Volume Title: College Choices: The Economics of Where to Go, When to Go, and How to Pay For It

Volume Author/Editor: Caroline M. Hoxby, editor

Volume Publisher: University of Chicago Press

Volume ISBN: 0-226-35535-7

Volume URL: http://www.nber.org/books/hoxb04-1

Conference Date: August 13-15, 2002

Publication Date: September 2004

Title: Peer Effects in Higher Education

Author: Gordon Winston, David Zimmerman

URL: http://www.nber.org/chapters/c10105

9

Peer Effects in Higher Education

Gordon C. Winston and David J. Zimmerman

9.1 Introduction

The existence and nature of peer effects are fundamental to understanding a variety of crucial issues facing both higher and lower education. Peer effects have played an important role in studies of primary and secondary education beginning when the prominent Coleman Report of 1966 claimed their centrality in the determination of childrens' schooling outcomes (Coleman, Campbell, Hobson, McPartland, Mood, Weinfeld, and York 1966). Arguments based on peer effects have been used to justify busing and have entered the debates on educational costs, on tracking, on distance learning, affirmative action, and on the effects of voucher systems (U.S. Supreme Court 1971; Summers and Wolfe 1977; Hanushek 1986; Robertson and Symons 1996; Epple and Romano 1998; Lazear 1999; Hoxby 2000). The relevance of peer effects to the economics of higher education has only recently been acknowledged (Rothschild and White 1995; Winston and Yen 1995; Winston 1999; Epple, Romano, and Sieg 2001) and has only a small, if growing, empirical basis (Zimmerman 1999; Stinebrickner

Gordon C. Winston is Orrin Sage Professor of Political Economy at Williams College. David J. Zimmerman is professor of economics and department chair at Williams College, and a research associate of the National Bureau of Economic Research (NBER).

We want to thank Atlantic Philanthropic Services and especially the Andrew W. Mellon Foundation for its continuing support of the Williams College Project on the Economics of Higher Education and for making their unique College and Beyond data set available for this study. Al Goethals, Georgi Zhelev, and Adam Sischy contributed a good deal to the chapter, while Caroline Hoxby went way beyond the usual volume editor in her interest in and contribution to the work. Finally, the participants in the NBER seminar at Bretton Woods—especially Tom Dee—were very helpful in spotting ambiguities and dubious assertions in an earlier draft.

and Stinebrickner 2000; Goethals 2001; Sacerdote 2001). In this chapter we will describe the importance of peer effects in some detail and will then offer new empirical evidence on their existence.

Estimating peer effects is difficult. First, we must decide on the appropriate set of educational outcomes believed to be sensitive to peer attributes. Second, we must specify the relevant peer attributes. Third, and perhaps most difficult, we must contend with the fact that *selection bias* is rampant in the estimation of peer effects. In a laboratory setting, we might randomly assign a set of subjects to different peer environments and gauge any resulting effects.[1] In the world of nonexperimentally derived observations, however, we must recognize that people choose their peers. And if people tend to associate with others with similar traits (many of which are likely to be unobservable to the researcher), then it is exceedingly difficult to determine whether we are observing peer effects or simply observing similar people behaving similarly. In this paper we use a unique data set that combines data for three schools from the Andrew W. Mellon Foundation's College and Beyond data for the entering class of 1989, along with phonebook data identifying roommates, to implement a quasi-experimental empirical strategy aimed at measuring peer effects in academic outcomes. In particular, we use data on individual student's grades, Scholastic Aptitude Test (SAT) scores, and the SAT scores of their roommates to estimate the effect of roommates' academic characteristics on an individual's grades. The schools selected for the analysis were chosen because their housing assignment protocols appear (near) random for first-year students. The results suggest that, for two of the three schools used, students in the middle of the SAT distribution may do somewhat worse in terms of grades if they share a room with a student who is in the bottom 15 percent of the SAT distribution. Students in the top of the SAT distribution are typically not affected by the SAT scores of their roommates.

Sections 9.2 through 9.4 of this chapter define peer effects and indicate the importance of peer effects in understanding several fundamental characteristics of higher education. Sections 9.6 through 9.9 review the existing empirical evidence and provide new evidence. Section 9.10 concludes.

9.2 Peer Effects in the Economics of Higher Education

Peer effects *exist* when a person's behavior is affected by his or her interaction with peers—"equals"—so in higher education peer effects result from interactions between students.[2] While peer quality is often included

1. For an interesting strategy to estimate peers in an experimental context, see Goethals (2001).

2. Peer effects among faculty (and administrators) can be important, too, of course, to recruiting, teaching, and scholarly productivity, but they are not the issue here (Rosovsky 1990; Kennedy 1997).

as an argument in an educational production function, it is useful to put it more directly as an interaction between two (for simplicity) students,

$$B_1 = f(B_2, C_2, \mathbf{X}), \tag{1}$$

where B_i is behavior, C_i is characteristics for students $i = 1, 2$, and $\mathbf{X}$ is a vector of other things relevant to 1's behavior. Peer effects *exist* if the partial derivatives of equation (1) are not zero and they are asymmetric (or nonlinear) if those partials differ at different levels of B and C.

Peer effects, we will argue, are relevant to the economics of higher education in several important ways:

1. They eliminate awkward anomalies in the institutional behavior of colleges and universities and in the economic structure of higher education as an industry if they exist.
2. They might justify, as economically efficient, the observed segmentation of student quality and resources if peer effects are appropriately nonlinear.
3. They lead to trade in peer quality in an input market inextricably linked with that for educational services. Both of those markets and their interaction appear essential to understanding pricing, admissions, and competition in higher education (Rothschild and White 1995; Winston 2003).

We will focus on the first of these claims. We will examine the second but express our frustration because, while they are potentially important, the empirical evidence gives only hints about their nature. And we will do little more than note the third—the blended markets for educational services and peer quality—because its discussion would require a chapter to itself.

9.3 Peer Effects and Economic Anomalies in Higher Education

9.3.1 Anomalies

Higher education looks much like a normal industry that makes a product (educational services) using purchased inputs that it sells to customers for a price in a quite competitive market. There are, however, some fundamental differences:

- Colleges always charge a price that fails—significantly—to cover their production costs.
- They turn away a majority of potential customers who are willing and able to buy their product if they can.
- They don't expand output to meet persistent excess demand.
- They lower the price to attract one customer, replacing another who would pay a higher price.

- They judge institutional quality by how many customers they can turn away, and they may manipulate sales-admission policies to increase that number.[3]
- They require elaborate application procedures before one is allowed to make a purchase.
- They practice extensive price discrimination not only to increase sales revenues but often to redistribute income among their customers.

9.3.2 Economic Characteristics

Those anomalies disappear and higher education becomes an economically coherent industry if four economic characteristics are, in fact, typical of colleges, universities, and higher education:

1. If charitable donations significantly augment schools' commercial (sales, tuition) revenues (Hansmann 1980)
2. If those donated resources are unevenly distributed, supporting a hierarchy of schools based on their independence from sales revenues
3. If colleges and universities are less interested in profits than in a "mission" and in "achieving excellence or prestige"
4. If students provide an input critical to the production of higher education and if peer effects are important to educational output

9.3.3 Evidence of Those Characteristics

Donations—Noncommercial Resources

It is well established that colleges and universities charge prices that are less than the costs of production. National Integrated Postsecondary Education Data System (IPEDS) data for more than 2,800 U.S. colleges and universities support the data in table 9.1. Averaged over colleges and universities in both public and private sectors in 1995–1996, tuition revenues support only a fraction of the cost of producing a student's education; the rest was covered by donations (past and present, public and private). The price-cost ratio averaged 0.32 and ranged from an average of 0.13 in the public sector to 0.45 among private schools. In a more complete analysis of IPEDS data that also recognized collegiate saving, it appeared that 75 percent of the economic resources used in higher education came from charitable contributions—only 25 percent came from commercial sales (Winston, Carbone, forthcoming).

