EXPANDING SCHOOL ENROLLMENT BY
SUBSIDIZING PRIVATE SCHOOLS:
LESSONS FROM BOGOTÁ

Claudia Uribe
Richard J. Murnane
John B. Willett
Marie-Andrée Somers

Expanding School Enrollment by Subsidizing Private Schools: Lessons from Bogotá
Claudia Uribe, Richard J. Murnane, John B. Willett and Marie Andrée Somers
NBER Working Paper No. 11670
September 2005
JEL No. I2

ABSTRACT

Many countries use tax revenues to subsidize private schools. Whether these policies meet social objectives depends, in part, on the relative quality of education provided by the two types of schools. We use data on elementary school students and their teachers in Bogotáá, Colombia to examine difference in resource mixes and differences in the relative effectiveness of public and private schools. We find that, on average, the schools in the two sectors are equally effective. However, they produce education using very different resource combinations. Moreover, there are large differences in the effectiveness of schools in both sectors, especially in the private sector. The results of our analysis shed light on the quantity-quality tradeoff that governments in many developing countries face in deciding how to use scarce educational resources.

Claudia Uribe
Inter-American Development Bank
1300 New York Avenue NW
Washington, DC 20577
claudiau@iadb.org

Richard J. Murnane
Harvard Graduate School of Education
Appian Way
Cambridge, MA 02138
and NBER
richard_murnane@harvard.edu

John B. Willett
Harvard Graduate School of Education
Appian Way
Cambridge, MA 02138
john_willett@harvard.edu

Marie-Andrée Somers
Harvard Graduate School of Education
Appian Way
Cambridge, MA 02138
somersma@gse.harvard.edu

Expanding School Enrollment by Subsidizing Private Schools:

Lessons from Bogotá

I.   Introduction

In many countries, government policies that use tax revenues to pay all or part of the cost of educating children in private schools have blurred the distinction between public and private schools. For example, Chile and New Zealand have implemented educational voucher programs to stimulate competition among schools.[1] The government of the Netherlands pays the costs of educating children at schools run by religious organizations.[2] Recently, some countries have introduced policies of subsidizing private schools as a way of increasing supply.[3] For example, the city of Bogotá in Colombia subsidizes private schools that enroll low-income students, as part of a strategy to meet its commitment to universal access to primary schooling.

Underlying these private sector strategies is the assumption that participating private schools will provide a quality education to their students, and will do so without

---

[1]   Edward B. Fiske and Helen F. Ladd, *When schools compete: A cautionary tale* (Washington, DC: Brookings, 2000); Chang-Tai Hsieh and Miguel Urquiola, "When Schools Compete, How Do They Compete? An Assessment of Chile's Nationwide School Voucher Program," *National Bureau of Economic Research (NBER) Working Paper* 10008 (October 2003).

[2]   Estelle James, "Public Subsidies for Private and Public Education: The Dutch Case." In *Private Education: Studies in Choice and Public Policy,* ed. D.C. Levy (Oxford, UK: Oxford University Press, 1986).

[3]   Indeed, Estelle James ("Why Do Different Countries Choose a Different Public-Private Mix of Educational Services?" *Journal of Human Resources* 28 (Summer 1993): 571-592) finds that countries with the largest percentage of private school enrollments are those with excess demand for schooling.

reducing the quality of public schools.[4] Understanding whether this assumption is valid requires investigation of the manner in which public schools and private schools "produce" student achievement, especially because, in many settings, the student body composition, teacher characteristics, and class sizes of public schools and private schools differ markedly.[5] If these factors play a role in determining children's skills, then these differences may compromise the success of "private sector" strategies in providing a high-quality education to the children they are meant to help.

Understanding the effects of teacher quality, class size, and peer group composition on school success is relevant to the design of public policies regarding the rules under which private schools operate. For example, the larger the role of peer group composition in determining students' achievement, the more important it is to design public policies that prevent the concentration of low income children in particular schools.

Disentangling the relative contributions of class size, peer group, and teacher quality is not possible in typical data sets in which each teacher works with only one group of children. In this study, however, we make use of an unusual property of a dataset from Bogotá – some teachers teach math to more than one classroom of students – to estimate the roles of teacher quality, peer groups, and class size. Because students appear to be randomly assigned to these classrooms, we can identify the effects of class

---

[4]   There is some evidence that private schools are more cost-effective than private schools. For example, using data from Indonesia, Estelle James, Elizabeth M. King, and Ace Suryadi ("Finance, Management, and Costs of Public and Private Schools in Indonesia," *Economics of Education Review* 15 (October 1996): 387-398) find that the cost per student of achieving a given level of academic performance is lower in private schools than public schools.

size and peers by comparing student achievement across classrooms taught by the same teacher. In these analyses, we obtain unbiased estimates of the causal effects of these classroom characteristics on student mathematics achievement by fitting statistical models that include teacher fixed effects. Also, we implement a new methodology based on excess variance contrasts to estimate the impact of peer group composition on student mathematics achievement. Finally, we fit an additional series of statistical models in which classrooms are treated as random effects to examine the extent to which differences between the public and private sectors in teacher quality, class size, and peer composition explain differences in the mathematics achievement of students in the two sectors.

Our results show that teacher quality, peer groups, and class size all influence student achievement. We also show that the average attributes of public schools and private schools differ markedly in dimensions that affect student achievement. Finally we show that differences in student achievement among schools in the public sector, and especially among schools in the private sector, are much larger than the average difference between the achievement of children attending public schools and those attending private schools, net of family background influences. The implication for parents is that the decision to use public or private schools is much less important than the choice of school within either sector. The implication for governments is that the details of subsidy programs are extremely important.

Section 2 provides a brief summary of Bogotá's recent strategy of subsidizing

---

5   Patrick McEwan & Martin Carnoy, "The Effectiveness and Efficiency of Private Schools in Chile's Voucher System, *Educational Evaluation and Policy Analysis* 22 (Autumn 2000): 213-239.

private schools. Section 3 discusses previous research on the effects of class size, peers, and teacher characteristics, particularly in Latin America. We focus on studies that have used the most plausible research designs. Section 4 presents the data, sample, measures and analytical strategies that we use to answer our research questions. Section 5 presents our findings, and Section 6 concludes with a discussion of the implications of these findings for school policy in Bogotá, and in other countries that provide public funds to private schools.

II. The Bogotá Setting

Since the 1950s, Bogotá has experienced a six-fold growth in population as migrants have streamed into the city in search of better economic opportunities or a refuge from the armed conflict that has ravaged large parts of the country. Providing

access to schooling for the growing population has been a major challenge for the Bogotá government. In 1997 when a new mayoral administration came into office, it found that more that 140,000 children in the city from the lowest income level were not attending school. Of those students from low income families who were in school, close to one third were enrolled in private unsubsidized schools of dubious quality.[6]

Committed to providing access to publicly funded schooling for all school-aged children, the Department of Education (Secretaría de Educación del Districto—SED) has used a two-pronged strategy to meet the demand for primary schooling. One prong is expansion of the public sector, by increasing class sizes,[7] reassigning teachers from

---

[6]   SED, *Informe de Gestión : 1998-2000* (Bogotá D.C.: Secretaría de Educación, Alcaldía Mayor, 2000).
[7]   SED established a minimum of 35 students per class in grades 1-5, and of 40 in grades 6-11.

administrative posts to classrooms, and constructing or rehabilitating classrooms in run-down schools. A related policy is to give students from low income families preferential access to the limited number of places available in public schools.

The second prong is subsidizing private schools that enroll low income students. As a result, in the year 2000, 53 percent of students in the city were enrolled in the city's 2,900 private schools.[8] For-profit schools, typically started by individuals, constitute 90 percent of all private schools in the city. The second most common type are religious schools, which make up 8 percent of the total, and are usually either subsidized or "contracted" by the government to provide educational services in low income areas. Finally, cooperative schools are established by neighborhood organizations typically in areas of unmet demand. They derive funding from both tuitions and public subsidies.

Public and private school sectors in Bogotá differ in a number of ways, particularly with respect to student composition, teacher qualifications, and class size. Consequently, even though SED's strategy of subsidizing private schools may have increased access to primary schooling in the city, the schooling experiences of low-income students attending schools in the two sectors may be quite different.

Private schools in Bogotá can set their own student admission policies, which are mostly based on academic criteria and/or on families' ability to pay.[9] Although private

---

[8] In comparison, only 21 percent of all pre-university level students in the country were enrolled in private schools in the same year, from SED, *Boletin Estadístico. Versión Borrador* (Bogotá, Colombia: Secretaría de Educación Distrital, 2001).

[9] The yearly tuition fee for private schools in Bogotá in the 2002 school-year ranged from $30 to $2,400, with an average of $285 (SED, *Estadísticas Educativas, 1998-2003* (Bogotá, Colombia: Secretaría de Educación Distrital, 2003). This is considerably less than the $439 average expenditure per student in public schools. However, these figures are not completely comparable because the public sector expenditure figure includes the costs of school supplies, transportation, and nutritional supplement, while the private school tuitions do not. Also, many private schools impose additional fees for school

schools in Bogotá enroll students from the whole range of the socio-economic spectrum, on average they serve a more affluent segment of the population than do public schools. For example, 43 percent of the private schools in our sample primarily enroll students from the lowest two (of six) socio-economic strata[10] while 10 percent primarily enroll students from the highest three strata. In comparison, 75 percent of public schools in our sample primarily enroll students from the lowest two strata, and none primarily enrolls students from the upper income strata.

Public and private schools also differ in the characteristics of their teachers. Most public primary school teachers in Bogotá have four-year degrees in education and post-graduate training.[11] As a result of relatively low turnover (3 percent annually) among Bogotá public school teachers and restrictions on the opening of new teaching posts in the public sector, the public school teaching force is mostly middle-aged and experienced, and relatively costly.[12] In contrast, private school teachers in Bogotá have much less teaching experience and fewer formal educational credentials than public school teachers.[13]

There are three complementary explanations for these differences. First, as a

activities and expect parents to make "voluntary" donations (bonos).

[10] Bogotá has a six-level system of socio-economic stratification, with 6 as the highest and 1 as the lowest. Roughly, estrato 1 and 2 correspond to families earning up to two times the minimum wage (the 2003 minimum wage is approximately US$100/month). Estrato 4 and 5 families earn approximately between 5 to 10 times the minimum wage.

[11] Based on data from A. Sarmiento, B. L. Caro, J. I. González, E. Castaño, E., and J. Espinosa, *Evaluación de la Calidad de la Educación Primaria en Santa Fé de Bogotá 1998, Factores Asociados al Logro, Informe Final Versión Definitiva* (Santa Fé de Bogotá: Departamento Nacional de Planeación, Secretaría de Educación de Bogotá, Corporación Mixta para el Desarrollo de la Educación Básica, 1999).

[12] New regulations that came into effect in 2002 introduce performance accountability mechanisms for promotions and reduce the weight of academic credentials and experience in determining promotions. These changes are likely to bring about changes in the city's teaching force over the coming years.

[13] Sarmiento et al.; Alejandro Gaviria and Jorge Hugo Barrientos, "Determinantes de la Calidad de la

result of a 1995 constitutional court ruling, private schools are required to pay teachers according to the public teacher salary scale that rewards experience and educational credentials.[14] However, private schools have greater flexibility than public schools in hiring teachers with modest credentials and in dismissing teachers at will. Consequently private schools have more control of personnel costs than do public schools. Second, in recent years a weak economy has made it difficult for many private schools to collect tuition payments and has created great pressure to reduce operating costs. Third, government regulations restrict the rate of private school tuition increases to the rate of inflation.

There are a number of concerns about the impact that the recent reforms aimed at increasing access have had on the quality of education provided to students. One stems from the increase in class sizes in public sector schools. A second is that schools may be becoming more segregated by socioeconomic status as students from the lowest income groups are given preferential assignment to public and private subsidized schools, leaving middle class children to attend unsubsidized private schools. A third is that as the city increases its subsidies program, some private schools accustomed to serving a relatively affluent population are increasingly serving children from low income families. Serving these children poses new challenges that are difficult to meet, especially by schools that contain costs by hiring as teachers adults with little pedagogical training or teaching experience.

If class size, peer groups, and teacher qualifications affect student achievement,

---

Educación en Colombia," *Planeación y Desarrollo* 32 (Autumn 2001): 339-386.
[14] Article 198, 1991 Constitution.

these reforms may jeopardize quality in both sectors. This provides our motivation for examining the effects that these factors have on student achievement and the roles they play in explaining the relative effectiveness of public and private schools. These issues are important not only to Bogotá, but also to other countries and cities that subsidize their private schools.

