SMALLER CLASSES AND STUDENT ACHIEVEMENT:

THREE PAPERS EXPLORING THE CLASS SIZE EFFECT

A Dissertation

by

COURTNEY AMANDA COLLINS

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

May 2010

Major Subject: Economics

SMALLER CLASSES AND STUDENT ACHIEVEMENT:

THREE PAPERS EXPLORING THE CLASS SIZE EFFECT

A Dissertation

by

COURTNEY AMANDA COLLINS

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Approved by:

Chair of Committee,	Li Gan
Committee Members,	Timothy J. Gronberg
	Thomas R. Saving
	Lori L. Taylor
Head of Department,	Timothy J. Gronberg

May 2010

Major Subject: Economics

ABSTRACT

Smaller Classes and Student Achievement:

Three Papers Exploring the Class Size Effect. (May 2010)

Courtney Amanda Collins, B.A., Rhodes College

Chair of Advisory Committee: Dr. Li Gan

This dissertation analyzes the effect of smaller classes on student performance using student-level test score data from the state of Texas, focusing on three specific issues: heterogeneity in the returns to smaller classes across a score distribution of students, the relationship between class size and students' moving decisions, and the connection between smaller classes and schools' class division procedures.

I first examine evidence of heterogeneity in the returns to class size reductions across a score distribution of students. I divide students into decile groups based on their previous year test scores, and I estimate the returns to smaller classes for each of the deciles. The empirical evidence supports the hypothesis that there are significant differences in students' responses to class size, based on their previous test scores.

I then model the class size effect simultaneously with students' decisions to switch schools, which is important because movers compose a substantial fraction of the dataset, and because class size effects vary between movers and nonmovers.

Recognizing that students move for different reasons, only some of which are school-related, I present a two-type moving model in which students are categorized as endogenous movers or exogenous movers. I estimate the model estimated using maximum likelihood. The results reveal key biases in traditional estimates of the moving effect and suggest significant differences in the class size effect across mover types.

I also explore the class size effect in conjunction with schools' decisions to sort students into different classes. Using student-level data in which students are linked to specific classes, I disentangle the class size effect from the sorting effect. Including a variable indicating the sorting index of a school decreases the magnitude and significance of the class size effect. I also examine different types of sorting. The findings suggest that sorting students into more homogeneous groups is beneficial for both high and low scoring students.

For Gree,

who asked for a dissertation.

And for Paw,

who will always be my hero.

ACKNOWLEDGEMENTS

I would like to thank my advisor, Dr. Li Gan, for his guidance and patient encouragement throughout my research and graduate work at A&M. Thank you also to my committee—Dr. Timothy Gronberg, Dr. Thomas Saving, and Dr. Lori Taylor—for their many comments and suggestions, and for *not* living up to every horror story I ever heard about a graduate committee. Thanks for making the process as smooth and efficient as writing a dissertation can be.

I gratefully acknowledge the Private Enterprise Research Center and the Bradley Foundation for giving me several opportunities to pursue my research. I would like to thank the staff at PERC, especially Dr. Andy Rettenmaier, for the chance to collaborate on projects and for giving me constructive feedback on my own work. A big thank you goes to Juanice for her persistent encouragement and attentiveness to my dry erase board. It looks like they're all marked off now.

Thank you to the countless others who helped define the A&M experience for me—my classmates and fellow grad students, who helped me fight my daily battle with Matlab and humored my attempts to learn Chinese; the students in my principles and metrics classes, who gave me an exceptionally enjoyable first teaching experience and taught me what it means to be an Aggie; and the secretaries, who navigated the administrative channels for me and who were always ready for a trip to Lenny's. Thanks for helping me enjoy the journey. And thank you to the unnamed ones who taught me the most important lesson I learned in grad school: it is important to be nice.

Thanks to my Rhodes professors who lured me into economics—especially Marshall and Teresa Gramm and Mark McMahon. They invested in the lives of their students and showed us what it is to be a truly incredible teacher.

Thanks especially to the Family Rope, for helping me remember what is important and for providing much spirited encouragement along the way.

Thanks to Mom and Dad, who taught me to wear the hat made especially for me. And thanks to Savannah, who called out data for my research projects, helped me memorize the Second Fundamental Welfare Theorem, and decorated all my qualifier study material with her fourth grade artwork. ZZA.

Thank you to Milton, the noble steed and wise research assistant.

A special thanks to Jasmine, who played a most spirited troll in search of billy goats; to Charlie, who counted down from ten and shot the rocket to the moon; and to the eighteen other four-year-olds in Mrs. Webb's 2008-2009 Pre-K class who reminded me that education is nothing if it does not sparkle.

And most important, I want to thank God—who is mighty even when I am not, and who understood everything, econometrics and all, before the world began. Thank you for the παρακλητος.

TABLE OF CONTENTS

		Page
ABSTRACT		iii
DEDICATION		v
ACKNOWLEDGEMENTS		vi
TABLE OF CONTENTS		viii
LIST OF FIGURES		x
LIST OF TABLES		xi
CHAPTER		
I	INTRODUCTION	1
II	CLASS SIZE AND STUDENT ACHIEVEMENT: EXPLORING HETEROGENEITY IN RETURNS TO CLASS SIZE REDUCTIONS	6
	II.1 Introduction	6
	II.2 Current Literature	8
	II.3 A Model of Heterogeneity in the Class Size Effect Across a Score Distribution	11
	II.4 Texas Student-Level Data	12
	II.5 Potential Endogeneity of Class Size	15
	II.6 Results	17
	II.7 Conclusions and Future Work	25
III	SWITCHING TO BETTER SCHOOLS: A MODEL OF CLASS SIZE AND MOVER ENDOGENEITY	27
	III.1 Introduction	27
	III.2 Current Literature	29
	III.3 A Two-Type Model of Moving	33
	III.4 Empirical Implementation of the Type-Specific Model	39
	III.5 Student-Level Data from Texas	45
	III.6 Results	50

CHAPTER Page

III.7 Conclusions and Future Work 72

IV DOES SORTING STUDENTS IMPROVE SCORES? AN ANALYSIS OF CLASS SIZE AND CLASS COMPOSITION 75

IV.1 Introduction 75
IV.2 Empirical Model 77
IV.3 TAKS Data from Dallas ISD 82
IV.4 Empirical Results 85
IV.5 Conclusions 104

V CONCLUSIONS 106

REFERENCES 109

APPENDIX A 112

APPENDIX B 116

VITA 145

LIST OF FIGURES

FIGURE Page

II.1 Class size effect by decile groups (math score) 22

II.2 Class size effect by decile groups with fixed effects (math score) 23

II.3 Class size effect by decile groups (reading score) 24

II.4 Class size effect by decile groups with fixed effects (reading score) 25

III.1 Two-type mover model 35

III.2 Three-type mover model 40

III.3 Percent of students in bottom deciles 67

IV.1 Summary of sorting status by sorting type, based on regressions from Appendices B1 and B2 90

LIST OF TABLES

TABLE		Page
II.1	Summary statistics for TAKS test	15
II.2	Effects of class size on math score	18
II.3	Effects of class size on reading score	20
III.1	Summary statistics of fourth grade students in 2004-2005	45
III.2	Mean comparison tests between nonmovers and movers	51
III.3	Mean comparison tests (difference in school characteristics between current school and previous school)	52
III.4A	Probit of moving decision on student characteristics (unconditional campus, district, and region moves)	55
III.4B	Probit of moving decision on student characteristics (conditional campus, district, and region moves)	56
III.5	Effects of class size on math score with and without move dummy	58
III.6	Effect of class size on math score (effects by moving types)	60
III.7	Effect of class size on math score (MLE with move dummies)	63
III.8	Moving decision with neighborhood characteristics (from MLE)	64
III.9	Effect of class size on math score with decile groups (2SLS)	68
III.10	Comparison of the move effect	70
IV.1	Summary statistics for Dallas ISD	84
IV.2	Class size effect on math score (all schools)	91
IV.3	Class size effect on math score by sorting type (G/T)	93

TABLE | Page

TABLE		Page
IV.4	Effect of class size on math score with sorting index	94
IV.5	Score sorting effect on math score	97
IV.6	G/T sorting effect on math score	99
IV.7	Alpha score sorting effect on math score	101
IV.8	Alpha score sorting effect on math score by score type	103
B1.1	Score sorting for schools 101-113	117
B1.2	Score sorting for schools 114-125	118
B1.3	Score sorting for schools 126-135	119
B1.4	Score sorting for schools 136-148	120
B1.5	Score sorting for schools 149161	121
B1.6	Score sorting for schools 162-173	122
B1.7	Score sorting for schools 174-187	123
B1.8	Score sorting for schools 189-199	124
B1.9	Score sorting for schools 200-210	125
B1.10	Score sorting for schools 211-220	126
B1.11	Score sorting for schools 222-233	127
B1.12	Score sorting for schools 235-263	128
B1.13	Score sorting for schools 264-274	129
B1.14	Score sorting for schools 275-284	130
B2.1	G/T sorting for schools 101-113	131
B2.2	G/T sorting for schools 114-125	132

TABLE Page

B2.3 G/T sorting for schools 126-135 133

B2.4 G/T sorting for schools 136-148 134

B2.5 G/T sorting for schools 149-161 135

B2.6 G/T sorting for schools 162-173 136

B2.7 G/T sorting for schools 174-186 137

B2.8 G/T sorting for schools 187-198 138

B2.9 G/T sorting for schools 199-209 139

B2.10 G/T sorting for schools 210-219 140

B2.11 G/T sorting for schools 220-232 141

B2.12 G/T sorting for schools 233-262 142

B2.13 G/T sorting for schools 263-273 143

B2.14 G/T sorting for schools 274-284 144

CHAPTER I

INTRODUCTION

The efficacy of school inputs has been a thoroughly studied topic, both in the economics literature and in other fields. For decades, policymakers, researchers, and school administrators have attempted to understand the determinants of student achievement—to uncover the inputs that improve performance and those that have little or no effect at all. The study of class size has been of particular interest, both in the academic literature and in statewide implementations. Many have touted smaller class sizes as a key factor in raising the academic achievement of children, emphasizing that smaller classes allow for more individualized attention for each student and a more manageable classroom for teachers. These factors may eventually lead to better performance on achievement tests, among other positive outcomes.

These potential benefits are associated with substantial costs, however. Smaller class sizes mean hiring new teachers and building additional classroom space. Several states have poured millions of dollars into class size reduction policies. In 1996, the state of California implemented statewide legislation that funneled $1 billion per year into class size reduction efforts. In 2002, Florida passed an amendment that would require class size reduction efforts in excess of

---

This dissertation follows the style of the Journal of Public Economics.

$16 billion. Many other states have restrictions that govern maximum class sizes for some or all grades and students.

Because class size reduction policies are so costly, it is important to have a clear understanding of their benefits to students. Although much research has been conducted relating to this topic, there are still several unaddressed issues which may affect the overall estimation of the class size effect. This dissertation explores three specific questions regarding class size.

First, I examine whether or not evidence exists of heterogeneity in the class size effect across a score distribution of students. I assume that the marginal productivity of class size may vary across students, depending on their previous year testing score. While I allow the marginal product of smaller classes to vary, I make no *a priori* assumption for its functional form. It is possible that smaller classes are most beneficial to students at the top of the testing distribution; alternatively, students with very low previous year testing scores may gain the most from a reduction in class size. It may also be the case that students near the middle of the distribution benefit most from smaller classes.

In order to determine which of these assumptions best fits empirically, I examine student-level test score data from the Texas Assessment of Knowledge and Skills (TAKS) exam, the standardized test given to students in the Texas public school system each year. I group students into deciles based on their previous year testing score and allow the returns to class size to vary across the score distribution. I evidence supporting heterogeneity in the class size effect. OLS and

2SLS estimates suggest that smaller classes are most effective for students whose scores fell in the lower end of the overall distribution in the previous year. This result holds for both the math and the reading scores from the exam.

The second question I examine relates the class size effect to students' decisions to switch schools from one year to another. A surprising number of students in the Texas dataset can be classified as a "mover," or someone who transfers to a new school in the following year. One third of students move at least once between grades 3 and 6, excluding any moves caused by a transition to middle school. Fifteen percent of students in any given cohort are movers. Because a substantial amount of mobility intrinsically affects school enrollment, and thus class size, it is important to model the moving effect simultaneously with the class size effect.

I create a two-type mover model in which students are classified into two groups: endogenous movers, who move schools because of some school-related reason, and exogenous movers, who move schools because of a reason unrelated to school. An endogenous mover might switch schools because his parents are unhappy with his current teacher, whereas an exogenous mover might switch schools because his father got a new job in a different city.

The student-level TAKS data allows me to link a child to his school campus, his district, and his region of the state. From this information, I create three groups of movers: campus movers, who switch schools but remain in the same district; district movers, who switch districts but remain in the same region; and region movers, who move regions across the state of Texas. I use these three transfer types to identify endogenous and exogenous movers. My prior belief is that students who are more likely to be endogenous movers will be campus or district movers, while students who move regions will be more likely to be exogenous movers. Using maximum likelihood estimation, I estimate the two-type moving effect simultaneously with the class size effect.

The results suggest strong heterogeneity in the class size effect across moving types. Most of the class size effect is driven by students who switch districts but remain in the same region of Texas, a result which is partially explained by the fact that district movers are more heavily composed of low scoring students. I also find that simple OLS or 2SLS estimates of the moving effect are biased relative to the simultaneous MLE model.

In the third paper, I attempt to disentangle the class size effect from schools' ability to sort students into classes. I assert that because schools with smaller classes necessarily have more classes, schools that use strategic sorting mechanisms will be able to sort more efficiently if they have smaller classes. Therefore, what is often labeled a class size effect may be confounded with schools' sorting ability.

In order to separate these two effects, I use a unique student-level data from Dallas ISD, which allows students to be tracked to their specific classes. This is important because students can be linked to their classmates, and score distributions are available for individual classes, rather than at the grade level. I create a sorting index based on the difference in the standard deviations of scores of individual classes. I then include this sorting index in a typical class size regression. I find that adding the sorting variable significantly decreases the magnitude of the class size coefficient.

I also explore different types of sorting and examine how they affect student test scores. I find that sorting students into homogeneous groups based on test score is advantageous for all students—both high scoring students and low scoring students. This mechanism appears to be significantly more beneficial than creating more heterogeneous groups with a balance of students from both the high and low end of the score distribution.

CHAPTER II

CLASS SIZE AND STUDENT ACHIEVEMENT:

EXPLORING HETEROGENEITY IN RETURNS TO CLASS SIZE REDUCTIONS

II.1 Introduction

The emphasis on student achievement as measured by test scores is becoming increasingly important in the United States, especially since the implementation of the No Child Left Behind (NCLB) Act of 2001. Although most states had some form of standardized student assessment prior to the law, the legislation expanded statewide testing by mandating that each public school student in grades three through eight be tested in both reading and math every year. The outcome of these tests play a major role not only in determining student promotion, but also in influencing schools' federal funding and right to self-administration.

Although a few schools with exceptionally high-scoring student populations are largely unaffected by the new assessment standards, NCLB provides important incentives for most schools and districts. For example, in 80 percent of Texas schools, at least half of all students in the seventh grade scored within 100 points of the passing threshold or below it.[1] Similar statistics hold for other grades, indicating that a significant number of schools are strongly affected by incentives implemented by NCLB.

[1] The mean of the test scores for the 7th graders is 2093 and the standard deviation is 246.

Such an emphasis on scores as a measure of achievement leads educators and policymakers to carefully evaluate how different measures of school inputs affect test scores. If educators are able to examine school inputs (such as class size, teacher education, and spending) and evaluate which ones affect student performance the most, then they can use that information in a practical way to benefit actual achievement. One of the most hotly-debated school input variables within both education circles and fields of economic analysis is class size. Class size reduction has been touted by policymakers as a key component in raising measures of student achievement, and politicians and interest groups have advocated channeling funds toward programs that emphasize smaller classes, especially for disadvantaged students. Currently, 40 states in the US have implemented programs or laws to limit class size or student/teacher ratios, at least for certain grades or groups of students. Several states have funneled large amounts of money into statewide class size reduction programs. The state of Florida, for example, is currently implementing a 2002 amendment that legislates class size maximums for every class in the state; to date, more than $16 billion have been spent to hire the additional teachers needed to comply with the law.

However, despite the apparent popularity of this initiative, researchers have not come to a consensus on the merits and efficiency of smaller classes. While some studies suggest class size reductions may increase student performance for some grades[2], others conclude that there is no solid evidence that smaller classes affect achievement in a meaningful way[3]. I revisit the class size question in this paper by adding a key assumption to the typical model. While most previous work has examined the *average* effect of class size (or any school input) on student achievement, I allow for the possibility of heterogeneity in returns to class size across a testing distribution of students. I explore whether inputs that have little or no effect on average may have significant impacts for some parts of the distribution. I find evidence that class size returns do vary based on students' previous score.

II.2 Current Literature

A large portion of the class size discussion centers around meta-analyses of the existing research. Perhaps the most well-known of these studies is the work conducted by Hanushek (1986, 1997). In the more recent study (1997), he analyzes 277 estimates from 59 studies that explore the effect of smaller classes on student achievement and that meet minimal quality requirements. Giving equal weight to each estimate, he finds no evidence of a significant, systematic relationship between class size and student performance.

Krueger (2003) takes issue with Hanushek's method of "vote counting" and contends that the proper procedure would be to give equal weight to each *study* rather than to each *estimate*. He reevaluates Hanushek's analysis and determines

---

[2] Angrist and Lavy (1999), for example, find that smaller classes increase performance for 4th and 5th grade students, but not for 3rd graders.

[3] See Hoxby (2000) and Rivkin et al. (2005).

that if each published study were given equal weight, the results would support a significant relationship between smaller classes and student achievement.[4]

Outside of the meta-analysis debate, there are several important studies whose methodologies produce interesting (and again, often conflicting) results. Angrist and Lavy (1999) address the potential endogeneity of class size by exploiting the discontinuity caused by maximum class size rules in Israel. They create an instrument for class size based on the Israeli school system's practice of limiting class size to 40 students (see section II.4.2 for a description of the instrument). Their results indicate class size effects for some (but not all) students. They find, for example, that a one student reduction in class size is associated with an increase in average math score of 0.05 points (on a 100 point scale) and an increase in average reading score of 0.13 points. They find significant effects for fourth and fifth graders, but not for third graders.

Hoxby (2000) uses maximum and minimum class size rules in Connecticut, combined with exogenous variation in the population of school-aged children, to identify the class size effect. She finds no evidence that smaller classes increase student achievement, nor does she find that smaller classes are beneficial for schools with a higher proportion of minority or low-income students.

While much of the class size literature uses econometric techniques to deal with the endogeneity problems present in existing data, there are a few studies that

---

[4] For a detailed description of both Hanushek and Krueger's methodologies and arguments, see *The Class Size Debate* (Mishel and Rothstein, eds.).

are experimental in nature and involve random assignment. The most important of these is the Tennessee Student Teacher Achievement Ratio (STAR) program, which marked the first large-scale study with an experimental design.[5] Schools in Tennessee were given the option of participating in the program, which would randomly assign students to either small classes (13-17 students) or large classes (22-25 students). These students were tracked from kindergarten through third grade and remained in either small or large classes. Researchers found that students in small classes performed significantly better on a standardized test at the end of kindergarten and that the score gap continued, but did not widen, as those students continued through third grade.

The results of the study have been interpreted in several different ways. Many researchers point to the significant class size effect as strong evidence that smaller classes make a substantial difference in improving student performance. In fact, this study was (and continues to be) one of the driving forces behind much of the state legislation aimed at reducing class size. However, other researchers highlight the fact that while class size may be important for children in kindergarten, if the effect persisted in higher grades, the score gap should increase as children remained in small classes.

[5] See Word et al. (1990) for a complete description of the Tennessee STAR project.

II.3 A Model of Heterogeneity in the Class Size Effect Across a Score Distribution

A typical model of the effect of school input, such as class size, on student achievement is given by the following equation:

$$s_{ijt} = \rho s_{ijt-1} + \eta C_{jt} + X_{ijt}\beta + \varepsilon_{ijt}, \quad \text{(II.1)}$$

where s_{ijt} is the test score of student *i* in school *j* at time *t*, s_{ijt-1} is student *i*'s score in the previous year, C_{jt} is the class size (or other school input) within a grade at school *j*, and X_{ijt} is a vector of student-specific demographic variables. The error term ε_{ijt} is assumed to be independent and identically distributed. The parameter η estimates the returns to a change in class size, on average. A common result in the literature is that school inputs like class size have no significant effect on score gains, after controlling for relevant demographic variables.

The key contribution of this paper is that I allow for the possibility of heterogeneity in returns to class size across the score distribution. I assume that the marginal productivity of a class size reduction varies across a distribution of students, allowing students in different percentile groups to experience differential gains from the same class size reduction. Intuitively, a student scoring at the 10^{th} percentile level responds differently to a smaller class size than a student scoring at the 50^{th} percentile level, and they both respond differently than a student scoring at the 90^{th} percentile level.

Although I assume that the marginal productivity of a class size reduction varies across a testing distribution, I do not make any assumptions about *how* it varies. Several possibilities may arise. The marginal productivity may be increasing in test score, so that students at the top of the score distribution respond the most to a reduction in class size. The opposite may also be true; if the marginal productivity function is decreasing in test score, then low scoring students have the most to gain from a class size reduction. However, it is not necessary for marginal productivity to be linear in score; it may be concave or convex so that students in the center or at the ends of the distribution respond the most. For now, I will only allow that differences in returns to class size exist, rather than assuming a specific functional form for the marginal productivity.

I incorporate this assumption into the model by dividing students into decile groups based on their previous year scores. I add dummy variables for each group into a typical model like equation (II.1):

$$s_{ijt} = \beta s_{ijt-1} + \sum_{k=1}^{10} \eta_{1k} C_{jt} 1\left[p_{1k} \leq s_{ijt-1} \leq p_{2k}\right] + X_{ijt}\eta_2 + \varepsilon_{ijt} \qquad \text{(II.2)}$$

where student i is sorted into group k if his score falls between two threshold scores, p_{1k} and p_{2k} . Returns to class size are then allowed to vary by decile group. For the empirical results that follow, I divide students into ten groups, but it is also possible to analyze returns to fewer or more groups.

II.4 Texas Student-Level Data

I use student-level test score data obtained from the Student Assessment Division of the Texas Education Agency (TEA). The dataset contains student mathematics and reading test scores from the Texas Assessment of Knowledge and Skills (TAKS) for grades three through eleven in 2004 and 2005. Because Texas class size laws only apply to classes up to grade four, I use student test scores in grades three and four for the empirical analysis. Each student in the dataset is assigned a unique student identification number, so that third graders' individual scores in 2004 can be tracked to the corresponding fourth grade scores in 2005. Students' grades, schools, and districts are known, although students cannot be linked to a specific class within a school.

I merge the TAKS score dataset by school with the Academic Excellence Indicator System (AEIS) report, also available from the TEA. The AEIS report includes average class size for each grade in each Texas public school. Although most classes are around 19 and 20, there is considerable variation across schools. The combined dataset also contains detailed demographic data at the student level, including gender, ethnicity, free or reduced lunch eligibility, migrant status, and ESL status.

Two scores for each student are reported in the dataset—a raw score and a scale score. The raw score, which falls between 0 and 40, indicates the number of questions the student answered correctly on the exam. The scale score is used to adjust for test difficulty across testing administrations and is calculated by the TEA using a Rasch Partial-Credit Model (RPCM). Scale scores fall between 1228 and 2697 for the math exam and 1319 and 2614 for the reading exam. Scale scores map one-to-one with the raw scores. Although scales scores are meant to control for changes in difficulty across testing administrations, they are not meant to be vertically linked from one year to the next. That is, a student's third grade scale score in 2004 is not necessarily comparable to his fourth grade score in 2005.

Because neither of the provided scores can be vertically linked, I transform the scale scores into z-scores so that they can be compared from one year to the next. The z-scores are defined in the following way:

$$Z_{ijt} = \frac{Scale_{ijt} - \mu(Scale_t)}{\sigma(Scale_t)} \qquad \text{(II.3)}$$

where Z_{ijt} is the z-score of student i at school j in year t, $Scale_{ijt}$ is the corresponding scale score, and μ and σ are the mean and standard deviation of the scale scores, calculated across the entire dataset.

Summary statistics for all variables are reported in Table II.1.

Table II.1
Summary statistics for TAKS test

	Obs	Mean	St Dev	Min	Max
Class Size	137389	19.7549	3.1857	3	48.4
Math Score	131371	2263.45	187.15	1280	2684
Previous Math Score	131371	2251.97	178.048	1228	2699
Reading Score	139362	2241.47	170.467	1319	2614
Previous Reading Score	139362	2294.18	154.169	1356	2588
Female	139362	0.51239	0.49985	0	1
Asian	139362	0.01447	0.1194	0	1
Black	139362	0.1286	0.33476	0	1
Hispanic	139362	0.37217	0.48339	0	1
Free Lunch	139362	0.39368	0.48857	0	1
ESL	139362	0.01651	0.12743	0	1
Bilingual	139362	0.02611	0.15947	0	1
GT	139362	0.03338	0.17963	0	1
Special Ed	139362	0.00477	0.06891	0	1
Average School Math	131371	2212.97	113.746	1324.81	2515.83
Var School Math	131371	64829.8	46268	112.5	419740
Average School Read	139362	2238	74.6046	1791	2519.33
Var School Read	139360	24514.7	5999.54	0	225121
Acc Rating	139362	2.54027	0.67914	1	4

II.5 Potential Endogeneity of Class Size

There are at least two sources of potential endogeneity in the model described in equation (II.2)—measurement error and sorting error. Endogeneity from measurement error may arise because the observed variable C_{jt} is average class size across the grades within a school, not actual class size for each student. (Even if actual class size data were available through the TEA, they would not be useful because individual students cannot be linked to specific classes.) Endogeneity is present if actual class size varies systematically with student scores.