The Uneven Distribution of Donations

The bottom part of table 9.1 indicates how unevenly those donations, and hence the student subsidies they support, are distributed among

3. On early decision, see Avery, Fairbanks et al. 2001; Fallows 2001.

Table 9.1 The Distribution of Average Cost, Price, and Student Subsidies, 1996

	Subsidy per Student ($)	Average Educational Cost ($)	Average Net Tuition ($)
All college and universities	8,423	12,413	3,989
Public	8,590	9,896	1,305
Private	8,253	14,986	6,734
Schools ranked by student subsidies			
Decile 1	20,991	27,054	6,063
Decile 2	11,865	15,801	3,936
Decile 3	10,009	13,310	3,301
Decile 4	8,752	11,831	3,080
Decile 5	7,855	10,565	2,710
Decile 6	7,020	9,820	2,799
Decile 7	6,250	9,464	3,214
Decile 8	5,447	8,848	3,401
Decile 9	4,262	9,297	5,035
Decile 10	1,736	8,084	6,348

Source: Based on U.S. Department of Education IPEDS data. Includes 2,791 institutions, of which 1,411 are public and 1,380 are private. All dollar amounts are per full-time equivalent student averaged over institutions. See Winston (forthcoming) and Winston and Yen (1995) for details on the derivation of these data from the IPEDS Finance Survey (medical schools are omitted here).

schools. The average student at a school in the top decile got a subsidy of $21,000 a year—paying $6,063 for a $27,054 education—while a typical student in a bottom-decile school got $1,700—paying $6,348 for an $8,084 education.

A Nonprofit Objective Function

The third economic characteristic—institutional objective function—can't be supported by data, of course. However, the existence of nonprofit behavior like need-blind admissions with need-based financial aid, along with policies like Berea College's zero tuition combined with its family income cap, strongly imply it. And economists—including Hansmann, James (1990), Clotfelter (James 1978; Hansmann 1980; Clotfelter 1996), and others—have described the objective function for a college in terms of excellence (or prestige) and mission.

9.4 The Existence of Peer Effects

The fourth characteristic—that peer effects are important to educational output—is central. If peer effects exist, they could motivate the stratification of students and resulting concentration of student quality in those schools with the most noncommercial resources per student. Stratification, if peer effects exist, is the result of an efficiency wage (Akerlof and Yellen 1986; Winston forthcoming) in the form of a student subsidy paid to generate a queue of applicants from which the best, in terms of peer quality, are selected. All schools may value the educational quality that is provided by good students through peer effects, but peer quality is scarce, and those schools that are able to pay the most for it get the most of it. The uneven distribution of noncommercial resources evident in table 9.1 creates a hierarchy that supports the stratification of student quality and motivates the long-run supply restrictions on which that selectivity rests. So the existence of peer effects—in a world of unevenly distributed noncommercial revenues and an institutional devotion to excellence—would produce the industry structure we see.

It is worth noting that there are other production externalities of student quality, aside from the peer interactions on which we are concentrating—like an instructor's ability to assign more advanced readings to better students, to give more intense and efficient lectures, or to have more productive seminars. These might also be thought of as "peer effects" and may have the same sort of effect of making educational production a function of student quality and hence motivate segmentation as efficient. Our focus, however, is not on these types of production externalities but, rather, as previously mentioned, on the direct impact that peer behavior or characteristics have on academic outcomes.

9.5 Efficiency and the Asymmetry of Peer Effects

Hoxby (2000) noted that the existence, per se, of peer effects may leave any regrouping of students as a largely distributional matter. Resorting students creates winners and losers to the same extent under strictly symmetric peer effects. But if those peer effects are asymmetric so that students at different levels of behavior or characteristics are influenced differently by their interaction with others, then peer effects introduce an issue of economic efficiency, too. How students are grouped will affect the total amount of learning produced in given participants from given resources.

If weak students gain more from proximity to strong peers than the strong students lose from that association, overall learning would be increased by reducing stratification—a point made by McPherson and Schapiro (1990) in suggesting random assignment of students to colleges. But if asymmetries in peer effects run the other way so that strong students interacting with other strong students are also more sensitive to peer influence—gaining more in learning than would weak students in those circumstances—then stratification and segmentation could increase, not decrease, aggregate learning. In the extreme, stratification would be supported on grounds of efficiency if strong students were sensitive to peer quality at all levels while weak students were unaffected by peers at any level.

Yet framing the issue as one of "strong students" and "weak students," while it fits the empirical work that's been possible so far, masks a potentially important question of peer "distance"—how far apart the peers are in their behavior and characteristics. Are peer responses very different outside a "neighborhood" of proximity so that a slightly different peer is influential but someone very different is not?[4] It's certainly a question at the center of the stratification question—a strong student might typically inspire somewhat weaker peers, while intimidating those more distant from his or her abilities. For the strong student, moderately weaker peers might represent a challenge and a chance to learn by teaching, while much weaker peers would overwhelm the strong student. And numbers would play a role not captured in either our framing or our evidence; given differences and distances among peers, a student would likely respond differently to one such peer than to a whole school of them. So the shape of nonlinearities in peer effect responses would depend on both peer distance and numbers.

Finally, whatever the efficiency or inefficiency of higher education's existing stratification in producing aggregate learning, social policy would have to address the question recently raised with some force by Nicholas Lemann (1999a,b) about whether those high-ability students, after learning more from their expensive educations with strong and sensitive peers, *use* all that learning to do anything useful for society and whether their social marginal product justifies that selectivity. He argued that it doesn't.

9.6 The Evidence

9.6.1 The Ideal Data

As a transition from the potentially central economic role we have suggested—that student peer effects might play in higher education to the more modest empirical results we are able to report on next and add to—it's useful to describe the ideal data whose analysis would persuasively support that role. Inevitably, of course, the actual evidence must fall far short of perfection, but it is useful to see how and where.

The empirical test of the existence and shape of peer effects in colleges would ideally, in terms of equation (1), deal with the following:

- Student behavior, B_1, that is centrally relevant to the purposes of higher education, broadly defined to include, inter alia, the development of intellectual curiosity, persistence, acquisition of facts, humane values, aesthetic sensitivities, analytical and technical sophistication, social responsibility, and so on.

4. This, of course, is in keeping with the Manski and Wise observation that students "preferred to enroll in colleges where the average academic ability of the enrolled students was slightly higher than their own. Schools where the average SAT scores of entering freshmen were either too low or too high were relatively disfavored" (Manski and Wise 1983, 159).

- His or her behavior, B_1, and the characteristics and behavior of peers, B_2 and C_2, that were unambiguously measurable in order to investigate not only the sign of peer differences and response but also their magnitudes
- A large population of students that generated a good deal of variation in B_1, B_2, and C_2 and their interaction, describing different distances between peers to reveal neighborhood asymmetries and nonlinearities
- Truly random assignment of associations between students that eliminated preferences in peer association
- Variations in peer characteristics of communities to reveal any social critical mass in conditioning peer interactions

Data meeting these conditions would allow an effective test of the existence of peer effects and their nonlinearities or asymmetries. And they would eliminate misgivings about the importance of the peer behaviors and characteristics studied so far to higher education.

Inevitably, of course, the studies described in the rest of the chapter fall short of the ideal. Although selection bias has largely been avoided through the use of randomly assigned roommates and experimental groups, and the results consistently show the existence of peer influences on behaviors that are relevant to education, it remains that in measuring a student's grade point average (GPA) response (or test scores or retention or fraternity membership) to his roommate's SATs (or income or fraternity membership), we're looking at a fairly thin slice of student behaviors and characteristics that leaves out a whole lot of what is happening to shape higher education.

But we find optimism in that thinness. If evidence of student peer effects can be found in so narrow a range of academic characteristics and behaviors, it's hard not to believe that with a wider and more appropriate range they would appear with a good deal more strength. Indeed, in having to use such limited evidence for so broad an influence (and so sweeping a hypothesis), we didn't initially expect peer effects to be significantly evident. But we could neither conjure up more appropriate data nor convince ourselves that we could adequately account for selection effects in a more general population with broader behaviors.[5] We were trying to see if we could find an iceberg and feel confident that we've located the tip.