Before describing our approach to answering these questions, we review research from other countries. We limit our review to studies that have employed compelling strategies for estimating the unbiased causal impacts of class size, peers, and teacher quality on student achievement, and the relative effectiveness of public and private schools.

III.   Prior Research

A substantial literature examines the effects of class size, peer effects, and teacher quality on student achievement, both in the US and in developing countries.[15] Most studies are problematic, however, because they lack compelling strategies to account for the non-random assignment of teachers and students to classrooms. For instance, most studies of the effects of class size on student achievement do not adequately account for the assignment of weaker students (or stronger students) to the smallest classes. As a result the estimates of class size effects are likely to be biased. Similarly, it is likely that highly motivated parents choose their child's school based on the characteristics of its students. Consequently, in many studies, estimates of peer group effects may be biased because researchers do not adequately control for the unobserved influences of parental

8

motivation on children's achievement. Following this same logic, if more able students

(or less able students) are assigned to better teachers, then the application of conventional

methods will result in biased estimates of the effects of teacher characteristics on student

achievement.

It follows that a powerful strategy for obtaining unbiased estimates of the causal

effects of class size, peers, and teacher characteristics on achievement is to randomly

assign teachers and students to classrooms. Randomized controlled trials are expensive,

however, and remain few in number.[16] The next best alternative is to learn from "natural

experiments" in which policies or policy changes create exogenous variation in the

assignment of students to classrooms.

Miguel Urquiola,[17] for example, estimates the effects of class size by exploiting

the results of a policy rule in Bolivia: when class enrollment size reaches 30 students, a

school can apply to the Ministry of Education for another teacher, such that in the

---

[15] For an older review of research in developing countries, see Bruce Fuller, "What School Factors Raise Achievement in the Third World?" *Review of Educational Research* 57 (Autumn 1987): 255-292.

[16] The best known class size experiment is Project STAR in Tennessee. See A.B. Krueger, "Experimental Estimates of Education Production Functions," *Quarterly Journal of Economics* 114 (May 1999): 497-532. In terms of peer effects, the *Moving to Opportunity* project randomly awards housing vouchers to residents of low-income neighborhoods in various cities in the US. For results from Boston, see Lawrence F. Katz, Jeffrey R. Kling, and Jeffrey B. Liebman, "Moving to Opportunity in Boston: Early Results from a Randomized Mobility Experiment*,*" *Quarterly Journal of Economics* 116 (May 2001): 607-654. The effects of peers in tertiary education have been evaluated using random assignments of students to college dorms or rooms within dorms (e.g. Bruce Sacerdote, "Peer Effects with Random Assignment: Results for Dartmouth Roommates," *Quarterly Journal of Economics* 116 (May 2001): 681-704.). The one random assignment experiment on teacher quality that we are aware of finds that Teach for America participants are somewhat more effective in teaching math (but not reading) to low-income children than the other teachers serving in these schools. See Paul Decker, Daniel Mayer, and Steven Glazerman, *The Effects of Teach for America on Students: Evidence from a National Evaluation* (Princeton, NJ: Mathematica Policy Research, June 2004). One other relevant random assignment study finds that tutors have a large positive impact on the achievement of low-income students (see Abhijit Banerjee, Shawn Cole, Esther Duflo, and Lee Linden, "Remedying Education: Evidence from Two Randomized Experiments in India," *MIT Working Paper* (September 2003).

following year there will be an additional classroom, and both will have fewer students. The characteristics of students in class sizes of 30 should be equal on average to the characteristics of students after the new classrooms are added, with the only difference being the number of students in the classroom. Consequently, the difference in student achievement before and after the application of the class size rule should provide an unbiased estimate of the class size effect. Using a regression-discontinuity research design, Urquiola finds that the mathematics achievement of third grade students in a class of 20 students is 0.24 to 0.45 standard deviations (SD) higher than the achievement of students in a class of 30 students. Using a similar analytic strategy in Israel, Joshua Angrist and Victor Lavy[18] find a comparable class size effect of 0.46 SD for 5th grade mathematics.[19] A remaining potential source of bias in these estimates, however, is that teachers hired to accommodate extra classrooms may be less effective. In this instance, the difference in achievement between the pre- and post-rule group of students may be biased downward by the lower average quality of teachers in the latter group. In other words, this design cannot control for the effect of unobserved teacher characteristics.

Another recent approach to identifying class size effects has been to use panel datasets containing measures of student achievement at several points in time and information on teachers and their classes in several school years. Such data permit estimation of class size effects by comparing the performances of students taught by the

[17] Miguel Urquiola, "Identifying Class Size Effects in Developing Countries: Evidence from Rural Schools in Bolivia," *World Bank Working Paper* 2711 (October 2000).

[18] Joshua D. Angrist and Victor Lavy, "Using Maimonides' Rule to Estimate the Effect of Class Size on Scholastic Achievement," *Quarterly Journal of Economics* 114 (May 1999): 533-575.

[19] Such policy rules have also been used in research from the US, but contrary to research from developing countries, the effect of class size is not statistically significant. See Caroline M. Hoxby,

same teacher in different school years. Studies using this methodology find smaller class size effects than those reported in studies using the regression discontinuity methodology.[20] One possible explanation concerns differences in setting. Panel studies of class size effects utilize datasets from the U.S, while studies using the regression discontinuity design utilized data from Bolivia and Israel.

Researchers have also used panel data on teachers and students to examine the impact of teaching experience on teaching performance. One striking finding from these studies is that the effectiveness of teachers increases markedly with experience during their first years on the job.[21]

As for peer group effects among elementary or secondary school students, to our knowledge only two studies make use of data from a random assignment experiment. Bryan Graham, using data from the Tennessee STAR experiment, reports that the predicted math achievement of a primary school student placed in a class in which the average academic ability of classmates was one standard deviation above the average value for classrooms in the sample is 0.15 standard deviations higher than the predicted math achievement of the same student placed in a class with peers of average academic

---

"The Effects of Class Size on Student Achievement: New evidence from Population Variation," *Quarterly Journal of Economics* 115 (November 2000): 1239-1285.

[20] By including dummy variables for every student and every teacher in the sample, studies using such datasets are able to eliminate all unobserved time-invariant student characteristics and teacher characteristics as sources of bias. Eric A. Hanushek, John F. Kain, and Stephen G. Rivkin ("Teachers, Schools, and Academic Achievement," *NBER Working Paper* 6691 (August 1998)) find that a 10-student difference in class size is associated with only a 0.04 to 0.06 SD difference in student achievement. Using panel data from New Jersey, Jonah E. Rockoff ("The Impact of Individual Teachers on Student Achievement: Evidence from Panel Data," *American Economic Review* 94 (May 2004): 247-252) finds that the effect of class size on mathematics achievement is not statistically significant.

[21] See Hanushek, Kain, and Rivkin; and Rockoff.

ability.[22]   In a study using this same dataset but a different methodology, Boozer and Cacciola report similar findings.[23]

To date, the best estimates of peer effects in the developing world have been obtained from regression models that use *school* fixed effects to account for unobserved differences among schools.   In this approach the effect of peers is estimated by comparing achievement differences among students in different classrooms within the same school. Applying this strategy to data from Chile, Patrick McEwan[24] finds that students in an eighth grade classroom in which the average level of mother's education of classmates is one standard deviation above the average for classrooms in the sample score 0.287 SD higher on a test of mathematics achievement than do students in a class in which the average level of peers' mother's education is at the sample average. This estimate may still be biased by unobserved teacher characteristics, however, unless students are randomly allocated to classrooms within schools, which does not seem be true for the Chilean dataset.[25]

In summary, class size effects have been relatively well estimated in developing countries.   However, reliable estimates of peer effects and teacher quality effects on

---

[22]   Bryan Graham, *Using Conditional Excess Variance Contrasts to Identify Social Interactions: Theory and an Application on the Relationship Between Peer Groups and Academic Achievement* (Cambridge, MA: Harvard University Department of Economics, unpublished manuscript, September 2004).

[23]   Michael A. Boozer and Stephen E. Cacciola, "Inside the 'BlackBox' of Project Star: Estimation of Peer Effects Using Experimental Data," *Economic Growth Center Discussion Paper* 832 (June 2001).

[24]   Patrick J. McEwan, "Peer Effects on Student Achievement: Evidence from Chile," *Economics of Education Review* 22 (April 2003): 131-141.

[25]   McEwan, "Peer Effects on Student Achievement", also estimates the effects of peer composition by comparing the test scores of siblings and twins who are assigned to different classrooms, i.e. by using siblings/twins fixed effects. He finds that a one standard deviation difference in the mean mother's education of one's classroom peers is associated with a difference in mathematics achievement of 0.371 and 0.536 SD, for the sample of siblings and twins respectively. These estimates will be biased, however, if siblings differ in ability, and/or students are non-randomly assigned to classrooms.

student achievement in developing countries are still unavailable. Given the slow progress in understanding roles that different resources play in determining why some schools are more effective than others, it is not surprising that progress has also been slow in understanding the relative effectiveness of public schools and private schools, and the differences between them.[26] In most studies, the private school effect is defined as the difference in achievement between public and private schools, net of the pre-existing characteristics of their students. This effect is estimated by regressing achievement on a private school indicator and an array of student background variables. Research from Colombia using this methodology finds a private school advantage in the range 0.20 to 0.29 SD for secondary school achievement.[27] However, these models fail to control for important student and school characteristics such as parents' expectations and school admissions policies, which may result in selection bias in OLS estimates. Attempts have been made to correct this bias using standard Heckman selectivity bias corrections, but their application fails to meet critical assumptions.[28]

More reliable estimates of the relative effectiveness of private schools come from randomized voucher experiments in which a "treatment" group of students is given vouchers to attend a private school, and their post-treatment achievement is compared to

---

[26]   For reviews of the evidence, see Patrick J. McEwan, "The potential impact of large-scale voucher programs," *Review of Educational Research* 70 (Summer 2000): 103-149; and Richard J. Murnane, "The Role of Markets in American K-12 Education," in *The Limits of Market Organization*, ed.,Richard R. Nelson (New York: Russell Sage Foundation, 2005). For evidence pertaining to Latin America, see Marie-Andrée Somers, Patrick J. McEwan, and J. Douglas Willms, "How Effective Are Private Schools in Latin America?" *Comparative Education Review* 48 (February 2004).

[27]   George Psacharopoulos, "Public vs. Private Schools in Developing Countries: Evidence from Colombia and Tanzania," *International Journal of Educational Development* 7 (January 1987): 59-67; Donald Cox and Emmanuel Jimenez, "The Relative Effectiveness of Private and Public Schools: Evidence from Two Developing Countries," *Journal of Development Economics* 34 (November 1991): 99-121.

that of students who were included in the lottery but did not receive a voucher. Most relevant to this paper is a random assignment experiment conducted for secondary school students in Colombia that was evaluated by Joshua Angrist et al.[29] These authors report a private school advantage of 0.29 SD.

Although estimates from randomized experiments are clearly informative, they do suffer from the limitation of treating schools as black boxes. As a result they do not shed light on the reasons why private schools are more effective than public schools. This may matter, especially if peer group composition is an important determinant of student achievement. A combination of self-selection by parents and the operation of private school admission policies may create academically stronger peer groups that enhance the effectiveness of private schools. However, if this is an important part of the explanation for the relative effectiveness of private schools, public policies designed to increase student enrollments in private schools may result in a reduction in their effectiveness. Somers, McEwan and Willms[30] present evidence indicating that the private school effect for 3rd and 4th grade language and mathematics achievement is no longer statistically significant after controlling for peer group socioeconomic status (SES). In Colombia,

---

[28] See Somers, McEwan, and Willms for further discussion.

[29] Joshua Angrist, Eric Bettinger, Erik Bloom, Elizabeth King, and Michael Kremer, "Vouchers for Private Schooling in Colombia: Evidence from a Randomized Natural Experiment," *American Economic Review* 92 (December 2002): 1535-1558. For experiments in the US, see W.G. Howell, P.J. Wolf, D.E. Campbell, and P.E. Peterson, "School Vouchers and Academic performance: Results from Three Randomized Field Trials," *Journal of Policy Analysis and Management, 21* (Spring 2002): 91-217; Cecilia, E. Rouse, "Private School Vouchers and Student Achievement: An Evaluation of the Milwaukee Parental Choice Program," *Quarterly Journal of Economics* 113 (May 1998): 553-602; and Eric Bettinger, *The Effects of Educational Vouchers on Academic and Non-academic Outcomes: Using Experimental Economic Methods to Study a Randomized Natural Experiment* (Unpublished manuscript, 2003).

