PUBLIC EDUCATION IN PUERTO RICO:
DOES CLASS SIZE MATTER?

Eddy V. Leal

Professor Thomas Nechyba, Faculty Advisor

Honors Thesis submitted in partial fulfillment of the requirements for Graduation with Distinction in Economics in Trinity College of Duke University

Duke University
Durham, North Carolina
2007

Acknowledgments

I am especially grateful to Thomas Nechyba, my advisor for his guidance and patience. His continuous support and general encouragement were crucial for the completion of this project. Alison Hagy's deadlines and feedback helped me stay on target and improve the writing of the paper. Joel Herndon helped me organize my data and become fluent in the language of STATA. Craig Burnside taught me the valuable lesson of examining your data before doing anything with it. Paul Dudenhefer's guide to writing in economics focused my writing skills. In addition, I want to thank officials from the Department of Education of Puerto Rico who helped me gain access to the data for this project. I would also like to thank Duke University's Undergraduate Research Support office for providing the funds needed for this research project. Last but not least, I want to thank Maribel Garcia for introducing me to academic research and my parents for their support and unique comments.

Abstract

Even though there is a large literature concerning the effects of class size on educational achievement, no previous research has formally examined the class size reduction policy in Puerto Rico. The evidence in this paper suggests that class size does not have a causal effect on student achievement in Puerto Rico. As a result, this paper points to a failure of the policy that Puerto Rico's government has invested heavily in for the last few decades in order to improve the quality of public education. Policy makers in Puerto Rico should seek alternatives in order to improve the quality of public education and consider innovations such as incentive based reforms now prevalent in the United States.

I. Introduction

Puerto Rico's public education system is in a state of fiscal crisis.[1] Puerto Rico's expenditures on primary and secondary education per student relative to per capita personal income (33.2 percent in 2001-02) greatly exceed the average in the United States (25.4 percent) (Ladd and Rivera-Batiz, 2006). This is largely due to the government's education policy that has been largely based on increasing educational inputs in order to improve educational outputs. A significant part of this spending has concentrated on decreasing class size, now smaller than in most states in the United States. In fact, while the student teacher ratio for Puerto Rico stood at 49.5 in the 1952-53 school year, it fell to 15.4 by 2003-04.[2] This is lower than the 16.1 student teacher ratio for the U.S. in 2002-03. Despite the large level of expenditure, the public education system suffers from a variety of deficiencies. Ladd and Rivera-Batiz (2006) estimate that the dropout rate for Puerto Rico is 21.3% (compared to 16.2% in the United States) and the majority of students in Puerto Rico are scoring in the lowest level of standardized tests.

The literature on the economics of education concerning class size has paid a great deal of attention to the United States, but little work has been done in this area with respect to Puerto Rico. This is surprising given the widespread belief among policy makers and in media circles that public education in Puerto Rico has undergone an

[1] In early May 2006, for instance, the local government closed most central government agencies because it had reached its borrowing capacity. Public school students and teachers were released a full month before the end of the school year ("Trouble on Welfare Island," 2006). The Department of Education is by far the country's largest employer with nearly half of those on the education department's payroll employed as non-teachers (Marlow-Ferguson, 2002). In 2003-04, there were 219,835 government employees, and the Department of Education accounted for 31.3 percent, or almost one out of every three (Ladd and Rivera-Batiz, 2006).

[2] In this paper, the data used is from the San Juan school district only. Relative to the rest of the island, this school district is highly urbanized and thus average class size is larger.

extreme deterioration ("Trouble on Welfare Island," 2006).[3] School production functions have been the predominant foundation of the traditional literature on the economics of education. This approach formulates an econometric model that relates the dependent variable (such as test results) to independent variables including characteristics from students, parents, and schools.

The effectiveness of inputs to education in the absence of additional reforms to improve the quality of education is a controversial topic in the literature. Results in this literature are inconclusive. Hanushek (1986; 1995; 1997; 2003) finds little evidence that class size reductions matter. Meanwhile, Krueger (1999) examines the Tennessee Student/Teacher Achievement Ratio experiment (known as Project STAR) and finds that class size does matter. While there have been many other studies, the school production function research has addressed the class size issue primarily by utilizing U.S. data. Even though Puerto Rico has spent heavily in order to decrease class size, no published research has formally examined the marginal benefit of decreasing class sizes in terms of student outcomes in Puerto Rico.

This paper will therefore use original data that I collected from Puerto Rico's Department of Education in order to investigate the marginal benefits of decreasing class size. I will estimate an education production function using ordinary least squares (OLS) regressions with fixed effects. In particular, I ask to what extent student achievement on standardized tests is systematically related to class size in Puerto Rico. My initial estimated coefficients on average class size suggest a counter-intuitive relationship: a positive relationship between class size and student achievement. If anything, the

[3] This is also unexpected in light of recent analyses that suggest that education has a powerful effect on the growth of national economies (Hanushek, 2002).

coefficients therefore imply that as average class size increases, the number of students expected to score in the basic (below proficient) level decreases and the number of students expected to score in the advanced level increases. However, after I adjust for the non-random allocation of school resources by using fixed effects, I find that class size does not have a causal effect on student achievement in Puerto Rico.

This paper will proceed as follows. Section II reviews the relevant literature on the class size debate. Section III describes the relevant theoretical framework that argues the importance of class size. Section IV describes the data used in this study. Section V analyzes the data. Finally, Section VI summarizes the findings and discusses the policy implications of this research.

II. Literature Review

The literature on the economics of education concerning class size in the United States is large and inconclusive. But it definitely raises questions for policy makers in Puerto Rico. School production functions that link school inputs to the quality of education have been the predominant foundation of the literature.[4] The effectiveness of decreasing student teacher ratios in the absence of any additional reform to improve the quality of education is a much debated topic in this literature. While there have been many studies, the school production function research has addressed this issue by utilizing primarily U.S. data. This paper contributes to the literature by examining this debate within the unique context of Puerto Rico.

