Economics of Education Review 31 (2012) 77–95

Contents lists available at SciVerse ScienceDirect

Economics of Education Review

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

Do reductions in class size raise students' test scores? Evidence from population variation in Minnesota's elementary schools☆

Hyunkuk Cho[a], Paul Glewwe[b,*], Melissa Whitler[b]

[a] Samsung Economic Research Institute, South Korea
[b] Department of Applied Economics, University of Minnesota, United States

ARTICLE INFO

Article history:
Received 25 June 2010
Received in revised form 23 January 2012
Accepted 29 January 2012

JEL classification:
I21

Keywords:
Class size
Student learning
Test scores
Elementary schools
Minnesota

ABSTRACT

Many U.S. states and cities spend substantial funds to reduce class size, especially in elementary (primary) school. Estimating the impact of class size on learning is complicated, since children in small and large classes differ in many observed and unobserved ways. This paper uses a method of Hoxby (2000) to assess the impact of class size on the test scores of grade 3 and 5 students in Minnesota. The method exploits random variation in class size due to random variation in births in school and district catchment areas. The results show that reducing class size increases mathematics and reading test scores in Minnesota. Yet these impacts are very small; a decrease of ten students would increase test scores by only 0.04–0.05 standard deviations (of the distribution of test scores). Thus class size reductions are unlikely to lead to sizeable increases in student learning.

1. Introduction

Policymakers, parents, school principals, and many pundits are all concerned about how much, or in some cases how little, students learn in school. This is true not only in the U.S. but also in many developed and developing countries. Despite substantial research in recent years, much is unknown about the impacts of specific education policies on student learning (see Hanushek, 2006; Hanushek & Rivkin, 2006, for recent reviews of the literature). One education policy that has received much attention is reductions in class size. Intuitively, smaller classes should allow teachers to provide more attention to each student, and to reduce time spent disciplining disruptive students, and thus should increase learning. Indeed, there appears to be a consensus among parents, teachers and school administrators that small classes improve students' academic achievement, especially among elementary (primary) school students.

Yet this consensus is at odds with academic research, which has found conflicting evidence on the impact of class size on learning. The basic problem is that students in small and large classes may differ in both observed and unobserved ways, and while it is relatively easy to account for observed differences, unobserved differences can lead to biased estimates of the impact of class size on student learning. For example, parents who take extra efforts to ensure that their children are enrolled in small classes may provide other, often unobserved, assistance to their children that helps them learn. This would lead to overestimation of the impact of reduced class size on student learning. On the other hand, students with learning difficulties may be assigned to smaller class sizes; if these learning difficulties are not observed, comparisons

☆ We would like an anonymous referee for very useful comments on a previous draft of this paper. We also thank the Center for Urban and Regional Affairs (CURA) at the University of Minnesota for a grant to support this research.

* Corresponding author.
E-mail address: pglewwe@umn.edu (P. Glewwe).

0272-7757/$ – see front matter © 2012 Elsevier Ltd. All rights reserved.
doi:10.1016/j.econedurev.2012.01.004

of learning across children in small and large classes will underestimate the impact of class size on student learning.

Perhaps the best method to measure the impact of class size reductions (and many other types of education policies) on student learning is to randomly assign students to "treatment" and "control" groups. That is, randomly assign some children to small classes and other children to large classes, and compare in later years the educational outcomes of interest across the two groups. Random assignment ensures that, on average, the two groups have the same observed and unobserved characteristics. In the U.S., there is only one study that has implemented this research design on a large scale, Project STAR (Student/Teacher Achievement Ratio Experiment), which was conducted in Tennessee from 1985–86 to 1988–89 and has been closely studied in the last 20 years.