A second source of endogeneity stems from sorting based on test score. For example, schools may sort low-ability students into smaller classes and high-ability students in larger classes, resulting in a positive correlation between ability and class size. Since ability is positively correlated with score gain, the coefficient on class size in a typical OLS regression would be upward-biased. The opposite result would transpire if schools sorted high-scoring students into smaller classes and low-scoring students into larger classes.

I propose two strategies to deal with potential endogeneity. The first is to use maximum class size rules to develop an instrument for actual class size. This instrument, proposed by Angrist and Lavy (1999), exploits the discontinuity in class size caused by maximum class size rules. In the state of Texas, classes in kindergarten through grade 4 are allowed to have up to 22 students, but no more. When the number of students in a grade is equal to 22, then average class size is 22. However, one additional enrolled student triggers the maximum class size rule. A school with 23 students in a grade is forced to add an additional class, and the average class size would be 11.5. The actual instrument used in the analysis, predicted class size, is calculated as follows:

$$PCS_{jt} = \frac{enroll_{jt}}{\text{int}\left(\frac{enroll_{jt} - 1}{22}\right) + 1} \qquad \text{(II.4)}$$

where $enroll_{jt}$ is equal to the total enrollment within a grade at school j in time t, and int(x) is the smallest integer greater than or equal to x. Predicted class size is positively correlated with average class size, and there is no reason to believe that it should be correlated with the error term in equation (II.2).

I also include school-fixed effects as another method to try to reduce problems caused by endogeneity. These fixed effects will account for any student-invariant school-level heterogeneity, such as administrative style or overall school efficiency.

II.6 Results

II.6.1 Average Class Size Effects

The ultimate question to be examined is whether changes in class size affect students differently across the score distribution. Before exploring any evidence of a differential impact on students, I first consider the average effect of a class size change. Table II.2 shows the results of the regression of math score on class size. The OLS column represents the base regression, before including school fixed effects or instrumenting for class size. In the absence of fixed effects, average school score and variance of school score are included to partially control for school quality. Student demographic controls include race, gender, and economic disability.

Table II.2
Effect of class size on math score

	OLS			2SLS		
Class Size	-0.0004	(0.99)		-0.0044	(1.97)	**
Previous Score	0.4027	(237.83)	***	0.4027	(237.73)	***
Female	-0.0149	(5.83)	***	-0.0149	(5.83)	***
Asian	0.1527	(14.48)	***	0.1531	(14.51)	***
Black	-0.0767	(16.68)	***	-0.0747	(15.82)	***
Hispanic	-0.0132	(3.87)	***	-0.0124	(3.61)	***
Free Lunch	-0.0522	(16.39)	***	-0.0540	(16.21)	***
ESL	0.0853	(8.75)	***	0.0862	(8.83)	***
Bilingual	-0.0195	(2.44)	**	-0.0176	(2.19)	**
Special Ed	-0.1250	(7.43)	***	-0.1238	(7.35)	***
GT	0.2576	(34.56)	***	0.2585	(34.59)	***
Average Score	0.0017	(90.54)	***	0.0017	(90.36)	***
Var Score	2.36E-06	(65.40)	***	2.35E-06	(64.30)	***
Acc Rating	-0.0110	(4.53)	***	-0.0108	(4.45)	***
R-squared	0.4941			0.4937		
Obs	131371			131371		

Recall that the dependent variable is measured as a z-score. The class size coefficient indicates that a one-student reduction in class size results in a statistically insignificant 0.0004 point increase in math z-score, on average. In addition to being insignificant, the point estimate itself is very small.[6] Most other coefficients have their expected signs.

Table II.2 also shows the results of the regression using predicted class size as an instrument for average class size. The Donald-Cragg F-statistic for weak instruments is 4444.5, which is substantially higher than the critical value for a

[6] The mean scale score is 2231.72, and the standard deviation is 210.58.

weak instrument. As predicted, the coefficient on class size is lower than the corresponding OLS coefficient, indicating that endogeneity causes an upward bias in the original estimate. The 2SLS class size coefficient indicates that a one-student reduction in class size leads to a 0.0044 higher math z-score (.4 percent of a standard deviation) on average.

Table II.3 reports the same regressions for the reading exam. The results are similar, except that reading scores seem to be even less responsive to class size reductions that math scores. The baseline OLS results suggest that class size does not significant impact reading score. (The point estimate itself is -0.00029, which is even smaller than the corresponding math effect.) The magnitude of the coefficient increases as expected when predicted class size is used as an instrumental variable, but the estimate is still not significantly different from zero.

Table II.3
Effect of class size on reading score

	OLS			2SLS		
Class Size	-0.0003	(0.47)		-0.004	(1.08)	
Previous Score	0.57588	(240.66)	***	0.57585	(240.60)	***
Female	0.06946	(15.84)	***	0.06945	(15.83)	***
Asian	0.18872	(10.23)	***	0.18915	(10.25)	***
Black	-0.092	(11.70)	***	-0.0906	(11.33)	***
Hispanic	-0.0094	(1.61)		-0.0085	(1.43)	
Free Lunch	-0.1177	(21.45)	***	-0.1191	(21.02)	***
ESL	-0.0428	(2.34)	**	-0.0416	(2.27)	**
Bilingual	-0.0895	(5.86)	***	-0.0879	(5.72)	**
Special Ed	-0.2475	(7.82)	***	-0.2455	(7.75)	***
GT	0.48022	(38.52)	***	0.48134	(38.46)	***
Average Score	0.00322	(71.98)	***	0.00323	(71.30)	***
Var Score	8.51E-07	(2.29)	**	8.68E-07	(2.33)	**
Acc Rating	-0.0153	(3.60)	***	-0.0151	(3.53)	***
R-squared	0.4599			0.4599		
Obs	137389			137389		

The results from Tables II.2 and II.3 suggest that average class size effects are, at best, small in their impact on measures of students' math achievement, and insignificant in their effect on reading scores. This is consistent with evidence found in much of the literature (see Hoxby 2000 and Rivkin et al. 2005). However, effects which are small or insignificant on average may be more important for certain groups of students. Smaller classes may be beneficial for students who are performing particularly poorly or well in school, or they may be especially helpful for students near the center of the distribution. I now explore the data to test for these possibilities.

II.6.2 Class Size Effects by Decile Group

I examine the effect of class size across a distribution of students by first dividing the students into decile groups based on their 2004 score. I rank all third grade students within a campus based on their previous year scores, and then I divide students into ten decile groups based on those rankings. (Note that because students' scores tend to clump into groups, the deciles do not contain all contain exactly the same number of students.) Decile 1 contains the bottom ten percent of students in every campus, decile 2 contains the next ten percent of students in every campus, and so on. The deciles are included in the regression as dummy variables, so returns to class size are allowed to vary from group to group. The additional controls included in the base regressions are included here as well.

Figure II.1 shows heterogeneity in class size returns by decile group for the math exam. Each point on the table represents the coefficient estimate for returns to class size for each decile group. Consider the OLS results first. The largest (in magnitude) effects are for students in decile 1, or students at the low end of the previous year's score distribution. The effect decreases for the higher deciles, showing that the effect of smaller classes seems to be decreasing in previous test score. The coefficient estimates are all significantly different from each other using an F test comparison, with the exception of groups 9 and 10.

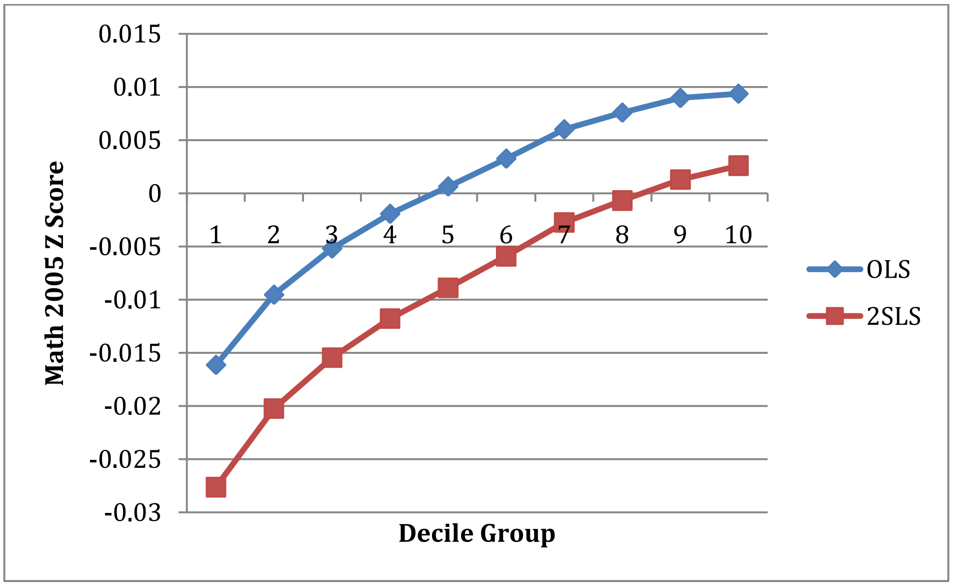

Figure II.1. Class size effect by decile groups (math score).

Figure II.1 also compares heterogeneity in class returns from the OLS regression with the corresponding IV estimates. The IV estimates reflect the same shape illustrated in the OLS estimates; the main difference lies in the magnitude of the results. Like the IV class size results without decile groups, these results are significantly larger in magnitude when compared with their OLS counterparts. All of the estimates for the deciles groups are significantly different from each other.

Figure II.2 reports the returns to class size by decile groups for the 2SLS regression with fixed effects. Because I am using school fixed effects in these regressions, I must exclude one of the decile groups to avoid perfect multicolinearity. This means that the coefficients reflect the *difference* in class size returns between deciles, but I cannot directly determine the *level* of returns. In order to report actual returns for each group, I use the original 2SLS estimates from the base regressions (given in Table II.2) and add the estimated coefficients from the current regression. That is, the shape of the returns in Figure II.2 is based on the fixed effects regression, but the level is only correct if we assume that adding the fixed effects does not change the base return. The fixed effects estimates reveal a similar pattern to the previous results, although there do not seem to be pronounced differences for students in deciles 7-9. These results also suggest that very high scoring students may benefit more than the OLS or 2SLS results indicate.

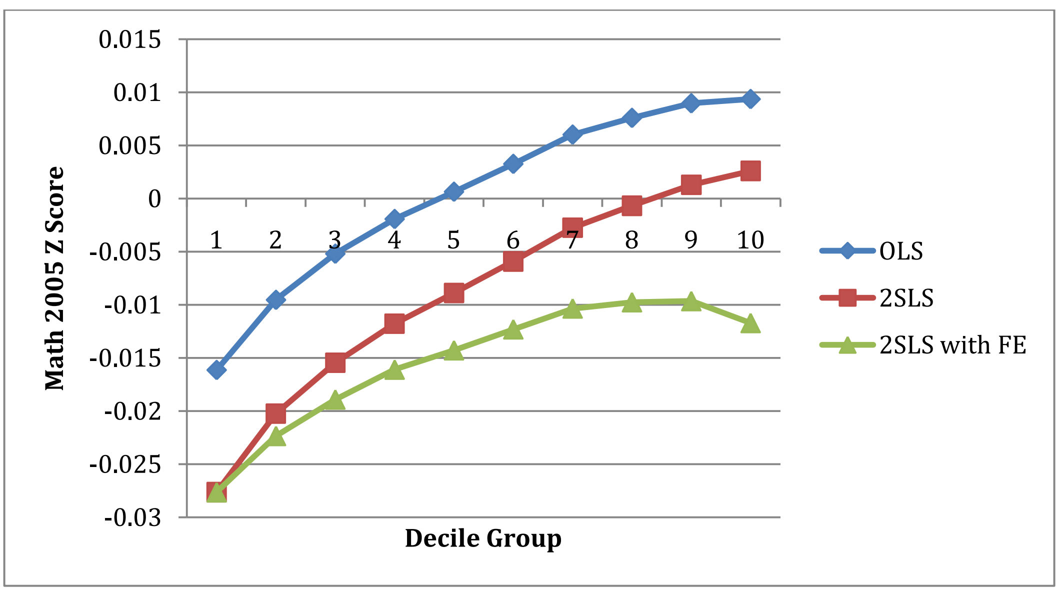

Figure II.2. Class size effect by decile groups with fixed effects (math score).

Figures II.3 and II.4 report results for the class size heterogeneity estimations for the reading exam. The conclusions are qualitatively similar to the math results. Again, I find evidence of heterogeneity across the score distribution, with students in the lowest decile being the prime beneficiaries of smaller classes. All OLS decile estimates are statistically different from each other, except deciles 9 and 10. All 2SLS decile estimates are significantly different from each other, except deciles 8 and 9. The fixed effect results reveal more of an inverse-U pattern, suggesting that smaller classes are most effective for students at the high end of the distribution, in addition to students at the very bottom of the distribution.

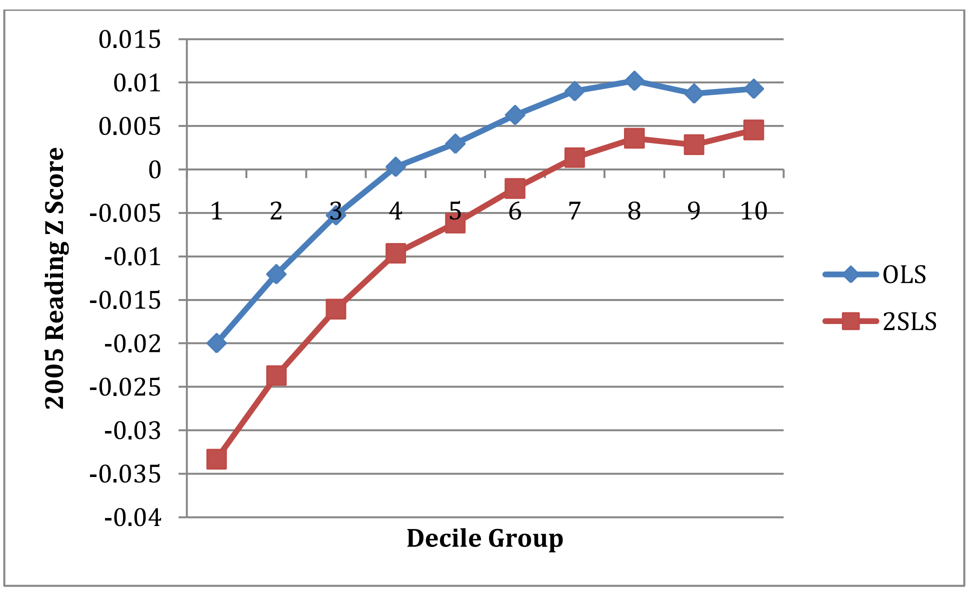

Figure II.3. Class size effect by decile groups (reading score).

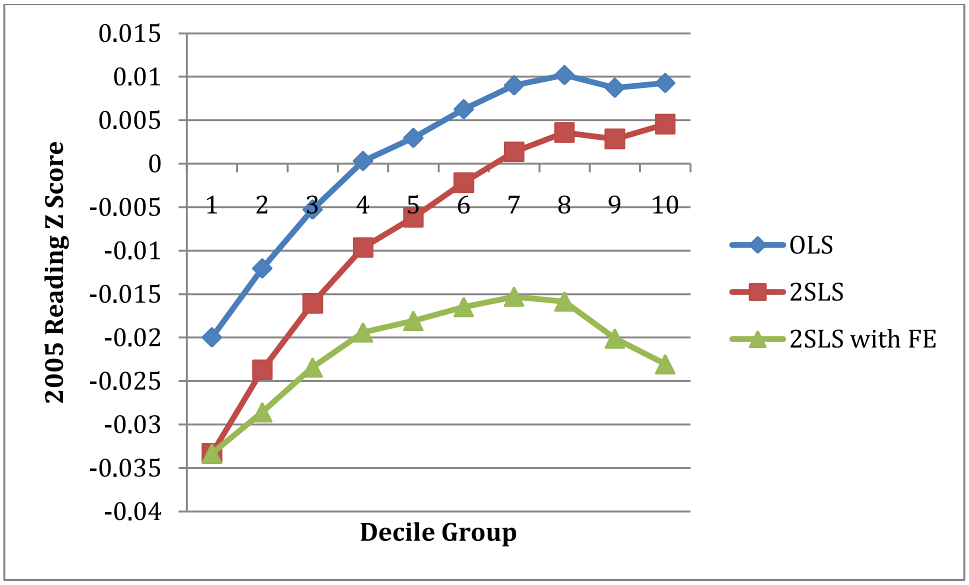

Figure II.4. Class size effect by decile groups with fixed effects (reading score).

II.7 Conclusions and Future Work

A clear analysis of the effects of class size reductions is necessary for researchers to be able to make policy recommendations regarding the allocation of school researchers. While many states have implemented large-scale programs aimed at reducing class size, there is no strong consensus in the economics literature about the efficacy of such programs. One reason may be that the marginal productivity of a class size reduction varies across students, or that smaller classes are more beneficial for certain groups of students than others.

I propose a model that allows the returns to smaller classes to vary across a student score distribution. Using student-level data from Texas and employing maximum class size rules as an instrument, I find strong evidence of heterogeneity in the returns to smaller classes across the distribution. I find that while the average effect of class size is small, students who are in the lowest previous score deciles are particularly responsive to decreases in class sizes. This result holds for both the math and reading versions of the exam.

This study provides evidence that the marginal productivity of a particular input (in this case, class size), is heterogeneous in its effect on different students. Future work will examine the possibility of heterogeneous effects for other school inputs, such as teacher education, teacher experience, or school expenditure.

CHAPTER III

SWITCHING TO BETTER SCHOOLS:

A MODEL OF CLASS SIZE AND MOVER ENDOGENEITY

III.1 Introduction

In 2002, the Florida state legislature passed an amendment requiring schools across the state to reduce class sizes. The new requirements, which created class size caps of 18 for kindergarten through third grade, 22 for fourth grade through eighth grade, and 25 for high school, were designed to be implemented in stages. Since 2002, the state has spent $16 billion in an effort to reduce class sizes and will need to spend billions more if it intends to fully comply with the law by the 2010-2011 school year.

The Florida legislation, while it is the largest statewide effort to reduce class size, is certainly not the first. In 1996, the state of California voted to allocate $1 billion per year to set a maximum class size of 20 for students in kindergarten through third grade. Other states, such as Wisconsin and Tennessee, have participated in large-scale class size reduction efforts, and forty states currently have some type of policy limiting class size or student-teacher ratios for some or all grades.

Although extensive state and federal funds have been allocated for class size reduction policies, the economics literature identifying the benefits of these policies does not provide a strong basis of support for them. Both the individual studies of smaller classes and the meta-analyses of these studies yield contradictory results. Some of these contradictions are rooted in the endogeneity caused by students who switch from one school to another. Because students and their families are able to transfer schools by choosing to locate in a different area, students are not randomly distributed across a region. Moreover, movers account for a substantial subset of the population. In Texas, almost 33 percent of students transfer schools at least once between third and sixth grades, excluding those who move because of a forced transition to middle school,[7] and about 15 percent of students in any given grade are movers. Given that such a large portion of the population engages in school transition and many of these moves may be driven by school characteristics, it is important to fully consider the moving decision when analyzing the class size effect, or the effect of any school input.

In this paper, I propose a model that allows for two types of movers: endogenous movers who switch schools because of a desire to increase school quality, and exogenous movers who switch schools for some reason unrelated to school (such as a change in family structure or parents' employment). Different types of school changes—transfers within district, transfers across districts, and transfers across regions—are used to identify endogenous and exogenous movers. Using student-level standardized test data from Texas and school-level zip code characteristics, I implement a maximum likelihood model to simultaneously

[7] This statistic is based on Texas public school students who are third graders in 2003-2004 and sixth graders in 2006-2007. It is consistent with Hanushek et al. (2004), who claim that one third of students switch schools at least once between fourth grade and seventh grade.

estimate the class size effect with the decision to switch from one school to another. This simultaneous estimation allows me to actually model the moving decision, rather than to simply control for movers, as previous studies have done. I find that the endogeneity of the moving decision leads to a substantial bias in the estimation of the moving effect in a simple OLS model. I also find that while the class size effect is small on average, it is heterogeneous in its effect on different types of movers. Smaller classes are most beneficial for students who move across districts but within region.

III.2 Current Literature

III.2.1 Class Size Effects

A large portion of the class size discussion centers around meta-analyses of the existing research. Perhaps the most well-known of these studies is the work conducted by Hanushek (1986, 1997). In the more recent study (1997), he analyzes 277 estimates from 59 studies that explore the effect of smaller classes on student achievement and that meet minimal quality requirements. Giving equal weight to each estimate, he finds no evidence of a significant, systematic relationship between class size and student performance.

Krueger (2003) takes issue with Hanushek's method of "vote counting" and contends that the proper procedure would be to give equal weight to each *study* rather than to each *estimate*. He reevaluates Hanushek's analysis and determines that if each published study were given equal weight, the results would support a significant relationship between smaller classes and student achievement.[8]

Outside of the meta-analysis debate, there are several important studies whose methodologies produce interesting (and again, often conflicting) results. Angrist and Lavy (1999) address the potential endogeneity of class size by exploiting the discontinuity caused by maximum class size rules in Israel. They create an instrument for class size based on the Israeli school system's practice of limiting class size to 40 students (see section III.4.2 for a description of the instrument). Their results indicate class size effects for some (but not all) students. They find, for example, that a one student reduction in class size is associated with an increase in average math score of 0.05 points (on a 100 point scale) and an increase in average reading score of 0.13 points. They find significant effects for fourth and fifth graders, but not for third graders.

Hoxby (2000) uses maximum and minimum class size rules in Connecticut, combined with exogenous variation in the population of school-aged children, to identify the class size effect. She finds no evidence that smaller classes increase student achievement, nor does she find that smaller classes are beneficial for schools with a higher proportion of minority or low-income students.

While much of the class size literature uses econometric techniques to deal with the endogeneity problems present in existing data, there are a few studies that

[8] For a detailed description of both Hanushek and Krueger's methodologies and arguments, see *The Class Size Debate* (Mishel and Rothstein, eds.).

are experimental in nature and involve random assignment. The most important of these is the Tennessee Student Teacher Achievement Ratio (STAR) program, which marked the first large-scale study with an experimental design. Schools in Tennessee were given the option of participating in the program, which would randomly assign students to either small classes (13-17 students) or large classes (22-25 students). These students were tracked from kindergarten through third grade and remained in either small or large classes. Researchers found that students in small classes performed significantly better on a standardized test at the end of kindergarten and that the score gap continued, but did not widen, as those students continued through third grade.

The results of the study have been interpreted in several different ways. Many researchers point to the significant class size effect as strong evidence that smaller classes make a substantial difference in improving student performance. In fact, this study was (and continues to be) one of the driving forces behind much of the state legislation aimed at reducing class size. However, other researchers highlight the fact that while class size may be important for children in kindergarten, if the effect persisted in higher grades, the score gap should increase as children remained in small classes.

III.2.2 Moving Effects

Much of the moving literature suggests that moving entails a substantial disruption cost for students and therefore has a negative effect on test score.

Ingersoll et al. (1989) find lower average achievement for movers when compared with nonmovers, especially for students in early grades. They find that although the effect is mitigated by controls for socioeconomic status, substantial differences in achievement still exist between the two groups. Kain and O'Brien (1999) estimate the effect of different types of moving on reading score using data from Texas. They examine five types of moves (voluntary campus moves, structural campus moves, district moves, into sample state moves, and out of sample state moves), and find that most types of moves have a negative impact on score. Voluntary campus moves and state moves have the largest effects.

Although many papers suggest a negative mobility effect, some research, particularly work on school choice, does report the opposite result. For example, Cullen et al. (2005) explore the effects of mobility caused by the open enrollment system in Chicago Public Schools. They find that high school students who move schools are more likely to graduate than their counterparts who remain in their previous schools. However, the students who are more likely to switch schools appear to be systematically different from those who stay, leading to a spurious correlation. (One exception is students who transfer to career academies; these students seem to genuinely benefit from the transfer choice.)

Hanushek et al. (2004) consider students who transfer schools and attempt to disentangle the disruption cost of moving from the (presumably positive) effect of Tiebout movers, who move to better schools. They use student-fixed effects and prior-year moves to identify the two effects. They find that within-district movers incur the highest disruption costs from moving, while students who move across districts but remain in the same region benefit from significantly higher school quality. They also examine the negative externalities movers impose on other students, a problem which appears to be greater for minorities and economically disadvantaged students.

III.3 A Two-Type Model of Moving

A significant concern in modeling the class size effect arises because, to a large degree, students and their parents are able to choose which school they attend by choosing where they want to live. According to the Tiebout choice model, individuals sort themselves into communities based on their preferences, and—at least for most families—schooling options are a significant component of the local community. As a result, students are not randomly distributed across a region; instead, many choose to move to a particular area specifically because of the school characteristics it offers.

