This article was downloaded by: [Virginia Tech Libraries]
On: 11 July 2013, At: 09:37
Publisher: Routledge
Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK

School Effectiveness and School Improvement: An International Journal of Research, Policy and Practice
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/nses20

Are teacher effects larger in small classes?

Spyros Konstantopoulos [a] & Min Sun [b]
[a] Michigan State University , East Lansing , MI , USA
[b] Virginia Tech , Blacksburg , VA , USA
Published online: 08 Jul 2013.

To cite this article: School Effectiveness and School Improvement (2013): Are teacher effects larger in small classes?, School Effectiveness and School Improvement: An International Journal of Research, Policy and Practice, DOI: 10.1080/09243453.2013.808233

To link to this article: http://dx.doi.org/10.1080/09243453.2013.808233

PLEASE SCROLL DOWN FOR ARTICLE

Taylor & Francis makes every effort to ensure the accuracy of all the information (the “Content”) contained in the publications on our platform. However, Taylor & Francis, our agents, and our licensors make no representations or warranties whatsoever as to the accuracy, completeness, or suitability for any purpose of the Content. Any opinions and views expressed in this publication are the opinions and views of the authors, and are not the views of or endorsed by Taylor & Francis. The accuracy of the Content should not be relied upon and should be independently verified with primary sources of information. Taylor and Francis shall not be liable for any losses, actions, claims, proceedings, demands, costs, expenses, damages, and other liabilities whatsoever or howsoever caused arising directly or indirectly in connection with, in relation to or arising out of the use of the Content.

This article may be used for research, teaching, and private study purposes. Any substantial or systematic reproduction, redistribution, reselling, loan, sub-licensing, systematic supply, or distribution in any form to anyone is expressly forbidden. Terms & Conditions of access and use can be found at http://www.tandfonline.com/page/terms-and-conditions

School Effectiveness and School Improvement, 2013
http://dx.doi.org/10.1080/09243453.2013.808233

Routledge
Taylor & Francis Group

Are teacher effects larger in small classes?

Spyros Konstantopoulos[a]* and Min Sun[b]

[a]*Michigan State University, East Lansing, MI, USA;* [b]*Virginia Tech, Blacksburg, VA, USA*

(*Received 3 October 2011; final version received 9 March 2013*)

Teachers spend most of their time in school in classrooms, and their instruction and teaching practices may be affected by classroom context such as class size. We examine whether teacher effects interact with classroom context such as class size. Specifically, we seek to determine whether teacher effects are more pronounced in small classes than in regular size classes in early grades. The results point to some evidence that teacher effects are more variable in small classes than in regular size classes. However, the pattern does not seem to be systematic.

Keywords: teacher effects; small classes; multilevel models; student achievement

Downloaded by [Virginia Tech Libraries] at 09:37 11 July 2013

Introduction

In many different education systems across the world, teachers play a fundamental role in educating students at the elementary and secondary levels. Folk knowledge as well as our own experiences in elementary and secondary schools suggest that the effects teachers have on students can be remarkable in some cases. Specifically, teachers' profound effects on students' learning and school experiences may last well into adulthood (see Pedersen, Faucher, & Eaton, 1978). It is not uncommon to attribute academic success to particular teachers during our school years and refer to them as great or influential teachers. Nonetheless, teachers may differ considerably in their effectiveness to promote their students' academic achievement, and this variation in teacher effectiveness seems to be large (Nye, Konstantopoulos, & Hedges, 2004; Rivkin, Hanushek, & Kain, 2005; Rowan, Correnti, & Miller, 2002). While some teachers are very effective in facilitating student learning, others are not as effective. Still, improving teacher effectiveness is an important mechanism through which schools promote student learning. For instance, in the US, the No Child Left Behind Act (NCLB) of 2001 has stressed the importance of teacher effectiveness and has mandated plans to improve teacher effectiveness and produce more effective teachers.

Of course, teachers do not spend their time in a context-free environment. Instead, teachers spend most of their school time in classrooms, and their instruction and teaching practices may be affected by classroom context. One important aspect of classroom context is classroom size. The last 3 decades, class size reduction programs have been identified by many researchers as a promising school mechanism that can increase achievement for all students (e.g., Finn & Achilles, 1990; Konstantopoulos, 2008; Krueger, 1999). As a result, in the US many states have passed legislation for the implementation of such programs, and in Europe countries such as the UK and Greece

*Corresponding author. Email: spyros@msu.edu

© 2013 Taylor & Francis

have enacted policies about the maximum number of students in a classroom (Blatchford, Basset, & Brown, 2011).

The purpose of the present study is to examine whether teacher effects are moderated by classroom context such as class size. Specifically, we are interested in investigating whether the effects teachers have on student achievement differ considerably by class size (e.g., small or regular size classes). To that end, we use data from a high-quality 4-year field experiment, Project STAR (Student Teacher Achievement Ratio), where students and teachers in early grades (i.e., kindergarten through third grade) were randomly assigned to classrooms of different sizes within schools. We follow previous work by Nye et al. (2004), and operationalize teacher effects as classroom-specific residuals that are adjusted for student effects such as gender, race, socioeconomic status (SES), and school fixed effects. In line with Nye et al. (2004), we use the term *teacher effects* instead of classroom effects to stress the central role of the teacher in the classroom processes. Because of random assignment, we assume that, on average, student and peer effects are similar across classes. Teachers were also randomly assigned to classes, which helps address potential selection biases in the baseline. That is, we assume that class size reduction is the only difference among classes in the baseline and that teachers in different types of classes are the driving force in terms of classroom practices, learning activities, and interactions among students and teachers. In turn, we assume that possible changes in student achievement are due to teachers, other things being equal.