Studies of school production functions include different measures of resources devoted to schools. The general approach formulates an econometric model that relates the dependent variable (such as test results) to independent variables including characteristics of students, households, and schools. One of the measures of resources is the *real resource* category – teacher education, teacher experience, and class size or student teacher ratios – as opposed to *financial aggregates* of resources (expenditure per student and teacher salary) and measures of other resources in schools (specific teacher characteristics, administrative inputs, and facilities). Hanushek (2002) points out that the *real resource* category accounts for most of the variations at the classroom level. These variables are readily available and well-measured and capture the most salient changes in Puerto Rico's public schools.

[4] Over the past decade, the literature has expanded to investigate more structural reforms aimed at raising the marginal product of school inputs. These include choice based reforms that were first suggested by Friedman (1962) and accountability systems such as those exemplified by the No Child Left Behind act.

The appropriate class size is a matter of debate in school systems throughout the world. The basic argument is straightforward. In smaller classes, teachers can provide more attention, even individualized work, for each student. Consequently, student achievement on tests should be higher in smaller classrooms. However, within the large body of existing research, one can find support for conflicting conclusions on the impact of class size on student performance. The attempts to test this hypothesis highlight the difficulty in determining how to improve resource use within schools.

Hanushek (1997), which updates the analysis in Hanushek (1986), reviews the large body of U.S. studies (available through 1994). The estimates of key production function parameters came from 376 separate published estimates, found in 89 separate articles or books.[5] The estimated relationships vary in a variety of substantive ways (by measure of student performance, by grade, by included measures of resources). These studies also vary widely in quality, as generally captured by methodology and adequacy of data.

In terms of the *real resource* category, fourteen percent of the estimates for teacher student ratios show positive and statistically significant effects on student performance. The individual studies tend to measure each of the inputs in different ways. For example, while many studies include an indicator variable for whether or not the teacher has a master's degree, some will include measures of the graduate credits. Some studies measure actual class size while the majority measures the teacher student ratios. A variety of functional forms has been used, ranging from simple linear relationships to different nonlinear forms with thresholds and quadratics (Hanushek 2002). The relatively

[5] All studies meet rudimentary quality standards (published in a refereed journal or book, including some measure of family background, and presenting information about the statistical properties of estimates).

small number of statistically significant positive results is balanced by another set finding statistically significant negative results – fourteen percent in the case of teacher student ratios, the same percentage that found an expected positive effect. In seventy two percent of the studies, estimated coefficients for the teacher student ratio are statistically insignificant.

In addition, Hanushek (1995) surveyed almost one hundred studies on the determinants of student test scores in developing countries. Of the thirty that looked at class size, eight found a statistically significant positive effect (the fewer students per teacher, the better students performed); eight found a statistically significant negative effect (the fewer students per teacher, the worse students performed); and fourteen found no significant effect at all. The results from the analysis of developing countries are consistent with the results in Hanushek (1997): no systematic relationship exists between class size and student performance.

Hanushek (2003) concludes that the input-based policy approach has been a failure. He points out that between 1960 and 2000 the United States lowered the student-teacher ratio from 25.8 to 16 (i.e., by more than a third). In addition, he notes that real spending per student in schools was 240 percent higher in 2000 than in 1960. Despite the substantial increases in education expenditure, test scores in math and English remained relatively flat from 1960 to 2000. Consequently, he finds a discrepancy between the money spent on education and the results obtained.

In the mid-1980s, because of ambiguity about the effects of class size on student performance, the State of Tennessee conducted a random assignment experiment in reducing class sizes (the project known as the STAR experiment).[6] In this specific experiment, a group of kindergarten students in Tennessee was randomly assigned to either large classes (22-24) or small classes (14-16). The student performance was monitored through testing as the students progressed from kindergarten to third grade. According to the test results, the children in smaller classes did better at the end of kindergarten and maintained their superior performance through third grade.

The findings of the STAR experiment seem to contradict the segments of the econometric evidence that suggests that student teacher ratios are irrelevant to student performance. However, two points are crucial to consider. First, while the use of experimentation is an important research approach, Hanushek (1999) and Mosteller (1995) emphasize that the actual implementation of this experiment is open to question. A number of questions have arisen regarding the quality of the randomization in the STAR experiment. Specifically, because of lack of data, it is difficult to assess the randomization of students or teachers into the experiments. There was substantial nonrandom attrition from the experiment; 51% of students initially in the experiment left before the end of the experiment. A substantial number of students (up to 12% by test) did not take annual tests; significant proportions of them changed experimental group (with the largest number of students going from large to small classes) during the experiment. Consequently, it is difficult to infer causation by looking solely at differences in mean performance across experimental and control groups.

Second, the findings, which pertain to a very large policy change in very specific circumstances, yielded results that were small and difficult to interpret. Because students

[6] The use of random-assignment experiments is an appealing alternative to econometric estimation in order to understand the effects of specific resources on student performance. Such an approach can potentially overcome a variety of concerns raised about selection and causation.

continue to get more resources (smaller classes), these resources should, according to the general hypothesis, keep producing a growing advantage. However, the initial achievement differences found in the year students entered a small class were maintained but did not become wider through the grades (Krueger 1999). Thus, if the benefit from class size reductions in Tennessee remain constant so long as small classes are maintained for the next twelve years, the intervention would be unlikely to pass a cost-benefit analysis.

The findings of U.S. studies of school production functions should be applied cautiously when considering Puerto Rico. They do not suggest that class size never matters; the results say only that there is no evidence of a robust and significant correlation between class size and student performance. In some instances, reducing class sizes may be warranted; in others, the extra expenditure on teachers and additional classrooms may yield no perceptible benefit. It may furthermore be the case that other structural reforms – such as incentive based reforms recently introduced in the United States – could make class size reduction more effectual. Regardless, it is important to undertake a research study that examines the effectiveness of reduced class sizes using data from Puerto Rico.