This endogeneity problem is complicated by the fact that students move schools for many different reasons. While some families locate in an area to allow their children to attend a preferred school, others move because of a change completely unrelated to schooling. Students may switch schools because one of their parents gets a new job in a different part of the state or because of a change in family structure, such as a divorce. Such moves are likely to be exogenous to school characteristics. A complete model of class size should allow for different types of movers in a simultaneous estimation of students' decisions to switch schools, along with the class size effect.

The following model allows students to belong to one of two types—the endogenous type, whose families move schools because of an unobserved school characteristic that may affect test score; and the exogenous type, whose families move schools because of a reason unrelated to school characteristics. A picture of this two type model is shown in Figure III.1. Student *i* belongs to the endogenous type (type N) and moves schools between periods *t-1* and *t* if the following equation holds:

$$move_{Nit} = 1(Z_i \eta_N + \varepsilon_{Ni} > 0), \quad \text{(III.1)}$$

where Z_i is a vector of school and neighborhood characteristics that effect student *i*'s decision to move and ε_{Ni} is the error term. A similar equation describes the moving decision of a student belonging to the exogenous type (type X):

$$move_{Xit} = 1(Z_i \eta_X + \varepsilon_{Xi} > 0), \quad \text{(III.2)}$$

The effect of class size *cs* of student *i* on his test score *s* in period *t* is given by

$$s_{it} = \rho s_{it-1} + X_{it}\beta + \gamma_1 cs_i + (\gamma_{2N} move_{Nit} + \gamma_{2X} move_{Xit}) cs_i + u_i, \quad \text{(III.3)}$$

where s_{it-1} is his score in the previous period, X_{it} is a vector of student-specific characteristics affecting test score, and u_i is the error term. The class size effect is allowed to differ between movers and non-movers, and also between movers of the two types.

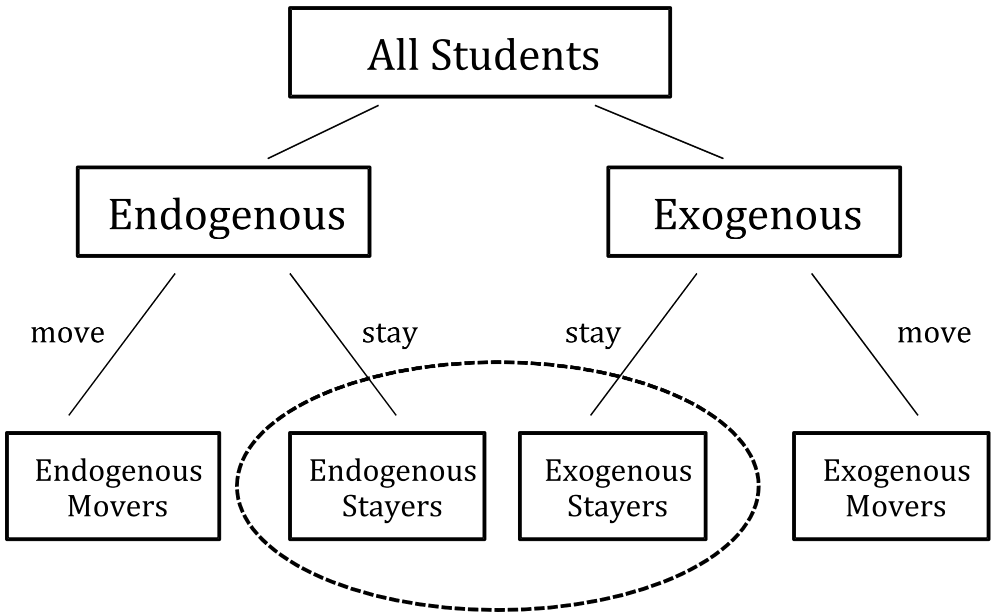

Figure III.1. Two-type mover model.

If type N movers switch schools based on unobserved schooling characteristics that affect test score, then the error from the moving equation ε_{Ni} and the error from the score equation u_i will be correlated; however, the correlation between the two error terms should be zero for type X movers:

$$cov(\varepsilon_{Ni}, u_i) \neq 0 \quad \text{(III.4)}$$

$$cov(\varepsilon_{Xi}, u_i) = 0 \quad \text{(III.5)}$$

The relationship between the two errors for the endogenous movers is given by

$$\varepsilon_N = \lambda_N u + v_N, \quad \text{(III.6)}$$

where $v_N \sim N(0, \sigma^2_{vN})$ and the variance of the error is

$$\sigma^2_{vN} = 1 - \lambda^2_N \sigma^2_u \quad \text{(III.7)}$$

Although I assume type X movers are exogenous, I test equation (III.5) empirically by allowing the errors to be correlated. Therefore,

$$\varepsilon_X = \lambda_X u + v_X, \quad \text{(III.8)}$$

where $v_X \sim N(0, \sigma^2_{vX})$ and the variance of the error is

$$\sigma^2_{vX} = 1 - \lambda^2_X \sigma^2_u. \quad \text{(III.9)}$$

When type X movers are truly exogenous, $\lambda_X = 0$, and the model collapses to the original assumption in equation (III.5).

As shown in equations (III.1) and (III.2), both of the two types of students can choose to either move schools or to remain in their current school. Therefore, each student falls into one of four categories: endogenous movers, endogenous stayers, exogenous movers, and exogenous stayers. The probability that a type N student i chooses to switch schools conditional on his test score s_{it} is given by the following equation:

$$\Pr(move_{Nit} = 1 \,|s_{it}) = \Pr(Z_i\eta_N + \varepsilon_{Ni} > 0|s_{it}). \tag{III.10}$$

Substituting with equation (III.6) and assuming v_1 is normally distributed as described above yields

$$\Pr(move_{Nit} = 1 \,|s_{it}) = \Pr(Z_i\eta_N + \lambda_N u + v_N > 0|s_{it})$$
$$= \Phi\left(\frac{Z_i\eta_N + \lambda_N u}{\sqrt{1-\lambda_N^2\sigma_u^2}}\right). \tag{III.11}$$

Finally, I substitute the error term u_i from equation (III.3) to obtain:

$$\Pr(move_{Nit} = 1 \,|s_{it})$$
$$= \Phi\left(\frac{Z_i\eta_N + \lambda_N(s_{it} - \rho s_{it-1} - \gamma_1 cs_i - (\gamma_{2N}move_{Nit})cs_i)}{\sqrt{1-\lambda_N^2\sigma_u^2}}\right). \tag{III.12}$$

The probability that a type N student *i* chooses *not* to switch schools is

$$\Pr(move_{Nit} = 0 \,|s_{it})$$
$$= 1 - \Phi\left(\frac{Z_i\eta_N + \lambda_N(s_{it} - \rho s_{it-1} - \gamma_1 cs_i - (\gamma_{2N}move_{Nit})cs_i)}{\sqrt{1-\lambda_N^2\sigma_u^2}}\right). \tag{III.13}$$

The other two conditional probabilities can be found in the same way. The probability that a type X student *i* switches schools is given by

$$\Pr(move_{Xit} = 1 \,|s_{it})$$
$$= \Phi\left(\frac{Z_i\eta_X + \lambda_X(s_{it} - \rho s_{it-1} - \gamma_1 cs_i - (\gamma_{2X}move_{Xit})cs_i)}{\sqrt{1-\lambda_X^2\sigma_u^2}}\right), \tag{III.14}$$

and the probability that a type X student *i* chooses to stay at his current school is

$$\Pr(move_{Xit} = 0 \,|s_{it})$$

$$= 1 - \Phi\left(\frac{Z_i\eta_N + \lambda_N(s_{it} - \rho s_{it-1} - \gamma_1 cs_i - (\gamma_{2N} move_{Nit})cs_i)}{\sqrt{1-\lambda_N^2\sigma_u^2}}\right). \quad \text{(III.15)}$$

I can now multiply each of the conditional probabilities by the density of u to obtain each of the final densities:

Case 1: $f(s_{it}, move_{Nit} = 1)$

$$= \Phi\left(\frac{Z_i\eta_N + \lambda_N(s_{it} - \rho s_{it-1} - \gamma_1 cs_i - (\gamma_{2N} move_{Nit})cs_i)}{\sqrt{1-\lambda_N^2\sigma_u^2}}\right) \quad \text{(III.16)}$$

$$* \frac{1}{\sigma_u}\phi\left(\frac{s_{it} - \rho s_{it-1} - \gamma_1 cs_i - (\gamma_{2N} move_{Nit})cs_i}{\sigma_u}\right)$$

Case 2: $f(s_{it}, move_{Nit} = 0)$

$$= \left[1 - \Phi\left(\frac{Z_i\eta_N + \lambda_N(s_{it} - \rho s_{it-1} - \gamma_1 cs_i - (\gamma_{2N} move_{Nit})cs_i)}{\sqrt{1-\lambda_N^2\sigma_u^2}}\right)\right] \quad \text{(III.17)}$$

$$* \frac{1}{\sigma_u}\phi\left(\frac{s_{it} - \rho s_{it-1} - \gamma_1 cs_i - (\gamma_{2N} move_{Nit})cs_i}{\sigma_u}\right)$$

Case 3: $f(s_{it}, move_{Xit} = 1)$

$$= \Phi\left(\frac{Z_i\eta_X + \lambda_X(s_{it} - \rho s_{it-1} - \gamma_1 cs_i - (\gamma_{2X} move_{Xit})cs_i)}{\sqrt{1-\lambda_X^2\sigma_u^2}}\right) \quad \text{(III.18)}$$

$$* \frac{1}{\sigma_u}\phi\left(\frac{s_{it} - \rho s_{it-1} - \gamma_1 cs_i - (\gamma_{2X} move_{Xit})cs_i}{\sigma_u}\right)$$

Case 4: $f\left(s_{it}, move_{Xit} = 0\right)$

$$= \left[1 - \Phi\left(\frac{Z_i \eta_N + \lambda_N (s_{it} - \rho s_{it-1} - \gamma_1 cs_i - (\gamma_{2N} move_{Nit}) cs_i)}{\sqrt{1 - \lambda_N^2 \sigma_u^2}}\right)\right]$$

$$* \frac{1}{\sigma_u} \phi\left(\frac{s_{it} - \rho s_{it-1} - \gamma_1 cs_i - (\gamma_{2X} move_{Xit}) cs_i}{\sigma_u}\right) \tag{III.19}$$

Given equations (III.16)-(III.19), the log likelihood function for the maximum likelihood estimation is

$$\ell_i = \sum_{k=1}^{4} \sum_{i=1}^{N_k} \log(Case_k). \tag{III.20}$$

III.4 Empirical Implementation of the Type-Specific Model

III.4.1 Moving Model

In order to empirically estimate this model, I must first identify which students belong to the endogenous type and which ones belong to the exogenous type. My dataset, which is described in section III.5, allows me to link a student to his campus, district, and region in each period. Using this information, I create three groups of students: campus movers (students who move schools within a district), district movers (students who move across districts but remain in the same region), and region movers (students who move across regions of the state).

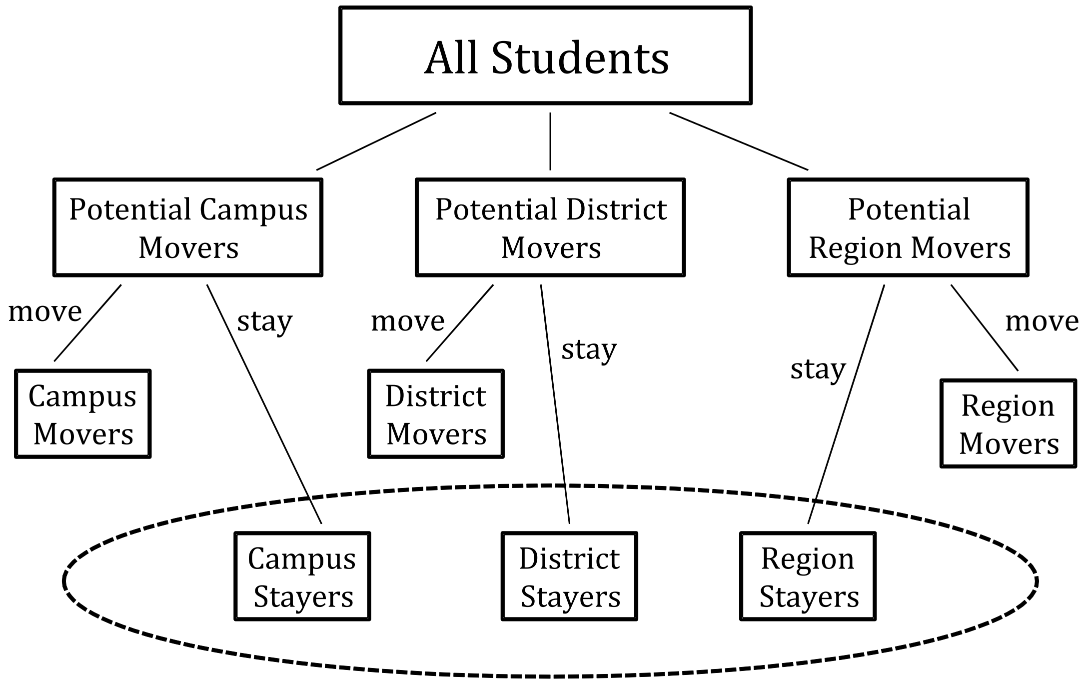

Figure III.2. Three-type mover model.

It is likely that students in these different groups move for different reasons. Endogenous movers, who switch schools because of school-related characteristics, are likely to be students who move within region. A student who is unhappy with his current school will probably look first for a different option nearby; it is unlikely that he and his family will move across the state to find a suitable alternative. Exogenous movers who switch schools because of some reason unrelated to school, such as a parent's new job or a change in family structure, are more likely to include those students who move across regions.

Students who move within a district are probably endogenous movers, but as Hanushek et al. (2004) conclude, the effects of moving for these students are different than the effects for across-district, within-region movers. This is because students who change schools but remain in the same district are still subject to the same common administration and financing they experienced prior to the move.

In the empirical implementation, I alter the two-type model described above to include instead the three groups of students described above: campus movers (type 1), district movers (type 2), and region movers (type 3).[9] A picture of this three-type model is shown in Figure III.2. I simultaneously estimate students' decisions to make these three different types of moves with the class size effect. The moving decisions are given by the following three equations:

[9] Note that the three types are mutually exclusive. Although district movers obviously also move campuses, I define "campus mover" to mean students who move campuses within a district, "district mover" to mean students who move districts within a region, and "region mover" to mean students who move across regions.

Campus Mover: $$move_{1it} = 1(Z_i\eta_1 + \varepsilon_{1i} > 0) \quad \text{(III.21)}$$

District Mover: $$move_{2it} = 1(Z_i\eta_2 + \varepsilon_{2i} > 0) \quad \text{(III.22)}$$

Region Mover: $$move_{3it} = 1(Z_i\eta_3 + \varepsilon_{3i} > 0) \quad \text{(III.23)}$$

The score equation is given by

$$\begin{aligned} s_{it} &= \rho s_{it-1} + X_{it}\beta + \gamma_1 cs_i \\ &\quad + (\gamma_{21} move_{1it} + \gamma_{22} move_{2it} + \gamma_{23} move_{3it})cs_i + u_i, \end{aligned} \quad \text{(III.24)}$$

where returns to class size are allowed to vary across moving type. The densities for the empirical estimation are derived as they are for the two type model above. Although I assume that region movers are exogenous, I allow for the correlation of the moving error and the score error for all three types so that I can test the exogeneity assumption empirically. (See Appendix A for a complete description of the three type model.)

III.4.2 Class Size

An additional concern in the estimation process is that the variable I use for class size is *average* class size within a grade, within a school. This introduces measurement error, because a student's actual class size is not equal to the average class size unless schools divide students into classes of exactly equal size. Consider the true class size model, represented by cs_{ij}^*:

$$cs_{ij}^* = \overline{cs_j} + \varepsilon_{ij} \quad \text{(III.25)}$$

Then the true score equation, simplified for convenience, is given by

$$\begin{aligned} s_{ijt} &= \rho s_{ijt-1} + X_{it}\beta + \gamma cs_{ij}^{*} + u_{ij} \\ &= \rho s_{ijt-1} + X_{it}\beta + \gamma(\overline{cs_j} + \varepsilon_{ij}) + u_{ij} \qquad \text{(III.26)} \\ &= \rho s_{ijt-1} + X_{it}\beta + \gamma\overline{cs_j} + \gamma\varepsilon_{ij} + u_{ij} \end{aligned}$$

The measurement error will cause the typical attenuation effect if the following condition holds:

$$cov\left(\overline{cs_j}, \varepsilon_{ij}\right) < 0 \qquad \text{(III.27)}$$

The covariance between these terms would be negative if schools with higher average class sizes have less ability to create individual classes of different sizes. This is likely true because Texas has a maximum class size rule of 22 for all classes in kindergarten through grade 4. A school with an average class size of 21 or 22 is constrained in the way it can move students between classes. There is not much leeway to create any small classes within the grade because the extra students would push the larger classes over the maximum size. However, a school with a smaller average class size of 17 or 18 would have more room to create relatively larger and smaller classes, causing a negative covariance between average class size and the error term. Under this assumption, the class size coefficient will be biased towards zero.

To mitigate the endogeneity caused by this problem, I create an instrument for class size based on maximum class size rules. This type of instrument was first introduced by Angrist and Lavy (1999), who use Israel's maximum class size rule of

40 to create a class size IV. The maximum class size rule of 22 in Texas creates a discontinuity that can be exploited to create a predicted class size variable, based on school enrollment. I define predicted class size in school *j* as

$$pcs_j = \frac{e_j}{\text{int}\left(\frac{e_j - 1}{22}\right) + 1}, \tag{III.28}$$

where e_j is the total enrollment of school j within a grade. The expression $\text{int}\left(\frac{e_j-1}{22}\right)$ represents the largest integer that is less than or equal to the value in parentheses.

Discontinuities arise in pcs_j when enrollment increases to the point where the maximum class size rule is triggered and the school must create an additional class. For example, if e_j is equal to 44, then pcs_j is equal to 22. The students are divided into two classes, each with 22 students. However, as soon as e_j increases to 45, the maximum class size rule is triggered, and a third class must be added. Now pcs_j is equal to 15; the 45 students are divided into three classes, each with 15 students. This variable can be used as an instrument for class size; even if the maximum class size rules are not strictly enforced for every class, pcs_j should be correlated with cs_j, yet there is no reason to believe it will be correlated with the error in the score equation.

To simplify the model, I run the first stage of the 2SLS procedure first and obtain the predicted value of class size, $\widehat{cs_j}$. I then use $\widehat{cs_j}$ as the key variable of interest in the maximum likelihood estimation.

Table III.1
Summary statistics of fourth grade students in 2004-2005

	Mean	St Dev	Min	Max	Obs
Class Size	18.612	4.44	1	48.4	131371
Move	0.1490	0.3561	0	1	131371
Move Campus	0.0744	0.2624	0	1	131371
Move District	0.0543	0.2266	0	1	131371
Move Region	0.0203	0.1411	0	1	131371
Math Scale Score	2263.45	187.15	1280	2684	131371
Previous Math Scale Score	2251.97	178.05	1228	2699	131371
Female	0.5086	0.4999	0	1	131371
Asian	0.0153	0.1229	0	1	131371
Black	0.1317	0.3382	0	1	131371
Hispanic	0.3599	0.4800	0	1	131371
Free Lunch	0.3904	0.4879	0	1	131371
Gifted/Talented	0.0320	0.1759	0	1	131371
Special Ed	0.0058	0.0761	0	1	131371
ESL	0.0186	0.1351	0	1	131371
Bilingual	0.0282	0.1654	0	1	131371
Average School Score	2212.97	113.746	1324.81	2515.83	131371
Variance School Score	64829.78	46268	112.5	419740	131371
Charter	0.0044	0.0660	0	1	131371
Accountability Rating	2.5516	0.6797	1	4	131371

III.5 Student-level Data from Texas

III.5.1 Moving Data

For the estimation, I use student-level data from the Texas Assessment of Knowledge and Skills (TAKS) obtained from the Texas Education Agency (TEA). Summary statistics for all variables are shown in Table III.1. I use the universe of fourth grade students who attended public school in Texas during the 2004-2005 school year and who also appear in the dataset during the previous school year.

The data allows a researcher to track a student to his grade, campus, district, and region (but not to his specific class). I use the campus, district, and region identifiers from the 2003-2004 and 2004-2005 school years to construct the following moving variables:

$$move_{1it} = 1 \text{ if } campus_{it} \neq campus_{it-1} \text{ and } district_{it} = district_{it-1} \quad \text{(III.29)}$$

$$move_{2it} = 1 \text{ if } district_{it} \neq district_{it-1} \text{ and } region_{it} = region_{it-1} \quad \text{(III.30)}$$

$$move_{3it} = 1 \text{ if } region_{it} \neq region_{it-1} \quad \text{(III.31)}$$

I exclude a student from the type 1 moving group (campus movers) if he is forced to move campuses due to a middle school transition and remains within his district; that is, if the highest grade in $campus_{it-1}$ was third grade. I do not make this adjustment for the other two groups; if a student moves across districts or regions, he is considered a mover even if he was forced out of his previous campus.

The region variable included in the dataset refers to a student's Educational Service Center (ESC) region. Each school district falls into one of the twenty ESC regions across the state. It is important to note that although the twenty regions are general approximations of broad geographic areas across Texas, their specific boundaries are somewhat arbitrary and do not match exactly with more conventional measures of regions, such as MSAs. An alternative method of division would be to link districts directly to MSAs, which might provide a more reasonable approximation of regional labor markets. In the current estimation, however, "region" refers to ESC region as provided in the dataset.

III.5.2 Score Data

I use a student's math score on the TAKS test as the dependent variable in the score regression. The dataset includes a unique student identifier which allows students' scores to be linked from one year to the next, so I also control for a student's previous year test score. By controlling for previous score, my specification becomes a "value added" model.

The TEA reports two types of scores for each student. The first is the raw score, which simply reports the number of questions the student answered correctly. The second score is the scale score, which is a transformation of the raw score used control for the difficulty of the exam across administrations so that scores can be compared across years. For example, scale scores could be used to compare the scores of third graders who took the exam in 2004 with the scores of the next cohort of third graders who took the exam in 2005.

Neither the raw score nor the scale score is meant to be vertically linked; that is, they should not be used to compare a third grader's 2004 score with that same student's fourth grade score in 2005. Because that is exactly the comparison I want to make, I convert scale scores into z scores by subtracting the mean and dividing by the standard deviation. Therefore, the dependent variable in the model is

$$s_{it} = \frac{scale_{it} - \mu_t}{\sigma_t}, \tag{III.32}$$

where $scale_{it}$ is student *i*'s scale score in period *t*, and μ_t and σ_t represent the mean and standard deviation of the scale scores. A student's score is now a representation of where he lies along the distribution of scores. I generate z scores for both the current year and the previous year.

III.5.3 Class Size

As explained in section III.4.2, I use average class size within a grade, within a school to capture class size effects. The TEA only reports average class size; actual class size for each student is not available. Even if comprehensive data on actual class size were reported, I would not be able to identify a specific student's actual class size because the TAKS data tracks a student to his grade and campus, but not to his specific class or teacher. Therefore, I use average class size as reported in the Academic Excellence Indicator System (AEIS). This data, which is reported by campus, is merged with the TAKS student-level data. Enrollment data is also available through AEIS, which I use to generate an instrumental variable for class size, as described in section III.4.2.

Other student-level controls available through the TAKS data include a student's gender, race, ESL status, and bilingual status. There are also variables indicating whether a student is enrolled in a gifted and talented program or a special education program, and whether or not he is eligible for free or reduced lunch. Additional school-level controls, such as a school's accountability rating, are available through the AEIS dataset. In Texas, schools are divided into one of four categories: exemplary, recognized, academically acceptable, and academically unacceptable. This rating, as well as average school score and variance of school score are included as campus-level controls in the score regressions.

III.5.4 Neighborhood Characteristics

The moving decisions described in equations (III.21)-(III.23) create the need for information on neighborhood characteristics, in addition to school characteristics. I use zip code data from the US census, and merge this data with the school characteristics by the zip code of the campus. (One limitation of the study is that I do not have information on students' addresses or zip codes, so I cannot link the zip code data to students and their families directly. However, because most schools are neighborhood schools, I assume that characteristics of the school neighborhood will be similar to characteristics of the student's own neighborhood, even if the zip codes differ.)

The dataset includes demographic and economic variables such as total population, population broken down by race groups, median age, average housing value, and average income per household. I merge these data to a student's campus in both period t and period $t-1$, which allows me to explore which kinds of neighborhoods students leave and which kinds of neighborhoods students enter when they switch schools.

III.6 Results

III.6.1 Differences Between Movers and Nonmovers

Before beginning the regression analysis, I first examine basic differences between movers and nonmovers. Table III.2 reports selected summary statistics by moving type and shows mean comparison tests between movers and nonmovers. The full dataset includes 131,371 fourth graders whose scores can be matched between the 2003-2004 school year and the 2004-2005 school year.[10] Of these students, about 85 percent remain in the same school both years (or switch to a different campus within their district because their 2003-2004 school only served students through the third grade). Of the 15 percent of students who move schools, about 7.5 percent are classified as campus movers (who switch campuses within their district), 5.5 percent are classified as district movers (who switch districts within their region), and 2 percent are classified as region movers (who switch regions within the state of Texas).