Related literature

Teacher effects

The literature about teacher effects encompasses different research strands. One research strand discusses the association between teacher characteristics and student achievement. Within that strand, some reviewers of the teacher effects literature have argued that measured teacher characteristics such as educational preparation, experience, or salary are only slightly related to student achievement (Hanushek, 1986). Other reviewers, however, have reported positive effects of teacher characteristics such as experience and teacher education on student achievement (Greenwald, Hedges, & Laine, 1996). More recent work has provided additional support about the positive association between teacher experience and student achievement (Clotfelter, Ladd, & Vigdor, 2006; Nye et al., 2004). There is also evidence that National Board certified teachers seem to be more effective than other teachers (Goldhaber & Anthony, 2007). Typically, regression or multilevel models are used to determine such associations between teacher variables and achievement.

A second research strand includes studies about the variance in average student achievement across classrooms (or teachers) controlling for student background. Within this strand, teacher effects are defined broadly as the proportion of variance in student achievement that is attributed to teachers. Regression or multilevel methods have been used to estimate the part of the variance of student achievement that is due to teacher effects. The underlying assumption is that the between-classroom or teacher variation in student achievement is caused by variation in teacher effectiveness. Family background is typically taken into account in such models in order to determine the value added of the teachers. Results from these studies have suggested that there is significant variation in teacher effectiveness (Goldhaber & Brewer, 1997; Murnane & Phillips, 1981; Nye et al., 2004; Rowan et al., 2002). That is, teachers differ significantly in how they impact student

Downloaded by [Virginia Tech Libraries] at 09:37 11 July 2013

achievement, with some teachers being more effective, whilst others are not as effective. In this study, we follow a similar approach to that used in Nye et al. (2004).

A third research tradition focuses more on what goes on in the classroom. In contrast with the previous two traditions of research that rely on statistical modeling and have a macro, black box, perspective on the effects of teachers on student achievement, this strand has a micro, within-classroom perspective about the mechanism of teacher effects. Specifically, in this line of research the objective is to identify classroom practices or processes that facilitate student learning. Pioneering work by Good and Brophy (1987) examined the association between good teaching practice and student achievement. Such studies are often referred to as process-product studies because they identify classroom processes and teaching practices that are associated with student outcomes or products such as student achievement. Parameters that have been discovered to be related with student achievement include teacher confidence in teaching students successfully, efficient allocation of classroom time to instruction and academic tasks, effective classroom organization and group management, and active/engaging teaching that emphasizes understanding of concepts (Good & Brophy, 1987). These processes have been documented as positive correlates of student achievement (Good & Brophy, 1987). All three research strands complement one another and provide a more comprehensive picture of what teacher effects are.

Downloaded by [Virginia Tech Libraries] at 09:37 11 July 2013

Class size effects

The effects of class size reduction on student achievement have been examined empirically via various research designs and analyses over the past few decades. Numerous small-scale experimental and quasi-experimental studies have investigated the effects of class size on student achievement. Meta-analytic reviews of early work on class size effects have suggested that class size reduction has positive effects on student achievement and that these effects become greater as class size becomes smaller (Glass, Cahen, Smith, & Filby, 1982; Glass, & Smith, 1979). The benefits of small classes appeared to be more pronounced in classes with 20 or 10 students as opposed to 30 or 40 students. The class size achievement association was more pronounced in randomized studies (Glass & Smith, 1979).

Other studies have examined the effects of class size reduction on student achievement using data from non-experimental studies. Typically, these studies compute the association between class size and student achievement, adjusting for important factors in student background such as gender, race/ethnicity, SES, and previous academic achievement. The interpretation of these studies' results has been controversial because of selection bias. That is, this body of research has not yielded clear evidence about the effects of small classes. For instance, Angrist and Lavy (1999) found that reducing class size increased fourth and fifth graders' scores of Israeli students significantly. In France, Bressoux, Kramarz, and Prost (2009) have reported benefits for students in small classes, especially for low achievers. Similarly, Pong and Pallas (2001) found positive small class effects on eighth-grade achievement using the Trends in International Mathematics and Science Study (TIMSS) data across different countries. In contrast, Hoxby (2000) reported that smaller classes had little to no effect on student achievement, and Milesi and Gamoran (2006) found no evidence of class size effects on student achievement.

Perhaps the most conclusive evidence about class size effects has been produced from analyses of Project STAR data. Early analyses of Project STAR data have indicated that small classes had positive effects on student achievement in early grades (Finn & Achilles, 1990). More recent analyses of the same data also demonstrated that small classes increase student achievement for all students (Hanushek, 1999; Krueger, 1999; Nye, Hedges, & Konstantopoulos, 2000). The average class size in small classes was 15 to 16 students and in regular size classes 22 to 23 students. Overall, the findings from studies that used high-quality experimental data have been consistent and have provided evidence about the positive immediate effects of small classes on student achievement. Follow-up data have also been analyzed, and the results have provided strong evidence that the positive effects of small classes on student achievement persist through the eighth grade (Finn, Gerber, Achilles, & Boyd-Zaharias, 2001; Konstantopoulos & Chung, 2009). These findings have also indicated that the lasting benefits of small classes become greater the longer students were in small classes in earlier grades. For example, the lasting benefits for students who were in small classes for all 4 years (i.e., kindergarten through third grade) were nearly twice as large as the benefits for students who were in small classes for only one year (Konstantopoulos & Chung, 2009).