III. Theoretical Framework

The framework of analysis of educational performance considers a production function that assumes student achievement outcomes (such as test scores) depend on individual characteristics (such as innate ability), family background (such as parent's education and income), and classroom features (such as class size or teacher's experience and educational attainment). The general production function is:

$$O_i = f\{F_i, S_i, A_i\},$$

where O_i – performance of student i, F_i – family inputs cumulative to time t, S_i – cumulative school inputs, and A_i – innate ability.

The performance of a student is usually measured by test results because they are correlated with the expected gains in the labor market. It is important to point out that the influences outside of schools enter into the production of achievement. Given that families frequently have preferences over the schools that their children will attend, it is important to account for external influences on performance in order to mitigate issues of selection bias. Family inputs are usually measured by the parent's educational and income level.[7] School inputs refer to class size and teacher's experience and educational attainment. The innate ability of students is difficult to measure and no data is available for Puerto Rico. Averett and McLennan (2004) point out that most studies do not include this in their OLS regressions.

The relevant theory suggests that class size matters for student performance – with bigger classes having a negative impact on student achievement. Mulligan (1984) is one of the few authors to develop a theoretical basis for analyzing the relationship

[7] Omission of family backgrounds will almost certainly lead to biased resource estimates (Hanushek 2006). Family backgrounds have been shown to be quite generally correlated with school resources and have been shown to have strong effects on student outcomes.

between class size and achievement. However, his approach is not based on standard economic assumptions about the rational behavior of teachers and students. Instead, he assumes that students form a queue to receive individual attention from their teacher. They do not learn while they are in the queue and their time in it increases as class size increases. It follows that students in larger classes are likely to achieve less than those in smaller classes.

Correa (1993) utilizes the assumptions and methods of economic theory to arrive at the same conclusion by studying the behavior of a rational teacher. He presents a theoretical model of the effect of class size on student outcomes based on the time allocation of rational teachers. He hypothesizes that teachers are constrained to divide their time between activities that affect the entire class (e.g. preparation), those that affect individual students (e.g. answering questions) and those outside the classroom. A teacher's utility function includes the average performance of the class as well as the time in non-educational activity. The average performance of the class is increasing in the time spent by the teacher on educational activities. However, performance is more strongly increasing in the time spent on activities that affect the whole group. For this reason, teachers will focus on the overall performance of the class and this is more likely to harm individual students as class size increases. In short, this framework suggests that the coefficient on class size should be negative, indicating that higher class size has a negative effect on student achievement on tests.

Lazear (2001) provides another theoretical model suggesting that class size is important in the education function because of its effects on classroom atmosphere and is likely to be more effective for improving outcomes for disadvantaged students. At the same time, he shows how rational allocation of resources may still lead one to not find class size effects in empirical work.

Researchers have addressed the endogeneity of class size using a variety of approaches including random and natural experiments.[8] The most well documented random experiment is the Tennessee Student Teacher Achievement Ratio (STAR) project. Researchers using non-experimental data use a number of approaches including evaluating "natural experiments," controlling for a wide array of background variables, and employing instrumental variables or fixed effects techniques or some combination of these methods. For this study, I will use background variables and employ fixed effects and will argue that grade size within schools is more plausibly exogenous than class size.

[8] Hoxby (2000), Angrist and Lavy (1999) and Case and Deaton (1999) all take advantage of variation in class size due to exogenous circumstances in various settings.

IV. Data

The original data for this study comes from the Puerto Rico Department of Education. The school system in Puerto Rico is divided into four levels: preschool, which includes prekindergarten and kindergarten; elementary school, which comprises first to sixth grade; middle school, encompassing the seventh to ninth grades; and high school, which includes the tenth to twelfth grades.

Test Results

The absence of any systematic student testing in schools until the mid-1990s makes it difficult to evaluate student achievement in Puerto Rico’s public school system. The implementation of schoolwide testing occurred only gradually. The first tests, called General Index of Scholastic Competence (Índice General de Competencia Escolares), were administered from 1996 to 2002. However, these tests cannot be compared over time because they were changed drastically from year to year.

The requirements of the federal No Child Left Behind legislation (2002) led Puerto Rico to establish a more reliable testing battery, designed and managed by the Educational Testing Service in Princeton, New Jersey. The tests, called Puerto Rican Tests of Academic Achievement (PRTAA) (Pruebas Puertorriqueñas de Aprovechamiento Académico), were first administered in April 2003 to all public school students in the third, sixth, eighth, and eleventh grades that accounted for the 2002-03 academic year. In April 2004, the fourth, fifth, and seventh grades were also included. Following the No Child Left Behind guidelines, the test scores were converted into three levels: basic, proficient, and advanced. Students who scored at the basic level were considered not to be proficient in the subject matter. In Puerto Rico, the tests have been administered in three subject areas: mathematics, English (as a second language), and Spanish (reading).

The data consists of the percentage of students in the San Juan district who scored in the categories of basic, proficient, or advanced in each subject area on the PRTAA. The test results for 2004 and 2005 are aggregated at the grade level by school (i.e., not at class level). Tables 1, 2, and 3 present the average percentage of students scoring in the basic, proficient, and advance categories for the PRTAA in 2003 to 2005 in the mathematics, English, and Spanish subject areas, respectively. In all three subject areas, more than half of the students scored in the basic (below proficient) level (except Math for 2004-05).

Table 1. Public School Students' Performance on the PRTAA for Mathematics 2003-2005

Academic Year 2003-2004	Obs.	Mean	Std. Dev.	Min	Max
Basic (below proficient)	398	58.11%	23.09%	0	97
Proficient	398	34.73%	16.59%	3	84
Advanced	398	7.17%	13.79%	0	96
Academic Year 2004-2005	Obs.	Mean	Std. Dev.	Min	Max
Basic (below proficient)	400	47.86%	25.81%	0	97
Proficient	400	39.25%	17.54%	2	92
Advanced	400	12.93%	18.66%	0	95

Source: Data from Commonwealth of Puerto Rico, Department of Education 2005.
Note: Students from the San Juan district tested in grade 3-8 and 11.