[10] Students whose scores cannot be matched may have moved out of the state of Texas or entered a private school. Alternatively, they may still be in Texas public schools but were absent for testing.

Table III.2

Mean comparison tests between nonmovers and movers

	Nonmovers	Campus Movers			District Movers			Region Movers		
	Mean	Mean	Difference from Nonmovers		Mean	Difference from Nonmovers		Mean	Difference from Nonmovers	
Score	2271.1	2224.7	-46.5	***	2203.2	-68.0	***	2244.0	-27.1	***
			(23.61)			(29.87)			(7.43)	
Female	0.507	0.515	0.008		0.518	0.012	*	0.530	0.023	**
			(1.52)			(1.88)			(2.39)	
Asian	0.017	0.008	-0.009	***	0.007	-0.010	***	0.006	-0.010	***
			(6.61)			(6.60)			(4.16)	
Black	0.118	0.204	0.086	***	0.247	0.129	***	0.116	-0.003	
			(24.64)			(31.94)			(0.41)	
Hispanic	0.356	0.433	0.077	***	0.347	-0.009		0.305	-0.050	***
			(15.27)			(1.46)			(5.35)	
White	0.509	0.354	-0.155	***	0.398	-0.112	***	0.571	0.062	***
			(29.43)			(18.29)			(6.35)	
Free Lunch	0.373	0.513	0.139	***	0.486	0.113	***	0.414	0.041	***
			(27.27)			(19.06)			(4.33)	
ESL	0.019	0.023	0.004	***	0.013	-0.006	***	0.010	-0.008	***
			(3.10)			(3.65)			(3.12)	
Bilingual	0.029	0.036	0.008	***	0.015	-0.013	***	0.009	-0.020	***
			(4.27)			(6.68)			(6.08)	
Gifted/Talented	0.036	0.015	-0.021	***	0.006	-0.030	***	0.012	-0.024	***
			(10.99)			(13.41)			(6.66)	
Special Ed	0.006	0.005	-0.001	*	0.002	-0.004	***	0.003	-0.003	**
			(1.73)			(4.69)			(2.11)	
Observations	111800	9772			7131			2668		

Note: Absolute values of t-statistics appear in parentheses beneath the difference.

Table III.3

Mean comparison tests (difference in school characteristics between current school and previous school)

	Nonmovers	Campus Movers			District Movers			Region Movers		
	Current-Previous	Current-Previous	Difference from Nonmovers		Current-Previous	Difference from Nonmovers		Current-Previous	Difference from Nonmovers	
Class Size	0.5661	0.5549	-0.0112		0.5296	-0.0365		0.4478	-0.1183	*
			(0.3102)			(0.8720)			(1.8108)	
Acc Rating	-0.2666	-0.1860	0.0805	***	-0.2141	0.0524	***	-0.2235	0.0430	***
			(10.9971)			(6.0739)			(3.1638)	
Average Score	-25.0487	-23.5163	1.5324	*	-18.7330	6.3157	***	-18.5553	6.4934	***
			(1.6819)			(5.9093)			(3.8617)	
Per. Asian	0.0002	0.0003	0.00003		0.0002	-0.0001		-0.0012	-0.0014	***
			(0.1449)			(0.2281)			(4.2892)	
Per. Black	-0.0003	-0.0033	-0.0030	***	-0.0290	-0.0287	***	-0.0100	-0.0097	***
			(5.4592)			(37.8927)			(11.2458)	
Per. Hispanic	0.0120	0.0056	-0.0065	***	-0.0113	-0.0233	***	-0.0102	-0.0222	***
			(9.9246)			(27.5166)			(19.0056)	
Per. Free Lunch	0.0173	0.0028	-0.0145	***	-0.0244	-0.0418	***	-0.0099	-0.0272	***
			(16.5892)			(33.8047)			(17.0370)	
Per. G/T	0.0125	0.0105	-0.0021	***	0.0083	-0.0043	***	0.0105	-0.0021	***
			(5.1234)			(8.7359)			(2.8932)	
Per. Special Ed	0.0032	0.0023	-0.0008		0.0047	0.0015	**	0.0049	0.0017	*
			(1.5854)			(2.4976)			(1.7808)	

Note: This table examines the mean difference between the current school characteristic and its previous year value. This value is then compared against the same change for nonmovers. Absolute values of t-statistics appear in parentheses beneath the difference.

The mean comparison tests show that, on average, movers of all types have lower testing scores than nonmovers. Campus movers are more likely to be black or Hispanic (when compared to nonmovers), district movers are more likely to be black, and region movers are more likely to be white. Movers of all types are more likely to qualify for free or reduced lunch.

In addition to exploring differences between movers and nonmovers, I also examine which types of schools movers are likely to choose as their destination campuses. Table III.3 provides mean comparison tests for the differences in school characteristics across the two years between movers and nonmovers. For example, change in average school score for campus movers is compared to change in average school score for nonmovers to explore whether students switch to schools that are relatively better or worse in terms of score than their nonmover counterparts.

The data suggest that students move to schools with better accountability ratings and higher scores than their previous schools. (Because average accountability ratings and average test scores both decrease across the period, this means that the destination schools have ratings and scores that are "less low" than they would be had the student stayed in the previous school.) The gap in test score is largest for district and region movers, who move to schools with an average test score drop of only about 19 points, compared with nonmovers, whose schools sustain a 25 point drop on average. Examining differences in race and economic composition is also interesting. Across the period, there is an increase in the percent of Hispanic students across schools, but a relatively smaller increase in the percent of Hispanic students at the destination schools of all types of movers (compared to their origin schools). The percent of students who qualify for free lunch increases across all schools, but the increase is relatively smaller for movers' destination schools. In fact, for district movers, there is actually a net *decrease* in the percent of students who qualify for free lunch of 2.4 percentage points.

Tables III.4A and III.4B show the results of a probit regression for the moving decision using student characteristics as regressors. These more formal results support the generalities suggested by the comparison tests. Table III.4A shows the moving decision in general, in addition to campus, district, and region moves. Movers in general are more likely to have lower previous scores, be black or Hispanic, and qualify for free lunch. Specifically, campus movers are more likely to be black or Hispanic, district movers are more likely to be black, and region movers are more likely to be white. Table III.4B shows campus, district, and region move probit regressions conditional on moving to allow comparisons between moving types. Table III.4B suggests that students who are more likely to move districts have particularly low previous scores, when compared to campus or region movers.

Table III.4A

Probit of moving decision on student characteristics (unconditional campus, district, and region moves)

	Move			Campus Move			District Move			Region Move		
Previous Score	-0.1005	(18.11)	***	-0.0530	(7.89)	***	-0.1149	(15.41)	***	-0.0628	(6.06)	***
Female	0.0104	(1.21)		0.0055	(0.53)		0.0032	(0.28)		0.0275	(1.71)	
Asian	-0.2453	(5.69)	***	-0.1296	(2.48)	**	-0.2069	(3.34)	***	-0.3591	(3.98)	***
Black	0.3029	(22.09)	***	0.3063	(18.30)	***	0.3173	(18.14)	***	-0.2269	(8.15)	***
Hispanic	0.0728	(6.66)	***	0.1778	(13.28)	***	0.0214	(1.45)		-0.1880	(9.22)	***
Free Lunch	0.1311	(12.99)	***	0.1193	(9.72)	***	0.0709	(5.28)	***	0.1124	(5.88)	***
G/T	-0.4095	(12.33)	***	-0.2159	(5.64)	***	-0.5519	(9.47)	***	-0.3672	(5.53)	***
Special Ed	-0.2417	(3.74)	***	0.0036	(0.05)		-0.4793	(4.27)	***	-0.3146	(2.36)	**
ESL	-0.1862	(5.66)	***	-0.0343	(0.93)		-0.2834	(5.79)	***	-0.2633	(3.55)	***
Bilingual	-0.1426	(5.24)	***	0.0482	(1.61)		-0.3185	(7.41)	***	-0.4057	(5.46)	***

Note: Absolute values of z-statistics are in parentheses.

Table III.4B

Probit of moving decision on student characteristics (conditional campus, district, and region moves)

	Campus Move			District Move			Region Move		
Previous Score	0.0670	(5.79)	***	-0.0722	(6.12)	***	0.0041	(0.28)	***
Female	-0.0019	(0.11)		-0.0117	(0.63)		0.0257	(1.12)	
Asian	0.2135	(2.00)	**	-0.0310	(0.28)		-0.3014	(2.19)	**
Black	0.1445	(5.33)	***	0.1665	(6.09)	***	-0.5821	(15.95)	***
Hispanic	0.2752	(12.06)	***	-0.0709	(3.05)	***	-0.3532	(12.53)	***
Free Lunch	0.0488	(2.37)	**	-0.0617	(2.95)	***	0.0160	(0.61)	
G/T	0.4707	(5.27)	***	-0.4619	(4.69)	***	-0.1372	(1.27)	
Special Ed	0.5965	(3.72)	***	-0.5330	(3.03)	***	-0.2241	(1.13)	
ESL	0.3283	(4.57)	***	-0.2629	(3.50)	***	-0.1963	(1.95)	*
Bilingual	0.5168	(8.23)	***	-0.3795	(5.81)	***	-0.4436	(4.51)	***

Note: Absolute values of z-statistics are in parentheses. Campus, district, and region moves are conditional on moving.

These regressions, along with the simple t-test comparisons, reveal stark differences between students who remain in their current schools and those who switch. In addition, it appears that the schools to which they transfer are systematically different from their previous schools. This suggests that rather than simply controlling for students who move, researchers should carefully model the moving decision in conjunction with any school input effects, such as class size.

III.6.2 Class Size Effect: OLS and 2SLS

Before running the full MLE model where the moving decision is modeled, I first report results for the simple OLS and 2SLS using the predicted class size instrument. The first column in Table III.5 reports the results from a simple regression of math score on average class size, controlling for student and school level variables. (In this regression, and in all following specifications, score is measured as a z score, as described in section III.5.2.) The class size effect is negative, small, and insignificant in this specification. Other variables have their expected values. The third column reports results for an identical regression with an added dummy variable equal to 1 if the student moves schools in the previous period. Again, class size is insignificant. The moving variable, however, is significant and negative, suggesting that, on average, movers' scores are 0.0397 standard deviations lower than non-movers.

Table III.5

Effect of class size on math score with and without move dummy

	OLS			2SLS			OLS			2SLS		
Class Size	-0.0018	(0.99)		-0.0212	(2.16)	**	-0.0016	(0.90)		-0.0229	(2.34)	**
Move							-0.0397	(10.99)	***	-0.0394	(10.89)	***
Prev Score	0.4027	(237.84)	***	0.4027	(237.73)	***	0.4019	(237.20)	***	0.4019	(237.07)	***
Female	-0.0149	(5.82)	***	-0.0149	(5.82)	***	-0.0148	(5.79)	***	-0.0148	(5.79)	***
Asian	0.1530	(14.52)	***	0.1535	(14.55)	***	0.1516	(14.38)	***	0.1521	(14.42)	***
Black	-0.0763	(16.58)	***	-0.0742	(15.70)	***	-0.0735	(15.95)	***	-0.0711	(15.05)	***
Hispanic	-0.0133	(3.88)	***	-0.0124	(3.61)	***	-0.0128	(3.77)	***	-0.0119	(3.46)	***
Free Lunch	-0.0523	(16.42)	***	-0.0543	(16.30)	***	-0.0512	(16.06)	***	-0.0533	(16.01)	***
GT	0.2576	(34.55)	***	0.2585	(34.59)	***	0.2554	(34.26)	***	0.2564	(34.31)	***
Sp Ed	-0.1251	(7.43)	***	-0.1237	(7.34)	***	-0.1268	(7.54)	***	-0.1253	(7.44)	***
ESL	0.0853	(8.75)	***	0.0863	(8.84)	***	0.0835	(8.57)	***	0.0847	(8.68)	***
Bilingual	-0.0195	(2.45)	**	-0.0174	(2.17)	**	-0.0209	(2.62)	***	-0.0186	(2.31)	**
Avg Score	0.3329	(90.43)	***	0.3335	(90.26)	***	0.3325	(90.34)	***	0.3331	(90.17)	***
Var Score	0.1460	(65.32)	***	0.1453	(64.23)	***	0.1456	(65.17)	***	0.1449	(64.05)	***
Charter	-0.0331	(1.71)	*	-0.0321	(1.65)	*	-0.0278	(1.43)		-0.0267	(1.38)	
Acc Rating	-0.0072	(4.55)	***	-0.0071	(4.46)	***	-0.0073	(4.59)	***	-0.0071	(4.49)	***
Constant	0.1462	(17.34)	***	0.2324	(5.34)	***	0.1507	(17.87)	***	0.2450	(5.63)	***
R-squared	0.4941			0.4937			0.4945			0.4941		
Cragg-Donald				4523.65						4531.04		
N=131,371												

Note : Absolute values of t-statistics are in parentheses. All non-binary variables are measured in standard deviations from the

Because of the potential measurement error that develops when using average test score, I use Texas' maximum class size rules to generate an instrument for class size, which should correct for the bias in the OLS estimates. The second and fourth columns in Table III.5 report the results from a 2SLS regression of test score on class size, using the predicted class size instrument described in section III.4.2. The fourth column includes an additional moving dummy. As expected, the estimates reveal that OLS appears to be biased upward; estimation with the instrumental variable produces a class size effect that is substantially more negative than the OLS estimation. While these effects are still small, they are both statistically significant and about 12-14 times larger in magnitude than the OLS results. A one standard deviation in class size, or about a four student reduction, results in a score increase of 0.0229 standard deviations.[11]

I also examine class size effects using interactions with different types of movers. Table III.6 reports OLS and 2SLS estimates for class size interacted with three different dummy variables—students who move within their own district (type 1), students who move across districts but within their own region (type 2), and students who move across regions (type 3).

[11] The standard deviation of average class size in the sample is 3.18 students.

Table III.6
Effect of class size on math score (effects by moving types)

	OLS			2SLS		
Class Size	-0.0017	-0.87		-0.0125	-1.13	
CS*MoveC	-0.0017	-0.27		-0.0003	-0.01	
CS*MoveD	-0.0008	-0.1		-0.1277	-3.27	***
CS*MoveR	0.0091	-0.69		-0.0456	-0.87	
MoveC	-0.0155	-0.54		-0.0215	-0.15	
MoveD	-0.0601	-1.72	*	0.502	-2.9	***
MoveR	-0.0776	-1.32		0.1618	-0.7	
Prev Score	0.4017	-237.1	***	0.4019	-236.69	***
Female	-0.0148	-5.8	***	-0.0149	-5.82	***
Asian	0.1514	-14.37	***	0.1521	-14.41	***
Black	-0.0733	-15.9	***	-0.0705	-14.85	***
Hispanic	-0.0132	-3.88	***	-0.0121	-3.52	***
Free Lunch	-0.0514	-16.11	***	-0.0533	-15.96	***
GT	0.2551	-34.22	***	0.2559	-34.19	***
Sp Ed	-0.1276	-7.58	***	-0.1261	-7.48	***
ESL	0.083	-8.52	***	0.0842	-8.61	***
Bilingual	-0.0217	-2.72	***	-0.0203	-2.51	***
Avg Score	0.3323	-90.27	***	0.3325	-89.25	***
Var Score	0.1456	-65.16	***	0.1447	-63.87	***
Charter	-0.0208	-1.07		-0.0068	-0.34	
Acc Rating	-0.0073	-4.6	***	-0.0073	-4.61	***
Constant	0.1514	-16.6	***	0.1991	-4.07	***
R-squared	0.4947			0.4932		
Cragg-Donald				1066.57		
N=131,371						

Note: Absolute values of t-statistics are in parentheses. All non-binary variables are measured in standard deviations from the mean.

Note again that, in most cases, the 2SLS estimates are larger in magnitude than the simple OLS estimates, suggesting that OLS is biased upward. Both the OLS and the 2SLS regressions suggest that the class size effect for non-movers is not significantly different from zero, but that at least one type of movers benefit from smaller classes. The only significant effects are for district movers; a one standard deviation decrease in class size increases math score by 0.128 standard deviations for students who move across districts but remain in the same region. These results suggest the presence of heterogeneity in the effects between movers and non-movers, but also among movers of different types.

III.6.3 Move Effect: OLS and 2SLS

In addition to the class size effect, I also examine the effect of moving on test score. I include dummy variables for each type of move in these regressions to allow a different intercept for each type. The district moving dummy is the only one that significantly affects score. The total move effect, when evaluated at the average level of class size, is -0.036, suggesting that district movers have scores that are 0.036 standard deviations below average. Effects for the other two mover types are similar, although they are not precisely estimated and are not significantly different from zero. These results, which suggest that students incur a substantial cost from moving, are consistent with much of the current literature.[12] However, this relationship may be confounded by the fact that movers are systematically different from nonmovers, as seen in Table III.2. A class size model that incorporates the moving decision will not only provide a clearer picture of the class size effect; it will also allow further exploration into the consequences of the moving choice.

[12] See Kain and O'Brien (1998).

III.6.4 Type-Specific Mover Model

Because students are able to choose which schools they attend through their choice of neighborhood, it is important to model the moving decision along with the class size effect. I allow there to be three types of movers—students who move schools within their district (type 1), students who move schools across districts but remain in the same region (type 2), and students who move schools across regions (type 3). While I hypothesize that type 1 and type 2 students are likely to be endogenous, while type 3 students are likely to be exogenous, I allow for the possibility that all types are endogenous.

I use maximum likelihood estimation to model both the class size effect and the moving decision as described in section III.4.1. The results from the regression are reported in Tables III.7 and III.8 (note that these tables include results from a single estimation, rather than two separate estimations). I first use the predicted class size instrument to obtain $\widehat{cs}_j$ from the first stage regression. I then use $\widehat{cs}_j$ as the key explanatory variable of interest in the MLE estimation.

Table III.7
Effect of class size on math score
(MLE with move dummies)

Class Size	-0.0151	(1.029)	
CS*MoveC	-0.0138	(0.382)	
CS*MoveD	-0.1029	(2.867)	***
CS*MoveR	-0.0397	(0.684)	
MoveC	0.1480	(0.930)	
MoveD	0.3498	(2.149)	**
MoveR	0.1433	(0.558)	
Sigma	0.4630	(454.252)	***
Lambda 1	-0.3927	(8.071)	***
Lambda 2	0.1708	(2.317)	**
Lambda 3	2.0249	(0.323)	
Previous Score	0.4021	(225.237)	***
Female	-0.0149	(5.667)	***
Asian	0.1528	(14.216)	***
Black	-0.0737	(14.486)	***
Hispanic	-0.0147	(4.110)	***
Free Lunch	-0.0554	(14.996)	***
Gifted/Talented	0.2568	(33.052)	***
Special Ed	-0.1276	(7.366)	***
ESL	0.0837	(8.390)	***
Bilingual	-0.0220	(2.667)	***
Average Score	0.3331	(84.656)	***
Variance Score	0.1461	(63.147)	***
Charter	-0.0307	(1.566)	
Acc Rating	-0.0076	(4.572)	***
Constant	0.2070	(3.190)	***
N=131,371			

Note: This regression also contains neighborhood characteristics, shown in Table III.8. Absolute values of t-statistics are in parentheses. All non-binary variables are measured in standard deviations from the mean.

Table III.8
Moving decision with neighborhood characteristics (from MLE)

	Campus Movers			District Movers			Region Movers	
Student Characteristics								
Previous Score	-0.0374	(3.436)	***	-0.1125	(8.898)	***	2.4743	(0.705)
Female	0.0135	(0.873)		0.0096	(0.487)		-0.5705	(0.869)
Asian	-0.2336	(3.243)	***	-0.2466	(2.895)	***	-25.4315	(0.636)
Black	0.3215	(8.059)	***	0.3159	(9.455)	***	8.3019	(0.676)
Hispanic	0.1627	(5.646)	***	0.0150	(0.563)		5.9089	(0.622)
Free Lunch	0.1969	(9.340)	***	0.0499	(2.549)	***	-2.4853	(0.654)
GT	-0.2904	(5.792)	***	-0.6743	(9.191)	***	-4.0099	(0.409)
Special Education	0.0067	(0.066)		-0.6115	(4.148)	***	-11.0801	(0.648)
ESL	-0.1787	(3.124)	***	-0.4554	(6.320)	***	-9.8134	(0.572)
Bilingual	-0.0442	(0.844)		-0.4543	(6.017)	***	-6.2099	(0.747)
School Characteristics: Current								
Average House Value	0.1509	(4.752)	***	-0.1338	(6.195)	***	-1.9376	(0.613)
Income Per Household	0.2400	(6.271)	***	0.4597	(15.188)	***	-7.5867	(0.665)
Median Age	-0.0330	(0.671)		0.1709	(4.183)	***	5.3571	(0.672)
Percent Black (Zip)	1.1191	(11.641)	***	-0.5038	(2.217)	**	-32.6034	(0.621)
Percent Hispanic (Zip)	0.5538	(0.791)		-0.0529	(0.405)		18.3096	(0.778)
Percent Asian (Zip)	-0.4788	(0.325)		-2.5051	(6.022)	***	74.7939	(0.652)
Percent Other (Zip)	1.8751	(1.023)		-0.4275	(2.387)	**	-69.7933	(0.680)
Average Score	0.0371	(1.393)		-0.1279	(4.799)	***	-0.7013	(1.063)
Variance Score	-0.0126	(0.868)		-0.0744	(4.876)	***	-0.3447	(0.348)
Accountability Rating	0.0043	(0.430)		0.0097	(0.921)		-0.6927	(0.509)
Class Size	0.1363	(2.876)	***	0.0029	(0.034)		4.1436	(0.687)

Table III.8 (Continued)

	Campus Movers			District Movers			Region Movers	
<u>School Characteristics: Previous</u>								
Average House Value	-0.1854	(5.921)	***	0.1392	(7.008)	***	-2.9800	(0.630)
Income Per Household	-0.1386	(3.207)	***	-0.5607	(19.917)	***	4.9065	(0.624)
Median Age	-0.5067	(11.100)	***	-0.4309	(10.525)	***	-9.6879	(0.644)
Percent Black (Zip)	-1.1290	(9.002)	***	0.1511	(0.848)		3.3603	(0.383)
Percent Hispanic (Zip)	-0.7650	(1.236)		-0.6656	(3.450)	***	-27.7386	(0.701)
Percent Asian (Zip)	-0.4512	(0.290)		2.6119	(8.577)	***	-30.6996	(0.507)
Percent Other (Zip)	-2.3662	(1.499)		0.5550	(1.216)		68.5499	(0.627)
Average Score	-0.0348	(1.791)	*	-0.1396	(5.641)	***	0.0472	(0.028)
Variance Score	0.0185	(1.346)		-0.0155	(1.065)		-1.4400	(0.966)
Accountability Rating	-0.1183	(11.950)	***	-0.0197	(1.936)		1.4152	(0.765)
Class Size	0.0222	(2.979)	***	-0.0349	(4.897)	***	1.4287	(0.653)
N=131,371								

Note: These results are from the MLE estimation reported in Table III.7. Absolute values of t-statistics are in parentheses. All non-binary variables are measured in standard deviations.

III.6.4.1 Class Size Effect in the Type-Specific Model

Table III.7 shows the heterogeneity in the effect of smaller classes that exists across different types of movers. Like the 2SLS results, the MLE suggests that the only students for whom smaller classes are effective are district movers. A one standard deviation decrease in class size increases expected score by 0.103 standard deviations. [13] While the effects for nonmovers, campus movers, and district movers are all statistically insignificant, it is useful to examine the point estimates. They are all negative and fairly small, suggesting a negative overall effect and slight additional effects for the other two mover types. However, even if these effects were significant, they would still be dwarfed by the much larger effect of the district movers.

This large class size effect for district movers is robust to specification; it is present in both the 2SLS and the MLE. A natural question is why this should be true. What is different about district movers that makes them any more sensitive to smaller classes than other types of students? One possible explanation lies in the composition of these students. Other research suggests that students along different points of the score distribution respond differently to school inputs, such as class size. Previous work shows that students at the lowest end of the distribution gain the most from a decrease in class size. If district movers are composed more heavily of students from the lower end of the distribution, then a greater response to class size should be expected.

[13] Note that all variables (excluding dummy variables) are measured in deviations from the mean.

To examine this type of distribution, I divide students into ten deciles of equal size based on their previous year score. Figure III.3 shows the percent of each group (nonmovers, campus movers, district movers, and region movers) composed of students from the bottom three deciles. This graph suggests that district movers include a disproportionate amount of these bottom decile students. For example, while 10.6 percent of the total population is in the first decile, 16.7 percent of district movers are in this group. (Only 11.9 percent of region movers and 14.2 percent of campus movers are in the first decile.)