Downloaded by [Virginia Tech Libraries] at 09:37 11 July 2013

Teaching practices and class size

Over time, published research has provided some evidence about classroom practices in smaller and larger classes. Early work by Smith and Glass (1980) indicated that in small classes there is more individualized instruction, more student participation, higher frequency of interactions among students and teachers, and higher quality of instruction overall. Along the same lines, another contemporary study found that students in small classes received more attention through questions and probes than students in other classes (Bourke, 1986). However, teachers in small classes were more likely to teach the class as a whole (Bourke, 1986). Small classes require less management and are less noisy, which should promote student learning (Bourke, 1986). Early reviewers of the literature (e.g., Odden, 1990) have reported that, in small classes, students are more attentive and teachers have more opportunities for curriculum expansion activities. One study that analyzed data in second and third grades from Project STAR has found that teachers did not modify their teaching practices considerably in small and regular size classes in these grades (Evertson & Randolph, 1989). The authors attributed this finding at least partly to the mandated curriculum and its goals about student assessment. More recent work has used large-scale data to examine how class size changes teaching practices and have provided additional evidence that, in small classes, teachers spent more time in individualized instruction and less time on group instruction (Betts & Shkolnik, 1999). In addition, teachers in small classes were more likely to devote time to review activities. The authors concluded that, overall, teachers in small classes did not change their instruction practices dramatically.

Previous work has argued that class size offers the context in which teachers and students interact (Zahorik, 1999). This notion implies that the effects of class size are "mediated" by what teachers and students do in the classrooms. Along the same lines, Anderson (2000) points out that small classes provide opportunities for teachers to modify their instruction accordingly to maximize learning. However, in order to take advantage of the opportunities that small classes provide, teachers need first to comprehend what types of instructional changes should be enacted when they teach small classes (Anderson, 2000). Anderson provided a useful theory of change model about class size effects that stresses three main factors: in small classes, (a) there should be fewer discipline problems, (b) teachers should have greater knowledge of their students, and (c) teacher satisfaction and enthusiasm should be greater. Now, these factors should result in more instructional time and individualized instruction which should result in more in-depth content coverage. In small classes, teachers should put greater efforts in their teaching, students should be more engaged, and thus there should be greater opportunities to learn, which ultimately will improve student achievement.

More recent work has reported that students in larger classes are more likely to be distracted from classroom work and are more frequently off task (Blatchford, Edmonds, & Martin, 2003). In that study, the authors also found no support that the relations among students are better in smaller classes. Blatchford and colleagues (2011) have also found that students in small classes receive more individual attention from their teachers, and that they have more opportunities to interact with their teachers. However, in small classes students were less likely to work in groups and engage in collaborative group work, a result consistent with the results reported by Betts and Shkolnik (1999). These results were by and large replicated in other studies that have reported that in small classes there were more frequent interactions between students and teachers with respect to knowledge building and learning activities (Blatchford, Russell, Bassett, Brown, & Martin, 2007). Further, teachers in smaller classes were more likely to focus their attention on specific students and provide individualized instruction and feedback (Blatchford, Moriarty, Edmonds, & Martin, 2002).

Downloaded by [Virginia Tech Libraries] at 09:37 11 July 2013

Research hypothesis

Given the findings of previous work about small classes especially in early grades, it seems more likely that management- and behavior-related issues are addressed in a shorter time in small classes than in regular size classes. In turn, less time on management and discipline should translate to more time spent on learning activities in small classes than in regular size classes. If according to Anderson's (2000) model, there is more individualized instruction, student engagement, and teacher enthusiasm in small classes, one would expect ultimately to find larger learning gains within a school year in small classes than in regular size classes.

However, is reducing the numbers of students in classes enough to produce a positive change in student outcomes? Indeed, class size provides the context in which learning takes place, but can a simple alteration of the size of a classroom change student achievement? It seems that certain conditions need to be met in order for that to happen. Class size reduction can offer opportunities to teachers to facilitate learning. That is, small classes provide the context that could enable teacher effects, but the degree to which this facilitation takes place is unclear. For example, some teachers may teach a larger class of 30 students and a smaller class of 15 students the same way (see Betts & Shkolnik, 1999). Alternatively, other teachers may modify their instruction practices to take advantage of the fact that there are fewer students in the classroom. That is, the important question is what do teachers do in classrooms with fewer students? Smaller classes may provide opportunities for more effective instruction, but such opportunities need to be exploited by the teacher in order to promote student learning. For instance, some teachers who are assigned to small classes may modify their instructional practices accordingly, and in that case perhaps the teacher effects on student achievement may be pronounced. Whenever teachers take advantage of the small class context, it seems plausible to expect larger increases in student achievement. Other teachers assigned to small classes, however, may not change their classroom practices, and as a result, increases in student achievement may not be as apparent.

It should be noted that simply altering instruction and classroom practices may not be enough; it would have to involve the "right" changes in teaching practices. For instance, individualized attention is an improvement in small classes, but to exercise repetitive teaching, over-individualize, or provide feedback instantly to students may be disruptive (Blatchford et al., 2007; Evertson & Randolph, 1989). Specifically, Blatchford et al. (2007) have argued that although individual attention in small classes is beneficial, teachers should consider modifying their instruction to include group work when needed.

Given the findings from prior work, one could expect that teachers' instructional practices may vary more in small classes than in regular size classes. If these differences translate to differences in achievement between classrooms, then one would expect to observe larger variability in achievement between small classes than between regular size classes. This variability is what we call teacher effects in this article (see also Nye et al., 2004). Along these lines, our main research hypothesis is that the variance in achievement between smaller classes will be larger than the variance in achievement between regular size classes (or that teacher effects will be larger in small classes).

In addition, given that our data come from a randomized experiment, and assuming that random assignment was successful, one would not be as concerned about average differences in student or classroom/teacher variables that could have influenced teacher effects. That is, on average, students across classes were similar, and thus the data from Project STAR likely satisfy the assumption that small and regular size classrooms are equally teachable on average (see Beckerman & Good, 1981, for a discussion about teachable classes).

The Tennessee class size experiment

Project STAR, also known as the Tennessee class size experiment, is a 4-year large-scale experiment that took place in Tennessee in the mid-1980s. The experiment was commissioned in 1985 by the Tennessee state legislature and was implemented by a consortium of Universities and the Department of Education in Tennessee. The experiment lasted for 4 years from Kindergarten to Grade 3, and the total cost, including hiring teacher and teacher aids, was about $12 million. The state of Tennessee paid for hiring additional teachers and classroom aides. Project STAR is considered one of the greatest experiments in education in the U.S.