[30] See Somers, McEwan, and Willms for further discussion.

more notably, the estimated private school effect is 0.32 SD, but is reduced to 0.02 SD when controls for peer SES are included in the model. Unfortunately, estimated effects from the Somers, MacEwan, and Willms study and other studies are likely to be biased by the omission of important student and school characteristics – hence the need for new research.

One other limitation of the literature on private school effects is that it does not explore the extent to which there is variation in quality among schools *within* each sector. This is critical because the greater the variation in quality among schools in a sector relative to the difference in average quality between sectors, the less important is the decision about which sector in which to enroll a child, and the more important is the decision about which *school* within a sector in which to enroll a child.

The effects of class size, peer group composition, and teacher characteristics on student achievement, and the roles of these resources in explaining differences in the relative effectiveness of public and private schools are all relevant to understanding the impacts of Bogotá's recent policies. Our approach to investigating these effects is based on a special property of our dataset that allows us to address some of the limitations of prior work.

IV. Research Design

Dataset

To address our research questions, we designed a two-wave study of 5th grade classrooms in public and private schools in Bogotá, Colombia. Beginning with the

population of schools serving fifth grade students in the city, we stratified schools by sector (public & private) and by school socio-economic status (SES).[31] We then selected schools from these strata randomly. Figure A1 in the appendix contains a flowchart that describes how we arrived at the final analytic sample. In 1999, out of the 102 schools selected for the study, 14 refused to participate. One year later, we contacted the 88 remaining schools for follow-up. On this occasion, 10 schools refused to participate, one school did not offer 5th grade, and in four schools the 5th grade mathematics teacher did not answer the questionnaire.

Of the 3,095 5th grade students tested in 2000, 714 were either "new" to the school or had been absent from school the previous year on the day the fourth grade tests were administered. For these students, we imputed a pretest score using a hot-decking procedure.[32,33] The final sample includes 3,095 students from 73 schools, grouped in 97 5th grade classrooms and taught by 77 mathematics teachers. Of these students, 1,660 (54 percent) were enrolled in private schools and 1,435 students (46 percent) in public

---

[31] For administrative purposes, Bogotá is stratified into 6 distinct socio-economic strata with 6 being the highest (see footnote 10). Schools in Bogotá are highly segregated by socio-economic level.

[32] The personal and family characteristics used in this procedure included student's age, gender, number of days absent from school, mother's education, mother head of household, family's ownership of an encyclopedia, computer, car, refrigerator, washing machine, vacuum cleaner, phone, and having access to electricity in the home. All of the latter have been used as indicators of the student's socio-economic level. We also fitted our final models on the smaller sample of students for whom we had complete data. The results were not substantively different from those reported in the paper.

[33] In order to evaluate the generalizability of our study's findings to the population from which our sample was drawn, we assessed whether the schools and students that did not respond in the second wave of data collection differed in any systematic manner from those that did. On average, non-matched students are more likely to be male, to have obtained lower scores on the tests, to come from single mother households and to have been absent from school more often. Although proportionally more public schools dropped out of the study than did private schools, we found no statistically significant differences in school average pre-test scores or school SES between public schools that participated both years and those that dropped out of the study after the first year. However, among private schools, those schools that participated in both years had, on average, higher pre-test scores and student SES than those schools that dropped out after the first year.

schools. Our sample includes 37 for-profit schools, nine religious schools, and one cooperative school.

   A particular characteristic of this dataset, which drives an important part of our analysis, is that of the 77 teachers in the sample, 13 taught fifth grade mathematics to multiple groups of children. As we discuss below, this characteristic of the data allows us to estimate the impacts of class size and peer group quality on student achievement, controlling for unobserved teacher characteristics.

Measures

Our data analyses involve three types of variables (outcome variables, question predictors and control predictors) measured at four different levels (the student, the classroom, the teacher, and the school). Table A1 in the appendix lists the variables and provides brief descriptions.

Student-Level Outcome and Control Predictors

Table 1 presents the means and standard deviations of the variables included in our data-analysis for the full sample of students and separately for students in public schools and private schools. We discuss differences between the sectors later in this section.

Outcome Variable

Our outcome variable is each student's score (MSCORE) on an IRT-scaled mathematics test developed by the United Nations Educational, Scientific and Cultural

Organization (UNESCO) and the Oficina Regional de Educación de la UNESCO para América Latina y el Caribe (OREALC),[34] and administered to students in the final weeks of their 5th grade school year.[35] The average mathematics score across all students in the sample was almost 300 points, with a standard deviation of about 40.

Control Predictors

We included as control predictors the average of each student's scores on achievement tests in mathematics and language taken at the end of the fourth grade (PRETEST) and its square.[36] In doing this, we assume that the pretest scores capture students' total educational experience up to that point in time, as well as the effects of other earlier student background factors that may be influencing their achievement. Note that the pre-test is scaled differently than the post-test, and as a result we were unable to

---

[34] This test was developed in 1997 for an international comparative study of 3rd and 4th graders organized by UNESCO's Laboratorio Latinoamericano de la Calidad de la Educación (LLECE). Thirteen Latin American countries, including Colombia, participated in this study. Curriculum and testing experts from all participating countries, with the help of the Educational Testing Service (ETS), developed and pilot tested the instruments to make them reflect the curriculum of the first four years of primary education in the participating countries.

[35] This test has two parallel forms (Form A and Form B), each containing 32 questions. We used both forms of this test in the study, randomly distributing each form to approximately half the students in each classroom. We included in all models a dichotomous predictor to distinguish the form of the fifth grade test taken by each student. The reported Cronbach's alpha reliability for Form A of the test is .86 and for Form B is .88, as reported in United Nations Educational, Scientific, and Cultural Organization (UNESCO), *Primer Estudio Internacional Comparativo sobre Lenguaje, Matemática y Factores Asociados en Tercero y Cuarto Grado: Informe Preparado por el Laboratorio de Evaluación de la Calidad de la Educación* (Santiago, Chile: UNESCO/OREALC, 1998), p.67.

[36] The 4th grade tests used in this study were developed by Universidad Nacional for Bogotá's Department of Education 1998 student assessment program where they were used to test 3rd graders. We verified that the substantive results were not different when the models were fit with both the 4th grade reading score and math score included *separately* as control variables. Since the fourth grade tests were different from the fifth grade tests and it was not possible to create a meaningful common metric, it would have been inappropriate to define the outcome variable as the difference between the fourth and fifth grade math score.

use a change score as our outcome variable.[37]

We included two student demographic characteristics as control predictors: (a) AGE, the student's chronological age, measured in years, and (b) FEMALE, student gender, measured as a dichotomous variable (coded 1 for female, and 0 for male). Notice, in Table 1, that of the 3,095 students in the sample, 57 percent are girls. This is due to the inclusion of 19 all-female classes (and two all-male classes) in the sample of 97 classes. Also, boys were more likely to change schools and to be absent from school, preventing us from matching them on the two years of this study. Students in the sample are on average 10.8 years old, and their ages range from eight to seventeen.

We used two sets of predictors as indicators of student SES: (a) MHH, a dichotomous indicator of whether the mother was head of the household (coded 0=no, 1=yes), and (b) indicators of the mother's educational attainment (MOTHED).[38,39] Approximately one quarter of the students in the sample lived in mother-headed households. On average (median and mode) mothers had completed some high school.

---

[37] The pre-tests had 3 parallel forms, each containing 17 questions on mathematics and 17 on language. We administered all three forms of the pretest, randomly distributing the forms to approximately one third of the students in each classroom. Items were scored separately by subject, on a scale ranging from 0 to 306. Students achieved an average of 200.2 points (st. dev. 50) with scores ranging from 12 to 306 points (6 students in the sample obtained perfect scores). There are no published reliability or validity estimates for the pretests.

[38] In our sample, 356 observations (9.9 percent of the student sample) were missing information on mother's education. We imputed these missing values by substituting the median value for the student's classmates (we imputed the median rather than the mean because we did not want the imputed values to be sensitive to outliers). Likewise, 178 observations (4.9 percent of the student sample) were missing data on whether their mother was head of the household (MHH). Given that we constructed this variable based on students' responses to whether their father lived with them or not, in the cases where they failed to answer we assumed that their father did not.

[39] We chose to use mother's education rather than father's education for two reasons. First, there are a number of studies that suggest that mother's education is more important than father's education in determining IQ, e.g. Arleen Leibowitz, "Home Investments in Children," *Journal of Political Economy* 82 (March 1974): S111-S131. Second, there was a high percentage of missing data on father's education in our sample.

We also included as control predictors measures of the student's school attendance, the availability of books in the home, and the support provided at home for the completion of homework, which previous studies in Bogotá had found to be related to student achievement.[40] We measured these variables, as follows:

- DAYABS, the numbers of days that the student was absent from school over the previous two weeks.[41]

- HELPHW, the frequency with which a student receives help with his or her homework at home, measured on a Likert-type scale (where 1=never receives help, 2=sometimes receives help, 3=always receives help). In the data analyses, we converted this ordinal variable into a system of dichotomous (dummy) predictors, with HWHELP3 (always receives help) being omitted from the fitted models as a reference category.[42]

- BOOKS, measuring the availability of books in the student's home, a composite variable standardized to a mean of 0 and standard deviation of 1.[43]

In spite of the low levels of mother's education among students in the sample, 97 percent of students reported having books in their homes and 93 percent reported receiving help sometimes or always with their homework at home. Students in the sample were absent from school an average of 0.65 days in the two weeks preceding completion of the questionnaire.

---

[40] Sarmiento et al., "Evaluación de la Calidad".
[41] In the 77 cases (2.8 percent of the student sample) in which information was missing on this variable, we substituted the sample mean for the missing value.
[42] In the 19 cases (0.7 percent of the student sample) in which information was missing for this variable, we substituted the value 2, which was the modal response.

Question Predictors

To answer our research questions on the influence of classrooms, teachers and peers on student achievement, we included two classroom-level question predictors and two teacher-level predictors in the data-analyses.

Classroom Characteristics

The predictor that we used to represent peer-group composition is the "leave-out" mean[44] of the student's classmates' mothers' educational level (CLMOTHED). This is defined as the classroom mean of mother's education, leaving out each student's own value from the calculation. As can be seen in Table 1, students in the sample are likely to have a peer-group whose average level of mother's education is an incomplete secondary school education, but there is considerable variation across classrooms.

We measure class size as the number of students enrolled in each classroom as reported by the teacher. On average, classrooms in the sample have 37 students (std. dev. 8.5) although the range extends from 10 to 53 students.[45] Based on exploratory analyses that revealed a non-linear relationship between class size and student achievement, we defined our class size predictor (CLASSIZE) as the natural logarithm of the sample maximum class size (53) minus the student's own class size, or ln[53 - class size +

---

[43] For the 15 cases (0.5 percent of the student sample) in which information was missing on this variable, we assumed that there were no books in the home.

[44] See Michael A. Boozer and Stephen E. Cacciola, "Inside the 'BlackBox' of Project Star: Estimation of Peer Effects Using Experimental Data," *Economic Growth Center Discussion Paper* 832 (June 2001).

[45] Averaged over students rather than classrooms, the mean is 36 students.

0.167].[46]   Note that we followed Mosteller's and Tukey's[47] advice and "started" the

transformation by adding a small amount (0.167) to the difference in order to avoid

infinities when the student's class size equaled the maximum.

Teacher Characteristics

We included as predictors two sets of teacher characteristics that are rewarded by

the existing salary structure. The first measures educational preparation to teach. The

original ordinal variable recorded the teacher's highest educational degree obtained

(coded: 1= Pedagogical high school; 2= Education technology (2- or 3-year post

secondary certificate); 3= A four-year university program in education; 4=Post graduate

courses, including Masters in education (2 year + thesis), or specialization (1 or 2 years

graduate level program); 5= Graduate from a five-year university-level program in a field

other than education). In our analyses, we converted this ordinal variable into a system

of dichotomous (dummy) predictors, with completion of a four-year university program

in education (the modal value in our analytic sample) omitted from the fitted models as a

reference category.

We measured teacher experience (TEAEXP) by the number of years that the 5th

---

[46]   We chose a log function over a quadratic function of class size for several reasons. First, a quadratic
functional form results in predicted values that first increase and then decrease very sharply with class
size. Based on previous research on class size, we deemed this pattern to be implausible. We
therefore chose a log functional form because it constrains achievement to be a monotonic function of
class size. Second, we could not reject the null hypothesis that the predicted values of student
achievement from models with the alternative functional forms for class size were not different, when
evaluated at either the 10th or 90th percentile of the class size distribution. Third, the model's fit was
slightly better with the log specification than with the quadratic specification.

