogenous sorting by using a particular structure of fixed effects, it is closely related to the strand of

[12]For example, peer effects in drinking and smoking may arise directly by those behaviors, that is, drinking and smoking together with peers. However, how peers would help or hinder one's learning is not obvious.

the literature based on network analysis (as an application for peer effect, see Boucher, Bramoullé, Djebbari, and Fortin (2014)).

As aforementioned, our identification is based on repeated interactions with the same peers. The setup is natural in classrooms. Suppose that students take two exams, a midterm (m) and a final (f). A student's score depends upon his or her own characteristics (X) and unobservable characteristics (u) as well as his or her peer's. Here, u can represent any unobservable characteristics, but for convenience, we call it ability.[13] For simplicity, we consider the case in which a peer group consists of two classmates, i and j, in class c, although the results below can be extended to a general case where there are more than two members.[14] We assume that the midterm score is determined in the manner

$$Y_{icm} = X_{ic}\alpha_{xm} + X_{jc}\widetilde{\alpha}_{xm} + u_i + \widetilde{\alpha}_{um}u_j + \mu_{cm} + \varepsilon_{icm},$$
$$Y_{jcm} = X_{jc}\alpha_{xm} + X_{ic}\widetilde{\alpha}_{xm} + u_j + \widetilde{\alpha}_{um}u_i + \mu_{cm} + \varepsilon_{jcm},$$

where ε_{icm} and ε_{jcm} are individual transitory shocks and they are not correlated conditional on class-by-test fixed effects (μ_{cm}). We assume that the shock is not autocorrelated as it is transitory. We would like to emphasize that our model does not assume that students are randomly paired. We allow that their observable and unobservable characteristics (Xs and us) can be correlated.

In the above model, peer effects may arise in two channels: (i) via observable characteristics (X_{jc}) as contextual effects and (ii) via unobservable characteristics (u_j). We are mainly interested in the second, the so-called ability peer effect. We believe that test scores are improved only if the student exerts more effort. Thus, any spillover effect from the peer's unobservable ability should result from the student's increase in (or reduction in) effort induced by his or her peer via the peer's unobservable ability. For example, a high-ability peer can directly help or motivate the student. However, a low-quality peer might misbehave during lectures and hinder the student's learning.

The final exam score is determined in the same way except that the midterm score may directly affect the final score by a factor of ρ:[15]

$$\begin{aligned} Y_{icf} &= \rho Y_{icm} + X_{ic}\alpha_{xf} + X_{jc}\widetilde{\alpha}_{xf} + \alpha_{uf}u_i + \widetilde{\alpha}_{uf}u_j + \mu_{cf} + \varepsilon_{icf}, \\ Y_{jcf} &= \rho Y_{jcm} + X_{jc}\alpha_{xf} + X_{ic}\widetilde{\alpha}_{xf} + \alpha_{uf}u_j + \widetilde{\alpha}_{uf}u_i + \mu_{cf} + \varepsilon_{jcf}. \end{aligned} \tag{1}$$

[13]We follow the convention in labor economics of labeling bias from omitting time-invariant unobservables as the ability bias. In our model, u may imply cognitive ability or motivation as long as it is relevant for academic achievement and does not change over time. For example, some students might be more motivated about economics and they might positively influence their classmates.

[14]In our empirical analysis, we focus on the next-seat peer so as to identify peer effects on a student. Although students sitting on two consecutive seats are most likely to interact with each other, it is possible that actual peers who influence the student are different from the next-seat peer. This measurement error will attenuate our estimates for peer effects (Foster (2006)).

[15]The effect of the student's own midterm score might reflect his or her effort choice conditional on the midterm score. That is, it might reflect the persistent effect of knowledge stock accumulated until the midterm on performance in the final examination.

Like the midterm score, the final score is affected by the pair's unobservable ability.[16] We are interested in identifying $\widetilde{\alpha}_{um}$ and $\widetilde{\alpha}_{uf}$, the effects of the pair's unobservable ability. We observe student outcomes only at two points during the semester, whereas peer effects should occur continuously. Therefore, $\widetilde{\alpha}_{um}$ represents the sum of the effects of the pair's ability that occurred until the midterm examination. Likewise, $\widetilde{\alpha}_{uf}$ represents that of the effects of the pair's ability that occurred between the midterm and final examinations.

Note that the above equations are not estimable because abilities are not directly observable. To derive estimable equations, we solve the first two equations for unobservables, u_i and u_j, and plug them into the equations for final scores. After some algebra, we obtain

$$\begin{aligned} Y_{icf} = {} & \left\{\rho + \alpha_{uf} + \frac{\widetilde{\alpha}_{um}(\widetilde{\alpha}_{uf} - \widetilde{\alpha}_{um}\alpha_{uf})}{\widetilde{\alpha}_{um}^2 - 1}\right\} Y_{icm} + \left\{\frac{\widetilde{\alpha}_{um}\alpha_{uf} - \widetilde{\alpha}_{uf}}{\widetilde{\alpha}_{um}^2 - 1}\right\} Y_{jcm} \\ & + \lambda X_{ic} + \widetilde{\lambda} X_{jc} + \mu_c + e_{icf}, \end{aligned} \tag{2}$$

where

$$\begin{aligned} \lambda &= \alpha_{xf} - \alpha_{uf}\alpha_{xm} - \frac{\widetilde{\alpha}_{um}\alpha_{uf} - \widetilde{\alpha}_{uf}}{\widetilde{\alpha}_{um}^2 - 1}(\widetilde{\alpha}_{xm} - \alpha_{xm}\widetilde{\alpha}_{um}), \\ \widetilde{\lambda} &= \widetilde{\alpha}_{xf} - \alpha_{uf}\widetilde{\alpha}_{xm} - \frac{\widetilde{\alpha}_{um}\alpha_{uf} - \widetilde{\alpha}_{uf}}{\widetilde{\alpha}_{um}^2 - 1}(\alpha_{xm} - \widetilde{\alpha}_{xm}\widetilde{\alpha}_{um}), \\ \mu_c &= -\left[\frac{\widetilde{\alpha}_{um}\alpha_{uf} - \widetilde{\alpha}_{uf}}{\widetilde{\alpha}_{um}^2 - 1}(1 - \widetilde{\alpha}_{xm}) + \alpha_{uf}\right]\mu_{cm} + \mu_{cf}, \\ e_{icf} &= \left[\frac{\widetilde{\alpha}_{um}\alpha_{uf} - \widetilde{\alpha}_{uf}}{\widetilde{\alpha}_{um}^2 - 1}\widetilde{\alpha}_{xm} - \alpha_{uf}\right]\varepsilon_{icm} - \frac{\widetilde{\alpha}_{um}\alpha_{uf} - \widetilde{\alpha}_{uf}}{\widetilde{\alpha}_{um}^2 - 1}\varepsilon_{jcm} + \varepsilon_{icf}. \end{aligned}$$

We estimate the simple linear equation

$$Y_{icf} = \beta Y_{icm} + \widetilde{\beta} Y_{jcm} + \lambda X_{ic} + \widetilde{\lambda} X_{jc} + \mu_c + e_{icf}, \tag{3}$$

where μ_c is the class-level fixed effect, which is a linear combination of the class-midterm fixed effect and class-final fixed effect. The equation resembles the reduced-form equation that is popularly used in literature to identify peer effects by using peers' lagged achievement. The coefficient for the peer's midterm score can be interpreted as an exogenous effect via the peer's unobservable time-invariant ability (Manski (1993)).

A few points are worth noting here regarding our final estimation equation. First, we control for the student's past achievement (i.e., midterm score). In this sense, it is a value-added specification (Hanushek (1979)). In particular, note that we allow the effects of educational inputs (X as well as u) to vary "with the temporal distance between the

[16] Our basic model excludes the possibility that the peer's midterm performance affects one's final score. In Appendix A, we show that the model can be extended to allow for the direct effect of the peer's midterm score. We cannot allow for contemporaneous peer effects because it introduces a simultaneity bias. Thus, our model is restricted in terms of peer effect channels.

time the inputs were applied and the time of the test score measure" (Todd and Wolpin (2003)).

Second, suppose that the effects of unobservables do not vary over time, similar to standard fixed effects; that is, $\alpha_{uf} = 1$ and $\widetilde{\alpha}_{um} = \widetilde{\alpha}_{uf}$ in Equation (2). In this case, we have that $\widetilde{\beta} = 0$ (and $\beta = \rho + 1$) in Equation (3). This means that even though peer effects do exist (i.e., $\widetilde{\alpha}_{um} = \widetilde{\alpha}_{uf} \neq 0$), it is possible that we fail to identify the effect. We would like to emphasize that our identification is conservative in that we may not always conclude there is a peer effect when it is actually present.

Third, we want to emphasize that one of the most annoying problems in estimating peer effects, which is the endogenous selection into peer groups, is not a problem in our model. Endogenous choice of peers implies that students choose their peers based on unobservables. In short, students self-sort into peer groups based on their unobservable characteristics. In our model, such action means that u_i and u_j are correlated. However, the correlation is not a problem since those unobservables are removed in the final estimation equation. Our estimation is still valid even if $E(u_i u_j) \neq 0$. The result is intuitive because we identify peer effects from changes in how students interact with each other from one point of time to another. Thus, initial sorting should not matter. Of course, this result depends on the assumption that students sort into peer groups based on their time-invariant unobserved ability conditional on observable characteristics. This assumption is likely to make sense in our setting because seat peers are determined in the beginning of the semester. If unobserved ability changes over time and students are sorted based on their expectation about unobserved ability, our identification strategy does not work.