Table 2. Public School Students' Performance on the PRTAA for English 2003-2005					
Academic Year 2003-2004	Obs.	Mean	Std. Dev.	Min	Max
Basic (below proficient)	398	59.91%	22.54%	0	100
Proficient	398	29.86%	16.89%	0	93
Advanced	398	10.26%	11.69%	0	76
Academic Year 2004-2005	Obs.	Mean	Std. Dev.	Min	Max
Basic (below proficient)	399	51.96%	25.22%	0	100
Proficient	399	31.44%	16.12%	0	83
Advanced	399	16.67%	17.63%	0	96
Source: Data from Commonwealth of Puerto Rico, Department of Education 2005. Note: Students from the San Juan district tested in grade 3-8 and 11.					

Table 3. Public School Students' Performance on the PRTAA for Spanish 2003-2005					
Academic Year 2003-2004	Obs.	Mean	Std. Dev.	Min	Max
Basic (below proficient)	398	57.86%	21.32%	0	98
Proficient	398	29.23%	13.36%	1	85
Advanced	398	12.93%	16.14%	0	92
Academic Year 2004-2005	Obs.	Mean	Std. Dev.	Min	Max
Basic (below proficient)	400	54.18%	22.99%	0	100
Proficient	400	28.70%	12.43%	0	68
Advanced	400	17.16%	18.95%	0	93
Source: Data from Commonwealth of Puerto Rico, Department of Education 2005. Note: Students from the San Juan district tested in grade 3-8 and 11.					

School Resources, Student (including parent), and Teacher Characteristics

The Puerto Rico Department of Education required that each public school submit information about school resources, students (including information about parents), and teachers. Table 4 contains grade level information on student (including parent) characteristics and average class size for the Fall 2003 to Spring 2004 period. The majority of students lived with families below the poverty line, in urban areas, and their parents had gained a high school preparation. The data for 2004-05 was similar and was therefore placed in the appendix (Table 1A).

The average class size variable was created from information reported to the Puerto Rico Department of Education. The data contain information on the number of students enrolled in the grade and the number of classrooms for a particular grade in each school. Given this information, the average class size variable is the number of students enrolled in a grade divided by the number of classrooms for that grade in each school. Grade size is arguably exogenous – and with limited numbers of classrooms in each school – this implies that average class size by grade is exogenous. It leaves open the possibility that class size within a grade is chosen unequally for different students – even as average class size is exogenous. Nevertheless, an exogenous increase in grade size would imply an increase in class size for all students even if schools rationally choose different class sizes within a grade (Lazear, 2001). This is captured by the average class size variable.

In addition, I created a variable to measure parents' average years of education for a particular grade. The data set contains the percentage of parents with an elementary, middle, high school, technical, vocational, undergraduate, masters, and PhD education. In order to create the variable, I multiplied the percentage of parents with a particular education times the approximate number of years that it takes to reach that educational level and then added all of them up. For instance, let us assume that for a particular grade, 80% of parents had a high school education and the rest had a middle school preparation. Given this case, parents average years of education for this particular grade is (.80*12) + (.20*9) = 11.4.

Table 4. Public School Students' Characteristics, 2003-2004

Independent Variable	Obs	Mean	Std. Dev.	Min	Max
Average Class Size	404	23.37	5.37	4	57
Students with Families Below Poverty Line	397	83.52%	16.03%	0	1
Students in Urban Zones	404	75.40%	37.95%	0	1
Parents' Average Years of Education	405	11.55	2.28	0	15.17

Source: Data from Commonwealth of Puerto Rico, Department of Education 2005.
Note: Students from the San Juan district in grades 3-8 and 11.

Regarding teacher characteristics for 2003-04, the mean years of experience was 14.25 (with a standard deviation of 3.2) and the mean academic preparation was 16.25 (with a standard deviation of .347). The majority of teachers had an undergraduate degree. The data for 2004-05 was very similar.

Limitations

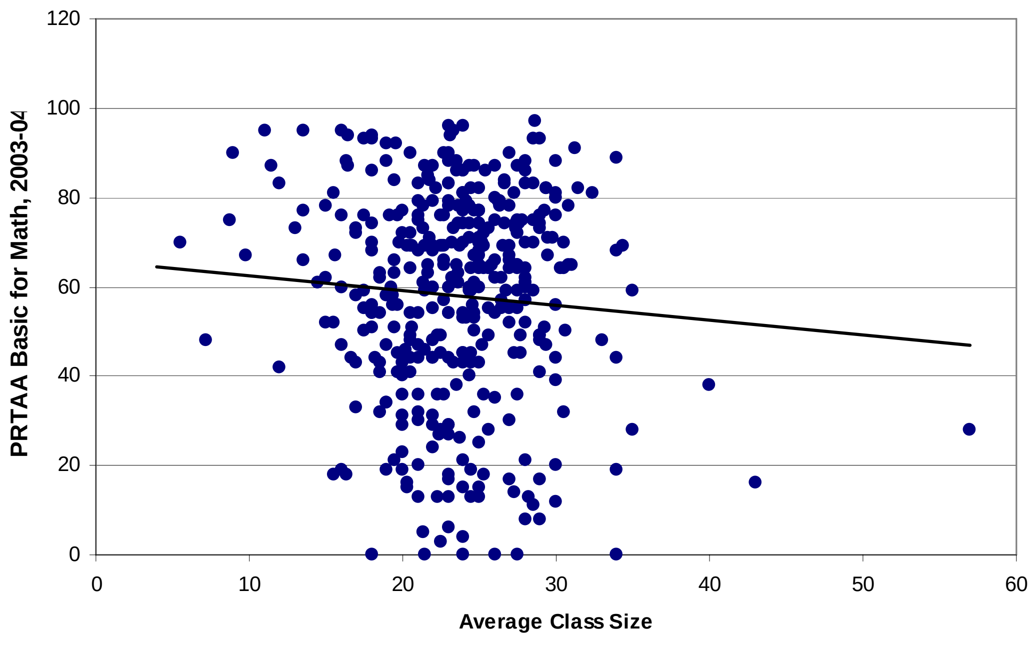
Two potential issues were present in my data. First, public schools do not uniformly submit the information solicited by the Puerto Rico Department of Education. Thus, the data sets have missing information that might make them susceptible to sample selection bias. For instance, schools failing to complete the forms might be more susceptible to having students that perform poorly on standardized tests. If this is the case, then the data provide an incomplete picture of the public schools in Puerto Rico. Second, the Educational Testing Service (ETS) does not provide the scores for classrooms with fewer than ten students. Thus, I will not have data for schools that have classrooms below ten students. This is an important issue because smaller classrooms might not be properly represented in the data set.