But clearly, it's been easier to find evidence of the existence of peer effects than to learn much about their nonlinearities. So these results do more to support the idea that peer effects help to *explain* industry structure and selectivity in higher education—their positive role—than to support the more demanding idea that asymmetries in peer effects can *justify* that structure on efficiency grounds—their normative role.

5. This makes it very difficult to document peer effects within athletic teams, for instance (Shulman and Bowen 2001).

9.6.2 What We Have Learned So Far

In an earlier study, one of us (Zimmerman 1999) investigated peer effects associated with a student's own GPA and the academic strength (as measured by SAT scores) of his peers. That study attempted to overcome the selection bias issue by assembling a unique set of data comprised of twelve classes of students at Williams College containing information on their grades, major, gender, race, and so on, along with information on where and *with whom* they were housed in their freshman year.[6] In that paper, Zimmerman argued that freshman housing at Williams College closely resembled random assignment. That being the case, it was meaningful to contrast students with high, medium, and low SAT scores who, by chance, had roommates with high, medium, or low SAT scores. This allowed, for example, comparisons between the grades of low-SAT students who roomed with other low-SAT roommates to the grades of low-SAT students who roomed with high-SAT roommates. Any differences in the outcomes could, because of the quasi-random assignment, be attributed to peer effects. The basic findings of that effort suggested that students in the middle of the SAT distribution did somewhat worse in terms of grades if they shared a room with a student who was in the bottom 15 percent of the verbal SAT distribution. Interestingly, students in the top and the bottom of the SAT distribution were not affected by the SAT scores of their peers. The effects for the middle group weren't large but were statistically significant in many models. Furthermore, peer effects were almost always linked more strongly with verbal SAT scores than with math SAT scores.

These results, however, were estimated in the context of a highly selective liberal arts college. In that study, the low-SAT students would, on average, still rank at about the top 15th percentile of the national SAT score distribution. The results could also have been idiosyncratic to Williams College.

Recent research has given additional support to the claim that peer effects exist in higher education (cf. Stinebrickner and Stinebrickner 2000; Goethals 2001; Sacerdote 2001). All of these studies have examined the influence the characteristics or behavior of one student has on the behavior of another. The peer characteristics observed were, for the most part, variants on academic ability—SAT or American College Test (ACT) scores or more nuanced evaluations of academic promise generated in the admission process—while the influenced behavior was largely grades or performance on a written test. These characteristics were broadened to include gender and income, and behaviors were broadened to include dropout behavior, choice of major, and fraternity membership.

Sacerdote (2001), using data from Dartmouth and also using a room-

6. See Zimmerman (forthcoming). This paper contains a broad overview of the academic literature considering peer effects.

mate-based strategy, found evidence of a peer impact of a student on his roommate's GPA as well as on his participation in fraternities. Sacerdote's results suggest a nonlinear relationship with both weaker and stronger students performing better when their roommate was in the top 25 percent of the academic index distribution. In addition, Sacerdote found evidence of peer effects in fraternity participation but no evidence of peer effects in choice of college major.

Stinebrickner and Stinebrickner (2000) employed a data set from Berea College. Like Zimmerman (1999) and Sacerdote (2001), they used the random assignment of roommates to identify the peer effect. Berea College targets low-income students (capping family income at about $65,000) and so provides a useful complement to the highly selective schools used in the other studies. Stinebrickner and Stinebrickner found no evidence at Berea College that either first-semester grades or retention are associated with roommates' ACT scores. They did, however, find evidence that roommate income had a positive impact on both grades and retention, holding ACT scores constant, but only for women.

Goethals (2001) employed a unique and innovative experimental framework to measure peer effects. The study explored whether students would perform better writing about newspaper articles they read and discussed in academically homogenous or heterogeneous groups of three. He found that students' performance was not linked to their own academic rating but was affected by whether they were placed with academically homogeneous or heterogeneous peers. He found that groups composed of students who all had a low academic rating and groups composed of students who all had a high academic rating perform similarly—with both groups of these types out performing groups in which some students had high ratings and some low ratings.[7] These results were stronger for men than for women. So he found that peers' academic characteristics influenced others' behavior but not with straightforward nonlinearities.

In sum, there is a growing—though still small—body of evidence suggesting that peer effects exist in higher education. The evidence is not clear on the nature of any nonlinearities or interactions based on gender. It also suggests that nonacademic peer characteristics may be important.

In this chapter we next add to the empirical evidence by employing data from the College and Beyond (C&B) database—created by the Andrew W. Mellon Foundation—along with matched housing data for three schools in the C&B data. This allows us to apply the same empirical roommate-based approach to measure the peer effects previously described. In so doing, this work adds further evidence on the impact of peer characteristics in higher education.

7. Should these results hold up on further study, they have clear implications for sorting, stratification, and hierarchy among colleges.

9.7 Empirical Strategy: New Evidence

To estimate academic peer effects from the College and Beyond data in terms of equation (1), we follow the now traditional path of relating the cumulative GPA of a student (B_1) to his own SAT scores and to the SAT scores of his first-year roommate (C_2). More formally, we estimate regression models specified as

$$\text{GPA}_i = \alpha + \beta_1\text{SAT}_i + \beta_2\text{SAT}_i^{\text{RM}} + \beta_3\mathbf{X}_i + \varepsilon_{ic}, \tag{2}$$

where GPA is the student's grade point average measured cumulatively to graduation,[8] SAT is the student's own SAT score (sometimes entered separately for math and verbal scores), SAT^{RM} is the student's freshman roommate's SAT score (sometimes entered separately for math and verbal scores), and **X** is a vector of other characteristics (such as race and gender) of the student.[9] If students are randomly assigned their roommates, then the estimated peer effect (β_2) will be unbiased. More generally, the estimate will be unbiased if it is plausible that the error term is uncorrelated with the explanatory variables.

In addition, we estimate models that allow for nonlinearities in the peer effect. In particular, we allow the peer effect to vary based on whether the student or his roommate is in the lowest 15 percent, the middle 70 percent, or the top 15 percent of the SAT distribution. Formally, we estimate

$$\text{GPA}_{ij} = \alpha + \beta_1\text{SAT}_i + \sum_{g=1}^{3}\beta_g\text{SAT}_{ig}^{\text{DRM}} + \beta_3\mathbf{X}_i + \varepsilon_{ic}; \quad j = 1, 2, 3, \tag{3}$$

where $\text{SAT}_{ig}^{\text{DRM}}$ are dummy variables for each SAT score range (indexed by g) and β_g is the peer effect associated with that range.

As previously discussed, if a school's objective is to maximize learning (which we proxy with GPA), then the relative magnitude of β_g for the various ability groupings of students and their peers will be important in efficiently allocating roommates. And, by analogy, they would be suggestive in how students would be sorted across colleges of differing quality. Suppose, for example, that strong students enhance the academic performance of weaker students. Further, suppose stronger students' grades are not affected by having a weak roommate. Then, mixing students may yield higher aggregate learning than would grouping weak students with weak students and strong students with strong students. The weaker students' grades would increase as a result of mixed ability groupings while the

8. Grade performance for the first year alone was not available in C&B data, but analysis of the Williams College data where both cumulative and freshman-year GPA could be used showed that they yielded the same results (Zimmerman 1999).

9. An appealing alternative strategy would be to include the roommate's GPA in the regression. Such a variable might better measure actual rather than potential performance. The problem with including such a variable is that it is simultaneously determined within the roommate context. Using such a measure would introduce simultaneous equation bias.

stronger students grades would not suffer. If, on the other hand, stronger students' grades did suffer we would have to ask whether their decline was sufficient to offset gains to the weaker students. Thus, the coefficients on the β_g parameters for the various groups (along with the relative numbers of students in the different ability categories) are critical in thinking about optimal groupings. The evidence, as we will see, is still mixed on this important issue.

9.8 Data

The C&B data used in this study were created and made available to us by the Andrew Mellon foundation. The C&B data contain both institutional and survey data for over 90,000 students enrolled in thirty-four mostly selective colleges and universities in the United States for the entering classes of 1951, 1976, and 1989. The present study uses data from three of the schools in the C&B population for the entering class of 1989—the graduating class of 1993. Institutional data in College and Beyond provide information on the students' grades, major, race, gender, and so on. These data were combined with housing information extracted from college phonebooks to form a unique data set that allowed us to identify college roommates.