[47]   Frederick Mosteller and John W. Tukey, *Data Analysis and Regression* (Boston, MA: Addison-
Wesley, 1977), ch.4. To verify the appropriateness of this recommendation, we estimated a non-linear
model of student achievement against ln[Max Class Size + k – Class size] in order to determine the

grade mathematics teacher had been in the teaching profession. Given that previous research suggests that the relationship between student achievement and teachers' years of experience is non-linear,[48] we included the natural logarithmic transformation of this variable (L_TEAEXP) as a predictor in the regression analyses.[49] On average, teachers in the sample have 16 years of teaching experience (std. dev. 9.8), while 11 percent have 3 years or less of experience.

School Sector

In order to examine the effect of school sector on student achievement, we included a dichotomous predictor (PRIVATE), coded 1 = Private, 0 = Public). Fifty four percent of the students in the sample attended private schools. A total of 59 private school classrooms located in 47 private schools are included in the sample.

Data-Analytic Strategy

The goal of our analysis is to investigate to what extent differences in class size, peers, and teacher quality account for differences between the effectiveness of public and private elementary schools in Bogotà. Before proceeding to a comparison of achievement across sectors, however, we address two baseline questions. The first is whether student achievement varies across classrooms, net of student characteristics and

value of k that maximized the model fit. This procedure did indeed return a value very close to the recommended value.

[48] Hanushek, Kain, and Rivkin; Richard J. Murnane and Barbara R. Phillips, "Learning by Doing, Vintage, and Selection: Three Pieces of the Puzzle of Relating Teaching Experience and Teaching Performance," *Economics of Education Review* 1 (Autumn 1981): 453-465; Rockoff.

[49] As in the case of class size, we chose the log functional form over the quadratic functional form because the pattern of predicted values resulting from the latter seemed unlikely, given previous research (such as that by Rivkin, Hanushek, and Kain). The difference in predicted values between the log and quadratic functional forms is not statistically significant at either the 10th or 90th percentile of experience. Moreover, model fit is virtually identical for the two functional forms.

prior achievement.   The second is whether achievement, net of background characteristics, varies among students in different classrooms taught by the same teacher. These questions are important because the focus of our analysis is on predictors that vary at the classroom or teacher level (e.g. class size, peer group, teacher characteristics). If student achievement, net of background characteristics, did not vary among classrooms, our research could not provide any results. Fortunately, as explained below, student achievement is higher in some classrooms than in others, and this is even true among classrooms taught by the same teacher.   Our second step is to investigate whether class size and peer characteristics explain differences in achievement among classrooms, irrespective of sector.   Finally, we estimate the extent to which class size, peer and teacher quality explain differences in achievement between the public and private sector.

In these analyses, to ensure that standard errors are estimated correctly, we fit all regression models using either fixed effects or random effects to account for the clustering of students within classrooms. The following section provides greater detail on each of the statistical models fitted in our analysis.

Research Question 1

On average, do children enrolled in some 5ᵗʰ grade classrooms (or assigned to some 5ᵗʰ grade teachers) have higher end-of-year mathematics achievement than children in other classrooms (or children assigned to other teachers), after controlling for student SES and end of 4ᵗʰ grade achievement?

Do children enrolled in some 5ᵗʰ grade Bogotá classrooms have higher end-of-

year mathematical achievement, on average, than children enrolled in other 5th grade

classrooms taught by the same teacher?

We address these "baseline" research questions using three complementary strategies. First, we use fixed-effects regression analysis to model the relationship between a student's 5th grade mathematics achievement score and a set of dummy predictors representing *classroom* membership (Model 1a) or *teacher* membership (Model 1b), controlling for a vector of student background characteristics. In these models, we allow all students in each classroom or assigned to each teacher to share a unique intercept parameter. Inspection of estimates of these intercepts allows us to assess whether student mathematics achievement differs by classrooms or by teachers, net of selected student background characteristics. Note that the structure of our data does not allow us to estimate both teacher and class effects at the same time. However, any differences in "fit" between the teacher and classroom models suggest the existence of classroom effects on student achievement that are unique from those of the teacher, an issue that is a focus of attention in this analysis. The hypothesized fixed-effects models are as follows:

$$(1a) \quad Y_{ij} = \sum_J \beta_{0j} D_j + \beta_1 S_{ij} + \varepsilon_{ij}$$

$$(1b) \quad Y_{ik} = \sum_k \beta_{0k} D_k + \beta_1 S_{ik} + \varepsilon_{ik}$$

The variables and parameters are defined as follows:

$Y_{ij/k}=$ 5th grade mathematics achievement of student i in classroom j or with

   teacher k.

$D_{j/k}=$ vector of classroom or teacher dummy predictors (one per classroom or

   teacher).

$S_{ij/k}=$ vector of grand-mean-centered individual-level student characteristics

   for student i in classroom j or with teacher k.

$\beta_{0j/k}=$ intercept ("fixed effect") for classroom j or teacher k (the average

   mathematics achievement of students in classroom j or with teacher k,

   adjusted for the included student control characteristics).

$\beta_1=$ parameter vector representing the effects of the student characteristics.

$\varepsilon_{ij/k}=$ residual for student i in classroom j or with teacher k.

We test the null hypothesis that all classroom fixed effects (intercept parameters β_{0j}) or teacher fixed effects (intercept parameters β_{0k}) are simultaneously equal. If we reject these null hypotheses, we can conclude, in Model 1a, that the fifth grade classroom (or, in the case of Model 1b, the fifth grade math teacher) to which the student was assigned influenced his or her end of 5th grade mathematics achievement score, net of included student control characteristics.

It is important to note, however, that the obtained fixed effects coefficients are estimates, and are therefore fallible (contain error). Consequently, the sample variance of the estimated fixed effects is inflated (biased upward), relative to the estimated variance of the *true* fixed effects. To estimate the variance of the *true* fixed effects, σ_μ^2, we

26

therefore adopt a second strategy, using random effects regression analysis:

$$(1a^*) \quad Y_{ij} = \beta_0 + \beta_1 S_{ij} + (u_j + \varepsilon_{ij})$$

$$(1b^*) \quad Y_{ik} = \beta_0 + \beta_1 S_{ik} + (u_k + \varepsilon_{ik})$$

where $\mu_{j/k}$ is the deviation of the average mathematics achievement of students in classroom j or with teacher k from the grand mean of achievement (the $\mu_{j/k}$ are also called the random-effects). The random effects approach represented in Models 1a* and 1b* embeds deviations of the mean achievement for teachers/classrooms into the error structure, rather than retaining them as fixed effects in the structural part of the model as Models 1a and 1b .

The advantage of using a random-effects regression approach is that it estimates the variance of the *true* classroom/teacher effects directly, rather than obtaining an error-inflated sample variance of estimated fixed effects. We can therefore use the random-effects approach to test the null hypothesis that the variance of the true fixed effects is zero (H$_0$: $\sigma_\mu^2 = 0$), using a Breusch-Pagan Lagrangian Multiplier Test.

The disadvantage of random effects regression, however, is that it does not eliminate bias in model coefficients (the β_1's) caused by the unobserved and non-random sorting of students to teachers, as does the fixed effects regression approach. Therefore, in our analyses, we use random effects regression to test whether there is statistically significant variation in teacher and classroom effects, but use the parameter estimates from fixed-effects regression analyses to address any subsequent research questions

pertaining to the impact of student, classroom and teacher characteristics on achievement.

Finally, in a third exercise, we return to fixed effects analysis. In order to examine whether there are classroom-specific characteristics that exert an influence above and beyond the impact of the teacher, we test the null hypothesis that all classroom coefficients β_{0j} *that pertain to classrooms taught by the same teacher* are simultaneously equal. If we can reject these null hypotheses, we can conclude that even among classrooms taught by the same teacher, average student achievement differs across classrooms, net of included student control characteristics.

While the above strategy can detect the presence of classroom or teacher effects, it does not reveal which specific classroom and teacher characteristics are responsible for the effects. Our next research questions examine this issue.

Research question 2

Do the student-body composition and number of students in the class in which a student spent the 5th grade influence his or her 5th grade mathematics achievement score?

To answer this question, we fit a set of regression models in which we add class size and peer group composition as predictors to the baseline teacher fixed-effects model (Model 1b). With this strategy, we examine whether differences in class size and peer groups among classes taught by the same teacher explain differences in student achievement. The regression model has the following form:

$$(2) \qquad Y_{ijk} = \sum_{K} \beta_{0k} D_k + \beta_1 S_{ijk} + \beta_2 C_{jk} + \varepsilon_{ijk}$$

28

Where additional variables and parameters are defined as follows:

Y_{ijk} = 5th grade mathematics achievement of student i in classroom j with teacher k.

D_k = vector of teacher dummy predictors (one per teacher).

S_{ijk} = vector of grand-mean-centered individual-level student characteristics for student i in classroom j with teacher k.

C_{jk} = vector of grand-mean-centered classroom and peer-group characteristics of classroom j with teacher k.

β_1 = parameter vector representing the effects of the student characteristics.

β_2 = parameter vector representing the effects of the classroom and peer-group characteristics.

ε_{ijk} = residual for student i in classroom j with teacher k.

Our strategy of including a set of dummy predictors to represent all observed and unobserved differences in the outcome among teachers (the fixed effects of teachers) resolves a problem that has plagued previous studies, namely, the non-random assignment of students to teachers. Once teacher fixed effects are included in the model, we are still able to estimate the impact of the classroom-level class-size and peer-group predictors on student achievement because our dataset contains some teachers who taught multiple classes, thereby providing classroom-level outcome variation that is not eliminated by the inclusion of the teacher fixed effects. However, our strategy will produce unbiased estimates of class size and peer group effects (β_2) only if students are

randomly assigned to different classes taught by the same teacher. To learn about assignment processes, we interviewed a teacher or administrator in each of the 13 schools in which individual teachers taught mathematics to more than one fifth grade class. Our respondents reported that in all but two cases, students were assigned to classes to obtain a mix of student backgrounds and abilities in each class and to roughly equalize class sizes. In two cases students with especially low or high skills were assigned to a particular class. We eliminated these two classrooms from the analytic sample.

To further verify that students taught by the same teacher were not assigned to classes on the basis of ability or socioeconomic status, we utilized a strategy proposed by McEwan,[50] in which "virtual" classrooms are created by randomly reassigning students into "new" classrooms. Comparing the distribution of class size and peer group composition in the "virtual classrooms" with those in the actual classrooms, we found almost no differences in the two distributions.[51] Of course, this means that there is only modest variation in class size and peer composition among classrooms taught by the same teacher, and as a result our tests of class size and peer group effects have little statistical power.

In summary, we believe that our method of making use of a unique property of our Bogotá dataset provides new insights into the roles that class size and peer group composition play in explaining the relative effectiveness of public and private schools. However, our empirical results are by no means definitive because they are based on

---

[50] Patrick J. McEwan, "Peer Effects on Student Achievement: Evidence from Chile," *Economics of Education Review* 22 (April 2003): 131-141.

[51] This simulation exercise allowed us to compare not only the *mean* characteristics of the classrooms taught by a given teacher, but also their variances.

estimates from student achievement in the modest number of classrooms (33) taught by
the 13 teachers who taught more than one fifth grade class.

Research Questions 3a, 3b, and 3c

*Do students enrolled in private schools have higher mathematics achievement
than students enrolled in public schools, after controlling for (a) student background
characteristics; (b) student background characteristic and peer group composition; and
(c) student background characteristics, peer group composition, and class size?*

To address question 3a we fit the random effects regression model below (Model
3). This model is similar to our Model 1a, except that the classroom-specific intercepts
(the classroom fixed effects) have been converted into the corresponding random effects,
u_j, in order that school-sector (PRIVATE) can be included as a predictor:

$$(3) \qquad Y_{ijk} = \beta_0 + \beta_1 S_{ijk} + \beta_3 PRIVATE_{jk} + u_j + \varepsilon_{ijk}$$

where:

$\beta_0 =$ intercept parameter representing the average student 5th grade mathematics
achievement (across all classrooms and teachers), controlling for all
predictors in the model.

$\beta_3 =$ effect of the school sector of classroom j with teacher k on student 5th grade
mathematics achievement, controlling for all other predictors in the model.

$u_j =$ random effect of classroom j on student 5th grade mathematics achievement.

To address question 3b, we add the predictor representing peer group composition to the model described above (to provide Model 3b); to address question 3c, we further add the class size predictor (to provide Model 3c).