Last, another common problem in the peer effect literature is that it is almost impossible to define the peer group accurately. In this model, we estimate the effects from the pair. In our estimation model, we allow for the effects from tier-one and tier-two peers. However, it is still possible that there is a third student with whom the student interacts despite large physical distance between these students in the classroom. Then, the third student's ability is an omitted variable in our specification. However, this does not necessarily bias our estimates; that is, bias occurs if the omitted peer's ability is correlated with the pair's score.[17]

What do we estimate by the coefficient in our model? Following Fruehwirth (2014), we assume that $\widetilde{\alpha}_{um} < 1$. Suppose that unobservables represent abilities. Then the assumption means that a student's own test score is more affected by his or her own ability than the peer's ability. Likewise, we assume that $\alpha_{uf} > \widetilde{\alpha}_{uf}$. Then the estimation of $\widetilde{\beta}$ will provide a lower bound for the effect of the peer's unobservable ability:

$$\widetilde{\beta} = \frac{\widetilde{\alpha}_{um}\alpha_{uf} - \widetilde{\alpha}_{uf}}{\widetilde{\alpha}_{um}^2 - 1} = \frac{1}{\widetilde{\alpha}_{um} + 1}\widetilde{\alpha}_{uf} + \frac{\widetilde{\alpha}_{um}\overbrace{(\alpha_{uf} - \widetilde{\alpha}_{uf})}^{(+)}}{\underbrace{\widetilde{\alpha}_{um}^2 - 1}_{(-)}} \leq \widetilde{\alpha}_{uf}.$$

[17]Another possible source of bias is that students might create a social network in response to the pair's ability. For example, they might try to organize a high-ability study group if the pair's ability is low. In this case, our estimate for the pair's ability is underestimated.

The magnitude of the bias depends on $\widetilde{\alpha}_{um}$ and $(\alpha_{uf} - \widetilde{\alpha}_{uf})$. It will be small if the effect of the peer's unobservable ability is of a similar size as that of own ability. We obtain the unbiased estimate if there is no peer effect on the midterm exam, that is, $\widetilde{\alpha}_{um} = 0$.

The ordinary least squares (OLS) estimates of β and $\widetilde{\beta}$ will be biased since the composite error, e_{icf}, is correlated with both Y_{icm} and Y_{jcm}. The error term can be simplified to

$$e_{icf} = (\widetilde{\alpha}_{um}\widetilde{\beta} - \widetilde{\alpha}_{uf})\varepsilon_{icm} - \widetilde{\beta}\varepsilon_{jcm} + \varepsilon_{icf}.$$

The first component in the composite error is correlated with the student's own midterm score, and the second term is correlated with the pair's midterm score. It is likely that the OLS estimate for $\widetilde{\beta}$ is biased toward zero since Y_{jcm} is positively correlated with ε_{jcm}, whereas the latter is negatively correlated with the dependent variable, Y_{icf}. This implies that the OLS estimate is likely to be a lower bound of the peer effect. Also since Y_{icm} is positively correlated with ε_{icm}, the OLS estimate for β is likely to be biased. It is also possible that ε_{icm} and ε_{jcm} are correlated even after controlling for class-level fixed effects, μ_{cm} and μ_{cf}. A subclass shock that affects only i and j may occur. This may bias the OLS estimates for β and $\widetilde{\beta}$. So as to address the potential endogeneity bias, we use the instrumental variables (IV) strategy. The IV method will consistently estimate $\widetilde{\beta}$, which is a lower bound of $\widetilde{\alpha}_{uf}$. We will explain our instrumental variables in Section 5.1.

4.2 *Endogenous seat selection*

Endogenous peer group formation is clearly one of major identification problems for estimating peer effects. In our setting, students select their seat on the first-come–first-served basis. Typically there is a long waiting line outside the classroom on the first day, and seats are taken by one after another in order. The students cannot prevent others from sitting next to them and they are not allowed to reserve seats for their friends.

It is still possible that students are endogenously matched. For example, two friends can arrive simultaneously and take two adjacent seats, or a student comes late, looks around, and takes one of the available seats next to her close friends. The data indeed show that there is endogenous sorting on their characteristics. For example, male students are more likely to sit together with male students, and seniors are more likely to sit next to other seniors. Also they are sorted by major. However our main concern is the possibility that students are sorted in terms of academic ability. There may also be endogenous sorting of students based on academic ability or aspiration, with aspiring students competing for front seats. Classes are large sized, mostly more than 100 students, and so it is quite difficult to focus from the back. Thus it is true that those with high aspiration are clustered in the front-center area. In fact, the data reveal that the average GPA of those in the first row is 3.2, while that of those in the last row (15th) is 2.9. To address this, in our regression analysis, we control for seat row.

To address any bias owing to initial endogenous sorting of pairs, we conduct some robustness checks. Here, we exclude those pairs who are likely endogenously matched from the sample. First, we check correlation in initial GPA between paired students and exclude classes with high correlation. A similar approach was taken in Clotfelter, Ladd, and Vigdor (2007) and Vigdor and Nechyba (2007). They selected schools where the hypothesis of random assignment was not rejected. Second, we exclude pairs who have shared classes in the past. Third, we exclude those who were paired in prior classes. Last, we exclude pairs with the same major. Note that the results should be robust to these exclusions because identification based on our model does not require $\mathrm{Cov}(u_i, u_j) = 0$ because the time-invariant unobservable ability variables are eliminated in the final regression equation. It is still possible that there exists time-varying unobservable ability. Our approach is limited to the extent that students' general academic ability varies within a semester.

5. Empirical results

5.1 *Average effects*

Table 2 presents the regression results for Equation (3). In column 1, we estimate our main equation without any peer variables. The results show that, not surprisingly, a student's own midterm score is a strong predictor for the student's final test score. We also find that the row of the seat matters. As one's seat is further from the front, the student's final exam score is lower even after controlling for the midterm score. The same result was previously found in Benedict and Hoag (2004). As the seat is further back by one row, the score is decreased by 0.01 standard deviations. This is because the distance to the instructor could impede student's learning or reduce student's attention. On the other hand, this might result from the selection of seats. Competent students might prefer and select front seats. Both stories imply that some students might as well compete for the front-row seats during the period of seat assignment at the beginning of the semester. To control for students' seat preferences, which are correlated with abilities or learning aspiration/efforts, in column 3, we control for row-by-class fixed effects (FE). The results are basically the same; thus, we only control for class fixed effects.

In column 2, we include the pair's midterm score and tier-one neighbors' characteristics as well as their average score. In column 4, we restrict our sample to those students who have a pair. In column 5, we add tier-two peers' variables. Across specifications, the estimates for peer effects—the coefficients for the pair's midterm score—are very similar, ranging from 0.0224 to 0.0254.

We find that the average score of tier-one neighbors does not matter. The estimates are actually larger than the estimates for the pair's score. But they are not directly comparable since there are often seven students in the tier-one neighbor group. Also, the neighbors' gender composition is not significant except in column 4. The grade composition matters: as students are surrounded by more heterogeneous neighbors in terms of grade, they do worse on their final exam. This supports the so-called focus model that homogeneity of peers is good for students (Sacerdote (2011)). The characteristics of tier-two neighbors turn out to be insignificant in column 5, suggesting that physical distance between students matters for their interactions. Thus, hereafter, we focus on spillover effects from the pair and tier-one neighbors.

As mentioned earlier, under the FSS system, students' seats are assigned on the first-come–first-served basis; that is, students choose their seats according to their arrival

TABLE 2. Estimates of the effect of pair's midterm score on own final score (dependent variable: own final score).

	(1) Without Peer Effect	(2) Class FE	(3) Row-by-Class FE	(4) Paired Students	(5) Adding Tier-Two Peers
Class FE	Y	Y	N	Y	Y
Row-by-Class FE	N	N	Y	N	N
Pair or Peer Characteristics					
Pair's midterm score		0.0224*	0.0254*	0.0240*	0.0229*
		(0.0129)	(0.0154)	(0.0130)	(0.0128)
No pair		−0.0886*	−0.0700		−0.0912*
		(0.0453)	(0.0464)		(0.0460)
Tier-one peers					
Avg. midterm score		0.0374	0.0334	0.0277	0.0385
		(0.0301)	(0.0290)	(0.0378)	(0.0305)
Grade composition		−0.0303**	−0.0400**	−0.0154	−0.0320**
		(0.0138)	(0.0187)	(0.0172)	(0.0141)
Gender composition		−0.0843	−0.0839	−0.1692**	−0.0905
		(0.0639)	(0.0580)	(0.0737)	(0.0633)
Tier-two peers					0.0052
Avg. midterm score					(0.0305)
Own Characteristics					
Midterm score	0.6128***	0.6097***	0.6073***	0.5925***	0.6089***
	(0.0199)	(0.0204)	(0.0186)	(0.0233)	(0.0205)
Grade	0.0189	0.0283**	0.0309***	0.0311**	0.0287**
	(0.0127)	(0.0134)	(0.0119)	(0.0143)	(0.0136)
Boy	0.0143	0.0024	0.0052	−0.0069	0.0010
	(0.0277)	(0.0262)	(0.0330)	(0.0316)	(0.0259)
Row of seat	−0.0095**	−0.0065		−0.0085*	−0.0060
	(0.0041)	(0.0040)		(0.0046)	(0.0040)
No pair	0.0126	0.0385	0.0308	0.0153	0.0415
	(0.0283)	(0.0281)	(0.0361)	(0.0336)	(0.0282)
Major: Business	0.0189	0.0242	0.0426	0.0343	0.0217
	(0.0508)	(0.0509)	(0.0494)	(0.0440)	(0.0508)
Humanity or social science	0.0001	0.0026	0.0248	0.0094	0.0030
	(0.0493)	(0.0489)	(0.0445)	(0.0497)	(0.0492)
Natural science or engineering	0.0395	0.0490	0.0723	0.0338	0.0481
	(0.0545)	(0.0543)	(0.0562)	(0.0580)	(0.0544)
Retaking	0.0090	0.0118	0.0205	−0.0041	0.0098
	(0.0444)	(0.0442)	(0.0476)	(0.0557)	(0.0442)
Exchange student	−0.2450**	−0.2389**	−0.2956**	−0.1277	−0.2125**
	(0.0979)	(0.0955)	(0.1320)	(0.0961)	(0.0867)
Constant term	−0.0143	0.0330	−0.0140	0.0634	0.0346
	(0.0476)	(0.0626)	(0.0458)	(0.0638)	(0.0627)
Observations	4155	4155	4155	3437	4144
Adj. R-squared	0.3702	0.3724	0.3726	0.3592	0.3712

Note: We estimate Equation (3). Robust standard errors are given in parentheses and are clustered by class for models (1), (2), (4), and (5) or by row-by-class group for model (3). The asterisks *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels.

time at the classroom. Therefore, they cannot always select their pairs when the seat location is selected. To check randomness of pair assignment, we examine correlations of observable characteristics between pairs. It turns out that gender correlation is very strong: the correlation coefficient is higher than 0.2 for 31 of 36 classes. However, the correlation in GPA, which is a proxy for ability, is not significant. The average is just 0.04 and furthermore it varies by class. This is not surprising. Due to the first-come–first-served rule, the first "comer" selects a seat without knowing who will be his or her pair. The second comer can select his or her pair, but typically they do not know each other given the large size of classes in our sample. Last, GPA is not observable unless they are very close friends.