My examination of the data suggests that public schools in the San Juan district had a high completion rate of the forms and only a few data points were eliminated because of ETS reporting policy. Consequently, these are both important reasons to utilize the San Juan school district rather than the whole island.

Class size and student performance

The raw correlation between average class size and the proportion of students scoring in the below proficient level for math in the 2003-04 year is -0.0738. Figure 1 illustrates this weak negative correlation through a scatter plot. This correlation implies that as the average class size increases the number of students scoring in the basic (below proficient) level decreases. Thus, the raw correlation goes against the hypothesis that a large class size adversely impacts student performance on tests. In fact, by itself the correlation suggests the opposite. In the next section of the paper, I will explore this relationship in more detail to ascertain possible causal channels.

Figure 1. Correlation between Average Class Size and PRTAA Basic (below proficient) level for Math, 2003-2004



V. Empirical Specification

The hypothesis in this paper is that class size is systematically related to student achievement. Holding everything else constant, a large class size should adversely impact student performance. I begin by estimating an education production function using ordinary least squares (OLS) regression to test this hypothesis:

$$test_{ij} = \beta_0 + \beta(F) + \delta(S) + \upsilon_{ij}. \quad (1)$$

The dependent variable $test_{ij}$ is the percentage of students in grades three to six, seven to eight, and eleven scoring in the basic (below proficient), proficient, and advanced levels in mathematics, English, and Spanish for the San Juan district on the Puerto Rican Test of Academic Achievement. F is a vector of student (including family) characteristics variables, such as family average educational attainment. It is essential to account for family background characteristics otherwise the OLS estimator of the average class size effect on test scores would be biased. S is a vector of school characteristics, including the average class size, teacher's experience and academic preparation per grade.

The OLS regressions treat the grade, not the individual student, as the unit of analysis. These OLS regressions ideally should be cast at the individual level because it is pupils who are affected by class size. However, Angrist and Lavy (1999) argue that little is lost in statistical precision from this aggregation because class size is naturally fixed within classes and student test scores are correlated within classes. In this paper, there is no option other than a grade-level analysis because micro-level data are unavailable.

A key concern in the estimation of the effect of class size on student achievement using the education production function is the endogeneity of class size. Averett and McLennan (2004) state that it is generally acknowledged that non-random assignment sorting of students into classes does occur and should be accounted for in an investigation of the effects of class size on achievement. Without taking into account the endogeneity of class size, the approach often leads to spurious results.

As mentioned before, I address this concern by controlling for background variables and, later on, utilizing fixed effects. The purpose of utilizing school fixed effects is to eliminate variations outside of a particular school and thus, focus on within school variations. As with most public schools in Puerto Rico, they face fixed resources including the number of teachers and classrooms in each grade. So, the average class size fluctuates according to school enrollment for that grade. In short, in the case of Puerto Rico grade size is largely exogenous as is, therefore, average class size per grade.

While average class size within a grade is effectively exogenous, class size may vary within a grade. For instance, Lazear (2001) argues that principals should optimally assign students within grades to differently sized classes, with weaker or more disruptive students ending up in smaller classes. If the principal does this perfectly to maximize achievement, the marginal gain per student in a given year would then be the same across classes of similar size within a grade. Unfortunately, I only have data on average class size per grade. However, if the principal is faced with an exogenously larger grade size, this implies that all classes within that grade will be larger (unless additional classrooms are added, which is unlikely). Thus, all students within a grade should experience an effective increase in class size in line with the average class size increase for that grade.

Table 5 presents the estimated coefficients from twenty seven different regressions for average class size while controlling for student and parent characteristics and a time fixed effect. Each estimate is the result of a single regression. Given my

Table 5. Estimated Coefficients of Average Class Size and Student Performance				
Dependent Variable		Grades 3-6	Grades 7-8	Grade 11
Basic (+)	Math	-0.366 (1.83)^	-0.458 (1.50)	2.627 (3.46)**
	English	-0.232 (1.17)	-0.271 (0.82)	2.446 (3.01)**
	Spanish	-0.310 (1.77)^	-0.154 (0.60)	2.297 (3.22)**
Proficient (?)	Math	0.377 (2.74)**	-0.155 (0.66)	-1.646 (2.68)*
	English	0.045 (0.31)	0.248 (1.30)	-0.418 (0.90)
	Spanish	0.149 (1.43)	0.063 (0.33)	-1.262 (3.57)**
Advanced (-)	Math	-0.006 (0.04)	0.622 (2.53)*	-1.002 (2.66)*
	English	0.185 (1.59)	0.026 (0.12)	-2.081 (4.15)**
	Spanish	0.161 (1.10)	0.083 (0.49)	-1.076 (2.47)*

Absolute value of t statistics in parentheses.
^ significant at 10%; * significant at 5%; ** significant at 1%
Controls for students (including parents) and teachers' characteristics and a time fixed effect.

hypothesis, I expect that the estimated coefficients for average class size will be positive for the dependent variables in the basic category and negative for the advanced category. The theoretical impact on proficient is ambiguous.[9] For grades three to six, the coefficients on average class size have the opposite (from expected) sign in almost all cases, with half the coefficients for basic and advanced coming close to statistical significance. A similar trend appears to hold for grades seven to eight, although only one coefficient reaches statistical significance.