The schools selected for our subsample were chosen because (1) they house their first year students together, and (2) the assignment mechanism of students to rooms (as indicated by their housing descriptions on the World Wide Web and conversations with their housing offices) seems roughly random. It was necessary to use schools that group first-year students together because the C&B data do not provide information on other classes. If, for example, a school allowed first- and second-year students to live together, we would have no information on the second-year students given C&B's restriction to the three cohorts. Selection bias, as previously noted, can be serious when students are allowed to choose their roommates or if the housing office groups students in such a way that under- or overperformers are more likely to be housed together. In this case, the requirement that the error term be uncorrelated with the explanatory variables would be violated. In Zimmerman's earlier study (1999) of Williams College's freshmen, he was able to utilize data from the housing application forms to conduct some relatively simple analyses to check whether the assumption of random assignment was plausible, and it was.[10] The schools in this sample employed a similar protocol to that used by Williams College in using housing forms indicating sleep preferences, smoking behavior, and so on in assigning students to rooms

10. Similarly, estimates in Sacerdote (2001) were unaffected by the inclusion of housing preference variables.

and roommates—though the underlying housing-form data were not obtained.[11]

9.9 Empirical Results

Table 9.2 provides summary statistics for the sample. The number of observations for the samples from each of the three schools ranged from 1,458 to 2,116. Individual SAT scores ranged from a low of 360 on the verbal test and 420 on the math test to a maximum of 800 on both tests. The average combined SAT score ranged from 1,344 to 1,409. These scores are high, putting the average student in the top 10 percent of the population of test takers. Each school had between 7 percent and 9 percent African-American students and 2 percent and 5 percent Hispanic students.

Table 9.3 presents estimates of equation (2). The results for each school are reported in a separate column where a student's cumulative GPA is regressed on his own SAT score (divided by 100), race, gender, major, and roommate's SAT score. The model includes controls for a student's major (which is selected in his or her junior year) to provide some control for grade differentials arising from students' taking different courses (Sabot 1991).

The effect of a student's own SAT score is large and statistically significant, with each 100 point increase resulting in between a 0.116 and a 0.132 increase in GPA. After controlling for SAT scores, black and Hispanic students score between one-quarter and one-third of a grade point below white students. Female students score between 0.082 and 0.127 grade points higher than male students. Finally, a roommate's SAT score is found to have a positive and statistically significant effect only for school 2—where a 100 point increase in a student's roommate's combined SAT score translates into a 0.02 increase in the student's own GPA. This effect is about 17 percent as large as that of a 100 point increment in the student's own SAT score.[12]

Tables 9.4 through 9.6 report estimates of equation (3), allowing the peer effect to depend on the student's own position in the SAT distribution. The top sections of tables 9.4 through 9.6 allow us to see whether weak, average, or strong students (as measured by their SAT scores) are more or less

11. See Zimmerman (forthcoming) for a mathematical model that illustrates the possibility of bias in the estimated peer effects flowing from the use of housing forms in assigning students to rooms. Chi-squared tests indicate that we cannot reject independence between the SAT scores of roommates for schools 1 and 3 in the sample. For school 2, independence is rejected. The rejection is driven by a somewhat high fraction of low-SAT students living together and a somewhat low fraction of low-SAT students living with high-SAT students. The distribution of low, medium, and high students is as expected under independence for the middle-SAT students. In total, there are about 100 of the 2,116 students that show signs of selection.

12. It is worth noting here that models allowing for differential effects for math and verbal SAT scores were also estimated, but standard F-tests indicated no measurable difference in their impact. Accordingly, only models using combined SAT scores are reported.

Table 9.2 Descriptive Statistics

	Mean	Standard Deviation	Minimum	Maximum
School 1				
Sample size	1,863	0	1,863	1,863
Own SAT score—verbal	714	66	420	800
Own SAT score—math	695	69	480	800
Own SAT score—combined	1409	112	1090	1600
Black	.079	.270	0	1
Hispanic	.052	.223	0	1
Native American	.004	.069	0	1
Asian	.151	.358	0	1
Not a citizen of the United States	.03	.169	0	1
Female	.432	.495	0	1
School 2	.430	.494	0	1
Sample size	2,116	0	2,116	2,116
Own SAT score—verbal	668	68	360	800
Own SAT score—math	676	68	450	800
Own SAT score—combined	1344	110	950	1600
Black	.086	.282	0	1
Hispanic	.044	.206	0	1
Native American	n.a.	n.a.	n.a.	n.a.
Asian	.160	.367	0	1
Not a citizen of the United States	.095	.292	0	1
Female	.430	.494	0	1
School 3				
Sample size	1,458	0	1,458	1,458
Own SAT score—verbal	687	61	450	800
Own SAT score—math	681	68	420	800
Own SAT score—combined	1368	106	880	1600
Black	.072	.258	0	1
Hispanic	.022	.148	0	1
Native American	.001	.036	0	1
Asian	.079	.270	0	1
Not a citizen of the United States	.03	.148	0	1
Female	.466	.499	0	1

Note: n.a. = not available.

affected by roommates. The results in these panels suggest that strong students at all three schools are unaffected by the SAT scores of their roommates. Students in the bottom 15 percent of the SAT distribution benefit from higher SAT scoring roommates at school 1—though not at schools 2 and 3. Students in the middle 70 percent of the distribution are unaffected by the SAT scores of their roommates at schools 1 and 3—though they benefit from higher-scoring roommates at school 2. Students in the middle 70 percent of the SAT distribution at school 2 experience, on average, a 0.02 increase in their cumulative GPA when their roommates' SAT scores increase by 100 points.

The bottom sections of tables 9.4 through 9.6 allow the peer effect to be

Table 9.3 Your Grades and Your Roommate's SAT Scores

	Cumulative GPA (School 1)	Cumulative GPA (School 2)	Cumulative GPA (School 3)
Own SAT score/100	0.131*	.116*	.132*
	(0.01)*	(.013)*	(.012)*
Black	–.264	–.306	–.380
	(.068)	(.060)	(.054)
Hispanic	–.172	–.080	.005
	(.085)	(.055)	(.046)
Native American	–.268	n.a.	.145
	(.157)		(.071)
Not a citizen of the United States	n.a.	–.047	n.a.
		(.065)	
Asian	–.011	–.071	–.033
	(.031)	(.031)	(.042)
Female	.127	.082	.112
	(.028)	(.024)	(.024)
Major dummy variables	Yes	Yes	Yes
Roommate's SAT score/100	0.013	0.020*	.013
	(0.007)	(0.008)*	(.009)
Sample size	1,863	2,116	1,458
R^2	.303	0.215	0.2475

Note: Standard errors (in parentheses) are corrected for correlation within roommate cluster. n.a. = not available.

*Significant at the 10 percent level.

nonlinear. That is, it allows us to see whether weak, average, or strong students (as measured by their SAT scores) are more or less affected by having roommates who are weak, average, or strong in terms of their combined SAT scores. For this model, at school 1 we find low-SAT students performing somewhat worse when roomed with a similarly weak peer. The coefficient shows grades for this group would increase by 0.156 points if they had a high-SAT roommate. The coefficient is significant at the 10 percent level. At school 2, neither the strongest nor the weakest students are affected by the SAT scores of their roommates. Students in the middle 70 percent of the SAT distribution, however, perform somewhat worse when their roommates are in the bottom 15 percent of the SAT distribution. The estimates suggest that a student with a bottom 15 percent roommate in this part of the SAT distribution would, on average, have a cumulative GPA that is lower by 0.086 points than that of a similar student whose roommate was in the top 15 percent of the SAT distribution. Similar results are found at school 3 where, in addition, there is evidence that the strongest students perform better when their roommates are academically stronger. It is worth noting that these results are robust to moderate variations in the percentile cutoffs used to define the groups.

