European Economic Review 98 (2017) 144–168

Contents lists available at ScienceDirect

European Economic Review

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

Bad Company: Understanding negative peer effects in college achievement

Ryan R. Brady, Michael A. Insler*, Ahmed S. Rahman

U.S. Naval Academy, 589 McNair Rd., Mail Stop 10D, Annapolis, MD 21402, United States

ARTICLE INFO

Article history:
Received 27 June 2016
Accepted 19 June 2017
Available online 4 July 2017

JEL classification:
D85
I21
I23
I26
J24

Keywords:
Peer effects
Social network formation
Academic achievement
Homophily

ABSTRACT

Existing peer effects studies produce contradictory findings, including positive, negative, large, and small effects, despite similar contexts. We explore these results using U.S. Naval Academy data covering a 17-year history of the random assignment of students to peer groups. Coupled with students' limited discretion over freshman-year courses, our setting affords an opportunity to better understand peer effects in different social contexts. We find negative effects at the broader "company" level – students' social and residential group – and positive effects at the narrower course-company level within small peer groups. We suggest that peer spillovers change direction because of differences in the underlying mechanism of peer influence.

Published by Elsevier B.V.

1. Introduction

Economists have closely studied the role of peer effects at all levels of schooling but, so far, have had difficulty extending these insights into the policy arena. Research has demonstrated the prevalence of peer effects in higher education in particular. Studies of institutional data from Dartmouth College, Williams College, the University of Maryland, Berea College, the University of Amsterdam, the United States Military Academy (USMA), and the United States Air Force Academy (USAFA) have revealed peer effects of various sizes, both positive and negative, on a range of academic outcomes.[1] For example, Carrell et al. (2009) estimate that a 100-point increase in the peer-group average SAT verbal score increases freshman students' GPAs, on average, by 0.4 grade points on a 4.0 scale (for a sample of USAFA students).[2] In a follow-up study, Carrell et al. (2013) analyze a direct intervention in which the researchers themselves allocate incoming students into peer groups designed to positively influence academic marks, as predicted by their 2009 article's estimates of peer effects. The intervention, however, backfires; the targeted beneficiaries of the experiment experience statistically significant *reductions* in their

* Corresponding author.
E-mail addresses: rbrady@usna.edu (R.R. Brady), insler@usna.edu (M.A. Insler), rahman@usna.edu (A.S. Rahman).

[1] See Sacerdote (2001), Zimmerman (2003), Foster (2006), Stinebrickner and Stinebrickner (2006), Lyle (2007), Carrell et al. (2009), and Guryan et al. (2009).

[2] The researchers suggest that their estimates are larger than previous findings because of the size (approximately 30 students) and critical role of peer groups at USAFA compared to more narrow roommate linkages seen in other studies. But they also find peer group SAT math scores to be statistically insignificant to individual achievement.

http://dx.doi.org/10.1016/j.euroecorev.2017.06.013
0014-2921/Published by Elsevier B.V.

grades. In another intervention at the University of Amsterdam, Booij et al. (2017) find that low-ability students perform better when grouped with similar-ability peers rather than high-ability peers. The various findings from the literature—positive, negative, or non-existent effects—suggest that policy makers who seek to affect the composition of students, with the intent to improve student performance, are as of yet still peering into a black box.

In this paper, we seek to better understand why researchers have measured so many disparate peer effects, positive and negative, large and small. To do so, we examine how peer interaction depends on the size and purpose of the peer group, which may then lead to either positive or negative peer effects in different situations. To identify such situations, we examine data on college freshmen at the United States Naval Academy (USNA), using a dataset that includes more than 65,000 fall semester grades from over 18,000 freshmen. In employing college level data, we follow in the tradition of studies cited above (in particular, Lyle (2007) and Carrell et al. (2009)). Unlike previous studies on college students, however, the size of our dataset, which includes seventeen graduating classes spanning 1999 to 2015, is the largest examined thus far at the college level.[3] Our data and setting allow us to observe differences in peer group interaction at both a broad group level (defined as the "company," which we explain in Section 3), and subgroup level (which we call "course-company" subgroups, in which companymates have been assigned to the same course). Across various peer subgroups, the scope and nature of student interaction can differ and thus potential spillovers can vary in direction and intensity (see Manresa (2016) and Booij et al. (2017) for discussion of such possibilities).