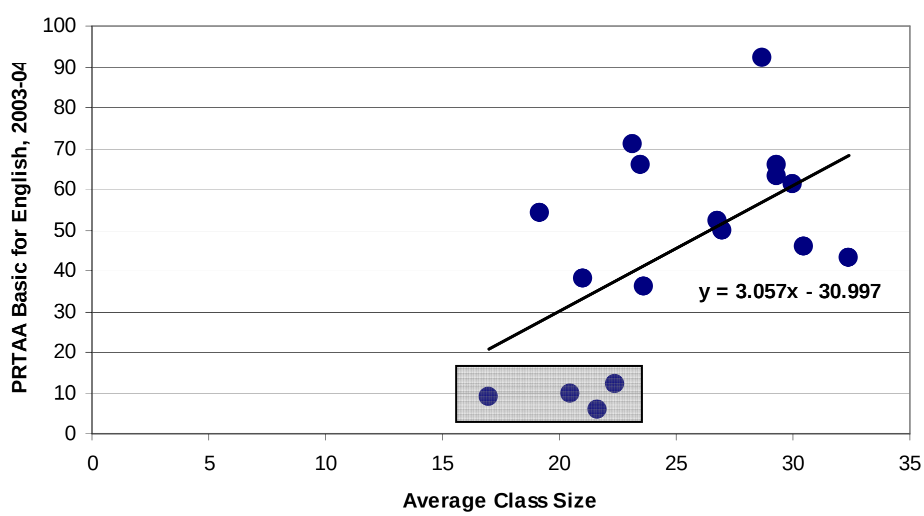
The estimated coefficients for average class size therefore suggest a positive relationship between class size and student performance for grades three to six and seven

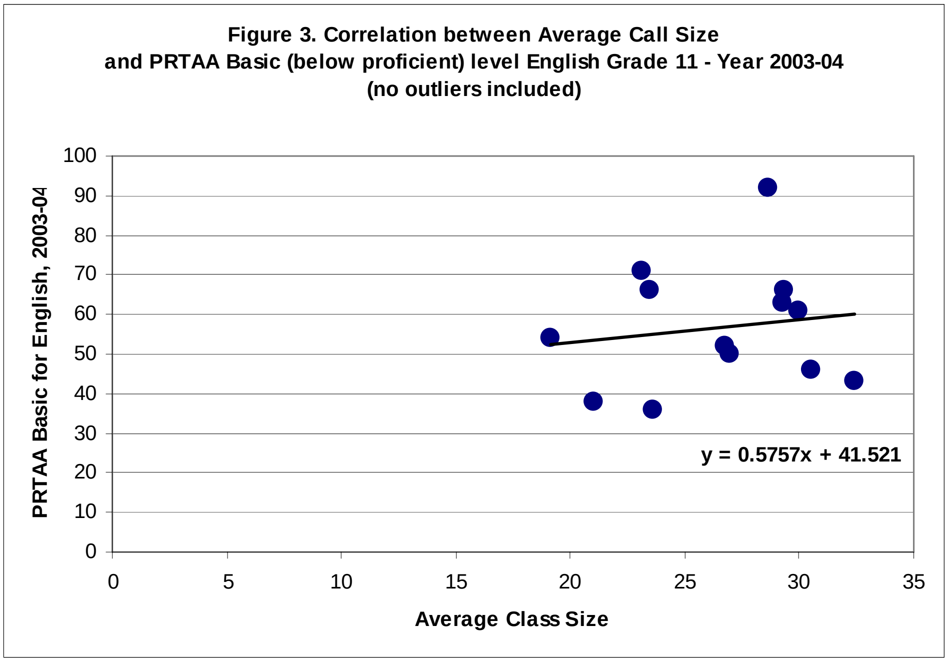
[9] Consider a class where the class size has been increased, theory suggests that some students scoring in the proficient level will move to the basic (below proficient) level and some students scoring in the advance level will move to the proficient level. However, theory does not suggest which movement outweighs the other and thus, the theoretical impact on proficient is ambiguous.

to eight. These regressions taken at face value imply that a large class size does not adversely impact student performance on tests. If anything, the coefficients imply that as average class size increases, the number of students expected to score in the basic (below proficient) level decreases and the number of students expected to score in the advanced level increases.

In the case of grade eleven, the estimated coefficients for average class size are in line with my expectations and the majority are statistically significant. However, there are relatively few 11th grade observations in my data and the result appears to be driven by four outliers. Figure 2 presents a scatter plot with a regression of the percentage of students that scored in the basic (below proficient) level for English in 2003-04 on average class size. The four schools that are in the shaded box skew the results and make the coefficient statistically significant. After removing these four schools that appear to be outliers, the estimated coefficient for average class size is no longer statistically significant (See Figure 3). Given the relative lack of 11th grade data, I will therefore focus on grades three to eight from this point forward.

Figure 2. Correlation between Average Call Size and PRTAA Basic (below proficient) level English Grade 11 - Year 2003-04





The estimated coefficients on the control variables have the expected sign in almost all cases. For instance, the estimated coefficients for average parents' educational attainment have a negative relationship with student performance in the basic (below proficient) level. The coefficients imply that as average parents' educational attainment increases, the number of students expected to score in the basic (below proficient) level decreases. Similarly, the estimated coefficients for parental education on the percentage scoring in the advanced category are positive. The remaining household control variables (i.e., fraction below poverty, fraction residing in urban areas) also consistently have the expected sign, but are not always statistically significant. The coefficients on teacher experience were rarely significant and typically had the opposite from expected sign, while the coefficients on teacher education usually had the expected sign but were not consistently significant.