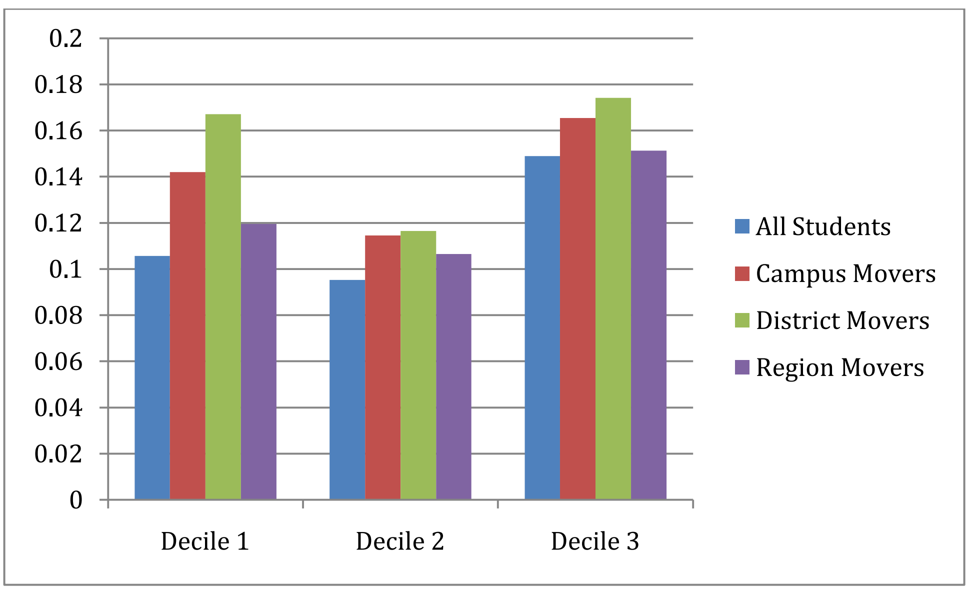

Figure III.3. Percent of students in bottom deciles.

Table III.9
Effect of class size on math score with decile groups (2SLS)

	Without Groups		With Groups	
Class Size	-0.0125	-1.13	0.1159	(5.54) ***
CS*MoveC	-0.0003	-0.01	0.0004	(0.01)
CS*MoveD	-0.1277	-3.27 ***	-0.0878	(2.35) **
CS*MoveR	-0.0456	-0.87	-0.0442	(0.85)
CS*Group1			-0.1838	(10.77) ***
CS*Group2			-0.1605	(9.29) ***
CS*Group3			-0.1414	(8.12) ***
CS*Group4			-0.1267	(7.22) ***
CS*Group5			-0.1148	(6.51) ***
CS*Group6			-0.1040	(5.84) ***
CS*Group7			-0.0943	(5.26) ***
CS*Group8			-0.0976	(5.37) ***
CS*Group9			-0.0871	(4.75) ***
CS*Group10			-0.1081	(5.70) ***
MoveC	-0.0215	(0.15)	-0.0206	(0.15)
MoveD	0.5020	(2.90) ***	0.3297	(1.99) **
MoveR	0.1618	(0.70)	0.1581	(0.69)
Prev Score	0.4019	(236.69) ***	0.3197	(37.28) ***
Female	-0.0149	(5.82) ***	-0.0044	(1.79) *
Asian	0.1521	(14.41) ***	0.1487	(14.60) ***
Black	-0.0705	(14.85) ***	-0.0436	(9.47) ***
Hispanic	-0.0121	(3.52) ***	0.0014	(0.41)
Free Lunch	-0.0533	(15.96) ***	-0.0451	(13.94) ***
GT	0.2559	(34.19) ***	0.2560	(35.32) ***
Sp Ed	-0.1261	(7.48) ***	-0.1087	(6.67) ***
ESL	0.0842	(8.61) ***	-0.0172	(1.74) *
Bilingual	-0.0203	(2.51) ***	-0.0204	(2.56) ***
Avg Score	0.3325	(89.25) ***	0.3097	(85.67) ***
Var Score	0.1447	(63.87) ***	0.1359	(61.96) ***
Charter	-0.0068	(0.34)	-0.0133	(0.69)
Acc Rating	-0.0073	(4.61) ***	-0.0100	(6.50) ***
R-squared	0.4932		0.5322	
Cragg-Donald	1066.57		300.033	
N=131,371				

Note: Absolute values of t-statistics are in parentheses. All non-binary variables are measured in standard deviations from the mean.

Table III.9 presents the 2SLS result from Table III.6, compared to the same regression with interaction effects between class size and the decile groups. An examination of the group coefficients confirms that students in the lowest deciles do gain the most from smaller classes. In addition, the class size coefficient for district movers has decreased in magnitude from -0.1277 to -0.0878, a reduction of almost one third. This suggests that a substantial portion of the class size effect for district movers can be accounted for by the composition effect, although there is clearly still a separate district effect.

III.6.4.2 Moving Effect in the Type-Specific Model

The chief advantage of the type-specific model I propose is that it allows the moving decision to be estimated simultaneously with the class size effect while accounting for different types of moves. I categorize students into endogenous and exogenous movers based on the types of transfers they make (campus moves, district moves, or region moves). Recall that the original hypothesis is that type 1 and type 2 movers should be endogenous, while type 3 movers should be exogenous. While I make this assumption, I allow for the possibility that all types are endogenous.

Table III.7 reports estimates for λ_1, λ_2 , and λ_3, which show how the errors of the score equation and the moving equation are correlated. I find that estimate for λ_1 is significant and negative, indicating endogeneity for type 1 movers. This estimate suggests that students who are likely to mover campuses perform poorer than expected on the standardized math test. The estimate for λ_2 is significant and positive, suggesting that students who are likely to be district movers perform better than expected on the standardized test. This is consistent with the findings in Hanushek et al. (2004), which suggest that district movers are Tiebout movers. I find that λ_3 is not significantly different from zero, which suggests that type 3 movers are likely to be exogenous. This makes sense intuitively; families who move regions across the state of Texas are more likely to be motivated by some outside factor like a change in job or family structure, rather than because of schooling.

Table III.10
Comparison of the Move Effect

	2SLS			MLE		
Class Size	-0.0125	(1.13)		-0.0151	(1.03)	
CS*MoveC	-0.0003	(0.01)		-0.0138	(0.38)	
CS*MoveD	-0.1277	(3.27)	***	-0.1029	(2.87)	***
CS*MoveR	-0.0456	(0.87)		-0.0397	(0.68)	
MoveC	-0.0215	(0.15)		0.148	(0.93)	
MoveD	0.502	(2.90)	***	0.3498	(2.15)	**
MoveR	0.1618	(0.70)		0.1433	(0.56)	
Overall Move Effect Evaluated at Average Class Size						
MoveC	-0.0228			0.0867		
MoveD	-0.0650			-0.1071		

The endogeneity of campus and district movers causes the estimates of the moving effect to be biased in the previous OLS and 2SLS regressions. Table III.10 compares the overall moving effect as estimated in the 2SLS and the MLE models and shows substantial bias in the 2SLS results. Consider first the district movers. Using the 2SLS estimates, the overall moving effect when evaluated at the average value of class size[14] is -0.065. Hanushek et al. (2004) find a similar negative effect of district move on current year score.[15] However, simply controlling for students who move may yield a biased result. The results from the MLE suggest that district movers are endogenous; students who are more likely to be district movers perform better than expected on the math exam. Because of this endogeneity, the moving coefficient of -0.065 is overestimated compared to the MLE, which is -0.1071.

The opposite result occurs for the campus movers, although the overall effects in both the OLS/2SLS and MLE are not significantly different from zero. Campus movers are endogenous and negative; students who are more likely to be campus movers perform worse than expected on the math exam. This endogeneity causes the overall moving effect in the 2SLS (-0.0228) to be underestimated relative to the MLE (0.0867), although both estimates are statistically insignificant.

There is not a large difference in the results for the regional movers, whose moving decisions seem to be exogenous. Both estimates are insignificant, but

[14] Class size here is measured in standard deviations. The average value is 4.44.
[15] Hanushek et al. (2004) find that district movers have scores that are 0.095 standard deviations lower than nonmovers.

similar in magnitude. The 2SLS point estimate of moving is -0.0407, compared with an estimate of -0.033 in the MLE.

III.6.4.3 Student and Neighborhood Characteristics in the Moving Decision

Table III.8 includes the student and neighborhood characteristics from the moving decision, which I model simultaneously with the class size effect. (Table III.7 and Table III.8 report results from the same estimation.) A different moving equation is estimated for each of the three moving types. In addition to student characteristics, the explanatory variables include the neighborhood characteristics described in section III.5.4 for both the current campus and the previous campus, as well as school characteristics (average score, accountability rating, and class size) from both campuses. Column 1 reports results for campus movers. Students who switch schools within their own districts are more likely to be low scoring, low income minority students. District movers are similar; they are likely to be black, to qualify for free lunch, and to have low previous scores. None of the characteristics are strong predictors for region movers, which provides further evidence that those students are exogenous movers.

III.7 Conclusions and Future Work

Although movers constitute a large proportion of the student population and although the class size effect is heterogeneous in its impact on movers and nonmovers, previous class size studies have not carefully modeled the moving decision simultaneously with the class size effect. As a result, simple OLS models may result in biased estimates of the moving effect. In this study, I propose a two-type model that estimates students' moving decisions as well as the class size effect. I assume that students may be exogenous or endogenous movers and allow the class size effect to vary across movers of different types.

Using student-level test score and moving data from Texas, I identify three groups of movers—campus movers, district movers, and region movers. I find that while the class size effect is small and insignificant on average, it is particularly important for district movers who switch to schools in another district but remain in the same region. At least part of this explanation lies in the fact that district movers are composed more heavily of lower scoring students, who tend to respond more to smaller classes. The maximum likelihood model which estimates the moving decision reveals that campus and district movers are likely to be endogenous, while students who move across regions are exogenous movers. Most important, the model reveals that the endogeneity of the moving decision causes OLS models that simply control for moving to be biased. The results suggest that the move effect for campus movers is typically underestimated in the OLS, while the effect is overestimated for district movers.

Creating a careful model of the moving decision is necessary not only in understanding school input effects, such as class size, but also in understanding the effect of moving itself on students' academic performance. This paper provides a first step in creating a model that incorporates both class size and the moving decision. Future work will model the endogeneity of movers in a more flexible way by allowing type of move (campus, district, or region) to serve as one of several identifying factors to determine the probability of endogeneity.

CHAPTER IV

DOES SORTING STUDENTS IMPROVE SCORES? AN ANALYSIS OF CLASS SIZE AND CLASS COMPOSITION

IV.1 Introduction

While several studies have found that smaller classes are an effective tool for improving student achievement[16] and many states have implemented class size reduction legislation, a related but lesser-studied issue is how students are actually divided into classes. Schools may use several different strategies to allocate students among different classrooms. Some schools may choose to sort students by ability level and create classes of relatively homogeneous students. Alternatively, schools may choose to sort students with varying abilities evenly across classes. Other schools may try to match students and teachers, while taking into account individual students' learning styles. In some situations, variables such as parents' preferences or students' behavior records may play a part in the class composition process.

Many of these allocation strategies will more effective when schools have more classes into which they can divide students. This is directly related to schools' class size policies. Holding enrollment constant, schools that choose to offer smaller classes must also offer more classes. This may increase the

---

[16] See Hanushek (1997), Krueger (2003), and Mishel and Rothstein (2002) for a general analysis of the class size literature.

effectiveness of schools' sorting strategies, if they choose to track students into groups based on ability or previous testing scores. Therefore, what some studies have classified as a class size effect may be confounded by schools' abilities to sort students into classes.

The purpose of this study is twofold. First, I examine the class size effect across different types of schools—some that appear to sort students into more homogeneous groups, and some who do not. Using a unique dataset from Dallas Independent School District that allows a student to be tracked not only to his school and grade, but also to his actual classroom, I can precisely measure both actual class size and the student dispersion within a class. I attempt to disentangle the class size effect from the sorting effect by constructing several sorting indices which measure the dispersion of students based on observable characteristics, such as previous score and Gifted and Talented (G/T) classification. Second, I explore the basic impact of sorting on student performance. I create a model that analyzes the effect of sorting for several different types of students in order to determine if sorting is beneficial, and, if so, for whom. Using a different grade's sorting index as an instrument for the sorting index within a given school-grade, I find that a higher degree of sorting significantly improves scores for all types of students.

IV.2 Empirical Model

IV.2.1 Class Size and Sorting

With equal enrollments, a school that divides its students into smaller classes necessarily has more classes than a school that has larger classes. Because schools that sort students into different group may be able to sort more efficiently with smaller classes, it is possible that what is typically identified as a class size effect is actually a sorting effect. Consider the following equation:

$$s_{ijt} = \rho s_{ijt-1} + X_{ijt}\beta + \eta cs_{jt} + \varepsilon_{ijt}, \tag{IV.1}$$

where s_{ijt} represents the test score of student *i* in class *j* in time *t*, s_{ijt-1} represents the same student's previous year score, cs_{jt} represents the number of students in class *j*, and X_{ijt} is a vector of student-level controls. In this equation, the degree of sorting within class *j* is unobserved. An endogeneity problem arises if sorting and class size are correlated and sorting affects student performance. If an increased level of sorting, or the creation of more homogeneous classes, improves students' test scores, and if class size and sorting are positively correlated, then the class size coefficient in equation (IV.1) will be biased downward. Alternatively, if more heterogeneous classes benefit students, then the class size coefficient will be biased upward.

To control for the degree of sorting in a specific class, I consider the following equation:

$$s_{ijt} = \rho s_{ijt-1} + X_{ijt}\beta + \eta cs_{jt} + \phi\gamma_{jt} + \varepsilon_{ijt}, \tag{IV.2}$$

where γ_{jt} is a sorting index for a class *j*, describing the dispersion of the students in the classroom based on observable characteristics, such as score. In addition to disentangling the effect of class size and sorting, I am also interested in the sorting effect on its own. In order to examine how sorting may affect different types of students, I allow ϕ to vary by students' observable characteristics. I used several different sorting indices in the analysis, each of which is described in the following section.

IV.2.2 A Sorting Index

A school's decision to sort students, or to track them into different groups, may have different implications for different groups of students. For example, sorting high-scoring students into one class and low-scoring students into another class may allow the classes to move at different paces, which may benefit both groups of students. The teacher in the low-scoring class may be able to focus on foundational skills necessary to the improvement the students, while the teacher in the high-scoring class may have the opportunity to move on to new, more challenging material without the fear of losing the understanding of the class.

However, this type of sorting may not necessarily benefit both groups. An alternative hypothesis is that by creating evenly distributed groups, students with more understanding of the material may be able to help those with less understanding. In this situation, low-scoring students might benefit without causing a cost for high-scoring students. (It may even be plausible that this situation could benefit both high scorers and low scorers.)

The same possibilities hold for sorting based on G/T classification. Some schools may group all G/T students into a single class to allow them to move at their own pace, while other schools may divide them into several classes with other non-G/T students.[17] Having G/T students included in a regular classroom could potentially help or hurt non-G/T students in the same ways that high-scoring students could affect low-scoring students.

To empirically determine the effects of both types of sorting, I first construct a measure defining how "sorted" a class is. I define the following index for each class:

$$\gamma_j = \left|\frac{\sigma_j - \sigma_k}{\sigma_k}\right|, \quad \text{(IV.3)}$$

where σ_j is the standard deviation of the scores within class *j*, and σ_k is the standard deviation of the scores within school *k*, of which class *j* is a member. In the extreme case in which a class is completely sorted, every student in the class has the same prior year score, so σ_j is equal to zero. The parameter σ_k is a measure of the variation in the school as a whole. As σ_j approaches zero, γ_j approaches 1.

In the opposite case, in which students are not sorted at all, there should be no difference in the dispersion of scores within class *j* and the dispersion of scores

[17] Even if G/T students are divided into classrooms with many non-G/T students, they still may be "pulled out" for several hours during the school day or during the week. Unfortunately, the Dallas ISD data contains only one classroom per student, so it is not possible to tell if the students participate in this type of program.

within school k. In this situation, σ_j is equal to σ_k and γ_j is equal to zero. Therefore, $\gamma_j \epsilon [0, 1]$ measures the dispersion of scores within a given class while controlling for overall potential dispersion at the school level.

I define a similar measure to gauge the measure of G/T sorting within a class. I construct the following index:

$$\gamma_j^{GT} = \left| \frac{\sigma_j^{GT} - \sigma_k^{GT}}{\sigma_k^{GT}} \right|, \tag{IV.4}$$

where σ_j^{GT} is the standard deviation within class j of the binary variable indicating G/T status, and σ_k^{GT} is the standard deviation of the same variable within school k. In the extreme case in which all the G/T students are placed into the same class, σ_j^{GT} is equal to zero, and γ_j^{GT} becomes 1.

Alternatively, if G/T students are divided evenly among all classes, then the standard deviation of the G/T variable in each class will be equivalent to the standard deviation of the G/T variable in the school. In this case, γ_j^{GT} becomes zero. Therefore, $\gamma_j^{GT} \epsilon [0, 1]$, where a higher number indicates more sorting.

IV.2.3 Endogeneity of the Sorting Index

It is essential to consider not only the effect of sorting on students' scores but also why they are sorted into their given classes at the outset. Although the dataset allows identification of characteristics such as previous score and G/T status, teachers and principals observe many other variables which may be used to divide students into different classrooms. Principals may attempt to "match" certain students with certain teachers, or they may have policies whereby parents can request a certain teacher for their children.

Unobserved variables like behavior may also play an important role in the classroom assignment process. For example, if a principal observes that several students have had behavior problems in the past, he may try to divide those students evenly across the classes within a grade, or he may assign them to a particular teacher who has had success with behavioral problems in the past. In this case, behavior is an unobserved variable that affects a school's sorting index. However, a student's behavior may also affect his test score, causing an endogeneity problem.

In order to deal with this endogeneity, I create an instrument for the sorting index using other another grade's sorting index. If the administration at school *k* uses certain guidelines in assigning students to classes in grade *g*, it is likely that those guidelines are also used for other grades in school *k*. Therefore, the sorting indices for classes in grade *g* should be correlated with the sorting indices for grade *g-1*. However, there is no reason to believe that the way in which classes are sorted in grade *g-1* should impact the scores of students in grade *g*. Therefore, sorting indices in grade *g-1* should provide valid instruments for sorting indices in grade *g*.

The problem that arises when trying to match indices from individual classes across grades is that there is no way to map the classes from third grade to specific fourth grade classes. Instead, I create a grade-specific sorting measure that can be used for all classes within a grade. I define the following two parameters:

$$\alpha_{1k} = \sqrt{\frac{1}{N}\sum\left(s_{ijk} - \overline{s_k}\right)^2} \tag{IV.5}$$

and

$$\alpha_{2k} = \sqrt{\frac{1}{J}\sum\frac{1}{N_j}\sum\left(s_{ijk} - \overline{s_j}\right)^2} \tag{IV.6}$$

where $\overline{s_k}$ is the score average in school k, $\overline{s_j}$ is the score average in school j, N_j represents the total number of students in class j, N represents the total number of students in school k, and J represents the total number of classes in school k. The parameter α_{1k} is a measure of score dispersion in school k, while the parameter α_{2k} is a measure of score dispersion in classes $j=1,\ldots,J$ of school k. I define the following variable as the sorting index for school k:

$$sort_k = \frac{\alpha_{1k}}{\alpha_{2k}}. \tag{IV.7}$$

Higher values of $sort_k$ indicate less dispersion of scores within classes relative to score dispersion with the school, which means more sorting. Lower values of $sort_k$ indicate more dispersion of scores within classes, which means less sorting. In the empirical estimation, I use $sort_k$ for the third grade to instrument for $sort_k$ for the fourth grade.

IV.3 TAKS Data from Dallas ISD

One drawback to many datasets used to explore the class size effect is that students cannot typically be linked to their actual classes. For example, the Texas Education Agency (TEA) collects student-level testing data from the Texas Assessment of Knowledge and Skills (TAKS) for all public school students starting in third grade. However, while students' schools and grade levels are available in the dataset, their specific classes are not. Therefore, any measure of class size must be an average across all grades within a given school. While class size effects may be identified from across-school variation, the structure of the data leaves little room to explore within-school class composition effects.

While students are not linked to specific classes in the statewide dataset, several school districts do collect student-level data that may be linked to a class variable. I employ a unique dataset from Dallas Independent School District that contains both class and grade identifying information. The dataset includes student-level math TAKS scores for two school years. I examine all third grade students in the 2003-2004 school year who become fourth graders in 2004-2005, a total of 9,325 children from 138 different schools in Dallas ISD. In addition to achievement scores for both years, the dataset contains race and gender variables and identifiers for students qualifying for programs such as free or reduced lunch, Gifted and Talented, Special Education, and Limited English Proficiency. Because the data is available at the class level, I construct actual class size instead of using grade-level averages. Summary statistics are shown in Table IV.1.

Table IV.1
Summary statistics for Dallas ISD

	Obs	Mean	St Dev	Min	Max
Scale Score	9325	2191.02	192.84	1280	2684
Previous Score	9325	2217.07	177.61	1708	2697
Class Size	9325	19.45	3.08	3	27
Black	9325	0.2894	0.4535	0	1
Hispanic	9325	0.6446	0.4787	0	1
Asian	9325	0.0119	0.1085	0	1
Gifted/Talented	9325	0.2479	0.4318	0	1
Free Lunch	9325	0.8571	0.3500	0	1
Special Ed	9325	0.0432	0.2034	0	1
Enrollment	9325	105.28	37.55	9	181
Number of Classes	9325	5.38	1.87	1	12
Gamma (Score)	9322	0.1437	0.1187	0	0.7579
Gamma (G/T)	9264	0.1879	0.2256	0	1

Texas reports students' scores in two ways. The first score is a student's raw score, which corresponds to the number of questions he answered correctly on the exam. For the 2004-2005 exam, the maximum raw score is 42 points. The second score measure is a student's scale score, which is scaled using the Rasch partial credit method to control for the difficulty of the exam across different administrations of the test. Scale scores are used to compare two different cohorts' scores. For example, scale scores could be used to compare fourth graders in 2004 with the following group of fourth graders, who took the exam in 2005.

Although the scores allow for direct comparison in this way, they are not meant to be vertically linked. That is, a third grader's 2004 score should not be directly compared to his fourth grade 2005 score in order to gauge improvement. Because that is precisely the comparison I want to make, I convert the scale scores into z scores, by subtracting out the mean score and dividing by the standard deviation in a given year. A student's z score is given by

$$s_{it} = \frac{scale_{it} - \mu_t}{\sigma_t}, \qquad \text{(IV.8)}$$

where $scale_{it}$ is student *i*'s scale score in period *t*, and μ_t and σ_t represent the mean and standard deviation of the scale scores. A student's score is now a representation of where he lies along the distribution of scores. I generate z scores for both the current year (2004-2005) and the previous year (2003-2004).

IV.4 Empirical Results

IV.4.1 Score Sorting

Before examining any effects of sorting or class size, it is first important to determine whether any schools appear to sort students based on observable characteristics and how prevalent this type of sorting is. I explore potential sorting based on two observable characteristics: previous TAKS math score and a student's Gifted/Talented status. To investigate sorting based on students' previous scores, I create dummy variables for each class and compare the mean scores by running the following regression:

$$s_{ijt-1} = \beta_1 + \sum_{j=2}^{J} \beta_j D_j + \varepsilon_{ij}, \qquad \text{(IV.9)}$$

where s_{ijt-1} is student *i*'s test score in the previous year and D_j is a dummy variable for class *j*. Therefore, β_1 gives the mean score for the first class and $\beta_2, \beta_3, \dots, \beta_J$ show the differences in score relative to the first class. If schools divide their students into classes randomly, then there should be no difference in the previous year score means for any of the classes. That is, $\beta_2, \beta_3, \dots, \beta_J$ should not be significantly different from zero or from each other.

Alternatively, if schools do divide students into classes based on their previous year scores, then there should be significant differences in the average scores. Consider the case in which a school has three classes within a single grade. The administration may choose to sort students into three groups—low-scoring students who need additional math assistance to improve their grades, average-scoring students who are achieving at grade-level, and high-scoring students who are ready to move on to more challenging material. In this case, β_2 and β_3 would be significantly different from zero, as well as different from each other.

I run regression (?) for each of the 138 schools in the district to determine which schools potentially sort by previous year score. The results are reported in Tables B1.1-B1.14 of Appendix B. Consider, for example, the results for school 186, which are given in Table B.1.7. This school has four classes of fourth graders—two with lower average math scores and two with higher average math scores. The average score for class 1, given by the constant, is 26.1 (the maximum raw score is

42 points). The coefficient for class 2 is not significantly different from zero, and the point estimate is only 1.2 points, suggesting that there is no substantial score difference between the two classes. However, the estimates for class 3 and 4 are both statistically significant and indicate a 4.9 point and 6.1 point difference in score from class 1. At least one class dummy variable is significant in 44 of the 138 schools (about 32 percent of schools).

For the other 94 schools, there is no significant difference between the average previous scores. School 109's results, reported in Table B1.1, suggest that there is no statistical difference in the scores of the four classes. The average score for class 1 is 28.8, and the score differences for the other classes range from .09 points to 2.3 points. None of these differences is statistically different from zero. It is important to note that even if score averages are not significantly different, schools may still be considering score in a strategic division of students into classes. Some schools may be purposefully allocating students of different abilities equally among classes. If administrators believe that an equal division of student ability is beneficial to some or all students, then there should be no significant score average score difference between classes, even if the school is acting strategically.

IV.4.2 Gifted/Talented Sorting

In addition to sorting by previous test score, schools may also sort by other observable characteristics, such as whether a student qualifies for a Gifted and Talented (G/T) program. Some schools may try to group all of their G/T students together in a single class, while others may try to disperse them evenly among a number of classes. In order to determine whether being in a certain class predicts the likelihood that a student qualifies as G/T, I run the following probit regression:

$$GT_{ij} = \beta_1 + \sum_{j=2}^{J} \beta_j \, D_j + \varepsilon_{ij}, \qquad \text{(IV.10)}$$

where GT_{ij} is a dummy variable equal to one if a student qualifies for a G/T program. The right hand side of this equation is analogous to equation (IV.9), where D_j is a dummy variable for class *j*.