Unlike the corresponding estimates obtained from our teacher fixed effects model (Model 2), it is possible that the estimated effect of school sector obtained in our classroom random effects model (Model 3) may be biased by the omission of unobserved student and classroom characteristics. However, we can reduce the magnitude of the bias in the school sector effect in Models 3b and 3c by *constraining the effects of the class size and peer group predictors to the estimates that we obtained in our teacher fixed effects specification (Model 2)*, and by controlling for students' pre-test scores. The logic to constraining these coefficients in this way is that, by virtue of its fixed effects specification, Model 2 takes into account all unobserved teacher quality differences with which class size and peer attributes may be correlated. [52]

Research Question 4:

Does a student's math achievement at the end of 5th grade depend on his or her 5th grade math teacher's qualifications and years of teaching experience?

To address this research question, we supplement the hypothesized random effects model in Model 3, above, as follows:

---

[52] We also tested whether including interactions between PRIVATE and the other predictors in the model substantially improved model fit. We could not reject the null hypothesis that the interaction terms are

32

$$(4) \quad Y_{ijk} = \beta_0 + \beta_1 S_{ijk} + \beta_2 C_{jk} + \beta_3 PRIVATE_{jk} + \beta_4 TEACHER_{jk} + u_j + \varepsilon_{ijk}$$

Where:

$\beta_4 =$ effect of a vector of teacher qualifications and teaching experience, controlling for vectors of student and classroom characteristics.

Again, and for the same reasons, we constrain the coefficients on the class size and peer group quality predictors to their values estimated from fitting Model 2.

V. Findings

Do Classrooms and Teachers Matter to Student Mathematics Achievement?

The 5th grade classroom and teacher to which a student was assigned makes a difference to his or her end of 5th grade mathematics achievement score. Based on tests of the equality of the fixed-effects across the 97 classrooms in Model 1a and across the 77 teachers in Model 1b, we can reject the null hypotheses that there are no differences among classrooms (F=4.30; p<.001) and no differences among teachers (F=4.77; p<.001).[53]

It is important to remember that the sample variance of the estimated teacher and classroom fixed-effects is inflated (biased upward) by the fallibility of the estimates. As explained earlier, one way of dealing with the fallibility of the estimated classroom/teacher fixed effects is to fit a random-effects model, and to estimate the

---

jointly equal to 0 (χ^2=11.91, p=0.4528, df=12, based on Model 3b). We therefore include only the main effect of PRIVATE in our reported analyses.

variance of the true teacher or classroom effects directly (Models 1a* and 1b*). After controlling for the effects of the usual student background characteristics, we find that the estimated variance of *true* classroom effects is 69.3, and statistically significant (p<0.001).[54]

Note that the estimated variance of the *estimated* classroom fixed-effects is 159.8, which is more than twice the size of the estimated *true* variance obtained directly from the random-effects model. This implies a high degree of error in the fixed-effects estimates, and indicates that the estimated reliability of our measurement of the classroom fixed-effects is less than 0.5 (69.4/159.8 = 0.434). The estimated inter-quartile range of the *true* classroom fixed-effects (after disattenuating for the fallibility of the estimates) is 9.8 points. This is about half of the difference between the median scores of 3rd graders and 4th graders when the same test was administered to large samples of students in these grades in Colombia in 1999.[55] Applying random-effects regression analysis to teachers rather than classrooms reveals that the estimated variance of true teacher effects is 51.5 (p<0.001),[56] which is less than half the estimated variance of the estimated teacher effects from the fixed effects analysis (139.5). In this case, the estimated inter-quartile spread in *true* fixed-effects is 9.2 points.

These findings suggest two complementary lessons. First, consistent with parents' beliefs, some teachers are much more effective than others in helping students learn mathematics. Second, consistent with the beliefs of many teachers, average student

---

[53] $F_{96,2986}=4.30$***.
[54] The χ^2 test statistic (1 degree of freedom) is 337.2.
[55] UNESCO, p.258.
[56] The χ^2 test statistic (1 degree of freedom) is 418.2.

test scores, adjusted for student background characteristics, are quite unreliable measures of teacher effectiveness.

Is It Teachers Or Classrooms That Matter?

To answer this question, we tested for the equality of the classroom fixed effects among groups of classrooms taught by the *same teacher* (after controlling for student background characteristics). We reject the global null hypothesis that there were no differences in the average achievement (net of student background characteristics) of students in different classes taught be the same teacher. [57] This is important because it means that it is possible, in subsequent analyses, for class size and peer group effects to make a difference in explaining why student achievement is higher in some classrooms than in others taught by the same teacher.

Before turning to the roles of specific classroom attributes in predicting student math achievement, we summarize briefly the impacts of those student-level characteristics that have the strongest impacts on student achievement within-classrooms (see Appendix Table A2, first column of estimates in top half). First, as many other studies have shown, a student's achievement score at the end of one grade (4th) is a strong predictor of the student's achievement at the end of the next grade (5th). Second, on average, boys outperform girls by an average of 8 points (p<.001). Third, younger students perform slightly better on average than older students (p<.001), a likely result of accumulated academic failure among the latter. Fourth, 5th grade mathematics achievement also depends upon school attendance. Students who were absent one

additional day in the two weeks before the math test was administered have scores that are 1.9 points less, on average (p<.001).

Do peer-group composition and class size matter to student mathematics achievement?

As other researchers have pointed out,[58] the estimation of peer effects and class size effects has often been biased as a result of failing to account for the non-random assignment of students to teachers. To minimize this problem, we rely on variation in peer groups among classes taught by the same teacher in the same school.

Columns 2 through 4 of Table 2 contain estimates of parameters from Models 2a, 2b, and 2c, which include fixed effects for teachers and different combinations of the two class-specific predictors, class size and peer group composition. As indicated in Models 2b and 2c in columns 3 and 4, the point estimates suggest that students in classrooms in which their peers have more educated mothers achieve higher mathematics scores. While the relevant coefficients in the teacher fixed effects models are not statistically significant, *they are similar in value* to the statistically significant coefficients in the analogous models fitted with random effects in columns 6 and 7 of Table 2.[59]

Based on the fitted fixed effects from Model 2b in column 3, a one SD difference

---

[57] F=7.1, p<0.001.

[58] Caroline M. Hoxby, "Peer Effects in the Classroom: Learning from Gender and Race Variation," *NBER Working Pape*r 7867 (August 2000); McEwan, "Peer Effects on Student Achievement"; Boozer and Cacciola.

[59] In essence, the parameter estimates from the fixed effects models are less biased than those from the analogous random effects models. On the other hand, because there are only 13 teachers in our sample who teach multiple classes, the standard errors of the estimates from fixed effects models will be larger than those from random-effect models, resulting in a loss of power when making statistical inferences. Fortunately, the estimated peer effects from the fixed effects and random effects models (in Table 2) are very similar, so we can conclude that the random effects model provides unbiased estimates after

in the average value of mothers' education is associated with a 0.150 std. dev. difference in individual students' math achievement. This is a little more than half the size of the analogous peer group effect of 0.287 SD estimated by McEwan[60] using data from Chile. The difference may stem from differences in the age of the students, or differences in model specification -- our model controls for unobserved teacher characteristics. Our estimate of the magnitude of the peer effect is almost identical to that found by Graham in his analysis of the Tennessee STAR experiment.[61]

Bryan Graham[62] provides a more powerful method of detecting the contribution of peer effects to achievement. His rationale begins with the observation that, if peer effects on student achievement do exist, there will be greater between-classroom variation in student achievement than one would expect in the absence of peer effects. In other words, "excess" between-group variation will be present. Of course, in a setting in which each teacher teaches only a single classroom, the presence of excess variance could simply reflect the impact of heterogeneity in teacher quality and not peer effects. But, in our sample, some teachers teach multiple classes of students and so we can use a variant of Graham's method to separate peer and teacher effects providing we assume

---

all. As such, we can use the (more precise) standard errors from the analogous random effects model to make statistical inferences.

[60] McEwan, "Peer Effects on Student Achievement". Given our identification strategy, it is possible that the "peer effect" actually reflects differences in the effectiveness with which the same teacher teaches groups of students with different characteristics.

[61] We also tried estimating peer effects using an alternate measure of peer composition: the percentage of students in classroom j with two parents. In the fixed effects specification (Model 2b), a 1 percent difference in the number of children from a two-parent family results in a 35.3 point difference in a classroom's average test score (and is statistically significant at the 5 percent level). Thus using this alternate measure also leads us to conclude that peer effects are statistically significant in our dataset.

[62] Bryan Graham, "Using Conditional Excess Variance Contrasts to Identify Social Interactions: Theory and an Application on the Relationship Between Peer Groups and Academic Achievement," Harvard University Department of Economics (September 2004).

that "teaching quality" is constant across classrooms taught by the same teacher.[63]   In this application, excess between-classroom variation *across classrooms taught by the same teacher* can be attributed to peer group effects. To implement Graham's approach, we therefore use data from the students of teachers with multiple classrooms to estimate the *actual* between-classroom variance in student achievement, as well that which would be *expected* were there no peer effects.[64] If the ratio of these two variances is greater than 1, then positive peer effects are present.   We find an excess variance ratio of 2.2, suggesting the presence of peer effects. However, due to the small sample size and the modest variation in peer composition among classrooms taught by the same teacher, we cannot reject the null hypothesis that the magnitude of the ratio is different from unity.[65]

Returning to Table 2, the estimates from Models 2a and 2c (in columns 2 and 4) indicate that class size is a statistically significant predictor of student achievement. (Figure A2 in the appendix presents this fitted relationship.) Based on this fitted model, students in a classroom with 30 students are predicted to score 0.35 standard deviations higher (13.5 points) than students in a classroom with 40 students. The magnitude of the impact is quite similar to that reported by Urquiola in Bolivia and Angrist and Lavy in Israel (who find that mathematics achievement is 0.24-0.45. and 0.46 of a standard deviation higher in a class that is 10 students smaller, respectively).

Notice in Table 2 that while class size is a strong predictor of student achievement in the models fitted with teacher fixed effects, this is not the case in the models fitted with

---

[63]   Although a reasonable assumption, it is nonetheless restrictive; for example, teachers may be more or less effective when teaching material for the first time, or in the morning, etc.

[64]   See Graham's paper for details of the estimation of the actual and expected variances.

[65] $F = 0.67$; df=1, 31; p=0.4193.

classroom random effects. The explanation is that, in the Bogotá schools included in our sample, the most effective teachers were assigned to work with the largest classes. Consequently, we argue that estimates of the class size effect from the random effects models – that do not fully control for differences in teacher quality – are biased.

One implication of this pattern is that estimates of the relative effectiveness of individual teachers taken from models that control for class size and peer group effects are quite different from estimates taken from models that do not control for these factors. For example, one teacher in our sample ranked 43rd in terms of average student achievement when only the student level characteristics were controlled; this teacher ranked first when both peer-group composition and class size were included as predictors. Figure 1 displays plots that present the sample distribution of the estimated coefficients on the teacher fixed-effects from fitted Model 1b, and Models 2b and 2c. Notice that variation in estimated teacher quality is much larger when estimated from fitted Model 2c, which controls for class size, than when derived from the fitted models that do not. The explanation is that more effective teachers were assigned to the largest classes in the data set. The net result is that differences in class size, when not taken into account, reduce the estimated variation in teacher quality. This finding raises serious questions about school accountability initiatives that attribute differences in average student achievement across classrooms solely to differences in teacher quality.

Does School Sector Matter to Student Achievement?

As shown in Table 1, students who attended private schools in our Bogotá dataset

are more economically advantaged, on average, than those students who attended public schools. Private school students are less likely to live in mother-headed families than are public school students. Their mothers have more education, on average, than the mothers of public school students. The differences in individual student characteristics between public and private schools naturally result in corresponding differences in the characteristics of peer-groups by sector. As a result, public school students have less advantaged peer groups than do private school students. Public school students are also likely to study in classrooms with larger class sizes than are private school students.

One respect in which public school students have a potential advantage over students in private schools is in the academic preparation and experience of their teachers. No public school teacher in the sample has less than six years of teaching experience in contrast to 28 percent of private school teachers. While almost 70 percent of public school teachers have a Master's Degree or are working towards one, the corresponding figure for private school teachers is 12 percent. As explained earlier, this pattern stems primarily from the public school teacher salary scale and its applicability to private schools. The salary schedule provides strong incentives for public school teachers to gain educational credentials and to remain in teaching. The requirement that private schools pay teachers according to this same schedule creates incentives for private schools to avoid highly experienced teachers with significant educational credentials because they are very expensive. One-third of private school teachers are on the first step of the salary scale. No public school teachers are on the first step and more than 95 percent are on the top half of the scale.