As noted in the previous section, the midterm scores in Equation (3) are endogenous. We deal with the endogeneity bias by the instrumental variable (IV) estimation method. As IVs for the midterm scores, we use the student's and his or her peer's grade point average (GPA) from the previous semester. Information on GPA is missing if they are freshmen, transfer students, or exchange students. Hence, we include a dummy variable indicating whether a GPA is missing. Therefore, we have two endogenous variables—the student's own and the pair's midterm score—and four instrumental variables—the student's GPA, an indicator of whether the student's GPA is missing, the pair's GPA, and an indicator of whether the pair's GPA is missing. The validity assumption is that the previous semester's GPA affects the final score only through its effect on the midterm score. Specifically, the assumption is that the IVs should not be correlated with the composite error in the second stage equation, which consists of three error terms: ε_{icm}, ε_{jcm}, and ε_{icf}. Therefore, the validity assumption is equivalent to the assumption that the error terms are not serially correlated.[18]

Table 3 presents the results. Panel A presents reduced-form results; that is, we include IVs directly instead of the midterm scores. Assuming that the GPA proxies one's unobservable academic ability (u_i and u_j), by estimating the reduced-form equation, we directly estimate Equation (1) before we eliminate the unobservable ability variables. In this case, the coefficients for GPA variables directly indicate peer effects. The results across specifications show that peer effects are present. A one-point increase in the pair's GPA increases the final exam score by 0.06. Again the effect is larger when students are not sorted in terms of GPA.

Panel B presents the first- and second-stage results. The first-stage results are as expected. A student's GPA is a strong predictor for his or her own midterm score but not for the pair's. Statistics and test results from the first stage show that the IVs pass weak instrument and overidentification tests. The second-stage results are consistent with our

[18]It is unlikely that ε_{jcm} is serially correlated because the pair in the current semester is likely different from that in the previous semester. But ε_{icm} and ε_{icf} might be serially correlated. The sign of the resulting bias is ambiguous. The previous semester's GPA is determined by the previous semester's midterm score, which contains the previous semester's ε_{icm}, and the final score, which contains the previous semester's ε_{icf}. The IV estimate is upwardly biased if ε_{icf} is serially correlated. On the other hand, it is downwardly biased if ε_{icm} is serially correlated since the composite error term, e_{icf}, is likely to be negatively correlated with ε_{icm} ($\alpha_{uf} > \widetilde{\alpha}_{xm}$ and the fraction part in front of $\widetilde{\alpha}_{xm}$ is positive and less than 1). We formally check the validity condition in Appendix B.

Table 3. Instrumental variable (IV) estimates of peer effects.

Model	(1) Class FE	(2) Row-by-Class FE	(3) Students With Pairs	(4) Adding Tier-Two Peers
Class FE	Y	N	Y	Y
Row-by-Class FE	N	Y	N	N
Panel A: Reduced-Form Analysis (dependent variable = own final score)				
Pair's GPA	0.0606**	0.0661**	0.0603**	0.0624**
	(0.0284)	(0.0306)	(0.0287)	(0.0288)
Own GPA	0.6668***	0.6679***	0.6304***	0.6669***
	(0.0328)	(0.0368)	(0.0355)	(0.0329)
Constant term	−2.1610***	−2.2387***	−2.0573***	−2.1678***
	(0.1612)	(0.1661)	(0.1819)	(0.1631)
R-squared	0.1480	0.1406	0.1326	0.1474
Panel B: 2SLS Analysis				
First stage I (dependent variable = own midterm score)				
Own GPA	0.6112***	0.6138***	0.5820***	0.6112***
	(0.0383)	(0.0380)	(0.0448)	(0.0384)
Pair's GPA	−0.0056	−0.0008	−0.0023	−0.0044
	(0.0243)	(0.0320)	(0.0247)	(0.0245)
R-squared	0.133	0.1263	0.114	0.1323
First stage II (dependent variable = pair's midterm score)				
Own GPA	−0.0061	−0.0016	−0.0057	−0.0069
	(0.0199)	(0.0263)	(0.0252)	(0.0201)
Pair's GPA	0.5600***	0.5586***	0.5609***	0.5598***
	(0.0436)	(0.0395)	(0.0437)	(0.0434)
R-squared	0.1077	0.1076	0.1112	0.1082
Second stage (dependent variable = own final score)				
Pair's midterm score	0.1171***	0.1217**	0.1140***	0.1186***
	(0.0446)	(0.0542)	(0.0442)	(0.0456)
Own midterm score	1.0915***	1.0873***	1.0832***	1.0916***
	(0.0528)	(0.0540)	(0.0549)	(0.0532)
Constant term	−0.0690	−0.0646	−0.0939	−0.0672
	(0.0586)	(0.1092)	(0.0629)	(0.0596)
R-squared	0.1458	0.2337	0.1148	0.1432
Additional statistics				
F-statistics for Wald test	109.36	101.73	88.9	108.66
p-value for underidentification test	0.0001	0.0000	0.0002	0.0001
p-value for Hansen statistics	0.9064	0.9481	0.9665	0.9054
Observations	4155	4155	3437	4144

Note: The table reports the key results of the IV estimation. We use own and pair's GPA for the previous semester as the IVs for own and pair's midterm scores. In addition, we include the indicators for missing GPA of the student and his/her pair. GPA is missing for transfer, exchange, and first-semester students. Panel A uses GPAs instead of midterm scores (i.e., a reduced-form estimation); furthermore, Panel B uses two-stage least squares (2SLS) estimations using GPAs and indicators for missing GPA as IVs. Robust standard errors, given in parentheses, are clustered by class for models (1), (3), and (4) or by row-by-class group for model (2). The asterisks *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels. The full results for the second stage in Panel B are reported in Table A.1.

previous OLS results and expectation about magnitude. The estimate of the coefficient for the pair's midterm score is about 0.12. The results support our prediction in Section 4.1; the OLS estimates are likely to be underestimated. Recall that even the IV estimates are the lower bounds for true peer effects. (The full results of our IV regressions are presented in Table A.1 in the Appendix.)

Our estimates of peer effects are within the range of estimates found in previous literature. Sacerdote (2011) summarized these estimates from previous literature (Table 4.2) and converted them to the marginal effects of a one-point increase in average peer score. The estimates of Hoxby (2000) range from 0.3 to 6.8. Hanushek et al. (2003) found the value of math to be 0.17. Burke and Sass (2013) found 0.04 for math among grades 3–10 students after controlling for student and teacher fixed effects. Lefgren (2004) found a value of 0.03 for grade 6 students. Lavy and Schlosser (2011) found 0.84 for females and 1.06 for males. Few studies have estimated peer effects in postsecondary education. Brunello, De Paola, and Scoppa (2010) found that for university students in Italy, a 1 standard deviation in class ability increases own grades by 0.08 standard deviation.

Moreover, we find that the estimates for one's own midterm score are much larger than the OLS estimates. In Table 2, we found that the estimates range around 0.6, which are significantly less than 1. The IV estimates are slightly larger than 1.

As mentioned earlier, endogenous sorting of pairs may be a concern. To address this concern, we conduct a few robustness checks. We exclude those pairs that are likely to be endogenously matched and run the same regression. Table 4 presents the results. First, we drop eight classes with significant and positive correlations between paired students in initial GPA. For the remaining classes, the correlation coefficient is lower than 0.15. The results in column 1 show that the peer effect is significant and positive. The new estimate is larger, suggesting that peer effects might be stronger when students are randomly matched. In column 2, we exclude those pairs who shared classes prior to the semester. In column 3, we exclude those paired in classes prior to the semester. Last, in column 4, we drop all pairs with the same major.[19] Panels A and B present the OLS and IV estimates, respectively. We find that the results are overall quite similar to our previous results. There are not many pairs who took the same classes before (7%) or were paired before (5%). Excluding pairs with the same major reduces the sample size drastically, but the results remain similar.

5.2 *Heterogeneous effects across subsamples*

In this subsection, we attempt to identify which factors mediate (reinforce or attenuate) peer effects. We relax the assumption that every student is affected by his or her peers in the same manner (Hoxby and Weingarth (2005), Arcidiacono et al. (2012)). Table 5 presents OLS and IV results over different subsamples. Overall, as with the results for the whole sample, the IV estimates for peer effects are larger than the corresponding OLS estimates. We find that peer effects differ across student groups and classroom

[19] Students with the same major are likely to know more about each other, and they might strategically cooperate over multiple classes.

Table 4. Peer effects among pairs without previous interaction (dependent variable: own final score).

	(1)	(2)	(3)	(4)
	Results After Excluding			
	Students in Classes With High Correlations Between Own and Pair's GPA	Pairs Who Took Any Same Classes Before the Semester	Pairs Who Were Paired in Any Classes Before the Semester	Pairs With the Same Major
Panel A: OLS Regression				
Peer characteristics				
Pair's midterm score	0.0328**	0.0226*	0.0189	0.0140
	(0.0156)	(0.0126)	(0.0134)	(0.0277)
No pair	−0.0636	−0.0850*	−0.0895*	−0.1018**
	(0.0525)	(0.0440)	(0.0458)	(0.0491)
Tier-one peers	0.0158	0.0452	0.0362	0.0535
Avg. midterm score	(0.0340)	(0.0305)	(0.0305)	(0.0381)
Own characteristics				
Midterm score	0.5982***	0.6150***	0.6088***	0.6291***
	(0.0248)	(0.0189)	(0.0210)	(0.0288)
Constant term	0.0359	0.0432	0.0203	−0.0258
	(0.0747)	(0.0665)	(0.0651)	(0.1254)
Adj. R-squared	0.3500	0.3755	0.3706	0.3818
Panel B: 2SLS Regression				
Peer characteristics				
Pair's midterm score	0.1802***	0.1034**	0.1155**	0.2205***
	(0.0622)	(0.0502)	(0.0459)	(0.0844)
No pair	0.0127	−0.0235	−0.0287	−0.0264
	(0.0721)	(0.0588)	(0.0606)	(0.0622)
Tier-one peers	0.0081	0.0516	0.0395	0.0676*
Avg. midterm score	(0.0432)	(0.0360)	(0.0366)	(0.0349)
Own characteristics				
Midterm score	1.1524***	1.0826***	1.1012***	1.1037***
	(0.0677)	(0.0608)	(0.0552)	(0.0629)
Constant term	−0.1053	−0.0559	−0.0553	0.0444
	(0.0790)	(0.622)	(0.0605)	(0.1273)
R-squared	0.0426	0.1666	0.1345	0.1691
Observations	3104	3854	3950	1709

Note: In Panel A, we use the model of column 2 in Table 2. In Panel B, we use the model of column 1 in Table 3. In all regressions, we control for class fixed effects. Robust standard errors are given in parentheses and are clustered by class. The asterisks *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels.

settings. First, when we divide the sample by gender, we only find significant peer effects among boys. Second, we compare students who are seated in the front and those in the back. We divide a class by using the fifth row as a cutoff row between the front and back. The cutoff row is so chosen to divide the sample almost equally. We find that peer effects exist among students seated in the back. As mentioned earlier, according to our surveys, students who are seated in the back are unlikely to have an acquaintance in the

Table 5. Peer effects by gender, seat row, language, and major.