Tables 9.7 through 9.9 report estimates of equation (3) separately for

Table 9.4 **Your Grades and Your Roommate's SAT Scores by SAT Group: School 1 (dependent variable is cumulative GPA)**

	Combined SAT Score (lowest 15%)	Combined SAT Score (middle 70%)	Combined SAT Score (top 15%)
	A. Linearity in Roommate's Scores		
Own SAT score—verbal/100	.065	.223*	.036
	(.087)	(.029)*	(.124)
Own SAT score—math/100	.024	.172*	.124
	(.127)	(.033)	(.148)
Black	–.174	–.297	–.758
	(.186)	(.079)	(.165)
Hispanic	.0402	–.311	–.024
	(.086)	(.142)	(.116)
Native American	–.045	–.356	(dropped)
	(.160)	(.251)	
Not a citizen of the United States	n.a.	n.a.	n.a.
Asian	.226	–.004	–.040
	(.230)	(.039)	(.052)
Female	.233	.138	.012
	(.110)	(.032)	(.056)
Major dummy variables	Yes	Yes	Yes
Roommate's SAT score/100	.032*	.011	–.009
	(.010)*	(.008)	(.014)
Sample size	269	1,281	313
R^2	.0288	0.295	0.154
	B. Nonlinearity in Roommate's Scores		
Own SAT score—verbal/100	.060	.223*	.021
	(.089)	(.02856)*	(.125)
Own SAT score—math/100	.021	.172*	.100
	(.128)	(.033)*	(.151)
Black	–.175	–.297	–.805
	(.183)	(.079)	(.163)
Hispanic	.043	–.312	–.022
	(.086)	(.141)	(.114)
Native American	–.075	–.352	(dropped)
	(.169)	(.251)	
Not a citizen of the United States	n.a.	n.a.	n.a.
Asian	.233	–.004	–.039
	(.231)	(.039)	(.051)
Female	.220	.137	.022
	(.110)	(.032)	(.051)
Major dummy variables	Yes	Yes	Yes
Roommate's SAT score—lowest 15%	–.156*	–.044	–.002
	(.086)	(.032)	(.050)
Roommate's SAT score—middle 70%	–.131	–.023	–.038
	(.086)	(.025)	(.043)
Sample size	269	1,281	313
R^2	0.295	0.295	0.154

Note: Standard errors (in parentheses) are corrected for correlation within roommate cluster. n.a. = not available.

*Significant at the 10 percent level.

Table 9.5 **Your Grades and Your Roommate's SAT Scores by SAT Group: School 2 (dependent variable is cumulative GPA)**

	Combined SAT Score (lowest 15%)	Combined SAT Score (middle 70%)	Combined SAT Score (top 15%)
	A. Linearity in Roommate's Scores		
Own SAT score—verbal/100	.162	.142*	–.109
	(.088)	(.025)*	(.098)
Own SAT score—math/100	.077	.166*	.063
	(.101)	(.027)*	(.112)
Black	–.235	–.341	–.117
	(.079)	(.085)	(.160)
Hispanic	–.036	–.060	–.071
	(.127)	(.070)	(.095)
Native American	n.a.	n.a.	n.a.
Not a citizen of the United States	–.204	–.016	.026
	(.243)	(.079)	(.065)
Asian	.102	–.083	–.111
	(.145)	(.033)	(.081)
Female	.067	.099	–.109
	(.077)	(.026)	(.129)
Major dummy variables	Yes	Yes	Yes
Roommate's SAT score/100	.017	.020*	.0438
	(.021)	(.009)*	(.026)
Sample size	280	1,500	336
R^2	0.286	0.181	0.178
	B. Nonlinearity in Roommate's Scores		
Own SAT score—verbal/100	.167	.143*	–.110
	(.088)	(.025)*	(.098)
Own SAT score—math/100	.088	.166*	.059
	(.100)	(.027)*	(.111)
Black	–.238	–.340	–.086
	(.079)	(.085)	(.168)
Hispanic	–.035	–.050	–0.05
	(.127)	(.069)	(.102)
Native American	n.a.	n.a.	n.a.
Not a citizen of the United States	–.174	–.009	–.109
	(.242)	(.078)	(.128)
Asian	.108	–.082	–.110
	(.142)	(.033)	(.081)
Female	.061	.102	.015
	(.077)	(.026)	(.064)
Major dummy variables	Yes	Yes	Yes
Roommate's SAT score—lowest 15%	–.042	–.086*	–.099
	(.088)	(.034)*	(.102)
Roommate's SAT score—middle 70%	–.066	–.022	–.079
	(.072)	(.023)	(.057)
Sample size	282	1,505	337
R^2	0.286	0.181	0.172

Note: Standard errors (in parentheses) are corrected for correlation within roommate cluster.
*Significant at the 10 percent level.

Table 9.6 Your Grades and Your Roommate's SAT Scores by SAT Group: School 3 (dependent variable is cumulative GPA)

	Combined SAT Score (lowest 15%)	Combined SAT Score (middle 70%)	Combined SAT Score (top 15%)
A. Linearity in Roommate's Scores			
Own SAT score—verbal/100	.214*	.114*	.183
	(.061)*	(.032)*	(.085)
Own SAT score—math/100	.146*	.101*	.236
	(.065)*	(.031)*	(.106)
Black	–.309	–.498	–.186
	(.082)	(.112)	(.076)
Hispanic	.028	–.021	.191
	(.086)	(.064)	(.131)
Native American	(dropped)	.120	(dropped)
		(.087)	
Not a citizen of the United States	n.a.	n.a.	n.a.
Asian	.310	–.097	.045
	(.164)	(.049)	(.090)
Female	.108	.088	.122
	(.078)	(.030)	(.068)
Major dummy variables	Yes	Yes	Yes
Roommate's SAT score/100	–.016	.019	.036
	(.025)	(.011)	(.026)
Sample size	221	975	262
R^2	0.3560	(0.1151)	0.1215
B. Nonlinearity in Roommate's Scores			
Own SAT score—verbal/100	.207*	.114*	.186*
	(.056)*	(.032)*	(.083)*
Own SAT score—math/100	.148*	.100*	.238*
	(.065)*	(.031)*	(.102)*
Black	–.303	–.498	–.145
	(.078)	(.111)	(.079)
Hispanic	.031	–.014	.193
	(.082)	(.059)	(.116)
Native American	(dropped)	.110	(dropped)
		(.085)	
Not a citizen of the United States	n.a.	n.a.	n.a.
Asian	.314	–.094	.058
	(.165)	(.049)	(.090)
Female	.110	.090	.139
	(.078)	(.030)	(.066)
Major dummy variables	Yes	Yes	Yes
Roommate's SAT score—lowest 15%	.069	–.092*	–.175*
	(.096)	(.041)*	(.077)*
Roommate's SAT score—middle 70%	.004	–.038	–.127*
	(.081)	(.031)	(.061)*
Sample size	223	981	263
R^2	0.3585	0.1173	0.1377

Note: Standard errors (in parentheses) are corrected for correlation within roommate cluster.
*Significant at the 10 percent level.

Table 9.7 Your Grades and Your Roommate's SAT Scores by SAT Group and Gender: School 1 (dependent variable is cumulative GPA)

	Combined SAT Score (lowest 15%)	Combined SAT Score (middle 70%)	Combined SAT Score (top 15%)
	A. Men		
Own SAT score—verbal/100	.048	.266*	–.006
	(.108)	(.034)*	(.172)
Own SAT score—math/100	.113	.163*	–.002
	(.122)	(.043)*	(.002)
Black	.041	–.438	–.817
	(.124)	(.132)	(.206)
Hispanic	.067	–.128	.006
	(.096)	(.134)	(.091)
Native American	(dropped)	–.717	(dropped)
		(.254)	
Not a citizen of the United States	n.a.	n.a.	n.a.
Asian	.926	.039	–.075
	(.220)	(.056)	(.112)
Major dummy variables	Yes	Yes	Yes
Roommate's SAT score—lowest 15%	–.167	–.054	.078
	(.117)	(.046)	(.060)
Roommate's SAT score—middle 70%	–.108	–.042	–.022
	(.088)	(.035)	(.033)
Sample size	137	739	187
R^2	0.637	0.323	0.309
	B. Women		
Own SAT score—verbal/100	.117	.187*	–.101
	(.166)	(.057)*	(.182)
Own SAT score—math/100	–.062	.192*	.095
	(.200)	(.046)*	(.227)
Black	–.436	–.228	(dropped)
	(.347)	(.085)	
Hispanic	–.057	–.474	(dropped)
	(.161)	(.251)	
Native American	–.242	–.064	(dropped)
	(.185)	(.130)	
Not a citizen of the United States	n.a.	n.a.	n.a.
Asian	.105	–.073	–.040
	(.149)	(.052)	(.086)
Major dummy variables	Yes	Yes	Yes
Roommate's SAT score—lowest 15%	–.104	–.026	–.020
	(.124)	(.040)	(.084)
Roommate's SAT score—middle 70%	–.143	–.006	.028
	(.124)	(.034)	(.101)
Sample size	132	543	128
R^2	0.279	0.325	0.441

Note: Standard errors (in parentheses) are corrected for correlation within roommate cluster.
*Significant at the 10 percent level.