In the 1st year of the experiment, a cohort of more than 6,000 students in more than 300 classrooms in 79 elementary schools in 42 districts in Tennessee participated. The sample included a broad range of schools and districts (urban, rural, wealthy, and poor). Districts had to agree to participate for 4 years, allow school visits for verification of class sizes, interviewing, and data collection, and include extra student testing. They also had to allow research staff to assign pupils and teachers randomly to class types within schools and to maintain the assignment of students to class types from kindergarten through third grade.

The cohort of kindergarten students was assigned randomly to different types of classrooms within each school: small classes (with 13 to 17 students), regular size classes (with 22 to 26 students), or regular size classes with a full-time aide. Teachers were also assigned randomly to classes of different types. Some students entered the study in the first grade and subsequent grades and were assigned randomly to classes at

Downloaded by [Virginia Tech Libraries] at 09:37 11 July 2013

that time. Teachers at each subsequent grade level were also assigned randomly to classes as the experimental cohort passed through the grades.

The validity of Project STAR

The internal validity of the Project STAR data depends on whether random assignment effectively eliminated preexisting differences between students and teachers assigned to different types of classrooms. One could examine empirically whether there were any differences on pre-existing observed characteristics of teachers or students. Unfortunately, no pretest scores were collected in Project STAR, so it is not possible to examine differences in pre-kindergarten achievement. Other student variables such as age, race, and SES, however, were available. Krueger (1999) examined the effectiveness of the randomization among treatment groups, small, regular, and regular size classes with a full-time aide, and found that across three observed variables such as SES, minority group status, and age, there were no significant differences between classroom types once school differences were taken into account. Krueger also found that there were no significant differences across classroom types with respect to teacher characteristics such as race, experience, and education. He concluded that it did not appear that random assignment was compromised by these observed characteristics. In addition, a more recent study examined whether randomization was successful across classrooms within treatment types within schools (Nye et al., 2004). The results produced from that study were consistent with what would be expected if randomization were successful. That is, no systematic differences were found for specific observed student characteristics between classrooms that were in the same treatment type (e.g., small classes) within schools.

Another important threat to the internal validity of Project STAR was attrition. Approximately 28% of the students who participated in Project STAR in kindergarten were not part of the study in the first grade. The attrition rate was nearly 25%, for students who participated in the study in the first grade but were not present in the second grade. Twenty percent of the students dropped out of the study after the second grade, and thus they were not present in the third grade. Across all grades, about 50% of the students who were part of the experiment in kindergarten were still part of Project STAR in the third grade. The effects of differential attrition on the estimates of class size have been discussed in previous work (Krueger, 1999; Nye et al., 2000). For example, Krueger (1999) examined whether differential attrition among types of classrooms biased the effects of class size on student achievement. He concluded that it seemed unlikely that differential attrition biased the class size estimates. The same conclusion was reached by Nye et al. (2000) independently, who used slightly different methods and examined differences in achievement for dropouts and stayers.

Statistical analysis

Distribution of classes

In order to identify the nested structure of the data, we computed first the number of classes across all schools for each grade separately. Second, we computed the number of classes for each school and grade separately. Third, we computed the number of small, regular, and regular size with a full-time aide classes for each school and grade. All three

Downloaded by [Virginia Tech Libraries] at 09:37 11 July 2013

analyses provided a clear picture of the distribution of classes overall and the types of classes across schools and grades.

Computation of teacher effects in each grade

The main objective of the study was to examine whether teacher effects in a specific grade (e.g., kindergarten) are more pronounced in small than in regular size classes. To address this objective, we conducted analysis for each grade separately (i.e., kindergarten to third grade). The outcomes were the Stanford Achievement Test (SAT) reading and mathematics scores collected as part of Project STAR at the end of each school year. SAT is a widely used test that measures academic achievement of elementary and secondary school students that is designed to measure, among other things, word reading, reading comprehension, and mathematics computation and application. We operationalized teacher effects as classroom-specific residuals, and the variance of these residuals indicated the magnitude of the effects of how different the effects of teachers can be (see Nye et al., 2004).

First, we computed teacher effects for all students across all classrooms and schools using two-level models. At the first (i.e., student) level, we included student characteristics such as gender, race, and SES. This is common practice in the teacher and school effects literature where student achievement is typically regressed on student demographic characteristics such as SES, race/ethnicity, and gender (Lee, 2000). At the second (i.e., class) level, we included school fixed effects (i.e., dummies to capture differences across schools) in order to explain the between-school variance in achievement that is due to school differences. To compute the teacher effects, we used the same specification for mathematics and reading achievement for each grade (i.e., kindergarten, first, second, or third grade) separately. Hence, for each grade g (g = k, 1, 2, 3) the first-level regression equation for student *i*, in class *j* is

$$Y_{ij} = \beta_{0j} + \beta_{1j}FEMALE_{ij} + \beta_{2j}LOWSES_{ij} + \beta_{3j}MINORITY_{ij} + \varepsilon_{ij}\,. \quad (1)$$

At the second level, the regression model for the classroom intercepts is

$$\beta_{0j} = \gamma_{00} + \mathbf{\Gamma SCHL}_j + \xi_{0j,} \quad (2)$$

where **SCHL** are school fixed effects (i.e., dummies). It is important to include the school dummies in the second level in order to tease out the proportion of the between-classroom variance that is attributed to differences between schools. All Level 1 slopes were treated as fixed at the second level, that is, $\beta_{1j} = \gamma_{10,} \beta_{2j} = \gamma_{20,} \beta_{3j} = \gamma_{30.}$. In the above equations, Y_{ij} represents student achievement (mathematics or reading), γ_{00} is the intercept, γ_{10} is the overall gender effect, *FEMALE* is a dummy variable for being a female student, γ_{20} is the overall low SES effect, *LOWSES* is a dummy variable for free or reduced price lunch eligibility, γ_{30} is the overall minority effect, *MINORITY* is a dummy variable for minority group membership indicating that the student was Black, Hispanic, or Asian, ε_{ij} is a student-specific random effect or residual, and ξ_{0j} is a classroom-specific random effect or residual. In this model, the variance of the error term is divided into two parts: the within-classroom and the between-classroom variance. The classroom-specific random effects, ξ, represent the teacher effects, and the variance of these residuals is adjusted by student gender, SES, minority group status, and school fixed effects. The variance of these random effects indicates differences in teacher effects or how large the effects can be