We turn now to the question of whether the differences in peer groups, class sizes, and teacher characteristics contribute to differences in the achievement of students attending schools in the two sectors. The average achievement difference between the two sectors can be found in Table 3, as indicated by the estimated coefficient of PRIVATE in the different specifications of Model 3. We also illustrate the variability in average student achievement among groups of students taught by different teachers in each sector by plotting the estimated teacher fixed effects from Models 1 and 2 by sector (see Figure 2).

Panel A of Figure 2 displays the distribution of average teacher-specific fifth grade student math scores in public schools and private schools in the sample. The mean of the distribution for students in private schools is 8 points higher than the mean for students in public schools (based on the estimated coefficient of PRIVATE for Model 3a* in Table 3, which does not include the student-level controls). This is not a surprise given the difference between the average socioeconomic status of private school students and public school students. What is striking is the wide variation in teacher-specific average achievement among teachers in both sectors. This highlights that private schools in the sample do not all serve high achieving students.

Panel B of Figure 2 displays the distribution of predicted average student achievement for different teachers, after controlling for student background characteristics including pre-score, by sector (i.e. the teacher fixed effects from Model 1b). The difference in the adjusted average achievement of students in the two sectors is 0.8 points (0.02 SD) and not statistically significantly different from zero, based on the

results of Model 3a in Table 3. This estimate is much smaller than the Colombian estimate from Somers, McEwan and Willms (0.32 SD), even though both estimates were obtained from a similar model, and from datasets of students of comparable age. It is also smaller than that reported by Angrist et al. in their evaluation of the voucher experiment in Colombia (0.29 SD). A plausible explanation is that the mix of private schools in the country as a whole is different from the mix of private schools in our Bogotá sample. (The statistically significant variation in predicted average student achievement among students taught by different private school teachers indicates how sensitive the size of the "private school effect" is to the composition of the private school sample.)

Panel C of Figure 2 displays the distribution of predicted average achievement of students taught by different teachers taken from fitted Model 2b, which controls for peer group as well as for student background, by sector. Net of student background and the peer group effect, the adjusted achievement of students in private school classes is four points *lower* than that of students in the public sector (or 0.10 SD), based on the results of Model 3b in Table 3. This highlights that a high SES peer group is one of the relative advantages many private schools offer. However, again it is striking how great the variation in average student achievement is among students enrolled in the classes of different private school teachers.

Finally, Panel D of Figure 2 displays the distribution of predicted average achievement of students taught by different teachers taken from fitted Model 2c, which controls for class size as well as for peer group and student background characteristics.

42

The adjusted mean achievement for private school students is now 14.5 points *lower* than the analogous mean for public school students, based on the results of Model 3c in Table 3. The explanation is that class size and peer groups matter to student achievement and, on average, class sizes are smaller and peer groups are more advantaged in the private sector.

Interpreting the estimated teacher fixed effects displayed in Panel D as our best estimates of teacher quality, the wide range of quality estimates for private school teachers is remarkable, especially relative to the magnitude of the variation in estimated effectiveness among public school teachers. The best private school teachers appear to be better than the best public school teachers, while the worst private school teachers appear to be considerably less effective than the worst public school teachers.[66] Unfortunately, we do not know whether the least effective private school teachers are more likely to be dismissed for poor performance than the weakest public school teachers.

These findings raise the question of whether there is a statistically significant difference in the *variation* of teacher quality between the public and private sectors. As previously discussed, the estimated variance of the estimated teacher fixed effects is inflated by the errors of estimation, such that we must once again use random effects regression analysis to better estimate the variance in *true* teacher effects. To achieve this, we fit the following regression model, in which we permit the variance of the teacher

---

[66] It is important to point out that we do not know for sure that the average differences in achievement among students in classrooms taught by different teachers represent teacher effects. Although we do control for class size and one measure of peer group composition, it is possible that there are also

random effects to differ by sector:

$$Y_{ijk} = \beta_0 + \beta_1 S_{ijk} + \beta_2 PRIVATE_{jk} + (\eta_k I_k + v_k D_k + \varepsilon_{ijk})$$

where η_k and v_k are sector-specific *teacher* random-effects, and I_k and D_k are indicator variables such that $I_k=1$ if a student's school is public, and $D_k=1$ if a student's school is private, otherwise they are set to zero. By comparing the -2*(log likelihood) statistic of this sector-heterogeneous teacher random effects model to that of a model with sector-homogeneous teacher random effects, we can test whether permitting a sector-specific error structure significantly contributes to model fit. This is equivalent to testing the null hypothesis that the population variances of teacher random effects in the public and private sectors are identical, $\sigma_\eta^2 = \sigma_v^2$. We find that the sector-specific variances of the teacher random effects are 64.8 in the public sector and almost double – 123.2 – in the private sector. Again, because of low statistical power, we cannot reject the null hypothesis that the variance of the true teacher random effects is the same across sectors.[67]

In conclusion, the panels of Figure 2 shed light on differences in the ways that public schools and private schools "produce" student achievement. As indicated in Panel B, on average, private schools in our Bogotá sample are as effective as public schools in

---

unmeasured classroom-specific differences that account for some of the classroom-specific achievement differences.

[67] $\chi^2=2.1$, df=1, p=0.147.

producing student achievement. However, schools in the two sectors use very different mixes of resources. On average, private schools have less skilled teachers than public schools, but they compensate by having smaller classes and somewhat more advantaged student peers.

Do Teacher Qualifications and Experience matter to Student Achievement?

The difference between the two sectors in the average quality of teachers and the extraordinary variation in teacher quality among teachers in the private sector prompts the question: do teaching experience and educational credentials, the attributes rewarded in the national teachers' contract, also predict teaching performance?

To answer this question, we fitted the set of regression models presented in Table 4 that estimate the effects of teacher education and years of experience on student mathematics achievement, with classrooms treated as a random effect. An important disadvantage of using a random effects specification, however, is that the estimated effects of class size and peer composition will be biased, given that these variables may be correlated with unobserved teacher quality. In estimating Models 4a and 4b, we therefore constrain the parameters associated with the peer group and class size predictors to their values estimated in the fixed effects analysis (Model 2c), which are *not* biased by unobserved teacher quality. For comparative purposes, we display the results from fitting these models with the peer group and class size parameters left unconstrained (Models 4c and 4d).

Figure 3 illustrates the fitted relationship between student achievement and teacher experience from Model 4b (which does not control for sector). Notice that the

slope of this relationship is quite steep during a teacher's first years in the classroom, after which the fitted relationship becomes flatter. This supports the findings from other recent studies which suggest that the greatest gains from additional experience occur during the first years of teaching.[68] Notice, however, that in Models 4c and 4d the coefficient on this predictor is close to zero, suggesting that more experienced teachers appear to be assigned to larger classes and the most needy groups of students, thus exerting a compensatory effect that neutralizes the effects of experience.

Notice in Models 4a and 4b that teachers who have less than a four-year college education are less effective, on average, than those with a four year degree. Nine percent of the teachers in private schools in the sample have less than a four-year degree compared to two percent of public school teachers. Thus, differences in educational credentials and years of teaching experience help to explain the difference between the average quality of private school teachers and those of public school teachers.

It is also important to point out that the estimates for Models 4a and 4b indicate that teachers with educational credentials beyond a four-year degree (EDLEVEL4) are not more effective, on average, than are those with a four-year degree. This supports the view of many Bogotá school officials that a salary schedule that rewards post-graduate educational credentials is not an efficient use of scarce resources.

A final pattern to notice is that the percentage of the variation in student achievement explained by Model 4b (20.6) – a random effects model that includes predictors for teacher experience and educational credentials – is much lower than the

---

[68]   Other studies that show this same pattern include Hanushek, Kain and Rivkin; Murnane and Phillips; and Rockoff.

percentage of variation explained by Model 2c (35.1), a model that includes fixed effects for teachers. The explanation is that, while information on educational credentials and years of teaching experience do indeed play a role in predicting teacher quality, they explain only a modest part of the variation in effectiveness among teachers in enhancing student achievement.

VI.   Lessons

We conclude by describing four lessons about the consequences of initiatives by the Bogotá government to greatly expand access to elementary school education. First, it is likely that expanding access led to a decline in the average quality of education. One mechanism through which this took place is that class sizes in public school classrooms increased. Another is that private schools hired teachers with relatively low skills. The likely decline in average quality does not mean that expanding access was a mistake – only that, given severe resource constraints, it was not accomplished without cost.

The second lesson is that the government's two-pronged strategy for expanding access seems quite efficient, given the assumption that it was unable politically to change the structure of public school teachers' compensation schedule or prevent its extension to private schools. Public schools, with very experienced and highly educated teachers, were told to serve more students. They responded by having the most skilled teachers teach the largest classes. It was probably less expensive to subsidize private schools that served low-income students than it was to educate the children in public schools. While schools in the two sectors produced achievement with quite different input mixes, on average, they were equally effective in producing student achievement.

The third lesson concerns the teacher salary schedule. Our evidence indicates that teachers with education beyond a four-year college degree were not more effective in producing student achievement than teachers with a four-year degree were. Thus, unless it could be demonstrated that the most highly educated teachers made contributions greater than other teachers in respects not measured by test scores, challenging the structure of the salary scale may be a way to improve efficiency.

Finally, the importance of peer groups in influencing student achievement creates a difficult public policy problem. Peer group effects create strong incentives for parents with resources to try to place their children in schools with advantaged peers and for schools interested in attracting students from such families to specialize in serving such families. The resulting sorting by socioeconomic status widens inequality in access to good schooling. Evidence from New Zealand and Chile demonstrates that such sorting by socioeconomic status does take place in large scale voucher plans in which the value of the voucher is the same for all students.[69]

One way to reduce the likelihood of sorting by socioeconomic status is to use differentially valued vouchers that create incentives for schools to attract students whom they would otherwise spurn.[70] This idea has a long history.[71] To date, there has been no significant trial of a school choice system with differentially valued vouchers.

The value of vouchers is only one of many critical details in the design of a voucher system. Others concern whether schools can charge families more than the value

---

[69] On New Zealand, see Fiske and Ladd ; on Chile, see Chang-Tai Hsieh and Miguel Urquiola.
[70] Caroline M. Hoxby, *Ideal Vouchers* (Cambridge, MA: Harvard University unpublished manuscript, 2001); Dennis Epple and Richard Romano, "Education Vouchers and Cream-Skimming," *NBER Working Paper* 9354 (November 2002).

of a child's voucher, the quality of information available on school performance, the process through which spaces are allocated in schools for which demand exceeds supply, rules about dismissing difficult students, and the availability of free transportation. By influencing the options available to low income families and the attractiveness of participation to potential suppliers of school services, these details are likely to influence how an educational voucher system would influence the quality of education provided to children from different backgrounds.

Perhaps the most important lesson from this paper for the debate over school vouchers is that the concept of unregulated markets is of little value in predicting outcomes. The reason is that, in the language of economics, peer group effects are externalities, and unregulated, competitive markets do not produce an efficient allocation of resources when externalities are present. Any school choice system will be a system with regulations. It is important to pay attention to how, in the face of peer group effects, parents choosing schools and entrepreneurs with different motivations for starting schools respond to the detailed regulations that define a choice system. These responses will determine how the system influences the quality of education provided to children from different backgrounds.

---

[71] See the discussion in Murnane, "The role of markets in American K-12".

Table 1
Descriptive Statistics on Outcome and Predictors, by Sector

VARIABLE	FULL SAMPLE n students =3,095 n classrooms=97 n teachers=77 n schools=73	PUBLIC n students=1,435 n classrooms=38 n teachers=29 n schools=26	PRIVATE n students=1,660 n classrooms=59 n teachers=48 n schools=47
OUTCOME VARIABLE			
MSCORE	297.6 (39.1)	292.3 (38.1)	302.2 (39.4)
STUDENT-LEVEL CONTROL PREDICTORS			
PRETEST	200.2 (50.1)	191.2 (49.4)	207.9 (49.4)
FEMALE	.57	.59	.56
AGE	10.8 (.97)	11.0 (1.1)	10.7 (.83)
DAYABS	.65 (1.4)	.73 (1.5)	.58 (1.3)
MHH	.25	.27	.22
MOTHED1	.06	.09	.04
MOTHED2	.26	.32	.20
MOTHED3	.37	.40	.35
MOTHED4	.31	.19	.41
BOOKS	0 (1.00)	-.16 (1.16)	.14 (.81)
HWHELP1	.07	.07	.07
HWHELP2	.81	.8	.82
HWHELP3	.12	.13	.11
CLASS-LEVEL PREDICTORS			
CLMOTHED	2.9 (.48)	2.7 (.37)	3.1 (.47)
CLASS SIZE	36.8 (8.5)	40.2 (3.8)	33.9 (10.1)
TEACHER PREDICTORS			
EDLEVEL1	.05	.02	.07
EDLEVEL2	.01	0	.02
EDLEVEL3	.49	.30	.65
EDLEVEL4	.38	.68	.12
EDLEVEL5	.07	0	.13
TEAEXP	16.2 (9.8)	19.9 (7.7)	12.9 (10.1)

Table 2

Fixed and Random Effects Regression Models Showing the Estimated Effects of Peer-Group Composition and Class Size on Students' 5th grade Mathematics Achievement Score

(n students=3,095; n classrooms =97; n teachers=77)

	TEACHER FIXED EFFECTS			CLASSROOM RANDOM EFFECTS		
	2a	2b	2c	2a (r.e.)	2b (r.e.)	2c (r.e.)
CLASS-LEVEL PREDICTORS						
CLMOTHED		12.199	7.932		13.805***	12.698***
CLASSIZE	26.534*		23.919~	-0.483		-1.011
Goodness of Fit						
R^2 "overall"	0.351	0.350	0.351	0.290	0.312	0.312

~p<.10; *p<.05; **p<.01; ***p<.001

Notes: Class-size is represented by a modified log transformation described in Section 3 of the text. Under this transformation, smaller values of class size correspond to larger values in the transformed predictor. Coefficients associated with the teacher fixed effects and the control predictors have been omitted from the table to save space. The regression coefficients associated with the control predictors can be found in Appendix Table A2.