Consider the example of a school with three classes. In an extreme case, the school may create a single class for only G/T students, in which case, that class dummy would predict G/T status with certainty. However, even if there are some G/T students in all classes, sorting may still exist if they are grouped allocated more heavily in some classes.

It should be noted that schools may face constraints related to which teachers are certified to teach G/T students. For example, if a principal's strategy included dispersing G/T students equally among all the classes within a grade, he would be forced to deviate from that strategy if some of the fourth grade teachers were not certified. Ideally, teacher characteristics would be included in the analysis to reveal potential sorting constraints. However, because the data allows linkage to a specific class but not to a teacher, this is not possible.

The results from this analysis are reported in Tables B2.1-B2.14 of Appendix B. Each column shows the results from an individual school. Almost all of the schools (134 of 138) serve at least one G/T student. (Of the four schools for which there is no variation in G/T status, three have no G/T students, and one is composed only of G/T students.) As in the previous section, I find that many schools appear to sort based on G/T status and many schools do not. School 108, for example, appears to sort G/T students into different classes. The results for this school are reported in Table B2.1. The school has three classes of fourth graders, one of which does not contain any G/T students. Being enrolled in class 3 reduces the likelihood that a child is classified as a G/T student by a statistically significant 32 percent, when compared to the base outcome (class 1), suggesting that this school groups its G/T students more heavily into class 1.

Other schools appear to divide their G/T students more evenly across classes. School 163 (results reported in Table B2.6) also has three classes of fourth graders, each of which contain G/T students. None of the class dummy variables is significant for this school, indicating that no class assignment significantly increases the likelihood that a student is classified as G/T over the base outcome. Again, lack of significance in this situation does not necessarily mean that schools do not consider G/T status when assigning students to classes. Schools may be purposefully dividing students evenly across classes.

Of the 134 schools that serve G/T students, 40 have at least one class dummy that significantly changes the likelihood of a student's G/T classification. In addition, 19 schools have at least one class with no G/T students at all. (These classes are dropped from the regressions because they perfectly predict failure of the dependent variable.) When including these schools, the percentage of schools sorting by G/T status is about 44 percent.

Figure IV.1 shows a summary of the results from sections IV.4.1 and IV.4.2. Of the 138 schools, 44 (about 32 percent of all schools) sort by previous score and 59 (about 43 percent of all schools) sort by G/T status. There are 68 schools that do not sort using either characteristic (or choose to create evenly distributed classes) and 33 schools that sort using both G/T status and previous score.

		Score Sorting: No	Score Sorting: Yes	
G/T Sorting	No	68 schools	11 schools	Total G/T Non-Sort: 79 schools
	Yes	26 schools	33 schools	Total G/T Sort: 59 schools
		Total Sc Non-Sort: 94 schools	Total Sc Sort: 44 schools	

Figure IV.1 Summary of sorting status by sorting type, based on regressions from Appendices B1 and B2.

IV.4.3 Class Size Effects and Sorting

Because of the potential relationship between class size and sorting, I estimate the class size effect separately for "sorting schools" and "non-sorting schools." I define two different types of sorting schools. A school is a score sorting school if it is as such in section IV.5.1. This requires that at least one of its classes has a statistically different previous score average than another. A school is a G/T sorting school if it is identified as such in section IV.5.2, which requires that at least one of its classes is a significant predictor of a student's G/T status.

Table IV.2
Class size effect on math score (all schools)

Class size	-0.00543
	(-1.300)
Previous score	0.583***
	(68.62)
Black	-0.405***
	(-11.09)
Hispanic	-0.221***
	(-6.285)
Asian	0.0391
	(0.521)
G/T	0.463***
	(24.57)
Free lunch	-0.0722***
	(-3.213)
Special Ed	-0.206***
	(-5.631)
Enroll	0.00158**
	(2.133)
Number Classes	-0.0466***
	(-3.240)
Obs	9325
R-sq	0.494

Table IV.2 shows the baseline results for all schools. I estimate the effect of class size on math score, controlling for student and school characteristics. As explained in section IV.4, the dependent variable in these estimations is a z score constructed from students' scale scores from the math TAKS test. The class size coefficient is negative, but small and not statistically significant. The point estimate indicates that a one student class size reduction increases predicted score by .005 standard deviations.

Table IV.3 divides the sample into schools that use G/T sorting and schools that do not. I estimate the same regression for both samples and find different class size effects for each group. While the class size effect for non-sorting schools is small and completely insignificant, the effect for schools that divide students based on G/T status is larger and magnitude and statistically significant. For sorting schools, a one student class size reduction is associated with a predicted score increase of .013 standard deviations.

Table IV.3
Class size effect on math score by sorting type (G/T)

	Non-Sorting Schools	Sorting Schools
Class size	0.00401	-0.0132**
	(0.623)	(-2.416)
Previous score	0.572***	0.598***
	(46.62)	(50.73)
Black	-0.453***	-0.334***
	(-8.874)	(-6.234)
Hispanic	-0.236***	-0.183***
	(-4.900)	(-3.504)
Asian	0.0712	0.0267
	(0.669)	(0.251)
G/T	0.454***	0.471***
	(17.15)	(17.59)
Free lunch	-0.0985***	-0.0381
	(-2.938)	(-1.262)
Special Ed	-0.205***	-0.202***
	(-3.863)	(-4.023)
Enroll	0.000593	0.00210**
	(0.482)	(2.281)
Number Classes	-0.0273	-0.0647***
	(-1.169)	(-3.507)
Obs	4843	4482
R-sq	0.464	0.530

Table IV.4

Effect of class size on math score with sorting index

	Full Sample			Score Sorters Only		G/T Sorters Only	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Gamma (Sc Sort)		-0.255***			-0.389***		
		(-4.038)			(-4.012)		
Gamma (G/T Sort)			-0.0358				-0.108***
			(-1.072)				(-2.922)
Class Size	-0.00543	-0.00782*	-0.00628	0.00738	0.00295	-0.0132**	-0.0157***
	(-1.300)	(-1.846)	(-1.480)	(1.095)	(0.430)	(-2.416)	(-2.850)
Previous Score	0.583***	0.582***	0.584***	0.607***	0.608***	0.598***	0.597***
	(68.62)	(68.54)	(68.51)	(44.54)	(44.66)	(50.73)	(50.64)
Black	-0.405***	-0.408***	-0.407***	-0.463***	-0.459***	-0.334***	-0.335***
	(-11.09)	(-11.20)	(-10.99)	(-8.184)	(-8.139)	(-6.234)	(-6.258)
Hispanic	-0.221***	-0.225***	-0.224***	-0.298***	-0.297***	-0.183***	-0.178***
	(-6.285)	(-6.404)	(-6.290)	(-5.518)	(-5.508)	(-3.504)	(-3.395)
Asian	0.0391	0.0323	0.0263	0.154	0.148	0.0267	0.0281
	(0.521)	(0.431)	(0.349)	(1.247)	(1.202)	(0.251)	(0.265)
G/T	0.463***	0.462***	0.459***	0.453***	0.455***	0.471***	0.461***
	(24.57)	(24.57)	(24.18)	(14.75)	(14.86)	(17.59)	(17.12)
Free Lunch	-0.0722***	-0.0733***	-0.0731***	-0.0307	-0.0299	-0.0381	-0.0351
	(-3.213)	(-3.263)	(-3.233)	(-0.902)	(-0.879)	(-1.262)	(-1.165)
Special Ed	-0.206***	-0.206***	-0.207***	-0.272***	-0.267***	-0.202***	-0.202***
	(-5.631)	(-5.633)	(-5.656)	(-4.508)	(-4.415)	(-4.023)	(-4.041)
Enroll	0.00158**	0.00167**	0.00169**	-0.00031	-9.75e-05	0.00210**	0.00236**
	(2.133)	(2.245)	(2.271)	(-0.233)	(-0.0727)	(2.281)	(2.556)

Table IV.4 (continued)
Effect of class size on math score with sorting index

	Full Sample			Score Sorters Only		G/T Sorters Only	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Number Classes	-0.0466***	-0.0475***	-0.0484***	-0.0256	-0.0247	-0.0647***	-0.0691***
	(-3.240)	(-3.287)	(-3.341)	(-1.047)	(-1.003)	(-3.507)	(-3.736)
Obs	9325	9322	9264	3546	3545	4482	4482
R-sq	0.494	0.495	0.491	0.528	0.530	0.530	0.531

Such a striking difference in the class size effect based on whether or not a school sorts students by G/T classification indicates a relationship between the class size effect and schools' sorting mechanisms. Table IV.4 presents class size results with the sorting indices described in equations (IV.3) and (IV.4). I include results for the full sample, only score sorters, and only G/T sorters. Because sorting appears to have a negative effect on score, omitting the sorting index biases the class size effect upward, so that the true class size effect is actually more negative when the sorting index is included.[18]

IV.4.4 Effect of Sorting on Score

It is not immediately clear whether sorting students will be beneficial for them or which types of sorting will be most beneficial for different types of students. As described earlier, an intuitive argument can be made for the benefits of tracking students into homogenous classes, as well as for evenly dividing them into heterogeneous classes. To explore this issue empirically, I create a sorting index for each class within a school measuring how dispersed its students are when compared to the overall school population at a single grade level. Following the formulas described in equations (IV.3) and (IV.4), I construct two indices: γ_j, which measures sorting by previous math score, and γ_j^{GT}, which measures sorting by G/T status. I include these variables in regressions for both types of sorting schools and

---

[18] It is important to note that the sorting index itself may be endogenous, which would change the results of the interaction between class size and class composition. See section 4.5 for a discussion of the endogeneity of the sorting index.

measure their effects on all students, as well as on high and low scoring students and G/T and non-G/T students.

Table IV.5
Score sorting effect on math score

	Overall Effect	Effect by Score Type
Gamma	-0.389***	
	(-4.012)	
Gamma*(High Scorers)		-0.126
		(-1.011)
Gamma*(Low Scorers)		-0.626***
		(-5.215)
Class Size	0.00295	0.00390
	(0.430)	(0.568)
Previous Score	0.608***	0.577***
	(44.66)	(35.37)
Black	-0.459***	-0.464***
	(-8.139)	(-8.232)
Hispanic	-0.297***	-0.298***
	(-5.508)	(-5.532)
Asian	0.148	0.149
	(1.202)	(1.207)
G/T	0.455***	0.447***
	(14.86)	(14.57)
Free Lunch	-0.0299	-0.0288
	(-0.879)	(-0.848)
Special Ed	-0.267***	-0.263***
	(-4.415)	(-4.355)
Enroll	-9.75e-05	-0.000252
	(-0.0727)	(-0.188)
Number Classes	-0.0247	-0.0227
	(-1.003)	(-0.924)
Obs	3545	3545
R-sq	0.530	0.531

Table IV.5 presents the results for the score sorting index. The first column shows the overall sorting effect for all students in a score sorting school. The overall effect is negative and significant, suggesting that students in classes with a higher degree of sorting score lower than those in classes with less sorting. The point estimate shows that a 0.1 increase in the sorting index decreases predicted score by 0.038 standard deviations. This might indicate that heterogeneous classes are more beneficial for students in terms of increasing test score, although the magnitude of the effect is fairly small.

The second column divides the sorting effect between high scorers and low scorers. The sample is divided roughly in half by previous year score; high scorers compose the top half of the distribution, and low scorers compose the bottom half.[19] Interestingly, the negative overall effect of sorting is being driven the low scoring students. The sorting coefficient for the low scorers is negative, significant, and about 1.6 times the magnitude of the overall coefficient. The effect for high scoring students is negative but insignificant. Intuitively, this suggests that a more even distribution of students within a class benefits students who typically perform poorly without hurting the scores of high performers.

[19] Low scorers have previous year scale scores between 1228 and 2198. High scorers have previous year scale scores between 2224 and 2697.

Table IV.6
G/T sorting effect on math score

	Overall Effect	Effect by G/T
Gamma	-0.108***	
	(-2.922)	
Gamma*(G/T)		0.0195
		(0.172)
Gamma*(Non-G/T)		-0.122***
		(-3.146)
Class Size	-0.0157***	-0.0155***
	(-2.850)	(-2.812)
Previous Score	0.597***	0.596***
	(50.64)	(50.53)
Black	-0.335***	-0.335***
	(-6.258)	(-6.258)
Hispanic	-0.178***	-0.177***
	(-3.395)	(-3.383)
Asian	0.0281	0.0297
	(0.265)	(0.280)
G/T	0.461***	0.432***
	(17.12)	(11.88)
Free Lunch	-0.0351	-0.0345
	(-1.165)	(-1.144)
Special Ed	-0.202***	-0.203***
	(-4.041)	(-4.051)
Enroll	0.00236**	0.00230**
	(2.556)	(2.486)
Number Classes	-0.0691***	-0.0681***
	(-3.736)	(-3.679)
Obs	4482	4482
R-sq	0.531	0.532

Table IV.6 presents the analogous results for the G/T sorting index. I include the results for all students and then split the sample and examine the effect of sorting for G/T students and for non-G/T students. Like the score sorting effect, the overall G/T sorting effect is negative and significant, revealing that a higher degree of sorting hurts students' scores. When I divide the effect between students of different G/T classifications, I find that this result is explained by the negative effect of sorting on non-G/T students. While the sorting coefficient for non-G/T students is significant and negative, the coefficient for G/T students is actually positive, although it is insignificant.

IV.4.5 Endogeneity of the Sorting Index

Schools choose how to divide students into classes, and it is likely that they make this determination using variables that are unobserved to the researcher. As described in section IV.2.3, unobservable characteristics such as behavior may affect both schools' sorting decisions and student performance. To control for this, I create $sort_k$, a school-grade sorting index for each school and use the third grade index as an instrument for the fourth grade index. The two indices should be correlated if schools' sorting guidelines are similar across grades, but the third grade index should not directly impact the scores of fourth grade students.

Table IV.7
Alpha score sorting effect on math score

	OLS	2SLS
Alpha Sort	-0.255**	2.555**
	(-2.406)	(2.300)
Class Size	0.00786	-0.0104
	(1.165)	(-1.003)
Previous Score	0.623***	0.614***
	(44.98)	(39.01)
Black	-0.374***	-0.418***
	(-6.088)	(-5.975)
Hispanic	-0.218***	-0.213***
	(-3.643)	(-3.216)
Asian	0.214*	0.291**
	(1.739)	(2.097)
G/T	0.465***	0.352***
	(14.66)	(6.227)
Free Lunch	-0.0689**	-0.0174
	(-1.991)	(-0.403)
Special Ed	-0.215***	-0.240***
	(-3.449)	(-3.446)
Enroll	-0.00281*	0.00350
	(-1.937)	(1.184)
Number Classes	0.0184	-0.127**
	(0.709)	(-1.985)
Obs	3262	3262
R-sq	0.549	0.451
Cragg-Donald		36.311

Table IV.7 shows the effect of the $sort_k$ on test score. The first column reports estimates using OLS. The coefficient on the sorting index is negative and significant, as it was in for the class sorting index in the previous table. This suggests that more sorting actually decreases students' test scores. However, when the third grade sorting index is used as an instrument, the 2SLS estimates reveal that a higher sorting index *increases* students' scores. The 2SLS results in the second column show that the sorting coefficient is positive, significant, and relatively large in magnitude. An increase in the sorting index of a school of 0.1 points is associated with a predicted score increase of 0.255 standard deviations.

I also examine the difference in the sorting effect between high scorers and low scorers. Table IV.8 presents OLS and 2SLS results by score type. The OLS results are similar to the findings in Table IV.4, which show that a higher degree of sorting decreases scores for low scoring students but has no significant effect on high scoring students. After using the instrument in the 2SLS results, however, it appears that more sorting significantly *increases* the test scores of both low scoring students and high scoring students. An increase in the sorting index of 0.1 points increases predicted score for low scoring students by 0.268 standard deviations and increases predicted score for high scoring students by 0.243 standard deviations.

Table IV.8
Alpha score sorting effect on math score by score type

	OLS	2SLS
Alpha Sort*(High Scorers)	-0.131	2.681**
	(-1.229)	(2.425)
Alpha Sort*(Low Scorers)	-0.372***	2.431**
	(-3.493)	(2.203)
Class Size	0.00986	-0.00827
	(1.471)	(-0.806)
Previous Score	0.519***	0.506***
	(25.72)	(22.00)
Black	-0.393***	-0.437***
	(-6.434)	(-6.283)
Hispanic	-0.230***	-0.225***
	(-3.859)	(-3.414)
Asian	0.204*	0.281**
	(1.673)	(2.038)
G/T	0.445***	0.331***
	(14.07)	(5.872)
Free Lunch	-0.0680**	-0.0165
	(-1.979)	(-0.385)
Special Ed	-0.214***	-0.238***
	(-3.446)	(-3.441)
Enroll	-0.00320**	0.00308
	(-2.221)	(1.052)
Number Classes	0.0243	-0.120*
	(0.942)	(-1.900)
Obs	3262	3262
R-sq	0.555	0.458
Cragg-Donald		18.15

IV.5 Conclusions

While several studies have found that smaller class sizes significantly increase students' test scores, one confounding factor may be the way in which schools sort students into classes. Schools may sort specific types of students into larger or smaller classes, and schools that have smaller classes or more classes may be able to more effectively sort students into groups. This study attempts to disentangle the class size effect from the sorting effect by creating a sorting index for schools, which captures how "sorted" its classes are. I use a school's previous grade sorting index as an instrument for the index of the grade of interest.

While adding this sorting index does seem to affect the magnitude of the class size effect, it also presents several interesting implications of its own. The OLS estimates of the effect of sorting indicate that more sorted schools actually hurt the scores of lower scoring students and that sorting G/T and non-G/T students into separate classes decreases the scores of non-G/T students. However, after controlling the possible endogeneity of the sorting index caused by unobserved variables, I find that more sorting is actually helpful for all students, regardless of previous score or G/T classification.

This study has valuable policy implications because unlike many school policy variables, the composition of classes can be changed with little need for increased funds. A school with three classrooms and three teachers can increase efficiency by sorting students in such a way that they all benefit. This study suggests that classes that are more sorted are beneficial for all students and that schools may improve overall scores by sorting students into more homogeneous groups.

CHAPTER V

CONCLUSIONS

This dissertation examines three issues related to the effect of smaller classes on student achievement. Using student-level data from the Texas Assessment of Knowledge and Skills (TAKS), I explore heterogeneity in the effect of smaller classes across a score distribution of students, model class size effects simultaneously with moving effects, and disentangle the class size effect from schools' decisions to divide students into classes. The overall results from the paper emphasize the importance of carefully modeling these effects while simultaneously considering other potential confounding or related issues.

I find strong evidence of differences in the class size effect across different groups of students. By dividing students into decile groups based on their previous testing performance and examining marginal effects for the different groups, I find that smaller classes are most helpful for student with low scores in the previous year. While score gains for these students are significant, gains for higher scoring students are small or nonexistent. Knowing how this effect varies across a score distribution is valuable information for teachers and school administrators who are faced with allocating students across classes.

In my analysis of moving decisions and class size, I also find differences in the class size effect between movers and nonmovers, and among movers of different types. I find that movers respond more to class size reductions than nonmovers, and that the strongest results exist for students who switch districts but remain in the same region of the state.

This simultaneous examination of the class size effect and the moving effect reveals a bias in the typical OLS or 2SLS estimates of the moving effect. I create a two-type model that includes endogenous movers, who switch schools because of a school-related reason, and exogenous movers, who transfer schools because of a reason unrelated to school. I find that students who are campus or district movers are more likely to be endogenous movers and that regional movers are more likely to be exogenous. The endogenous campus movers perform worse than expected on the standardized exam, and the endogenous district movers perform better than expected on the exam. This causes a bias in the overall moving effect for both campus and district movers. These estimates show the importance of a simultaneous model of moving and class size.

My third paper is an analysis of the relationship between class size and class composition. Using data from Dallas ISD, I estimate the class size effect along with the effect of sorting students into homogeneous groups. The data is unique because it allows me to create the score distribution of an individual class, whereas most datasets only allow for the creation of a distribution of an entire grade within a school.

I create a sorting index for each school, which is a measure of how similar or dissimilar the scores within the classes of the school are. Because of potential endogeneity of the sorting index due to the fact that students are not divided into classes randomly, I create an instrument for the fourth grade index using the third grade index. It is reasonable to believe that the two indices are correlated if schools have common sorting guidelines across grades, but that the division of third grade students should not directly affect fourth grade scores.

I include the sorting index in a typical class size regression and find that the magnitude of the class size coefficient decreases significantly, suggesting that what would have been labeled a class size effect is actually attributable to the sorting mechanism. The coefficient of the sorting index itself suggests that students at schools with a more homogeneous sorting process have higher test scores. This is true for both high scoring students and low scoring students; the results still hold when the effect is allowed to vary by previous test score. This evidence indicates that students benefit from being divided into groups with classmates who are similar in academic achievement to themselves.

REFERENCES

Angrist, Joshua D., Lavy, Victor, 1999. Using Maimonides' Rule to estimate the effect of class size on scholastic achievement. The Quarterly Journal of Economics 114 (2), 533-575.

Cullen, Julie B., Jacob, Brian A., Levitt, Steven D., 2005. The impact of school choice on student outcomes: an analysis of the Chicago Public Schools. Journal of Public Economics 89 (5-6), 729-760.

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

Hanushek, Eric A., 1997. Assessing the effects of school resources on student performance: an update. Education Evaluation and Policy Analysis, 19 (2), 141-164.

Hanushek, Eric A., Kain, John F., Rivkin, Steven G., 2004. Disruption versus Tiebout improvement: the costs and benefits of switching schools. Journal of Public Economics 88 (9-10), 1721-1746.

Hoxby, Caroline M., 2000. The effects of class size on student achievement: new evidence from population variation. The Quarterly Journal of Economics 115 (4), 1239-1285.

Ingersoll, Gary M., Scamman, James P., Eckerling, Wayne D., 1989. Geographic mobility and student achievement in an urban setting. Educational Evaluation and Policy Analysis 11 (2), 143-149.

Kain, John F., O'Brien, Daniel M., 1999. A Longitudinal Assessment of Reading Achievement: Evidence from the Harvard/UTD Texas Schools Project. University of Texas at Dallas, UTD Texas Schools Project.

Krueger, Alan B., 2003. Economic considerations and class size. The Economic Journal 113 (485), F34-F63.

Mishel, Lawrence, Rothstein, Richard (Eds.), 2002. The Class Size Debate. Economic Policy Institute, Washington, D.C.

Rivkin, Steven G., Hanushek, Eric A., Kain, John F., 2005. Teachers, schools, and academic achievement. Econometrica 73 (2), 417-458.

Word, Elizabeth, Johnston, John, Bain, Helen Pate et al., 1990. The State of Tennessee's Student/Teacher Achievement Ratio (STAR) Project Technical Report 1985-1990. Tennessee Department of Education, Nashville, TN.

Supplemental Sources Consulted

Alexander, Karl L., Entwisle, Doris R., Dauber, Susan L., 1996. Children in motion: school transfers and elementary school performance. Journal of Educational Research, 90 (1), 3-12.

Brown, Byron W., Saks, Daniel H., 1975. The production and distribution of cognitive skills within schools. The Journal of Political Economy 83 (3), 571-594.

Coleman, James S., Campbell, Ernest Q., Hobson, Carl F., McPartland, James, Mood, Alexander M., 1966. Equality of Educational Opportunity (EEOS). US Government Printing Office, Washington, D.C.

Cullen, Julie, Reback, Randall, 2006. Tinkering toward accolades: school gaming under a performance accountability system, in: Gronberg, Timothy J., Jansen, Dennis W. (Eds.), Advances in Applied Microeconomics. Improving School Accountability, vol. 14. Elsevier Science, Amsterdam, The Netherlands, 1-34.

Deere, Donald, Strayer, Wayne, 2002. Competitive Incentives: School Accountability and Student Outcomes in Texas. Working Paper.

Dubin, Jeffrey A., McFadden, Daniel L., 1984. An econometric analysis of residential electric appliance holdings and consumption. Econometrica 52 (2), 345-362.

Eide, Eric, Showalter, Mark H., 1998. The effect of school quality on student performance: a quantile regression approach. Economic Letters 58 (3), 345-350.

Ehrenberg, Ronald G., Brewer, Dominic J., 1994. Do school and teacher characteristics matter? Evidence from High School and Beyond. Economics of Education Review 13 (1),1-17.

Figlio, David, Getzler, Lawrence S., 2006. Accountability, ability, and disability: gaming the system?, in: Gronberg, Timothy J., Jansen, Dennis W. (Eds.), Advances in Applied Microeconomics. Improving School Accountability, vol. 14. Elsevier Science, Amsterdam, The Netherlands, 35-49.

Jacob, Brian A., 2005. Accountability, incentives, and behavior: the impact of high-stakes testing in the Chicago Public Schools. Journal of Public Economics 89 (5-6), 761-796.

Jacob, Brian A., Levitt, Steven, 2003. Rotten apples: An investigation of the prevalence and predictors of teacher cheating. Quarterly Journal of Economics 118 (3), 843-878.