The unexpected signs on the average class size coefficients in the OLS regressions for grades three to eight may be the result of non-random allocation of school resources in Puerto Rico. For instance, school funding might be differently targeted to particular grades that are thought to be especially responsive to class size. Alternatively, the Department of Education may differentially target lower performing schools in order to reduce class size where such reductions can have the largest impact. Such targeting might then result in a correlation between class size and student achievement – not because increased class size causes achievement to rise but rather because low achievement leads to increased resources and thus smaller classes.

Since I observe multiple grades within schools and schools across time, I can also make use of grade and school fixed effects to control for omitted variables. Time fixed effects can be used to account for variations in tests across time.

First, I estimate an education production function using an OLS regression with grade fixed effects:

$$test_{ij} = \beta_0 + \beta(F) + \delta(S) + \sum \gamma_k G_{kij} + \upsilon_{ij},$$

where $_i$ = school, $_j$ = grade, and G_{kij} = 1 if k = j, 0 otherwise. The grade fixed effects are used to control for omitted variables in the data that are constant over time but vary across grades. I arbitrarily omit a binary variable for one of the grades in order to avoid perfect multicollinearity. In this regression, I estimate the effect on test scores of class size holding constant family background, school characteristics, and unobserved grade characteristics which vary across grades, but not across time. I then estimate the production function using school fixed effects:

$$test_{ij} = \beta_0 + \beta(F) + \delta(S) + \sum_{k=2}^{N} \gamma_k D_{kij} + \upsilon_{ij},$$

where $_i$ = school, $_j$ = grade, and D_{kij} = 1 if k = i, 0 otherwise.

Table 6 then provides the estimated coefficients from fifty six different regressions for average class size for grades three through six using different variants of these specifications. Each estimate is the result of a single regression. At the bottom of the table, I indicate whether or not demographic controls as well as different types of fixed effects were included in each of the six types of regressions. The basic pattern identified in the initial OLS specification from Table 5 continues to hold throughout specification (1) through (4). The results are therefore immune to the inclusion of demographic controls, time fixed effects and grade fixed effects. This suggests there is no different targeting of resources for grades across schools.

Table 6. Estimated Coefficients for Average Class Size for Grades 3-6

Dependent Variable		(1)	(2)	(3)	(4)	(5)	(6)
Basic (+)	Math	-0.373 (1.90)^	-0.355 (1.74)^	-0.366 (1.83)^	-0.398 (2.14)*	0.044 (0.20)	0.084 (0.40)
	English	-0.242 (1.25)	-0.226 (1.12)	-0.232 (1.17)	-0.286 (1.44)	0.114 (0.54)	0.167 (0.81)
	Spanish	-0.419 (2.44)*	-0.309 (1.76)^	-0.310 (1.77)^	-0.404 (2.33)*	0.213 (1.19)	0.218 (1.26)
Proficient (?)	Math	0.383 (2.86)**	0.372 (2.68)**	0.377 (2.74)**	0.401 (3.03)**	0.055 (0.34)	0.023 (0.14)
	English	-0.004 (0.03)	0.043 (0.30)	0.045 (0.31)	0.085 (0.59)	-0.193 (1.16)	-0.188 (1.15)
	Spanish	0.237 (2.36)*	0.152 (1.46)	0.149 (1.43)	0.151 (1.57)	0.017 (0.14)	0.026 (0.21)
Advanced (-)	Math	-0.003 (0.02)	-0.011 (0.08)	-0.006 (0.04)	-0.003 (0.02)	-0.096 (0.59)	-0.102 (0.65)
	English	0.244 (2.07)*	0.181 (1.50)	0.185 (1.59)	0.199 (1.70)	0.080 (0.59)	0.024 (0.17)
	Spanish	0.180 (1.27)	0.156 (1.06)	0.161 (1.10)	0.253 (1.83)^	-0.233 (1.38)	-0.247 (1.52)
Time Fixed Effect		NO	NO	YES	YES	YES	YES
Grade Fixed Effect		NO	NO	NO	YES	NO	NO
School Fixed Effect		NO	NO	NO	NO	YES	YES
Demographic Controls		NO	YES	YES	YES	YES	NO

Absolute value of t statistics in parentheses.
^ significant at 10%; * significant at 5%; ** significant at 1%

However, the previous pattern of results disappears with the inclusion of school fixed effects in specifications (5) and (6). The signs of the coefficients for average class size become consistent with the hypothesis but are neither economically nor statistically significant. This suggests that the Department of Education conducts class reduction programs in poor performing schools. This result also holds for grades seven and eight (See Table 7). When controlling for this, class size does not have a causal effect on achievement.

Table 7. Estimated Coefficients for Average Class Size for Grades 7-8

Dependent Variable		(1)	(2)	(3)	(4)	(5)	(6)
Basic (+)	Math	-0.382 (1.27)	-0.454 (1.46)	-0.458 (1.50)	-0.451 (1.47)	0.049 (0.12)	-0.002 (0.01)
	English	-0.309 (0.98)	-0.270 (0.82)	-0.271 (0.82)	-0.275 (0.82)	0.372 (0.78)	0.417 (0.94)
	Spanish	-0.161 (0.57)	-0.148 (0.54)	-0.154 (0.60)	-0.166 (0.65)	0.044 (0.12)	-0.023 (0.07)
Proficient (?)	Math	-0.032 (0.14)	-0.154 (0.65)	-0.155 (0.66)	-0.183 (0.78)	0.091 (0.21)	0.091 (0.24)
	English	0.225 (1.30)	0.248 (1.30)	0.248 (1.30)	0.252 (1.31)	-0.204 (0.62)	-0.196 (0.64)
	Spanish	0.150 (0.74)^	0.060 (0.31)	0.063 (0.33)	0.062 (0.33)	-0.002 (0.01)	-0.008 (0.03)
Advanced (-)	Math	0.424 (1.74)	0.617 (2.40)*	0.622 (2.53)**	0.643 (2.61)**	-0.158 (0.37)	-0.104 (0.26)
	English	0.083 (0.38)	0.025 (0.11)	0.026 (0.12)	0.026 (0.12)	-0.165 (0.52)	-0.218 (0.74)
	Spanish	0.010 (0.06)	0.080 (0.45)	0.083 (0.49)	0.096 (0.56)	-0.055 (0.18)	0.011 (0.04)
Time Fixed Effect		NO	NO	YES	YES	YES	YES
Grade Fixed Effect		NO	NO	NO	YES	NO	NO
School Fixed Effect		NO	NO	NO	NO	YES	YES
Demographic Controls		NO	YES	YES	YES	YES	NO