Figure 1
Sample distribution of estimated regression coefficients on the teacher fixed-effects, by
model specification
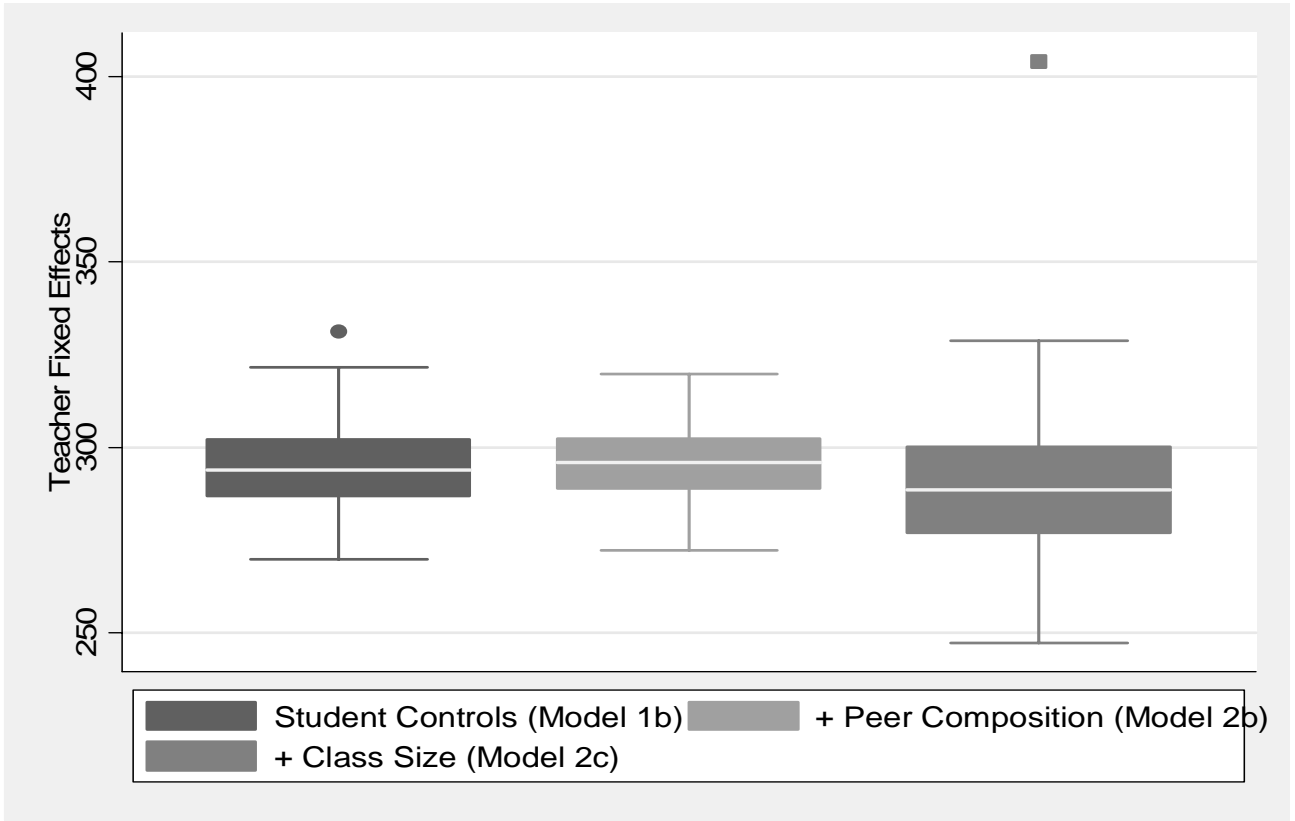
Table 3
Taxonomy of Classroom Random effects Regression Models Estimating the
Effects of School Sector on Students' 5th grade Mathematics Achievement Score
(n students=3,095; n classrooms=97; n teachers=77)

	3a* (no student-level controls)	3a	3b (coefficient on CLMOTHED constrained to its value in Model 2b)	3c (coefficients on Class Size and CLMOTHED constrained to their values in Model 2c
CLASS-LEVEL PREDICTORS				
CLMOTHED			(12.199)	(7.932)
CLASSIZE				(23.919)
SCHOOL SECTOR (PRIVATE=1) **PRIVATE**	**8.239***	**0.754**	**-3.789~**	**-14.502***
Goodness of Fit				
R² "within"	0.009	0.200	0.201	0.201
R² "between"	0.050	0.650	0.681	0.294
R² "overall"	0.025	0.290	0.316	0.195

~p<.10; *p<.05; **p<.01; ***p<.001

Notes: Regression coefficients for control variables can be found in Appendix Table A2.

Figure 2
Distribution of Estimated Teacher Fixed Effects by Sector and Model Specification
(Panel B is based on fitted Model 1; Panel C is based on fitted Model 2b; and Panel D is
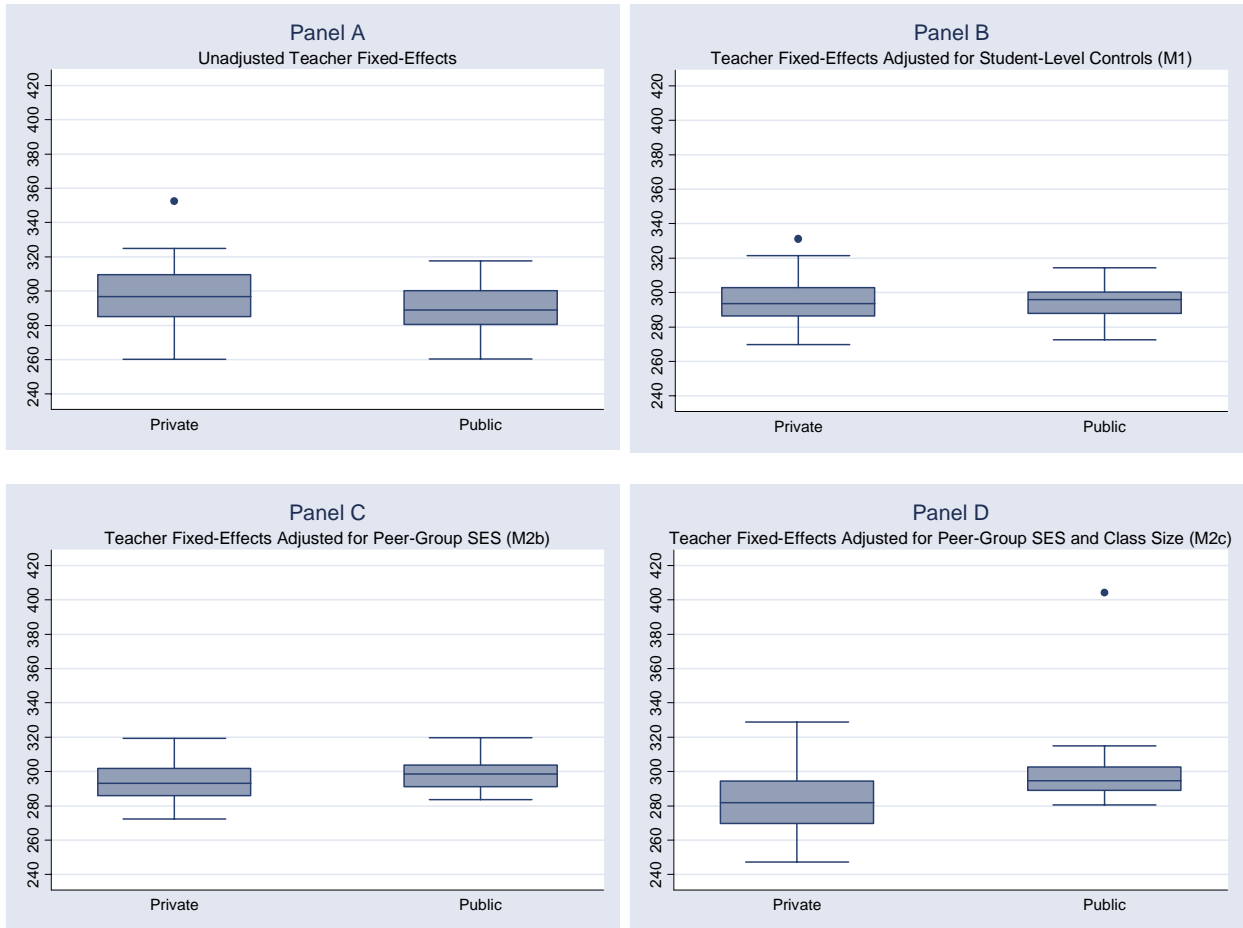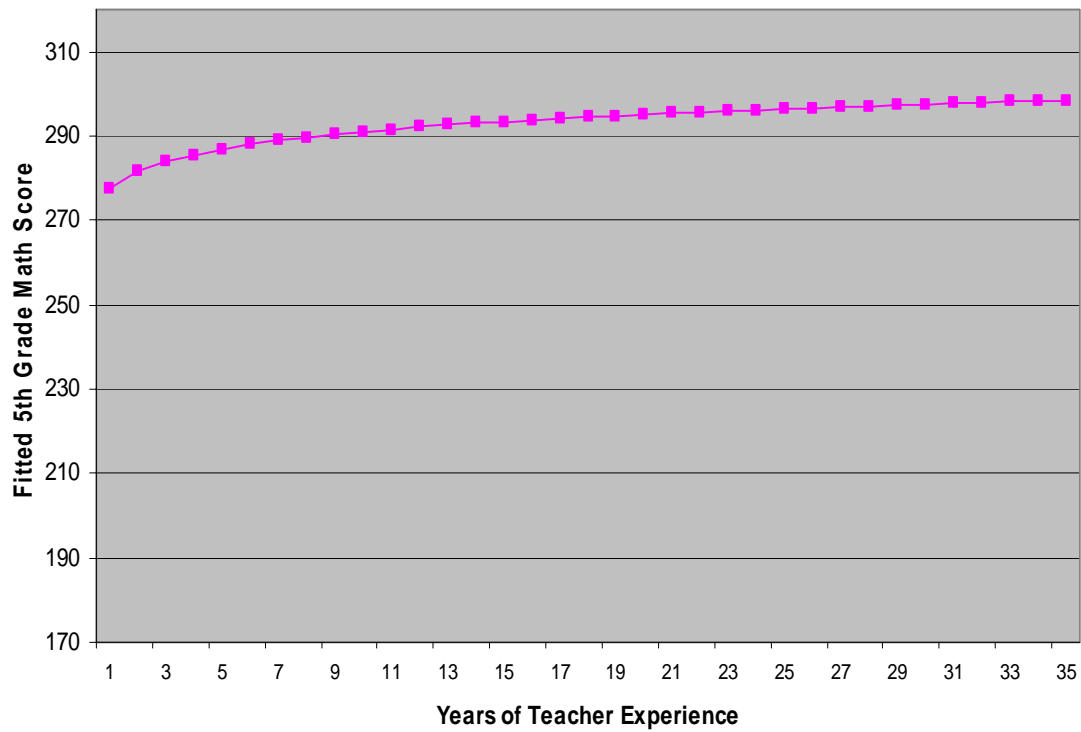based on fitted Model 2c)

Table 4
Taxonomy of Fitted Constrained and Unconstrained Random effects
Models Showing the Estimated Effects of Teacher Characteristics on
Students' 5th grade Mathematics Achievement Score
(n students=3,095; n classrooms =97; n teachers=77)

	CLASSROOM RANDOM EFFECTS (with Class Size and CLMOTHED constrained to the values of their estimates in Fitted Model 2c)		CLASSROOM RANDOM EFFECTS (Unconstrained Models)	
	4a	4b	4c	4d
CLASS-LEVEL PREDICTORS				
CLMOTHED	(7.932)	(7.932)	16.154***	15.694***
CLASSIZE	(23.919)	(23.919)	-0.802	-1.033
SCHOOL SECTOR (PRIVATE=1)				
PRIVATE	-12.695*		-1.603	
TEACHER PREDICTORS				
Teacher Years of Experience				
L_TEAXP	3.796	5.918*	0.097	0.327
Teacher Level of Education (Licensure or EDLEVEL3 omitted category)				
EDLEV1	-10.673	-11.842	-2.766	-2.840
EDLEV2	-23.592~	-27.053*	-2.345	-2.683
EDLEV4	-3.479	2.405	4.837~	5.438*
EDLEV5	5.107	3.233	-2.405	-2.838
Test of joint significance of EDLEV (Chi-square)	5.55	7.05	4.70	7.98~
Goodness of Fit				
R^2 "overall"	0.251	0.206	0.318	0.318

~p<.10; *p<.05; **p<.01; ***p<.001

Note: Estimated regression parameters associated with the control predictors can be found in Appendix Table A2.