Table 9.8 Your Grades and Your Roommate's SAT Scores by SAT Group and Gender: School 2 (dependent variable is cumulative GPA)

	Combined SAT Score (lowest 15%)	Combined SAT Score (middle 70%)	Combined SAT Score (top 15%)
	A. Men		
Own SAT score—verbal/100	.230	.194*	–.164
	(.166)	(.034)*	(.114)
Own SAT score—math/100	.105	.212*	.038
	(.165)	(.038)*	(.127)
Black	–.239	–.281	(dropped)
	(.187)	(.131)	
Hispanic	–.134	.055	–.087
	(.233)	(.077)	(.112)
Native American	n.a.	n.a.	n.a.
Not a citizen of the United States	–.068	.027	–.163
	(.377)	(.093)	(.141)
Asian	.188	–.053	–.166
	(.270)	(.048)	(.112)
Major dummy variables	Yes	Yes	Yes
Roommate's SAT score—lowest 15%	–.132	–.132*	–.092
	(.194)	(.056)*	(.121)
Roommate's SAT score—middle 70%	–.093	–.036	–.082
	(.109)	(.029)	(.068)
Sample size	110	839	245
R^2	0.258	0.209	0.238
	B. Women		
Own SAT score—verbal/100	.126	.074	.093
	(.094)	(.041)	(.179)
Own SAT score—math/100	.165	.118*	.119
	(.123)	(.040)*	(.269)
Black	–.226	–.375	–.477
	(.083)	(.113)	(.166)
Hispanic	.046	–.273	(dropped)
	(.124)	(.116)	
Native American	n.a.	n.a.	n.a.
Not a citizen of the United States	–.358	–.087	(dropped)
	(.403)	(.070)	
Asian	.030	–.102	–.065
	(.133)	(.048)	(.145)
Major dummy variables	Yes	Yes	Yes
Roommate's SAT score—lowest 15%	.102	–.014	.139
	(.112)	(.043)	(.129)
Roommate's SAT score—middle 70%	.072	.022	–.018
	(.095)	(.036)	(.080)
Sample size	172	666	92
R^2	0.439	0.204	0.209

Note: Standard errors (in parentheses) are corrected for correlation within roommate cluster.
*Significant at the 10 percent level.

Table 9.9 **Your Grades and Your Roommate's SAT Scores by SAT Group and Gender: School 3 (dependent variable is cumulative GPA)**

	Combined SAT Score (lowest 15%)	Combined SAT Score (middle 70%)	Combined SAT Score (top 15%)
	A. Men		
Own SAT score—verbal/100	.079	.136*	.154
	(.073)	(.048)*	(.099)
Own SAT score—math/100	.255*	.174*	.176
	(.105)*	(.049)*	(.134)
Black	–.261	–.632	–.077
	(.151)	(.159)	(.077)
Hispanic	.006	–.170	.112
	(.124)	(.087)	(.170)
Native American	(dropped)	.043	(dropped)
		(.088)	
Not a citizen of the United States	n.a.	n.a.	n.a.
Asian	.236	–.158	–.008
	(.219)	(.071)	(.105)
Major dummy variables	Yes	Yes	Yes
Roommate's SAT score—lowest 15%	.161	–.085	–.107
	(.120)	(.069)	(.093)
Roommate's SAT score—middle 70%	.105	–.063	–.107
	(.112)	(.045)	(.063)
Sample size	104	464	204
R^2	0.4625	0.1634	0.1396
	B. Women		
Own SAT score—verbal/100	.292*	.110*	.460*
	(.081)*	(.044)*	(.127)*
Own SAT score—math/100	.200*	.031	.350*
	(.098)*	(.039)	(.123)*
Black	–.192	–.377	–.335
	(.107)	(.135)	(.055)
Hispanic	.0190	.070	.429
	(.145)	(.073)	(.233)
Native American	(dropped)	(dropped)	(dropped)
Not a citizen of the United States	n.a.	n.a.	n.a.
Asian	.128	–.050	.212
	(.150)	(.072)	(.084)
Major dummy variables	Yes	Yes	Yes
Roommate's SAT score—lowest 15%	.018	–.059	–.266*
	(.179)	(.048)	(.133)*
Roommate's SAT score—middle 70%	–.124	.003	–.149*
	(.114)	(.039)	(.076)*
Sample size	119	517	59
R^2	0.4546	0.1172	0.6660

Note: Standard errors (in parentheses) are corrected for correlation within roommate cluster.
*Significant at the 10 percent level.

Table 9.10 **Peer Coefficients from Stacked Data for All Three Schools (dependent variable is cumulative GPA)**

	Combined SAT Score (lowest 15%)	Combined SAT Score (middle 70%)	Combined SAT Score (top 15%)
Men and women			
Roommate's SAT score—lowest 15%	–.070	–.067*	–.063
	(.057)	(.021)*	(.044)
Roommate's SAT score—middle 70%	–.073	–.029*	–.066*
	(.050)	(.015)*	(.030)*
Men			
Roommate's SAT score—lowest 15%	–.004	–.077*	–.029
	(.088)	(.031)*	(.062)
Roommate's SAT score—middle 70%	–.042	–.044*	–.067*
	(.068)	(.020)*	(.035)*
Women			
Roommate's SAT score—lowest 15%	–.091	–.036	–.062
	(.075)	(.027)	(.057)
Roommate's SAT score—middle 70%	–.101	.006	–.021
	(.078)	(.022)	(.055)

Notes: Other controls are own SAT scores, gender, ethnicity, major, and school fixed effect. Standard errors are in parentheses.

*Significant at the 10 percent level.

men and women. Perhaps due to smaller sample sizes, peer effects are not statistically significant for most groups. Exceptions are found at schools 2 and 3. At school 2, male students in the middle of the SAT distribution are found to perform worse when their roommate is in the lowest 15 percent of the SAT distribution; at school 3, academically strong women perform better when given academically strong peers.

Table 9.10 presents estimates using data stacked for the three schools. School fixed effects are included in these models. The main advantage of stacking the data is that there are larger cell sizes—giving us more precise estimates—with which to gauge any nonlinearities. These results are presented pooled by gender and also separately for male and female students. The results mirror the foregoing ones with students in the middle showing lower grades if their roommate is in the bottom 15 percent of the SAT distribution. The estimates suggest that this result is driven by the male sample as the coefficients for women are not significant for any of the SAT groups. There is also some evidence—again particularly for men—that strong students perform somewhat worse if their roommate is in the middle of the SAT distribution rather than in the top.

To put the myriad results in context it is useful to summarize the existing research more succinctly. The research to date, including the evidence re-

Table 9.11 Recent Students of Academic Peer Effects

Study	Peer Characteristic	Coefficient on Grades	Comments
Zimmerman (1999)	Roommate's verbal SAT in bottom 15%	–.077 (.027)	Impact on middle 70% of SAT distribution, Williams College.
Zimmerman (as reported in this chapter)	Roommate's verbal SAT in bottom 15%	–.086 (.034)	Impact on middle 70% of SAT distribution, three schools from College and Beyond.
Sacerdote (2001)	Roommate in top 25% of Academic Rating Index	.060 (.028)	Dartmouth. Controls for housing questions. Also peer effect on fraternity membership but none on major.
Steinbrickner and Stinebrickner (2001)	ACT score	.001 (.004)	Controls for roommate's family income. Roommate income is significant with grades, rising .052 per $10,000 income, for women.
Goethals (2000)	Admissions office academic rating	n.a.	Finds performance increases with group homogeneity in academic rating.