Downloaded by [Virginia Tech Libraries] at 09:37 11 July 2013

(see Nye et al., 2004). Notice again that the variance of the classroom-specific residuals is adjusted for school differences (i.e., school fixed effects), and therefore it should represent only classroom effects that we call teacher effects.

Second, we computed teacher effects within small and regular size classes separately to determine whether there were differences in the estimates of teacher effects within each type of classroom. We focus our within-class type analysis to comparisons between small and regular size classes. That is, our contrast of interest involved the estimates of teacher effects in small and regular size classes. The two-level model used in the within type of classroom analysis was identical to the one illustrated in Equations (1) and (2). This analysis was conducted for mathematics and reading scores separately for each grade.

To determine whether the variability in achievement differed significantly between small and regular size classes, we used a typical F test for equality of two variances (see Snedecor & Cochran, 1989), namely

$$F = \frac{S^2_{sml}}{S^2_{reg}} \tag{3}$$

where S^2 indicates the sample variance, subscript *sml* indicates small classes, and subscript *reg* indicates regular size classes. This F test has $n_{sml} - 1$, $n_{reg} - 1$ degrees of freedom, where n indicates sample size. We set the level of statistical significance according to typical standards (i.e., 0.05).

Downloaded by [Virginia Tech Libraries] at 09:37 11 July 2013

Sample description

A cohort of approximately 6,500 kindergartners participated in the study in the 1st year. Nearly 50% of the students in the sample were female and low-SES students. Approximately one third of the students were minorities (more than 90% of them were Black), and about 30% of the students were in small classes. Nearly 3,000 students participated in the study for all 3 years, and more than 11,000 students were part of Project STAR in the 4-year period. The outcomes of interest were mathematics and reading scores that were standardized to have a mean of zero and a standard deviation of one. On average, there were 15 to 16 students in small classes and 22 to 24 students in regular size classes across grades. The range in small classes was between 11 and 20 students across grades, whilst in regular classes the range was between 15 and 29 students across grades.

The number of classes by type of classroom and grade is reported in Table 1. On average and across grades, there are nearly 100 regular size and 100 regular size with a full-time aide classes, and nearly 130 small classes. Overall, there were nearly 330 classes across grades. Approximately 80% of the schools had three to five classes in each grade

Table 1. Classroom type by grade.

Type of Classroom	Grade K	Grade 1	Grade 2	Grade 3
Small	127	124	133	140
Regular	99	115	100	89
Regular with Full-Time Aide	99	100	107	107
Total	325	339	340	336

Downloaded by [Virginia Tech Libraries] at 09:37 11 July 2013

Table 2. Number of classes by school and grade.

Classes Per School	Number of Schools	Classes Per School	Number of Schools
Grade K		Grade 1	
2	1	3	22
3	28	4	25
4	26	5	11
5	15	6	13
6	4	7	4
7	4	12	1
8	1		
Grade 2		Grade 3	
3	13	3	14
4	33	4	31
5	15	5	17
6	8	6	8
7	5	7	4
11	1	9	1

(see Table 2). However, there were also a few schools that had more than five classes in each grade. For instance, in kindergarten four schools had six classes and another four schools had seven classes. In the second and third grades, the number of schools that had six classes was twice as large (i.e., 8), and in the first grade the number of schools with six classes tripled (i.e., 13). Still, the overwhelming majority of schools had less than six classes in each grade. In addition, more than 90% of schools had one or two small, regular, or regular with full-time aide classes in each grade (see Table 3). For instance, in kindergarten 41 schools (52%) had two small classes and 35 schools (44%) had one small class. Similar patterns were observed across grades and for all types of classes.

Table 3. Number of classes by type of class, school, and grade.

Grade K		Grade 1		Grade 2		Grade 3	
Classes Per School	Number of Schools	Classes Per School	Number of Schools	Classes Per School	Number of Schools	Classes Per School	Number of Schools
Small Classes							
1	35	1	32	1	23	1	20
2	41	2	40	2	46	2	45
3	2	3	4	3	6	3	10
4	1						
Regular Classes							
1	59	1	43	1	49	1	50
2	17	2	29	2	22	2	18
3	2	3	3	3	1	3	1
		5	1	4	1		
Regular Classes with Aide							
1	59	1	47	1	46	1	46
2	20	2	23	2	27	2	27
		3	1	3	1	3	1
		4	1	4	1	4	1

Downloaded by [Virginia Tech Libraries] at 09:37 11 July 2013

Results

Teacher effects were computed separately for four grades: kindergarten, first, second, and third grade. Because the teacher effects were defined as classroom residuals, the mean of the teacher effects distribution is zero. The distribution of teacher effects is, in this case, defined well by its variance, which indicates the magnitude of teacher effects. That is, larger variances indicate larger differences among teachers in terms of effectiveness. If differences in teacher effectiveness are larger in smaller classes, then one would expect larger between-classroom variances in small classes. The estimates of the Level 1 and Level 2 variances produced by the two-level model analyses using data across all students, classes, and schools for each grade are given in Table 4. Results for reading achievement are given in the left columns of the table and those for mathematics achievement are given in the right columns. We report results produced by a two-level model that controls for student characteristics (e.g., gender, race, etc) as well as school fixed effects. Sample sizes for students and teachers/classes are also reported. Below, we discuss the results.