Absolute value of t statistics in parentheses.
^ significant at 10%; * significant at 5%; ** significant at 1%

VI. Conclusion

The government of Puerto Rico for the last few decades has invested heavily in decreasing class size in order to improve the quality of public education. Despite substantial expenditures, the public education system suffers from a variety of deficiencies. For instance, the dropout rate is particularly high and the majority of students score in the lowest level of standardized tests. Even though there is a very large literature on the economics of education concerning class size, no previous research has formally examined the class reduction policy in Puerto Rico. Consequently, I ask to what extent student achievement on standardized tests is systematically related to class size in Puerto Rico.

My initial regressions suggest a positive relationship between class size and student achievement. If anything, the coefficients therefore imply that as average class size increases, the number of students expected to score in the basic (below proficient) level decreases and the number of students expected to score in the advanced level increases. However, after adjusting for the non-random allocation of school resources by using fixed effects, I find that class size does not have a causal effect on student achievement in Puerto Rico. This finding is consistent with the literature.

This paper suggests a failure of the class reduction policy in the absence of any additional reforms. As a result, policy makers in Puerto Rico should seek alternatives in order to improve the quality of public education. For instance, they should examine market-based reforms, such as school vouchers, public school choice, and accountability system. In the past decade, such choice and incentive based reforms have become a major focus for education policy makers around the world, with an aim of increasing the marginal impact of additional and current spending on education.

Appendix

Table 1A. Public School Students' Characteristics, 2004-2005					
Independent Variable	Obs	Mean	Std. Dev.	Min	Max
Average Class Size	402	23.38	5.38	4	57
Students with Families Below Poverty Line	395	83.68%	15.91%	0	1
Students in Urban Zones	402	75.31%	38.02%	0	1
Parents' Average Years of Education	403	11.56	2.29	0	15.17
Source: Data from Commonwealth of Puerto Rico, Department of Education 2005 Note: Teachers from the San Juan district in grades 3-8 and 11.					

References

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

Averett, S. L., & McLennan, M. C. (2004). International Handbook on the Economics of Education. In G. Johnes & J. Johnes (Eds.), *Exploring the effect of class size on student achievemen: What have we learned over the past two decades?* (p. 329-367). Northampton, MA: Edward Elgar Publishing Ltd.

Case, A., & Deaton, A. (1999). School inputs and educational outcomes in South Africa. *Quarterly Journal of Economics*, *114*(3), 1047-1084.

Commonwealth of Puerto Rico, Department of Education. 2005. Data. San Juan.

Correa, H. (1993). An economic analysis of class size and achievement in education. *Education Economics*, *1*(2), 129-135.

Friedman, M. (1962). *Capitalism and freedom*. Chicago and London: University of Chicago Press.

Hanushek, E. A. (1986). The economics of schooling: Production and efficiency in public schools. *Journal of Economic Literature*, *24*(3), 1141-1177.

Hanushek, E. A. (1995). Interpreting recent research on schooling in developing countries. *World Bank Research Observer*, *10*(2).

Hanushek, E. A. (1997). Assessing the effects of school resources on student performance: An update. *Educational Evaluation and Policy Analysis*, *19*(2), 143-163.

Hanushek, E. A. (1999). Some findings from an independent investigation of the Tennessee STAR experiment and from other investigations of class size effects. *Educational Evaluation and Policy Analysis*, *21*(2), 143-163.

Hanushek, E. A. (2002). Handbook of Public Economics. In A. J. Auerbach & M. Feldstein (Eds.), *Publicly provided education* (Vol. 4, p. 2045-2141).

Hanushek, E. A. (2003). The failure of input-based schooling policies. *Economic Journal*, *113*(485), 64-98.

Hoxby, C. M. (2000). The effects of class size and composition on student achievement: new evidence from natural population variation. *Quarterly Journal of Economics*, 1239-1285.

Krueger, A. B. (1999). Experimental estimates of education production functions. *Quarterly Journal of Economics*, *114*(2), 497-532.

Ladd, H. L., & Rivera-Batiz, F. L. (2006). Puerto Rico: Prospects for growth. In *Education and economic development in Puerto Rico.*

Lazear, E. P. (2001). Educational production. *Quarterly Journal of Economics*, *116*(3), 777-803.

Marlow-Ferguson, R. (Ed.). (2002). World education encyclopedia: A survey of educational systems worldwide. (2^{nd} ed.). Farmington Hills, MI: Gale Group.

Mosteller, F. (1995). The Tennessee study of class size in the early school grades. *The Future of Children*, *5*(2), 113-127.

Mulligan, J. G. (1984). A classroom production function. *Economic Inquiry*, *22*(2), 218-226.

National Center for Education Statistics (NCES). (2003). *Common core of data (CCD) 2002-03*. National Public Education Financial Survey, 2002-03. Washington, D.C.: US Department of Education.

Trouble on Welfare Island. (2006). *The Economist*, May 25.

Unger, H. G. (2001). *Encyclopedia of American education* (Vol. 2). New York City, NY: Facts on File Inc.