Notes: Coefficient on grades data taken from table 4 in Zimmerman (1999), table 3 in Sacerdote (2001), table 3 in Stinebrickner and Stinebrickner (2000), and tables 9.5 and 9.10 in this chapter. Standard errors are in parentheses. n.a. = not available.

ported in this paper, on the effect of peer academic characteristics on a "grade type" outcome is summarized in table 9.11.

These studies differ in a variety of ways: the selectivity of the school surveyed, the measurement and detection of nonlinearities, the outcome considered, the existence of differences by gender, and so on. The evidence found thus far suggests that the existence of peer effects at the most basic level has been confirmed in each of the studies. Sacerdote (2001) finds that grades are higher when students have unusually academically strong roommates. Zimmerman (1999, 2003) finds that weak peers might reduce the grades of middling or strong students. Stinebrickner finds that peer ACT scores are insignificant after controlling for roommate's family income, which is significant. Goethals (2001) finds that homogeneity per se matters—students perform better when grouped with others of like ability.

Additional studies using data closer to the ideal described previously would certainly be useful. Most of the results, thus far, pertain to highly selective institutions. Data from multiple, more diverse, and more representative schools would provide greater variation in the variety of differing academic peer environments we observe and, thus, a chance to better evaluate the impact of peers across the spectrum of student abilities. In short, such data would enable us to better estimate the functional form relating outcomes, peer characteristics and behaviors, and their interactions.

9.10 Conclusion and Implications

Evidence on peer effects in higher education now exists at the most basic level for six colleges and universities—covering some 12,000 students (across published studies)—with interactions measured for randomly assigned roommates and participants in psych lab experiments. It seems clear that peer effects exist—students' characteristics and behavior do, indeed, influence other students' behavior with conventionally measured academic characteristics (like SAT) influencing conventionally measured academic performance (like GPA). New evidence presented in this chapter adds to our confidence that peer effects exist and that the signs of those effects are in the direction that would motivate institutional selectivity—strong students tend to increase peers' academic performance, and weak students tend to reduce it. Combined with a sharply skewed distribution of resources across colleges, the broad question "Can peer effects in educational production help explain the unusual economic structure and behavior of higher education?" is answered "Yes." The models of Winston (1999) and Epple, Romano, and Sieg (2001)—data-driven and formally derived, respectively—fit both the data and the peer effect evidence.[13]

But beyond that key question, the facts become less clear, and the agenda for investigation of peer effects becomes larger. So there are often different results by gender, as in Hoxby's (2000) K–12 results, even in these data that rest on individual interactions rather than on those between groups. On nonlinearities—whether peer influences operate equally and symmetrically across characteristics and behaviors—the evidence is puzzling, with strong or weak homogeneous groupings sometimes performing significantly better than those with peers of different abilities. Students of middling ability are usually more susceptible to peer influence than those at either end of the ability distribution (keeping in mind that the student populations reported on here represent very narrow ability ranges). And because our data are based on pairwise interactions, a similar analysis might well be extended to those interactions that are electronically mediated to see if a "distance learning" environment generates any evidence of peer effects.

13. Note that there is no evidence of a "teaching effect" in which strong students gain from association with weaker students whom they can teach (as implied by Zajonc's [1976] analysis of older siblings) nor is there strong evidence of an "intimidation effect," though that might help explain Goethals' (2001) finding that weak students do better when grouped with other weak students.

The range of peer characteristics and behaviors should be extended, too, wherever possible. The work reported here sticks, by and large, to the most measurable and obvious aspects of education—grade performance and academic ability—with occasional departure into fraternity membership, family income, and dropout behavior. But while these are clearly the right place to start, they capture a small part of the behaviors influenced by higher education and of interest to colleges in their selection of student peer quality—it may be possible to get closer to our "ideal data" with other measurable academic behaviors among randomly associated peers. Like Heckman (1999), Bowles, Gintis, and Osborn (2001) point out that a small part of the variance in wages attributable to education is explained by the cognitive skills we measure with tests and GPAs—the rest is attributable to behaviors learned before, after, in, and outside of school that may escape cognitive measurement but influence job performance, nonetheless, like reliability, attitude, discipline, fatalism, impatience, and so on. To the extent that these characteristics and behaviors can be identified and measured, they need to be included in studies of peer effects in higher education.[14]

So we conclude that evidence on the existence of peer effects in higher education is strong, consistent with an understanding of its economic structure—selectivity, skewed resources, and the resulting stratification—that relies on them, but that there remains a rich set of questions on how and how broadly peer effects operate among students in colleges[15] and especially on the shape of the nonlinearities that would help us evaluate that structure.

References

Akerlof, G. A., and J. L. Yellen. 1986. Introduction. *Efficiency wage models of the labor market.* Cambridge: Cambridge University Press.

Avery, C., A. Fairbanks, and R. Zeckhauser. 2001. What worms for the early bird? Early admissions at selective colleges. Faculty Research Working Paper Series no. RWP 01-049. Harvard University, Kennedy School of Government.

Bowles, S., H. Gintis, and M. Osborn. 2001. The determinants of earnings: A behavioral approach. *Journal of Economic Literature* 39 (4): 1137–176.

14. On the basis of evidence that a student's impatience (his time-discount behavior) influences his own academic performance (students with lower discount rates get better grades, holding SATs constant (Kirby, Winston et al. 2002), we tried, in a very small sample, to find evidence of peer influence such that one roommate's discount rate affected the other's academic performance. But while the sign of the relationship was right, it was decidedly insignificant.

15. Our discussion has not even touched on negative peer effects like binge drinking and date rape.

Clotfelter, C. T. 1996. *Buying the best: Cost escalation in elite higher education.* Princeton, N.J.: Princeton University Press.

Coleman, J. S., E. Q. Campbell, C. F. Hobson, J. McPartland, A. M. Mood, F. D. Weinfeld, and R. L. York. 1966. *Equality of educational opportunity.* Washington, D.C.: Office of Education, U.S. Department of Health, Education, and Welfare.

Epple, D., and R. E. Romano. 1998. Competition between private and public schools, vouchers, and peer-group effects. *American Economic Review* 88 (1): 33–62.

Epple, D., R. E. Romano, and H. Sieg. 2001. Peer effects, financial aid, and selection of students into colleges and universities: An empirical analysis. Working Paper no. 00-02. Duke University, Department of Economics.

Fallows, J. 2001. The early-decision racket. *The Atlantic Monthly* 288 (2): 37–52.

Goethals, G. R. 2001. Peer effects, gender and intellectual performance among students at a highly selective college: A social comparison of abilities analysis. Williamstown, Mass.: The Williams Project on the Economics of Higher Education.

Hansmann, H. 1980. The role of nonprofit enterprise. *The Yale Law Journal* 89 (5): 835–901.

Hanushek, E. 1986. The economics of schooling: Production and efficiency in public schools. *Journal of Economic Literature* 24 (3): 1141–177.

Heckman, J. J. 1999. Doing it right: Job training and education. *Public Interest* (Spring): 86–109.

Hoxby, C. M. 2000. Peer effects in the classroom: Learning from gender and race variation. NBER Working Paper no. 7867. Cambridge, Mass.: National Bureau of Economic Research.

James, E. 1978. Product mix and cost disaggregation: A reinterpretation of the economics of higher education. *Journal of Human Resources* 12 (2): 157–186.

———. 1990. Decision processes and priorities in higher education. *The Economics of American Universities,* ed. S. A. Hoenack and E. L. Collins, 77–106. Buffalo: State University of New York Press.

Kennedy, D. 1997. *Academic duty.* Cambridge, Mass.: Harvard University Press.

