Contents lists available at ScienceDirect

Economics of Education Review

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

Disruptive peers and the estimation of teacher value added

Irina Horoi*, Ben Ost

Department of Economics, University of Illinois at Chicago, Chicago, IL, United States

ABSTRACT

Classroom disruption is often cited as an obstacle to effective teaching, yet little is known regarding how disruptive students influence classroom learning and teacher evaluation. In this study, we show that students with serious behavioral difficulties substantially reduce the academic performance of their peers. Since standard value-added models fail to account for these peer effects, we find that some teachers' value added is penalized because of the students she is assigned. Importantly, we show that the assignment of disruptive students to teachers is non-random, so these peer effects do not impact the evaluation of all teachers equally.

© 2015 Elsevier Ltd. All rights reserved.

1. Introduction

Understanding classroom peer effects is important both for determining optimal student grouping patterns and for generally understanding the educational production function. While classroom peer effects have been studied extensively, most research has focused on how the existence or absence of peer effects influences whether students should be tracked or placed in heterogeneous classrooms. While these considerations are first order, the existence of peer effects also implies that the educational production functions typically estimated in the literature omit an important input. To the extent that these unmeasured peer inputs are correlated with other school and classroom inputs, estimates of non-peer inputs will be biased. This point is illustrated theoretically by Lazear (2001) in the context of estimating the returns to class size, but little research has examined how peer effects influence the estimated impact of non-peer inputs empirically.

In this study, we consider the extent to which peer effects bias the estimated impact of other inputs by showing how students who are likely to be disruptive influence the estimation of teacher value added. While teachers are just one input whose estimated impact could be biased by peer effects, the use of value-added estimates in high-stakes personnel decisions makes it particularly important to correctly estimate teachers' impacts.[1]

Many different forms of peer interactions have the potential to bias value-added estimation; we illustrate the issue in the context of disruptive students for several reasons. First, surveys of teachers and administrators frequently mention disruption as a major obstacle to learning (Figlio, 2007). Research confirms that serious class disruption is a common occurrence, particularly in urban schools (Johnston, 2013; OECD, 2013). Second, while it is common for researchers to control for average peer demographic and peer academic performance when estimating teacher value added, it is rare to control for measures of disruption. Similarly, to the best of our knowledge, none of the value-added models currently in use to make high-stakes personnel decisions control for classroom disruption. Third, while there is a large literature on classroom peer effects, most of this research

---

* Corresponding author. Tel.: +1 517 505 4579.
  E-mail address: irina.horoi1@gmail.com, ihoroi2@uic.edu (I. Horoi).

[1] As of 2013, 40 states require that a teacher's annual evaluation is based in part on her value added (Doherty and Jacobs 2013).

focuses on how peer academic performance impacts one's academic performance, and fewer studies explore how the non-cognitive attributes of one's peers impact one's academic performance.

While disruption is frequently reported as an issue by teachers and administrators, datasets typically do not include direct measures of disruption, so researchers necessarily use student characteristics that proxy for disruption (Carrell & Hoekstra, 2010; Figlio, 2007; Fletcher, 2009a, 2009b; Friesen, Hickey, & Krauth, 2010). We follow this approach by using the diagnosis of an emotional disability to proxy for disruption. In the institutional context that we study, emotional disabilities are diagnosed primarily because students exhibit disruptive behaviors in school, and we show that emotional disability correlates strongly with disciplinary action such as suspension.[2] That said, without data on actual in-class behavior, it remains possible that the peer impacts that we document are due to some non-disruptive characteristic of emotionally disabled students.[3]

This article expands the literature on classroom peer effects in several ways. First, we provide carefully identified evidence that peer non-cognitive attributes can influence academic achievement. Second, we use matched longitudinal data on students and teachers over a 6-year period to show that the existence of these non-cognitive peer effects systematically influences the estimation of teacher value added. We show that for a variety of value-added models currently being used in policy, teaching emotionally disabled (ED) students reduces a teacher's estimated value added.