The results indicate that the majority of the variance estimates in reading and mathematics achievement is between students within classes, not between classes. In fact, the between-student within-class variances are 6 to 7 times larger than the variances between classes. Still, the estimated between-teacher variance components in reading achievement ranged from 0.081 to 0.110 and were significantly different than zero suggesting significant teacher effects (see Nye et al., 2004). In mathematics, the second-level variance components estimates were slightly larger than those in reading. In particular, the estimated between-teacher variance components in mathematics achievement ranged from 0.111 and 0.131. The largest estimates were observed in kindergarten for reading and in first grade for mathematics. Overall, across grades there is systematic evidence about significant teacher effects.

The results of the Level 1 and Level 2 variance component estimates produced by the two-level model used in the within small or regular size class analyses are reported in Table 5. The structure of Table 5 is identical to that of Table 4. First, we discuss the results in small classes. The estimated between-teacher variance components in reading achievement ranged from 0.081 and 0.157. In mathematics, the second-level variance components estimates were slightly larger. In particular, the estimated between-teacher variance components in mathematics achievement ranged from 0.084 and 0.167. The largest estimates were observed in kindergarten for reading and in first grade for mathematics. Overall, across grades there is systematic evidence of significant teacher effects.

Now, we discuss the results in regular size classes. The estimated between-teacher variance components in reading achievement ranged from 0.066 and 0.108. In mathematics, the second-level variance components estimates were slightly larger. In particular, the estimated between-teacher variance components in mathematics achievement ranged from 0.093 to 0.127. The largest estimates were observed in kindergarten for reading and in first and third grades for mathematics. Overall, across grades there is systematic evidence of significant teacher effects.

The results suggest by and large more pronounced differences in classroom achievement in small classes than in regular size classes, especially in kindergarten and first grade for reading and mathematics. In reading, the variance estimates of classroom achievement were also larger between small classes than between regular size classes in Grades 2 and 3. The results in Grades 2 and 3 for mathematics indicated that the variance estimates between regular size classes were larger than between small classes. Overall, in six out of a total of eight comparisons the variances between small classes were greater than those between regular size classes (see Table 6). We determined significant differences between

Downloaded by [Virginia Tech Libraries] at 09:37 11 July 2013

Table 4. Two-level residual variance components estimates by grade: total sample.

	Reading				Mathematics			
	Number of Students	Within-Classroom Variance	Number of Classes	Between-Classroom Variance	Number of Students	Within-Classroom Variance	Number of Classes	Between-Classroom Variance
Kindergarten	5770	0.676*	325	0.110*	5852	0.663*	325	0.123*
First Grade	6393	0.677*	339	0.081*	6596	0.647*	339	0.131*
Second Grade	6056	0.695*	340	0.088*	6044	0.671*	340	0.119*
Third Grade	5987	0.748*	336	0.086*	6065	0.700*	336	0.111*

*$p < 0.05$.

Downloaded by [Virginia Tech Libraries] at 09:37 11 July 2013

Table 5. Two-level residual variance components estimates by grade and classroom type (small or regular).

	Reading				Mathematics			
	Number of Students	Within-Classroom Variance	Number of Classes	Between-Classroom Variance	Number of Students	Within-Classroom Variance	Number of Classes	Between-Classroom Variance
Small Classes								
Kindergarten	1734	0.725*	127	0.157*	1757	0.733*	127	0.150*
First Grade	1823	0.713*	124	0.092*	1867	0.668*	124	0.167*
Second Grade	1791	0.699*	133	0.081*	1786	0.698*	133	0.084*
Third Grade	1913	0.781*	140	0.107*	1937	0.726*	140	0.096*
Regular Classes								
Kindergarten	2001	0.610*	99	0.108*	2027	0.647*	99	0.119*
First Grade	2434	0.648*	115	0.074*	2505	0.610*	115	0.121*
Second Grade	2048	0.682*	100	0.068*	2046	0.657*	100	0.093*
Third Grade	1814	0.718*	89	0.066*	1832	0.674*	89	0.127*

*$p < 0.05$.

Table 6. *F* tests about equality of variances in small and regular size classes by grade.

	Small Class		Regular Class			
	N	Variance	*N*	Variance	*F* test	*p* value
Reading						
Kindergarten	127	0.157	99	0.108	1.457	0.025
First Grade	124	0.092	115	0.074	1.249	0.114
Second Grade	133	0.081	100	0.068	1.190	0.180
Third Grade	140	0.107	89	0.066	1.627	0.007
Mathematics						
Kindergarten	127	0.150	99	0.119	1.262	0.113
First Grade	124	0.167	115	0.121	1.388	0.038
Second Grade	133	0.084	100	0.093	0.898	0.721
Third Grade	140	0.096	89	0.127	0.758	0.929

Note: *N* = sample size.

Downloaded by [Virginia Tech Libraries] at 09:37 11 July 2013

variance estimates using *F* tests, and *F* test values greater than one indicate that the variance between small classes was larger than that between regular size classes. Specifically, in kindergarten the *F* test was significant in reading achievement indicating significant differences in variance estimates between small and regular size classes. That is, the between-classroom variance among small classes was significantly larger than that among regular size classes. In kindergarten mathematics, however, we did not detect any differences in the variances between small and regular size classes.

In first grade, the *F* test indicated that in first-grade mathematics the between-classroom variance estimate among small classes was significantly higher than that among regular size classes. In first grade reading, however, we did not detect any differences in the variances between small and regular size classes. Similarly, in the second grade we did not detect any differences in variances between small and regular size classes. Finally, in third grade the *F* test indicated that the between-classroom variance estimates in reading among small classes were significantly larger than those among regular size classes. By and large, these results provide some evidence that teacher effects were larger in small classes than in regular size classes. However, the results do not seem to be systematic. That is, we constructed eight variance estimates comparisons between small and regular size classes, and although in six of the eight contrasts the *F* tests were greater than one, only in three of those occasions were the *F* tests significant at the 0.05 level.