	(1) Boys	(2) Girls	(3) Up to Fifth Rows	(4) After Fifth Rows	(5) Class in English	(6) Class in Korean	(7) Economics Major	(8) Non-Economics Major
Panel A: OLS								
Peer characteristics								
Pair's midterm score	0.0296	0.0124	0.0056	0.0356*	0.0061	0.0285*	0.0071	0.0411**
	(0.0177)	(0.0160)	(0.0186)	(0.0182)	(0.0270)	(0.0145)	(0.0173)	(0.0156)
Tier-one peers	0.0550	−0.0016	−0.0010	0.0535	0.0503	0.0183	0.0010	0.0759**
Avg. midterm score	(0.0332)	(0.0510)	(0.0399)	(0.0342)	(0.0485)	(0.0374)	(0.0381)	(0.0370)
Own characteristics								
Midterm score	0.6016***	0.6172***	0.5983***	0.6151***	0.6103***	0.6073***	0.6287***	0.5833***
	(0.0303)	(0.0266)	(0.0328)	(0.0220)	(0.0280)	(0.0293)	(0.0280)	(0.0235)
Constant term	−0.0016	0.1013	0.0228	−0.0227	0.1355	0.0042	0.0469	0.0453
	(0.0790)	(0.0844)	(0.0593)	(0.0960)	(0.1277)	(0.0679)	(0.0763)	(0.0900)
Adj. R-squared	0.3762	0.3676	0.3381	0.3953	0.3834	0.3686	0.3971	0.3530
Panel B: 2SLS								
Peer characteristics								
Pair's midterm score	0.1186**	0.1126	−0.0283	0.1890***	0.1877	0.0949**	0.0434	0.1621**
	(0.0557)	(0.0776)	(0.0744)	(0.0618)	(0.1154)	(0.0392)	(0.0569)	(0.0789)
Tier-one peers	0.0679**	−0.0190	0.0135	0.0920**	0.0767	0.0253	0.0024	0.0877***
Avg. midterm score	(0.0275)	(0.0650)	(0.0508)	(0.0374)	(0.0490)	(0.0461)	(0.0511)	(0.0317)
Own characteristics								
Midterm score	1.0659***	1.1312***	1.0954***	1.0884***	1.2490***	1.0250***	1.1715***	1.0360***
	(0.0530)	(0.0819)	(0.0672)	(0.0704)	(0.0864)	(0.0602)	(0.0840)	(0.0776)
Constant term	−0.1050	−0.0421	0.0192	−0.2476***	0.0807	−0.1085	−0.2053***	0.0032
	(0.0878)	(0.0952)	(0.0683)	(0.0961)	(0.1001)	(0.0709)	(0.0753)	(0.0695)
R-squared	0.1653	0.1377	0.1274	0.1673	−0.0163	0.2020	0.1259	0.1594
Observations	2449	1706	1953	2202	1471	2684	2223	1932

Note: In Panel A, we use the model of column 2 in Table 2. In Panel B, we use the model of column 1 in Table 3. In all regressions, we control for class fixed effects. Robust standard errors are given in parentheses and are clustered by class. The asterisks *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels.

neighborhood of their seat. Thus, the finding that peer effects arise among these students seated in the back implies that preexisting friendship is not a necessary condition for peer effects.

In columns 5 and 6, we separate the sample at the class level by whether the class is taught in Korean or English. One-third of the classes were taught in English during the sample period. Here, the results are somewhat different by the estimation method. The OLS results show that peer effects exist for classes in Korean, whereas the IV results show that the effects are stronger for classes in English, although the effects for English classes are marginally significant. If we take the IV results as unbiased, it is conceivable that students are more likely to help each other in classes in English. Informal conversation with students reveals that they often have difficulty understanding lectures in English. Last, in columns 7 and 8, we divide the sample by major, economics versus non-economics. The results show that peer effects exist only for students with a non-economics major. This may be another piece of evidence that economics students are more selfish and less cooperative (Frank, Gilovich, and Regan (1993)). (The full IV regression results are reported in Table A.2.)

One possible theoretical explanation about heterogeneous effects across subsamples is that the peer effect is the consequence of students' rational behavior. For example, we find that the peer effect is larger among students sitting in the back, in English-taught classes, and among non-economics students. Presumably, students under these environments might need more help from peers. Those students sitting in the back might have more difficulty concentrating on lectures or even note-taking (recall that most classes in our sample are large and have more than 80 students). Similarly, students might need more help from each other in English-taught classes or when they are not economics majors. It makes sense that students interact more when there is a higher return to cooperation.

We also think that peer effects may arise through peers' disruptive behavior, as Lazear (2001) pointed out. This might explain why we find a significant peer effect among boys but not among girls. Boys are more likely to exhibit some disruptive behavior during classes. Lavy and Schlosser (2011) found that much of the peer effect arises through changes in classroom violence and disruption. Furthermore, it is likely that more disruptive students are in the back, possibly because of sorting or because they get less attention from teachers. Students may be more likely to be disruptive in English-taught classes, because it is more costly to concentrate. Non-economics students might be more disruptive because they are less motivated in economics classes.[20]

5.3 *Nonlinear effects over the test score distribution*

In this subsection, we apply a Chernozhukov–Hansen instrumental variables quantile regression (IVQR) estimator to estimate the heterogeneous peer effects over the distribution of the students' final exam scores (Chernozhukov and Hansen (2005)). Looking

[20]Thus far, we have used standardized test scores within classes. As a robustness check, we estimated the peer effect in Table A.3 using raw scores (with the full score set to 1). We found that the results are similar to those using standardized scores.

at heterogeneous effects over the distribution rather than the average effect makes more theoretical sense. Consider an educational production function, which depends on one's own ability and his or her peer's. The presence of peer effects indicates that the partial derivative with respect to the peer's ability is not zero (likely positive). Moreover, the cross-partial should not be zero as long as there is complementarity (or substitutability) between the two inputs; that is, the effect of the peer's ability should differ over the distribution of one's own ability.

Figure 2 plots the entire quantile process for the whole sample. The dashed line shows the standard quantile estimates, while the solid line illustrates the IVQR estimates. We present the 90% percent confidence interval for IVQR. The left panel presents IVQR estimates for $\widetilde{\beta}$, and the right panel shows those for β.[21] First, we find that the effect of one's own score is always positive, but the effect gets smaller as the score is higher. This suggests that one's final score is a quadratic function of his or her midterm score. Second, and more interestingly, we find that peer effects are significant at two different parts of the score distribution: among those who are below the average midterm score and among those at the top end. There is no significant peer effect among upper-middle students. A similar pattern is found in Hoxby and Weingarth (2005). Sacerdote (2011) reviews recent papers' finding that peer effect differs by one's own ability and the type of the peer group.

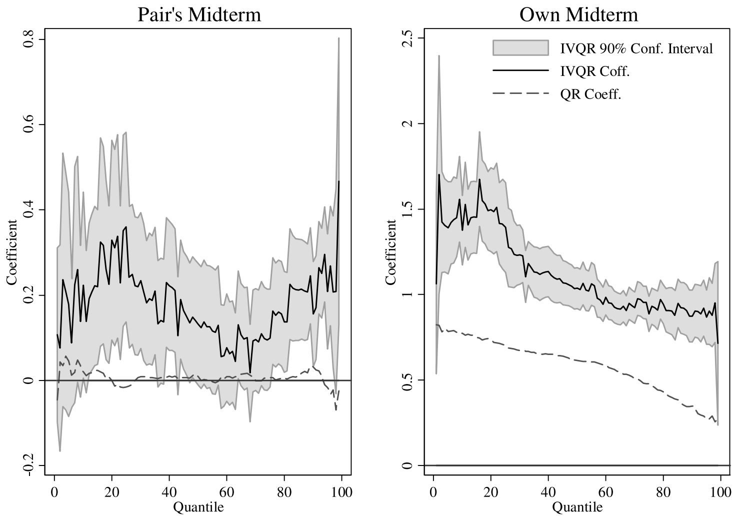

FIGURE 2. Instrumental variable quantile regression estimation results. *Note*: To test the nonlinearity of peer effects over the test score distribution, we ran a quantile regression (QR) and IV quantile regression (IVQR). These figures graphically describe the estimation results for two key variables: pair and own midterm score. In each regression, we used the same control variables as those in the baseline models in Tables 2 and 3, but class fixed effects are dropped. The dashed line indicates the QR estimation coefficients. The solid line indicates the IVQR estimation coefficients, and their 90 percent confidence interval is also presented.

[21] For the results for ρ, see Figure A.2.

The underlying mechanisms for peer effects in the two groups—peer effects among top students and peer effects among those just below average but not at the bottom tail—might differ: top students might help and stimulate each other better than others. Furthermore, top students might be better able to minimize negative effects from lower-quality peers. On the other hand, below-average students might be easily bothered by lower-quality peers. Low-quality students might be incapable of capitalizing on positive effects from high-quality peers.

Figure 3 plots the quantile regression estimates for different subsamples. The results are consistent with our previous findings: significant peer effects appear among male students, among those who are seated in the back, among those taking classes conducted in English, and among non-economics students. Consistent with the patterns in Figure 1, for the subsamples where significant effects are present, they appear either at the lower tail or at the upper tail of the score distribution.