Kirby, K. N., G. C. Winston, and M. Santiesteban. 2002. Impatience and grades: Delay-discount rates correlate negatively with college GPA. Williamstown, Mass.: The Williams Project on the Economics of Higher Education.

Lazear, E. P. 1999. Educational production. NBER Working Paper no. 7349. Cambridge, Mass.: National Bureau of Economic Research.

Lemann, N. 1999a. *The big test: The secret history of the American meritocracy.* New York: Ferrar, Straus & Giroux.

———. 1999b. The kids in the conference room: How McKinsey & Company became the next big step. *The New Yorker,* October 18 and 25, 209–216.

Manski, C. F., and D. A. Wise. 1983. *College choice in America.* Cambridge, Mass.: Harvard University Press.

McPherson, M. S., and M. O. Schapiro. 1990. Selective admission and the public interest. New York: The College Board Press.

Robertson, D., and J. Symons. 1996. Do peer groups matter? Peer group versus schooling effects on academic attainment. London School of Economics Centre for Economic Performance Discussion Paper.

Rosovsky, H. 1990. *The university: An owner's manual.* New York: W. W. Norton.

Rothschild, M., and L. J. White. 1995. The analytics of pricing in higher education and other services in which customers are inputs. *Journal of Political Economy* 103 (3): 573–586.

Sabot, R., and J. Wakeman-Linn. 1991. The implications of grading policies for student course choice. *Journal of Economic Perspectives* 5 (1): 159–170.
Sacerdote, B. 2001. Peer effects with random assignment: Results for Dartmouth roommates. *Quarterly Journal of Economics* 116 (2): 681–704.
Shulman, J. L., and William G. Bowen. 2001. *The game of life: College sports and educational values.* Princeton, N.J.: Princeton University Press.
Stinebrickner, T. R., and R. Stinebrickner. 2000. Peer effects among students from disadvantaged backgrounds. Working Paper Series no. 2001-3. University of Western Ontario, Department of Economics.
Summers, A. A., and B. L. Wolfe. 1977. Do schools make a difference? *American Economic Review* 67 (4): 639–652.
Swann v. Charlotte-Mecklenburg Board of Education, 402 U.S. 1, U.S. Supreme Court (1971).
United States v. Brown University, F.3d 658:1993 U.S. App. LEXIS 23895, United States Court of Appeals for the Third District (1993).
Winston, G. C. 1999. Subsidies, hierarchy, and peers: The awkward economics of higher education. *Journal of Economic Perspectives* 13 (1): 13–36.
———. 2002. College prices, peer wages, and competition. Williamstown, Mass.: The Williams Project on the Economics of Higher Education.
———. 2003. Differentiation among U.S. colleges and universities. *Review of Industrial Organization.*
Winston, G. C., and Ivan C. Yen. 1995. Costs, prices, subsidies, and aid in U.S. higher education. Williamstown, Mass.: The Williams Project on the Economics of Higher Education.
Zajonc, R. B. 1976. Family configuration and intelligence. *Science* 192:227–236.
Zimmerman, D. J. 1999. Peer effects in academic outcomes: Evidence from a natural experiment. Williamstown, Mass.: The Williams Project on the Economics of Higher Education.
———. 2003. Peer effects in academic outcomes: Evidence from a natural experiment. *Review of Economics and Statistics* 85 (1): 9–23.

Comment Thomas S. Dee

Gordon Winston and David Zimmerman motivate their chapter with a provocative outline of how several "awkward" features of higher education are difficult to understand from a conventional economic perspective on firm behavior (e.g., the screening and rebuking of customers, subsidizing accepted customers, requiring that customers live together). They go on to discuss how the existence of peer effects and other traits (i.e., nonprofit objectives and the role of donative resources) provide a coherent framework for understanding these institutional behaviors. In particular, the possibility that the quality of peers influences student achievement could help to explain both the sharp concentration of student ability in relatively few

Thomas S. Dee is assistant professor of economics at Swarthmore College, and a faculty research fellow at the National Bureau of Economic Research.

schools as well as the receipt of larger, implicit subsidies among those students. However, they acknowledge that establishing the existence and possibly heterogeneous pattern of peer effects is a notoriously difficult empirical problem. First and foremost, identifying the mere existence of peer effects with conventional data sets is complicated by the fact that students often choose their peers in response to their own unobserved determinants of achievement. Second, the policy implications of peer effects depend in large measure on the pattern of response heterogeneity (e.g., how peer effects vary for students of different backgrounds and for sharp changes in peer quality). The central focus of this chapter is on extending and discussing the recent developments with regard to both of these empirical issues.

Specifically, Winston and Zimmerman use student-level data from three schools in the College and Beyond (C&B) data to evaluate the effect of peer quality (as measured by the SAT scores of freshman roommates) on cumulative GPA. The central contribution of this approach, which has also been adopted in other recent studies (Sacerdote 2001; Zimmerman 2003), is both novel and important. By relying on the putatively random assignment of freshman roommates, this research design may eliminate the confounding biases in estimated peer effects. Like the related studies, their results suggest the existence of peer effects, at least for students who did not have low SAT scores themselves. However, there are at least two reasons to interpret these new estimates with some degree of caution. First, the case for roommate assignments at these schools being completely random is not as clear as in the earlier studies. Second, the sizes of the estimated effects seem suspiciously large given that the dependent variable is cumulative, not freshman, GPA. Sacerdote (2001) found that freshman roommates had similarly sized effects only on freshman, not senior, GPA. Some lack of persistence in these effects seems reasonable since student interaction with freshman roommates is often limited and relatively short term.

Nonetheless, in conjunction with the studies by Sacerdote (2001) and Zimmerman (2003), the new empirical results presented here constitute a compelling case for the existence of some peer group effects in higher education. But, while that is a meaningful statement, it is not clear that the available evidence supports any further conclusions. For example, Winston and Zimmerman note that the weak statistical power implied by the available data make it difficult to ascertain how peer effects might vary across students and peers of different backgrounds. Furthermore, even if these data did indicate certain response heterogeneities, their external validity would be unclear at best. The segmentation of students across colleges and universities implies that the students at these three elite schools have a relatively narrow range of ability. This clearly implies that these estimates may not speak to the policy-relevant question of how sharp changes in peer quality would influence student achievement.

There is also little evidence that the peer effects associated with freshman roommates extend to other peer environments (most notably, classrooms). While we might be tempted to assume that the existence of roommate effects implies the existence of peer effects in other settings, there are two reasons to proceed more cautiously. First, Sacerdote (2001) found that peers' effects on GPA occurred only at the roommate level, not at the dormitory or floor level. Second, the broader existence of peer effects, combined with the sharp segmentation of students across colleges by ability, would clearly imply that there is a substantial return to attending elite schools. However, the evidence on the returns to college selectivity is decidedly mixed (e.g., Hoxby 1998; Dale and Krueger 2002).

Given these limitations, Winston and Zimmerman are correct to stress that many important questions about peer effects remain unanswered. In particular, despite important recent additions to the available research, we do not yet have evidence that could be used to justify existing educational policies or advocate policy changes. Additionally, it is also not entirely clear that peer effects explain the segmented structure of higher education. Even if peer effects exist, parents, students, and college administrators may still respond more to perceptions of prestige. Instead, a major contribution of recent evidence like that in this chapter is to motivate and direct further research in this area. For example, one possibly fruitful direction would be to test for peer effects in other settings (e.g., college classrooms). This could be done in a convincing manner by exploiting institutional mechanisms that may generate plausibly exogenous assignments. Other useful evidence about the scope of peer effects should also come from applying similar research designs to less elite schools that are more integrated by prior student ability.

References

Dale, S. B., and A. B. Krueger. 2002. Estimating the payoff to attending a more selective college: An application of selection on observables and unobservables. *Quarterly Journal of Economics* 117 (4): 1491–527.

Hoxby, C. M. 1998. The return to attending a more selective college: 1960 to the present. Harvard University. Unpublished Manuscript. Available at [http://post.economics.harvard.edu/faculty/hoxby/papers/whole.pdf].

Sacerdote, B. 2001. Peer effects with random assignment: Results for Dartmouth roommates. *Quarterly Journal of Economics* 116 (2): 681–704.

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