Discussion

We investigated whether the effects of teachers vary by classroom size (e.g., small or regular size classes). Specifically, we determined whether the teacher effects differ in small and regular size classes in specific grades (i.e., kindergarten through third grade). We used high-quality data from Project STAR where teachers and students were randomly assigned to classrooms within schools. The magnitude of teacher effects was captured by the between-classroom variance in achievement.

The results overall point to systematic evidence that teacher effects in early grades are significant both in mathematics and reading achievement. The analyses within classroom type suggested that teacher effects were more pronounced in small than in regular size classes. The variance estimates in small classes were typically larger, especially in kindergarten and first grade, than those in regular size classes. In kindergarten, first, and third grades, the teacher effects in small classes were typically significantly larger than the teacher effects in regular size classes. Thus, these results are to some degree in accord with the notion that teachers' instructional practices may differ more in small classes than in regular size classes, and that these differences result in larger differences in achievement between small classrooms. It is possible that some teachers in small classes change their instruction and design learning activities accordingly in order to take advantage of the fact that there are fewer students in the classroom. In contrast, other teachers may not modify their practices. These differences result in an increased variability of teacher effects in small classes in some grades and partially support the notion that while some teachers take advantage of the dynamics of small classes, others do not. In regular size classes, there is also some variability in teacher effects, but by and large the effects are more consistent in regular size classes than in small classes. The results that teacher effects estimates in small classes are typically larger in kindergarten and first grade are in line with the conjecture that, in the first 2 years of elementary school, teachers in small classes may spent less time on management and discipline issues and more time on productive instruction and learning activities.

Our findings are congruent with Anderson's (2000) model and previous empirical evidence (e.g., Evertson & Randolph, 1989). Indeed, small classes seem to facilitate teacher effects, but some teachers in small classes in particular, appear to take advantage of the opportunities presented in small classes to become more effective teachers. Unfortunately, classroom observations or teacher logs are not available to us, and therefore it is impossible to know the actual teaching practices that took place in each classroom over the 4-year period. Still, the results indicate differences in teaching practices between small and regular size classes.

Further, because of the computation of teacher effects as classroom specific residuals, it is possible that these estimates include some amount of measurement error. That is, the results reported here may be affected by measurement error and may overstate or understate the teacher effects. This is a potential limitation of the study. Another potential limitation of the study is that the teacher effects observed here may not be grade specific but may be cumulative. For example, second-grade students may experience both the effects of their kindergarten and first-grade teachers as well as their second-grade teachers. Thus, it is unclear whether the teacher effects observed in one grade are due to a specific teacher in that grade or whether they are a combination of teacher effects across grades. This could be more of a concern in the second and third grades. If the effects of previous teachers are included in the teacher effects we estimated, then we may have overestimated the teacher effects especially in the third grade. One could then argue that the best evidence about teacher effects is in kindergarten.

Moreover, one could argue that our definition of teacher effects is closely related to the definition of classroom effects, which may include peer effects and classroom dynamics (e.g., interactions among students) that are related to the size of the classroom. Previous work has indicated that in Project STAR the peer effects were overall small, did not affect teacher effects that much, and did not seem to interact with class size (Konstantopoulos & Sun, 2012). Still, it is unclear to what degree the teacher effects we have measured are "pure" teacher effects, because it is possible that they incorporate classroom dynamics that could vary by class size. Unfortunately, we do not have such data to test this possibility.

In conclusion, the present study showed that teachers matter in early grades and in classrooms of different sizes. The findings highlight the importance of teacher effects in early grades regardless of class size. There is, however, some evidence that teacher effects are more pronounced in small classes in some grades than in regular size classes. Teachers are

Downloaded by [Virginia Tech Libraries] at 09:37 11 July 2013

more likely to make a difference in small classes and promote student achievement when they understand the dynamics and opportunities in small classes and alter their teaching to maximize student learning.

Notes on contributors

Spyros Konstantopoulos is associate professor of measurement and quantitative methods in the Department of Counseling, Educational Psychology, and Special Education at Michigan State University. His research interests include teacher and school effects, class size effects, power analysis in experimental designs with nested data structures, and multilevel meta-analysis.

Min Sun is an assistant professor of Educational Leadership and Policy Studies at Virginia Tech. Her research focuses on educational policy analysis, the development and evaluation of teachers and principals, and quantitative research methodology.

Downloaded by [Virginia Tech Libraries] at 09:37 11 July 2013

References

Anderson, L. W. (2000). Why should reduced class size lead to increased student achievement? In M. C. Wang & J. D. Finn (Eds.), *How small classes help teachers do their best*. (pp. 3–24). Philadelphia, PA: Temple University Center for Research in Human Development and Education.

Angrist, J. D., & Lavy, V. (1999). Using Maimonides' rule to estimate the effect of class size on scholastic achievement. *Quarterly Journal of Economics*, *114*, 533–575.

Beckerman, T. M., & Good T. L. (1981). The classroom ratio of high- and low-aptitude students and its effect on achievement. *American Educational Research Journal*, *18*, 317–327.

Betts, J. R., & Shkolnik, J. L. (1999). The behavioral effects of variations in class size: The case of math teachers. *Educational Evaluation and Policy Analysis*, *21*, 193–213.

Blatchford, P., Bassett, P., & Brown, P. (2011). Examining the effect of class size on classroom engagement and teacher-pupil interaction: Differences in relation to prior pupil attainment and primary vs. secondary schools. *Learning and Instruction*, *21*, 715–730.

Blatchford, P., Edmonds, S., & Martin, C. (2003). Class size, pupil attentiveness and peer relations. *British Journal of Educational Psychology*, *73*, 15–36.