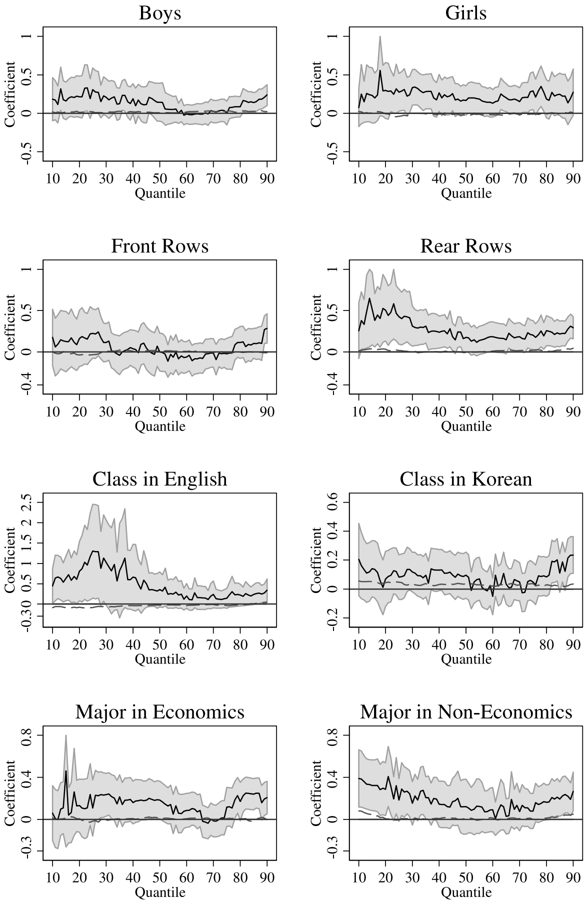

FIGURE 3. Instrumental variable quantile regression estimation results by subsample. *Note*: We ran QRs and IVQRs for eight subsamples. Each figure shows the coefficients of the pair's midterm score estimated by the QR (dashed line) and the IVQR (solid line). The shaded area indicates a 95 percent confidence interval of the IVQR estimation.

5.4 *Personality and peer effects*

Peer effects arise from students' interactions that depend on their interpersonal or social relationships. Thus, different peer effects may arise depending on the way students form social groups or friendships. For example, Carrell, Sacerdote, and West (2013) found that students choose whom to interact with, which depends on the ability distribution of their classmates. In this subsection, we examine whether personality plays a role in shaping a social structure that reinforces or hinders peer interactions.[22]

Admitted to Sogang University, most freshmen (more than 95%) take a personality test, called the Minnesota Multiphasic Personality Inventory (MMPI), conducted by the Student Counselling Center (SCC). This test is well known as the premier device for screening government security personnel and police officers.[23] The results of the test consist of three validity scales (L, F, and K) and 10 clinical scales (Hs, D, Hy, Pd, Mf, Pa, Pt, Sc, Ma, and Si). The validity scales are designed to detect respondents' tendencies to underreport or overreport psychological symptoms. Clinical scales are supposed to measure a variety of psychological symptoms, such as anxiety, depression, and suspiciousness. Higher scores indicate severe symptoms: respondents with 70 or higher points in any scale are typically diagnosed as those who need clinical counselling. According to the Sogang SCC, the measures help predict freshmen's adaptation to a new environment, the ability to cope with stress, and, more importantly, social relationship with others.

Unfortunately we cannot obtain the personality test data at the individual level. Instead, every year, the SCC publishes freshmen's MMPI results aggregated by gender and major, and we use the aggregate data. Surprisingly, the test results vary substantially by entering year, gender, and major. Table 6 presents the percentages of students with high scores for each scale by entering year, gender, and major along with how to interpret high scores in a clinical manner. For example, in the case of the clinical scale for depression (D), 2.28% of male students need clinical counseling compared to 0.61% of female students. There is also sufficient variation by entering year and major.[24] We collected and merged the aggregated data with our individual-level data. Thus, we do not know each individual's personality; yet, the distribution of personality of the individual's group is defined by entering year, gender, and major. Then we classified students according to whether the group they belong to is "social" or "unsocial." A group is defined as *unsocial* if the proportion of the members who are diagnosed as needing clinical counselling is

[22] To the best of our knowledge, no study has examined the role of personality in peer effects. A recent paper by Buechel, Mechtenberg, and Petersen (2014) found that self-control is an important mediator for positive spillover from peers.

[23] Butcher and Rouse (1996) conducted a review of the literature on clinical personality assessment for the 20 years since 1974 and found that out of 8905 articles they reviewed, 4542 used MMPI (or MMPI-2). Rorschach is the second most used assessment.

[24] We do not know why such substantial variation occurs by entering year and major. One possible explanation is that the university selects students based on their academic ability and that academic ability is orthogonal to personality.

Table 6. Percentage (%) of students with high MMPI scores by gender, entering year, and major.

Type of MMPI Scale (Original Label)	Clinical Interpretation of High Scores	Variable: Percentage of High MMPI Score										
				Mean by Subgroup								
				Gender		Entering Year			Major			
		Mean	S.D.	M	F	2008	2009	2011	Econ	Bus	Hu/SS	NS/EE
Hs (Hypochondriasis)	Unrealistically concerned with physical complaints	1.38	1.31	2.08	0.84	0.60	1.92	1.08	1.54	0.65	1.73	0.89
D (Depression)	Unhappy, depressed, and pessimistic	1.34	1.52	2.28	0.61	1.65	1.36	1.08	1.58	0.59	1.40	0.72
Hy (Hysteria)	Focusing on vague physical symptoms to avoid dealing with severe psychological stress	0.68	0.91	1.44	0.10	0.52	0.85	0.53	0.80	0.43	0.59	0.39
Pd (Psychopathic deviate)	Social interactions with emotional shallowness, rebelliousness, and disregard for law	0.74	1.12	1.12	0.45	1.30	0.88	0.07	0.44	0.70	2.19	0.71
Mf (Masculinity–feminity)	Showing interests and behaviors usually associated with opposite sex role	3.48	2.89	1.77	4.81	2.54	4.05	3.23	3.39	3.24	2.66	6.55
Pa (Paranoia)	Strong, irrational suspicions and overestimating own importance	1.04	1.60	1.74	0.50	1.65	1.23	0.24	0.78	1.77	1.40	1.10
Pt (Psychasthenia)	Tense, rigid, anxious and having obsessivethoughts and compulsive behaviors	1.02	1.44	2.12	0.16	0.34	0.53	2.37	1.10	0.94	1.04	0.36
Sc (Schizophrenia)	Experiencing distortions of realityand acting bizarrely	0.56	0.85	0.93	0.27	0.67	0.74	0.15	0.36	1.00	1.09	0.32
Ma (Hypomania)	Outgoing, impulsive, overly active, and excited	1.83	1.51	2.81	1.06	0.80	1.45	3.27	2.41	0.80	0.95	0.49
Si (Social introversion)	Shy, inhibited, and self-effacing	2.39	2.26	4.34	0.86	1.06	1.62	4.72	2.86	2.11	1.25	0.90
Observations		1118	1118	489	629	247	551	320	712	172	157	77

Source: Cox, Weed, and Butcher (2009) for the interpretation of high scores.

Note: The report by the Student Counseling Center classified a student as one who needs clinical counseling when his or her MMPI score is higher than 70 for each type of scale. The distribution presented in the table indicates the average percentage of students with high scores (>70) by gender, entering year, and major. The MMPI was not surveyed for those who entered in 2010. In the table, we classified 13 majors into 4 groups as follows: Econ, economics; Bus, business; Hu/SS, humanities, social sciences, international cultures 1 & 2, mass communications, and law; NS/EE, natural sciences, mechanical engineering, computer science and electronic engineering, and chemical and biomolecular engineering.

higher than the median; that is, for each scale, we divided the whole sample into two groups of equal size.[25] Obviously, some students from the social group may be unsocial while some from the unsocial group may be social. Such misclassification errors should attenuate our estimates below.

We divided students into four groups according to one's own and pair group's social personality traits. Then we ran the class-FE(fixed effects) regression for each social-personality combination group constructed in 10 MMPI clinical scales, respectively. Table 7 presents the coefficients of peer effect (i.e., pair's midterm score) and its standard errors. In column 1, both the student under study and his or her pair are from the social group. In columns 2 and 3, either the student or the pair belongs to the unsocial group. Last, in column 4, both are from the unsocial group.[26]

The first notable finding is that overall, the estimates in column 1 tend to be positive, large, and statistically significant (Hs, Sc, and Si). The results imply that peer effects arise when two students are likely to be social (as they belong to social groups). On the other hand, for Pd and Ma, the estimates in column 4 are statistically significant. The results suggest that the peer effect is significant when both students are from the unsocial groups.[27]

One possible interpretation of our results is that peer effects could be either positive or negative. It is likely that a positive effect may arise between two social students, while a negative effect can arise between two unsocial ones. In both cases, we should find a positive correlation among peers. Of course, there might be many other interpretations. For example, it is conceivable that peer effects or students' interactions occur only among students of a similar kind.

Although being far from conclusive due to data limitation, we believe the results here are the first bit of evidence in economics literature that social personality traits are some intermediary elements for peer effects to arise.[28] We believe that this is an interesting avenue for future research. Peer effects arise from students' interactions, which depend on their social relationships. We believe that this is an important point in that any sorting or tracking policies should consider the way that students join social groups or make friends in response to such policies.

[25]In Table A.4, we used alternative definitions of social and unsocial groups and checked the robustness of our results. We also present the IV results for those types that turn out to be significant in the class-FE models.

[26]The sample size is not equal across subsamples, although we equally divided the whole sample into social and unsocial groups. This is because students are sorted based on the variables that we used so as to define the group (gender, major, and entering year).

[27]Explaining why peer effects arise between social groups defined by Hs, Sc, and Si and between unsocial groups defined by Pd and Ma is beyond the scope of this paper. Some types (e.g., Pd and Si) seem to be directly related to social interactions, because they have direct implications for social relationships. However, for other types, it is difficult to explain why they would matter for peer effects.

[28]Some studies in education literature have found that personality matters for peer interaction, particularly in small groups. For example, an earlier study found that extroverted students were more likely to receive help from their peers than introverted students (Webb (1982)).

TABLE 7. Peer effects by personality.