We produce two primary empirical findings. First, we find *negative* peer effects at the company level (which defines a students' residential and social group). For both STEM (science, technology, engineering and math) and non-STEM courses, average peer ability across all freshman companymates, as measured by verbal SAT scores, negatively affects own grades. The magnitude of the peer effect is small: for example, a 100-point increase in companymates' average verbal SAT score reduces a student's STEM course grade, on average, by 0.16 grade points. Thus in attempting to replicate the findings of Carrell et al. (2009) company-level results, we instead obtain *negative* peer effects that are more similar to their 2013 hands-on experiment at USAFA (Carrell et al. (2013)). Our negative result comes from models with both linear and non-linear specifications and is robust to numerous sensitivity checks. We argue that the company level is where endogenous peer group formation is most likely to occur, thus increasing the prominence of negative spillovers.

Second, we find that at the course-company level—students in the same company *and* the same course—average peer ability *positively* affects student performance, but the effect is most prominent within small course-companymate peer groups. Specifically, an increase in course-companymates' average verbal SAT score increases own humanities or social science course grades when the size of the potential (*i.e.,* course-companymate) peer group is relatively small. For STEM course grades, course-company level peer effects are less precisely estimated; they appear to remain negative but smaller in magnitude than company level effects, but once we allow the peer effect to vary by group size, any perceptible spillovers vanish. Viewed alongside our company-level results, these findings conform with the idea that students may avoid constructive interaction in broader social settings, yet may benefit from closer interactions in smaller groups, which are more likely to perform specific common tasks.

While previous studies have looked at peer effects on college students at the dorm-room level, dorm-floor level, and at the service academy residential group level, ours is the first to examine peer effects within residential groups *and* at the course level. In such settings, there is a natural opportunity for students to collaborate on similar tasks with groups of varying sizes. Importantly, first-semester USNA freshmen not only have no discretion over their company assignment, but also—due to institutional requirements—very limited choice into their course enrollment. We show the core course enrollment mechanism is "close enough to random" to provide a suitable environment to estimate causal subgroup peer effects. Hence, we exploit the variation in pre-treatment peer ability at the company level (*i.e.*, the average SAT score across all freshman companymates), as well as the variation in pre-treatment peer ability at the *course-company* level to analyze how spillovers may differ across various observable subgroups of students. In other words, unique to other peer effects studies, we measure peer effects within more narrowly defined subgroups that are clearly designated to engage in common tasks, as opposed to interacting in a broader social group while taking disparate courses.

Our paper explores the seemingly contradictory findings in the literature, where each study observes a different social network and thus uncovers a different type of social interaction. The evidence we provide in this paper on the varying peer effects in different settings stresses the need to consider the context of each milieu, and how the scope and nature of interactions are linked to endogenous peer group formation. Failure to take these factors into consideration when attempting to manipulate peer groups can otherwise produce disastrous outcomes. Our broader picture of peer interaction has potential implications for policy changes, and we expound on these themes in the remaining sections of the paper.

In Section 2 we provide a theoretical motivation for our empirical findings: a simple conceptual model of peer interaction where a student is motivated both by homophily, which pulls her towards individuals of similar attributes, and by academic spillovers from peers, which may pull her towards a different group of individuals. The model illustrates why—in situations with broad social interaction—individuals may move towards more similar peers, while situations requiring "narrow" task-based interactions can induce individuals to collaborate with peers who have different attributes. These interactions in turn affect individual performance. The insights of our model are evocative of explanations found in a variety of peer group

[3] Notable datasets on primary and secondary students include those analyzed by Burke and Sass (2013) and Lavy et al. (2012). Also, Cornelissen et al. (2017) analyze a very large dataset of German retirees.