Blatchford, P., Moriarty, V., Edmonds, S., & Martin, C. (2002). Relationships between class size and teaching: A multi-method analysis of English infant schools. *American Educational Research Journal*, *39*, 101–132.

Blatchford, P., Russell, A., Bassett, P., Brown, P., & Martin, C. (2007). The effects of class size on the teaching of pupils aged 7–11 years. *School Effectiveness and School Improvement*, *18*, 147–172.

Bourke, S. (1986). How smaller is better: Some relationships between class size, teaching practices, and student achievement. *American Educational Research Journal*, *23*, 558–571.

Bressoux, P., Kramarz, F., & Prost, C. (2009). Teachers' training, class size, and students' outcomes: Learning from administrative forecasting mistakes. *Economic Journal*, *119*, 540–561.

Clotfelter, C. T., Ladd, H. F., & Vigdor, J. L. (2006). Teacher–student matching and the assessment of teacher effectiveness. *Journal of Human Resources*, *41*, 778–820.

Evertson, C. M., & Randolph, C. H. (1989). Teaching practices and class size: A new look at an old issue. *Peabody Journal of Education*, *67*, 85–105.

Finn, J. D., & Achilles, C. M. (1990). Answers and questions about class size: A statewide experiment. *American Educational Research Journal*, *27*, 557–577.

Finn, J. D., Gerber, S. B., Achilles, C. M., & Boyd-Zaharias, J. (2001). The enduring effects of small classes. *Teachers College Record*, *103*, 145–183.

Glass, G. V., Cahen, L. S., Smith, M. L., & Filby, N. N. (1982). *School class size: Research and policy*. Beverly Hills, CA: Sage.

Glass, G. V., & Smith, M. L. (1979). Meta-analysis of research on class size and achievement. *Educational Evaluation and Policy Analysis*, *1*, 2–16.

Goldhaber, D., & Anthony, E. (2007). Can teacher quality be effectively assessed? National board certification as a signal of effective teaching. *Review of Economics and Statistics*, *89*, 134–150.

Downloaded by [Virginia Tech Libraries] at 09:37 11 July 2013

Goldhaber, D. D., & Brewer, D. J. (1997). Why don't schools and teachers seem to matter? Assessing the impact of unobservables on educational productivity. *Journal of Human Resources*, *32*, 505–523.

Good, T., & Brophy, J. (1987). *Looking in classrooms*. New York, NY: Harper & Row.

Greenwald, R., Hedges, L. V., & Laine, R. D. (1996). The effect of school resources on student achievement. *Review of Educational Research*, *66*, 361–396.

Hanushek, E. A. (1986). The economics of schooling: Production and efficiency in public schools. *Journal of Economic Literature*, *24*, 1141–1177.

Hanushek, E. A. (1999). Some findings from an independent investigation of the Tennessee STAR experiment and from other investigations of class size effects. *Educational Evaluation and Policy Analysis*, *21*, 143–163.

Hoxby, C. M. (2000). The effects of class size on student achievement: New evidence from population variation. *Quarterly Journal of Economics*, *115*, 1239–1285.

Konstantopoulos, S. (2008). Do small classes reduce the achievement gap between low and high achievers? Evidence from Project STAR. *The Elementary School Journal*, *108*, 275–291.

Konstantopoulos S., & Chung, V. (2009). What are the long-term effects of small classes on the achievement gap? Evidence from the Lasting Benefits Study. *American Journal of Education*, *116*, 125–154.

Konstantopoulos, S., & Sun, M. (2012). Is the persistence of teacher effects in early grades larger for lower-performing students? *American Journal of Education, 118*, 309–339.

Krueger, A. B. (1999). Experimental estimates of education production functions. *Quarterly Journal of Economics*, *114*, 497–532.

Lee, V. E. (2000). Using hierarchical linear modeling to study social contexts: The case of school effects. *Educational Psychologist*, *35*, 125–141.

Milesi, C., & Gamoran, A. (2006). Effects of class size and instruction on kindergarten achievement. *Educational Evaluation and Policy Analysis*, *28*, 287–313.

Murnane, R. J., & Phillips, B. R. (1981). What do effective teachers of inner-city children have in common? *Social Science Research, 10*, 83–100.

Nye, B., Hedges, L. V., & Konstantopoulos, S. (2000). The effects of small classes on achievement: The results of the Tennessee class size experiment. *American Educational Research Journal*, *37*, 123–151.

Nye, B., Konstantopoulos, S., & Hedges, L.V. (2004). How large are teacher effects? *Educational Evaluation and Policy Analysis, 26*, 237–257.

Odden, A. (1990). Class size and student achievement: Research-based policy alternatives. *Educational Evaluation and Policy Analysis*, *12*, 213–227.

Pedersen, E., Faucher, T. A, & Eaton, W. W. (1978). A new perspective on the effects of first grade teachers on children's subsequent status. *Harvard Educational Review*, *48*, 1–31.

Pong, S. L., & Pallas, A. (2001). Class size and eighth-grade math achievement in the United States and abroad. *Educational Evaluation and Policy Analysis*, *23*, 251–273.

Rivkin, S. G., Hanushek, E. A., & Kain J. F. (2005). Teachers, schools, and academic achievement. *Econometrica*, *73*, 417–458.

Rowan, B., Correnti, R., & Miller, R. J. (2002). What large-scale, survey research tells us about teacher effects on student achievement: Insights from the Prospects Study of elementary schools. *Teachers College Record*, *104*, 1525–1567.

Smith, M. L., & Glass, G. V. (1980). Meta-analysis of research on class size and its relationship to attitudes and instruction. *American Educational Research Journal*, *17*, 419–433.

Snedecor, G. W., & Cochran, W. G. (1989). *Statistical methods*. Ames, IA: Iowa State University.

Zahorik, J. A. (1999). Reducing class size leads to individual instruction. *Educational Leadership*, *57*(1), 50–52.