Type of MMPI Scale	Own Group Personality / Pair's Group Personality	(1) Social Social	(2) Social Unsocial	(3) Unsocial Social	(4) Unsocial Unsocial
Hs (Hypochondriasis)	Pair's midterm	0.0732**	−0.0228	0.0345	−0.0028
		(0.0339)	(0.0859)	(0.0822)	(0.0311)
	Observations	570	116	117	315
D (Depression)	Pair's midterm	0.0333	0.0783	0.1068	0.0063
		(0.0476)	(0.0751)	(0.0771)	(0.0304)
	Observations	420	148	130	420
Hy (Hysteria)	Pair's midterm	0.0473	−0.0579	0.0342	0.0150
		(0.0357)	(0.0869)	(0.1017)	(0.0354)
	Observations	511	128	130	349
Pd (Psychopathic deviate)	Pair's midterm	0.0216	−0.1597	−0.0492	0.1441***
		(0.0158)	(0.1340)	(0.0368)	(0.0484)
	Observations	538	123	125	332
Mf (Masculinity–feminity)	Pair's midterm	0.0457	−0.0956	−0.0192	0.0392
		(0.0390)	(0.0923)	(0.0613)	(0.0380)
	Observations	568	128	126	296
Pa (Paranoia)	Pair's midterm	0.0311	−0.0810	−0.0014	0.0830
		(0.0208)	(0.1202)	(0.0506)	(0.0544)
	Observations	560	110	109	339
Pt (Psychasthenia)	Pair's midterm	0.0629	−0.0160	0.0885	−0.0151
		(0.0478)	(0.0689)	(0.0635)	(0.0300)
	Observations	486	131	133	368
Sc (Schizophrenia)	Pair's midterm	0.0321*	−0.0981	0.0036	0.0664
		(0.0162)	(0.0600)	(0.0496)	(0.0519)
	Observations	606	125	126	261
Ma (Hypomania)	Pair's midterm	0.0009	−0.0335	0.0404	0.0459**
		(0.0378)	(0.1328)	(0.1324)	(0.0218)
	Observations	520	89	89	420
Si (Social introversion)	Pair's midterm	0.0688*	0.0212	0.0402	0.0074
		(0.0384)	(0.0630)	(0.0661)	(0.0394)
	Observations	398	235	106	379

Note: The Student Counseling Center reports the distribution of students with high MMPI scores by gender, entering year, and 13 majors, and the summary statistics are presented in Table 6. We linked the reported distribution into this study's student records according to gender, entering year, and major. For each type of MMPI scale, we evenly divided students into social and unsocial groups in terms of their cohort's respective MMPI distribution. An unsocial group includes students whose cohorts are more likely to comprise those whose MMPI score was over 70. Only 1114 students and their pairs could be classified into the two groups. We divided these students into four MMPI-score groups according to own and pair group personality (i.e., social or unsocial). Then we ran the FE regression baseline model for each MMPI-score group across 10 types of MMPI clinical scales. The table reports only the coefficients of the pair's midterm scores and their standard errors clustered by class. The asterisks *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels.

6. Conclusions

Social interactions are of great interest in economics. They occur in many different contexts, such as schools, neighborhoods, and workplaces. In particular, understanding students' interactions with other students within classrooms is important not only for designing various education policies, such as tracking and school vouchers, but also for better understanding knowledge spillovers among economic agents. In this paper, we contribute to the literature by looking inside classrooms and by examining students' interactions at the individual level. We also present an identification strategy that exploits repeated interactions among peers.

Our findings show that students are influenced by their peers, particularly by the ones seated next to them. The effect is sizable; having a student with a midterm test score higher by 1 standard deviation increases the final exam score by 0.12. This supports the argument that peer effects are underestimated since irrelevant peers are included. In fact, when we define peer groups more widely (including distant peers), we find smaller or little effects. We also find that peer effects differ across students' characteristics and classroom settings as well as over the test score distribution. Also peer effects are heterogeneous and nonlinear over the achievement distribution and significant among lower-performing students and among top students. Last, our findings suggest that social personality traits may play some intermediary roles for peer effects to arise. This last point is an important topic for future research. Understanding channels of peer effects is needed to promote positive spillovers.

Appendix A: Extended model

In this Appendix, we allow that the pair student's midterm score can directly affect one's final examination score. The idea is that the peer's performance revealed as the midterm score in the middle of the course may motivate the student in the remaining part of the semester,

$$Y_{icf} = \rho Y_{icm} + \delta Y_{jcm} + X_{ic}\alpha_{xf} + X_{jc}\widetilde{\alpha}_{xf} + \alpha_{uf}u_i + \widetilde{\alpha}_{uf}u_j + \mu_{cf} + \varepsilon_{icf},$$

where δ captures the effect of the pair's midterm score. The equation for the final score is modified as

$$\begin{aligned} Y_{icf} = {} & \left\{\rho + \alpha_{uf} + \frac{\widetilde{\alpha}_{um}(\widetilde{\alpha}_{uf} - \widetilde{\alpha}_{um}\alpha_{uf})}{\widetilde{\alpha}_{um}^2 - 1}\right\} Y_{icm} + \left\{\frac{\widetilde{\alpha}_{um}\alpha_{uf} - \widetilde{\alpha}_{uf}}{\widetilde{\alpha}_{um}^2 - 1} + \delta\right\} Y_{jcm} \\ & + \lambda X_{ic} + \widetilde{\lambda} X_{jc} + \mu_c + e_{icf}. \end{aligned}$$

Then we have the relationship

$$\tilde{\beta} = \frac{\widetilde{\alpha}_{um}\alpha_{uf} - \widetilde{\alpha}_{uf}}{\widetilde{\alpha}_{um}^2 - 1} + \delta \leq \widetilde{\alpha}_{uf} + \delta.$$

Note that the estimation of $\tilde{\beta}$ now provides a lower bound of the combined effect of the peer's unobservable ability and the peer's observable performance.

Appendix B: Validity condition for IV

To think about the validity condition, suppose a reduced-form equation for course grade points as (without loss of generality, we ignore observable characteristics)

$$y_{ic} = \delta u_i + \tilde{\delta} u_{jc} + \mu_c + \varepsilon_{ic},$$

where y_{ic} is the grade points of student i for individual course c that the student took prior to the semester for which we measure the peer effect. Student jc is the student's peer at course c. This peer can be the same or different from the current peer.[29] Keeping our notations in the paper, u represents unobservable ability, μ_c is the course-specific common shock, and ε_{ic} is the individual-by-course error term. GPA is the average of grade points from all the courses that the student took previously (say, N courses):

$$GPA_i \equiv \frac{1}{N}\sum_c y_{ic} = \delta u_i + \tilde{\delta}\frac{1}{N}\sum_c u_{jc} + \frac{1}{N}\sum_c \mu_c + \frac{1}{N}\sum_c \varepsilon_{ic}.$$

There are two notable things in the above equation. First, it is obvious why GPA is a good proxy for own unobservable ability, u_i. The equation shows that GPA should satisfy the relevance condition for the IV. This is also why we used GPA as a direct measure of unobservable ability and estimated a reduced-form equation (see Panel A in Table 3).

Second, using the equation for GPA, we can write out the validity condition. For GPA to be a valid IV, it should be uncorrelated with ε_{icm}, ε_{jcm}, and ε_{icf} in the composite error term of Equation (3). The validity condition is therefore that those error terms are uncorrelated with the average of individual-course idiosyncratic shocks, $\frac{1}{N}\sum_c \varepsilon_{ic}$. This condition is satisfied if the three error terms are transitory shocks, not serially correlated after conditioning on time-invariant unobservable ability and class-specific fixed effects. Furthermore, the error term in GPA is the average of individual-course shocks, so the correlation between the error term in GPA and the composite error term is likely to be weak.

Appendix C: Supplementary tables and figures

In the following, we add supplementary tables and figures. Table A.1 reports the full results of the IV regressions whose key results are presented in Table 3 in the text. Table A.2 reports the results of the estimation where we apply IV regression models to the subsample analysis. Table A.3 reports the estimates of peer effects using raw test scores rather than standardized scores. In Table A.4, we use alternative definitions of social and unsocial groups as robustness check for Table 7 in the text. On the other hand, Figure A.1 provides some classroom photos, which show that pair students are physically close and are seated right next to each other. Figure A.2 plots the coefficients of own midterm score estimated by quantile and IV-quantile regressions for various subsamples.

[29] We checked the robustness of our results after excluding those pairs who have ever taken any of the same classes before.

TABLE A.1. Instrumental variable (IV) estimates of peer effects (dependent variable: own final score).

	(2) Class FE	(3) Row-by-Class FE	(4) Paired Students	(5) Adding Tier-Two Peers
Class FE	Y	N	Y	Y
Row-by-Class FE	N	Y	N	N
Pair or Peer Characteristics				
Pair's midterm score	0.1171***	0.1240**	0.1140***	0.1186***
	(0.0446)	(0.0544)	(0.0442)	(0.0456)
No pair	−0.0287	−0.0617		−0.0305
	(0.0602)	(0.0514)		(0.0606)
Tier-one peers:				
Avg. midterm score	0.0426	0.0808**	0.0346	0.0444
	(0.0357)	(0.0354)	(0.0479)	(0.0364)
Grade composition	−0.0306**	−0.0346*	−0.0079	−0.0311**
	(0.0150)	(0.0208)	(0.0211)	(0.0152)
Gender composition	−0.0547	−0.0259	−0.1616*	−0.0576
	(0.0708)	(0.0693)	(0.0835)	(0.0709)
Tier-two peers:				0.0080
Avg. midterm score				(0.0397)
Own characteristics				
Midterm score	1.0915***	1.0886***	1.0831***	1.0915***
	(0.0528)	(0.0543)	(0.0549)	(0.0532)
Grade	0.0531***	0.0508***	0.0467***	0.0533***
	(0.0130)	(0.0129)	(0.0152)	(0.0132)
Boy	−0.0527	−0.0480	−0.0506	−0.0540*
	(0.0322)	(0.0372)	(0.0377)	(0.0320)
Row of seat	0.0047		0.0023	0.0050
	(0.0051)		(0.0057)	(0.0052)
No pair	0.0778**	0.0931**	0.0655	0.0783**
	(0.0381)	(0.0414)	(0.0472)	(0.0382)
Major: Business	0.0205	0.0433	0.0305	0.0168
	(0.0550)	(0.0549)	(0.0518)	(0.0553)
Humanity or social science	−0.0274	−0.0066	−0.0052	−0.0270
	(0.0549)	(0.0521)	(0.0578)	(0.0553)
Natural science or engineering	0.0120	0.0469	0.0225	0.0118
	(0.0651)	(0.0637)	(0.0679)	(0.0654)
Retaking	0.0352	0.0347	0.0076	0.0341
	(0.0465)	(0.0526)	(0.0546)	(0.0467)
Exchange student	0.2557*	0.1760	0.4085*	0.2788**
	(0.1353)	(0.1756)	(0.2207)	(0.1364)
Constant term	−0.0690	−0.0646	−0.0939	−0.0672
	(0.0586)	(0.1092)	(0.0629)	(0.0596)
Observations	4155	4155	3437	4144
Adj. R-squared	0.1458	0.2337	0.1148	0.1432

Note: The table reports the full IV estimation results of the second stage in Panel B of Table 3. Robust standard errors are in parentheses and are clustered by class for models (1), (3), and (4) or by row-by-class group for model (2). *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels.