Identifying the impact of disruptive students on their peers is difficult because of the well-known issues of homophily, reflection and common shocks.[4] Our study addresses these concerns in several ways. First, we are able to address the possibility that students are non-randomly placed into classrooms by aggregating peer groups to the school-grade-year level and including a school-by-year fixed effect. Second, we focus on transfer students who were previously diagnosed as emotionally disabled to address concerns regarding reflection and common shocks (correlated effects). Finally, we test for non-random sorting into grades and find that the arrival of an emotionally disabled transfer student is uncorrelated with all observable predetermined characteristics, suggesting that homophily is unlikely to bias estimates of the peer effects we document.

Educational production functions invariably omit important inputs and we do not argue that this incompleteness necessarily leads to biased estimates of teacher quality. For example, parental and neighborhood inputs are rarely controlled for in value-added models, but since these inputs are likely to be highly correlated over time, controlling for lagged test score or student fixed effects plausibly addresses many concerns regarding these omitted inputs.

Compared to omitting family or neighborhood characteristics, failing to control for peer effects presents a potentially more serious issue for value-added modeling for several reasons. First, since classmates change each year, peer effects will be time varying, and thus lagged test score will not control for current peer effects. Second, the majority of value-added models emphasize individual rather than peer controls, and these individual controls are unlikely to be good proxies for peer characteristics. While some researchers have controlled for average peer achievement and demographics when estimating teacher value added, few school districts collect or use data on peer quality in measuring teacher quality (Kane, 2014).

If disruptive students were randomly assigned to teachers, then the peer effects we document would make the estimation of yearly teacher value added more noisy, but these estimates would remain unbiased. Conversations with principals suggest, however, that the classroom placement of disruptive students is a non-random decision, and our data bear this out. We find that within a school-grade-year, emotionally disabled transfer students are nearly six percentage points more likely to be placed with male teachers. More broadly, we document non-random teacher assignments of many types of transfer students, providing clear evidence that the overall assignment of transfer students to teachers is not random. The systematic assignment of students to certain teachers may be optimal for student learning, but our study suggests that the practice imposes a cost on these teachers, particularly if value added is being used for high-stakes personnel decisions.

While our study is focused on teachers, the tension we highlight between worker evaluation and task assignment is applicable to a variety of occupations. For example, financial analysts are often times rewarded for accurate forecasts, but some analysts are assigned more difficult markets than others. Similarly, universities evaluate professors based on teaching evaluations, but the material in certain courses may be more easily accessible and appealing to students. Though pay-for-performance compensation schemes are theoretically effective at eliciting optimal effort, a critical difficulty in implementation is adjusting for task assignment difficulty. In contexts where identifying task difficulty is imperfect, randomly assigning tasks to workers ensures a more fair assessment of worker productivity, but may reduce total productivity by failing to capitalize on the comparative advantage of workers when assigning tasks. Pay-for-performance schemes that fail to adjust for task difficulty create perverse incentives in which workers with a comparative advantage in difficult tasks aim to hide this information from employers.

Relative to the evaluation systems in many other occupations, value-added models include substantial adjustment for task difficulty. Teachers who are assigned low-achieving

---

[2] Emotional disability is not a DSM medical diagnosis, but rather a designation used by schools to identify students in need of services due to their behavior. As such, if a student's disability does not manifest itself through school behavior, it will not be captured in our measure.

[3] While many students who are not ED transfer students may be disruptive, the extent of the disruption might differ between ED transfer students and other disruptive students. As such, we view our study as providing evidence that classroom disruption has the potential to meaningfully impact teacher value added, but we cannot provide empirical evidence as to the total impact of all forms of disruption on teacher evaluation.

[4] "Homophily" refers to the idea that individuals may sort into groups based on their characteristics so that ego and peer outcomes will tend to be correlated in the absence of any peer effect. "Common shocks" refers to the idea that all of the individuals in a peer group may be exposed to the same inputs so that their outcomes will be correlated in the absence of any peer effect. "Reflection" refers to the difficulty between distinguishing the impact of peer characteristics on ego outcomes from the impact of ego characteristics on peer outcomes.