Kerbow, David, 1996. Patterns of urban student mobility and local school reform. Journal of Education for Students Placed at Risk 1 (2), 147-169.

Ladd, Helen F., Walsh, Randall, 2002. Implementing value-added measures of school effectiveness: getting the incentives right. Economics of Education Review 21(1), 1-17.

Medina, Jennifer, 2003. New York measuring teachers by test scores. New York Times (January 21, 2008).

Reback, Randall, 2008. Teaching to the rating: school accountability and the distribution of student achievement. Journal of Public Economics, 92 (5-6), 1394-1415.

Todd, Petra E., Wolpin Kenneth I., 2007. The production of cognitive achievement in children: home, school, and racial test score gaps. Journal of Human Capital 1 (1), 91-136.

APPENDIX A

THREE TYPE MOVING MODEL

In the full moving model, there are three types of movers—campus movers (type 1), district movers (type 2), and region movers (type 3). I define the moving decision in the following way:

Campus Move: $m = 1(Z_i\eta_1 + \varepsilon_1 > 0)$
District Move: $m = 1(Z_i\eta_2 + \varepsilon_2 > 0)$
Region Move: $m = 1(Z_i\eta_3 + \varepsilon_3 > 0)$

The score equation is given by

$$s_{it} = \rho s_{it-1} + \gamma_1 cs_i + (\gamma_{21} moveC + \gamma_{22} moveD + \gamma_{23} moveR)cs_i + u_i.$$

The covariances and errors of the three types of moves are described by the following equations:

Campus Move: $cov(\varepsilon_1, u_i) \neq 0$ $\varepsilon_1 = \lambda_1 + v_1$ and $\sigma_{v1}^2 = 1 - \lambda_1^2\sigma_u^2$
District Move: $cov(\varepsilon_2, u_i) \neq 0$ $\varepsilon_2 = \lambda_2 + v_2$ and $\sigma_{v2}^2 = 1 - \lambda_2^2\sigma_u^2$
Region Move: $cov(\varepsilon_3, u_i) \neq 0$ $\varepsilon_3 = \lambda_3 + v_3$ and $\sigma_{v3}^2 = 1 - \lambda_3^2\sigma_u^2$

Because potential movers of any type can move or stay, there are six possibilities:

- Pr(moveC=1|s_{it})

$$= \Pr(Z_i\eta_1 + \varepsilon_1 > 0|s_{it}) = \Pr(Z_i\eta_1 + \lambda_1 u + v_1 > 0|s_{it})$$
$$= \Phi\left(\frac{Z_i\eta_1 + \lambda_1 u}{\sqrt{1 - \lambda_1^2\sigma_u^2}}\right)$$
$$= \Phi\left(\frac{Z_i\eta_1 + \lambda_1(s_{it} - \rho s_{it-1} - \gamma_1 cs_i - (\gamma_{21} moveC)cs_i)}{\sqrt{1 - \lambda_1^2\sigma_u^2}}\right)$$

- Pr(moveC=0|s_{it})

$$= 1 - \Phi\left(\frac{Z_i\eta_1 + \lambda_1(s_{it} - \rho s_{it-1} - \gamma_1 cs_i - (\gamma_{21} moveC)cs_i)}{\sqrt{1 - \lambda_1^2\sigma_u^2}}\right)$$

- Pr(moveD=1|s_{it})

$$= \Pr(Z_i\eta_2 + \varepsilon_2 > 0|s_{it}) = \Pr(Z_i\eta_2 + \lambda_2 u + v_2 > 0|s_{it})$$
$$= \Phi\left(\frac{Z_i\eta_2 + \lambda_2 u}{\sqrt{1-\lambda_2^2\sigma_u^2}}\right)$$
$$= \Phi\left(\frac{Z_i\eta_2 + \lambda_2(s_{it} - \rho s_{it-1} - \gamma_1 cs_i - (\gamma_{22} moveD)cs_i)}{\sqrt{1-\lambda_2^2\sigma_u^2}}\right)$$

- Pr(moveD=0|s_{it})

$$= 1 - \Phi\left(\frac{Z_i\eta_2 + \lambda_2(s_{it} - \rho s_{it-1} - \gamma_1 cs_i - (\gamma_{22} moveD)cs_i)}{\sqrt{1-\lambda_2^2\sigma_u^2}}\right)$$

- Pr(moveR=1|s_{it})

$$= \Pr(Z_i\eta_3 + \varepsilon_3 > 0|s_{it}) = \Pr(Z_i\eta_3 + \lambda_3 u + v_3 > 0|s_{it})$$
$$= \Phi\left(\frac{Z_i\eta_3 + \lambda_3 u}{\sqrt{1-\lambda_3^2\sigma_u^2}}\right)$$
$$= \Phi\left(\frac{Z_i\eta_3 + \lambda_3(s_{it} - \rho s_{it-1} - \gamma_1 cs_i - (\gamma_{23} moveR)cs_i)}{\sqrt{1-\lambda_3^2\sigma_u^2}}\right)$$

- Pr(moveR=0|s_{it})

$$= 1 - \Phi\left(\frac{Z_i\eta_3 + \lambda_3(s_{it} - \rho s_{it-1} - \gamma_1 cs_i - (\gamma_{23} moveR)cs_i)}{\sqrt{1-\lambda_3^2\sigma_u^2}}\right)$$

While we can distinguish between potential movers who actually move, we cannot distinguish between stayer types. Therefore, the density functions are as follows:

- $f(s_{it}$,moveC=1)=

$$\Phi\left(\frac{Z_i\eta_1 + \lambda_1(s_{it} - \rho s_{it-1} - \gamma_1 cs_i - (\gamma_{21} moveC)cs_i)}{\sqrt{1-\lambda_1^2\sigma_u^2}}\right) * \frac{1}{\sigma_u}\phi\left(\frac{s_{it} - \rho s_{it-1} - \gamma_1 cs_i - (\gamma_{21} moveC)cs_i)}{\sigma_u}\right)$$

- $f(s_{it}$,moveD=1)=

$$\Phi\left(\frac{Z_i\eta_2 + \lambda_2(s_{it} - \rho s_{it-1} - \gamma_1 cs_i - (\gamma_{22} moveD)cs_i)}{\sqrt{1-\lambda_0^2\sigma_u^2}}\right) * \frac{1}{\sigma_u}\phi\left(\frac{s_{it} - \rho s_{it-1} - \gamma_1 cs_i - (\gamma_{22} moveD)cs_i)}{\sigma_u}\right)$$

- $f(s_{it}$,moveR=1)=

$$\Phi\left(\frac{Z_i\eta_3 + \lambda_3(s_{it} - \rho s_{it-1} - \gamma_1 cs_i - (\gamma_{23} moveR)cs_i)}{\sqrt{1-\lambda_0^2\sigma_u^2}}\right) * \frac{1}{\sigma_u}\phi\left(\frac{s_{it} - \rho s_{it-1} - \gamma_1 cs_i - (\gamma_{23} moveR)cs_i)}{\sigma_u}\right)$$

- $f(s_{it}$,move=0)=

$$\left[1 - \Phi\left(\frac{Z_i\eta_1 + \lambda_1(s_{it} - \rho s_{it-1} - \gamma_1 cs_i - (\gamma_{21} moveC)cs_i)}{\sqrt{1-\lambda_1^2\sigma_u^2}}\right) + 1 - \Phi\left(\frac{Z_i\eta_2 + \lambda_2(s_{it} - \rho s_{it-1} - \gamma_1 cs_i - (\gamma_{22} moveD)cs_i)}{\sqrt{1-\lambda_2^2\sigma_u^2}}\right) + 1 - \Phi\left(\frac{Z_i\eta_3 + \lambda_3(s_{it} - \rho s_{it-1} - \gamma_1 cs_i - (\gamma_{23} moveR)cs_i)}{\sqrt{1-\lambda_3^2\sigma_u^2}}\right)\right] * \frac{1}{\sigma_u}\phi\left(\frac{s_{it} - \rho s_{it-1} - \gamma_1 cs_i - (\gamma_{21} moveC + \gamma_{22} moveD + \gamma_{23} moveR)cs_i)}{\sigma_u}\right)$$

The log likelihood function for the maximum likelihood estimation is

$$\ell_i = \sum_{k=1}^{4}\sum_{i=1}^{N_k} \log(Case_k).$$

APPENDIX B

INDIVIDUAL SCHOOL REGRESSIONS FOR SCORE SORTING AND G/T SORTING

Table B1.1

Score sorting for schools 101-113

Dependent Variable: Raw Math Score (2004)										
School	101	103	104	105	106	108	109	110	112	113
Constant	28.79***	30.82***	29.24***	31.89***	27.60***	24.11***	28.80***	28.95***	30.85***	32.29***
	(13.47)	(24.01)	(19.88)	(23.34)	(14.34)	(11.16)	(17.00)	(20.11)	(19.86)	(28.06)
Class 2	-3.119	-0.529	0.174	-2.332	4.733*	6.333**	1.200	-0.619	-0.179	0.614
	(-1.050)	(-0.292)	(0.0790)	(-1.154)	(1.816)	(2.393)	(0.535)	(-0.278)	(-0.088)	(0.345)
Class 3	-5.016	1.301	-0.683	-1.784		4.278	0.0947	0.860	0.804	
	(-1.629)	(0.706)	(-0.315)	(-0.911)		(1.616)	(0.0418)	(0.393)	(0.403)	
Class 4	-0.661	1.093	-1.888				2.311	1.548	-1.199	
	(-0.226)	(0.548)	(-0.897)				(1.007)	(0.707)	(-0.581)	
Class 5	-0.319		0.429						-0.146	
	(-0.107)		(0.198)						(-0.073)	
Class 6	-1.357								-19.9***	
	(-0.449)								(-4.666)	
Obs	87	62	94	53	22	45	72	68	90	24
R-Sq	0.048	0.022	0.016	0.029	0.142	0.120	0.021	0.015	0.233	0.005

Table B1.2

Score sorting for schools 114-125

Dependent Variable: Raw Math Score (2004)										
School	114	115	116	117	118	119	120	121	124	125
Constant	30.32***	28.95***	31.29***	32.65***	28.41***	33.27***	29.67***	25.80***	30.69***	31.16***
	(468.4)	(19.43)	(24.46)	(23.03)	(18.28)	(26.09)	(17.02)	(17.31)	(24.92)	(25.71)
Class 2	3.433**	1.850	1.888	-0.923	1.882	0.733	0.167	4.494**	2.719	1.136
	(2.236)	(0.878)	(0.925)	(-0.471)	(0.857)	(0.413)	(0.0706)	(2.197)	(1.662)	(0.644)
Class 3	1.183	2.828	-1	-2.706		-2.443	1.649	3.700	2.367	0.898
	(0.731)	(1.307)	(-0.553)	(-1.314)		(-1.397)	(0.707)	(1.655)	(1.446)	(0.517)
Class 4		0.574	1.920	-5.496**		-2.425	2.667			2.564
		(0.276)	(1.009)	(-2.433)		(-1.422)	(1.082)			(1.476)
Class 5			1.143	-1.826		-0.642				0.430
			(0.622)	(-0.873)		(-0.362)				(0.244)
Class 6			0.143	-4.286**						1.405
			(0.0780)	(-2.188)						(0.784)
Class 7			1.306	0.739						
			(0.699)	(0.359)						
Class 8			-0.937							
			(-0.492)							
Obs	11273	79	120	130	34	83	67	44	47	105
R-Sq	0.000	0.027	0.043	0.096	0.022	0.068	0.026	0.114	0.067	0.025

Table B1.3

Score sorting for schools 126-135

Dependent Variable: Raw Math Score (2004)										
School	126	127	128	129	130	131	132	133	134	135
Constant	32.69***	30.50***	33.55***	32.83***	28.31***	28.73***	34.27***	32.82***	34.60***	32.59***
	(22.90)	(14.60)	(19.56)	(25.40)	(15.82)	(13.39)	(31.20)	(24.31)	(14.82)	(29.43)
Class 2	-1.288	-1.500	0.330	-2.721	0.217	-1.894	-1.267	-0.350	0.567	2.412
	(-0.672)	(-0.490)	(0.148)	(-1.416)	(0.0870)	(-0.638)	(-0.841)	(-0.188)	(0.179)	(1.562)
Class 3	-1.158			-1.445	1.500	-3.182		-1.824	-4.933	1.471
	(-0.582)			(-0.772)	(0.593)	(-1.049)		(-0.908)	(-1.294)	(0.939)
Class 4	0.312			-5.159***	-1.136	-1.035		-1.297	-1.100	
	(0.161)			(-2.645)	(-0.456)	(-0.355)		(-0.698)	(-0.314)	
Class 5	-2.902				1.570	-5.427*		-1.294	-2.200	
	(-1.389)				(0.630)	(-1.745)		(-0.678)	(-0.666)	
Class 6	-3.988*				-0.0903				-1.600	
	(-1.733)				(-0.037)				(-0.485)	
Class 7	-8.56***				1.621				-2.00	
	(-3.464)				(0.630)				(-0.606)	
Class 8	0.786				1.438				0.733	
	(0.406)				(0.568)				(0.192)	
Class 9					-1.839					
					(-0.757)					
Obs	123	15	27	75	151	57	32	86	36	52
R-Sq	0.155	0.018	0.001	0.000	0.030	0.066	0.023	0.015	0.103	0.048

Table B1.4

Score sorting for schools 136-148

Dependent Variable: Raw Math Score (2004)										
School	136	137	139	140	141	142	144	145	147	148
Constant	29.93***	29.40***	30.18***	28.67***	31.91***	32.88***	31.57***	31.50***	30.77***	33.42***
	(15.23)	(16.81)	(19.38)	(16.60)	(18.47)	(15.85)	(18.86)	(20.80)	(17.27)	(21.85)
Class 2	-1.823	2.306	0.718	0.256	-0.909	-0.952	0.929	0.857	1.842	-5.217*
	(-0.704)	(0.961)	(0.318)	(0.107)	(-0.353)	(-0.361)	(0.358)	(0.432)	(0.788)	(-1.850)
Class 3	-1.634	1.100	3.091	1		-0.375	1.012	-0.375	0.168	-2.639
	(-0.616)	(0.475)	(1.404)	(0.379)		(-0.144)	(0.411)	(-0.165)	(0.0701)	(-1.130)
Class 4	-4.693*		2.068			-0.625	-2.038	-2.300	-1.969	
	(-1.769)		(0.959)			(-0.213)	(-0.875)	(-1.074)	(-0.809)	
Class 5								0.750	2.159	
								(0.366)	(0.873)	
Class 6								1.643	0.0543	
								(0.828)	(0.0229)	
Class 7								2.233	-0.894	
								(1.142)	(-0.373)	
Class 8								1.794		
								(0.940)		
Obs	67	52	44	34	20	43	51	100	109	26
R-Sq	0.051	0.019	0.055	0.005	0.007	0.004	0.042	0.075	0.044	0.138

Table B1.5

Score sorting for schools 149-161

Dependent Variable: Raw Math Score (2004)

School	149	150	152	153	156	157	158	159	160	161
Constant	28.44***	30.85***	30.22***	34.07***	27.71***	28.63***	32.06***	30.69***	32***	26.95***
	(13.40)	(19.90)	(16.75)	(27.64)	(17.17)	(20.37)	(23.69)	(21.35)	(18.25)	(16.30)
Class 2	2.368	-0.275	-5.500**	0.656	2.571	2.563	0.938	2.912	1.273	0.598
	(0.892)	(-0.128)	(-2.155)	(0.353)	(1.126)	(1.290)	(0.473)	(1.256)	(0.513)	(0.265)
Class 3	3.006		-0.500	-0.571	4.345*	1.508	-1.262	4.131*		0.653
	(1.176)		(-0.196)	(-0.338)	(1.993)	(0.747)	(-0.649)	(1.834)		(0.283)
Class 4	3.079		-2.294	-3.148*	2.286		-0.562	0.236		-1.241
	(1.214)		(-0.841)	(-1.772)	(1.001)		(-0.258)	(0.110)		(-0.516)
Class 5			-2.696		4.050*		1	1.812		-0.526
			(-1.071)		(1.858)		(0.522)	(0.825)		(-0.225)
Class 6			-3.522				2.622	0.812		
			(-1.416)				(1.427)	(0.370)		
Class 7							2.838			
							(1.563)			
Obs	66	27	107	54	76	47	110	74	22	97
R-Sq	0.027	0.001	0.059	0.090	0.065	0.037	0.072	0.066	0.013	0.010

Table B1.6
Score sorting for schools 162-173

Dependent Variable: Raw Math Score (2004)

School	162	163	164	166	167	168	169	170	171	173
Constant	35.53***	31.44***	29.75***	27.24***	28.71***	31.67***	31.47***	32.53***	31.84***	31.18***
	(26.29)	(21.74)	(21.75)	(20.76)	(15.67)	(13.88)	(31.71)	(21.04)	(24.55)	(27.67)
Class 2	-0.304	0.614	2.250	-3.435	1.590	-5.167*	5.533***	1.400	1.325	0.369
	(-0.157)	(0.296)	(1.148)	(-1.593)	(0.684)	(-1.767)	(3.943)	(0.640)	(0.712)	(0.246)
Class 3	-0.241	3.489	2.039	4.820***	0.391	-1.917	1.950	-1.004	0.720	0.516
	(-0.129)	(1.626)	(1.040)	(2.635)	(0.162)	(-0.672)	(1.310)	(-0.473)	(0.375)	(0.301)
Class 4			1.050	3.965**	-0.714	-2	3.176**	-0.800	-0.0421	0.0543
			(0.502)	(2.222)	(-0.302)	(-0.620)	(2.224)	(-0.366)	(-0.023)	(0.0317)
Class 5			0.179	5.615***	1.230	-1.792	4.033***			0.585
			(0.0838)	(3.146)	(0.504)	(-0.628)	(2.709)			(0.386)
Class 6			0.309	4.065**		-2.810	1.200			4.118**
			(0.153)	(2.278)		(-0.961)	(0.740)			(2.584)
Class 7										3.706**
										(2.326)
Class 8										4.192***
										(2.703)
Obs	58	50	120	105	95	78	77	62	73	139
R-Sq	0.001	0.059	0.024	0.217	0.017	0.049	0.211	0.026	0.011	0.138

Table B1.7

Score sorting for schools 174-187

Dependent Variable: Raw Math Score (2004)										
School	175	178	180	181	182	183	184	185	186	187
Constant	29.71***	31.18***	29.49***	30.85***	29.54***	29.38***	33.25***	34.19***	26.10***	30.69***
	(17.35)	(21.43)	(29.96)	(22.43)	(17.65)	(19.80)	(27.20)	(27.08)	(13.18)	(18.69)
Class 2	-1.581	0.474	1.625	0.150	1.662	2.292	-0.558	-0.688	1.233	0.187
	(-0.643)	(0.239)	(0.944)	(0.0630)	(0.727)	(1.124)	(-0.305)	(-0.372)	(0.460)	(0.0808)
Class 3	-3.143	-0.176	2.980	-3.500*	-0.681	-1.606			4.900*	0.312
	(-1.278)	(-0.087)	(1.626)	(-1.799)	(-0.293)	(-0.725)			(1.890)	(0.125)
Class 4				-1.203	0.395	0.692			6.054**	1.455
				(-0.593)	(0.173)	(0.324)			(2.298)	(0.606)
Class 5				-0.750	-0.0679					
				(-0.386)	(-0.0305)					
Class 6					1.873					
					(0.843)					
Obs	49	55	70	87	91	62	29	30	49	58
R-Sq	0.034	0.002	0.041	0.048	0.025	0.055	0.003	0.005	0.143	0.008

Table B1.8

Score sorting for schools 189-199

Dependent Variable: Raw Math Score (2004)

School	189	190	192	193	194	195	196	197	198	199
Constant	32.86***	29.58***	27.94***	31.63***	26.09***	31.27***	28.67***	26.36***	33.87***	32.60***
	(24.92)	(22.58)	(17.99)	(18.35)	(24.45)	(22.33)	(23.51)	(15.90)	(21.93)	(27.69)
Class 2	-1.197	5.350***	-1.592	-3.807	0.0144	-1.267	4.500**	6.734***	-1.717	1.295
	(-0.578)	(3.044)	(-0.714)	(-1.410)	(0.00916)	(-0.588)	(2.609)	(2.695)	(-0.840)	(0.822)
Class 3	-2.564	5.512***	2.371	-3.092	1.623	3.358*	1.905	2.643	-3.631*	0.0667
	(-1.341)	(3.356)	(1.094)	(-1.248)	(1.063)	(1.723)	(1.033)	(1.083)	(-1.714)	(0.0377)
Class 4		4.226**	1.056	-8.125**	10.82***	-10.27**		2.00		-0.600
		(2.573)	(0.466)	(-2.462)	(5.854)	(-2.515)		(0.853)		(-0.366)
Class 5					11.65***	5.633**				
					(7.430)	(2.545)				
Class 6					10.99***	5.933***				
					(6.275)	(2.680)				
Obs	57	69	70	48	105	64	50	51	52	62
R-Sq	0.032	0.166	0.050	0.128	0.542	0.318	0.127	0.137	0.057	0.027

Table B1.9

Score sorting for schools 200-210

Dependent Variable: Raw Math Score (2004)										
School	200	201	202	203	204	205	206	207	209	210
Constant	31.57***	37.21***	28.83***	32.92***	30.39***	27.50***	30.93***	26.20***	30.61***	33.06***
	(11.59)	(30.60)	(15.18)	(17.81)	(15.77)	(20.90)	(19.76)	(12.09)	(17.34)	(24.01)
Class 2	-1.264	-2.481	-4.333	-4.462	-2.330	2.342	-0.227	1.752	-0.401	-1.184
	(-0.374)	(-1.467)	(-1.613)	(-1.670)	(-0.843)	(1.275)	(-0.106)	(0.665)	(-0.163)	(-0.599)
Class 3	2.429	-1.571	-0.611	0.0119	-3.222	2.111	0.844	6.326**	0.742	-0.614
	(0.719)	(-0.914)	(-0.211)	(0.00473)	(-1.182)	(1.134)	(0.398)	(2.362)	(0.293)	(-0.320)
Class 4	-5.143	0.695	0.917	0.0208	-2.514	1.600		6.700**	-2.842	-0.427
	(-1.335)	(0.379)	(0.341)	(0.00852)	(-0.895)	(0.882)		(2.523)	(-1.042)	(-0.225)
Class 5	-1.446	-2.00		-3.917	-0.0948	4.500**		4.943*	-1.361	-2.121
	(-0.443)	(-1.163)		(-1.060)	(-0.034)	(2.509)		(1.741)	(-0.428)	(-1.073)
Class 6		-1.786		-2.583		2.833				
		(-1.038)		(-0.807)		(1.580)				
Class 7				-7.917						
				(-1.619)						
Obs	56	82	45	65	86	117	50	84	75	103
R-Sq	0.096	0.059	0.097	0.114	0.028	0.058	0.006	0.123	0.027	0.017

Table B1.10

Score sorting for schools 211-220

Dependent Variable: Raw Math Score (2004)

School	211	212	213	214	215	216	217	218	219	220
Constant	28.7***	32.42***	30.44***	33.25***	32.27***	14.83***	37.58***	29***	32.14***	30.36***
	(20.78)	(23.33)	(20.97)	(17.91)	(18.64)	(6.286)	(89.17)	(18.43)	(21.23)	(23.31)
Class 2	1.872	-2.417	-0.490	-0.712	-1.267	10.11***	0.532	1.083	-4.078*	5.186***
	(0.879)	(-0.580)	(-0.249)	(-0.276)	(-0.547)	(3.319)	(0.881)	(0.459)	(-1.779)	(2.747)
Class 3	1.800	-0.0167	1.096	-1.517	1.633	10.77***	0.0526	2.00	-3.136	-2.056
	(0.880)	(-0.0050)	(0.525)	(-0.609)	(0.713)	(3.401)	(0.0883)	(0.883)	(-1.345)	(-0.962)
Class 4	-1.750	-1.560	2.125	-0.250	1.322	8.325***		0.125	-0.397	3.636*
	(-0.80)	(-0.533)	(1.035)	(-0.095)	(0.556)	(2.762)		(0.0571)	(-0.188)	(1.926)
Class 5	1.863	0.0119		-2.679		8.778***		0.647		-4.481**
	(0.941)	(0.00407)		(-1.059)		(2.881)		(0.300)		(-2.271)
Class 6				-1.139		12.54***		0.667		
				(-0.402)		(4.018)		(0.313)		
Class 7						12.50***		0.118		
						(4.103)		(0.0545)		
Class 8								-0.0625		
								(-0.029)		
Obs	75	46	66	75	71	116	56	125	78	92
R-Sq	0.055	0.014	0.032	0.021	0.032	0.167	0.017	0.011	0.058	0.260