Table A.2. Instrumental variable estimates of peer effects by subgroup (dependent variable: own final score).

	(1) Boys	(2) Girls	(3) Up to Fifth Rows	(4) After Fifth Rows	(5) Class in English	(6) Class in Korean	(7) Economics Major	(8) Non-Economics Major
Pair or Peer Characteristics								
Pair's midterm score	0.119**	0.113	−0.028	0.189***	0.188	0.095**	0.043	0.162**
	(0.056)	(0.078)	(0.074)	(0.062)	(0.115)	(0.039)	(0.057)	(0.079)
No pair	−0.052	0.027	0.019	−0.061	−0.075	0.013	−0.038	−0.024
	(0.066)	(0.077)	(0.075)	(0.079)	(0.105)	(0.074)	(0.072)	(0.079)
Tier-one peers:								
Avg. midterm score	0.068**	−0.019	0.013	0.092**	0.077	0.025	0.002	0.088***
	(0.027)	(0.065)	(0.051)	(0.037)	(0.049)	(0.046)	(0.051)	(0.032)
Grade composition	−0.058***	−0.009	−0.001	−0.038*	−0.079***	−0.015	0.016	−0.073***
	(0.019)	(0.029)	(0.032)	(0.019)	(0.023)	(0.017)	(0.027)	(0.025)
Gender composition	−0.031	−0.130	−0.131	0.012	0.100	−0.122	−0.041	−0.080
	(0.080)	(0.150)	(0.091)	(0.114)	(0.117)	(0.090)	(0.101)	(0.094)

(*Continues*)

Table A.2. *Continued.*

	(1) Boys	(2) Girls	(3) Up to Fifth Rows	(4) After Fifth Rows	(5) Class in English	(6) Class in Korean	(7) Economics Major	(8) Non-Economics Major
Own characteristics								
Midterm score	1.066***	1.131***	1.095***	1.088***	1.249***	1.025***	1.172***	1.036***
	(0.053)	(0.082)	(0.067)	(0.070)	(0.086)	(0.060)	(0.084)	(0.078)
Grade	0.072***	0.041	0.054**	0.053***	0.084***	0.038**	0.047***	0.073***
	(0.017)	(0.026)	(0.022)	(0.016)	(0.021)	(0.015)	(0.016)	(0.018)
Boy			−0.087**	−0.011	−0.086**	−0.034	−0.059	−0.039
			(0.044)	(0.039)	(0.040)	(0.046)	(0.049)	(0.037)
Row of seat	0.008	0.000	0.023*	0.010	0.012	0.004	0.005	0.002
	(0.006)	(0.010)	(0.014)	(0.010)	(0.011)	(0.005)	(0.007)	(0.006)
No pair	0.060	0.118*	−0.021	0.157***	0.080	0.070	0.068	0.083
	(0.053)	(0.061)	(0.052)	(0.060)	(0.073)	(0.046)	(0.068)	(0.056)
Major: Business	0.024	−0.014	0.080	−0.050	−0.098	0.088		0.035
	(0.066)	(0.084)	(0.078)	(0.071)	(0.094)	(0.068)		(0.049)
Humanity or social science	−0.017	−0.068	−0.002	−0.072	−0.203**	0.047		−0.029
	(0.064)	(0.080)	(0.064)	(0.086)	(0.096)	(0.071)		(0.057)
Natural science or engineering	−0.015	0.021	−0.099	0.060	−0.244**	0.110		
	(0.103)	(0.076)	(0.085)	(0.084)	(0.112)	(0.071)		
Retaking	0.023	0.035	0.017	0.069	−0.068	0.111**	−0.031	0.039
	(0.060)	(0.074)	(0.056)	(0.061)	(0.072)	(0.046)	(0.083)	(0.057)
Exchange student	0.586***	−0.062	0.402*	0.129	0.390**	0.007		0.226
	(0.227)	(0.175)	(0.223)	(0.173)	(0.173)	(0.440)		(0.183)
Constant term	−0.105	−0.042	0.019	−0.248***	0.081	−0.109	−0.205***	0.003
	(0.088)	(0.095)	(0.068)	(0.096)	(0.100)	(0.071)	(0.075)	(0.069)
Observations	2449	1706	1953	2202	1471	2684	2223	1932
Adj. R-squared	0.165	0.138	0.127	0.167	−0.016	0.202	0.126	0.159

Note: The table reports the full IV estimation results of Panel B of Table 5. Robust standard errors are in parentheses and are clustered by class. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels.

TABLE A.3. Estimates of the peer effect using raw test score (dependent variable: own final score (raw test score)).

		(1) All	(2) Boys	(3) Girls	(4) Up to Fifth Rows	(5) After Fifth Rows	(6) Class in English	(7) Class in Korean	(8) Economics Major	(9) Non-Economics Major
Panel A: Raw Test Score										
Final	Mean	0.639	0.642	0.635	0.658	0.623	0.660	0.628	0.650	0.627
	S.D.	0.232	0.234	0.228	0.224	0.237	0.230	0.232	0.232	0.230
Midterm	Mean	0.696	0.696	0.696	0.709	0.685	0.717	0.684	0.710	0.680
	S.D.	0.198	0.202	0.193	0.191	0.205	0.194	0.200	0.197	0.199
Panel B: OLS Regression										
Pair's midterm score		0.0383**	0.0518**	0.0111	0.0103	0.0577**	0.0371	0.0389*	0.0124	0.0643***
		(0.0165)	(0.0219)	(0.0214)	(0.0222)	(0.0228)	(0.0280)	(0.0221)	(0.0224)	(0.0230)
Adj. R-squared		0.5094	0.5097	0.5125	0.4917	0.5190	0.4936	0.5180	0.5092	0.5175
Panel C: 2SLS Regression										
Pair's midterm score		0.1278**	0.1076	0.1632*	−0.0321	0.2025***	0.2230	0.1328**	0.0713	0.1827*
		(0.0590)	(0.0682)	(0.0958)	(0.0956)	(0.0721)	(0.1448)	(0.0568)	(0.0749)	(0.0971)
R-squared		0.1937	0.2238	0.1399	0.1624	0.2092	−0.0538	0.2840	0.2160	0.1517
Observations		4155	2449	1706	1953	2202	1471	2684	2223	1932

Note: In the table, we estimate the peer effect using the raw test score (set max to 1) instead of the within-class standardized test score. In Panel B, we use the model of column 2 in Table 2. In Panel C, we use the model of column 1 in Table 3. In all regressions, we control for class fixed effects. Robust standard errors are in parentheses and are clustered by class. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels.

Table A.4. Alternative definitions of social/unsocial groups.

	Cutoff of High MMPI Score: >70				Cutoff of High MMPI Score: >60			
Personality	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Own	Social	Social	Unsocial	Unsocial	Social	Social	Unsocial	Unsocial
Pair	Social	Unsocial	Social	Unsocial	Social	Unsocial	Social	Unsocial
Panel A: Hs (Hypochondriasis)								
OLS	0.0732**	−0.0228	0.0345	−0.0028	0.0047	0.1187	0.0069	0.0602
	(0.0339)	(0.0859)	(0.0822)	(0.0311)	(0.0269)	(0.1271)	(0.0449)	(0.0446)
IV	0.3321**	−0.1822	0.0038	−0.0415	−0.0352	−0.0890	0.2920	0.2496
	(0.1382)	(0.3462)	(0.1517)	(0.1330)	(0.2470)	(0.2471)	(0.2533)	(0.2273)
Panel B: Pd (Psychopathic Deviate)								
OLS	0.0216	−0.1597	−0.0492	0.1441***	0.0200	−0.0813	−0.1013	0.1034*
	(0.0158)	(0.1340)	(0.0368)	(0.0484)	(0.0182)	(0.1828)	(0.0653)	(0.0526)
IV	0.0238	0.2562	0.3193	0.3110***	0.0632	0.0549	0.1240	0.3114**
	(0.1342)	(0.5528)	(0.3173)	(0.1057)	(0.1911)	(0.4674)	(0.1718)	(0.1394)
Panel C: Sc (Schizophrenia)								
OLS	0.0321*	−0.0981	0.0036	0.0664	0.0220	−0.0328	0.0619	0.0851*
	(0.0162)	(0.0600)	(0.0496)	(0.0519)	(0.0287)	(0.1445)	(0.0633)	(0.0495)
IV	0.1324	−0.0556	0.1392	0.0846	0.2761	−0.1310	0.3076	0.1438
	(0.1924)	(0.1814)	(0.2496)	(0.1397)	(0.2173)	(0.2291)	(0.3317)	(0.1438)
Panel D: Ma (Hypomania)								
OLS	0.0009	−0.0335	0.0404	0.0459**	0.0496	0.1148	−0.0045	0.0056
	(0.0378)	(0.1328)	(0.1324)	(0.0218)	(0.0474)	(0.0867)	(0.0868)	(0.0293)
IV	0.1342	−0.1946	0.4858	0.1259	0.1100	0.6605	0.0062	0.3417
	(0.1273)	(0.2298)	(0.6015)	(0.1670)	(0.1456)	(0.4211)	(0.2018)	(0.4829)
Panel E: Si (Social Introversion)								
OLS	0.0688*	0.0212	0.0402	0.0074	0.1022***	−0.0073	0.0734	−0.0358
	(0.0384)	(0.0630)	(0.0661)	(0.0394)	(0.0349)	(0.0750)	(0.0842)	(0.0320)
IV	0.1857	0.1353	−0.0039	−0.0958	0.2898**	0.3055	0.0021	−0.1635
	(0.1614)	(0.2876)	(0.1826)	(0.1478)	(0.1296)	(0.5001)	(0.1257)	(0.1171)

Note: We selected five MMPI scales (Hs, Pd, Sc, Ma, and Si) whose OLS estimated coefficients are statistically significant in Table 7. The OLS results in models (1)–(4) are the same as those in Table 7. In models (5)–(8), we use the alternative definition of social/unsocial groups using 60 as the cutoff of high MMPI scores rather than 70. In each model, we use the model of column (1) in Table 3 to conduct IV estimations. In all regressions, we control for class fixed effects. Robust standard errors are in parentheses and are clustered by class. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels.