Figure 3

Fitted Relationship between 5th Grade Mathematics Achievement Score and
Years of Teacher Experience for a Prototypical Student in a Classroom of Average Size
and Average Peer-Group SES
(from Model 4b)
(n students = 3,095; n classrooms = 97; n teachers=77)

Appendix Figures and Tables

Figure A1
Sampling Process for the Two-Wave
Data Collection

SELECTED SAMPLE
102 schools
(59 Private, 43 Public)

REFUSED 1999
14 schools
(5 Private, 9 Public

ACHIEVED SAMPLE 1999

88 schools (54 Private, 34 Public)
3,993 students

Non-matched students 1999
897

REFUSED 2000
15 schools
(7 Private, 8 Public)
663 Students

SAMPLE 2000
73 schools (47 private, 26 public)
77 mathematics teachers
99 Classrooms
3,149 5th grade students

"New" students 2000
714

Two classrooms excluded because of tracking

FINAL ANALYTIC SAMPLE
73 schools (47 private, 26 public)
77 mathematics teachers
97 Classrooms
3,095 5th grade students
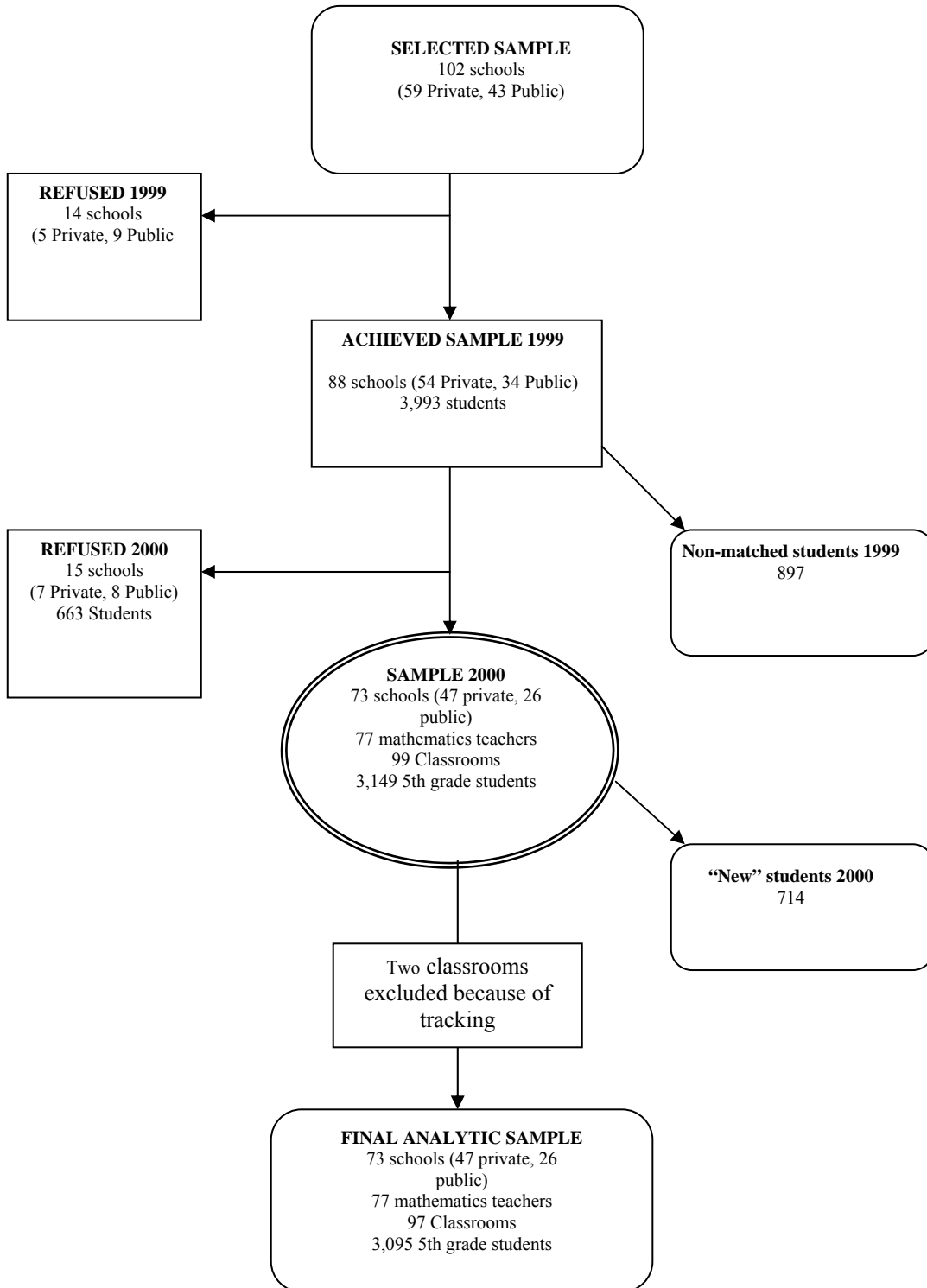
Table A1

Description of Outcome, Predictor and Control Variables

VARIABLE   DEFINITION

OUTCOME VARIABLE

   MSCORE   Student mathematics achievement score on a standardized test administered at the
                    end of 5th grade (possible score range = 0 to 492 points).

STUDENT-LEVEL CONTROL PREDICTORS

   FORM   Dichotomous variable indicating test form (0 = Form B, 1= Form A)

   PRETEST   Student achievement on a mathematics and language test administered at the end of
                    4th grade (possible score range = 0 to 306 points).

   AGE   Student age at the end of the 5th grade school year (years).

   FEMALE   Dichotomous variable indicating student's gender (0 = male, 1 = female)

   MOTHED   Ordinal variable indicating mother's educational level (0 = no education, 1 =
                    incomplete primary, 2 = complete primary, 3 = incomplete secondary and, 4 =
                    complete secondary and beyond).

   MHH   Dichotomous variable indicating whether the student's mother is the head of the
                    household (0=no,1=yes)

   DAYABS   Number of days that the student was absent from school in the last 2 weeks.

   HWHELP   System of dummy variables that indicate the frequency with which the student
                    receives help with homework at home (1 = never, 2 = sometimes, 3 = always). In
                    our analyses, HWHELP = 3 (always receives help) acts as the reference category.

   BOOKS   PCA-based composite measure constructed from multiple indicators of the
                    availability of books, story-books, and encyclopedias in a student's home.
                    Standardized to a mean of zero and a standard deviation of one, with positive
                    scores indicating better availability.

CLASSROOM PEER-GROUP COMPOSITION and CLASS SIZE

   CLMOTHED   "Leave-out" classroom average level of students' mother's education.

   CLMEAN   "Leave-out" classroom mean score on the pretest.

   CLASSIZE   Continuous variable indicating number of students enrolled in the classroom

TEACHER CHARACTERISTICS

EDLEVEL   Ordinal variable indicating the teacher's maximum educational level (either obtained or in process of obtaining). In our analyses, we converted this variable into a system of dichotomous (dummy) predictors, EDLEVEL1 through EDLEVEL5 (1 = Pedagogical high school, 2 = Educational Technology 3 = Licensure in Education, 4 = Postgraduate courses including Master's degree or specialization in education, 5=other professional qualification). EDLEVEL3 (licensure) was omitted from our analyses as a reference category.

TEAEXP   Continuous variable indicating the total number of years that the teacher has been in the teaching profession.

SSCALE   Ordinal variable indicating the teacher's step on the salary scale (1-14)

SCHOOL CHARACTERISTICS

PRIVATE   Dichotomous variable indicating whether the school is public or private (1 = private, 0 = public)

Estimated Regression Coefficients Associated with Control Predictors, by Model

	Model 1a	Model 1b	Model 2a	Model 2b	Model 2c	Model 2a (r.e.)	Model 2b (r.e.)	Model 2c (r.e.)
Constant	296.312***	296.546***	296.546***	296.586***	296.572***	296.027***	295.945***	296.091***
FORM	-6.277***	-6.325***	-6.328***	-6.335***	-6.334***	-6.440***	-6.403***	-6.457***
PRETEST	0.308***	0.301***	0.302***	0.302***	0.303***	0.329***	0.314***	0.310***
PRETEST SQ	0.001***	0.001***	0.001***	0.001***	0.001***	0.001***	0.001***	0.001***
FEMALE	-8.172***	-8.180***	-8.160***	-8.145***	-8.140***	-7.358***	-8.125***	-8.160***
AGE	-2.734***	-2.636***	-2.579***	-2.649***	-2.593***	-2.160**	-2.311**	-2.214**
AGE SQ	-0.841*	-0.896**	-0.906**	-0.885**	-0.897**	-1.003**	-0.840*	-0.869*
DAYABS	-1.853***	-1.758***	-1.807***	-1.762***	-1.805***	-1.939***	-1.888***	-1.832***
MHH	4.626***	4.371***	4.385***	4.395***	4.399***	4.521***	4.730***	4.702***
BOOKS	2.395***	2.479***	2.453***	2.465***	2.446***	2.693***	2.363***	2.404***
HWHELP1	7.734**	7.735**	7.792**	7.672**	7.745**	7.650**	7.695**	7.626**
HWHELP2	7.600***	7.627***	7.592***	7.528***	7.526***	7.374***	7.503***	7.463***
MOTHED	0.512	0.606	0.579	0.936	0.796	1.868**	0.991	0.974

	Model 3a	Model 3b	Model 3c	Model 3d	Model 4a	Model 4b	Model 4c	Model 4d
Constant	292.310***	295.540***	298.140***	300.424***	292.056***	277.394***	295.328***	293.681***
FORM	-6.782***	-6.447***	-6.401***	-6.359***	-6.353***	-6.351***	-6.389***	-6.392***
PRETEST		0.329***	0.318***	0.313***	0.311***	0.309***	0.313***	0.312***
PRETEST SQ		0.001***	0.001***	0.001***	0.001***	0.001***	0.001***	0.001***
FEMALE		-7.307***	-8.094***	-7.665***	-7.809***	-7.882***	-8.512***	-8.524***
AGE		-2.134**	-2.343***	-2.517***	-2.570***	-2.561***	-2.417***	-2.400***
AGE SQ		-1.005**	-0.873**	-0.922**	-0.898**	-0.883**	-0.827*	-0.826*
DAYABS		-1.938***	-1.896***	-1.866***	-1.860***	-1.865***	-1.894***	-1.899***
MHH		4.517***	4.660***	4.413***	4.469***	4.426***	4.801***	4.824***
BOOKS		2.683***	2.456***	2.460***	2.442***	2.405***	2.388***	2.381***
HWHELP1		7.603**	7.854**	7.697**	7.707**	7.662**	8.044**	8.024**
HWHELP2		7.348***	7.577***	7.531***	7.568***	7.544***	7.712***	7.695***
MOTHED		1.850*	1.244~	1.044	0.998	0.899	1.055	1.042

~p<.10; *p<.05; **p<.01; ***p<.001

Figure A2
Plot Displaying the Fitted Relationship between Students' 5th Grade Mathematics
Achievement Score and Class Size,
for a Prototypical Student in a Classroom of Average Peer-Group SES
(Based on Model 2c)
(n students=3,095; n classrooms=97; n teachers=77)