Early results from Project STAR suggested that students in small classes scored higher on mathematics and reading tests than did students in regular-size classes (see Finn & Achilles, 1990; Folger & Breda, 1990; Word, Johnston, & Bain, 1990). These results influenced policies in other U.S. states. For example, California's decision to reduce class size in grades K-3 to 20 or less, which was first implemented in the 1996–97 school year, was directly influenced by the results from Project STAR (Bohrnstedt & Stecher, 1999; see Funkhauser, 2009, for a recent assessment). Also in 1996–97, Wisconsin initiated a program to reduce class size in those grades to less than 15. The U.S. federal government also became involved; in 1999 it initiated the Class-Size Reduction Program, which provides funds to states for hiring new teachers to reduce class size in grades 1–3 (U.S. Department of Education, 2004).

Yet even if reducing class size does increase test scores, class size reductions may not be a wise policy choice because such reductions are very expensive. That is, there may be other policies that increase learning by an equivalent amount yet at a lower cost. Continuing with the above examples, California spent $11 billion and Wisconsin spent $463 million from 1996–97 to 2004–05 to attain their class size reduction goals. If reductions in class size are less effective, in terms of the increase in learning per dollar spent, than other policies, they imply a large waste of government resources. Thus there is a great need for more rigorous research on the impact of class size on student learning.

This paper uses a method introduced by Hoxby (2000) to assess the impact of class size on math and reading test scores of children in grades 3 and 5 in Minnesota public schools. The method exploits random variation in class size due to random variation in births from year to year within school catchment areas, and also within school district boundaries. The results show positive effects of reductions in class size on mathematics and reading test scores in Minnesota. Yet the impacts of class size reductions are very small; reducing class size by ten students would increase test scores by only about 0.04–0.05 standard deviations (of the distribution of students' test scores). This implies that reductions in class size alone are unlikely to lead to sizeable increases in student learning.

The remainder of this paper is organized as follows. The next section briefly reviews the recent literature on the impact of class size on learning, after which Section 3 presents the empirical strategy used in this paper to measure this impact, which is taken from Hoxby (2000). Section 4 describes the data, and Section 5 presents the main estimation results. Section 6 checks the robustness of the estimates, and the last section summarizes the findings and discusses their implications for education policy.

2. A review of the literature on the impact of class size on learning

Ever since the publication of the Coleman Report in 1966 (Coleman et al., 1966), many social science researchers have attempted to estimate the influence of a variety of factors on student learning, including the influence of class size. Yet the research up to the mid 1990s was of variable quality due to inadequate data and serious estimation problems. For example, many data sets used had few school and teacher variables, which suggests that their estimates likely suffer from omitted variable bias. The uneven quality of the research led to a lack of consensus on the impact of class size on learning, as seen in Hanushek's (1997) review of the literature from the late 1960s to the early 1990s. He points out that, of 277 studies that attempted to estimate the impact of class size on student performance, 15% found an unexpected statistically significant positive effect, 13% found the expected significantly negative effect, and the remaining 72% found no statistically significant effect.[1]

In the last 10 years much more careful analyses have been done of the impact of class size on student learning. The best studies focus on removing or at least reducing likely sources of bias in the estimates. Almost all sources of bias arise because class size is correlated with unobserved student, parent or school variables that directly affect student learning. For example, parents who are very concerned about their children's education are more likely to move to areas where schools have small classes, which leads to a negative correlation between class size and parents' educational aspirations and thus overestimation of the impact of class size reductions if parental aspirations are unobserved. Another problem is that parents may enroll their children in schools that they perceive to be of high quality, increasing class sizes in schools perceived to be of high quality. If some or all of the quality variables that parents use to form their perceptions are not in the data set, this behavior will induce positive correlation between unobserved components of school quality and class size, and thus will cause underestimation of the impact of reductions in class size on learning. A third problem is that educators can assign students to classes of different sizes depending on their abilities; estimates will be biased if student ability is not observed, and the direction of bias will depend on whether high ability or low ability students are assigned to small classes.

The "ideal" method to overcome this problem is to implement a randomized experiment, such as the STAR program implemented in Tennessee. Random assignment of students to classes of different size assures that class

[1] See Krueger (2003) for a criticism of these findings, and Hanushek (2003) for a response to Krueger.