Figure A.1. Classroom photos.

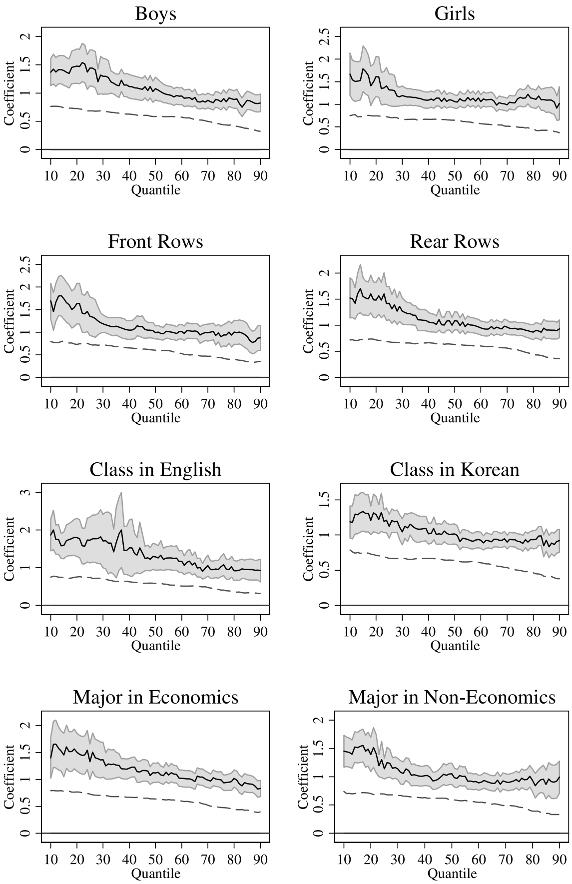

Figure A.2. Instrumental variable quantile regression estimation by subsample: coefficients of own midterm scores. *Note*: We ran QRs and IVQRs for the eight subsamples. Each figure shows the coefficients of own midterm score estimated using the QR (dashed line) and IVQR (solid line). The shaded area indicates 95 percent confidence interval of IVQR estimation.

References

Altonji, J. G., C.-I. Huang, and C. R. Taber (2010), "Estimating the cream skimming effect of school choice." NBER Working Paper 16579. [245]

Angrist, J. D. and K. Lang (2004), "Does school integration generate peer effects? Evidence from Boston's Metco Program." *American Economic Review*, 94 (5), 1613–1634. [239]

Arcidiacono, P., G. Foster, N. Goodpaster, and J. Kinsler (2012), "Estimating spillovers using panel data, with an application to the classroom." *Quantitative Economics*, 3 (3), 421–470. [241, 245, 254]

Arcidiacono, P. and S. Nicholson (2005), "Peer effects in medical school." *Journal of Public Economics*, 89 (2–3), 327–350. [240]

Benedict, M. E. and J. Hoag (2004), "Seating location in large lectures: Are seating preferences or location related to course performance?" *Journal of Economic Education*, 35 (3), 215–231. [250]

Betts, J. and A. Zau (2004), "Peer groups and academic achievement: Panel evidence from administrative data." Working Paper, University of California, San Diego. [240]

Boucher, V., Y. Bramoullé, H. Djebbari, and B. Fortin (2014), "Do peers affect student achievement? Evidence from Canada using group size variation." *Journal of Applied Econometrics*, 29 (1), 91–109. [246]

Brunello, G., M. de Paola, and V. Scoppa (2010), "Peer effects in higher education: Does the field of study matter?" *Economic Inquiry*, 48 (3), 621–634. [254]

Buechel, B., L. Mechtenberg, and J. Petersen (2014), "Peer effects and students' self-control." MPRA Paper 53658, University Library of Munich, Germany. [260]

Burke, M. A. and T. R. Sass (2013), "Classroom peer effects and student achievement." *Journal of Labor Economics*, 31 (1), 51–82. [240, 254]

Butcher, J. N. and S. V. Rouse (1996), "PERSONALITY: Individual differences and clinical assessment." *Annual Review of Psychology*, 47, 87–111. [260]

Carrell, S. E., R. L. Fullerton, and J. F. West (2009), "Does your cohort matter? Measuring peer effects in college achievement." *Journal of Labor Economics*, 27 (3), 439–464. [239]

Carrell, S. E., B. I. Sacerdote, and J. E. West (2013), "From natural variation to optimal policy? The importance of endogenous peer group formation." *Econometrica*, 81 (3), 855–882. [260]

Chernozhukov, V. and C. Hansen (2005), "An IV model of quantile treatment effects." *Econometrica*, 73 (1), 245–261. [257]

Clotfelter, C., H. Ladd, and J. Vigdor (2007), "Teacher credentials and student achievement: Longitudinal analysis with student fixed effects." *Economics of Education Review*, 26 (6), 673–682. [250]

Cox, A. C., N. Weed, and J. N. Butcher (2009), "The MMPI-2: History, interpretation, and clinical issues." In *Oxford Handbook of Personality Assessment* (J. N. Butcher, ed.), 250–276, Oxford University Press, New York. [261]

Ding, W. and S. F. Lehrer (2007), "Do peers affect student achievement in China's secondary schools?" *Review of Economics and Statistics*, 89 (2), 300–312. [239]

Duflo, E., P. Dupas, and M. Kremer (2011), "Peer effects and the impacts of tracking: Evidence from a randomized evaluation in Kenya." *American Economic Review*, 101 (5), 1739–1774. [240]

Epple, D. and R. E. Romano (2011), "Peer effects in education: A survey of the theory and evidence." In *Handbook of Social Economics*, Vol. 1B (J. Benhabib, A. Bisin, and M. O. Jackson, eds.), 1053–1164, North-Holland, Amsterdam. [240]

Foster, G. (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 (8–9), 1455–1475. [239, 246]

Frank, R. H., T. Gilovich, and D. T. Regan (1993), "Does studying economics inhibit cooperation?" *Journal of Economic Perspectives*, 7 (2), 159–171. [257]

Fruehwirth, J. C. (2013), "Identifying peer achievement spillovers: Implications for desegregation and the achievement gap." *Quantitative Economics*, 4 (1), 85–124. [241, 244, 245]

Fruehwirth, J. C. (2014), "Can achievement peer effect estimates inform policy? A view from inside the black box." *Review of Economics and Statistics*, 96 (3), 514–523. [241, 244, 245, 248]

Hanushek, E. A., J. F. Kain, J. M. Markman, and S. G. Rivkin (2003), "Does peer ability affect student achievement?" *Journal of Applied Econometrics*, 18 (5), 527–544. [240, 241, 254]

Hanushek, E. A. (1979), "Conceptual and empirical issues in the estimation of educational production functions." *Journal of Human Resources*, 14 (3), 351–388. [247]

Hoxby, C. M. (2000), "Peer effects in the classroom: Learning from gender and race variation." NBER Working Papers 7867. [239, 254]

Hoxby, C. M. and G. Weingarth (2005), "Taking race out of the equation: School reassignment and the structure of peer effects." Working paper. [254, 258]

Jaffe, A. B., M. Trajtenberg, and R. Henderson (1993), "Geographic localization of knowledge spillovers as evidenced by patent citations." *Quarterly Journal of Economics*, 108 (3), 577–598. [244]

Kang, C. (2007), "Classroom peer effects and academic achievement: Quasi-randomization evidence from South Korea." *Journal of Urban Economics*, 61 (3), 458–495. [239]

Keller, W. (2004), "International technology diffusion." *Journal of Economic Literature*, 42 (3), 752–782. [244]

Lavy, V. and A. Schlosser (2011), "Mechanisms and impacts of gender peer effects at school." *American Economic Journal: Applied Economics*, 3, 1–33. [254, 257]

Lazear, E. P. (2001), "Educational production." *Quarterly Journal of Economics*, 116 (3), 777–803. [257]

Lefgren, L. L. (2004), "Educational peer effects and the Chicago public schools." *Journal of Urban Economics*, 56 (2), 169–191. [254]

Manski, C. F. (1993), "Identification of endogenous social effects: The reflection problem." *Review of Economic Studies*, 60 (3), 531–542. [239, 247]

Sacerdote, B. (2001), "Peer effects with random assignment: Results from dartmouth roommates." *Quarterly Journal of Economics*, 116, 681–704. [239]

Sacerdote, B. (2011), "Peer effects in education: How might they work, how big are they and how much do we know thus far?" In *Handbook of the Economics of Education*, Vol. 3 (E. A. Hanushek, S. Machin, and L. Woessmann, eds.), 249–277, North-Holland, Amsterdam. [240, 250, 254, 258]

Sojourner, A. (2013), "Identification of peer effects with missing peer data: Evidence from project STAR." *Economic Journal*, 123 (569), 574–605. [240]

Stinebrickner, R. and T. R. Stinebrickner (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 (8–9), 1434–1454. [239]

Todd, P. E. and K. I. Wolpin (2003), "On the specification and estimation of the production function for cognitive achievement." *Economic Journal*, 113 (485), F3–F33. [248]

Vigdor, J. and T. Nechyba (2007), "Peer effects in North Carolina public schools." In *Schools and the Equal Opportunity Problem* (L. Woessmann and P. E. Peterson, eds.), 73–101, MIT Press, Cambridge, MA. [239, 250]

Webb, N. M. (1982), "Group composition, group interaction and achievement in cooperative small groups." *Journal of Educational Psychology*, 50, 266–276. [262]

Zimmerman, D. J. (2003), "Peer effects in academic outcomes: Evidence from a natural experiment." *Review of Economics and Statistics*, 85 (1), 9–23. [239]

Co-editor Petra E. Todd handled this manuscript.

Manuscript received 12 February, 2014; final version accepted 13 March, 2016; available online 11 April, 2016.