Table B1.11

Score sorting for schools 222-233

Dependent Variable: Raw Math Score (2004)										
School	222	223	224	225	226	228	229	230	232	233
Constant	24.69***	29.92***	32***	26.25***	24.38***	30.20***	32.93***	30.93***	28.82***	36.55***
	(12.21)	(20.33)	(13.36)	(14.14)	(12.34)	(22.32)	(27.87)	(21.02)	(19.54)	(30.00)
Class 2	2.262	1.896	-2.263	1.295	2.615		1.127	0.0667	1.353	-2.962*
	(0.834)	(0.974)	(-0.81)	(0.445)	(1.008)		(0.715)	(0.0302)	(0.649)	(-1.757)
Class 3	0.963		-1.789	-2.625	3.668		1.601	-3.206	0.598	-1.084
	(0.355)		(-0.64)	(-1.000)	(1.430)		(1.003)	(-1.417)	(0.294)	(-0.655)
Class 4	1.523			2.515	4.715*		1.214	3.924*	0.248	-1.636
	(0.555)			(0.972)	(1.857)		(0.727)	(1.853)	(0.113)	(-0.950)
Class 5	0.693			-1.917			2.134		1.954	
	(0.258)			(-0.676)			(1.319)		(0.950)	
Class 6	2.513						1.302		2.614	
	(0.926)						(0.765)		(1.234)	
Obs	116	28	45	72	70	15	92	52	101	47
R-Sq	0.012	0.035	0.015	0.070	0.053	0.000	0.021	0.171	0.024	0.070

Table B1.12

Score sorting for schools 235-263

Dependent Variable: Raw Math Score (2004)										
School	235	236	237	241	250	251	256	259	262	263
Constant	23.33***	29.87***	30.47***	27**	29.17***	33.38***	29.77***	26.11***	33***	32.13***
	(13.22)	(18.72)	(16.63)	(3.792)	(16.33)	(18.85)	(17.92)	(9.054)	(23.67)	(18.02)
Class 2	3.167	-2.067	-1.538	-1.500	4.657*	-0.108	-3.133		-1.846	0.337
	(1.289)	(-0.916)	(-0.583)	(-0.12)	(1.996)	(-0.043)	(-1.277)		(-0.96)	(0.138)
Class 3	0.881	1.210	-6.333**		0.141	-3.480			0.444	0.200
	(0.347)	(0.517)	(-2.444)		(0.0569)	(-1.448)			(0.209)	(0.0793)
Class 4		1.192	-2.967			-8.45***			1.600	
		(0.545)	(-1.163)			(-3.195)			(0.774)	
Class 5			2.462			-2.063			-0.0909	
			(0.934)			(-0.823)			(-0.05)	
Class 6			0.248			-4.691*			-1.455	
			(0.0939)			(-1.951)			(-0.72)	
Class 7			-1.252							
			(-0.475)							
Class 8			-3.114							
			(-1.239)							
Obs	45	60	119	7	42	98	24	9	66	47
R-Sq	0.041	0.048	0.115	0.008	0.125	0.136	0.069	0.000	0.059	0.000

Table B1.13
Score sorting for schools 264-274

Dependent Variable: Raw Math Score (2004)

School	264	265	267	268	269	270	271	272	273	274
Constant	29.94***	31.33***	32.50***	33.18***	28.33***	32.93***	32.50***	30.32***	33.28***	33.26***
	(14.63)	(18.65)	(12.04)	(39.12)	(16.69)	(24.19)	(25.97)	(21.51)	(21.09)	(25.60)
Class 2	-0.854	-2.976	2.722	0.218	1.667	-0.670	-2.286	-3.649	-2.468	-0.311
	(-0.273)	(-1.231)	(0.839)	(0.111)	(0.655)	(-0.368)	(-1.208)	(-1.611)	(-1.148)	(-0.173)
Class 3	-1.028	-2.833	2.900	0.532	4.417*	2.008	0.676	-0.941	-0.178	-4.93***
	(-0.321)	(-1.172)	(0.908)	(0.233)	(1.967)	(1.075)	(0.377)	(-0.451)	(-0.082)	(-2.646)
Class 4	-1.553		-0.800	0.318	2.258	0.305	-0.618	1.474	-7.17***	-2.368
	(-0.508)		(-0.251)	(0.161)	(1.026)	(0.171)	(-0.344)	(0.739)	(-3.258)	(-1.289)
Class 5	1.370		5.125	-2.848		-0.121	-1.437	1.906	-1.444	-1.541
	(0.448)		(1.550)	(-1.555)		(-0.064)	(-0.788)	(0.944)	(-0.647)	(-0.827)
Class 6			-0.500	0.374		-0.0510	0.250	1.518	-2.219	-1.541
			(-0.151)	(0.182)		(-0.027)	(0.137)	(0.751)	(-0.980)	(-0.827)
Class 7			-4.125	-4.455**		1.178	-0.222	-1.907		-1.854
			(-1.248)	(-2.349)		(0.639)	(-0.126)	(-0.991)		(-1.045)
Class 8				-2.807		0.773	0.0385			
				(-1.298)		(0.414)	(0.0199)			
Class 9				-4.727**						
				(-2.493)						
Obs	65	43	57	122	69	140	129	124	113	135
R-Sq	0.016	0.046	0.241	0.113	0.058	0.024	0.029	0.083	0.119	0.067

Table B1.14

Score sorting for schools 275-284

Dependent Variable: Raw Math Score (2004)

School	275	276	277	280	281	283	284
Constant	26.78***	29.89***	23.92***	32***	30.53***	31.37***	29.08***
	(20.76)	(20.01)	(12.82)	(25.73)	(17.55)	(20.89)	(19.63)
Class 2	6.022***	4.817**	3.294	-0.400	1.217	0.520	1.506
	(3.387)	(2.247)	(1.382)	(-0.21)	(0.502)	(0.255)	(0.705)
Class 3	7.139***	0.0397	5.655**	3.444*	0.467	1.275	3.923*
	(4.183)	(0.0176)	(2.224)	(1.771)	(0.186)	(0.633)	(1.793)
Class 4	5.813***	2.040	5.672**	0.231	-0.533	1.272	4.559**
	(3.342)	(0.903)	(2.328)	(0.131)	(-0.217)	(0.608)	(2.084)
Class 5	-1.232	4.799**	3.028	-1.714	-0.475		
	(-0.588)	(2.204)	(1.257)	(-0.99)	(-0.199)		
Class 6		2.549		1.889	1.400		
		(1.170)		(0.971)	(0.569)		
Class 7					1.921		
					(0.718)		
Obs	95	95	80	69	103	72	47
R-Sq	0.265	0.095	0.086	0.123	0.018	0.008	0.116

Table B2.1: Schools 101-113

G/T sorting for schools 101-113

Dependent Variable: Gifted/Talented (Probit Regression[1])

School	101	103	104	105	106	108†	109	110	112†	113
Class 2	-0.0845	-0.075	0.0954	0.127	0.318**		0.0300	-0.075		-0.036
	(-0.942)	(-0.63)	(0.768)	(0.952)	(1.978)		(0.209)	(-0.63)		(-0.23)
Class 3	-0.0270	-0.025	0.146	0.105		-0.32**	0.0300	0.225*	-0.0038	
	(-0.287)	(-0.21)	(1.159)	(0.818)		(-2.30)	(0.209)	(1.668)	(-0.031)	
Class 4	0.00646	-0.066	-0.028				0.126	0.259*	0.0120	
	(0.0683)	(-0.55)	(-0.23)				(0.869)	(1.932)	(0.100)	
Class 5	-0.0379		0.195						-0.147	
	(-0.418)		(1.518)						(-1.276)	
Class 6	-0.0327								-0.154	
	(-0.354)								(-1.352)	
Obs	111	77	112	60	28	43	82	82	108	30

[1]Reported with marginal effects

†These schools have at least one class with zero G/T students.

‡These schools have at least one class with all G/T students.

Table B2.2

G/T sorting for schools 114-125

Dependent Variable: Gifted/Talented (Probit Regression[1])

School	114	115	116†	117†	118	119	120†	121	124	125
Class 2	-0.083	0.0687			0.0151	-0.0237		0.216	0.0605	-0.0626
	(-0.71)	(0.517)			(0.160)	(-0.19)		(1.393)	(0.362)	(-0.64)
Class 3	-0.083	0.175	-0.0487	-0.118		-0.0368	0.269*	0.0596	0.225	-0.106
	(-0.71)	(1.330)	(-0.44)	(-1.011)		(-0.29)	(1.762)	(0.381)	(1.398)	(-1.11)
Class 4		0.0280	0.216*	0.00860		-0.0368	0.336**			-0.0562
		(0.208)	(1.831)	(0.0689)		(-0.30)	(2.163)			(-0.57)
Class 5			0.121	0.0438		-0.183				-0.0143
			(1.017)	(0.352)		(-1.48)				(-0.14)
Class 6			-0.0487	-0.0612						-0.0494
			(-0.44)	(-0.513)						(-0.49)
Class 7			-0.0487	0.180						
			(-0.44)	(1.392)						
Class 8			-0.0554							
			(-0.52)							
Obs	67	91	144	136	49	97	62	57	56	124

[1]Reported with marginal effects

†These schools have at least one class with zero G/T students.

‡These schools have at least one class with all G/T students.

Table B2.3

G/T sorting for schools 126-135

Dependent Variable: Gifted/Talented (Probit Regression[1])

School	126	127	128	129	130†	131	132	133	134†	135
Class 2	0.0548	-0.222	0.175	-0.09		-0.0212	0.0779	0.0224		0.0780
	(0.603)	(-1.36)	(1.238)	(-0.88)		(-0.188)	(0.710)	(0.173)		(0.547)
Class 3	-0.026			0.0309		0.0582		-0.194	0.263	-0.092
	(-0.32)			(0.298)		(0.483)		(-1.43)	(0.858)	(-0.63)
Class 4	0.0151			-0.042	0.0159	0.0472		0.0443	-0.101	
	(0.176)			(-0.41)	(0.154)	(0.401)		(0.348)	(-0.30)	
Class 5	-0.073				0.00486	0.00310		0.176	-0.200	
	(-0.95)				(0.0488)	(0.0258)		(1.231)	(-0.63)	
Class 6	-0.065				-0.0461				0.00	
	(-0.81)				(-0.469)				(-0.00)	
Class 7	-0.071				0.0223				-0.404	
	(-0.91)				(0.211)				(-1.28)	
Class 8	0.0476				0.159				0.0693	
	(0.535)				(1.368)				(0.189)	
Class 9					0.0101					
					(0.100)					
Obs	171	27	37	89	139	89	47	98	33	67

[1]Reported with marginal effects

†These schools have at least one class with zero G/T students.

‡These schools have at least one class with all G/T students.

Table B2.4

G/T sorting for schools 136-148

Dependent Variable: Gifted/Talented (Probit Regression[1])

School	136	137	139	140	141	142	144	145†	147†	148
Class 2	0.103	0.402***	0.173	0.430**	0.124	0.0850	0.0791			-0.38**
	(0.769)	(2.616)	(0.865)	(2.223)	(0.970)	(0.506)	(0.625)			(-2.50)
Class 3	0.0611	0.104	0.00	0.293		0.124	0.153	-0.091	0.201	-0.147
	(0.452)	(0.662)	(-0.00)	(1.544)		(0.745)	(1.142)	(-0.80)	(1.398)	(-1.04)
Class 4	0.0486		0.141			0.00	-0.073	0.0388	0.0147	
	(0.366)		(0.721)			(-0.00)	(-0.60)	(0.325)	(0.111)	
Class 5								-0.028	0.0719	
								(-0.25)	(0.539)	
Class 6								0.123	-0.012	
								(0.947)	(-0.10)	
Class 7								0.149	0.171	
								(1.171)	(1.242)	
Class 8								0.0388		
								(0.325)		
Obs	80	58	49	43	35	63	67	130	117	45

[1]Reported with marginal effects

†These schools have at least one class with zero G/T students.

‡These schools have at least one class with all G/T students.

Table B2.5

G/T sorting for schools 149-161

Dependent Variable: Gifted/Talented (Probit Regression[1])										
School	149	150	152†	153	156	157	158	159†	160	161
Class 2	0.0966	-0.143		0.0344	0.0964	0.0711	0.164		0.0523	0.0368
	(0.678)	(-0.92)		(0.332)	(0.598)	(0.479)	(1.100)		(0.551)	(0.334)
Class 3	0.292**		0.0151	0.0872	0.0654	0.300**	-0.00841	0.0421		-0.051
	(2.031)		(0.208)	(0.810)	(0.421)	(1.983)	(-0.061)	(0.400)		(-0.49)
Class 4	0.247*		0.0180	-0.120	0.143		-0.0765	0.168		-0.130
	(1.727)		(0.243)	(-1.12)	(0.895)		(-0.550)	(1.483)		(-1.23)
Class 5			0.0719		0.126		0.164	-0.051		-0.036
			(0.887)		(0.801)		(1.100)	(-0.52)		(-0.34)
Class 6			0.0125				0.108	0.0856		
			(0.174)				(0.747)	(0.827)		
Class 7							0.0965			
							(0.678)			
Obs	90	28	111	87	87	56	128	104	35	113

[1]Reported with marginal effects

†These schools have at least one class with zero G/T students.

‡These schools have at least one class with all G/T students.

Table B2.6

G/T sorting for schools 162-173

Dependent Variable: Gifted/Talented (Probit Regression[1])

School	162	163	164	166†	167†	168	169	170†	171	173†
Class 2	0.132	0.0652	0.0650			-0.0062	0.500***		-0.083	
	(0.930)	(0.492)	(0.474)			(-0.058)	(3.106)		(-0.64)	
Class 3	0.250*	0.0772	-0.0253	0.994***	0.00682	0.0605	0.123	-0.094	-0.059	-0.0720
	(1.795)	(0.572)	(-0.189)	(10.09)	(0.0585)	(0.572)	(0.737)	(-0.79)	(-0.44)	(-0.453)
Class 4			-0.131	0.994***	0.0823	-0.0820	0.241	-0.054	-0.053	-0.211
			(-0.974)	(10.09)	(0.710)	(-0.830)	(1.466)	(-0.46)	(-0.41)	(-1.297)
Class 5			-0.164	0.990***	0.0560	-0.0286	0.413**			-0.127
			(-1.273)	(9.429)	(0.475)	(-0.287)	(2.492)			(-0.785)
Class 6			-0.0569	0.957		-0.0771				0.413**
			(-0.425)			(-0.769)				(2.517)
Class 7			0.00797							0.447***
			(0.0575)							(2.809)
										0.640***
										(4.014)
Obs	71	57	145	105	91	108	61	91	141	53

[1]Reported with marginal effects

†These schools have at least one class with zero G/T students.

‡These schools have at least one class with all G/T students.

Table B2.7

G/T sorting for schools 174-186

Dependent Variable: Gifted/Talented (Probit Regression[1])

School	174	175	178	180	181	182†	183	184	185	186
Class 2	0.177	0.0409	-0.0573	0.0631	0.00		-0.21**	0.00294	-0.204	0.177
	(1.060)	(0.514)	(-0.40)	(0.560)	(-0.00)		(-2.16)	(0.0187)	(-1.64)	(1.070)
Class 3	0.231	0.00	0.114	0.241**	-0.00362		-0.21**			0.104
	(1.388)	(-0.00)	(0.784)	(2.018)	(-0.034)		(-2.23)			(0.638)
Class 4					0.133	-0.177*	-0.132			0.177
					(1.089)	(-1.74)	(-1.38)			(1.070)
Class 5					0.252*	-0.088				
					(1.939)	(-0.95)				
Obs	62	62	84	106	80	59	81	37	43	62

[1]Reported with marginal effects

†These schools have at least one class with zero G/T students.

‡These schools have at least one class with all G/T students.

Table B2.8

G/T sorting for schools 187-198

Dependent Variable: Gifted/Talented (Probit Regression[1])

School	187	189	190	192	193	194‡	195†	196	197†	198
Class 2	-0.044	0.0529	0.199	-0.0463	-0.0926			0.500***		-0.0127
	(-0.42)	(0.408)	(1.145)	(-0.469)	(-0.96)			(3.154)		(-0.095)
Class 3	-0.098	0.143	0.314*	0.103	-0.142			0.00648	-0.039	-0.0354
	(-0.95)	(1.171)	(1.766)	(1.002)	(-1.51)			(0.0393)	(-0.41)	(-0.258)
Class 4	-0.051		0.329*	0.00976	-0.21**	-0.191	0.609***		-0.099	
	(-0.49)		(1.892)	(0.100)	(-2.10)	(-1.255)	(3.788)		(-1.00)	
Class 5						0.151	0.628***			
						(1.112)	(3.367)			
Class 6						0.679***	0.372**			
						(3.819)	(2.049)			
Obs	78	70	87	81	70	81	72	57	52	63

[1]Reported with marginal effects

†These schools have at least one class with zero G/T students.

‡These schools have at least one class with all G/T students.

Table B2.9: Schools 199-209

G/T sorting for schools 199-209

Dependent Variable: Gifted/Talented (Probit Regression[1])

School	199	200	201	202	203†	204†	205	206	207	209
Class 2	-0.063	0.505**	-0.114	0.0842			0.0233	0.0471	0.0899	0.0297
	(-0.88)	(2.143)	(-0.90)	(0.657)			(0.175)	(0.343)	(0.556)	(0.232)
Class 3	-0.085	0.631***	-0.044	0.0185	-0.095	-0.26**	-0.024	0.125	0.193	-0.113
	(1.13)	(2.856)	(-0.33)	(0.142)	(-0.99)	(-2.49)	(-0.18)	(0.883)	(1.133)	(-0.96)
Class 4	-0.101	0.173	-0.031	0.148	-0.035	-0.149	-0.125		0.567***	-0.059
	(1.35)	(0.624)	(-0.23)	(1.117)	(-0.35)	(-1.51)	(-0.97)		(3.273)	(-0.48)
Class 5		0.370	0.0308		0.193*	-0.023	0.0689		0.261	-0.093
		(1.534)	(0.221)		(1.657)	(-0.21)	(0.513)		(1.505)	(-0.77)
Class 6			-0.105		-0.100	-0.085	0.182			-0.211*
			(-0.82)		(-1.04)	(-0.84)	(1.334)			(-1.86)
Obs	78	63	90	63	95	113	135	57	108	118

[1]Reported with marginal effects

†These schools have at least one class with zero G/T students.

‡These schools have at least one class with all G/T students.

Table B2.10

G/T sorting for schools 210-219

Dependent Variable: Gifted/Talented (Probit Regression[1])										
School	210	211	212†	213†	214	215	216†	217 [2‡]	218	219
Class 2	0.0284	0.0458			-0.067	0.0575			0.178	0.0147
	(0.194)	(0.486)			(-0.40)	(0.476)			(1.162)	(0.114)
Class 3	0.00	-0.087	0.0808	-0.053	0.00	0.0500	-0.0794		0.329**	-0.0478
	(-0.00)	(-0.99)	(0.343)	(-0.42)	(-0.00)	(0.420)	(-0.952)		(2.110)	(-0.371)
Class 4	0.0835	-0.083	-0.036	0.203	0.130	0.129	-0.121		0.165	0.369***
	(0.586)	(-0.95)	(-0.17)	(1.612)	(0.745)	(1.013)	(-1.467)		(1.097)	(2.766)
Class 5	0.0363	0.0123	0.109		-0.150		-0.00736		0.165	
	(0.258)	(0.133)	(0.530)		(-0.93)		(-0.081)		(1.097)	
Class 6	-0.053				0.0502		-0.0874		0.100	
	(-0.38)				(0.282)		(-1.071)		(0.678)	
Class 7							-0.0479		0.238	
							(-0.561)		(1.537)	
Class 8									0.178	
									(1.162)	
Obs	120	99	44	63	89	83	122		169	92

[1]Reported with marginal effects

[2]No variation in G/T status

†These schools have at least one class with zero G/T students.

‡These schools have at least one class with all G/T students.

Table B2.11

G/T sorting for schools 220-232

Dependent Variable: Gifted/Talented (Probit Regression[1])										
School	220	222	223	224	225†	226†	228[2]	229	230†	232
Class 2	0.504***	0.0660	0.00433	-0.089				0.158		0.199
	(3.238)	(0.528)	(0.0488)	(-0.75)				(1.046)		(1.458)
Class 3	0.00823	0.126		-0.089	-0.067	0.00386		0.156	0.201	0.248*
	(0.0558)	(0.971)		(-0.75)	(-0.59)	(0.0354)		(0.972)	(1.008)	(1.799)
Class 4	0.369**	0.161			0.0590	0.192*		0.370**	0.491***	0.210
	(2.373)	(1.242)			(0.492)	(1.718)		(2.250)	(2.609)	(1.511)
Class 5	0.0170	0.0209			0.0590			0.244		0.128
	(0.114)	(0.168)			(0.492)			(1.488)		(0.974)
Class 6		0.0752						0.193		0.136
		(0.592)						(1.160)		(1.020)
Obs	109	137	43	63	80	85		110	44	130

[1]Reported with marginal effects

[2]No variation in G/T status

†These schools have at least one class with zero G/T students.

‡These schools have at least one class with all G/T students.

Table B2.12

G/T sorting for schools 233-262

Dependent Variable: Gifted/Talented (Probit Regression[1])										
School	233	235	236	237†	241 [2]	250	251	256	259 [2]	262
Class 2	0.0410	0.0888	-0.0753			0.276*	-0.119	-0.0941		0.122
	(0.266)	(0.641)	(-0.584)			(1.812)	(-1.001)	(-0.614)		(0.665)
Class 3	-0.194	0.0455	-0.0121			-0.132	-0.251**			0.280
	(-1.283)	(0.327)	(-0.0920)			(-0.74)	(-2.438)			(1.434)
Class 4	-0.0147		-0.0232	-0.0927			-0.259**			-0.0394
	(-0.097)		(-0.178)	(-0.892)			(-2.505)			(-0.219)
Class 5				-0.147			-0.186*			0.0700
				(-1.421)			(-1.701)			(0.376)
Class 6				0.00			-0.305***			-0.0955
				(-0.00)			(-3.015)			(-0.509)
Class 7				-0.0370						
				(-0.338)						
Class 8				-0.00941						
				(-0.0864)						
Obs	70	69	73	108		50	119	32		75

[1]Reported with marginal effects

[2]No variation in G/T status

†These schools have at least one class with zero G/T students.

‡These schools have at least one class with all G/T students.

Table B2.13

G/T sorting for schools 263-273

Dependent Variable: Gifted/Talented (Probit Regression[1])

School	263	264	265	267†	268	269	270	271	272	273†
Class 2	0.137	0.0825	-0.124		0.0596	0.272*	0.123	0.00181	0.138	
	(1.051)	(0.544)	(-0.89)		(0.525)	(1.656)	(1.035)	(0.0176)	(1.037)	
Class 3	-0.054	-0.053	-0.255*		-0.0863	0.272*	0.184	0.0406	0.0822	
	(-0.41)	(-0.34)	(-1.80)		(-0.756)	(1.656)	(1.466)	(0.389)	(0.637)	
Class 4		0.00			0.0747	0.319*	0.123	-0.0537	0.0756	-0.065
		(-0.00)			(0.636)	(1.932)	(1.035)	(-0.551)	(0.595)	(-0.71)
Class 5		0.251		0.00	0.195		0.0713	0.00949	0.0822	-0.053
		(1.542)		(-0.00)	(1.639)		(0.623)	(0.0905)	(0.637)	(-0.57)
Class 6				0.0866	0.0747		0.123	-0.0430	0.194	-0.112
				(0.423)	(0.636)		(1.035)	(-0.427)	(1.398)	(-1.23)
Class 7				0.102	-0.0023		-0.048	0.00181	0.138	
				(0.469)	(-0.020)		(-0.44)	(0.0176)	(1.037)	
Class 8					0.00987		-0.048	0.0180		
					(0.0832)		(-0.44)	(0.168)		
Class 9					0.0920					
					(0.755)					
Obs	63	90	52	43	181	91	176	165	152	88

[1]Reported with marginal effects

[2]No variation in G/T status

†These schools have at least one class with zero G/T students.

‡These schools have at least one class with all G/T students.

Table B2.14

G/T sorting for schools 274-284

Dependent Variable: Gifted/Talented (Probit Regression[1])

School	274	275†	276	277†	280	281	283	284
Class 2	0.0561		0.0566		-0.0422	0.0739	-0.0083	0.0214
	(0.582)		(0.539)		(-0.418)	(0.521)	(-0.065)	(0.194)
Class 3	-0.0474	0.102	-0.053	-0.19*	-0.0976	0.150	0.0352	-0.027
	(-0.556)	(0.774)	(-0.55)	(-1.88)	(-0.994)	(1.023)	(0.276)	(-0.24)
Class 4	0.00417	-0.0640	0.0566	-0.100	0.0449	-0.097	0.138	0.0214
	(0.0466)	(-0.517)	(0.539)	(-1.05)	(0.416)	(-0.68)	(1.048)	(0.194)
Class 5	-0.0474	-0.0100	0.0566	-0.026	0.0543	0.0739		
	(-0.556)	(-0.080)	(0.539)	(-0.26)	(0.492)	(0.521)		
Class 6	0.00869		-0.053		-0.0066	0.0739		
	(0.0954)		(-0.55)		(-0.065)	(0.521)		
Class 7	0.134				-0.0365	0.0450		
	(1.288)				(-0.354)	(0.282)		
Class 8								
Class 9								
Obs	153	90	115	102	136	123	84	78

[1]Reported with marginal effects

[2]No variation in G/T status

†These schools have at least one class with zero G/T students.

VITA

Name: Courtney Amanda Collins

Address: Mercer University
Stetson School of Business and Economics
1400 Coleman Ave.
Macon, GA 31207

Email Address: collins_ca@mercer.edu

Education: Ph.D., Economics, Texas A&M University, 2010
B.A., Economics, Rhodes College, 2005

Fields: Economics of Education
Public Economics
Applied Microeconometrics